

УНИВЕРЗИТЕТ У НОВОМ САДУ
ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА У
НОВОМ САДУ



мр Бобан Бонцулић

**ПРОЦЕНА КВАЛИТЕТА СЛИКЕ И
ВИДЕА КРОЗ ОЧУВАЊЕ
ИНФОРМАЦИЈА О ГРАДИЈЕНТУ**

ДОКТОРСКА ДИСЕРТАЦИЈА

Нови Сад, 2015



КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, РБР:	
Идентификациони број, ИБР:	
Тип документације, ТД:	Монографска документација
Тип записа, ТЗ:	Текстуални штампани материјал
Врста рада, ВР:	Докторат
Аутор, АУ:	Мр Бобан Бонцулић, дипл. инж.
Ментор, МН:	ванр. проф. др Владимир Петровић
Наслов рада, НР:	Процена квалитета слике и видеа кроз очување информација о градијенту
Језик публикације, ЈП:	Српски
Језик извода, ЈИ:	Српски
Земља публикација, ЗП:	Република Србија
Уже географско подручје, УГП:	Војводина
Година, ГО:	2015.
Издавач, ИЗ:	Ауторски репринт
Место и адреса, МА:	Нови Сад, Трг Доситеја Обрадовића 6
Физички опис рада, ФО: (поглавља/страна/ цитата/табела/слика/графика/прилога)	8 поглавља / 321 страна / 284 цитата / 86 табела / 124 слике
Научна област, НО:	Електротехничко и рачунарско инжењерство
Научна дисциплина, НД:	Телекомуникације и обрада сигнала
Предметна одредница/Кључне речи, ПО:	Дигитална обрада слике, процена квалитета слике и видеа, амплитуде и оријентације градијента, очување информација о градијенту
УДК	
Чува се, ЧУ:	У библиотеци Факултета техничких наука у Новом Саду
Важна напомена, ВН:	
Извод, ИЗ:	У овој дисертацији разматране су објективне мере процене квалитета слике и видеа са потпуним и делимичним референцирањем на изворни сигнал. За потребе евалуације квалитета развијене су поуздане, рачунски ефикасне мере, засноване на очувању информација о градијенту. Мере су тестиране на великом броју тест слика и видео секвенци, различитих типова и степена деградације. Поред јавно доступних база слика и видео секвенци, за потребе истраживања формиране су и нове базе видео секвенци са преко 300 релевантних тест узорака. Поређењем доступних субјективних и објективних скорова квалитета показано је да је објективна евалуација квалитета веома сложен проблем, али га је могуће решити и доћи до високих перформанси коришћењем предложених мера процене квалитета слике и видеа.
Датум прихватања теме, ДП:	
Датум одбране, ДО:	
Чланови комисије, КО:	Председник: Др Владимир Црнојевић, редовни професор
	Члан: Др Дубравко Ђулибрк, ванредни професор
	Члан: Др Бранислав Тодоровић, научни саветник
	Члан: Др Миленко Андрић, ванредни професор
	Члан, ментор: Др Владимир Петровић, ванредни професор

Потпис ментора



KEY WORDS DOCUMENTATION

Accession number, ANO :	
Identification number, INO :	
Document type, DT :	Monographic publication
Type of record, TR :	Textual printed material
Contents code, CC :	PhD thesis
Author, AU :	Boban Bondžulić, MSc
Mentor, MN :	Dr Vladimir Petrović, associate professor
Title, TI :	Gradient-based image and video quality assessment
Language of text, LT :	Serbian
Language of abstract, LA :	Serbian
Country of publication, CP :	Republic of Serbia
Locality of publication, LP :	Vojvodina
Publication year, PY :	2015.
Publisher, PB :	Author's reprint
Publication place, PP :	Novi Sad, Dositeja Obradovića sq. 6
Physical description, PD : <small>(chapters/pages/ref./tables/pictures/graphs/appendixes)</small>	8 chapters / 321 pages / 284 references / 86 tables / 124 pictures
Scientific field, SF :	Electrical and computer engineering
Scientific discipline, SD :	Telecommunications and signal processing
Subject/Key words, S/KW :	Digital image processing, image and video quality assessment, gradient magnitude and gradient orientation, gradient information preservation
UC	
Holding data, HD :	The Library of Faculty of Technical Sciences, Novi Sad, Serbia
Note, N :	
Abstract, AB :	This thesis presents an investigation into objective image and video quality assessment with full and reduced reference on original (source) signal. For quality evaluation purposes, reliable, computational efficient, gradient-based measures are developed. Proposed measures are tested on different image and video datasets, with various types of distortions and degradation levels. Along with publicly available image and video quality datasets, new video quality datasets are made, with more than 300 relevant test samples. Through comparison between available subjective and objective quality scores it has been shown that objective quality evaluation is highly complex problem, but it is possible to resolve it and achieve high performance using proposed quality measures.
Accepted by the Scientific Board on, ASB :	
Defended on, DE :	
Defended Board, DB :	President: Dr Vladimir Crnojević, full professor
	Member: Dr Dubravko Čulibrk, associate professor
	Member: Dr Branislav Todorović, principal research fellow
	Member: Dr Milenko Andrić, associate professor
	Member, Mentor: Dr Vladimir Petrović, associate professor
	Mentor's sign

Захвалница

Захваљујем свима који су ми помогли при реализацији ове дисертације или учинили да она буде боља и квалитетнија.

Изражавам захвалност ментору ванр. проф. др Владимиру Петровићу и члановима комисије – без њихових сугестија дисертација не би изгледала као што сада изгледа. Такође захваљујем на подршци колегама с Војне академије јер су створили позитивну атмосферу за истраживачки рад.

Хвала Мирјани и Ањи на стрпљењу и љубави које су несебично уложили у наш заједнички успех.

Садржај

САДРЖАЈ.....	II
СПИСАК СЛИКА.....	I
СПИСАК ТАБЕЛА.....	I
СКРАЋЕНИЦЕ.....	I
1. УВОД.....	1
2. ПОСЕБНИ АСПЕКТИ ПРОЦЕНЕ КВАЛИТЕТА СЛИКЕ И ВИДЕА.....	10
2.1. СУБЈЕКТИВНА ЕВАЛУАЦИЈА КВАЛИТЕТА.....	10
2.2. ЗНАЧАЈ VQEG ГРУПЕ У ОБЛАСТИ ПРОЦЕНЕ КВАЛИТЕТА.....	14
2.3. ЈЕДНОСТАВНЕ ОБЈЕКТИВНЕ МЕРЕ ПРОЦЕНЕ КВАЛИТЕТА СЛИКА.....	16
2.4. СЛОЖЕНИЈЕ ОБЈЕКТИВНЕ МЕРЕ ПРОЦЕНЕ КВАЛИТЕТА СЛИКА.....	17
2.4.1. <i>Хибридне технике процене квалитета слика.....</i>	<i>24</i>
2.4.2. <i>Колор модели намењени процени квалитета слика.....</i>	<i>25</i>
2.5. ПРИМЕНА МОДЕЛА ВИЗУЕЛНОГ СИСТЕМА ЧОВЕКА У ПРОЦЕНИ КВАЛИТЕТА СЛИКА И ВИДЕА.....	27
2.6. ОБЈЕКТИВНА ПРОЦЕНА КВАЛИТЕТА ВИДЕО СИГНАЛА.....	29
2.6.1. <i>Процена квалитета секвенци са малим битским брзинама.....</i>	<i>32</i>
2.6.2. <i>Простор и време као фактори процене квалитета.....</i>	<i>34</i>
2.6.3. <i>Процена квалитета HEVC кодованих секвенци.....</i>	<i>36</i>
2.7. УВОЂЕЊЕ МАПА ВИЗУЕЛНОГ ЗНАЧАЈА У ЕВАЛУАЦИЈУ КВАЛИТЕТА.....	36
2.7.1. <i>Квалитетом вођена интеграција.....</i>	<i>38</i>
2.7.2. <i>Интеграција на основу садржаја.....</i>	<i>39</i>
2.7.3. <i>Корисићење реалних мапа субјективног значаја.....</i>	<i>39</i>
2.7.4. <i>Корисићење рачунарских модела за формирање мапа субјективног значаја.....</i>	<i>40</i>
2.8. ПРОЦЕНА КВАЛИТЕТА СА РЕДУКОВАНИМ РЕФЕРЕНЦИРАЊЕМ.....	42
2.9. ПРОЦЕНА КВАЛИТЕТА БЕЗ РЕФЕРЕНЦИРАЊА.....	44
2.10. ОЧУВАЊЕ ГРАДИЈЕНТА У ПРОЦЕНИ КВАЛИТЕТА СЛИКЕ И ВИДЕА.....	48
2.11. КРИТЕРИЈУМИ ЗА ЕВАЛУАЦИЈУ ПЕРФОРМАНСИ ОБЈЕКТИВНИХ МЕРА ПРОЦЕНЕ КВАЛИТЕТА СЛИКЕ И ВИДЕА.....	52
2.11.1. <i>Тачност предикције.....</i>	<i>54</i>
2.11.2. <i>Монотоност предикције.....</i>	<i>55</i>
2.11.3. <i>Конзистентност предикције.....</i>	<i>56</i>
2.12. ЗАКЉУЧАК.....	58
3. ПРОЦЕНА КВАЛИТЕТА СЛИКА ЗАСНОВАНА НА ОЧУВАЊУ ИНФОРМАЦИЈА О ГРАДИЈЕНТУ.....	60
3.1. Увод.....	60
3.2. ИЗДВАЈАЊЕ ПАРАМЕТАРА ГРАДИЈЕНТА.....	61
3.3. МЕРЕ ОЧУВАЊА ПАРАМЕТАРА ГРАДИЈЕНТА.....	64
3.4. РЕЗУЛТАТИ ПРИМЕНЕ МЕРЕ ОЧУВАЊА ГРАДИЈЕНТА.....	67
3.4.1. <i>Утицај избора базе на перформансе објективних мера.....</i>	<i>90</i>
3.4.2. <i>Однос перформанси на подкуповима и на глобалном плану.....</i>	<i>92</i>
3.4.3. <i>Анализа перформанси објективних мера процене на глобалном плану.....</i>	<i>93</i>
4. ПОСЕБНИ АСПЕКТИ ПРОЦЕНЕ КВАЛИТЕТА СЛИКА.....	96
4.1. ИНФОРМАЦИОНИ АСПЕКТ ОЧУВАЊА ИНФОРМАЦИЈА О ГРАДИЈЕНТУ.....	96
4.1.1. <i>Пренете (заједничке) информације.....</i>	<i>98</i>
4.1.2. <i>Изгубљене информације.....</i>	<i>99</i>
4.1.3. <i>Убачене (лажне) информације.....</i>	<i>99</i>
4.1.4. <i>Резултати.....</i>	<i>104</i>
4.2. ПРИМЕНА МАПЕ СУБЈЕКТИВНОГ ЗНАЧАЈА У ЕВАЛУАЦИЈИ КВАЛИТЕТА.....	106
4.2.1. <i>Одређивање мапе субјективног значаја.....</i>	<i>111</i>

4.2.2.	Примена мапе субјективног значаја у процени квалитета слике.....	116
4.3.	СЈЕДИЊАВАЊЕ ИНФОРМАЦИЈА У ПРОЦЕНИ КВАЛИТЕТА СЛИКЕ	119
4.3.1.	Сједињавање информација о очувању структуре.....	120
4.3.1.1.	Индекс структурне сличности.....	121
4.3.1.2.	Квалитетом вођена процена квалитета.....	122
4.3.1.3.	Адитивни индекс структурне сличности.....	126
4.3.1.4.	Перформансе адитивних SSIM мера.....	128
4.3.1.5.	Закључак.....	131
4.3.2.	Примена адитивног модела сједињавања информација код објективне процене засноване на очувању градијента.....	132
4.4.	ЗАКЉУЧАК	141

5. ПРОЦЕНА КВАЛИТЕТА ВИДЕА ЗАСНОВАНА НА ОЧУВАЊУ ИНФОРМАЦИЈА О ГРАДИЈЕНТУ 143

5.1.	Увод.....	143
5.2.	РАЗЛАГАЊЕ ВИЗУЕЛНИХ ИНФОРМАЦИЈА.....	145
5.3.	ОЧУВАЊЕ ПРОСТОРНИХ ИНФОРМАЦИЈА	147
5.4.	ОЧУВАЊЕ ВРЕМЕНСКИХ ИНФОРМАЦИЈА	149
5.5.	ОЧУВАЊЕ ХРОМАТСКИХ ИНФОРМАЦИЈА.....	152
5.6.	КРАЉЈА МЕРА КВАЛИТЕТА ВИДЕА	154
5.7.	ПРИМЕРИ ПРИМЕНЕ МЕРЕ КВАЛИТЕТА VQ^{AB}	157
5.8.	КВАЛИТЕТОМ ВОЂЕНА VQ^{AB} МЕРА.....	164
5.9.	АНАЛИЗА ПЕРФОРМАНСИ ПРЕДЛОЖЕНЕ МЕРЕ	165
5.9.1.	Анализа резултата на UoM бази секвенци.....	166
5.9.2.	Анализа резултата на FUB бази секвенци.....	173
5.9.3.	Анализа резултата на EPFL-PoliMI бази секвенци.....	178
5.9.4.	Анализа резултата на VQEG бази секвенци	183
5.9.5.	Анализа резултата на LIVE Video бази секвенци	187
5.9.6.	Анализа резултата на LIVE Wireless бази секвенци.....	194
5.10.	УВОЂЕЊЕ ЗНАЧАЈА ОБЈЕКТИВНИХ СКОРОВА КАДРОВА	198
5.11.	ПОРЕЂЕЊЕ ЛИНЕАРНОГ И НЕЛИНЕАРНОГ МОДЕЛА ОЧУВАЊА ИНФОРМАЦИЈА О ГРАДИЈЕНТУ... ..	207
5.12.	ОПИС STRIAL 0 И STRIAL 1 СЕТОВА НАМЕЊЕНИХ ПРОЦЕНИ КВАЛИТЕТА ВИДЕО СЕКВЕНЦИ СА МАЛИМ БИТСКИМ БРЗИНАМА.....	208
5.12.1.	Опис референтних видео секвенци.....	211
5.12.2.	Опис коришћених кодека	211
5.12.3.	Просторна резолуција видео секвенци.....	213
5.12.4.	Битске брзине.....	214
5.12.5.	Трајање секвенци.....	215
5.12.6.	Генерисање видео записа.....	215
5.12.7.	Опрема за испитивање.....	220
5.12.8.	Софтвер за тестирање	221
5.12.9.	Процедура тестирања	221
5.12.10.	Анализа прикупљених података.....	222
5.12.11.	Опис првог субјективног теста – тест STRial 0.....	224
5.12.12.	Опис другог субјективног теста – тест STRial 1	225
5.12.13.	Статистика субјективних скорова STRial 0 места.....	229
5.12.14.	Статистика субјективних скорова STRial 1 места.....	234
5.12.15.	Резултати субјективних тестова – коментари	240
5.13.	АНАЛИЗА РЕЗУЛТАТА ОБЈЕКТИВНИХ МЕРА ПРОЦЕНЕ НА STRIAL СЕТОВИМА.....	242
5.13.1.	Мере и софтвер за евалуацију објективних мера процене.....	243
5.13.2.	Мере за процену перформанси.....	245
5.13.3.	Анализа перформанси објективних мера процене на STRial 0 сету.....	246
5.13.4.	Анализа перформанси објективних мера процене на STRial 1 сету.....	251
5.14.	ЗАКЉУЧАК	256

6. ПРОЦЕНА КВАЛИТЕТА ВИДЕА У УСЛОВИМА РЕДУКОВАНОГ РЕФЕРЕНЦИРАЊА.....	261
6.1. Увод.....	261
6.2. УВОЂЕЊЕ МАПЕ СУБЈЕКТИВНОГ ЗНАЧАЈА.....	268
6.3. ЕКСПЕРИМЕНТАЛНИ ПОДАЦИ О КРЕТАЊУ ЗЕНИЦЕ ЉУДСКОГ ОКА.....	270
6.4. ФОРМИРАЊЕ МАПЕ ВИЗУЕЛНОГ ЗНАЧАЈА ИЗ РЕАЛНИХ УЗОРАКА.....	273
6.5. РЕАЛНО ПОНДЕРИСАНЕ ОБЈЕКТИВНЕ МЕРЕ ПРОЦЕНЕ КВАЛИТЕТА.....	281
6.5.1. Пондерисани вршни однос сигнал шум - $WPSNR$	281
6.5.2. Пондерисани индекс структурне сличности - $WSSIM$	282
6.5.3. Пондерисана мера $VQ^{AB} - WVQ^{AB}$	282
6.6. АНАЛИЗА ПЕРФОРМАНСИ ПОНДЕРИСАНИХ ОБЈЕКТИВНИХ МЕРА НА UoM СЕТУ.....	283
6.7. ЗАКЉУЧАК И ПРАВЦИ ДАЉЕГ РАДА.....	294
7. ЗАКЉУЧАК.....	297
8. ЛИТЕРАТУРА.....	303

Списак слика

Сл. 2.1 Изворна слика "Lena" и њене модификације настале: (а) променом средње вредности нивоа сивог (mean-shift), (б) развлачењем контраста, (в) замрљањем и (г) JPEG компресијом.	18
Сл. 2.2 Блок дијаграм одређивања индекса структурне сличности два сигнала.	19
Сл. 2.3 Пример везе између VQR и DMOS.	53
Сл. 2.4 Поређење мера објективне процене квалитета кроз тачност предикције.	54
Сл. 2.5 Поређење мера објективне процене квалитета кроз монотоност предикције.	56
Сл. 2.6 Поређење мера објективне процене квалитета кроз конзистентност предикције.	57
Сл. 3.1 Генерална структура градијентне објективне процене квалитета.	62
Сл. 3.2 Вертикална и хоризонтална Собел маска.	62
Сл. 3.3 (а) изворна слика, (б) амплитуде градијента и (в) оријентације градијента.	63
Сл. 3.4 Промена оријентације за све могуће комбинације оријентација градијента референтне и тест слике.	65
Сл. 3.5 (а) нелинеарне сигмоидне функције које се користе у очувању параметара градијента, (б) нелинеарни сигмоидни модел очувања информација који се користи за моделовање одзива биолошких система и линеарни модел очувања информација (модел репрезентације сигнала).	66
Сл. 3.6 (а) изворна слика, (б) (в) (г) тест слике са степеном компресије од 0.25 bpr, 0.58 bpr и 1.78 bpr, тим редом, (д) (ђ) (е) мапе квалитета добијене применом мере очувања градијента на слике из претходног реда.	68
Сл. 3.7 (а) (б) (в) изворне слике, (г) (д) (ђ) тест слике настале применом JPEG компресије са 0.27 bpr, 0.30 bpr и 0.78 bpr, тим редом, (е) (ж) (з) мапе квалитета/дисторзије тест слика.	70
Сл. 3.8 (а) (б) (в) изворне слике, (г) (д) (ђ) тест слике настале применом JPEG2000 компресије са 0.19 bpr, 0.20 bpr и 1.15 bpr, тим редом, (е) (ж) (з) мапе квалитета/дисторзије тест слика.	71
Сл. 3.9 (а) изворна слика из IVC базе и њене модификације настале: (б) JPEG компресијом, (в) JPEG2000 компресијом, (г) замрљањем и (д) LAR кодовањем.	74
Сл. 3.10 Тест слике CSIQ базе са различитим деградацијама: (а) JPEG компресија, (б) JPEG2000 компресија, (в) замрљање, (г) модификација контраста, (д) адитивни бели шум и (ђ) адитивни ружичасти (pink) шум.	75
Сл. 3.11 Дијаграми расипања субјективних (DMOS) и објективних вредности квалитета на LIVE бази слика.	76
Сл. 3.12 Дијаграми расипања субјективних (MOS) и објективних вредности квалитета на Тоуата бази слика.	80
Сл. 3.13 Дијаграми расипања субјективних (MOS) и објективних вредности квалитета на IVC бази слика.	83
Сл. 3.14 Дијаграми расипања субјективних (DMOS) и објективних вредности квалитета на CSIQ бази слика.	86
Сл. 3.15 Коефицијент линеарне корелације (LCC) и корелација рангова (SROCC) на подскуповима слика различитих база: (а) JPEG2000 подскуп слика и (б) JPEG подскуп слика.	90

Сл. 3.16 Изворна слика и њене модификације у различитим базама слика: (а) изворна слика, (б) JPEG деградирана слика из Тоуата базе (MOS=1.69), (в) JPEG деградирана слика из LIVE базе (DMOS=84.89), (г) JPEG2000 деградирана слика из Тоуата базе (MOS=1.19) и (д) JPEG2000 деградирана слика из LIVE базе (DMOS=82.73).	91
Сл. 3.17 Илустрација перформанси VSNR мере на CSIQ бази слика и на њеним подскуповима: (а) дијаграм расипања и интерполациона функција на комплетној бази, (б) интерполационе функције на подскуповима слика, (в) коефицијент линеарне корелације (LCC) на подскуповима и комплетној бази и (г) корелација рангова (SROCC) на подскуповима и комплетној бази.	93
Сл. 3.18 Анализа перформанси мера објективне процене на глобалном нивоу (нивоу комплетних база слика): (а) коефицијент линеарне корелације (LCC) и (б) корелација рангова (SROCC).	94
Сл. 4.1 Једноставни Венов дијаграма процеса преноса слике.	98
Сл. 4.2 (а) изворна слика, (б) тест слика настала додавањем Гаусовог шума изворној слици, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).	101
Сл. 4.3 (а) изворна слика, (б) тест слика настала JPEG компресијом изворне слике, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).	102
Сл. 4.4 (а) изворна слика, (б) тест слика настала замрљањем изворне слике, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).	103
Сл. 4.5 Дијаграми расипања субјективних импресија (DMOS) и уведених аспеката на LIVE бази слика: (а) заједничке информације, (б) изгубљене информације и (в) убачене информације.	106
Сл. 4.6 (а) изворна слика, (б) мапа визуелног значаја Сл. 4.6(а), (в) изворна слика и (г) мапа визуелног значаја Сл. 4.6(в).	108
Сл. 4.7 (а) изворна слика, (б) тест слика, (в) мапа квалитета добијена помоћу SSIM и (г) пиксели са најнижим локалним SSIM скоровима квалитета (бели пиксели).	110
Сл. 4.8 (а) (б) изворне слике, (в) (г) хистограми градијента изворних слика и функције значаја и (д) (ђ) мапе значаја.	113
Сл. 4.9 (а) изворна слика, (б) тест слика, (в) мапа очувања параметара градијента, (г) (д) мапе субјективног значаја добијене коришћењем (42) и (43), тим редом.	115
Сл. 4.10 Корелација рангова (SROCC) у функцији параметра p за SSIM (47), контраст c (45) и структуру s (46) на LIVE подскупу слика за обучавање.	126
Сл. 4.11 SROCC 3D оптимизациона површина тренинг скупа LIVE базе у функцији SSIM p_c и p_s вредности из (49) ($w_c=0.55$).	127
Сл. 4.12 Корелација рангова AM-SSIM тренинг скупа у функцији w_c ($w_s=1-w_c$).	128
Сл. 4.13 Расипање субјективних DMOS/MOS и објективних SSIM/AM-SSIM/AM-GSSIM скорова квалитета. Свака тачка одговара једној тест слици. Интерполационе криве су дате пуном линијом.	131
Сл. 4.14 SROCC 3D оптимизациона површина тренинг скупа LIVE базе у функцији p_g и p_a из (54) ($w_g=0.7$).	134
Сл. 4.15 Корелација рангова AM- Δ^{AB} тренинг скупа у функцији w_g ($w_a=1-w_g$).	135
Сл. 4.16 Расипање субјективних DMOS/MOS и објективних (Q^{AB} и AM- Δ^{AB}) скорова квалитета.	139

Сл. 4.17 Поређење перформанси објективних градијентних процена на четири базе слика: (а) коефицијент линеарне корелације (LCC) субјективних и објективних скорова и (б) корелација рангова (SROCC) субјективних и објективних скорова. . .	139
Сл. 5.1 Пренос видеа са становишта очувања и процене квалитета.	144
Сл. 5.2 Евалуација квалитета деградиране секвенце.	145
Сл. 5.3 Илустрација HSV колор простора.	146
Сл. 5.4 Трансформација RGB слике у HSV колор простор.	147
Сл. 5.5 Оператори коришћени за издвајање: (а) временског градијента и (б) просторног градијента.	150
Сл. 5.6 Евалуација очувања хроматских информација.	153
Сл. 5.7. Нелинеарна сигмоидна функција која се користи у очувању хроматских информација.	155
Сл. 5.8 Оптимизационе површине значаја изворних информација у објективној процени квалитета: (а) према коефицијенту линеарне корелације, LCC, (б) према корелацији рангова, SROCC.	156
Сл. 5.9 Пример очувања информација изворне секвенце за битски проток 96 kbps.	158
Сл. 5.10 Пример очувања информација изворне секвенце за битски проток 416 kbps.	159
Сл. 5.11 Очување информација током времена за примере анализиране на Сл. 5.9 и Сл. 5.10.	161
Сл. 5.12 Кадрови изворних и деградираних секвенци за различите параметре система преноса: (а) промена кодека, (б) промена битске брзине и (в) промена димензије кадра.	162
Сл. 5.13 Приказ кадрова изворних секвенци УоМ базе.	167
Сл. 5.14 Шематски приказ презентације видео садржаја.	168
Сл. 5.15 Дијаграми расипања објективних и субјективних скорова за секвенце у УоМ сету: (а) (в) (д) прикази са интервалима поузданости субјективних скорова, (б) (г) и (ђ) прикази са оптималним интерполационим кривама.	170
Сл. 5.16 Приказ резултата објективних мера из Таб. 5.5 (средње вредности по подскуповима и комплетан сет): (а) корелација рангова, SROCC, и (б) коефицијент линеарне корелације, LCC.	172
Сл. 5.17 Дијаграми расипања објективних и субјективних скорова за секвенце у УоМ сету према различитим улазним параметрима: (а) типу компресије, (б) димензији кадра и (в) битском протоку.	173
Сл. 5.18 Средње вредности и стандардне девијације: (а) субјективних скорова, (б) објективних VQ^{AB} скорова, (в) објективних SSIM скорова и (г) објективних PSNR скорова, по кодецима коришћеним у УоМ сету.	174
Сл. 5.19 Приказ илустративних кадрова изворних секвенци FUB сета.	175
Сл. 5.20 Дијаграми расипања субјективних и објективних (SMOVIЕ и VQ^{AB}) скорова (свака тачка одговара једној видео секвенци).	177
Сл. 5.21 Приказ илустративних кадрова изворних секвенци EPFL-PoliMI сета (CIF подскуп).	179
Сл. 5.22 Дисторзије добијене у преносу компримоване секвенце кроз канал са различитим нивоима пакетских губитака.	182
Сл. 5.23 Дијаграми расипања субјективних и објективних VQ^{AB} скорова (свака тачка одговара једној видео секвенци).	182

Сл. 5.24 Примери кадрова изворних секвенци из VQEG Phase I сета – 625 линија/50 Hz.....	184
Сл. 5.25 Примери кадрова изворних секвенци из VQEG Phase I сета – 525 линија/60 Hz.....	184
Сл. 5.26 Примери кадрова компримованих тест секвенци са различитим битским брзинама – сесија са секвенцама високог квалитета.....	185
Сл. 5.27 Примери кадрова компримованих тест секвенци са различитим битским брзинама – сесија са секвенцама ниског квалитета.....	186
Сл. 5.28 Графички приказ резултата предложене VQ^{AB} мере на VQEG Phase I сету - дијаграми расипања за: (а) 525/60Hz сет, (в) 625/50Hz сет и (д) комплетну базу, и прикази одговарајућих логистик функција (б), (г) и (ђ).....	188
Сл. 5.29 Примери кадрова изворних секвенци LIVE Video сета.....	189
Сл. 5.30 Примери деградација унетих у изворну видео секвенцу LIVE базе.....	191
Сл. 5.31 Дијаграми расипања објективних VQ^{AB} и субјективних скорова за све видео секвенце у LIVE Video бази: (а) приказ са интервалима поузданости субјективних скорова, (б) приказ са оптималном логистик кривом.....	193
Сл. 5.32 Примери кадрова изворних секвенци LIVE Wireless базе.....	195
Сл. 5.33 Деградације које се јављају у компримованој видео секвенци са 2 Mbps пренетој кроз канал са различитим нивоима пакетских губитака.....	196
Сл. 5.34 Дијаграми расипања објективних VQ^{AB} и субјективних скорова за све видео секвенце у LIVE Wireless бази: (а) приказ са интервалима поузданости субјективних скорова, (б) приказ са оптималном логистик кривом.....	197
Сл. 5.35 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на UoM сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.9784, $p=38\%$) и (г) најлошији случај (LCC=0.9502, $p=2\%$).....	199
Сл. 5.36 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на EPFL-PoliMI сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.8380, $p=2\%$) и (г) најлошији случај (LCC=0.6542, $p=100\%$).....	200
Сл. 5.37 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на VQEG 525/60Hz сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.8144, $p=100\%$) и (г) најлошији случај (LCC=0.7053, $p=2\%$).....	201
Сл. 5.38 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на VQEG 625/50Hz сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.8453, $p=94\%$) и (г) најлошији случај (LCC=0.7567, $p=2\%$).....	202
Сл. 5.39 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова за LIVE Video базу секвенци: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.6532, $p=6\%$) и (г) најлошији случај (LCC=0.5323, $p=100\%$).....	203
Сл. 5.40 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова за LIVE Wireless базу секвенци: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.9548, $p=20\%$) и (г) најлошији случај (LCC=0.9509, $p=2\%$).....	204
Сл. 5.41 (а) процена квалитета кадрова тест секвенце из EPFL-PoliMI базе, (б) процена квалитета кадрова тест секвенце из VQEG базе, (в) зависност мере VQ^{AB} од процента релевантних кадрова секвенце из EPFL-PoliMI базе и (г) зависност мере VQ^{AB} од процента релевантних кадрова секвенце из VQEG базе. Тест секвенце су настале од исте изворне секвенце – секвенце Mobile.....	205

Сл. 5.42 Кадрови анализираних тест секвенци: (а) кадар најбољег квалитета из EPFL-PoliMI базе, (б) кадар најлошијег квалитета из EPFL-PoliMI базе, (в) кадар најбољег квалитета из VQEG базе и (г) кадар најлошијег квалитета из VQEG базе.	206
Сл. 5.43 Примери кадрова изворних секвенци STrial сетова.	211
Сл. 5.44 Поједностављена блок шема поступка добијања видео материјала прилагођеног захтевима субјективног теста.	216
Сл. 5.45 Поједностављена блок шема MPEG-4 Part2 и H.264/AVC поступака кодовања са одговарајућим приказима екрана током кодовања.	217
Сл. 5.46 Поједностављена блок шема RealVideo 10 поступка кодовања са одговарајућим приказима екрана током кодовања.	218
Сл. 5.47 Поједностављена блок шема поступка издвајања репрезентативног дела секвенце са одговарајућим приказима екрана током рада.	219
Сл. 5.48 Изглед странице за приказ и навигацију (STrial 1 сет).	220
Сл. 5.49 Листинг фајлова са субјективним импресијама на STrial 1 сету.	223
Сл. 5.50 Садржај текстуалне датотеке за једног од учесника STrial 1 сета.	224
Сл. 5.51 Анализа резултата STrial 0 сета: (а) сортиране MOS вредности са интервалима поузданости ($MOS \pm SE$) и (б) хистограм MOS вредности.	231
Сл. 5.52 Расподела субјективних скорова за различите референтне секвенце у STrial 0 сету.	232
Сл. 5.53 Средње вредности и стандардне девијације субјективних скорова ($MOS \pm \sigma$ опсег) за сваког од посматрача STrial 0 сета.	234
Сл. 5.54 Анализа резултата STrial 1 сета: (а) сортиране MOS вредности са интервалима поузданости ($MOS \pm SE$) и (б) хистограм MOS вредности.	237
Сл. 5.55 Расподела субјективних скорова за различите референтне секвенце у STrial 1 сету.	238
Сл. 5.56 Средње вредности и стандардне девијације субјективних скорова ($MOS \pm \sigma$ опсег) за сваког од посматрача STrial 1 сета.	240
Сл. 5.57 Део листинга субјективних скорова STrial 1 сета.	245
Сл. 5.58 Дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета: (а) према приказаним сетовима и (б) према референтним секвенцама.	246
Сл. 5.59 (а) субјективни и (б) објективни видео квалитет добијен за различите кодеке STrial 0 сета.	248
Сл. 5.60 (а) субјективни и (б) објективни квалитет видео секвенци са различитим битским брзинама STrial 0 сета.	249
Сл. 5.61 (а) субјективни и (б) објективни видео квалитет кодовања са различитим енкодерима при различитим битским брзинама STrial 0 сета.	250
Сл. 5.62 Дијаграми расипања субјективних и објективних скорова на STrial 0 сету.	251
Сл. 5.63 Дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета: (а) према приказаним сетовима и (б) према референтним секвенцама.	252
Сл. 5.64 (а) субјективни и (б) објективни видео квалитет добијен за различите кодеке STrial 1 сета.	254
Сл. 5.65 (а) субјективни и (б) објективни квалитет видео секвенци са различитим битским брзинама STrial 1 сета.	254
Сл. 5.66 (а) субјективни и (б) објективни квалитет видео кодовања са различитим енкодерима при различитим битским брзинама STrial 1 сета.	255

Сл. 5.67 Дијаграми расипања субјективних и објективних скорова на STrial 1 сету.	257
Сл. 6.1 Генерални приступ при процени квалитета слике/видеа са делимичним референцирањем.	262
Сл. 6.2 Уобичајени приступ интеграције значаја визуелних информација у процес процене квалитета.	269
Сл. 6.3 (а) ASL 500 Eye Tracker систем и (б) експериментални услови друге студије у којима је спроведено прикупљање података.	270
Сл. 6.4 Приказ фиксација посматрача преко кадрова тест секвенци.	272
Сл. 6.5 (а) густина фоторецептора и ћелија ганглија на мрежњачи, (б) експоненцијални филтар који се користи за моделовање одзива ћелија мрежњаче и (в) 3-D приказ филтра који се користи у формирању континуалне мапе значаја.	274
Сл. 6.6 Кадрови изворне секвенце ‘Original_176x144_Road_50.avi’	276
Сл. 6.7 Кадрови секвенце са деградацијом – EV компресијом ‘EV_176x144_576_S3_014.avi’	276
Сл. 6.8 Региони фиксација за први кадар секвенце са деградацијом анализирани на Сл. 6.7: (а) бинарна мапа фиксација, (б) финална мапа значаја, (в) мапа значаја приказана у 3-D простору и (г) региони од значаја у анализираном кадру са деградацијом (црвеном бојом су оивичени региони за које је субјективни значај већи од нуле).	277
Сл. 6.9 Позиције фиксација за сваког од посматрача током посматрања комплетне тест секвенце (позиције су приказане преко 36.-тог кадра анализираних секвенци): (а) (б) посматрач 1 (сесија 4, 37-а секвенца), (в) (г) посматрач 2 (сесија 1, 46-та секвенца), (д) (ђ) посматрач 3 (сесија 3, 7-а секвенца) и (е) (ж) посматрач 4 (сесија 1, 31-а секвенца).	279
Сл. 6.10 Региони случајних фиксација за први кадар секвенце са деградацијом анализирани на Сл. 6.7: (а) бинарна мапа фиксација, (б) финална мапа значаја, (в) мапа значаја приказана у 3-D простору и (г) случајни региони од значаја у анализираном кадру (оивичени црвеном бојом).	280
Сл. 6.11 Дијаграми расипања објективних, пондерисаних објективних и субјективних скорова за секвенце у UoM сету: (а) PSNR, (б) SSIM и (в) VQ^{AB}	285
Сл. 6.12 Дијаграми расипања објективних, пондерисаних објективних и субјективних скорова за секвенце у UoM сету: (а) PSNR, (б) SSIM и (в) VQ^{AB} (анализа према типу кодека).	286
Сл. 6.13 Перформансе објективних мера процене квалитета (коэффициент линеарне корелације LCC) кроз: (а) очување просторних информација, (б) очување временских информација, (в) очување информација о боји и (г) очување информација у сва три канала.	290
Сл. 6.14 Перформансе објективних мера процене квалитета (корелација рангова SROCC) кроз: (а) очување просторних информација, (б) очување временских информација, (в) очување информација о боји и (г) очување информација у сва три канала.	291
Сл. 6.15 Дијаграми расипања субјективних и објективних скорова – мера очувања временских информација.	292
Сл. 6.16 Процентуално учешће изворних информација коришћених у процени квалитета са редукованим референцирањем UoM сета.	294

Списак табела

Таб. 3.1 Вредности субјективних и објективних процена тест слика са Сл. 3.6.	69
Таб. 3.2 Вредности субјективних и објективних процена тест слика са Сл. 3.7.	71
Таб. 3.3 Вредности субјективних и објективних процена тест слика са Сл. 3.8.	71
Таб. 3.4 Основни подаци о условима при којима су прикупљени резултати субјективних тестова референтних база слика.	72
Таб. 3.5 Поређење перформанси објективних мера процене квалитета на комплетној LIVE бази слика.	77
Таб. 3.6 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.	78
Таб. 3.7 Корелација рангова (SROCC) субјективних и објективних скорова на LIVE бази слика.	78
Таб. 3.8 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.	78
Таб. 3.9 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.	79
Таб. 3.10 Конзистентност објективних скорова (OR) након нелинеарне регресије на LIVE бази слика.	79
Таб. 3.11 Поређење перформанси објективних мера процене квалитета на комплетној Тоуата бази слика.	81
Таб. 3.12 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG подскуп.	81
Таб. 3.13 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG2000 подскуп.	81
Таб. 3.14 Поређење перформанси објективних мера процене квалитета на комплетној IVC бази слика.	82
Таб. 3.15 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.	84
Таб. 3.16 Корелација рангова (SROCC) субјективних и објективних скорова на IVC бази слика.	84
Таб. 3.17 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.	84
Таб. 3.18 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.	85
Таб. 3.19 Конзистентност објективних скорова (OR) (након нелинеарне регресије) на IVC бази слика.	85
Таб. 3.20 Поређење перформанси објективних мера процене квалитета на комплетној CSIQ бази слика.	87
Таб. 3.21 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.	88
Таб. 3.22 Корелација рангова (SROCC) субјективних и објективних скорова на CSIQ бази слика.	88
Таб. 3.23 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.	88

Таб. 3.24 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.....	89
Таб. 3.25 Конзистентност објективних скорова (OR) након нелинеарне регресије на CSIQ бази слика.	89
Таб. 4.1 Корелација рангова (SROCC) субјективних и објективних скорова на подскуповима и комплетној LIVE бази.	104
Таб. 4.2 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова на подскуповима и комплетној LIVE бази.....	104
Таб. 4.3 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.....	116
Таб. 4.4 Корелација рангова (SROCC) субјективних и објективних скорова на LIVE бази слика.	117
Таб. 4.5 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.	117
Таб. 4.6 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG подскуп.....	117
Таб. 4.7 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG2000 подскуп.....	117
Таб. 4.8 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.....	118
Таб. 4.9 Корелација рангова (SROCC) субјективних и објективних скорова на IVC бази слика.	118
Таб. 4.10 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.....	118
Таб. 4.11 Корелација рангова (SROCC) субјективних и објективних скорова на CSIQ бази слика.	119
Таб. 4.12 Перформансе различитих SSIM мера на подскупу слика за обучавање из LIVE базе.	124
Таб. 4.13 Перформансе различитих GSSIM мера на подскупу слика за обучавање из LIVE базе.	124
Таб. 4.14 Перформансе објективних мера процене на LIVE, CSIQ, IVC и Тоуата базама слика.	130
Таб. 4.15 Поређење перформанси мера процене квалитета на LIVE бази слика (подскуп слика за тестирање).	136
Таб. 4.16 Поређење перформанси мера процене квалитета на CSIQ бази слика.	136
Таб. 4.17 Поређење перформанси објективних мера процене квалитета на IVC бази слика.	136
Таб. 4.18 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.	137
Таб. 4.19 Оптималне вредности параметара адитивног модела очувања информација о градијенту.	140
Таб. 4.20 Поређење перформанси мера процене квалитета на LIVE бази слика (подскуп слика за тестирање).	140
Таб. 4.21 Поређење перформанси објективних мера процене квалитета на CSIQ бази слика.....	141
Таб. 4.22 Поређење перформанси објективних мера процене квалитета на IVC бази слика.....	141

Таб. 4.23 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.	141
Таб. 5.1 Вредности субјективних и објективних процена за секвенце анализиране на Сл. 5.9 и Сл. 5.10.	160
Таб. 5.2 Субјективни и објективни скорови за примере са Сл. 5.12.	163
Таб. 5.3 Параметри коришћени при формирању UoM базе секвенци.	166
Таб. 5.4 Перформансе три објективне мере на комплетном UoM сету.	169
Таб. 5.5 Анализа перформанси објективних мера на подскуповима и комплетној UoM бази секвенци.	171
Таб. 5.6 Параметри коришћени при формирању FUB базе секвенци.	174
Таб. 5.7 Коефицијент линеарне корелације (LCC) објективних и субјективних скорова на FUB бази секвенци.	176
Таб. 5.8 Корелација рангова (SROCC) објективних и субјективних скорова на FUB бази секвенци.	176
Таб. 5.9 Средња апсолутна грешка (MAE) између субјективних и објективних скорова на FUB бази секвенци.	176
Таб. 5.10 Корен средње квадратне грешке (RMSE) између субјективних и објективних скорова на FUB бази секвенци.	177
Таб. 5.11 Конзистентност (OR) објективних скорова на FUB бази секвенци.	177
Таб. 5.12 Перформансе оригиналних кодованих поворки бита из EPFL-PoliMI базе секвенци.	180
Таб. 5.13 Перформансе објективних мера на комплетном EPFL-PoliMI сету.	180
Таб. 5.14 Поређење перформанси мера објективне процене на VQEG Phase I сету (узете су у обзир секвенце из све четири сесије).	187
Таб. 5.15 Поређење перформанси алгоритама објективне процене видеа на LIVE Video бази – анализа SROCC за сваку од деградација и базу у целини.	192
Таб. 5.16 Поређење перформанси алгоритама објективне процене видеа на LIVE Video бази – анализа LCC за сваку од деградација и базу у целини.	192
Таб. 5.17 Перформансе мера објективне процене на LIVE Wireless бази секвенци.	197
Таб. 5.18 Перформансе предложене мере објективне процене квалитета кроз поређење линеарног и нелинеарног модела очувања информација.	208
Таб. 5.19 Опис изворних секвенци коришћених у Strial сетовима.	212
Таб. 5.20 Додатне информације о изворним секвенцама STrial 0 сета – назив секвенце, временски исечак секвенце коришћен у анализи, жељена и достигнута средња битска брзина.	226
Таб. 5.21 Додатне информације о изворним секвенцама STrial 1 сета – назив секвенце, временски исечак секвенце коришћен у анализи, жељена и достигнута средња битска брзина.	228
Таб. 5.22 Статистике субјективних скорова видео секвенци STrial 0 сета.	229
Таб. 5.23 Стандардне девијације субјективних скорова деградираних STrial 0 видео секвенци насталих од исте изворне секвенце.	233
Таб. 5.24 Статистике субјективних скорова видео секвенци STrial 1 сета.	234
Таб. 5.25 Стандардне девијације субјективних скорова деградираних STrial 1 видео секвенци насталих од исте изворне секвенце.	239
Таб. 5.26 Сагласност субјектата у STrial 1 сету.	239
Таб. 5.27 Средње вредности стандардних девијација и стандардних грешака субјективних скорова на комплетним STrial сетовима.	241

Таб. 5.28 Квантитативни показатељи перформанси VQ^{AB} мере на STrial 0 сету.....	247
Таб. 5.29 Квантитативни показатељи перформанси према референтним секвенцама.	248
Таб. 5.30 Перформансе објективних мера на STrial 0 сету.....	250
Таб. 5.31 Квантитативни показатељи перформанси VQ^{AB} мере на STrial 1 сету.....	252
Таб. 5.32 Квантитативни показатељи перформанси према референтним секвенцама.	253
Таб. 5.33 Перформансе објективних мера на STrial 1 сету.....	256
Таб. 6.1 Типичне битске брзине за додатни канал делимичног референцирања.	265
Таб. 6.2 Перформансе објективних мера на UoM сету.	283
Таб. 6.3 Перформансе мера репрезентације изворних информација у секвенцама са деградацијом UoM сета.	288
Таб. 6.4 Перформансе мера репрезентације изворних информација након примене нелинеарних сигмоидних функција у секвенцама са деградацијом UoM сета.	289
Таб. 6.5 Перформансе мера очувања изворних информација након примене нелинеарних сигмоидних функција и увођења значаја изведеног из амплитуде градијента у секвенцама са деградацијом UoM сета.	289

Скраћенице

3D TV	3D television
4G/LTE	Fourth-Generation/Long-Term Evolution
ACR	Absolute Category Rating
ANSI	American National Standards Institute
AVC	Advanced Video Coding
BER	Bit Error Rate
BIQI	Blind Image Quality Assessment Index
BLIINDS	Blind Image Integrity Notator using DCT Statistics
CIE	Commission Internationale de l'Eclairage
CIF	Common Intermediate Format
CPU	Central Processing Unit
CSF	Contrast Sensitivity Function
CT	Curvelet Transform
DCR	Degradation Category Rating
DCT	Discrete Cosine Transform
DMOS	Difference Mean Opinion Score
DNT	Divisive Normalization Transform
DSCQS	Double Stimulus Continuous Quality Scale
DSIS	Double Stimulus Impairment Scale
DWT	Discrete Wavelet Transform
EPSNR	Edge Peak Signal to Noise Ratio
FR	Full-Reference
GPU	Graphics Processing Unit
HD	High-Definition
HDTV	High-Definition Television
HSV	Hue-Saturation-Value
HEVC	High Efficiency Video Coding
HVS	Human Visual System
IP	Internet Protocol
IPS	In-Plane Switching

IPTV	Internet Protocol Television
ITU	International Telecommunication Union
JND	Just-Noticeable Distortion
JPEG	Joint Photographic Experts Group
JVT	Joint Video Team
LAR	Locally Adaptive Resolution
LCC	Pearson Linear Correlation Coefficient
LSVR	Linear Support Vector Regression
MAD	Most Apparent Distortion
MAE	Mean Absolute Error
MOS	Mean Opinion Score
MOVIE	Motion-based Video Integrity Evaluation
MPEG	Motion Picture Experts Group
MSE	Mean Squared Error
MS-SSIM	Multi-Scale SSIM
NIQE	Natural Image Quality Evaluator
NR	No-Reference
NSS	Natural Scene Saliency
NTIA	National Telecommunications and Information Administration
NTSC	National Television System Committee
OR	Outlier Ratio
PAL	Phase Alternating Line
PC	Pair Comparison
PC	Phase Congruency
PLR	Packet Loss Rate
PSNR	Peak Signal to Noise Ratio
QCIF	Quarter CIF
QoS	Quality of Service
QoE	Quality of Experience
RAM	Random Access Memory
RMSE	Root Mean Squared Error
RR	Reduced-Reference
SD	Standard Definition

SNR	Signal to Noise Ratio
SROCC	Spearman Rank-Order Correlation Coefficient
SSCQE	Single Stimulus Continuous Quality Evaluation
SSIM	Structural Similarity Index
SV	Support Vector
SVD	Singular Value Decomposition
SVM	Support Vector Machine
UoM	University of Manchester
UIQI	Universal Image Quality Index
VDP	Visible Differences Predictor
VIF	Visual Information Fidelity
VQA	Video Quality Assessment
VQEG	Video Quality Experts Group
VQM	Video Quality Metric
VQR	Video Quality Ratings
VSNR	Visual Signal-to-Noise Ratio

1. Увод

Последњих година дошло је до наглог развоја система за дигиталну обраду, пренос и приказ слике. Аналогна телевизија са коначном резолуцијом и карактеристичним шумом је замењена кристално јасним, великим форматом визуелних приказа, који се могу наћи у дневним собама, радњама, шеталиштима, кафићима, учионицама, ... Способност дигиталних камера, паметних телефона и таблет рачунара за снимање и приказ слика и видеа високе резолуције расте из дана у дан, а исто тако брзо расту и захтеви корисника. Друштвене мреже као што су Facebook, Youtube, Google+, Flickr итд. имају растући тренд преузимања, архивирања и прегледа слика и видеа различитих садржаја, трајања и степена квалитета. Четврта генерација (4G/LTE Fourth-Generation/Long-Term Evolution) бежичних мрежа омогућава мобилни веб видео велике брзине, IP телефонију, видео игре, мобилни HDTV, видео конференцију па чак и мобилну 3D TV [1]. Због погодности и слобода које пружају мобилни уређаји високих перформанси оптимизовани за видео пренос, бежични видео саобраћај је значајно порастао. Како количина видео саобраћаја наставља експоненцијално да расте, од посебног значаја је, како за провајдере тако и за крајње кориснике, избор начина испоруке садржаја доброг квалитета. Напредак у области дигиталног видеа поједини аутори називају "дигиталном видео револуцијом" [2].

Пре приказа посматрачу, слика, односно видео, пролази бројне фазе обраде и преноса при чему свака од фаза може унети деградације које утичу на квалитет крајњег приказа. Деградације слике и видеа у процесу снимања могу настати због карактеристика оптике, шума сензора, калибрације боје, времена експозиције, кретања камере и сл. Након снимања, слика или видео се најчешће кроз поступак компресије прилагођава пропусном опсегу система преноса. Увођење већег степена компресије спроводи се на штету веће деградације сигнала. Током преноса кроз комуникациони канал или током архивирања битске грешке такође могу довести до деградације. Коначно, уређаји за приказ крајњем кориснику могу утицати на субјективни утисак квалитета (лоша резолуција, лоша калибрација и сл.).

Како је човек посматрач и корисник највећег броја система за обраду слике, субјективна процена је најпоузданији метод за евалуацију квалитета визуелних

сигнала. Да би се добили статистички релевантни резултати, потребан је репрезентативан узорак посматрача, експерименте треба спровести у контролисаним условима, потребно је спровести додатну обраду одговора посматрача, што све заједно чини субјективну процену временски дугом, напорном и скупом. Додатно, субјективна евалуација квалитета није погодна не само због коначног броја слика или видеа који се користе у спровођењу тестова, већ и због тога што се најчешће техника мерења квалитета жели уградити у алгоритме обраде слике и видеа, како би се излазни квалитет максимизовао за расположиве ресурсе. На тај начин, субјективна евалуација је непрактична за примену у реалном времену, тј. у реалним апликацијама.

Да би се избегла субјективна процена, потребна је процедура за аутоматску, рачунску евалуацију квалитета слике/видеа. Процедура аутоматске процене назива се објективна процена и корисна је у многим применама где је потребно евалуирати визуелне ефекте који се јављају на сликама током снимања, обраде, компресије, преноса и архивирања. Објективне мере процене не захтевају опрему за спровођење тестирања, нема комплексне организације гледалаца и уз софтверску имплементацију време процене се смањује од дана и недеља до рада у реалном времену.

Основни циљ истраживања у области објективне процене квалитета је развој квантитативне мере која (алгоритамски, аутоматски) даје процене квалитета слика/видеа које су у доброј корелацији са просечном субјективном проценом (Mean Opinion Score, MOS). Идеална објективна мера за процену квалитета треба да буде применљива код различитих типова дисторзије, да квантитативно обухвати различите степене дисторзије и широк опсег садржаја изворног сигнала. Од овакве мере се очекује да да добре процене квалитета не само за дисторзије близу визуелног прага већ и за дисторзије које су далеко преко прага детекције. У практичним применама, због захтева за радом у реалном времену, осим набројаних услова битна је и рачунарска комплексност [3].

Објективне мере процене квалитета имају три врсте примене. Прва је праћење квалитета слике како би се контролисао систем за пренос. У том смислу, алгоритми процене квалитета се могу искористити за поправку квалитета слике кроз

перцептуалну оптимизацију процеса снимања, изменом брзине преноса, реалокацијом ресурса са циљем балансирања квалитета кроз мрежу, кроз пост-процесирање или комбиновањем оваквих приступа. Друго, могу се користити за одабир система и алгоритама за обраду слике. На пример, мере квалитета могу помоћи у избору једног од више система обраде слике у специфичним апликацијама. Треће, могу се уградити у систем за обраду слике како би се оптимизовали алгоритми и параметри који се користе. На пример, основни циљ видео компресије је остварити минимални степен дисторзије за задати битски проток, или прихватљиви ниво дисторзије остварити са што мањим битским протоком [3]. Комерцијална потреба за тачном и поузданом објективном мером процене квалитета слике/видео порасла је са новим дигиталним видео апликацијама и сервисима као што су Интернет видео, видео на захтев, видео телефонија, IPTV и сл.

У складу са комерцијалним аспектом квалитета апликација и сервиса појављују се два термина – квалитет сервиса, QoS (Quality of Service), и задовољство корисника, QoE (Quality of Experience). Традиционални QoS је релативно добро познат и истражен и односи се на перформансе мреже и преноса података. QoE је са друге стране још увек активно поље истраживања. Под овим термином се мисли на опис квалитета из перспективе корисника или потрошача (тј. гледаоца), са тежиштем на субјективно примећеном квалитету неког садржаја (субјективним осећајем корисника) [4].

Мере за процену квалитета могу се класификовати према различитим критеријумима, као што су: домен у коме је мера имплементирана (просторни или трансформациони), тип дисторзије код кога се примењује (шум, замрљање (blur), компресија, ...) и количина информација потребна за процену квалитета [1-4].

Према количини информација изворне (оригиналне, референтне) слике која се користи у поступку процене квалитета на пријемној страни (страни посматрача), објективне мере за процену квалитета слике се могу поделити у три категорије, без референцирања (No-Reference, NR), са потпуним референцирањем (Full-Reference, FR) и са делимичним референцирањем (Reduced-Reference, RR) [3-5]. Све три категорије имају предности и мана.

NR мере не захтевају познавање изворне слике и процена у потпуности зависи од људске перцепције тест слике (слике на пријемној страни). Перцепција слике је повезана са семантичким везама на слици, па је веома тешко дизајнирати идеалну NR објективну меру. Овакве мере се могу користити у свим апликацијама где се захтева мерење квалитета. Поуздана NR процена квалитета је тренутно могућа уколико су унапред позната знања о типу дисторзије (JPEG компресија, JPEG2000 компресија, замрљање), а у новије време се појављују и поуздане генералне технике за процену квалитета без референцирања [6]. За разлику од посматрача који на основу искуства веома лако може из садржаја слике одредити тип дисторзије/деградације, код NR техника постоји ризик да се садржај слике прогласи дисторзијом (нпр. шаховска табла да се интерпретира као блоковски ефекат). Такође, наш визуелни систем може веома лако разликовати слике доброг квалитета од слика лошег квалитета, односно знати шта је погрешно на њима, без гледања оригинала.

FR мере захтевају потпуно познавање информација изворне слике. У овом случају се систем за процену квалитета може посматрати као комуникациони систем у коме се изворна слика налази на предајној страни а тест слика (слика са деградацијом) на пријемној страни. Основу FR мера чини поређење двеју слика (изворне и тест слике) на нивоу пиксела, региона и/или фреквенцијских карактеристика [3]. Међутим, у неким реалним апликацијама познавање оригиналне слике на пријемној страни није могуће. Такође, FR мере за процену квалитета слике обично захтевају прецизна просторна усаглашавања (регистрацију). Код процене квалитета видео секвенци FR мере осим просторне захтевају и временску регистрацију. Изворна слика/видео обично долази са сензора високог квалитета и као таква захтева/заузима много више ресурса од слике/видеа након компресије. Због свега наведеног, FR мере процене квалитета се користе у лабораторијским тестирањима за избор техника обраде слике и видеа.

RR технике се налазе између претходне две категорије и дизајниране су да дају практична решења у процени квалитета док се задржава тачност процене. Код ових техника се само најбитније изворне информације шаљу од предајне ка пријемној страни. Како у овом случају додатна количина информација није велика, захтеви по питању пропусног опсега канала се значајно не мењају. У основи RR мера

је статистичка анализа промењених/модификованих обележја изворне слике и одговарајућих обележја деградираних слика. Главно интересовање је детекција битних а довољних обележја како би се идентификовале разлике између слика. И за RR мере важи услов регистрације али он није толико строг као код FR техника.

RR и NR алгоритми процене квалитета могу се користити као посредници (агенти) у преносу података кроз мреже за уградњу у рутерима, сет боксовима, паметним телефонима, таблетима, лаптоповима. Могу се користити као примарни алат QoS кроз који се могу добијати повратне информације за адаптацију извора и механизме контроле алокације ресурса, кодовање извора и друге мрежне параметре [6].

Објективне мере процене квалитета слике и видеа нашле су бројне примене. Већина примена односи се на ситуације у којима је потребно проценити квалитет модификоване верзије слике/видеа, али се користе и у ситуацијама где се поређење преводи у нешто што није квалитет сигнала, нпр. скуп мерних података или одлука.

Тако се процена квалитета слике/видеа користи у:

- стеганографији [7],
- уградњи дигиталног воденог жига [8],
- процени квалитета сједињавања слика [9, 10],
- потискивању шума [11, 12],
- побољшању визуелног квалитета слике [13],
- процени успешности супер-резулционих техника [14],
- процени квалитета резолуционо деградираних слика [15],
- конверзији динамичког опсега слика [16],
- ретаргетирању [17],
- кодовању [18, 19],
- даљинском осматрању [20],
- видео надзору [21],
- препознавању циљева [22],
- праћењу покретних објеката [23],

- класификацији [24],
- анализи квалитета сервиса [25], ...

Значај објективне процене квалитета слике и видеа може се уочити само кроз набројане примере примене (са списком примена који није коначан).

Постојеће објективне мере често нису у стању да адекватно процене цео опсег квалитета па су су потребне мере које адекватно евалуирају комплетан опсег, а нарочито сигнале нижег квалитета. Поред тога, тренутно најбољи модели за објективну процену квалитета видео сигнала користе неки од мултирезолуционих домена и сложене моделе визуелног система човека. Резултат су изузетно сложени алгоритми процене квалитета, непрактични за рад у реалном времену (видео секвенце од 10 секунди анализирају неколико десетина минута или часова). Стога постоји јасна потреба за ефикасним мерама квалитета које ће бити способне да раде у реалном времену (времену трајања сигнала).

Предмет истраживања представљеног у овој дисертацији јесу алгоритми за потпуну и делимично референцирану објективну процену квалитета (FR и RR алгоритми) и верности визуелних сигнала, слике и видеа, модификованих компресијом, губицима у преносу и другом обрадом. Развијене су робусне, рачунски ефикасне објективне мере (алгоритми) за аутоматску процену верности и квалитета слика и видеа у условима када је потпуно или делимично познат оригинални сигнал. Развијене мере дају нумерички скор квалитета слике односно видеа који је у доброј корелацији са просечном субјективном проценом квалитета истих сигнала. Мере су применљиве и тачне у присуству различитих типова деградације изворних сигнала, квантитативно обухватају широк опсег степена деградације и имају могућност одређивања просторне расподеле (локализацију) деградација.

Предложени алгоритми се надовезују на технике које у процени квалитета слике и видеа користе очување информација о структури. Структурне информације из сигнала се представљају кроз локалне параметре градијентних поља које се директно пореде између оригиналних и модификованих слика и видеа. Разлике између градијената су преточене у процене локалног квалитета дефинисањем биолошки инспирисаних модела за процену очувања верности и квалитета структура

у сигналу. Очување визуелних информација се анализира у просторном (на нивоу слика) и просторно-временском домену (на нивоу секвенци).

Посебни модели су дефинисани за мерење перцептуалних разлика у информацијама о боји, за које тренутно не постоје тачни модели. Познати концепти различитих утицаја на перцепцију квалитета, као што је "најнижи скор има највећи значај" су анализирани као оквир за интеграцију локалних процена квалитета у глобалне процене за целе слике односно видео.

Велика већина доступних објективних мера за процену слике и видеоа, као и база сигнала са субјективним проценама садрже слике и секвенце високе просторне резолуције намењене системима преноса са великим битским брзинама као што је телевизија [26]. Међутим, неке од примена видео преноса врше се са значајно ограниченим битским протоцима, као што је даљинско осматрање, где су ефекти на верност и квалитет снимка значајно другачији од оних који се срећу на већим битским брзинама. Због тога је истраживање обухватило спровођење субјективних тестова на видеу ниског квалитета (за ниске битске брзине). Резултати субјективних тестова су искоришћени за калибрацију и валидацију развијених објективних мера.

За валидацију предложених мера објективне процене квалитета коришћене су јавно доступне базе слика (LIVE, CSIQ, Toyama, IVC) и видео секвенци (VQEG, LIVE Video, LIVE Wireless, EPFL-PoliMI, FUB), са доступним субјективним скоровима. У расположивим базама постоји више од 1600 тест слика и више од 400 тест видео секвенци. Субјективним тестовима у оквиру истраживања, у сарадњи са Универзитетом у Манчестеру, додато је још преко 300 релевантних, субјективно процењених видео секвенци. Новоформиране базе садрже секвенце добијене при компресији изворних секвенци за мале битске брзине.

Алгоритми објективне процене квалитета и пратећи софтвер развијени су у програмском пакету Matlab[®]. Алгоритми објективне процене других аутора, са чијим мерама је вршено поређење, преузимани су у изворном облику.

За одређивање успешности мера, тј. поређење субјективних и објективних скорова квалитета, коришћене су мере које су прихваћене од стране ИТУ (кофицијент корелације, корелација рангова, средња квадратна грешка, средња апсолутна грешка и стандардна девијација процена) [5].

У складу са наведеним предметом истраживања било је потребно: (1) извршити прикупљање доступних база слика и видео секвенци, (2) развити мере за објективну процену квалитета слике и видеа у условима нарушавања квалитета, тј. деградације, (3) формирати базу оригиналних и тест видео секвенци и спровести субјективне тестове за оцењивање њиховог квалитета и (4) развити софтвер за анализу перформанси мера објективне процене, кроз поређење са субјективним проценама.

Дисертација је организована у седам поглавља. Након увода, други део дисертације кроз преглед релевантне литературе разматра неке од посебних аспеката процене квалитета слике и видеа. У трећем делу рада предложена је објективна мера за процену квалитета слика заснована на очувању информација о градијенту. Перформансе предложене мере су анализирани на различитим базама слика, за подскупове слика и на глобалном плану. Степен слагања (линеарна корелација) добијених објективних скорова и субјективних импресија квалитета зависи од базе слика са којом се ради и креће се на глобалном плану (нивоу комплетне базе) од 85% (CSIQ база слика) до 94% (LIVE база слика). Предложена мера се по перформансама налази међу прве три од седам анализираних објективних мера процене. Побољшање стандардних мера процене квалитета слика кроз увођење мапа субјективног значаја, квалитетом вођену интеграцију и адитивни модел сједињавања информација презентовано је у четвртом поглављу рада. Интеграција предложених мапа субјективног значаја са предложеном објективном проценом квалитета довела је до побољшања перформанси објективне мере. Добитак зависи од базе слика са којом се ради и креће се од 0.5% до 2%. Међутим, утицај предложених мапа субјективног значаја на перформансе мере у оквиру подскупова слика из различитих база није једнозначан, тј. није увек доводио до побољшања перформанси. Коришћењем дела најнижих објективних скорова квалитета (уместо свих локалних скорова) потврђена је хипотеза да су у субјективној евалуацији квалитета најбитнији региони са највећим степеном деградације. Адитивни модели сједињавања информација о очувању обележја изворне слике у тест слици су предложени као алтернатива стандардним мултипликативним формама. Адитивни модели су комбиновани са квалитетом вођеном интеграцијом, што је заједно донело побољшање перформанси објективних мера.

Предложена објективна мера процене квалитета слика је проширена за процену квалитета видеа и презентована у петом делу рада, заједно са описом спроведених субјективних тестова. Перформансе објективне мере су анализирани на седам база видео секвенци са доступним субјективним импресијама квалитета. Предложена мера процене квалитета видео сигнала се показала веома погодном за процену квалитета секвенци са малим битским брзинама. Од седам анализираних база видео секвенци мера је дала најбоље резултате на четири базе, за битске брзине до 576 kbps.

Током спровођења једног од субјективних тестова вршено је снимање кретања зенице ока посматрача. Добијени снимци су преведени у координатни систем видеа и искоришћени за пондерисање објективних скорова квалитета и за процену квалитета видеа са редукованим референцирањем. Добијени резултати су представљени у шестом делу рада. Коришћењем експерименталних података показано је да се преносом дела изворног сигнала на пријемну страну (свега 4%) врши редукација информација потребних за евалуацију квалитета тест сигнала, без нарушавања перформанси објективне мере. Наиме, перформансе мере са референцирањем на 4% изворног сигнала се не разликују од перформанси мере са потпуним референцирањем на изворни сигнал (100% изворног сигнала).

Преглед најзначајнијих резултата, најбитнијих закључака и даљих праваца рада дати су у последњем, седмом делу рада.

2. Посебни аспекти процене квалитета слике и видеа

У овом делу дисертације су кроз преглед најзначајније литературе разматрани различити аспекти процене квалитета слике и видеа. У обзир су узети аспекти који су најбитнији са становишта циљева израде дисертације. Део литературе покрива више разматраних аспеката, тако да су радови разврстани у област којој тежишно припадају.

2.1. Субјективна евалуација квалитета

Процедуре и стандарди за субјективну процену квалитета говора, аудио и видео сигнала постоје већ дужи низ година. Због обиља дигиталних аудио и видео садржаја, спровођење субјективних експеримената за мерење квалитета постало је релативно једноставно. Субјективни тестови за визуелну процену квалитета формализовани су у ITU препорукама [27] и [28], које сугеришу стандардне услове посматрања, критеријуме за избор посматрача и тест материјала, процедуре процене и методе анализа података. Прва препорука односи се на телевизију док се друга односи на мултимедијалне апликације. Експерименталне поставке и услови посматрања се разликују, али је процедуре потребно размотрити за било који експеримент.

Препоруке дефинишу неке од најчешће коришћених процедура за субјективну евалуацију квалитета:

- Double Stimulus Continuous Quality Scale (DSCQS), где субјекти евалуирају парове секвенци кратког трајања које се састоје од тест видеа и одговарајућег референтног видеа.
- Double Stimulus Impairment Scale (DSIS), која је позната и као Degradation Category Rating (DCR), где субјекти евалуирају степен усаглашености тест видеа у односу на познати референтни видео.
- Single Stimulus Continuous Quality Evaluation (SSCQE), где субјекти посматрају програм трајања од 20 до 30 минута и у континуитету евалуирају текући квалитет.

- Absolute Category Rating (ACR), спада у Single Stimulus методе, где субјекти евалуирају сваки од тест видеа без поређења са одговарајућом референцом.
- Pair Comparison (PC), где се пореде парови тест видео сигнала исте сцене добијени при различитим условима (са разним комбинацијама) и субјекти дају предност једној од побуда унутар пара.

Наведене процедуре тестирања имају различите примене. Такође, постоје и разлике у скалама оцењивања, које могу бити континуалне и дискретне, са нумеричким или категоријским оцењивањем и са различитим бројем нивоа – од 5 нивоа до 100 нивоа.

Изназ субјективних експеримената су оцене квалитета посматрача, које се након теста усредњавају и презентују кроз просечну субјективну оцену (MOS). MOS је најчешће коришћени метод генерализације субјективних скорова датих од стране више независних посматрача квалитету неког сигнала – аудио и/или видео сигнала и користи се као полазна основа за развој објективних мера процене. То је проста аритметичка средина скорова посматрача (учесника теста) датих неком сигналу:

$$MOS_i = \frac{1}{N_S} \sum_{n=1}^{N_S} SQ(n,i) \quad (1)$$

где су:

i – индекс (ознака) видеа са деградацијом у субјективном тесту,

$SQ(n,i)$ – субјективни квалитет дат од n -тог посматрача i -тој секвенци и

N_S – број посматрача у субјективном тесту.

Просечна субјективна оцена има исти динамички опсег као и скала квалитета коришћена током тестирања. Може се рећи да је MOS "демократска" мера која сваки субјективни скор третира једнако и стварно представља средње (просечно) мишљење. Добијене MOS вредности се могу директно користити за поређење са вредностима добијеним применом објективних мера процене.

Једна од алтернатива средњој вредности квалитета је медијана добијених субјективних импресија квалитета. То је вредност која дели доступне субјективне скорове на доњу и горњу половину, тј. може се рећи да је то средњи скор. Иако је

медијана често блиска средњој вредности, за скорове који нису униформно расподељени, медијана може бити бољи естиматор стварног квалитета. Без обзира на ово запажање, MOS је и даље универзалан начин генерализације субјективних скорова.

Како се MOS вредност добија од неколико појединачних скорова квалитета, процени квалитета може се придружити извештајни ниво статистичке неизвесности/неодређености. Уколико постоје значајне флукуације субјективних скорова ова неодређеност је велика. Непоузданост субјективних скорова може се мерити на разне начине али се најчешће користе стандардна девијација, варијанса и стандардна грешка.

Стандардна девијација се одређује на основу појединачних субјективних скорова видеа i , SQ_i (индекс n који одговара посматрачу је игнорисан) и просечне субјективне оцене тог видеа:

$$\sigma_i = \sqrt{E[SQ_i^2] - (E[SQ_i])^2}$$
$$\sigma_i = \sqrt{E[SQ_i^2] - MOS_i^2}$$
(2)

где је $E[X]$ очекивана вредност, односно средња вредност променљиве X .

Стандардна грешка процене MOS се одређује из стандардне девијације субјективних скорова и узима у обзир број мерења при одређивању мерене вредности. У овом случају то је број посматрача који су учествовали у субјективним тестовима, N_S :

$$SE_i = \frac{\sigma_i}{\sqrt{N_S}}$$
(3)

Стандардна грешка се најчешће приказује заједно са средњом вредношћу као њено позитивно и негативно одступање ($MOS_i \pm 2 \cdot SE_i$) – интервал поузданости, али она не говори о разлозима одступања.

Усклађеност скорова квалитета се разликује од посматрача до посматрача. Како су посматрачи слободни у избору опсега апсолутног квалитета за очекивати је

да се минималне и максималне вредности придружене видео секвенцама разликују од посматрача до посматрача. Великодушни посматрачи додељују релативно високе скорове док строжији посматрачи већину секвенци оцењују скоровима из доњег дела доступне скале квалитета. Иако апсолутни опсег субјективних скорова не утиче на ранг оцењиваних видеа, противречни опсези могу утицати на MOS.

Апсолутни опсези субјективних скорова прикупљених током теста могу значајно варирати. Иако релативне вредности и ранг скорова одсликавају реални квалитет приказаног видеа, бројни истраживачи користе нормализацију субјективних скорова пре одређивања MOS и других статистичких мера. Нормализација се најчешће спроводи на нивоу субјекта, где се врши манипулација индивидуалних оцена посматрача како би се искористио комплетан опсег субјективне скале квалитета.

Стандардни приступ је глобална статистичка нормализација где се средња вредност скорова које је посматрач дао (за све видео секвенце), поравнава са одређеном фиксном вредношћу, тј. средином скале квалитета. Стандардна девијација његових субјективних скорова се такође нормализује на фиксну вредност. Оваквим приступом се нарушава апсолутни опсег квалитета свих субјеката са циљем повећања релативних разлика. Иако се на овај начин добија сажетији сет субјективних скорова, није јасно каква манипулација скорова ће побољшати субјективну процену квалитета.

Осим кроз MOS, резултати субјективних тестова могу се дати и кроз диференцијалне MOS скорове (Difference Mean Opinion Score, DMOS), а који се добијају одузимањем просечне субјективне оцене видеа са дисторзијом од просечне субјективне оцене референтног видеа:

$$DMOS_i = MOS_i^{original} - MOS_i^{distorted} \quad (4)$$

Кроз субјективне тестове истраживачи су дошли до следећих закључака [29]: (1) слике веома доброг и јако лошег квалитета једноставније је субјективно евалуирати, (2) време одзива (процене) за слике које је теже евалуирати је веће, (3) постоји веза између поузданости посматрача и стандардне грешке и (4) поузданост посматрача се може проценити на основу расподеле скорова и времена одзива.

2.2. Значај VQEG групе у области процене квалитета

VQEG (Video Quality Experts Group) група је основана 1997. године од чланова ITU-T и ITU-R група. Група се састоји од експерата у области процене квалитета видеа из истраживачких центара, универзитета, владиних организација, организација за развој стандарда и индустрије. Чланство у VQEG групи је на добровољној основи. Генерални циљ VQEG је унапређење области процене квалитета видеа кроз евалуацију објективних мера квалитета и истраживања нових метода субјективне евалуације.

VQEG планови и тестови валидације објективних мера се доносе кроз заједнички рад предлагача алгоритама објективне процене квалитета и корисника, што је све подржано са независним лабораторијама у којима се врше тестирања.

Тестови валидације објективних мера обухватају и креирање видео база и спровођење субјективних експеримената. Тест секвенце нису познате предлагачима објективних мера. Резултати субјективних тестова се користе за оцену успешности предикције објективних мера квалитета. За ту сврху се користе бројни статистички критеријуми.

Први тест (FR-TV Phase I) VQEG групе је завршен 2000. године и односио се на мере са потпуним референцирањем намењене за дистрибуцију видеа стандардне дефиниције (Standard Definition, SD). Углавном су коришћене MPEG-2 кодоване секвенце. На жалост, резултати прве фазе били су неуверљиви; перформансе објективних модела и вршног односа сигнал шум (Peak Signal to Noise Ratio, PSNR) били су статистички еквивалентни, тако да ни један алгоритам није предлошен. Допринос истраживачкој заједници је јавно доступна база видео секвенци.

Наставак истраживања (FR-TV Phase II) за SD TV апликације завршен је 2003. године. Друга фаза је била усмерена на секундарну дистрибуцију дигиталног видеа и већи степен дисторзија. Коришћене су нове изворне секвенце и нови тест услови, тако да је укупно направљено 128 тест секвенци. Најбоља мера је достигла степен слагања од 94% са MOS, чиме су значајно превазиђене перформансе PSNR мере, код које је степен слагања 70%. Резултати ове фазе били су основа за две препоруке [30, 31].

Обе FR-TV фазе су биле посвећене мерама са потпуним референцирањем и усмерене на MPEG-2 компресију за дигиталну TV дифузију, тј. нису обухваћена IP оштећења. У 2008. години VQEG је комплетирао евалуацију мера у мултимедијалном сценарију (MM Phase I) [32], који је посвећен мањим битским брзинама и мањим димензијама слика, као и већем броју кодека и услова преноса. Тест је обухватио видео сервисе који се испоручују са брзинама до 4 Mb/s, са фокусом на широкопојасни Интернет и мобилни видео пренос. MM Phase I сет тестова је коришћен за валидацију мера са потпуним референцирањем, редукованим референцирањем и без референцирања. Као резултат публиковане су четири нове препоруке за процену квалитета мултимедије. Препорукама се дефинишу FR мере [33, 34] и RR мере [35, 36], док NR мере нису дале задовољавајуће резултате.

Редуковано референцирање и слепа процена квалитета SD телевизије (625-линија и 525-линија) били су предмет разматрања RR/NR-TV пројекта комплетираног јуна 2009. године. Сви NR модели су одбачени, док је ITU одлучио да је тачност дела RR мера довољна за стандардизацију [37]. Резултати су такође послужили за дефинисање референтног алгорита одређивања вршног односа сигнал шум [38].

Прва фаза HDTV пројекта односила се на валидацију мера објективне процене квалитета HD видеа. Мере процене без референцирања нису задовољиле постављене захтеве док су један FR модел и један RR модел задовољили да буду предмет стандардизације. Део видео материјала коришћеног у истраживањима је доступан, а резултати су прихваћени у две препоруке [39, 40].

Највећи број препорука проистеклих из резултата и извештаја VQEG групе публиковала је ITU-T Study Group 9. Додатно, она развија стандарде за субјективну евалуацију квалитета као и сродне захтеве, као што су калибрација и упаривање видео секвенци [41].

VQEG има централну улогу у независној валидацији објективних мера процене. Међутим, релативно спор прогрес не прати индустријске захтеве, тако да стандардизовани модели застаревају. Додатно, неки аутори сматрају да алгоритам објективне процене није потребно стандардизовати јер је његов примарни захтев мерење квалитета са поузданом тачношћу [42].

Поред VQEG групе постоје и други сродни комитети који разматрају субјективне и објективне мере процене квалитета – ATIS IIF (ATIS IPTV Interoperability Forum) [43], DSL Forum [44] и Video Services Forum [45].

Текући пројекти VQEG групе односе се на 3DTV, процену квалитета аудио и/или видеа HD резолуције, процену квалитета видеа са великим динамичким опсегом, хибридне мере процене квалитета, мониторинг аудио и визуелног квалитета, процену квалитета видеа са специфичним циљем, развој нових методологија субјективне процене квалитета и објективних мера интерактивних комуникационих сервиса, ултра HD резолуцију и процену квалитета слике/видеа у потрошачкој индустрији.

2.3. Једноставне објективне мере процене квалитета слика

Најчешће коришћене објективне мере за процену квалитета/дисторзије су средња квадратна грешка (Mean Squared Error, MSE), однос сигнал шум (Signal to Noise Ratio, SNR) и вршни однос сигнал шум (PSNR), али су ове мере често критиковане јер не дају резултате који су у потпуности у складу са субјективним проценама [46]. Међутим, ове мере се и даље користе јер су једноставне за имплементацију (хардверску и софтверску), не зависе од услова под којима се процена спроводи, не зависе од посматрача, имају јасно физичко тумачење и често дају решења оптимизационих проблема у затвореној форми.

За процену квалитета слике \mathbf{D} , тј. за процену одступања слике \mathbf{D} од референтне слике \mathbf{R} , средња квадратна грешка и однос сигнал шум се дефинишу као:

$$MSE(\mathbf{D}, \mathbf{R}) = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M (R(n, m) - D(n, m))^2 = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M e^2(n, m) \quad (5)$$

$$SNR(\mathbf{D}, \mathbf{R}) = 10 \log_{10} \frac{\sigma^2}{MSE(\mathbf{D}, \mathbf{R})} \text{ [dB]} \quad (6)$$

где је σ^2 варијанса референтне слике. Димензије слика су $N \times M$ пиксела.

Много погоднији за практичну примену је вршни однос сигнал-шум, који се такође изражава у dB. За представу осветљености пиксела са p бита ($L=2^p$ нивоа сивог), вршни однос сигнал-шум је дат као:

$$PSNR(\mathbf{D}, \mathbf{R}) = 10 \log_{10} \frac{L^2}{MSE(\mathbf{D}, \mathbf{R})} \text{ [dB]} \quad (7)$$

где је L динамички опсег слике (типично $[0,255]$).

Поређење (5), дато у форми разлике два сигнала, $e(n,m)$, се може користити и у општијем облику:

$$d_p(\mathbf{D}, \mathbf{R}) = \left(\frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M |e(n,m)|^p \right)^{1/p} \quad (8)$$

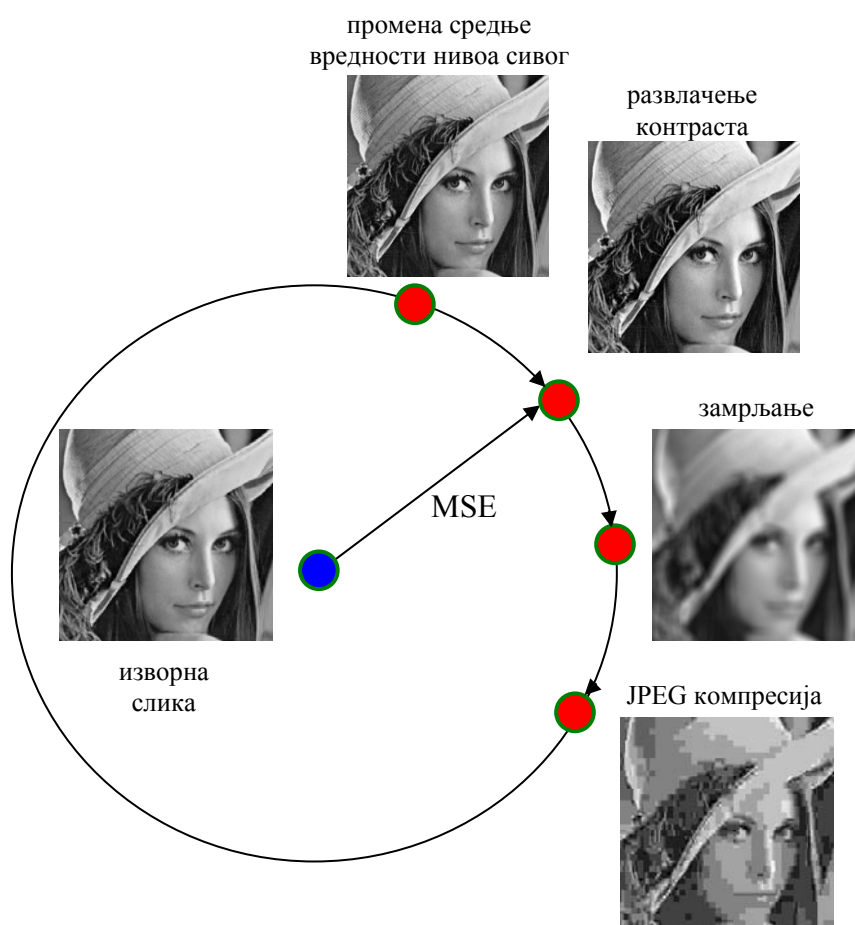
познатом као Минковски техника (Minkowski error pooling formula). Ова техника је независна од структурног садржаја слике, због чега се MSE, SNR и PSNR називају квази-стандардним мерама сличности [4]. Истраживачи који се баве мрежним QoS такође користе једноставне мере за квантификовање грешака у преносу, као што су степен битске грешке (Bit Error Rate, BER) или пакетске грешке (Packet Loss Rate, PLR). Наведене мере не узимају у обзир садржај, тј. значење, а самим тим ни визуелни значај пиксела или пакета.

2.4. Сложеније објективне мере процене квалитета слика

Индекс структурне сличности (Structural Similarity Index, SSIM) је померио процену квалитета са нивоа пиксела на ниво структуре [47, 48]. У позадини ове мере је нова филозофија у процени квалитета слика и видеа: *Основна функција људског ока је издвајање структурних информација из видног поља и људски визуелни систем је прилагођен овом захтеву. Због тога ће се мерењем структурних разлика добити добра апроксимација дисторзије слике* [47, 48]. Деградација слике посматра се као (визуелно приметни) губитак информација о структури, за разлику од традиционалних приступа код којих се посматра разлика интензитета на нивоу пиксела (MSE, SNR, PSNR). Предност новог приступа илустрована је кроз пример са Сл. 2.1.

Приказане слике имају значајне разлике у погледу визуелног квалитета а исту вредност MSE. Минковски технике веома тешко могу објаснити зашто се развлачењем контраста добија слика доброг субјективног квалитета, односно зашто

је субјективна разлика у односу на референтну слику веома мала. То је лако разумети са новом филозофијом јер су скоро све структурне информације очуване, чак се кроз инверзну линеарну трансформацију оригиналне информације могу вратити (осим у веома тамним или светлим регионима). Са друге стране структурне информације изворне слике су неповратно изгубљене у JPEG компримованој слици и замрљаној слици, па се за њих мора добити и мања вредност квалитета него код слике са модификацијом контраста.



Сл. 2.1 Изворна слика "Lena" и њене модификације настале: (а) променом средње вредности нивоа сивога (mean-shift), (б) развлачењем контраста, (в) замрљањем и (г) JPEG компресијом.

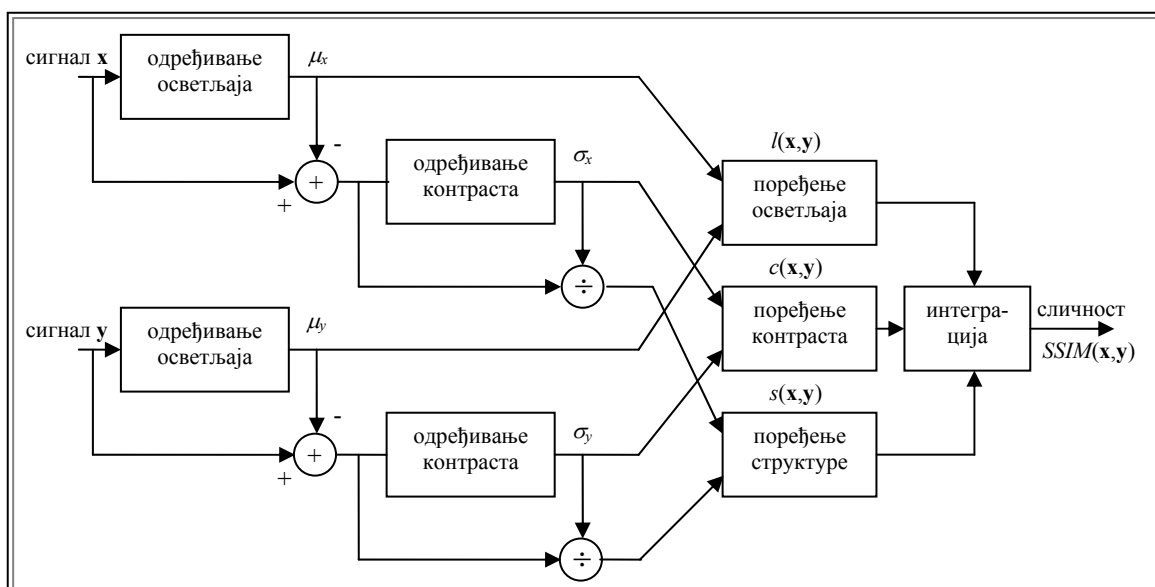
Индекс структурне сличности описује квалитет слике поређењем локалних сличности осветљаја (луминансе), контраста и структуре изворне и тест слике, као што је илустровано на Сл. 2.2. Нека су x и y делови референтне и тест слике који су међусобно просторно усаглашени – регистровани. Како је референтна слика

"идеалног квалитета", кроз поређења се квантитативно мери квалитет тест сигнала. Прво се врши поређење осветљаја два сигнала. За сигнале дате у дискретној форми $\mathbf{x}=\{x_i|i=1,2,\dots,N\}$ и $\mathbf{y}=\{y_i|i=1,2,\dots,N\}$ осветљај се одређује као средња вредност интензитета:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i, \mu_y = \frac{1}{N} \sum_{i=1}^N y_i \quad (9)$$

Поређење осветљаја, $l(\mathbf{x},\mathbf{y})$, зависи од μ_x и μ_y :

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \quad (10)$$



Сл. 2.2 Блок дијаграм одређивања индекса структурне сличности два сигнала.

Након првог корака елиминишу се средње вредности интензитета сигнала. У дискретној форми, резултујући сигнал $\mathbf{x}-\mu_x$ одговара пројекцији вектора \mathbf{x} на раван дефинисану са:

$$\sum_{i=1}^N x_i = 0 \quad (11)$$

За процену контраста користи се стандардна девијација која је у дискретној форми дата са:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2}, \sigma_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2} \quad (12)$$

Поређење контраста, $c(\mathbf{x}, \mathbf{y})$, се своди на поређење σ_x и σ_y :

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (13)$$

Структурно поређење се спроводи након одузимања средње вредности интензитета позадине и нормализације. Сигнали су нормализовани (подељени) са сопственом стандардном девијацијом, тако да имају јединичне стандардне девијације. Јединични вектори $(\mathbf{x}-\mu_x)/\sigma_x$ и $(\mathbf{y}-\mu_y)/\sigma_y$ придружују се структурама слика које се пореде. Корелација (унутрашњи производ) ових сигнала је једноставан и ефикасан начин за мерење структурне сличности. У дискретној форми σ_{xy} се одређује као:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (14)$$

Структурно поређење се спроводи над нормализованих сигналама $(\mathbf{x}-\mu_x)/\sigma_x$ и $(\mathbf{y}-\mu_y)/\sigma_y$:

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (15)$$

На крају, три поређења се комбинују у збирну меру сличности сигнала \mathbf{x} и \mathbf{y} :

$$SSIM(\mathbf{x}, \mathbf{y}) = f(l(\mathbf{x}, \mathbf{y}), c(\mathbf{x}, \mathbf{y}), s(\mathbf{x}, \mathbf{y})) = l(\mathbf{x}, \mathbf{y}) \cdot c(\mathbf{x}, \mathbf{y}) \cdot s(\mathbf{x}, \mathbf{y}), \quad (16)$$

која је стандардно дата у облику производа [48].

Треба запазити да су ова три члана међусобно концептуално независна јер прва два члана зависе од осветљаја и контраста слика које се пореде и потпуна промена осветљаја или контраста било које слике нема утицаја на трећи члан. Геометријски, $s(\mathbf{x}, \mathbf{y})$ одговара косинусу угла између вектора $\mathbf{x}-\mu_x$ и $\mathbf{y}-\mu_y$ који не зависи од дужине вектора.

Од локалних вредности добија се мапа квалитета која даје више информација о деградацији комплетне слике. Локалне вредности квалитета се интегришу у глобалну процену квалитета.

Приступ од локалног ка глобалном квалитету уведен је из више разлога. Као прво, статистичке карактеристике слике су често просторно нестационарне. Друго,

дисторзије слике могу, али и не морају зависити од локалне статистике слике. Треће, посматрач у једном тренутку са типичних даљина посматрања може видети само део слике у високој резолуцији.

Изворни индекс структурне сличности се може сврстати у технике имплементиране у просторном домену. Интересантни приступи просторног домена могу се наћи и у [49-52], где се поређење сигнала спроводи кроз теорију фази скупова [49], сингуларну декомпозицију (Singular Value Decomposition, SVD) [50], дискретне ортогоналне моменте нижег реда [51] и пирамидалну Гаусову мултивеличинску репрезентацију слике [52].

У трансформационом домену користе се Фуријеова трансформација [53-57], вејвлет трансформација (Discrete Wavelet Transform, DWT) [58-60], косинусна трансформација (Discrete Cosine Transform, DCT) [18], Чебишевљева трансформација [61], Riesz трансформација [62] и сл.

Слика са деградацијом се у [53] посматра као оригинална слика код које је изршена линеарна промена фреквенције и којој је додат адитивни шум. Како су ефекти дисторзије фреквенције и убацивања шума међусобно независни, аутори су ова два извора деградације посматрали независно и анализирали како сваки од њих утиче на визуелни систем човека. Утицај шума је анализиран кроз контрастну пирамиду док је девијација фреквенције разматрана у фреквенцијском (дводимензионална Фуријеова трансформација – 2D Fourier Transform) домену уз интеграцију са моделом фреквенцијског одзива визуелног система човека. Фуријеова трансформације је искоришћена за процену квалитета кроз сличност фаза [54] и спектрални остатак [55]. Спектрални остатак (резидуал) се добија од логаритамског спектра слике и представља иновациони део слике након елиминације статистички редувантних компоненти. Поређење амплитуда и фаза дводимензионалне дискретне Фуријеове трансформације слика представља основу мере [56], где је додатно кроз стратегију неуниформног пондерисања фреквенцијских компоненти узета у обзир нелинеарност визуелног система човека. Занимљив је закључак из [56] којим је потенцирано да у процени квалитета више информација носи очување информација о фази него очување информација о амплитуди. Аутори из [57] подржавају теорију да визуелни систем човека користи више стратегија у процени

квалитета слике, тј. сугеришу да једнострука стратегија (нпр. детекција визуелних разлика и издвајање структурних информација) процене квалитета можда неће бити довољна. Садржај квалитетних слика је јасно видљив и визуелни систем човека гледа "кроз" слику тражећи дисторзије (стратегија детекције). Сlike са јасно уочљивим дисторзијама визуелни систем човека посматра тражећи садржај слике (стратегија појављивања), тј. гледа се "кроз" дисторзије. За моделовање одвојених стратегија предложена је MAD техника (Most Apparent Distortion, MAD) [57]. Фуријеова трансформација се користи у обе стратегије – за имплементацију маскирања контраста у стратегији детекције и за декомпозицију слике на различите величине и различите оријентације у стратегији појављивања.

Две најпознатије технике које користе вејвлет трансформацију у процени квалитета слике су [58] и [59], а познате су под називима VIF (Visual Information Fidelity, [58]) и VSNR (Visual Signal-to-Noise Ratio, [59]). Квантитативни опис квалитета слике из [58] заснован је на статистици слика природних сцена (Natural Scene Statistics, NSS) и полази од премисе да је визуелни систем човека развијен на основу статистичких особина природног окружења. У складу са том претпоставком квалитет слике са дисторзијом може се добити на основу количине информација која она носи о оригиналу. Квалитет слике се одређује на основу заједничке информације коефицијената из одговарајућих вејвлет подопсега оригинала и тест слике.

VSNR мером се оцена квалитета слике добија кроз две фазе. У првој фази се кроз вејвлет модел визуелног маскирања и сумирања одређују прагови детекције дисторзије слика природних сцена (кроз контраст). Уколико дисторзија постоји али је испод прага детекције, може се рећи да је тест слика верна оригиналу ($VSNR=\infty$) и даља анализа није потребна. Уколико је интензитет дисторзије изнад прага детекције, пролази се кроз другу фазу коришћењем визуелних карактеристика нижег (контраст) и средњег нивоа (глобална импресија). Ове две особине су моделоване кроз Еуклидску дистанцу у дисторзија-контраст простору мултивеличинске вејвлет декомпозиције. Вејвлет трансформација се користи и у приступу представљеном у [60]. Разлика изворне и тест слике посматра се кроз губитак детаља и додатна слабљења. Губитак детаља односи се на губитак корисних визуелних информације које утичу на видљивост изворних садржаја, док додатна слабљења репрезентују

редундантне визуелне информације чија ће појава у тест слици одвратити пажњу посматрача од корисних садржаја узрокујући незадовољство приказаним садржајем (блоковски ефекти (blocking), ефекти прстена (ringing) и сл.).

Неке од техника процене квалитета представљају комбинацију обележја изведених из просторног и фреквенцијског домена. Мера предложена у [63] полази од чињенице да визуелни систем човека разуме слику углавном на основу њених обележја ниског нивоа (low level), па се користе комплементарна обележја – фазна конгруенција (Phase Congruency, PC) која је инваријантна на контраст и амплитуда градијента (због утицаја контраста на људску перцепцију слике). Додатно је мапа фазне конгруенције искоришћена за пондерисање локалних скорова квалитета. Од избора обележја значајно зависе резултати објективне процене квалитета [64]. Иако је вршни однос сигнал шум познат као лош предиктор квалитета, показало се да је веома ефикасан за процену квалитета слика са адитивним Гаусовим шумом. Аутори из [64] су показали да PSNR може бити ефикасан и за друге типове дисторзије ако се користе друга визуелна обележја, односно уколико се мери разлика вредности локалних обележја (осветљаја, боје, контраста, амплитуда ивица и оштрине слике).

До квантитативног скорa квалитета уз потпуно референцирање може се доћи и кроз различите врсте машинског обучавања [65-67]. Приступ предложен у [65] користи SVD декомпозицију и SV (Support Vector) регресију. Обележја изведена из дводимензионалне мел-кепстралне анализе сигнала (2D mel-cepstrum), [66], користе се за мапирање вектора обележја у перцептуални скор квалитета. Изведена обележја представљају структурне информације које су пресудне за процену квалитета а за мапирање обележја у скор квалитета такође се користи SV регресија. За сврставање тест слике у једну од 5 класа квалитета (од најлошијег (1) до најбољег (5)) користи се SVM (multi-Support Vector Machine) класификатор [67]. Обележја на основу којих се доноси одлука су изведена из просторног и фреквенцијског домена. Просторна обележја (13 карактеристика) су изведена по узору на индекс структурне сличности. Обележја из фреквенцијског домена (12 обележја) су изведена из пирамидалне трансформације (Steerable Pyramid Transform) како би се у што већој мери моделовао визуелни систем посматрача.

2.4.1. Хибридне технике процене квалитета слика

Бројни аутори наглашавају да не постоји објективна техника процене квалитета која ће бити употребљива у свим ситуацијама – за различите типове деградације, за различите степене деградације, за различите примене и сл. (универзална мера). Ово ограничење се може превазићи кроз комбиновање процена више објективних мера процене квалитета, тј. кроз хибридне моделе процене. Хибридним техникама процене квалитета се елиминишу недостаци појединачних мера процене, док се истовремено њиховим коришћењем заобилази потреба за уградњом сложених модела визуелног система човека у објективну процену. Ове технике најчешће захтевају претходну анализу о могућности примене различитих мера процене у оквиру хибридног модела, као и анализу значаја погодних мера процене (кроз обучавање). Сложеност хибридних мера расте са повећањем броја техника уграђених у сам модел. И док се у неким хибридним техникама процене квалитета користи двадесетак мера процене, аутори из [68] су показали да се нелинеарном комбинацијом три мере процене могу добити задовољавајуће перформансе.

Сједињавање скорова може бити на локалном нивоу – нивоу пиксела или на глобалном нивоу – нивоу слика. Сједињавање локалних скорова очувања квалитета може се наћи у [69, 70]. У [69] се комбинују мапе квалитета SSIM мере и S-CIELAB просторног колор модела, док се у [70] мапе квалитета SSIM мере двоструко пондеришу – са моделом хоризонталне осетљивости визуелног система човека (највећа осетљивост је за оријентације од 0 до 45°) и локалних спектралних резидуала (остатака).

Већина хибридних техника комбинује скорове квалитета различитих мера на глобалном нивоу [71-77]. У [71] се користи глобални хибридни мултипликативни приступ (са две, три и четири објективне мере) за процену квалитета слика са вишеструком дисторзијом. Зависност перформанси од типа деградације је у [72] превазиђена тако што је најпре кроз линеарну дискриминантну анализу извршена класификација тест слике, а након тога је процена квалитета извршена најпогоднијом техником. У процесу класификације као обележја коришћене су вредности добијених објективних скорова квалитета по различитим техникама.

Перформансе објективних мера процене значајно зависе и од опсега квалитета тест слика [73, 74]. На пример, SSIM и његово мултивеличинско проширење (Multi-Scale SSIM, MS-SSIM) су погодне за процену квалитетнијих слика док је VIF техника погодна за процену квалитета слика са већим степеном деградације. Аутори из [73, 74] су показали да се комбиновањем мера процене погодних за различите опсеге квалитета може добити значајно побољшање перформанси у односу на појединачне мере или актуелне објективне мере – SSIM, MS-SSIM и VIF. Интересантан начин комбиновања две различите технике процене презентован је у [75]. Глобалне вредности појединачних скорова представљају дводимензионално стање које се пресликава у процену визуелног квалитета. Првом мером се одређује губитак структурних информација док друга мера представља индикатор типа дисторзије тест слике.

Хибридне технике процене квалитета погодне су и код редукованог референцирања и процене квалитета без референцирања. Комбинацијом три мере дизајниране за мерење специфичних аспеката перцептуалног квалитета (блоковски ефекат, губитак просторних информација и промена контраста) уз адаптивни поступак пондерисања, у [76] је изведен композитни објективни скор квалитета са редукованим референцирањем, а који је погодан за предикцију субјективних импресија квалитета слика из различитих класа и широког опсега садржаја. Статистике слика природних сцена из три трансформациона домена су искоришћене у формирању објективне мере процене квалитета без референцирања [77]. Мера се користи за процену квалитета слика са четири типа дисторзије, а тачност предикције је у рангу мера са потпуним референцирањем.

2.4.2. Колор модели намењени процени квалитета слика

Иако је највећи број база намењених процени квалитета састављен од колор слика и видеа, поуздане мере процене не користе у пуној форми информације о боји, односно користе канал осветљаја или монохроматски сигнал. Постоје објективне мере које користе очување информација из канала осветљаја и канала боје [78-85], али је њихова популарност далеко иза популарности стандардних мера као што су SSIM, VIF, MAD, VSNR и сл. Разлози за то су бројни. Коришћење три канала колор

слика/видеа доводи до усложњавања модела процене. Због различитих учестаности одабирања хроминентни канали носе мање информација од канала осветљаја. Деградација боје у хроминентним каналима се значајно разликује од деградација канала осветљаја. Деградацију боје је веома тешко усагласити са субјективним утиском насталих промена. Колор координатни системи се значајно разликују и поставља се питање који би био оптималан за евалуацију квалитета. Динамички опсеzi канала осветљаја и канала боје се некада значајно разликују, а тиме и динамички опсеzi насталих деградација.

Примена индекса структурне сличности у каналима различитих колор координатних система предложених од стране CIE (Commission Internationale de l'Eclairage), довела је до закључка да мерење очувања информација у каналима осветљаја доноси значајно боље слагање објективних и субјективних скорова квалитета него очување информација у каналима боје [78]. Мере које користе сва три канала у процени квалитета углавном већи значај дају каналу осветљаја него каналима боје.

CIELAB је један од коришћених колор модела у процени квалитета [79, 80]. Како би се мерио просторни утицај, након конверзије колор слика у канал осветљаја и два хроминентна канала у [79] су колор компоненте пропуштене кроз просторне филтре који су одабрани у складу са просторном осетљивошћу људског ока. Добијени модел назван је S-CIELAB и коришћен је и у каснијим истраживањима [80]. Истраживањима је показано да се применом једноставних мера (MSE, PSNR, SSIM) у каналима S-CIELAB модела може остварити добитак у односу на класичне примене над монохроматским сликама.

У објективну процену квалитета колор слика могу се укључити модели визуелног система човека [81, 82]. Монохроматски модел је прилагођен процени квалитета колор слика и искоришћен за одређивање визуелне репрезентације изворне и тест слике у [81]. Разлика визуелних репрезентација представља меру квалитета тест слике. Визуелни систем човека користи се и у [82] за процену квалитета колор слика уз редуковано референцирање.

Индекс структурне сличности се на различите начине може прилагодити процени квалитета колор слика [83, 84]. Најчешће се посматра очување информација

колор канала [83], али се могу дефинисати и сложенији приступи [84]. Мера предложена у [84] представља векторско проширење SSIM индекса коришћењем кватерниона. У предложеном приступу сваки (колор) пиксел се у векторском домену посматра као један кватернион број, тако да има релативно малу комплексност.

Постојеће структурне мере одсликавају очување структуре у монохроматским сликама. За мерење квалитета може се посматрати и очување расподеле боје. Кроз статистичку анализу кумулативних хистограма у троканалном колор простору предложена је објективна мера која је независна од резолуције тест слике као и просторних операција као што су translација, ротација, скалирање и одсецање [85].

2.5. Примена модела визуелног система човека у процени квалитета слика и видеа

Како би се осигурало да се у евалуацији квалитета добија субјективно виђени квалитет а не сличност два сигнала, користе се објективне мере засноване на моделима визуелног система човека (Human Visual System, HVS). Један од првих радова који обухвата визуелни систем човека је рад од Далија (Daly) [86]. Уместо квалитета слике кроз скаларну вредност (апсолутна мера), предложена је релативна мера VDP (Visible Differences Predictor), кроз коју се представљају видљиве разлике између две слике. Полазећи од изворне и тест слике долази се до треће слике која даје њихове субјективно уочљиве разлике односно вероватноће детекције њихових разлика. Модел визуелног система човека се састоји од три компоненте кроз које се моделује осетљивост визуелног система човека на промене интензитета, просторне фреквенције и садржај сигнала. Осетљивост визуелног система човека на просторне фреквенције назива се функција осетљивости контраста (Contrast Sensitivity Function, CSF), док су варијације осетљивости у функцији садржаја сигнала познате као маскирање. Моделовањем визуелног система човека кроз Габорову пирамиду у [87] је показано да се модел може користити у процени квалитета. Кроз експерименте је анализирана осетљивост визуелног система човека на просторне фреквенције, оријентације и средњу вредност осветљаја. По узору на [86], тј. формирање матрице видљивих разлика, аутори из [88] су предложили на који начин се осетљивост

визуелног система човека на осветљај, просторну фреквенцију и контраст може уградити у вејвлет домену. Делотворност модела је након тога тестирана кроз анализу разлика оригиналне и модификованих слика (компримованих или са додатим шумом).

У објективној процени квалитета најчешће су коришћене две карактеристике визуелног система човека: ефекат визуелног маскирања и двострука стратегија перцепције [89]. Маскирање текстуре условљава да ће већи степен дисторзије бити толерисан од стране посматрача у регионима са израженом текстуром (статична или динамичка) него у униформним регионима. Двострука стратегија перцепције говори да посматрач на различите начине процењује квалитет у видеу високог квалитета (где тражи дисторзију – разлику) и видеу лошег квалитета (где детектује артефакте).

Истовремено присуство више маскирајућих фактора у локалном окружењу условљава отежану детекцију артефаката унетих у слику/видео у односу на случај присуства само једног маскирајућег фактора јер се ефекти маскирања не могу линеарно сабирати. Различити ефекти маскирања се користе како би се извели прагови максималне дисторзије коју визуелни систем човека не може приметити – прагови JND (Just-Noticeable Distortion). Најчешћи фактори који се узимају у обзир приликом одређивања JND су адаптација осветљаја, маскирање контраста и маскирање текстуре. JND профил за видео захтева да се поред просторних особина обухвати и време које додатно утиче на визуелни систем човека. JND просторно-временски модел се може одредити у просторном [90] или фреквенцијском домену [91].

Једноставан али ефикасан модел визуелног система човека који користи ефекте просторног маскирања, као што су маскирање осветљаја и текстуре, предложен је у [92]. Ефекти су имплементирани у просторном домену и комбиновани у један коефицијент видљивости. У анализи слика са блоковским ефектима илустрована је добра корелација изведене објективне мере са субјективним скоровима. У [93] је предложена перцептуално пондерисана мултивеличинска варијанта индекса структурне сличности (комплексна вејвлет варијанта – CWSSIM) која представља интеграцију структурног приступа са традиционалним JND перцептуалним приступима. Визуелно маскирање је комбиновано са индексом

структурне сличности за процену квалитета слика и у [94]. Поређење осветљаја је задржано из оригиналног SSIM индекса, поређење структуре је урађено на ефикаснији начин а перцептуалне карактеристике визуелног система човека су интегрисане у поређење контраста кроз ефекат маскирања контраста и маскирање од стране суседа. Сложеност описа квалитета слике илустрована је у [95], где је наглашена потреба за мултикомпонентним описом квалитета слике. Предложена је мера квалитета у којој се разматрају праг визуелне перцепције, визуелна импресија (кроз осветљај и боју) и визуелни значај. Тиме је визуелни систем човека разматран кроз три аспекта.

За евалуацију алгоритама објективне процене квалитета користе се јавно доступне базе са субјективним скоровима квалитета. Коришћењем доступних база, неколико група истраживача је побољшало своје мере кроз увођење одређених особина HVS или увођење статистичких модела слика. Међутим, веома тешко је знати да ли је побољшање настало због опонашања перцептуалних особина или прилагођења на неку од карактеристика базе [96]. Аутори [96] сматрају да постојеће базе не садрже довољан број узорака за тестирање значаја неке од особина HVS нити покривају разноврсност типова дисторзија које одражавају захтеве постојећих апликација. Да би илустровали своју тврдњу, аутори су показали да једноставна мера заснована на ефекту маскирања контраста може имати перформансе као и актуелне (значајно сложеније) мере процене. Оно што актуелне базе не обухватају су дисторзије у глатким регионима, дисторзије у тамним/светлим регионима, дисторзије у занимљивим регионима и сл. Такође, аутори се не слажу да већина особина HVS има тривијалну улогу у субјективној евалуацији квалитета. Бројне особине су небитне и могу се избацити из мере једино због тога што сама база не обухвата све особине HVS.

2.6. Објективна процена квалитета видео сигнала

Једноставан алгоритам за процену квалитета видеа може се добити директном применом процене квалитета слике кадар по кадар (фрејм по фрејм). Напредне технике процене квалитета видеа узимају у обзир и временску корелацију и информације о покрету у видео сигналима, али често доводе до тога да се

рачунарско време значајно увећава. Због тога процена квалитета видеа има три додатна аспекта која је потребно размотрити у односу на процену квалитета мирних слика: (1) поред просторних потребно је размотрити временске факторе, (2) потребно је одредити утицај сваког од фактора (просторног и временског), као и њихову интеракцију при одређивању крајње вредности квалитета и (3) водити рачуна о сложености резултујућег метода.

Проширења (једноставна или сложена) познатих објективних мера процене квалитета слика за процену квалитета видеа могу се наћи у [97-104]. Често критикована средња квадранта грешка, MSE, је у [97] комбинована са просторно-временском активношћу и искоришћена за поуздану процену квалитета секвенци CIF просторне резолуције компримованих за широк опсег битских брзина коришћењем H.264/AVC стандарда.

Индекс структурне сличности [48] је доживео бројна прилагођења процени квалитета видеа [98-102]. Прилагођење из [98] се састоји у томе да се структурна сличност одређује у три канала YCbCr колор координатног система након чега се формира збирни скор пондерисањем скорова квалитета из канала (канал осветљаја носи 80% значаја а канали боје 20%). Збирна вредност квалитета добија се усредњавањем скорова квалитета по кадровима. Аутори су такође показали да се робуствена процена квалитета видео секвенци може добити поређењем делова (случајно одабраних блокова) изворног и тест сигнала. Очување информација о структури се може посматрати и дуж временских путања (дуж путања кретања блокова) [99], где се ради ефикасности метода у одређивању временских путања користи компензација кретања позната у компресији видеа. У [100] се очување информација током времена мери применом индекса структурне сличности на слике разлика суседних кадрова изворног и тест видеа (временски индекс структурне сличности). Очување структурних информација у времену се комбинује са очувањем просторних информација (срачунатих кадар по кадар) и усредњавањем добија крајња мера квалитета. Просторни квалитет је посматран кроз скупове суседних кадрова а не кадар по кадар. Аутори су се на такав приступ одлучили због тога што тренутна импресија квалитета представља резултат и неколико претходних кадрова. Уведено је филтрирање просторних скорова квалитета кадрова тако да нови кадар тренутно нема значајан ефекат на процену квалитета и да се утицај претходних кадрова брзо

не изгуби. Видео сигнал се може посматрати као тродимензионални (3D) волумен (две просторне димензије и време), а индекс структурне сличности проширити и на трећу димензију – 3D SSIM [101]. Коришћењем блокова димензија 7x7x7 пиксела креирана је 3D мапа SSIM квалитета која је додатно пондерисана са локалним информационом садржајем и усредњена за добијање процене квалитета. MOVIE индекс [102] (MOtion-based Video Integrity Evaluation Index) представља веома познату меру процене квалитета видеа а користи принципе постављене кроз SSIM [48] и VIF [58] мере. Просторне и временске дисторзије посматрају се у више величина (резолюција), кроз трајекторије кретања од интереса, уз истовремено зарачунавање ефеката просторног и временског маскирања.

Аутори из [103] су увели просторно-временске слике које се добијају посматрањем интензитета редова и колона изворне и тест секвенце током времена. Артефакти кретања се у овим сликама манифестују као просторни артефакти, што омогућава да се за евалуацију њихове дисторзије користе класични алгоритми процене квалитета (у овом случају MAD [57]). Поређење (познато као ST-MAD) је додатно модификовано са тежинама изведеним из оптичког тока (optical flow), како би се већи значај дао брзопокретним регионима који се налазе у средишту видеа. Објективна мера процене квалитета мирних слика предложена у [60] је послужила као основа за прорачун очувања просторних информација у видео сигналу [104]. Додатно је у раду анализирано како кретање на сцени и карактеристике HVS утичу на људску перцепцију просторних дисторзија а самим тим и на објективну процену квалитета видеа.

Технике процене квалитета видеа не морају имати основу у објективним мерама процене квалитета мирних слика [105-107]. Објективном мером из [105] одређује се видљивост/невидљивост дисторзије, мери сличност блокова и верност садржаја. Видљивост дисторзије се одређује кроз ефекте маскирања осветљаја, маскирања текстуре, маскирања у времену и разлике суседних кадрова. Сличност блокова се одређује поређењем заснованим на детекцији блоковских ефеката док се верност садржаја мери поређењем ентропија на нивоу кадрова. Мера је тестирана на CIF (352x288 пиксела) и QCIF (176x144 пиксела) секвенцама при различитим битским брзинама (24-384 kbps) и различитом темпу (7.5-30 fps). Са перцептуалног становишта анализиран је квалитет H.264 и MPEG-4 компримованих секвенци. Како

би се квантификовао субјективни квалитет компримованог видеа аутори из [106] спровели су два експеримента. Резултати првог експеримента, са посматрачима који нису фамилијарни са обрадом слике, су показали да су за два кодека, AVS и H.264, перцептуално доминантна три типа артефаката: замрљање, блоковски ефекат и дисторзија боје. H.264 кодек се показао субјективно бољим за SD (720x576i) формат секвенци, док су резултати били слични за HD (1280x720p) и CIF формате. У другом експерименту, са посматрачима који су експерти из области обраде слике, показано је да се деградација видеа која настаје компресијом може предвидети кроз линеарну комбинацију атрибута кроз које се мере три доминантна перцептуална артефакта. F-MSE (Foveated Mean Squared Error) мера процене квалитета видеа узима у обзир променљиву резолуцију са којом визуелни систем човека анализира сцену [107]. Највећа оштрина је за тачку (фиксацију) која се пројектује на фовеу – део ретине са највећом густином фоторецептора. Визуелна оштрина брзо опада са удаљавањем региона слике од фиксационе тачке. Мера додатно узима у обзир и редукцију просторне оштрине због кретања у секвенци.

Сложеност објективне процене квалитета видео сигнала илустрована је у радовима [108, 109]. У [108] је презентована нова база са 200 видео секвенци – LIVE Mobile Video Quality Assessment database. Поред дисторзија које су и раније разматране, као што су компресија и пакетски губици, база садржи и динамички променљиве дисторзије, као што је замрзавање кадрова и временски променљив степен компресије. На бази је евалуирано неколико објективних мера процене квалитета слика и видеа и закључено је да мере нису у потпуности прихватљиве за процену квалитета дисторзија које се мењају током времена. Анализа објективних FR и NR мера спроведена у [109], показала је да већина техника није способна да обухвати различите комбинације просторних и временских дисторзија (резолуција). Резултати су значајно зависили и од садржаја видеа.

2.6.1. Процена квалитета секвенци са малим битским брзинама

Највећи број мера за објективну процену квалитета видео сигнала, укључујући и мере прихваћене од стране VQEG групе [30, 31], намењен је за процену квалитета видео секвенци са фиксном и високом просторном и временском

резолюцијом. Тако су објективне мере из VQEG FR-TV Phase I тестова евалуиране на PAL и NTSC телевизијским форматима, резолуција 720x576/50fps и 720x486/60fps, тим редом. Увођење квантизације видео сигнала неизбежно доводи до дисторзија унутар кадрова као што су блоковски ефекат, ефекат прстена, замрљање и сл. Бројне мере су предложене за евалуацију ових дисторзија и за предикцију њиховог перцептуалног квалитета.

Пренос видео сигнала кроз комуникациони канал ограниченог капацитета, као што је мобилни радио канал, веома је тешко остварити уз високу просторну и временску резолуцију. У оваквим ситуацијама је уобичајено да се поред грубе квантизације врши смањење како просторне тако и временске резолуције (темпа), са циљем смањења количине података. Смањење темпа се спроводи кроз одбацивање кадрова док се просторна резолуција смањује кроз избацивање редова и колона. Редукција доводи до нове, неизбежне деградације квалитета. Одбацивање кадрова доводи до подрхтавања (jitter/jerkiness) а избацивање редова и колона доводи до замрљања (уколико се изврши интерполација и видео прикаже у изворној резолуцији). Артефакти који се јављају између кадрова (inter-frame) су анализирани у нешто мањој мери него артефакти унутар кадрова (intra-frame). Здружени ефекти артефаката унутар кадрова и између кадрова нису изучавани у пуном обиму, тако да се може наћи јако мали број радова намењених процени квалитета видео секвенци са различитом просторном резолуцијом и темпом. Од посебног значаја је пронаћи оптималан однос димензије кадрова и темпа за ограничен капацитет канала.

Перформансе H.264/AVC стандарда видео кодовања у мултимедијалним (мобилним) сервисима са малим битским брзинама анализирани су у [110]. Рађено је са QCIF форматом а тежиште рада је дато на избору параметра квантизације и темпа. Аутори су закључили да је битска брзина од 64 kbps довољна за прихватљив квалитет док темпо од 10 fps задовољава потребе корисника. Занимљиво је да су корисници за секвенце са спортских дешавања бирали нешто већи темпо – 15 fps.

Полазећи од вишеканалног модела визуелног система човека, аутори из [111] су предложили меру за објективну евалуацију квалитета видеа при преносу кроз канал са малим битским брзинама. Мера је евалуирана на CIF видео секвенцама уз битске брзине од 100 до 800 kbps. У процени квалитета користи се пирамидална

декомпозиција изворне и тест секвенце, након чега се пореде локалне вредности декомпозиција за различите величине и фреквенције. Сличности локалних вредности се интеграле у два корака, кроз интеграцију на нивоу кадра (просторно-фреквенцијска интеграција) и интеграцију на нивоу секвенце (временска интеграција). Аутори су додатно показали да одбацивање поређења у најфинијим скалама, поред значајног смањења рачунарске комплексности, доводи и до повећања степена слагања субјективних и објективних скорова.

Резултати истраживања представљени у [112] наглашавају да на субјективни квалитет утичу тип енкодера, видео садржај, битска брзина, темпо и димензија кадра, по опадајућем значају. Независно од темпа и димензије кадра (QCIF или CIF), кодовањем помоћу H.264 кодера добијен је задовољавајући (добар или одличан) квалитет уколико се пиксел представља са 0.1 bpp. Ограничење битске брзине захтева оптималну комбинацију темпа и димензије кадра, где субјекти бирају мању димензију кадра уз спорији (бржи) темпо за видео секвенце са високом (ниском) временском активношћу.

Утицај крајњег уређаја у систему преноса на субјективну импресију квалитета анализиран је у [113]. Као крајњи уређаји коришћени су персонални рачунар и мобилни телефон. Коришћен је QCIF формат видеа, са битским брзинама од 48 kbps до 128 kbps и H.264 компресијом. У скоро свим сценаријима субјекти су дали предност персоналним рачунарима.

2.6.2. Простор и време као фактори процене квалитета

Пренос дигиталног видео сигнала кроз канал ограниченог капацитета у коме делује шум доводи до просторних и временских деградација које су уочљиве у декодованом видеу. Оба типа оштећења утичу на субјективну перцепцију квалитета видеа [114-118]. Истраживања презентована у [114] наглашавају да увођење деградације у једном модалитету утиче на перцепцију другог модалитета и ова размена утицаја је већа код квалитетнијих видео секвенци него код лошијих. Просторни артефакти су генерисани кодовањем видеа за различите битске брзине, а временски артефакти су генерисани увођењем замрзавања кадра различитог трајања. Такође, закључено је да просторни квалитет има већи значај при формирању крајње импресије квалитета. Варијација просторног квалитета током

времена има значајан утицај на глобалну вредност квалитета [115, 116]. Једноставни алгоритми процене квалитета кадар по кадар (PSNR, SSIM) се додатно могу учинити робуснијим увођењем фактора промене квалитета током времена [115]. Варијација квалитета током времена била је предмет истраживања описаних у [117, 118]. Варијација квалитета настаје због варијација (комплексности) видео садржаја, конфигурације кодека и мрежних услова. Еволуција просторних дисторзија се може искористити за побољшање мера објективне процене квалитета [117]. Субјективним експериментима из [118] анализиран је одзив посматрача на видео секвенце које се састоје од више сегмената приближно истог или значајно различитог квалитета. Анализа је показала да једноставни модели којима се сабира квалитет појединих сегмената, као што је усредњавање или пондерисање, нелинеарно минимално и медијан филтрирање и дисторзијом вођена интеграција, не одговарају процени квалитета видеа комбинованог од више секвенци. Због тога је предложен адаптивни модел који је асиметрично подешен за растући и опадајући квалитет. Предложени модел је изведен из резултата субјективних тестова који су показали да: (1) блага промена квалитета суседних секвенци даје и малу промену у субјективној оцени квалитета, (2) значајан пад квалитета између секвенци (после секвенце доброг се презентује секвенца лошег квалитета – негативна промена) доводи до много већег пада у субјективној евалуацији квалитета него што је добитак у обрнутој ситуацији, тј. после секвенце лошег квалитета се презентује секвенца значајно бољег квалитета (позитивна промена) – асиметрична стратегија корекције квалитета и (3) субјективна оцена квалитета значајно не зависи од квалитета последњег (или првог) приказаног сегмента. Асиметрична адаптација је по деловима линеарна, тј. за мале промене објективног квалитета суседних сегмената врше се благе корекције објективних скорова, док су корекције веће за негативне промене квалитета него за позитивне промене. Предложеном асиметричном адаптацијом су побољшане перформансе објективних процена (PSNR и SSIM) код којих је квалитет разматран независно за сваки од сегмената.

Утицај просторне, временске и амплитудске резолуције на перцептуални квалитет компримованог видеа истражен је у [119]. Коришћено је 5 просторних резолуција, 3 временске резолуције и 4 нивоа квантизације. Субјективни подаци су показали да се утицај просторне, временске резолуције и корака квантизације може обухватити са по једним параметром који зависи од садржаја видеа. Уочено је да временска резолуција утиче на квалитет независно од просторне и амплитудске, али

постоји значајна интеракција просторне резолуције и корака квантизације. Наиме, пад квалитета који настаје због смањења просторне резолуције линеарно зависи од корака квантизације.

2.6.3. Процена квалитета HEVC кодованих секвенци

H.264/AVC стандард за видео кодовање је компетиран од стране JVT тима (Joint Video Team) 2003. године и формално стандардизован као ITU-T H.264 и ISO/IEC MPEG-4 Part 10. То је и даље један од најчешће коришћених видео кодека за архивирање, компресију и дистрибуцију HD видеа у великом броју апликација (TV радио-дифузија, видео конференција, веб видео и Blu-Ray). У односу на претходне стандарде видео кодовања, као што су MPEG-2 и MPEG-4 Part 2, остварена је уштеда у битском протоку за више од 50%. Добитак је остварен кроз скуп напредних техника кодовања као што су компензација предикције кретања са променљивом димензијом блока, ентропијско кодовање засновано на садржају, временска предикција и адаптивни деблокинг филтар [120].

Последњу деценију обележио је рад на увођењу новог стандарда компресије видео сигнала – H.265/HEVC. Циљ увођења стандарда је био да се задржи субјективни квалитет видеа уз смањење битског протока за 50% у односу на H.264 [121]. Супериорност HEVC у односу на VP9 и H.264 компресију илустрована је у [122]. Исти субјективни квалитет обезбеђен је са 43.3% и 39.3% мањом битском брзином у односу на VP9 и H.264 кодеке. Радови у којима се анализира процена квалитета HEVC секвенци наглашавају да стандардне објективне мере потцењују субјективни добитак настао увођењем кодовања, па се сугерише увођење нових техника визуелног значаја, нових модела интеграције скорова и увођење процене временских артефаката [121, 123].

2.7. Увођење мапа визуелног значаја у евалуацију квалитета

Алгоритми за процену квалитета слике/видеа најчешће имају две фазе. У првој фази квалитет се евалуира на локалном нивоу и добија се мапа квалитета/дисторзије на нивоу пиксела. У другој фази користи се просторно

удруживање локалних скорова како би се од мапе квалитета/дисторзије добила збирна вредност квалитета слике/видеа. Најчешће коришћена техника просторног удруживања јесте усредњавање (просторно и временски) [124], где се крајња вредност може посматрати као мера просечне репрезентације изворног сигнала у тест сигналу.

Проблеми са просторно-временским усредњавањем се јављају када дисторзија није униформно распоређена у простору слике или се мења током времена. На пример, када у малом делу слике постоји дисторзија а остали делови су доброг квалитета, посматрач ће већу пажњу обратити региону лошег квалитета и даће процену нижу од средње вредности мапе квалитета/дисторзије. Због тога истраживачи покушавају да дефинишу оптималан начин комбиновања локалних скорова квалитета [124-148].

Пондерисање локалних скорова квалитета може се поделити на две класе (са различитим подкласама) [125]:

1. квалитетом вођена интеграција
 - Минковски интеграција
 - интеграција на основу локалних вредности квалитета/дисторзије (применом монотоне функције усредњавања или коришћењем најнижих скорова квалитета)
2. интеграција на основу садржаја
 - коришћењем информационог садржаја
 - коришћењем реалних мапа субјективног значаја
 - коришћењем мапа добијених применом рачунарских модела

Иако универзална техника интеграције која би одговарала различитим објективним мерама и атрибутима слике/видеа не постоји, увођењем напредних метода интеграције добијају се боља слагања субјективних и објективних скорова.

Просторна и временска интеграција су у блиској вези са визуелним значајем, који се последњих година веома широко проучава од стране психолога, неуробиолога и инжењера како би се истражило који делови слике привлаче већу пажњу визуелног система човека. Евалуацијом визуелног значаја идентификују се информације на слици које значајно утичу на посматрача када формира импресију о

квалитету. То омогућава да се повећа утицај битних информација на коначни резултат процене. Другим речима, значајније информације имају и већи значај у укупној процени.

2.7.1. Квалитетом вођена интеграција

Квалитетом вођена интеграција (Минковски интеграција и интеграција на основу локалне вредности квалитета) над локалним скоровима квалитета добијеним кроз апсолутне разлике нивоа сивог и одређивањем структурне сличности, истражена је у [124]. Остварени добитак је зависио од мере која се користи за процену очувања локалних информација. Генерално, добитак је већи уколико се у процени квалитета користи апсолутна разлика нивоа сивог, што је мера која је непоузданија од индекса структурне сличности. Добитак остварен квалитетом вођеном интеграцијом над локалним структурним разликама није био значајан. Аутори су поред квалитетом вођене интеграције користили и интеграцију на основу локалног информационог садржаја и закључили да је он погоднији за употребу (добитак је износио 6% за средњу апсолутну разлику – корелација рангова је повећана са 85% на 91%, и 1% за индекс структурне сличности – корелација рангова је повећана са 95% на 96%).

Генерално је прихваћено да снажне дисторзије у простору и/или времену имају велики ефекат на крајњу импресију квалитета [126-128]. Снажне дисторзије дају ниске вредности сличности изворног и тест сигнала, па се коришћењем најнижих скорова квалитета може формирати крајња импресија о квалитету. Анализа из [126] је показала да коришћењем најнижих скорова квалитета, независно од величине у којој се ради, могу побољшати перформансе индекса структурне сличности у процени квалитета слика. Слагање субјективних и објективних скорова квалитета се побољшава уколико се користи до 6% најнижих скорова квалитета [126]. Овај приступ – приступ најнижи скор има највећи значај, искоришћен је и за процену квалитета слика које се добијају сједињавањем [127]. Значај лошијих скорова квалитета дуж обе димензије видео секвенце (просторне и временске) илустрован је и кроз адаптивну стратегију удруживања локалних скорова квалитета [128]. Аутори су запазили да крива сортираних локалних објективних скорова

квалитета унутар кадрова обично има два уочљива сегмента. Први сегмент чине нижи скорови квалитета код којих крива има велики нагиб, док други сегмент чине високи скорови квалитета где крива сортираних вредности има zasiћење. Граница ова два сегмента се користи за поделу скорова на две класе. Скорови класе са вредностима испод прага имају већи значај у формирању објективне вредности квалитета кадра. Скорови квалитета кадрова су додатно поступком кластеризације подељени у две класе. Класа са нижим вредностима скорова кадрова такође има већи значај у формирању глобалне вредности квалитета. Предложена стратегија је анализирана на LIVE Video Quality и EPFL-PoliMI базама видео секвенци, у оквиру три алгорита процене (MSE, SSIM, MOVIE). Остварен је конзистентан добитак са све три објективне мере на различитим типовима дисторзија и на обе базе. Постигнути резултати су бољи од усредњавања или коришћења дела најнижих скорова.

2.7.2. Интеграција на основу садржаја

Рад [129] се може сврстати у радове који у пондерисању локалних скорова квалитета користе садржај слике. Аутори су кренули од претпоставке да је током посматрања слика природних сцена оптимални субјективни значај пропорционалан локалном информационом садржају слике (број бита потребних за његову интерпретацију). У раду су изведена три битна закључка: (1) пондерисање са информационом садржајем даје конзистентно побољшање перформанси објективних мера процене на свим базама слика, (2) пондерисање често критикованог вршног односа сигнал шум са информационом садржајем може дати меру која је по перформансама у рангу стандардних техника процене квалитета и (3) најбоље перформансе процене објективних мера се остварују кроз здружени ефекат мултивеличинске имплементације објективне мере праћен пондерисањем са информационом садржајем.

2.7.3. Коришћење реалних мапа субјективног значаја

За пондерисање локалних скорова квалитета могу се користити реалне мапе визуелног значаја – мапе изведене из експеримената у којима се кроз посматрање

кретања зенице људског ока у процесу евалуације квалитета одређују визуелно значајни региони у слици/видеу [130, 131]. У [130] су спроведена два експеримента. У првом експерименту посматрачима је дата инструкција да посматрају изворне слике по слободном нахођењу (free-looking стратегија), док је други експеримент спроведен са циљем процене квалитета слика. У првом експерименту 20 посматрача гледало је 29 оригиналних слика, након чега су изведене мапе значаја слика природних сцена (Natural Scene Saliency, NSS). У другом експерименту, 20 посматрача је процењивало квалитет слика са дисторзијом, изведених од оригинала из првог експеримента. Линеарним пондерисањем локалних скорова квалитета FR (PSNR, SSIM, VIF) и NR мера (GBIM [149], NRPB [150]) са реалним мапама значаја донело је побољшање перформанси у односу на просто усредњавање (добитак је већи уколико се користе мапе значаја изведене из оригиналних слика, тј. из првог експеримента). Добитак зависи од објективне мере, типа деградације и садржаја слике. Истраживањима је илустровано да линеарна стратегија комбиновања за неке слике и неке типове дисторзија није најбоља, па су потребне адаптивне стратегије пондерисања. У наставку истраживања, презентованих у [131], аутори су анализирали утицај садржаја слике на перформансе објективних мера пондерисаних са реалним мапама значаја. Показано је да варијација визуелног значаја између посматрача значајно зависи од садржаја слике и као таква се може користити за предикцију добитка увођења значаја у објективну процену квалитета. Што је мања варијација значаја добија се веће побољшање. За слике са великом варијацијом субјективног значаја постоји ризик да ће се деградирати перформансе предикције објективних мера. Преглед јавно доступних база које се могу искористити за извођење реалних мапа визуелног значаја може се наћи у [132].

2.7.4. Коришћење рачунарских модела за формирање мапа субјективног значаја

Моделовање визуелног значаја је веома тежак задатак, јер расподела визуелног значаја зависи од бројних фактора: величине објекта и његове позиције на слици [133], кретања камере [134], циља са којим посматрачи гледају слике/видео [135], степена дисторзије [136], типа деградације, а додатно је у моделовању потребно узети у обзир и меморијски ефекат током посматрања видеа [137].

Коришћење субјективних мапа значаја је непрактично у реалним апликацијама процене квалитета слике, па се уместо реалних мапа користе мапе изведене из рачунарских модела [138-146]. Често коришћене мапе субјективног значаја [138, 139] користе једноставна обележја као што су величина, интензитет, боја и оријентација [138], односно осветљај, контраст и осветљај и контраст након филтрирања слике филтрима пропусницима опсега [139]. Обележја предложена у [138] су имплементирана јер се веома добро препознају у примарној можданој кори. Обележја предложена у [139] су анализирана за делове слике око стварних фиксација и случајно изабраних фиксација. Аутори су закључили да су вредности изабраних обележја веће у регионима слике одређеним на основу стварних фиксација него у ситуацији када су фиксације изабране насумично. На основу добијених резултата, изведен је алгоритам селекције региона кандидата за фиксације (GAFFE модел). Селектовани региони показали су висок степен корелације са фиксацијама посматрача.

Мапе визуелног значаја изведене коришћењем приступа описаних у [138, 139] интегрисане су са три објективне мере процене квалитета (MSE, PSNR, SSIM) [140], чиме су побољшане перформансе предикције субјективних импресија квалитета слика. Добитак зависи од визуелног модела, типа дисторзије и карактеристика саме мере. Највећи добитак остварен је применом GAFFE модела са добитком који је нешто нижи од мапа изведених из субјективних експеримената. Добитак је већи код једноставних мера (PSNR, MSE) него код сложеније SSIM мере. Најмањи добитак је код слика са додатим шумом. Мапа субјективног значаја [138] искоришћена је и за интегрисање са резултатима локалних вредности квалитета четири мере са потпуним референцирањем (SSIM, MS-SSIM, VQM и MOVIE) у процени квалитета видеа [141]. Показано је да се увођењем мапа субјективног значаја побољшавају перформансе свих тестираних мера. Највећи добитак у перформансама је за просторне мере (SSIM, Spatial MOVIE), тј. за мере које разматрају просторне деградације.

Мапе визуелног значаја могу се извести из спектралног остатка слике [142], фазног спектра [143], кватернион репрезентације [144], контраста мултирезолуционе пирамиде [145], а може се користити и комбинација више различитих приступа

[146]. Кроз анализу амплитудског спектра изворне слике [142] резидуал слике је издвојен у спектралном домену, а након тога се из њега у просторном домену прави мапа значаја (просторно филтрирање инверзне Фуријеове трансформације спектралног резидуала). За разлику од приступа из [142] код кога се фазни спектар задржава, аутори из [143] су за одређивање мапе користили само фазни спектар слике (амплитудски спектар је на свим позицијама једнак јединици), при чему се након инверзне Фуријеове трансформације и преласка у просторни домен додатно спроводи просторно филтрирање. Приступом из [144] се сваки кадар видео записа своди на кватернион слику која се добија од осветљаја, хоризонталних и вертикалних промена и промена између кадрова. Мапа значаја се добија од фазног спектра кватернион слике. Полазећи од реалних мапа субјективног значаја, аутори из [145] су извели рачунарски модел који користи само канал осветљаја. У одређивању мапе значаја користи се декомпозиција слике помоћу Гаусове пирамиде са четири нивоа, након чега се на сваком нивоу одређује контраст. Интеграцијом контраста из више величина долази се до мапе значаја комплетне слике.

Осим у процени квалитета слике/видеа, мапе значаја користе се и за руковођење радом кодера [147, 148]. Визуелно битни региони представљају се са већим бројем бита, а мање бита придружује се регионима у којима се може толерисати већи степен дисторзије. Визуелна побуда се некада комбинује са аудио побудом, па значај пиксела зависи од оба извора. Тако се у [148] значај изводи из аудио-визуелне побуде, тј. региони из којих се емитује звук добијају на значају.

2.8. Процена квалитета са редукованим референцирањем

Процена квалитета слика са редукованим референцирањем представља практично решење за аутоматску евалуацију квалитета слике у различитим апликацијама где су на страни пријема (посматрача) доступне делимичне информације о оригиналној слици. То могу бити делови изворног сигнала или обележја изведена из њега. Количина доступних информација варира од приступа до приступа и у неким ситуацијама се своди на свега неколико скаларних вредности [151-154]. У [151] је илустрован приступ код кога се за редуковано референцирање

користи само једна скаларна вредност по слици. Ова вредност се добија издвајањем обележја која одговарају типичним артефактима присутним у слици (блоковски ефекат, замрљање, ефекат прстена, маскирање осветљаја (*intensity masking*) и губитак блокова (*lost blocks*)). На страну пријемника се преноси здружена вредност наведених пет обележја изворне слике. Аутори из [152] су посматрали хоризонтални и вертикални градијент слике и закључили да се њихови хистограми амплитуде могу моделовати генерализованом Лапласовом расподелом. На страну пријемника се преносе четири обележја по сваком смеру градијента – два параметра Лапласове расподеле, варијанса хистограма градијента и одступање моделованог од стварног хистограма градијента. Статистика слика природних сцена у градијентном домену користи се и у [153], где је по слици потребно пренети 6 вредности. Поред четири параметра коришћена у [152], додатно су уведени куртозис (*kurtosis*) и асиметрија (*skewness*) хистограма амплитуда градијента (аутори су посматрали само хоризонтални градијент). Аутори из [154] су кроз мултивеличинску вишеоријентациону трансформацију (*Divisive Normalization Transform, DNT*) представљање изворне слике свели на 36 скаларних обележја ((3 величине) \times (4 оријентације) \times (3 обележја)). За сваки од 12 подопсега одређују се стандардна девијација и куртозис DNT коефицијената и одступање GSM (*Gaussian scale mixture*) модела расподеле амплитуда DNT коефицијената од стварне расподеле. За формирање мере дисторзије искоришћена је филозофија по којој је изведен индекс структурне сличности (*SSIM*). Објективна мера са редукованим референцирањем је тестирана на шест јавно доступних база слика, на којима је постигнут висок степен корелације предложене мере како са SSIM тако и са субјективним евалуацијама квалитета.

Највећи број мера са редукованим референцирањем користи обележја изведена из неког од трансформационих домена јер се коришћењем малог броја обележја референтна слика може представити у облику погодном за анализу перцептуалног квалитета [155, 156]. Статистичке карактеристике подопсега у реорганизованом DCT домену искоришћене су за процену квалитета слика у [155], док је расподела коефицијената DCT искоришћена за редуковано референцирање и контролу битске брзине у систему за пренос путем бежичног линка у [156].

2.9. Процена квалитета без референцирања

Досадашње NR мере процене квалитета најчешће полазе од претпоставке да је тип дисторзије познат (JPEG компресија, JPEG2000 компресија, замрљање) [157-163].

Једна од најпопуларнијих техника за компресију слике и видеа јесте трансформационо кодовање на нивоу блокова пиксела коришћењем дискретне косинусне трансформације (block-based DCT). Како се кодовање обавља независно на нивоу блокова овакав вид компресије доводи до појаве блоковских артефаката, тј. веома уочљивих граница између суседних блокова. Овај ефекат је нарочито уочљив при високом степену компресије, тј. за пренос са малим битским брзинама. Субјективна импресија квалитета слике/видеа са оваквом компресијом се не слаже у потпуности са популарним објективним мерама процене квалитета, као што је PSNR. Због тога су у [157-159] предложени алгоритми за процену квалитета слика/видеа модификованих блоковском DCT трансформацијом. Код мера се најчешће посматрају разлике интензитета граница блокова а додатно се могу увести и ефекти адаптације и маскирања [158].

Истраживања [160, 161] у домену NR процене квалитета спроведена су са циљем процене квалитета JPEG2000 класе слика. У овом случају типични артефакти настали вејвлет компресијом су замрљање и ефекат прстена. Замрљање настаје због пригушења високих просторних фреквенција слике, док ефекат прстена настаје због квантизације високофреквенцијских коефицијената при трансформационом кодовању. Ови ефекти се манифестују кроз разливање ивица, губитак детаља и таласање у околини изражених граница. Ефекти се могу мерити и користити за процену квалитета у просторном [160] или трансформационом домену [161]. Степен замрљања се у [160] одређује на основу ширине ивица. Наиме, након одређивања истакнутих вертикалних ивица на слици анализира се њихово окружење и одређују локални максимуми и минимума (почетак и крај ивице). Разлика позиција максимума и минимума одговара локалном степену замрљања. Степен замрљања слике се добија усредњавањем локалних степена замрљања. Амплитуде коефицијената DCT се у [161] користе за одређивање статистичке мере – куртозиса, кроз који се индиректно мери степен замрљања, тј. са повећањем степена замрљања расте и вредност куртозиса.

Процена квалитета секвенци кодованих H.264 компресијом била је предмет разматрања у [162, 163]. Метод предложен у [162] је заснован на мерењу анизотропије (anisotropy) слике кроз варијансу дирекционе ентропије слике при чему се користе Renyi ентропије вишег реда. Процена квалитета без референцирања намењена дигиталном преносу H.264 видео секвенци из [163] садржи два корака: процена грешке кодовања и перцептуално пондерисање грешке. Процена грешке је прорачуната у трансформационом домену, прихватањем да су коефицијенти DCT искварени шумом квантизације. Расподела коефицијената DCT је моделована коришћењем Кошијеве (Cauchy) или Лапласове (Laplace) функције густине вероватноће, чији параметри зависе од корака квантизације и квантизираних података. За одређивање значаја коришћена је просторно-временска функција осетљивости контраста, која је имплементирана у DCT домену.

У највећем броју случајева тип дисторзије није познат, па није разумно усвојити претходно наведену претпоставку. Опште технике за процену квалитета без референцирања углавном имају фазу обучавања кроз коју се анализира статистика слика природних сцена у неком од погодних домена [164-166]. Различити типови дисторзије имају јединствену сигнатуру унутар статистике слика природних сцена, а која се може користити за класификацију типа дисторзије. Статистика слика природних сцена из вејвлет домена је искоришћена за одређивање вероватноће припадности тест слике одређеним класама слика у [164]. Добијене вероватноће су искоришћене за пондерисање објективних скорова који се добијају пресликавањем вектора обележја из вејвлет домена у квалитет. Мера из [165] захтева минималан тренинг а заснована је на посматрању статистика коефицијената DCT трансформације слика природних сцена. Како расподела енергије по оријентацијама и скалама зависи од типа дисторзије слике, СТ (Curvelet Transform) трансформација такође представља користан алат за обучавање модела процене квалитета слика без референцирања [166]. Приступ из [166] који се састоји од две фазе – класификација дисторзије праћена проценом квалитета, прихваћен је и у [167], где се свака слика пресликава у 12-елементни вектор који се користи за класификацију и процену квалитета. Изворна слика анализира се у три величине и при томе се у свакој од величина одређују по два обележја из просторног и спектралног домена. То су

средња вредност и коефицијент асиметрије (skew) локалних вредности нивоа сивог (просторни домен) и нормализованих локалних вредности DCT трансформације слике.

Актуелне мере за процену квалитета слика без референцирања [164-167] захтевају знања о очекиваним дисторзијама у форми обучавајућег скупа и познавање субјективних импресија квалитета (у форми MOS или диференцијалног MOS – DMOS). У процени квалитета слика/видеа без референцирања може се ићи и корак даље, тј. развијати перцептуални модели кроз које се може проценити квалитет тест слика са што мање знања о сликама или њиховим дисторзијама [168, 169]. Приступии презентовани у [168, 169] не захтевају процес обучавања, односно не захтевају субјективне импресије квалитета. Због тога се за њих може рећи да су у потпуности "слепи" (completely blind). Процена квалитета из [168] користи перцептуално битна обележја изведена из слике са дисторзијом – средња вредност фазне подударности (сличности) (PC), ентропија PC слике, ентропија слике са дисторзијом и средња вредност амплитуде градијента тест слике. Наведена обележја су усвојена јер су се показала перцептуално битним у истраживањима [170]. Модел познат под називом NIQE (Natural Image Quality Evaluator) такође не захтева тренинг на тест сликама [169]. Модел представља наставак истраживања модела из [171] (познат као BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator), а у процени квалитета користи статистике (средња вредност, варијанса, облик) локално нормализованих вредности осветљаја како би се квантификовали могући губици у природности слике захваљујући присуству дисторзије. Мера је мале рачунарске сложености па је погодна за рад у реалном времену, а додатно не захтева одређивање обележја карактеристичних за поједине дисторзије – замрљање, блоковски ефекат и ефекат прстена.

Иако аутори из [168, 169] наглашавају да њихови приступи не захтевају процес обучавања, у мере су уграђена знања стечена у ранијим истраживањима у области процене квалитета слика. Кроз тестове и тренинг на великом узорку слика, изабрана су битна обележја и одређене оптималне вредности неопходних параметара. Због тога мере не морају да виде део неког новог тест скупа како би се прилагодиле.

Генералне технике за процену квалитета без референцирања на изворни видео сигнал предложене су у [172, 173]. Разлике вектора кретања региона високе просторне комплексности између суседних кадрова пондерисани са временском активношћу видеа, представљају основу процене предложене у [172]. Процена просторних дисторзија спроводи се коришћењем вектора кретања а кроз временску активност се узима у обзир да визуелни систем човека може толерисати већи степен дисторзије у брзопокретним регионима. Крајња вредност квалитета добија се усредњавањем добијених просторних дисторзија по свим кадровима у видеу. Модел који није прилагођен конкретном типу дисторзије предложен је и у [173]. Допринос рада огледа се у предлогу просторно-временског модела статистика слика природних сцена за видео сигнал (изведен у DCT домену), предлогу модела промена на сцени којим се квантификује доследност кретања на сцени и анализи примене предложених модела у процени квалитета видеа. Предложени модел је назван video BLINDS (Blind Image Integrity Notator using DCT Statistics) и тестиран је на LIVE VQA и EPFL-PoliMi видео базама. Перформансе су веома близу перформанси најбољих мера са потпуним и редукованим референцирањем. Објективна процена квалитета видеа се добија полазећи од формирања разлика суседних кадрова, поделе слике разлике на блокове и одређивања DCT коефицијената за сваки од блокова. Расподела DCT коефицијената сваког од блокова се моделује са генерализованом Гаусовом функцијом густине вероватноће. Параметри функције густине вероватноће се разликују за референтни и видео са дисторзијом, а на параметре додатно утиче и степен дисторзије. Кроз процес обучавања (Linear Support Vector Regression, LSVR) предложен је начин пресликавања статистичких параметара модела у скор квалитета.

Занимљиви приступи и закључци из домена процене квалитета без референцирања могу се наћи у [174-176]. Процена квалитета без познавања оригиналног садржаја видеа кроз скривање података приказана је у [174]. У кадрове видеа утискује се бинарна слика (отисак). На страни пријемника отисак се издваја и његова деградација користи за процену квалитета видеа. Кроз експерименте је анализирана видљивост отиска и сметње које уноси убацивање отиска и естимиран је оптималан отисак за појединачни видео. Ћулибрк и остали [175] су показали како се кроз мултивеличинско одузимање слике позадине могу на ефикасан начин издвојити визуелно упадљиви региони у кадровима видео секвенце. Предложено је

17 нових обележја која узимају у обзир визуелни значај покретних објеката у видеу. Након тога су упоређене могућности 18 најчешће коришћених обележја и новоуведених, у процени квалитета MPEG-2 видео секвенци без референце. Експерименталним резултатима је потврђено да се информације издвојене из региона већег визуелног значаја (кретање) могу поправити резултати процене квалитета без референцирања. Квалитет секвенци без коришћења изворног сигнала се може добити и коришћењем једноставних параметара издвојених из секвенце са деградацијом – параметар квантизације, кретање у секвенци и фактор алокације бита [176].

2.10. Очување градијента у процени квалитета слике и видеа

Истраживачи у области процене квалитета слика и видеа често користе запажање да је визуелни систем човека осетљив на ивице и контуре, односно да региони у којима постоје промене нивоа сивог (градијент) имају већи субјективни значај. Тако су развијене бројне градијентне технике процене квалитета, од којих је вероватно најпознатија индекс структурне сличности (SSIM) [48]. Поређење стандардних девијација блокова изворне и тест слике може се сматрати поређењем амплитуда градијента док се крос-корелација блокова може посматрати као мера сличности оријентација градијента (структура). SSIM индекс је само потврдио значај градијента у процени квалитета слика/видеа, који је коришћен и у ранијим радовима за различите примене [177-179]. Структурне информације, заједно са временским информацијама се у [177] користе за објективну процену квалитета видеа. У процени квалитета слика са компресијом [178] користе се три фактора: промена контраста, промена структуре и грешке квантизације (промена амплитудске резолуције). Ова три фактора одређивана су у блоковима 3x3 пиксела након чега су локалне вредности комбиноване у једну меру квалитета. Контраст је процењен преко стандардне девијације нивоа сивог, просторна структура (облик) унутар блока је одређена кроз детекцију ивица, а резолуција нивоа сивог је одређена рачунањем броја различитих нивоа сивог унутар блока. Структурне информације се у [179] кроз градијентну репрезентацију користе у процени квалитета мултисензорског сједињавања слика.

Кроз наставак истраживања започетих у [177] NTIA (The National Telecommunications and Information Administration) је дошла до генералног модела за процену квалитета видеа – VQM (Video Quality Metric) [180]. Поред модела за објективну процену квалитета, описане су и бројне пратеће технике калибрације (просторна и временска регистрација, одређивање валидног региона у кадру, усаглашавање динамичког опсега и сл.). Модел је евалуиран од стране VQEG групе у другој фази тестова са потпуним референцирањем (Phase II Full Reference Television – FR TV). Како се генерални модел показао јединим естиматором квалитета који је дао добре резултате на оба подскупа (525 и 625 линија), ANSI (American National Standards Institute) је прихватио NTIA генерални модел и придружене технике калибрације као северноамерички стандард у 2003. години. Такође, ITU је NTIA модел прихватила као нормативни метод у више препорука. Генерални модел покрива широк опсег квалитета и битских брзина (иницијални тестови спроведени су за битске брзине од 10 kbit/s до 45 Mbit/s). У основи мере налази се издвајање перцептуалних обележја из блокова пиксела (блок пиксела представља тродимензионални скуп суседних пиксела – просторно и временски), одређивање вредности квалитета/сличности и комбиновање појединачних вредности квалитета у крајњу вредност квалитета. Обележја се добијају коришћењем једноставних математичких функција – средња вредност и стандардна девијација. У збирној процени се користи седам параметара. Четири параметра су изведена од просторног градијента осветљаја, два параметра су изведена из хроминентних компоненти, док је један параметар изведен из контраста и кретања на сцени. Параметри који се посматрају су: губитак просторних информација, промена оријентације хоризонталних и вертикалних ивица (у дијагоналне), промена оријентације дијагоналних ивица (у хоризонталне и вертикалне), промена расподеле боје на слици, добитак квалитета кроз појачање и истицање ивица, здружена анализа промене контраста и кретања на сцени и анализа значајних промена боје.

Данас се у литератури могу наћи бројне технике које у процени квалитета користе очување информација о градијенту [181-193]. Информације о оријентацији ивица слике се у [181] користе за мерење блоковских артефаката у слици и видеу. Блоковски артефакти доводе до дисконтинуитета нивоа сивог суседних блокова, због

чега се оријентације ивица слике концентришу око 0° (хоризонтална ивица) и 90° (вертикална ивица). Процент пиксела слике са оријентацијом од 0° или 90° говори о степену дисторзије. Харисов детектор се у [182] користи за добијање информација о градијенту и процену квалитета слика. Поређење обележја изведених коришћењем Харисовог детектора спроводи се по узору на SSIM, а поређење се спроводи на позицијама код којих постоји значајна разлика између изворне и тест слике. Промена амплитуда и фаза градијента је искоришћена и у [183]. Параметри градијента су издвојени коришћењем Собел просторне маске, извршена је њихова квантизација и пресликавање у кодне речи. Хемингова дистанца кодних речи је искоришћена за процену квалитета.

Осим параметара ивица у процени квалитета битну улогу има и очување позиције ивице [184, 185]. Објективна мера предложена за процену квалитета слика у [184] дата је у форми односа броја позиција на којима се подударују ивице изворне и тест слике и броја ивичних пиксела референтне слике. Што су мапе ивица сличније то тест слика више личи на референту слику, а самим тим је бољег квалитета. Промена позиције центра масе блокова тест слике (11x11 пиксела) се у [185] посматра заједно са очувањем средње вредности нивоа сивог, очувањем момената и одржањем ентропије при дефинисању објективне мере са потпуним референцирањем.

Значај градијента је потврђен и кроз истраживања презентована у [186], где се очување локалних вредности амплитуда градијента пондерише са амплитудом градијента тест слике. Амплитуде градијента одређују се коришћењем просторне маске димензија 3x3 пиксела, а очување амплитуда градијента изведено је по узору на SSIM индекс. Кроз пондерисање се већи значај даје регионима богатим ивицама и текстуром него глатким регионима. Међутим, предложени приступ тестиран је на свега једној изворној слици и њених седам модификација. Предност очувања информација о градијенту над очувањем информација о осветљају представљена је кроз адитивни модел процене квалитета [187]. Очување информација о осветљају дато је у форми разлике нивоа сивог, док се очување информација о градијенту добија кроз поређење амплитуда градијента по узору на SSIM индекс. Значај локалних поређења осветљаја и амплитуда градијента изводи се из сличности

градијента. Уколико је вредност сличности градијента изворне и тест слике мала, постоји велика вероватноћа да је осветљај маскиран са деградацијом контраста и/или структуре. Због тога се у оваквим ситуацијама очувању осветљаја даје много мањи значај него очувању амплитуде градијента. У раду је додатно разматран ефекат маскирања кроз праг видљивости промена. Нешто већи број особина визуелног система човека у домену градијента уграђен у мултивеличинском приступу [188]. Разматрају се нелинеарност одзива HVS на градијент, униформност квалитета, праг детекције градијента и осетљивост на просторне фреквенције. Мера очувања амплитуде градијента по узору на меру очувања контраста код индекса структурне сличности се користи за процену квалитета у [189]. Градијент сваког пиксела одређен је за четири угла (смера), након чега се посматра промена амплитуде градијента дуж једне осе и њој ортогоналне осе – локални континуитет дуж једне осе и локални дисконтинуитет дуж ортогоналне осе.

Амплитуда градијента се у [190] одређује за сваки од канала HSV колор координатног система изворне и тест слике и користи за одређивање заједничке информације у сваком од канала. Збирна мера представља пондерисану суму мера из сва три канала при чему је највећи значај дат каналу интензитета (54%). Веома једноставна, рачунарски ефикасна, мера за процену квалитета слика представљена је у [191]. У процени квалитета користи се глобална варијација сличности градијента изворне и тест слике (већа вредност мере одговара већем степену деградације). Иако се ради о веома једноставној мери, њене перформансе су у рангу актуелних мера са потпуним референцирањем. Како би што више поједноставили меру процене аутори су за издвајање градијента користили Превит просторну маску. Информације о градијенту су осим за процену квалитета са потпуним референцирањем искоришћене и за процену квалитета са редукованим референцирањем [192], односно процену квалитета без референцирања [193]. Редуковано референцирање из [192] добија се додатним преносом информација о амплитуди градијента референтног сигнала у околини центра слике, где се захтева додатни пренос од 1/1290 дела изворног сигнала. Хоризонталне и вертикалне разлике вредности осветљаја се користе за процену квалитета JPEG2000 кодованих слика без коришћења изворног сигнала [193].

2.11. Критеријуми за евалуацију перформанси објективних мера процене квалитета слике и видеа

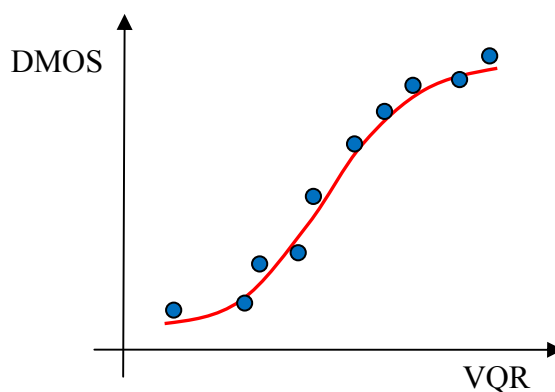
За описивање и поређење перформанси објективних мера процене квалитета слике/видеа могу се користити бројни атрибути. Најчешћи атрибути којима се одсликава успешност објективне мере у предикцији квалитета слике/видеа су: (1) тачност предикције, (2) монотоност предикције и (3) конзистентност предикције [5].

Метрике којима се мере ови атрибути се изводе из резултата објективних мера процене и резултата субјективних тестова на истом скупу тест слика/секвенци. Објективним и субјективним тестовима се за сваку изворну и тест секвенцу добијају нумеричке (бројне, скаларне) вредности које говоре о односу квалитета изворног видеа и тест секвенце. Подразумева се да се из резултата субјективних тестова користе средње вредности скорова (MOS/DMOS) и интервали поузданости кроз које се зарачунавају разлике унутар популације посматрача и разлике које могу постојати између лабораторија у којима су спроведени субјективни тестови.

Излази мера објективне процене квалитета – VQR (Video Quality Ratings), се пореде са резултатима субјективних тестова који могу бити дати кроз DMOS или MOS скорове квалитета. Веза између VQR и DMOS/MOS скорова не мора бити линеарна јер резултати субјективних тестова могу имати нелинеарну компресију (сабијање) скорова око екстремних вредности коришћеног опсега квалитета. Линеарност везе није критична, већ је битна стабилност њиховог односа и варијанса грешке њиховог односа кроз које се одређује успешност предикције објективних мера. Како би се елиминисале нелинеарности које су унете процесом субјективног оцењивања (Сл. 2.3) и олакшало поређење објективних мера у заједничкој анализи, веза између предикција објективне мере и субјективних скорова квалитета се посматра коришћењем нелинеарне регресије између VQR и DMOS скупова. Увођењем нелинеарног пресликавања тачност и конзистентност предикције се мењају, а монотоност предикције остаје иста.

Нелинеарна регресија се спроводи над [VQR,DMOS] скупом података уз услов да је монотона у опсегу VQR скорова. Облик нелинеарне регресије није критичан, уз услов да је монотон, генерално прихватљив и да има што мањи број

слободних параметара како би се олакшала интерполација података. Како природа нелинеарности није унапред позната, у регресији се користе различите функционалне форме за сваку од објективних мера и бира она која је најподеснија (са минималном средњеквадратном грешком) за дату меру.



Сл. 2.3 Пример везе између VQR и DMOS.

Најчешће функције које се користе у регресији су:

1. полином трећег реда са четири параметра

$$y = \text{Quality}(x) = A_0 + A_1x + A_2x^2 + A_3x^3 \quad (17)$$

2. логистик (logistic) функција са четири параметра

$$y = \text{Quality}(x) = \frac{\beta_1 - \beta_2}{1 + \exp\left(\frac{x - \beta_3}{|\beta_4|}\right)} + \beta_2 \quad (18)$$

3. логистик функција са пет параметара (са додатим линеарним чланом)

$$y = \text{Quality}(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(x - \beta_3))} \right) + \beta_4x + \beta_5 \quad (19)$$

где x представља скуп VQR вредности ($x \equiv \text{VQR}$) док y представља скуп предикција DMOS вредности, DMOS^p ($y \equiv \text{DMOS}^p$), а које се у даљем раду пореде са DMOS скоровима изведеним из субјективних тестова.

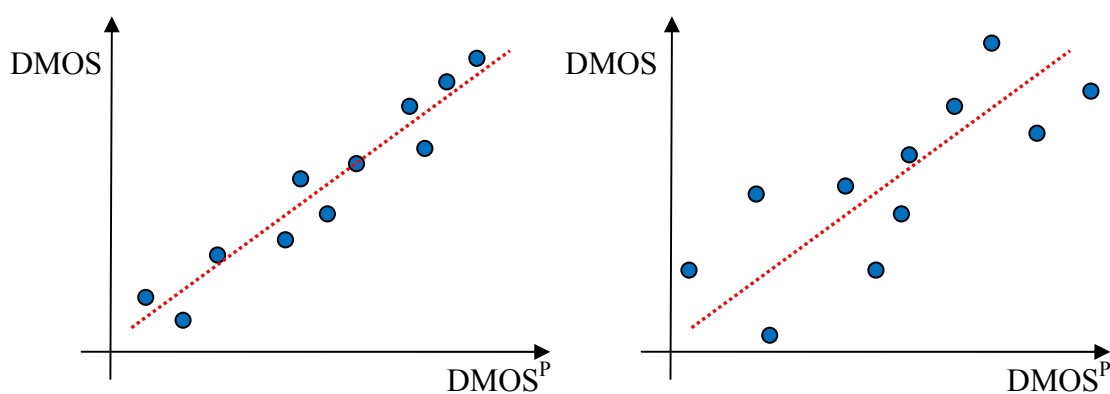
За нелинеарно пресликавање између објективних процена x и субјективних скорова y најчешће се користи логистик крива са четири параметра, која је прихваћена и у самој дисертацији. Параметри β_j ($j=1,2,3,4$) се одређују

оптимизацијом помоћу Nelder-Mead алгоритма како би се минимизирала сума квадрата грешака између објективних процена након нелинеарног пресликавања ($DMOS^P$) и субјективних процена ($DMOS$). Иницијалне вредности за процену параметара β_j су:

$$\beta_1 = \max(y), \beta_2 = \min(y), \beta_3 = \bar{x}, \beta_4 = 1. \quad (20)$$

2.11.1. Тачност предикције

Тачност представља способност објективне мере процене за предикцију $DMOS$ субјективних вредности са минималном просечном грешком [5]. На Сл. 2.4(а) приказани су резултати мере са мањом средњом грешком између $DMOS^P$ и $DMOS$ скорова у односу на меру са Сл. 2.4(б). Самим тим тачност предикције објективне мере са Сл. 2.4(а) је боља.



(а) објективна мера са већом тачношћу

(б) објективна мера са мањом тачношћу

Сл. 2.4 Поређење мера објективне процене квалитета кроз тачност предикције.

Бројне метрике се могу користити за одређивање средње грешке а уобичајено се користе средња апсолутна грешка (Mean Absolute Error, MAE) и корен средње квадратне грешке (Root Mean Squared Error, RMSE) између објективних процена након нелинеарног пресликавања и субјективних скорова. Ако се разлика између измерених $DMOS$ вредности и њихових предикција одреди као:

$$Error_i = DMOS_i - DMOS_i^P \quad (21)$$

где се индекс i односи на редни број анализираних тест секвенци, средња апсолутна грешка и корен средње квадратне грешке за скуп од N тест секвенци се добијају као:

$$MAE = \frac{1}{N} \sum_{i=1}^N |Perror_i| \quad (22)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Perror_i)^2} \quad (23)$$

Коефицијент линеарне корелације (Pearson Linear Correlation Coefficient, LCC), иако није директна мера средње грешке, је такође честа метрика која се користи за одређивање тачности предикције. Нижим вредностима MAE и RMSE (које су пожељне и у идеалном случају једнаке нули), одговарају више вредности коефицијента корелације (који је у идеалном случају једнак јединици). За скуп од N парова (x_i, y_i) коефицијент линеарне корелације се дефинише као:

$$LCC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (24)$$

где су \bar{x} и \bar{y} средње вредности подскупова x и y који у разматраном случају представљају DMOS и DMOS^P скупове тачака.

2.11.2. Монотоност предикције

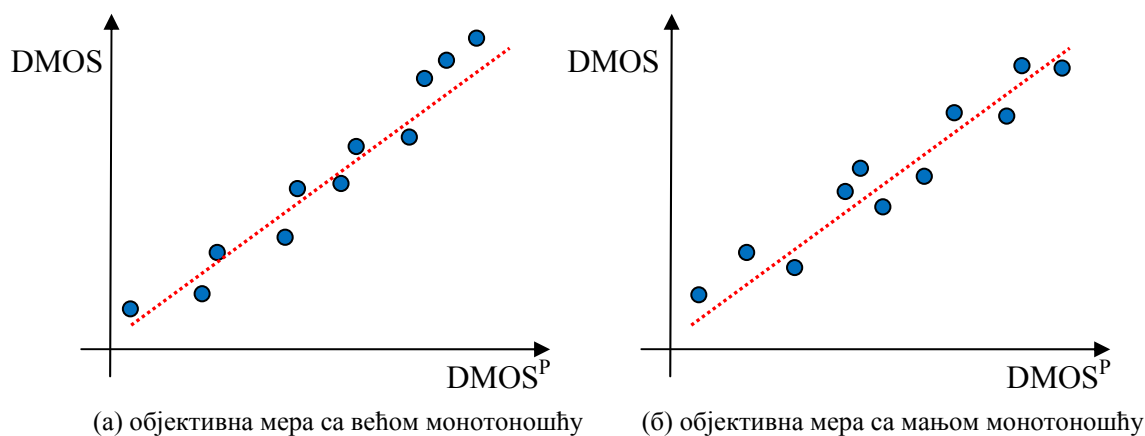
Монотоност представља степен у коме се предикције објективне мере слажу са релативним амплитудама субјективних процена [5]. DMOS^P вредности изведене из резултата објективне мере идеално би требало да буду потпуно монотоне у односу на упарене DMOS вредности, тј. промена DMOS^P вредности треба да има исти знак као и промена DMOS скорова.

На Сл. 2.5 илустроване су хипотетичке везе између DMOS^P и DMOS за две мере различитих монотоности. Мере имају приближно исту тачност гледану кроз просечну грешку али је монотоност предикције на Сл. 2.5(а) боља. Монотоност предикције на Сл. 2.5(б) је лошија што се види кроз пад DMOS^P вредности у односу на оно што стварно виде посматрачи а што је дато кроз повећање DMOS вредности.

Монотоност предикције се може мерити коефицијентом корелације рангова (Spearman Rank-Order Correlation Coefficient, SROCC), који се дефинише као:

$$SROCC = \frac{\sum_{i=1}^N (\chi_i - \bar{\chi})(\gamma_i - \bar{\gamma})}{\sum_{i=1}^N (\chi_i - \bar{\chi})^2 \sum_{i=1}^N (\gamma_i - \bar{\gamma})^2} = 1 - \frac{6 \cdot \sum_{i=1}^N (\chi_i - \gamma_i)^2}{N(N^2 - 1)} \quad (25)$$

где су χ_i и γ_i рангови вредности x_i и y_i а $\bar{\chi}$ и $\bar{\gamma}$ средње вредности рангова. Коефицијент корелације рангова је непараметарски тест, тј. не захтева унапред познату везу између x_i и y_i (линеарна, полиномијална, и сл.), односно не зависи од функције која се користи за интерполацију између VQR објективних и DMOS^P субјективних скорова. Кроз овај параметар се пореди промена између суседних парова DMOS вредности са променом унутар одговарајућих DMOS^P вредности. Како SROCC ради само са ранговима (редоследом) података и игнорише релативне дистанце између њих, он се узима као мера корелације са мањом осетљивошћу и типично се користи уколико је број тачака (одбирака) мали. Са повећањем вредности SROCC расте и монотоност објективне мере процене квалитета (у идеалном случају SROCC=1).

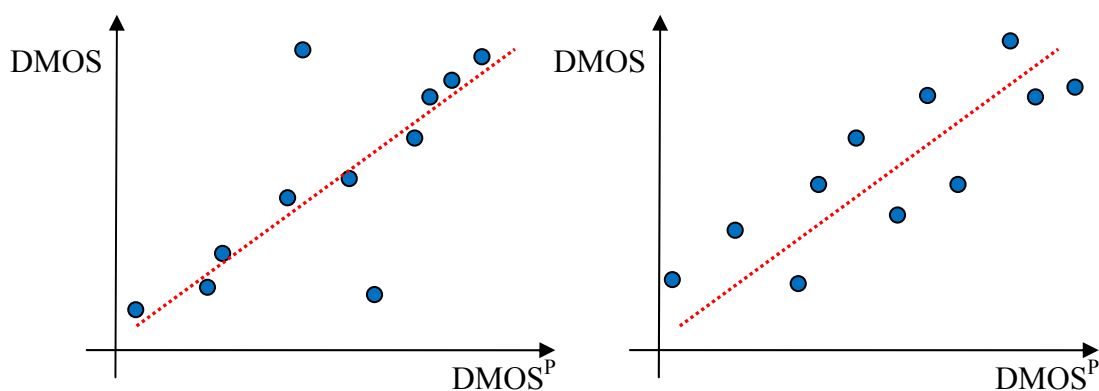


Сл. 2.5 Поређење мера објективне процене квалитета кроз монотоност предикције.

2.11.3. Конзистентност предикције

Овај атрибут је повезан са степеном у коме мера објективне процене задржава тачност предикције кроз цео скуп анализираних видео секвенци [5]. Од објективне мере се захтева доследност за све типове видео секвенци, односно да се не појави прекомерно одступање за део анализираних секвенци.

На Сл. 2.6 приказани су резултати изведени из две објективне мере процене са приближно истим вредностима MAE/RMSE између DMOS и $DMOS^P$ скупова. Сл. 2.6(a) је пример мере која за већину од секвенци има потпуно тачне предикције али има и велике грешке предикције за две тачке у средњем делу слике. Сл. 2.6(b) је пример мере која има уједначене (балансиране) грешке предикције – за већину секвенци она није тачна као мера са Сл. 2.6(a) али се уочава њена доследност кроз прихватљиве предикције за све секвенце.



(a) објективна мера са великим грешкама предикције (b) објективна мера са уједначеним грешкама

Сл. 2.6 Поређење мера објективне процене квалитета кроз конзистентност предикције.

Конзистентност објективне мере процене се може одређивати кроз број тачака за које је грешка предикције већа од заданог прага. За меру која даје мањи број оваквих тачака каже се да је доследнија. За праг се најчешће узима двострука вредност стандардне грешке DMOS скорова, $DMOS^{SE}$, тј. грешка предикције постоји ако је:

$$|Perror_i| > 2 \cdot DMOS_i^{SE} \quad (26)$$

Број грешака предикције у односу на укупан број тачака назива се OR однос (Outlier Ratio, OR) и користи се као мера конзистентности објективних процена квалитета. За објективну меру са мањим OR односом каже се да је доследнија.

Као што је извршено пресликавање VQR скорова квалитета у предикције DMOS скорова, може се извршити пресликавање VQR скорова у предикције MOS скорова (MOS^P). Како се референтни видео користи код FR и RR мера објективне процене, препорукама ITU се за њихову евалуацију сугерише коришћење DMOS и

DMOS^P скупова. Истовремено, поређење MOS и MOS^P скупова се предлаже у евалуацији NR објективних мера процене (код којих није доступна референтна секвенца).

Перформансе објективних мера процене квалитета се у дисертацији евалуирају су кроз сва три описана аспекта (атрибута) предикције:

- *Тачност предикције*, одређену рачунањем коефицијента корелације, LCC, средње апсолутне грешке, MAE, и корена средње квадратне грешке, RMSE, између објективних процена након нелинеарне регресије (DMOS^P/MOS^P) и субјективних скорова (DMOS/MOS);
- *Монотоност предикције*, одређену Спирмановим коефицијентом корелације рангова, SROCC, између објективних процена након нелинеарне регресије (DMOS^P/MOS^P) и субјективних скорова (DMOS/MOS); Како SROCC не зависи од функције која се користи у пресликавању VQR скорова у DMOS^P/MOS^P скорове (битан је редослед), корелација рангова се може директно одређивати кроз поређење рангова VQR и DMOS/MOS скорова (тј. без регресије);
- *Конзистентност предикције*, одређену рачунањем OR односа; За праг грешке процене користи се двострука вредност стандардне грешке DMOS/MOS скорова.

2.12. Закључак

Анализа јавно доступних база слика и видеа намењених објективној процени квалитета може се наћи у [26], док се компаративне анализе објективних мера и актуелни трендови у овој области могу наћи у [194-213]. У већини прегледних радова потенцирано је да вршни однос сигнал шум није најприкладнија техника процене квалитета. Додатно је наглашено да перформансе објективних мера процене квалитета значајно варирају од базе до базе и зависе од типа деградације, чиме универзална мера процене квалитета не постоји. Такође, пондерисањем локалних скорова квалитета може се остварити добитак у односу на стандардну имплементацију објективне мере.

Актуелни правци истраживања у области процене квалитета слика и видеа су [212, 213]: (1) процена квалитета слика са израженом текстуром, (2) анализа утицаја динамичког опсега и резолуције, (3) примене у медицини, (4) здружени аудио/визуелни модели процене, (5) процена квалитета сигнала са вишеструком дисторзијом и други.

3. Процена квалитета слика заснована на очувању информација о градијенту

3.1. Увод

Објективна процена квалитета слике има значајну улогу у различитим поступцима обраде слике и видеа, као што су компресија, интерполација, уклањање шума у слици и слично, при чему се успешност неког поступка обраде процењује на основу његовог утицаја на квалитет слике.

Данас се све више користе дигиталне слике као извор информација. Ове слике су током снимања, обраде, преноса и приказа подложне разним врстама изобличења, због чега је за систем преноса веома битно идентификовати и одредити степен деградације квалитета слике.

Објективна процена квалитета слике разликује се у аналогним и дигиталним системима. Код аналогних система објективна процена својстава телевизијских сигнала заснива се на мерењима коришћењем различитих тест сигнала. Тест сигнали су посебно обликовани сигнали који омогућавају контролу и мерење тачно одређеног параметра квалитета видео сигнала. Објективна мерења у дигиталним системима нису могућа таквим начином јер се грешке које настају могу најбоље уочити на самој тест слици. Због тога се и користе објективне мере које су засноване на разлици оригиналне и тест слике. Главни циљ објективних мера је што већа корелација са резултатима субјективних тестова, независно од врсте и степена изобличења у слици, и што већа брзина прорачунавања за примене у реалном времену.

За одређивање квалитета тест слике у овом поглављу је предложен FR приступ. Наведени приступ је проистекао из приступа који се користи у процени квалитета сједињавања изворних слика [179, 214], где се очување просторних информација евалуира кроз градијентну репрезентацију. Градијентна процена полази од идеје да је пренос (компресија, деградација) којим се пренесе више градијентних информација бољи. У процени се полази од издвајања параметара градијента и

коришћења субјективног модела очувања градијента. Разматрањем субјективног значаја региона слике, очување параметара градијента локалних региона се сабира у једну нумеричку вредност која одсликава квалитет тест слике и има вредност у опсегу од нула до један. Предложена мера је упоређена са PSNR, SSIM, VIF и VSNR мерама на четири референтне базе слика са доступним субјективним скоровима. Осим стандардне SSIM мере [48] у раду су дати и резултати мере из које је она изведена – универзалног индекса квалитета слике (Universal Image Quality Index, UIQI) [47] и њене мултивеличинске варијанте MS-SSIM [215].

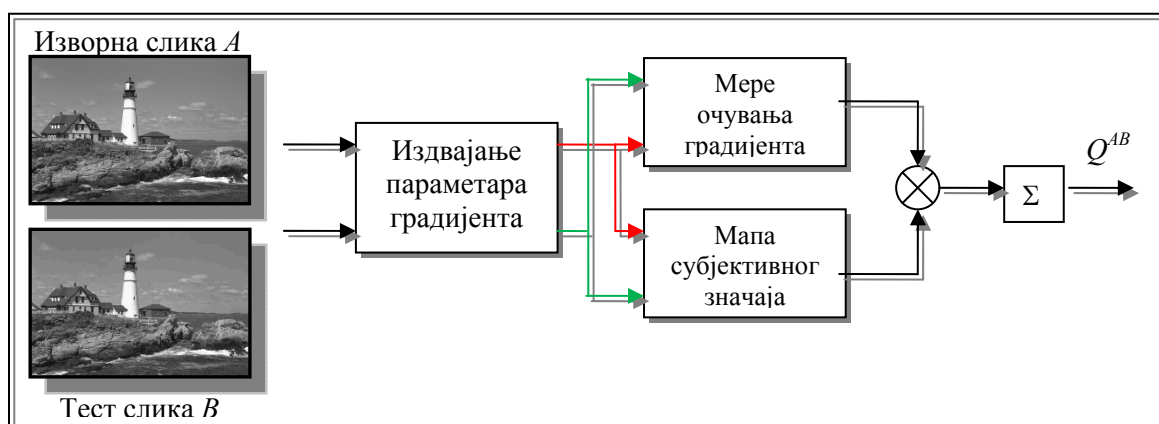
3.2. Издајање параметара градијента

Током визуелног побуђивања посматрачи су мотивисани да решавају неизвесности (uncertainties), које се на слици манифестују кроз промене нивоа сивог или промене боје. У реалним сликама, промене нивоа сивог нису концентрисане у унапред познатим регионима слике већ су распоређене у складу са садржајем слике. Посматрач тражи визуелне надражаје у којима постоје неизвесности и из њих издаја информације. Међутим, информације нису само детектоване промене нивоа сивог. Одсуство промене сигнала (нулта ивица – zero edge) такође носи малу али коначну количину информација, тј. информацију о томе да нема промена нивоа сивог. Због тога, како би се обухватиле све информације на слици, морају се све неизвесности узети у обзир. То је урађено одређивањем информација о градијенту на свим позицијама презентоване слике. Повезивање визуелних информација са градијентом је подржано са студијама о људском визуелном систему [179, 214, 216] и доста се користи у анализи слике и алгоритмима компресије. Због тога се квалитет слике процењује одређивањем верности репрезентације параметара градијента изворне слике у тест слици. При процени квалитета мора се водити рачуна о субјективном значају параметара градијента слике. Тест слика која сачува више битних ивица изворне слике има бољи квалитет од тест слике која очува мање битних ивичних информација.

Генерална структура градијентне (ивичне) објективне процене квалитета приказана је на Сл. 3.1. Прво се врши издајање параметара градијента изворне и

3. Процена квалитета слика заснована на очувању информација о градијенту

тест слике. Затим се одређују мере очувања параметара градијента, које приказују колико су верно пресликане изворне ивице у тест слици. Одређивањем субјективног значаја пиксела изворних слика добијају се нумеричке вредности којима се модулишу мере очувања градијента и резултат интегрални по свим пикселима. Као резултат добија се нумеричка вредност Q^{AB} која одсликава квалитет тест слике – тачност са којом су изворне информације представљене у тест слици.



Сл. 3.1 Генерална структура градијентне објективне процене квалитета.

Расподела градијента реалних слика зависи од садржаја слике. Због ове неизвесности, да би се издвојиле све информације слике, параметри градијента се рачунају за сваки пиксел понаособ. Визуелне информације слика, на свакој позицији, су представљене амплитудама и оријентацијама градијента. Параметри градијента се издвајају коришћењем Собел (Sobel) оператора приказаног на Сл. 3.2, димензија 3x3 пиксела. Мотивација за избор Собел оператора је да се свим пикселима обезбеди исти утицај у формирању градијента. У односу на једноставнију Превит (Prewitt) маску Собел маска је робуснија на утицај шума [217, 218]. Просторне маске представљају хоризонталне и вертикалне операторе помоћу којих се издвајају компоненте градијента по хоризонтали и вертикали.

1	0	-1
2	0	-2
1	0	-1

-1	-2	-1
0	0	0
1	2	1

Сл. 3.2 Вертикална и хоризонтална Собел маска.

За потребе објективне процене потребно је обе слике филтрирати са просторним маскама. У спроведеној анализи слика је представљена као матрица димензија $N \times M$, где је N број колона а M број редова. Резултате конволуције слика који дају информације о хоризонталним и вертикалним компонентама ивица (x и y оријентација) означимо са s^x и s^y , тим редом. Од ових компоненти добијају се амплитуда $g(n,m)$ и оријентација градијента $\alpha(n,m)$ пиксела на позицији (n,m) слике (рецимо слике A):

$$g_A(n,m) = \frac{\sqrt{(s_A^x(n,m))^2 + (s_A^y(n,m))^2}}{g_{\max}}, \quad g_{\max} = 4.472, \quad 0 \leq g_A(n,m) \leq 1 \quad (27)$$

$$\alpha_A(n,m) = \arctan\left(\frac{s_A^y(n,m)}{s_A^x(n,m)}\right), \quad -\pi \leq \alpha_A(n,m) \leq \pi \quad (28)$$

У (27) је са g_{\max} означена максимална вредност градијента која се може добити применом Собел маске ($g_{\max}=4.472$).

Оба параметра градијента су ограничена. Амплитуде g могу бити у опсегу $[0,1]$, где минимална вредност одговара ситуацијама када нема промена нивоа сивог, а максимална вредност одговара максималном контрасту. Оријентације α могу бити у опсегу $[-\pi, \pi]$, где вредности $-\pi$, 0 и π одговарају хоризонталним ивицама а вредности $-\pi/2$ и $\pi/2$ вертикалним ивицама. Оријентација је циклична променљива, тј. екстремне вредности $-\pi$ и π су у основи исте. Вредности оријентације једнаке су нули и код пиксела са нултим хоризонталним и вертикалним компонентама градијента.

На Сл. 3.3 приказане су амплитуде и оријентације градијента слике “Lena”.



Сл. 3.3 (а) изворна слика, (б) амплитуде градијента и (в) оријентације градијента.

На Сл. 3.3(б) се уочавају региони који имају изражене ивице (бели пиксели). Са Сл. 3.3(в) види се да оријентација градијента има случајан карактер, нарочито у регионима у којима нису изражене ивице.

3.3. Мере очувања параметара градијента

Параметри градијента се чувају у мапама (матрицама) амплитуда и оријентација и у другој фази се користе за процену њиховог очувања. На овај начин се добијају нове матрице које говоре о степену њиховог очувања.

Мера очувања градијента је заснована на поређењу параметара градијента изворне и тест слике, при чему се поређење врши за сваки пиксел понаособ. Ивица изворне слике је сасвим преликана у тест слику ако су амплитуде и оријентације градијента на свим пикселима преко којих се протеже остале непромењене. Промена једног или оба параметра је показатељ губитка информација. Ова мера показује релативне промене параметара градијента које су настале у току преноса (компресије).

Промена амплитуде градијента, Δ_g^{AB} , се дефинише као однос амплитуда изворне и тест слике:

$$\Delta_g^{AB}(n, m) = \begin{cases} \frac{g_B(n, m) + C}{g_A(n, m) + C}, & g_A(n, m) > g_B(n, m) \\ \frac{g_A(n, m) + C}{g_B(n, m) + C}, & g_A(n, m) \leq g_B(n, m) \end{cases}, \quad C = \frac{4}{256} = \frac{1}{64} \quad (29)$$

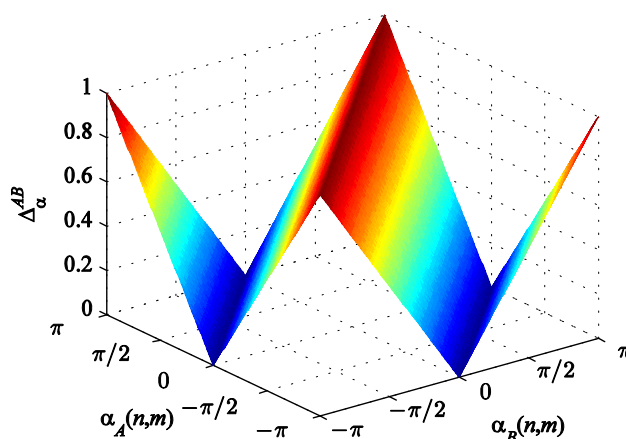
У (29) је са C означена константа која се додаје због ситуација у којима су оба градијента веома близу нуле или прага детекције ивице. Под нормалним околностима и условима посматрања људски визуелни систем може правити разлику између 64 нивоа осветљаја (нивоа сивог) на скали између максималног осветљаја (бело) и минималног осветљаја (црно). То значи да ће, игноришући ефекте просторне учестаности, на 8-битним монохроматским сликама, посматрачи у просеку разликовати два региона уколико се њихови средњи осветљаји разликују за четири или више нивоа сивог [219-221]. Због тога је у раду за константу C усвојена вредност $4/256$ ($1/64$).

Из израза (29) види се да Δ_g^{AB} има вредност један ако је $g_B=g_A$ и да линеарно пада ка нули уколико разлика амплитуда референтне и тест слике расте. Такође, није битно који градијент је већи, тест слике или референтне слике. Последица тога је да се било каква промена амплитуде градијента од изворне до слике са деградацијом разматра као губитак информација у преносу/компресији и даје резултат мањи од један.

Промена оријентације градијента слике A у односу на слику B , Δ_α^{AB} , се дефинише као нормализована разлика оријентација α_A и α_B :

$$\Delta_\alpha^{AB}(n, m) = \frac{\left| \alpha_A(n, m) - \alpha_B(n, m) \right| - \pi}{\pi} \quad (30)$$

Из израза (30) види се да Δ_α^{AB} има вредност један ако је $\alpha_B=\alpha_A$ и да линеарно пада како разлика оријентација референтне и тест слике расте. Промене оријентације за све могуће вредности α_A и α_B су приказане на Сл. 3.4.



Сл. 3.4 Промена оријентације за све могуће комбинације оријентација градијента референтне и тест слике.

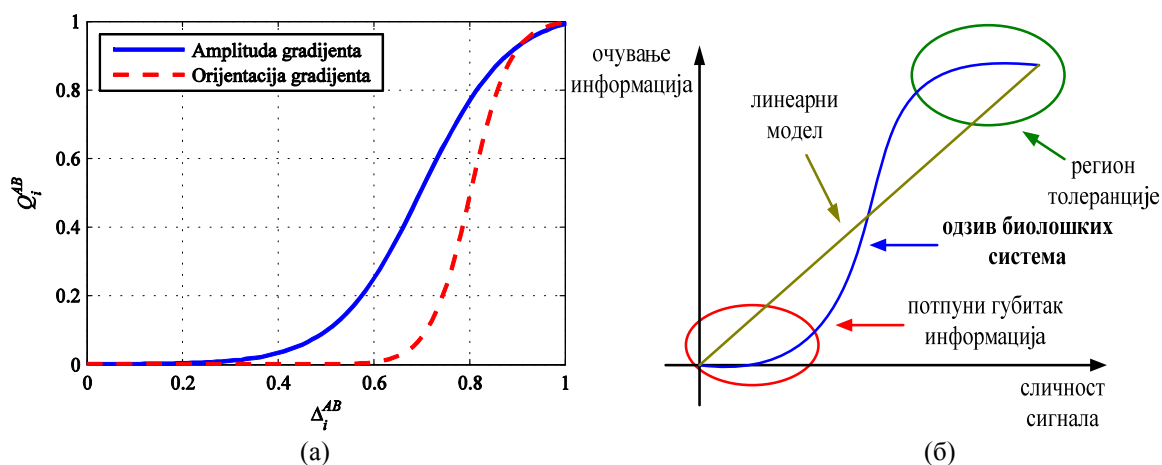
Са Сл. 3.4 се види да Δ_α^{AB} има максималну вредност дуж дијагоналне линије на средини графика ($\alpha_A=\alpha_B$), и у екстремним случајевима, тј. када α_A и α_B узимају једну од екстремних вредности (циклична променљива). Насупрот томе, вредност Δ_α^{AB} је једнака нули када је разлика између α_A и α_B максимална (π) – дуж две краће дијагонале на графику. Између ових екстремних вредности разлика оријентација линеарно пада са порастом разлике између α_A и α_B .

Вредности Δ_g^{AB} и Δ_α^{AB} описују линеарне промене параметара градијента. Међутим, како субјективни губитак информација није линеарно зависан од промене амплитуда и оријентација у обзир је узета и нелинеарна природа визуелног система човека. Субјективни губитак амплитуда и оријентација градијента моделован је са нелинеарним сигмоидним функцијама (релација (31) и Сл. 3.5). Коришћењем нелинеарних сигмоидних функција добијају се мере очувања параметара градијента, Q_g^{AB} и Q_α^{AB} , које говоре о субјективној верности улазних амплитуда и оријентација на позицији (n, m) у тест слици:

$$Q_i^{AB}(n, m) = \frac{\Gamma_i}{1 + \exp\{k_i[\Delta_i^{AB}(n, m) - \sigma_i]\}}, \quad i \in \{g, \alpha\}. \quad (31)$$

Сигмоидне функције су погодне за моделовање одзива биолошких система који је нелинеаран. Широко се користе у применама неуралних мрежа. У том контексту се могу користити и за моделовање субјективног губитка визуелних информација. Константе Γ_i у изразу (31) се добијају тако да је $Q_i^{AB}=1$ када је $\Delta_i^{AB}=1$.

Очување параметара градијента у односу на промене амплитуда и оријентација приказане су на Сл. 3.5(а). Приказане функције су добијене на основу параметара $[k_g, \sigma_g, k_\alpha, \sigma_\alpha] = [-11, 0.7, -24, 0.8]$. Дате вредности параметара нелинеарних сигмоидних функција су добијене оптимизацијом резултата мере за објективну процену квалитета сједињавања и субјективних процена [214, 221], а показале су се прихватљивим и у објективној процени квалитета слика [222].



Сл. 3.5 (а) нелинеарне сигмоидне функције које се користе у очувању параметара градијента, (б) нелинеарни сигмоидни модел очувања информација који се користи за моделовање одзива биолошких система и линеарни модел очувања информација (модел репрезентације сигнала).

Са Сл. 3.5 види се да посматрачи толеришу ситна одступања, односно мале промене параметара градијента изворне слике имају ограничен утицај на вредности одржања параметара градијента. Како разлике параметара градијента расту, промене постају уочљиве, тако да сличност брзо опада. Уколико се уочи значајна разлика параметара градијента, Q_i^{AB} постаје једнако нули. Треба приметити и да је визуелни систем човека мање осетљив на промене амплитуде градијента (које одговарају промени контраста) него на промене у оријентацији (које одговарају промени облика у слици). Чак и мале промене оријентације ивица условљавају промену облика, тако да карактеристике слика постају битно различите. Због тога Q_α^{AB} опада много брже, и за промену оријентације од 35 степени ($\Delta_\alpha^{AB}=0.61$) долази до губитка свих информација ($Q_\alpha^{AB}=0.006$).

Мере очувања амплитуда и оријентација $Q_g^{AB}(n,m)$ и $Q_\alpha^{AB}(n,m)$ се комбинују у меру очувања градијента $Q^{AB}(n,m)$, која представља меру верности којом слика B представља слику A :

$$Q^{AB}(n,m) = \sqrt{Q_g^{AB}(n,m) \cdot Q_\alpha^{AB}(n,m)} \quad (32)$$

Крајња мера квалитета тест слике B , Q^{AB} , се добија као средња вредност очувања параметара градијента свих пиксела изворне слике, $Q^{AB}(n,m)$:

$$Q^{AB} = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M Q^{AB}(n,m) \quad (33)$$

Вредности мере Q^{AB} се налазе у опсегу $[0,1]$, где вредност нула указује да је дошло до потпуног губитка информација изворне слике, док вредност један указује да се ради о идеалном преносу при коме су информације изворне слике у потпуности преликане у тест слику. У пракси, екстремне вредности се веома ретко срећу. За тест слику која има већу вредност Q^{AB} кажемо да је бољег квалитета.

3.4. Резултати примене мере очувања градијента

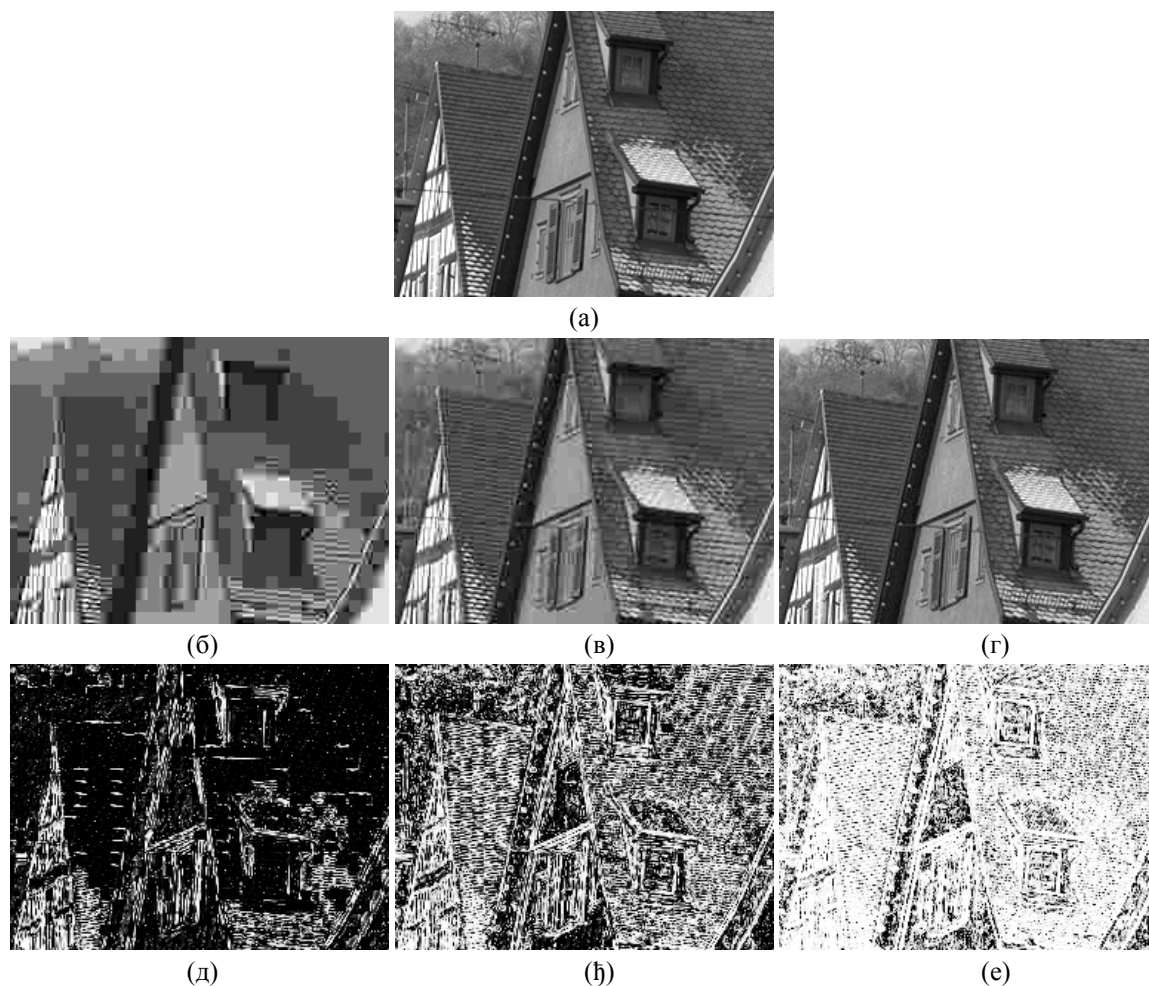
На Сл. 3.6 приказане су изворна слика “Buildings“, три њене модификације настале JPEG компресијом и мапе квалитета настале применом мере очувања градијента. Степен компресије тест слика је различит. Изворне слике су димензија

3. Процена квалитета слика заснована на очувању информација о градијенту

768x512 пиксела а због бољег приказа на Сл. 3.6 су приказани њихови делови димензија 256x192 пиксела.

Са Сл. 3.6 види се да тест слике имају различит визуелни квалитет, а са мапа квалитета види се да се са повећањем степена компресије (смањење бпр) очува и мање информација о градијенту изворне слике (црни пиксели \equiv 0, бели пиксели \equiv 1).

Степени компресије, субјективне импресије квалитета и вредности анализираних објективних мера процене дати су у Таб. 3.1. Субјективни квалитет дат је кроз DMOS, где ниже вредности одговарају бољем квалитету. Изузев MSE, више вредности објективних мера процене одговарају бољем квалитету тест слика.



Сл. 3.6 (а) изворна слика, (б) (в) (г) тест слике са степеном компресије од 0.25 бпр, 0.58 бпр и 1.78 бпр, тим редом, (д) (ђ) (е) мапе квалитета добијене применом мере очувања градијента на слике из претходног реда.

3. Процена квалитета слика заснована на очувању информација о градијенту

Таб. 3.1 Вредности субјективних и објективних процена тест слика са Сл. 3.6.

Слика	Степен компресије	Субјективни квалитет	Вредности мера објективне процене квалитета					
	<i>bpp</i>	<i>DMOS</i>	<i>MSE</i>	<i>PSNR (dB)</i>	<i>UIQI</i>	<i>SSIM</i>	<i>MS-SSIM</i>	Q^{AB}
Сл. 3.6(б)	0.24723	89.6422	623.07	20.18	0.45	0.56	0.83	0.23
Сл. 3.6(в)	0.58384	42.0432	207.45	24.95	0.71	0.79	0.96	0.46
Сл. 3.6(г)	1.7772	13.7032	40.72	31.96	0.90	0.94	0.99	0.75

Из Таб. 3.1 види се да промена степена компресије утиче како на субјективне импресије квалитета тако и на вредности објективних мера процене. Како је један од циљева објективне процене квалитета правилно рангирање слика по субјективном квалитету, из Таб. 3.1 се види да се рангирање које су дали посматрачи слаже са рангирањима свих анализираних објективних мера процене. Вредности добијене применом мере очувања градијента боље одсликавају стварни квалитет изворних слика (вредности процене су из опсега 0.23 до 0.75), за разлику од нпр. MS-SSIM (вредности процене су из опсега 0.83 до 0.99).

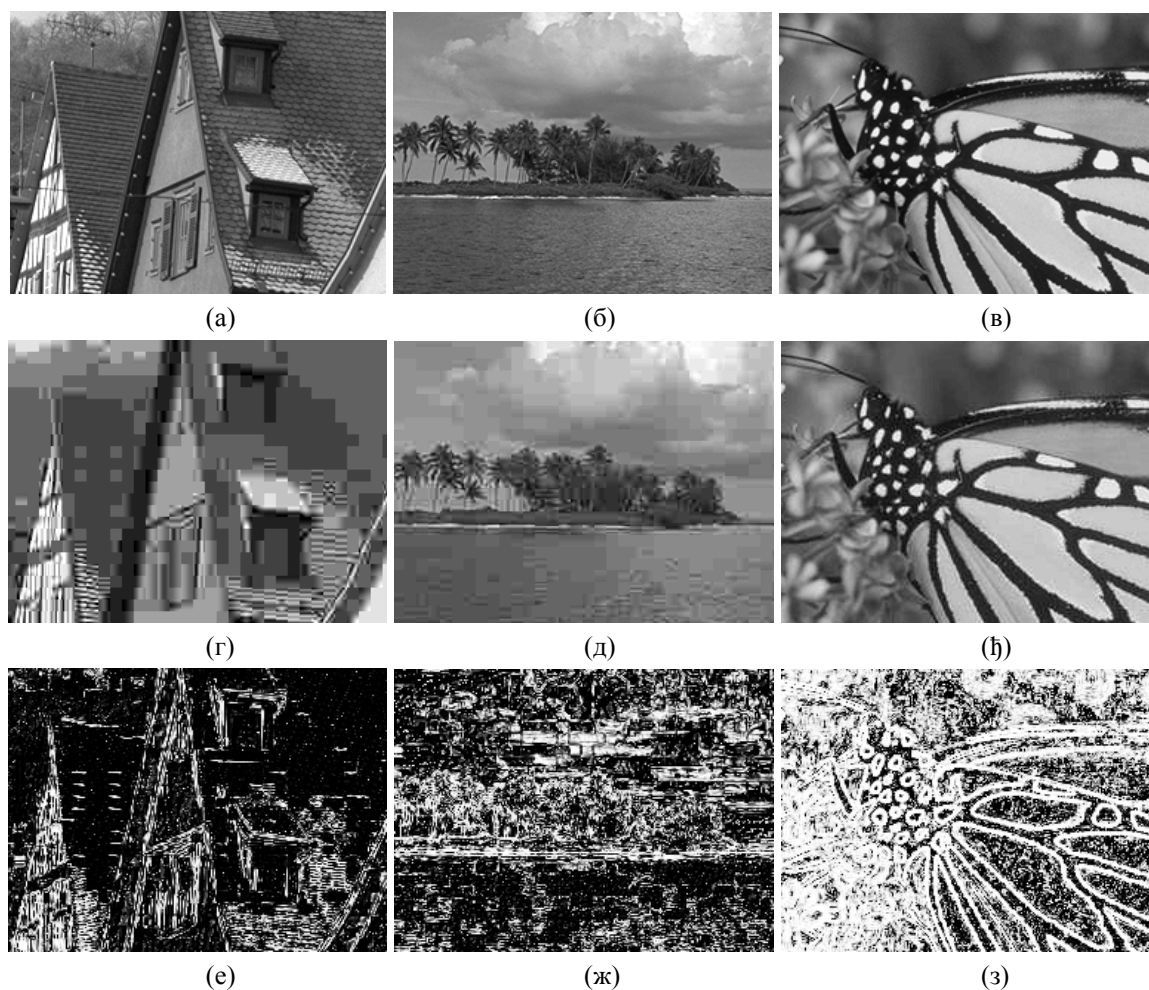
На Сл. 3.7 и Сл. 3.8 приказане су тест слике различитог квалитета настале применом JPEG и JPEG2000 компресије, тим редом. Степен компресије тест слика је различит. Треба запазити да груба квантизација настала коришћењем JPEG и JPEG2000 компресије доводи до губљења финих детаља изворних слика (цреп на Сл. 3.7, односно дрвеће на Сл. 3.8). Како су у овим регионима структурне информације скоро потпуно изгубљене то условљава и лошији визуелни квалитет.

У Таб. 3.2 и Таб. 3.3 дати су степени компресије, вредности субјективних и вредности објективних процена тест слика са Сл. 3.7 и Сл. 3.8, тим редом.

Из Таб. 3.2 види се да се рангирања слика са JPEG компресијом од стране посматрача у потпуности слажу са рангирањем добијеним објективним мерама квалитета. Код слика са JPEG2000 компресијом субјективни рангови се не слажу са

3. Процена квалитета слика заснована на очувању информација о градијенту

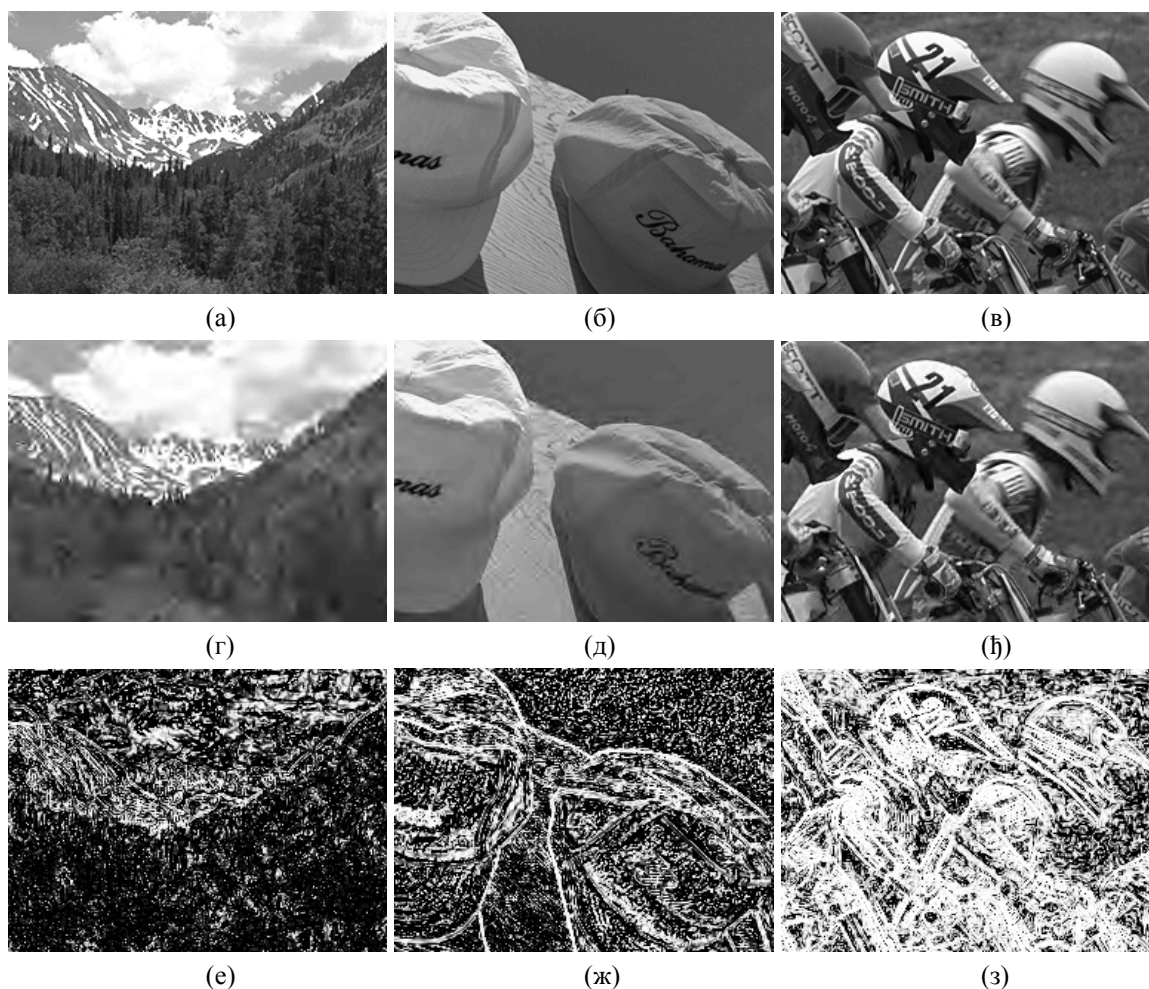
ранговима добијеним применом MSE (PSNR) и UIQI (Таб. 3.3). Такође, из Таб. 3.3 се види да садржај изворне слике значајно утиче на субјективну импресију квалитета. Наиме, исти степен компресије изворних слика, Сл. 3.8(а) и Сл. 3.8(б), даје значајно различит субјективни квалитет тест слика, Сл. 3.8(г) и Сл. 3.8(д).



Сл. 3.7 (а) (б) (в) изворне слике, (г) (д) (ђ) тест слике настале применом JPEG компресије са 0.27 бпр, 0.30 бпр и 0.78 бпр, тим редом, (е) (ж) (з) мапе квалитета/дисторзије тест слика.

Како се кроз структурна поређења SSIM (MS-SSIM) и Q^{AB} мера одсликава очување садржаја изворних слика, добијају се бољи резултати процене него код MSE (PSNR) где се поређење спроводи на нивоу пиксела, тј. не узима се у обзир структура садржана у изворним сигналима.

3. Процена квалитета слика заснована на очувању информација о градијенту



Сл. 3.8 (а) (б) (в) изворне слике, (г) (д) (љ) тест слике настале применом JPEG2000 компресије са 0.19 бпр, 0.20 бпр и 1.15 бпр, тим редом, (е) (ж) (з) мапе квалитета/дисторзије тест слика

Таб. 3.2 Вредности субјективних и објективних процена тест слика са Сл. 3.7.

Слика	Степен компресије	Субјективни квалитет	Вредности мера објективне процене квалитета					
	bpp	DMOS	MSE	PSNR (dB)	UIQI	SSIM	MS-SSIM	Q^{AB}
Сл. 3.7(е)	0.26731	58.2719	558.90	20.65	0.48	0.59	0.84	0.25
Сл. 3.7(д)	0.29797	45.0182	71.80	29.53	0.49	0.78	0.92	0.32
Сл. 3.7(љ)	0.77547	29.6278	18.29	35.36	0.71	0.95	0.99	0.56

Таб. 3.3 Вредности субјективних и објективних процена тест слика са Сл. 3.8.

Слика	Степен компресије	Субјективни квалитет	Вредности мера објективне процене квалитета					
	bpp	DMOS	MSE	PSNR (dB)	UIQI	SSIM	MS-SSIM	Q^{AB}
Сл. 3.8(е)	0.19204	72.8170	396.87	22.14	0.45	0.51	0.84	0.25
Сл. 3.8(д)	0.19944	46.5843	30.81	33.15	0.44	0.88	0.96	0.36
Сл. 3.8(љ)	1.1466	27.4472	39.64	32.08	0.89	0.92	0.99	0.65

3. Процена квалитета слика заснована на очувању информација о градијенту

Осим кроз претходна три примера, перформансе предложене објективне процене упоређене су са перформансама конвенционалне PSNR мере и перформансама мера које се због добре предикције субјективних процена најчешће користе: UIQI, SSIM, MS-SSIM, VIF и VSNR. Поређење је извршено на референтним базама слика са доступним субјективним проценама: LIVE [223], Тоуама [224], IVC [225] и CSIQ базама [226]. Основни подаци о базама слика дати су у Таб. 3.4 (део података о базама преузет је из [227]). Коришћење више база омогућава да се стекне шира слика о погодности мере објективне процене квалитета, њеним предностима и манама, ефектима које имају различити типови деградације на успешност мере и сл.

Таб. 3.4 Основни подаци о условима при којима су прикупљени резултати субјективних тестова референтних база слика.

	<i>LIVE database</i>	<i>Toyama database</i>	<i>IVC database</i>	<i>CSIQ database</i>
Монитор	CRT, 21"	CRT, 17"	CRT	Sceptre X24WG LCD, 24"
Резолуција монитора	1024x768	1024x768	-	1920x1200
Даљина посматрања	2-2.5 S (S – висина монитора)	4 H (H – висина слике)	6 S (S – висина монитора)	70 cm
Број референтних слика	29	14	10	30
Димензије изворних слика	углавном 768x512	768x512	512x512	512x512
Тип деградације	JPEG, JPEG2000, White noise, Gaussian blurring, Fast fading	JPEG, JPEG2000	JPEG, JPEG2000, Gaussian blurring, LAR coding	JPEG, JPEG2000, Gaussian blurring, global contrast decrements, additive pink noise, additive Gaussian noise
Број различитих деградација	5	2	4	6
Број нивоа деградације	5-9	6	5	4-5
Побуда	Single Stimulus (са скривеном референцом)	Single Stimulus	Double Stimulus Impairment Scale (DSIS)	Categorical Subjective Image Quality
Број посматрача	20-29	16	15	25
Просечан број посматрача по слици	23	16	15	5-7
Број слика са деградацијом	779	168	235	866
Формат слика	BMP	BMP	BMP	PNG

Из Таб. 3.4 се види да се услови у којима су прикупљени резултати субјективних тестова значајно разликују од базе до базе. Резолуције монитора и даљине посматрања тест слика, а од којих зависи субјективна импресија квалитета [215], се значајно разликују. Број референтних слика се креће од 10 до 30. Базе садрже од два до шест типова деградације изворног сигнала, при чему се број нивоа деградације креће од четири до девет. У свим базама слика постоје слике са JPEG и JPEG2000 компресијом, а неке од база садрже и нетипичне модификације изворних слика (нпр. LAR (Locally Adaptive Resolution) компресија у оквиру IVC базе). Посматрачи су импресију о квалитету углавном давали на основу приказа једне слике (Single Stimulus) на монитору (LIVE, Toyama, IVC), док су код CSIQ базе истовремено посматрали тест слике на четири монитора.

Изворне слике и њихове модификације приказане на Сл. 3.6, Сл. 3.7 и Сл. 3.8 су из LIVE базе слика. Toyama база слика садржи два типа деградације изворних слика (JPEG и JPEG2000 компресија) и примери слика из ове базе биће коришћени у илустрацијама наредног дела рада. На Сл. 3.9 приказане су слике из IVC базе – изворна слика и њене четири модификације.

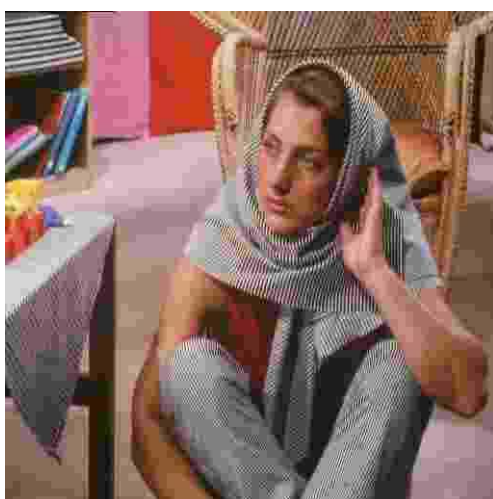
Са Сл. 3.9 се уочавају ефекти типични за поједине типове деградације (блоковски ефекат, одјек ивица, замрљање,...) и који се међусобно веома разликују. Такође, изузев слике са замрљањем код које се ефекат губитка високофреквенцијских компоненти уочава на комплетној слици, на преостале три тест слике се види да деградација није униформно расподељена.

На Сл. 3.10 приказане су неке од тест слика CSIQ базе. Кроз тест слике илустровано је свих шест деградација присутних у бази. Ефекти присутни на тест сликама се веома разликују. Присуство деградације је на неким сликама видљиво на свим позицијама (замрљање и модификација контраста), док је на делу слика присуство дисторзије маскирано у регионима са израженом текстуром (присуство адитивног шума и блоковских ефеката).

3. Процена квалитета слика заснована на очувању информација о градијенту



(a)



(б)



(в)



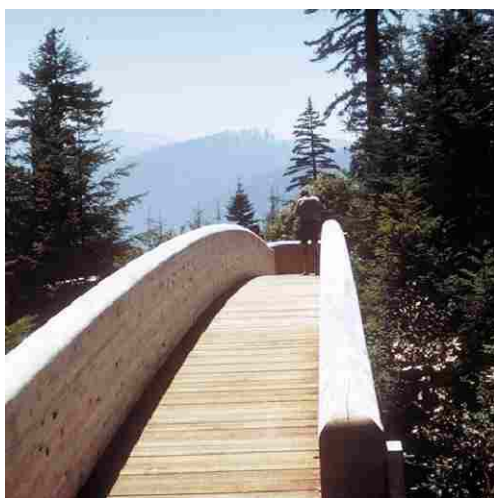
(г)



(д)

Сл. 3.9 (а) изворна слика из IVC базе и њене модификације настале: (б) JPEG компресијом, (в) JPEG2000 компресијом, (г) замрљањем и (д) LAR кодовањем.

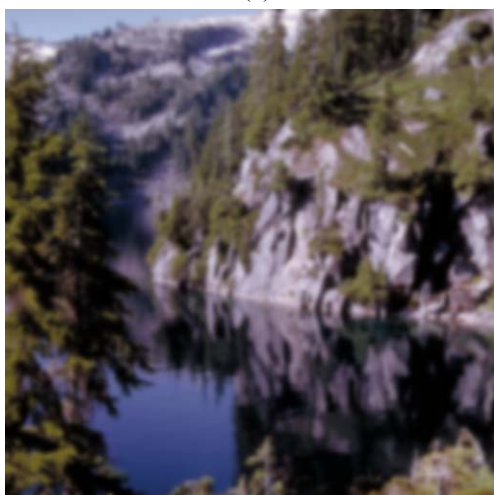
3. Процена квалитета слика заснована на очувању информација о градијенту



(a)



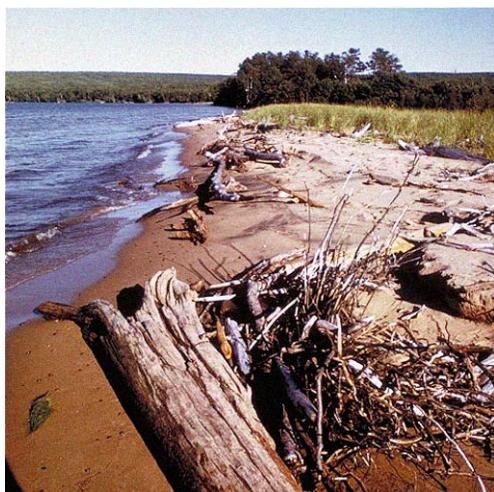
(б)



(в)



(г)



(д)

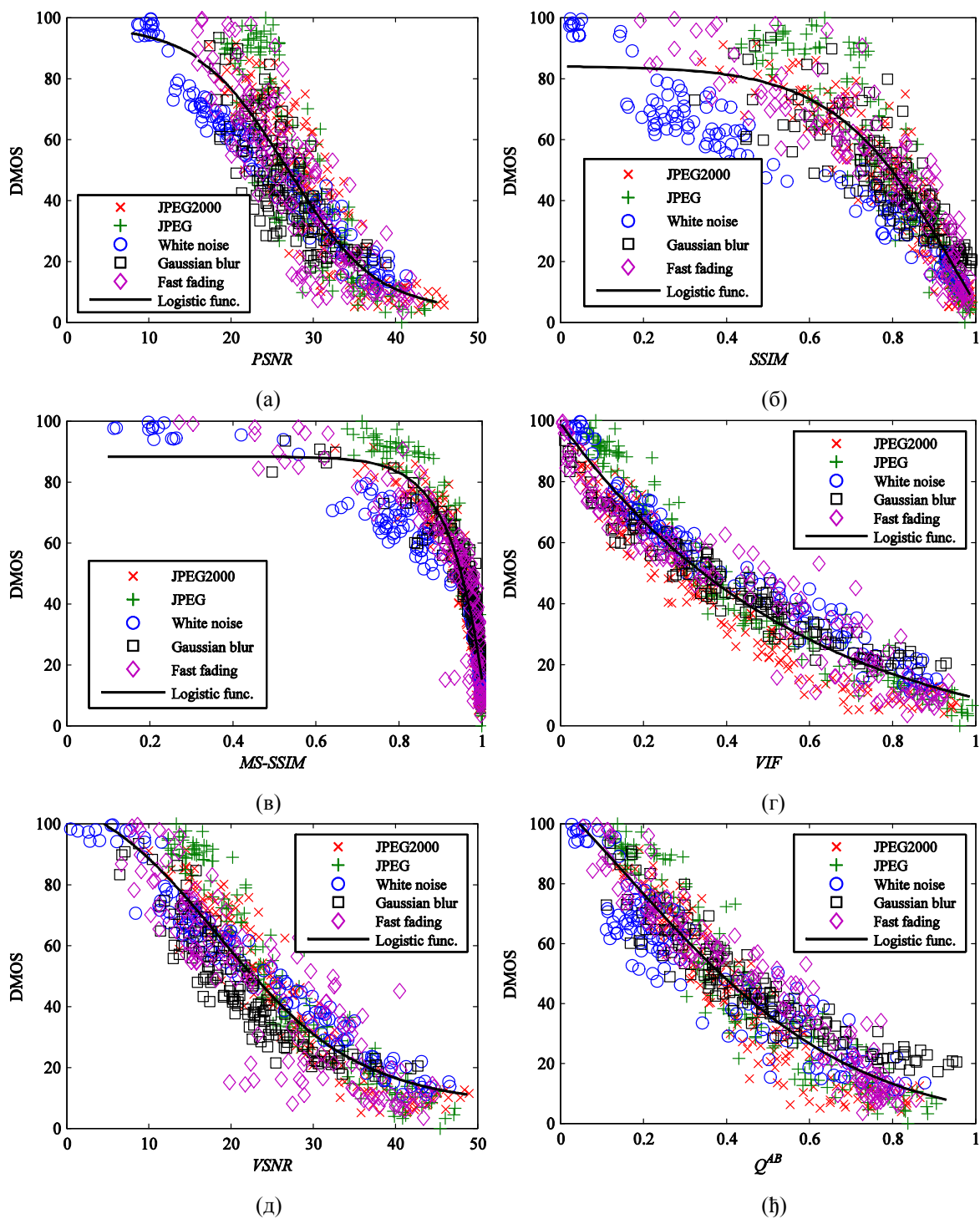


(ђ)

Сл. 3.10 Тест слике CSIQ базе са различитим деградацијама: (а) JPEG компресија, (б) JPEG2000 компресија, (в) замрљање, (г) модификација контраста, (д) адитивни бели шум и (ђ) адитивни ружичасти (pink) шум.

3. Процена kvaliteta слика zasnovana na очувању информација о градијенту

Перформансе објективних мера процене су на референтним базама евалуиране кроз тачност предикције (одређену кроз MAE, RMSE и LCC), монотоност предикције (одређену кроз SROCC) и конзистентност предикције (одређену кроз OR) [5].



Сл. 3.11 Дијаграми расипања субјективних (DMOS) и објективних вредности квалитета на LIVE бази слика.

3. Процена квалитета слика заснована на очувању информација о градијенту

На Сл. 3.11 приказани су дијаграми расипања (scatter plots) субјективних (DMOS) и објективних вредности квалитета на LIVE бази слика, где свака тачка представља једну тест слику. Вертикална и хоризонтална оса представљају DMOS и добијене објективне процене, тим редом.

Са Сл. 3.11(ђ) уочава се скоро линеарна веза између субјективних и објективних вредности квалитета, са константним расипањем Q^{AB} скорова око интерполационе криве у комплетном опсегу квалитета. Расипање око интерполационе криве је најмање код VIF мере, где се такође може говорити о линеарности објективних и субјективних импресија квалитета. Уколико се у процени користи SSIM мера, уочава се значајно расипање скорова код тест слика лошијег квалитета (ниже вредности SSIM). Овакав тренд се пресликао и у мултивеличински (MS-SSIM) приступ, где се додатно уочава гомилање објективних скорова квалитета у горњој половини расположивог динамичког опсега.

У Таб. 3.5–Таб. 3.10 дате су перформансе мера за процену квалитета на LIVE бази слика. Дате су перформансе објективних мера на комплетној бази (Таб. 3.5), као и перформансе објективних мера на подскуповима слика у оквиру базе (Таб. 3.6 – Таб. 3.10). Анализирано је седам објективних мера процене квалитета, при чему су перформансе три мере са најбољим резултатима обележене масним (подебљаним) словима.

Таб. 3.5 Поређење перформанси објективних мера процене квалитета на комплетној LIVE бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>PSNR</i>	0.8701	0.8756	10.5345	13.4671	10.2696
<i>UIQI</i>	0.8984	0.8941	9.4158	11.9991	6.2901
<i>SSIM</i>	0.9014	0.9104	9.3341	11.8323	7.3171
<i>MS-SSIM</i>	0.9382	0.9527	7.5734	9.4525	2.0539
<i>VIF</i>	0.9597	0.9636	6.2072	7.6753	0.6418
<i>VSNR</i>	0.9229	0.9273	8.0751	10.5213	4.1078
Q^{AB}	0.9378	0.9385	7.6159	9.4877	3.2092

3. Процена kvaliteta slika zasnovana na ochuvanju informacija o gradijentu

Из Таб. 3.5 се види да је по свим критеријумима предложена мера на глобалном нивоу боља од PSNR, UIQI, SSIM и VSNR. Перформансе предложене мере су у нивоу перформанси мултивеличинског индекса структурне сличности (MS-SSIM), са степеном слагања субјективних и објективних скорова од 94% (LCC=0.9378). На комплетној бази слика најбоље слагање субјективних и објективних скорова квалитета добија се коришћењем VIF мере.

Таб. 3.6 Коefицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
PSNR	0.8996	0.8879	0.9857	0.7835	0.8895	0.8701
UIQI	0.8684	0.9083	0.9367	0.9475	0.9487	0.8984
SSIM	0.9410	0.9504	0.9695	0.8743	0.9428	0.9014
MS-SSIM	0.9680	0.9804	0.9707	0.9486	0.9255	0.9382
VIF	0.9772	0.9864	0.9839	0.9743	0.9617	0.9597
VSNR	0.9629	0.9720	0.9775	0.9330	0.9018	0.9229
Q^{AB}	0.9560	0.9653	0.9466	0.9429	0.9655	0.9378

Таб. 3.7 Корелација рангова (SROCC) субјективних и објективних скорова на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
PSNR	0.8954	0.8809	0.9854	0.7823	0.8907	0.8756
UIQI	0.8620	0.9064	0.9100	0.9398	0.9362	0.8941
SSIM	0.9355	0.9449	0.9629	0.8944	0.9413	0.9104
MS-SSIM	0.9638	0.9814	0.9781	0.9542	0.9354	0.9527
VIF	0.9696	0.9846	0.9858	0.9728	0.9650	0.9636
VSNR	0.9560	0.9645	0.9787	0.9413	0.9044	0.9273
Q^{AB}	0.9503	0.9650	0.9298	0.9548	0.9639	0.9385

Таб. 3.8 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
PSNR	8.4564	11.3573	3.8081	9.0990	10.0425	10.5345
UIQI	9.8944	10.5392	7.9561	4.7049	6.7064	9.4158
SSIM	6.7460	7.6159	5.6985	6.7595	7.4447	9.3341
MS-SSIM	5.1180	4.8937	5.5801	4.7312	8.4649	7.5734
VIF	4.2199	4.0165	4.0395	3.3137	6.1517	6.2072
VSNR	5.2601	5.9408	4.6511	5.1742	9.6289	8.0751
Q^{AB}	5.8420	6.5551	7.4039	4.7537	5.8400	7.6159

3. Процена квалитета слика заснована на очувању информација о градијенту

Таб. 3.9 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
PSNR	11.0177	14.6526	4.7139	11.4778	13.0141	13.4671
UIQI	12.5105	13.3276	9.7961	5.9068	9.0058	11.9991
SSIM	8.5349	9.9070	6.8544	8.9643	9.4986	11.8323
MS-SSIM	6.3335	6.2766	6.7187	5.8436	10.7903	9.4525
VIF	5.3583	5.2329	5.0002	4.1576	7.8095	7.6753
VSNR	6.8068	7.4844	5.8950	6.6463	12.3112	10.5213
Q^{AB}	7.4008	8.3144	9.0184	6.1513	7.4192	9.4877

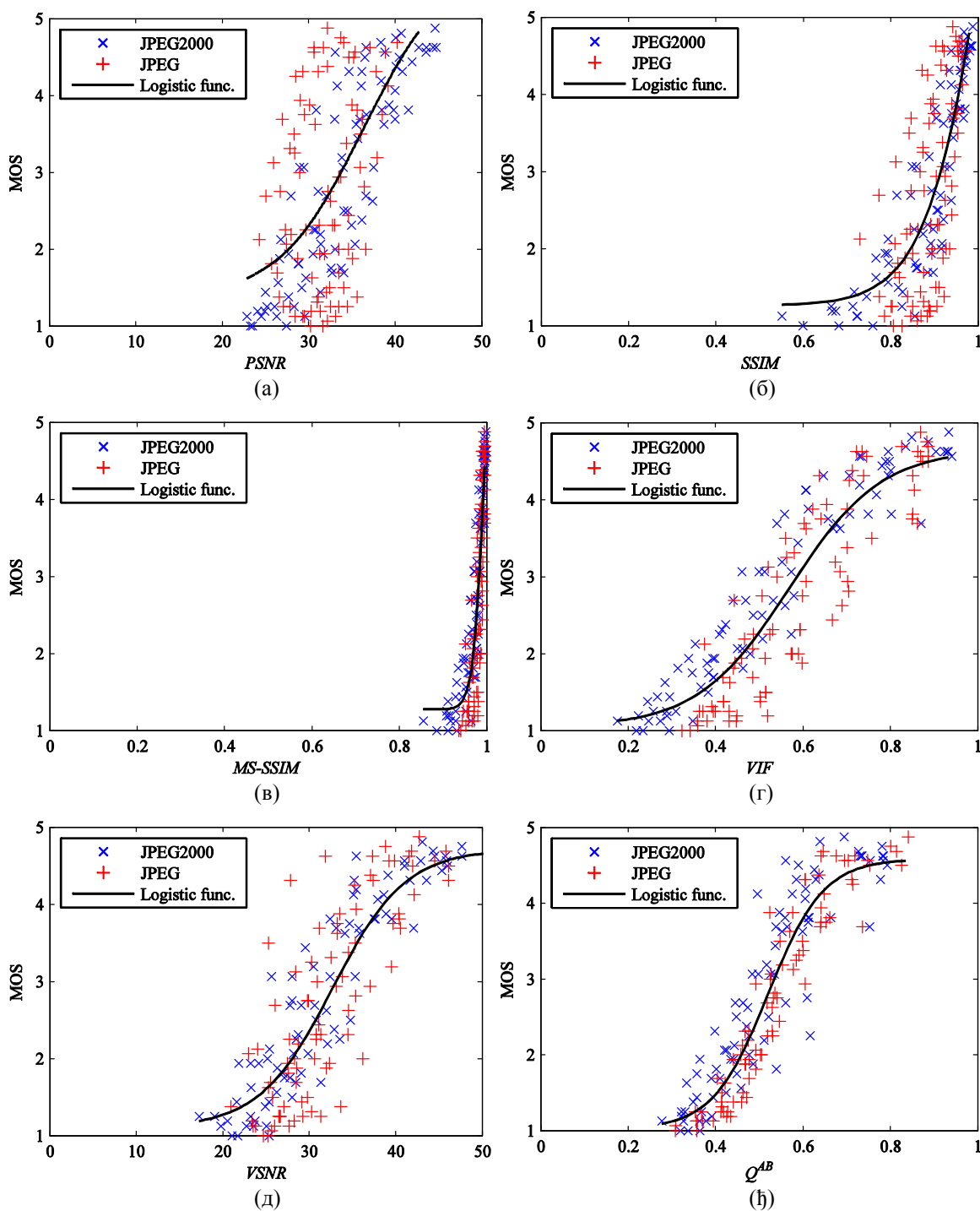
Таб. 3.10 Конзистентност објективних скорова (OR) након нелинеарне регресије на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
PSNR	5.3254	12.5714	0	11.7241	7.5862	10.2696
UIQI	4.7337	6.2857	0.6897	0	1.3793	6.2901
SSIM	1.1834	2.2857	0	4.1379	2.0690	7.3171
MS-SSIM	0	0	0	0	2.7586	2.0539
VIF	0	0	0	0	0.6897	0.6418
VSNR	0	0.5714	1.3793	0.6897	3.4483	4.1078
Q^{AB}	0	1.7143	2.7586	0.6897	0	3.2092

Из Таб. 3.6 – Таб. 3.10 уочава се доследност процена предложене мере по различитим класама слика. Коефицијент линеарне корелације субјективних и објективних Q^{AB} скорова је за све класе слика већи од 94%, при чему је за две класе слике коефицијент корелације већи од 96%. Степен слагања субјективних и објективних Q^{AB} скорова квалитета на нивоу комплетне базе (LCC=0.9378) је нижи од степена слагања на нивоу класа. Овакав резултат је последица различитих трендова објективних скорова квалитета на нивоу класа а који су последица различитих ефеката присутних у појединачним класама (адитивни шум, замрљање, блоковски ефекат, ...). Такође, може се закључити и да је по подскуповима слика из LIVE базе, VIF мера постигла најбоље резултате. Додатно се може уочити и да је на White noise класи слика изразито добре резултате постигла и PSNR мера (LCC>0.98).

На Сл. 3.12 приказани су дијаграми расипања субјективних (MOS) и објективних вредности квалитета на Тоуама бази слика, где свака тачка представља једну тест слику. Вертикална и хоризонтална оса представљају MOS и добијене објективне процене, тим редом. Ова база слика обухвата два типа деградације (компресије) изворних слика – JPEG2000 и JPEG компресије.

3. Процена квалитета слика заснована на очувању информација о градијенту



Сл. 3.12 Дијаграми расипања субјективних (MOS) и објективних вредности квалитета на Тоуата бази слика.

Са Сл. 3.12 се уочава да је најмање расипање објективних скорова квалитета око интерполационе криве уколико се користи предложена мера Q^{AB} . Расипање скорова је константно у комплетном динамичком опсегу квалитета. Као и код LIVE

3. Процена квалитета слика заснована на очувању информација о градијенту

базе слика, уочава се значајно расипање SSIM скорова код тест слика лошијег квалитета, што се пресликало и у мултивеличинску варијанту. MS-SSIM скорови се гомилају у близини максималне вредности квалитета.

У Таб. 3.11 – Таб. 3.13 дате су перформансе објективних мера за процену квалитета по анализираним аспектима на Тоуата бази слика. Дате су перформансе објективних мера на комплетној бази (Таб. 3.11), као и перформансе објективних мера на подскуповима слика у оквиру базе (Таб. 3.12 и Таб. 3.13).

Таб. 3.11 Поређење перформанси објективних мера процене квалитета на комплетној Тоуата бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
Мера	LCC	SROCC	MAE	RMSE	OR (%)
<i>PSNR</i>	0.6356	0.6132	0.7833	0.9662	22.0238
<i>UIQI</i>	0.7158	0.7020	0.6964	0.8739	14.8810
<i>SSIM</i>	0.7979	0.7870	0.5891	0.7545	14.2857
<i>MS-SSIM</i>	0.8948	0.8910	0.4305	0.5588	7.1429
<i>VIF</i>	0.9137	0.9077	0.4032	0.5087	5.3571
<i>VSNR</i>	0.8705	0.8608	0.4651	0.6160	9.5238
<i>Q^{AB}</i>	0.9322	0.9276	0.3482	0.4529	2.9762

Таб. 3.12 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG подскуп.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>PSNR</i>	0.3780	0.2868	0.9835	1.1449	39.2857
<i>UIQI</i>	0.8120	0.7937	0.5556	0.7217	8.3333
<i>SSIM</i>	0.6544	0.6263	0.7810	0.9351	25
<i>MS-SSIM</i>	0.8076	0.8659	1.0657	1.1998	41.6667
<i>VIF</i>	0.9000	0.9061	0.4430	0.5390	3.5714
<i>VSNR</i>	0.8082	0.7966	0.5655	0.7283	11.9048
<i>Q^{AB}</i>	0.9603	0.9543	0.2698	0.3450	2.3810

Таб. 3.13 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG2000 подскуп.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>PSNR</i>	0.8575	0.8605	0.4938	0.6497	7.1429
<i>UIQI</i>	0.6440	0.6330	0.7801	0.9660	17.8571
<i>SSIM</i>	0.9184	0.9148	0.3859	0.4997	4.7619
<i>MS-SSIM</i>	0.9416	0.9392	0.3237	0.4253	4.7619
<i>VIF</i>	0.9622	0.9559	0.2724	0.3440	3.5714
<i>VSNR</i>	0.9302	0.9246	0.3531	0.4635	3.5714
<i>Q^{AB}</i>	0.9234	0.9191	0.3515	0.4847	3.5714

3. Процена квалитета слика заснована на очувању информација о градијенту

Због значајног груписања објективних PSNR и MS-SSIM скорова квалитета на JPEG подскупу слика, њихове интерполационе функције су на датом подскупу добијене коришћењем логистик криве са пет параметара (са додатим линеарним чланом (19)).

Из Таб. 3.11 се уочава да је на комплетној бази слика, по свим анализираним аспектима, најбоље слагање субјективних и објективних Q^{AB} скорова квалитета. Исти закључак важи за JPEG подскуп (Таб. 3.12), док је за објективну евалуацију квалитета слика са JPEG2000 компресијом најбоље користити VIF меру (Таб. 3.13). Мањи степен слагања субјективних и Q^{AB} објективних скорова квалитета ($LCC_{JPEG}=0.96 > LCC_{JPEG2000}=0.92$) је највероватније последица десетак "проблематичних" JPEG2000 слика чије објективне вредности одступају од тренда осталих скорова (Сл. 3.12(ђ)).

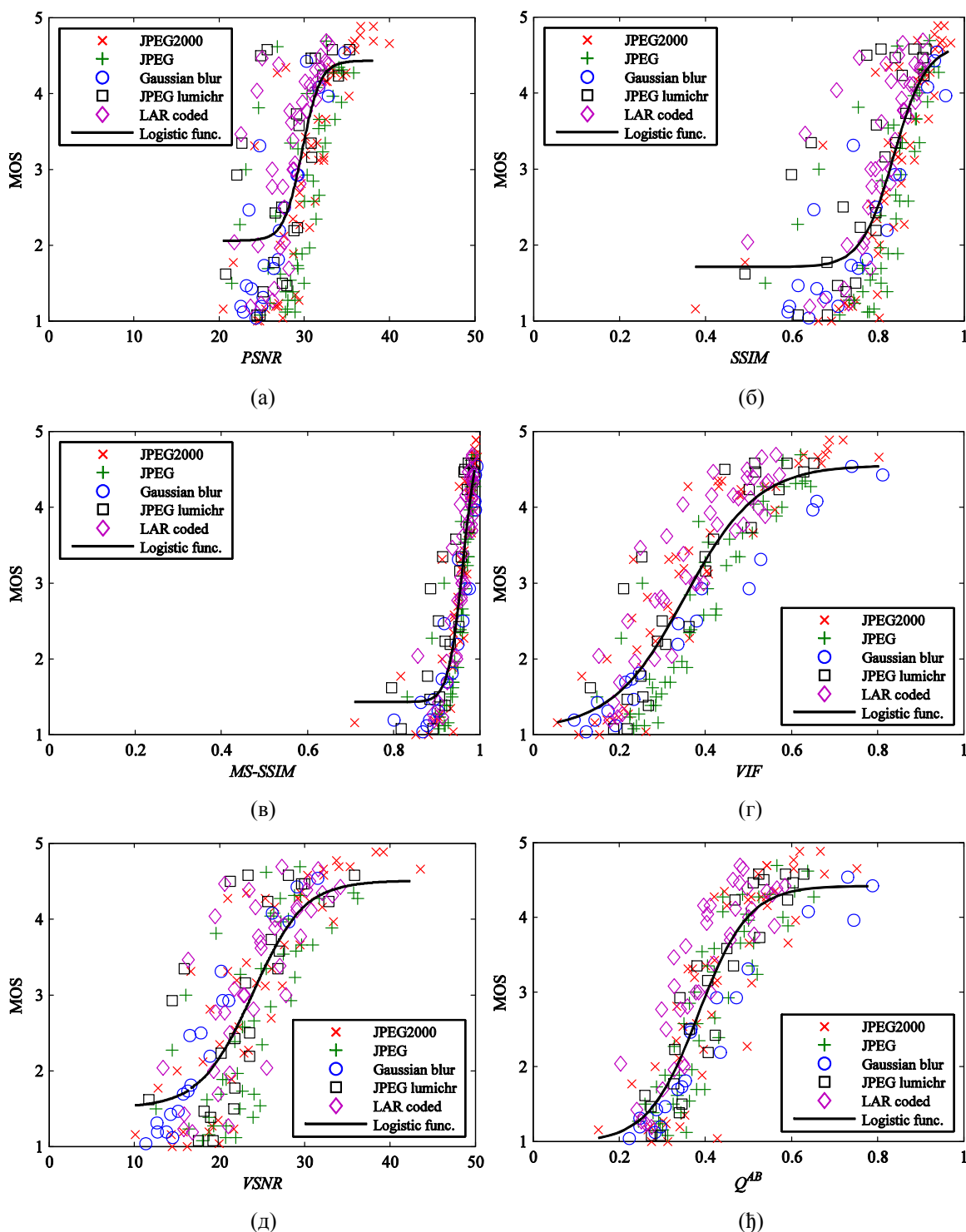
На Сл. 3.13 приказани су дијаграми расипања субјективних (MOS) и објективних вредности квалитета на IVC бази слика, где свака тачка представља једну тест слику. Вертикална и хоризонтална оса представљају MOS и добијене објективне процене, тим редом.

У Таб. 3.14 – Таб. 3.19 дате су перформансе објективних мера за процену квалитета по анализираним аспектима на IVC бази слика. Дате су перформансе објективних мера на комплетној бази (Таб. 3.14), као и перформансе објективних мера на подскуповима слика у оквиру IVC базе (Таб. 3.15 – Таб. 3.19).

Таб. 3.14 Поређење перформанси објективних мера процене квалитета на комплетној IVC бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>PSNR</i>	0.7192	0.6884	0.6688	0.8465	20.5405
<i>UIQI</i>	0.8302	0.8244	0.5261	0.6792	12.4324
<i>SSIM</i>	0.7924	0.7789	0.5547	0.7431	17.8378
<i>MS-SSIM</i>	0.9108	0.8971	0.3818	0.5031	5.9459
<i>VIF</i>	0.9026	0.8964	0.4096	0.5244	7.5676
<i>VSNR</i>	0.8027	0.7993	0.5585	0.7265	15.6757
Q^{AB}	0.8813	0.8699	0.4444	0.5757	5.4054

3. Процена kvaliteta слика zasnovana na очувању информација о градијенту



Сл. 3.13 Дијаграми расипања субјективних (MOS) и објективних вредности квалитета на IVC бази слика.

Из Таб. 3.14 се по добрим перформансама издвајају VIF, MS-SSIM и Q^{AB} мере. Интересантно је уочити да се по свим аспектима анализе објективних скорова,

3. Процена квалитета слика заснована на очувању информација о градијенту

мера UIQI показала бољом од њене "побољшане" верзије – индекса структурне сличности SSIM.

Таб. 3.15 Коefицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>PSNR</i>	0.8469	0.7391	0.8960	0.6675	0.6556	0.7192
<i>UIQI</i>	0.8064	0.8310	0.9400	0.9544	0.8531	0.8302
<i>SSIM</i>	0.8628	0.8330	0.9103	0.7994	0.7639	0.7924
<i>MS-SSIM</i>	0.9239	0.9467	0.9580	0.8927	0.8647	0.9108
<i>VIF</i>	0.9347	0.9387	0.9887	0.9149	0.9055	0.9026
<i>VSNR</i>	0.8626	0.8392	0.9803	0.7533	0.7373	0.8027
<i>Q^{AB}</i>	0.8629	0.9090	0.9781	0.9332	0.8998	0.8813

Таб. 3.16 Корелација рангова (SROCC) субјективних и објективних скорова на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>PSNR</i>	0.8500	0.6740	0.8051	0.5624	0.6988	0.6884
<i>UIQI</i>	0.7855	0.8247	0.9428	0.9422	0.7926	0.8244
<i>SSIM</i>	0.8501	0.8067	0.8691	0.7469	0.7118	0.7789
<i>MS-SSIM</i>	0.9264	0.9233	0.9458	0.8598	0.8587	0.8971
<i>VIF</i>	0.9356	0.9240	0.9729	0.8775	0.8878	0.8964
<i>VSNR</i>	0.8715	0.7786	0.9661	0.6992	0.7206	0.7993
<i>Q^{AB}</i>	0.8606	0.8905	0.9669	0.9083	0.8460	0.8699

Таб. 3.17 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>PSNR</i>	0.4700	0.6163	0.3607	0.7320	0.5841	0.6688
<i>UIQI</i>	0.6130	0.5163	0.2786	0.3051	0.4459	0.5261
<i>SSIM</i>	0.4538	0.4745	0.3407	0.6022	0.4968	0.5547
<i>MS-SSIM</i>	0.3327	0.2501	0.2475	0.4438	0.4184	0.3818
<i>VIF</i>	0.3382	0.2803	0.1388	0.3441	0.3528	0.4096
<i>VSNR</i>	0.4965	0.5161	0.1755	0.6307	0.5226	0.5585
<i>Q^{AB}</i>	0.5009	0.3999	0.1895	0.3679	0.3963	0.4444

3. Процена квалитета слика заснована на очувању информација о градијенту

Таб. 3.18 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>PSNR</i>	0.6861	0.7836	0.5070	0.9183	0.8005	0.8465
<i>UIQI</i>	0.7630	0.6471	0.3895	0.3683	0.5531	0.6792
<i>SSIM</i>	0.6523	0.6436	0.4727	0.7410	0.6841	0.7431
<i>MS-SSIM</i>	0.4937	0.3747	0.3273	0.5558	0.5325	0.5031
<i>VIF</i>	0.4586	0.4010	0.1711	0.4978	0.4498	0.5244
<i>VSNR</i>	0.6528	0.6326	0.2256	0.8112	0.7161	0.7265
<i>Q^{AB}</i>	0.6520	0.4849	0.2378	0.4432	0.4626	0.5757

Таб. 3.19 Конзистентност објективних скорова (OR) (након нелинеарне регресије) на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>PSNR</i>	14	14	10	16	22.5000	20.5405
<i>UIQI</i>	14	12	5	0	10	12.4324
<i>SSIM</i>	14	10	10	20	12.5000	17.8378
<i>MS-SSIM</i>	12	2	0	4	5	5.9459
<i>VIF</i>	8	2	0	12	2.5000	7.5676
<i>VSNR</i>	14	10	0	20	12.5000	15.6757
<i>Q^{AB}</i>	12	2	0	0	0	5.4054

Анализа перформанси по класама слика IVC базе, указује да се код неких класа алтернативно, уместо VIF, MS-SSIM и Q^{AB} , могу користити VSNR (Gaussian blurring класа) и UIQI (JPEG lumichr класа) мере.

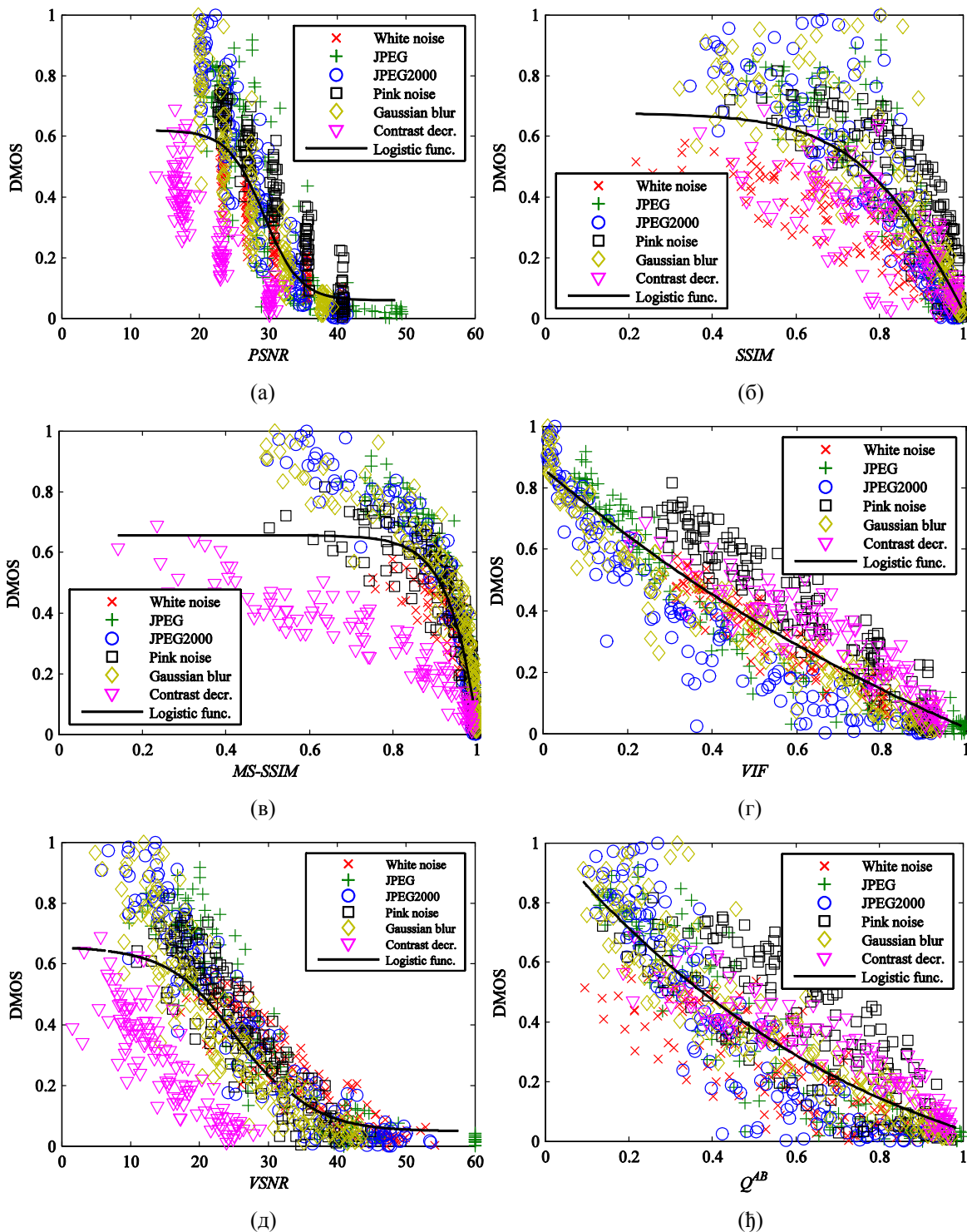
Због значајног груписања објективних PSNR и MS-SSIM скорова квалитета на подскупу слика са замрљањем и подскупу са LAR кодовањем, њихове интерполационе функције су на датим подскуповима добијене коришћењем логистик криве са пет параметара (са додатим линеарним чланом (19)).

Перформансе предложене мере значајно зависе од избора подскупа IVC базе слика. Тако се степен слагања субјективних и Q^{AB} објективних скорова квалитета (мерен кроз LCC) на подскуповима креће од 86% до 98%. Перформансе осталих објективних мера процене такође значајно варирају од подскупа до подскупа.

На Сл. 3.14 приказани су дијаграми расипања субјективних (DMOS) и објективних вредности квалитета на CSIQ бази слика, где свака тачка представља

3. Процена квалитета слика заснована на очувању информација о градијенту

једну тест слику. Вертикална и хоризонтална оса представљају DMOS и добијене објективне процене, тим редом.



Сл. 3.14 Дијаграми расипања субјективних (DMOS) и објективних вредности квалитета на CSIQ бази слика.

3. Процена квалитета слика заснована на очувању информација о градијенту

Са Сл. 3.14(ђ) уочава се скоро линеарна веза између субјективних и објективних Q^{AB} вредности квалитета. Добра линеарност Q^{AB} вредности квалитета на сликама са модификацијом контраста (Contrast decrement images) указује да су параметри нелинеарних сигмоидних функција (31) добро изабрани. У овом случају је изолован ефекат очувања амплитуде градијента и лоши параметри сигмоиде би дали нелинеарну зависност Q^{AB} и DMOS вредности.

Расипање око интерполационе криве је најмање код VIF мере, где се такође може говорити о линеарности објективних и субјективних импресија. Уколико се у процени користи SSIM мера, уочава се значајно расипање скорова око интерполационе функције. Овакав тренд се пресликао и у мултивеличински (MS-SSIM) приступ, где је расипање израженије код слика лошег квалитета. Код примене VSNR мере, уочава се одступање скорова добијених за слике са модификацијом контраста у односу на осталих пет класа слика. Дијаграм расипања субјективних и објективних PSNR скорова илуструје недостатке PSNR мере. PSNR вредности квалитета се групишу око пет карактеристичних вредности (три карактеристичне вредности за класу са модификацијом контраста), док истовремено постоје значајне разлике у субјективним импресијама (вертикално простирање скорова на дијаграму расипања Сл. 3.14(a)).

У Таб. 3.20–Таб. 3.25 дате су перформансе мера за процену квалитета по анализираним аспектима на CSIQ бази слика. Дате су перформансе објективних мера на комплетној бази (Таб. 3.20), као и перформансе објективних мера на подскуповима слика у оквиру CSIQ базе (Таб. 3.21 – Таб. 3.25).

Таб. 3.20 Поређење перформанси објективних мера процене квалитета на комплетној CSIQ бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>PSNR</i>	0.7999	0.8057	0.1195	0.1576	34.2956
<i>UIQI</i>	0.8289	0.8092	0.1127	0.1469	34.4111
<i>SSIM</i>	0.8151	0.8368	0.1161	0.1521	33.4873
<i>MS-SSIM</i>	0.8666	0.8774	0.0972	0.1310	27.7136
<i>VIF</i>	0.9252	0.9194	0.0753	0.0996	22.7483
<i>VSNR</i>	0.8018	0.8132	0.1152	0.1569	30.1386
<i>Q^{AB}</i>	0.8556	0.8520	0.1039	0.1359	31.1778

3. Процена квалитета слика заснована на очувању информација о градијенту

Из Таб. 3.20 се по добрим перформансама издвајају VIF, MS-SSIM и Q^{AB} мере. Перформансе предложене мере су у нивоу перформанси мултивеличинског индекса структурне сличности, са степеном слагања субјективних и објективних скорова од 86% (LCC=0.8556). На комплетној бази слика најбоље слагање субјективних и објективних скорова квалитета добија се коришћењем VIF мере.

Таб. 3.21 Коefицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.

<i>Mepa</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>PSNR</i>	0.9532	0.8906	0.9468	0.9532	0.9252	0.8986	0.7999
<i>UIQI</i>	0.7634	0.9179	0.9044	0.7190	0.9485	0.8479	0.8289
<i>SSIM</i>	0.9256	0.9400	0.9226	0.8935	0.9005	0.7435	0.8151
<i>MS-SSIM</i>	0.9430	0.9834	0.9788	0.9387	0.9620	0.9453	0.8666
<i>VIF</i>	0.9606	0.9883	0.9776	0.9568	0.9736	0.9416	0.9252
<i>VSNR</i>	0.9261	0.9555	0.9603	0.9198	0.9338	0.8732	0.8018
Q^{AB}	0.8433	0.9527	0.9396	0.8222	0.9297	0.9422	0.8556

Таб. 3.22 Корелација рангова (SROCC) субјективних и објективних скорова на CSIQ бази слика.

<i>Mepa</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>PSNR</i>	0.9363	0.8879	0.9361	0.9338	0.9291	0.8623	0.8057
<i>UIQI</i>	0.7599	0.9065	0.8816	0.7117	0.9430	0.8313	0.8092
<i>SSIM</i>	0.9256	0.9219	0.9205	0.8925	0.9245	0.7399	0.8368
<i>MS-SSIM</i>	0.9437	0.9643	0.9724	0.9327	0.9747	0.9448	0.8774
<i>VIF</i>	0.9575	0.9703	0.9671	0.9509	0.9744	0.9345	0.9194
<i>VSNR</i>	0.9254	0.9172	0.9485	0.9082	0.9445	0.8714	0.8132
Q^{AB}	0.8399	0.9378	0.9255	0.8170	0.9438	0.9437	0.8520

Таб. 3.23 Средња апсолутна грешка (MAE) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.

<i>Mepa</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>PSNR</i>	0.0387	0.0964	0.0731	0.0551	0.0777	0.0551	0.1195
<i>UIQI</i>	0.0866	0.0880	0.0988	0.1266	0.0667	0.0712	0.1127
<i>SSIM</i>	0.0493	0.0749	0.0885	0.0843	0.0850	0.0855	0.1161
<i>MS-SSIM</i>	0.0426	0.0418	0.0495	0.0634	0.0635	0.0440	0.0972
<i>VIF</i>	0.0358	0.0354	0.0479	0.0546	0.0493	0.0447	0.0753
<i>VSNR</i>	0.0495	0.0659	0.0650	0.0687	0.0738	0.0600	0.1152
Q^{AB}	0.0711	0.0649	0.0787	0.1041	0.0747	0.0439	0.1039

3. Процена квалитета слика заснована на очувању информација о градијенту

Таб. 3.24 Корен средње квадратне грешке (RMSE) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.

<i>Мера</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>PSNR</i>	0.0507	0.1391	0.1017	0.0684	0.1087	0.0739	0.1576
<i>UIQI</i>	0.1084	0.1214	0.1349	0.1572	0.0907	0.0893	0.1469
<i>SSIM</i>	0.0635	0.1044	0.1219	0.1015	0.1246	0.1126	0.1521
<i>MS-SSIM</i>	0.0558	0.0556	0.0646	0.0779	0.0782	0.0549	0.1310
<i>VIF</i>	0.0466	0.0467	0.0664	0.0658	0.0653	0.0567	0.0996
<i>VSNR</i>	0.0633	0.0902	0.0882	0.0887	0.1025	0.0821	0.1569
<i>Q^{AB}</i>	0.0902	0.0930	0.1081	0.1287	0.1055	0.0564	0.1359

Таб. 3.25 Конзистентност објективних скорова (OR) након нелинеарне регресије на CSIQ бази слика.

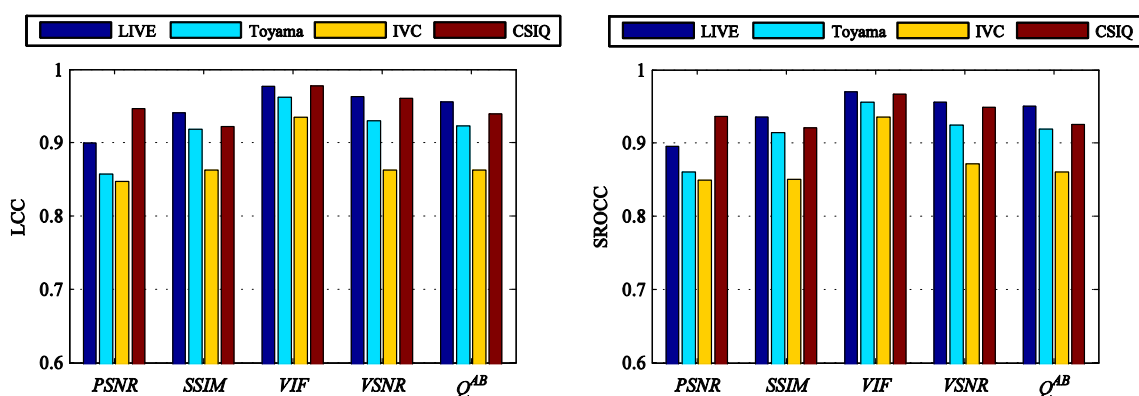
<i>Мера</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>PSNR</i>	6	32	38.6667	10.6667	22	8.6207	34.2956
<i>UIQI</i>	18.6667	30.6667	44	34.6667	19.3333	14.6552	34.4111
<i>SSIM</i>	6	26	43.3333	16.6667	18.6667	17.2414	33.4873
<i>MS-SSIM</i>	6	15.3333	30	14	16	6.8966	27.7136
<i>VIF</i>	6	14	28.6667	10	14	6.0345	22.7483
<i>VSNR</i>	8.6667	26	31.3333	12.6667	21.3333	9.4828	30.1386
<i>Q^{AB}</i>	13.3333	25.3333	40.6667	26.6667	18	6.8966	31.1778

Перформансе предложене мере Q^{AB} су најлошије на класама слика са додатим шумом (White noise и Pink noise класе). Степен слагања субјективних и Q^{AB} објективних вредности на овим подскуповима, мерен кроз линеарну корелацију и корелацију рангова, је бољи од 82%. Овај степен слагања је значајно лошији од слагања субјективних и Q^{AB} објективних скорова квалитета на зашумљеним сликама LIVE базе слика, где је слагање око 93%. Разлике у степену слагања су највероватније последица припреме тест слика (о чему ће бити речи у наредном – 3.4.3 делу рада), тј. лошије перцептуалне сепарације тест материјала CSIQ базе.

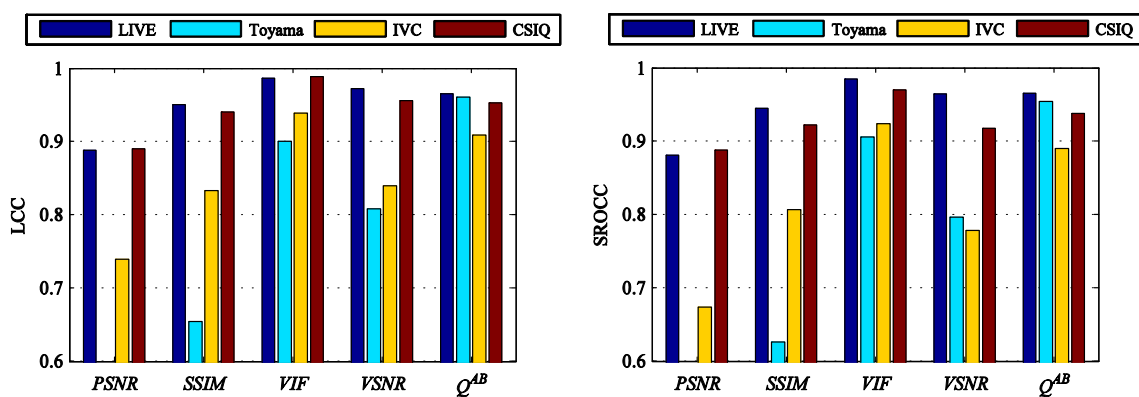
Истовремено, као и код LIVE базе, степен слагања субјективних и објективних PSNR скорова на зашумљеним сликама је висок ($LCC > 0.95$). Висок степен слагања по свим класама слика је присутан уколико се као објективне мере користе VIF и MS-SSIM.

3.4.1. Утицај избора базе на перформансе објективних мера

Како све четири анализирани базе садрже слике са JPEG и JPEG2000 деградацијама, занимљиво је анализирати перформансе објективних мера на овим подскуповима слика. Перформансе објективних мера на подскуповима слика из различитих база приказане су на Сл. 3.15. Као параметри поређења користе се коефицијент линеарне корелације (LCC) и корелација рангова (SROCC).



(a) JPEG2000 класа слика



(б) JPEG класа слика

Сл. 3.15 Коефицијент линеарне корелације (LCC) и корелација рангова (SROCC) на подскуповима слика различитих база: (а) JPEG2000 подскуп слика и (б) JPEG подскуп слика.

Са Сл. 3.15 се уочава сличан тренд перформанси мера од базе до базе. На JPEG класи слика, у свим базама и по оба параметра, најлошији резултати се добијају применом PSNR мере. Сличан закључак може се извести и за JPEG2000 класу, где су перформансе PSNR најлошије на три базе (изузетак је CSIQ). JPEG2000 класа слика је на свим базама најбоље евалуирана применом VIF мере, док је иста мера на JPEG класи најбоље резултате дала на три од четири анализирани базе.

3. Процена квалитета слика заснована на очувању информација о градијенту

Такође, са Сл. 3.15 се уочава да се на истом типу деградације изворних слика, од базе до базе перформансе мера значајно разликују, што је најочљивије код PSNR, SSIM и VSNR мера на JPEG класи слика. Разлике у перформансама мера објективне процене од базе до базе могу да буду последица динамичког опсега квалитета слика са деградацијом.



(a)



(б)



(в)



(г)



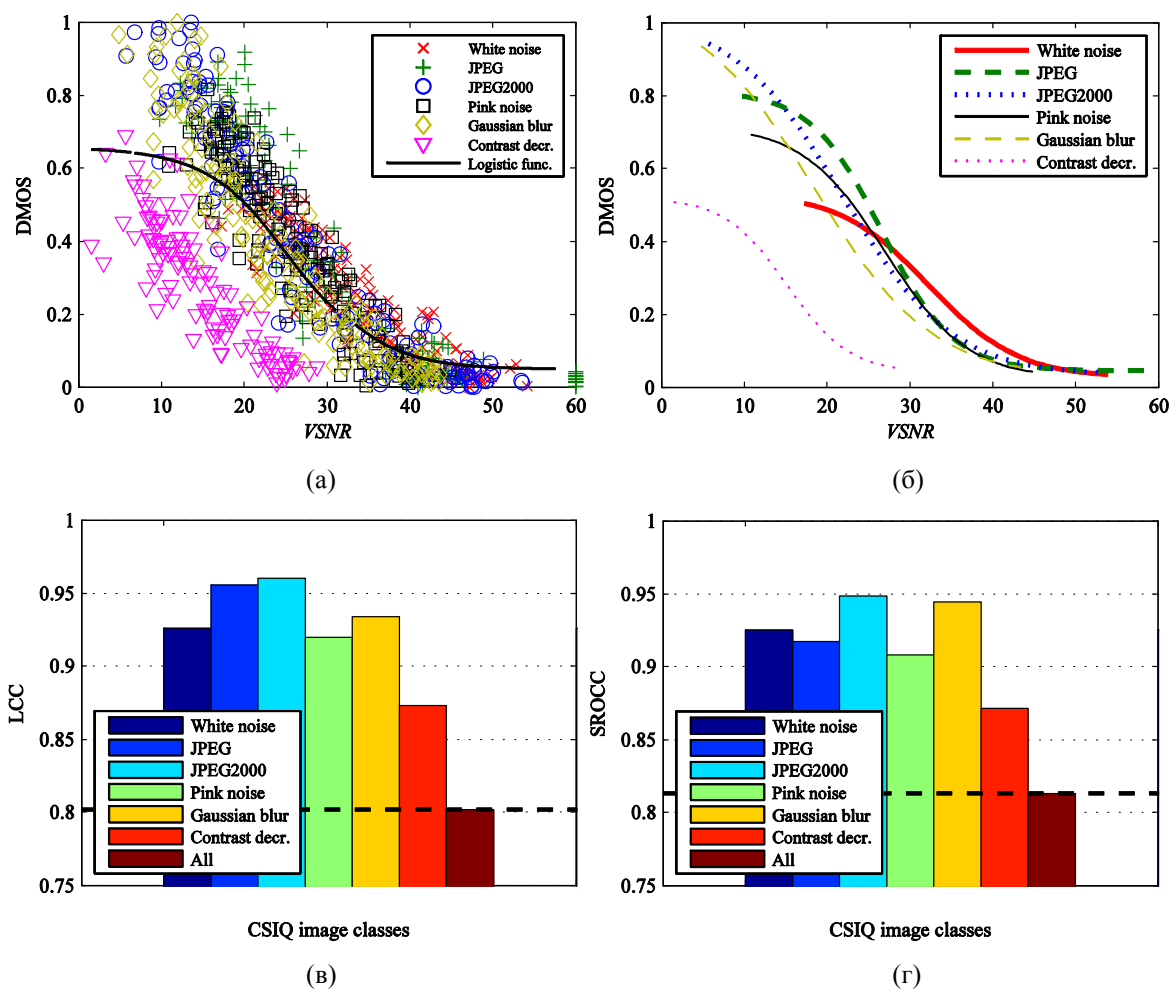
(д)

Сл. 3.16 Изворна слика и њене модификације у различитим базама слика: (а) изворна слика, (б) JPEG деградирана слика из Тоуата базе (MOS=1.69), (в) JPEG деградирана слика из LIVE базе (DMOS=84.89), (г) JPEG2000 деградирана слика из Тоуата базе (MOS=1.19) и (д) JPEG2000 деградирана слика из LIVE базе (DMOS=82.73).

На Сл. 3.16 приказана је изворна слика и њене модификације из две базе слика (Тоуата и LIVE базе). При томе су одабране модификације које у наведеним базама имају најлошији субјективни утисак. Уочава се да слике са дисторзијом у LIVE бази имају екстремно низак квалитет. Одговарајуће слике из Тоуата базе имају значајно бољи квалитет. Перформансе неких од анализираних објективних мера (PSNR, SSIM, VSNR) су значајно боље на LIVE и CSIQ базама него на базама са ужим динамичким опсегом квалитета, као што су Тоуата и IVC базе. Овакав резултат наводи на закључак да евалуација перформанси објективне мере на једној бази није довољна, јер перформансе значајно зависе од расположивог опсега квалитета. Уколико тест слике имају висок квалитет, субјективна импресија квалитета више зависи од карактеристика визуелног система посматрача него од статистике саме слике [228]. Стога је веома битно познавати разлике које постоје између база. Ради извођења валидних закључака додатно је потребно познавати и структуру (концепт) мера за процену квалитета. За очекивати је да ће објективне мере које користе само структурно поређење имати лоше перформансе на скупу слика са модификацијом контраста (без или уз минимално нарушавање структуре).

3.4.2. Однос перформанси на подскуповима и на глобалном плану

Из спроведене анализе по све четири базе се може уочити да успех мере на појединачним класама слика не води увек ка глобалној успешности мере. Оваква ситуација је уочена код VSNR мере. Ако се посматрају њене перформансе на подскуповима CSIQ базе слика (Сл. 3.17), уочава се веома висок степен корелације субјективних и објективних скорова. Изузев класе слика са модификацијом контраста ($LCC=0.8732$), на преосталих пет класа слика коефицијент корелације је већи од 92%. Међутим, на глобалном нивоу (нивоу комплетне базе) коефицијент корелације са субјективним скоровима је значајно нижи ($LCC=0.8018$). Овакав резултат је вероватно последица значајног одступања тренда објективних процена слика са модификацијом контраста од тренда процена остатка базе (Сл. 3.17(а) и Сл. 3.17(б)). Идеална објективна мера процене би требало да за све класе слика има приближно исти тренд процена, што би се могло остварити применом адаптивне мере која би пре давања коначне импресије квалитета одредила класу слике са деградацијом и евалуирала карактеристичне ефекте у оквиру класе.



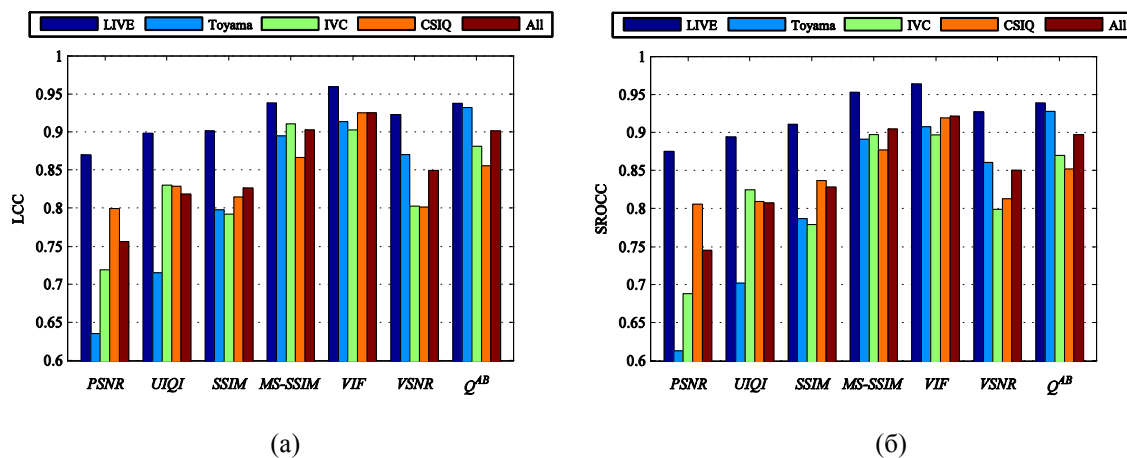
Сл. 3.17 Илустрација перформанси VSNR мере на CSIQ бази слика и на њеним подкуповима: (а) дијаграм расипања и интерполациона функција на комплетној бази, (б) интерполационе функције на подкуповима слика, (в) коефицијент линеарне корелације (LCC) на подкуповима и комплетној бази и (г) корелација рангова (SROCC) на подкуповима и комплетној бази.

3.4.3. Анализа перформанси објективних мера процене на глобалном плану

На Сл. 3.18 приказани су коефицијент линеарне корелације и корелација рангова субјективних и објективних скорова на свим анализираним базама слика (глобални показатељ перформанси) и дата је средња вредност датих параметара (по свим базама).

Са Сл. 3.18 се по оба параметра (LCC и SROCC) уочава да су на глобалном нивоу најлошије перформансе PSNR (по свим базама и укупно). Од три мере које у основи имају поређење структурних информација (UIQI, SSIM и MS-SSIM) најбољи

резултати се добијају применом мултивеличинског индекса MS-SSIM. Индекс структурне сличности SSIM, као наследник UIQI мере, се показао успешнијим од UIQI на три базе слика, међутим разлика у перформансама није значајна. Од анализираних мера најбоље слагање са субјективним скоровима има VIF мера. Ова мера је најуспешнија на две базе (LIVE и CSIQ), док су на по једној бази најбоље перформансе MS-SSIM (IVC база) и предложене мере Q^{AB} (Toyama база). Перформансе предложене мере су боље од перформанси VSNR, SSIM, UIQI и конвенционалне PSNR мере, док су у нивоу перформанси доста рачунарски комплексније MS-SSIM мере.



Сл. 3.18 Анализа перформанси мера објективне процене на глобалном нивоу (нивоу комплетних база слика): (а) коефицијент линеарне корелације (LCC) и (б) корелација рангова (SROCC).

Најбоље перформансе на глобалном плану су по свим анализираним објективним мерама остварене на LIVE бази слика. Такав резултат је највероватније последица квалитетне припреме тест слика. Наиме, пре спровођења субјективних тестова кроз које се за ову базу дошло до вредности DMOS скорова, аутори [223] су спровели широк спектар субјективних тестова кроз које су за све типове дисторзија и изворних слика одредили степене деградације који омогућавају перцептуално уочавање разлика између тест слика. Посматрачи су помоћу покретног клизача имали могућност да континуално мењају степен деградације изворне слике, а додатно су имали захтев да одреде тренутак када уочавају разлике између две суседне побуде (тест слике). Насупрот оваквој припреми тест слика налазе се тест материјали CSIQ базе. Аутори CSIQ базе [226] су у припреми тест слика кренули од

PSNR. То се може видети са Сл. 3.14(a), где се уочавају дискретне вредности PSNR. Задавањем фиксних вредности PSNR аутори су долазили до степена деградације изворне слике, односно до тест слике. Међутим, апсолутна промена PSNR од нпр. 5 dB (Сл. 3.14(a)), не мора да буде увек субјективно уочљива. Кроз овакву припрему тест материјала аутори су показали да истој вредности PSNR одговарају значајно различите субјективне импресије квалитета тест слика (Сл. 3.14(a)). Додатно су оваквом припремом тест материјала отежали посматрачима оцењивање, а што се види и кроз значајно расипање скорова око интерполационих кривих у субјективно-објективном простору (Сл. 3.14). Самим тим је и конзистентност (OR) објективних мера процене значајно лошија него код LIVE базе слика (Таб. 3.5 и Таб. 3.20). У доступној литератури није било могуће доћи до информација о начину припреме тест слика Тоуата [224] и IVC [225] база. Осим начина припреме тест слика, у делу рада 3.4.1 показано је да динамички опсег квалитета (степен деградације) такође значајно утиче на перформансе објективних мера (динамички опсег квалитета изворних слика Тоуата и IVC база слика је значајно ужи од динамичког опсега квалитета LIVE и CSIQ база).

У овом делу дисертације анализирана је објективна процена квалитета слика заснована на очувању локалних информација о градијенту. Свим пикселима је у финалној мери очувања изворних информација дат исти значај, тј. мапа субјективног значаја са Сл. 3.1 је униформна. У наредном делу дисертације биће анализирани посебни аспекти процене квалитета слике, у оквиру кога ће један део бити посвећен формирању мапа субјективног значаја и њиховом доприносу у финалној процени квалитета.

4. Посебни аспекти процене квалитета слика

Процена квалитета слике је сложен процес трансфера информација и њихове репрезентације. Због тога се у литератури могу наћи бројни покушаји да се обухватањем различитих аспеката процене квалитета слика и видеа додатно побољшају перформансе објективних мера процене. Аспекти који су разматрани обухватају: мултивеличинску анализу [126, 215], примену боје у процени квалитета [78-85], увођење мапа субјективног значаја [124-148], адитивни модел сједињавања информација у евалуацији квалитета [229], избор колор-координатног система [230] и сл.

У оквиру дисертације разматрана су три аспекта процене квалитета. Први аспект третира пренос слике кроз модел трансфера информација по узору на комуникациони канал. На основу мере очувања информација о градијенту предложен је метод за свестранију анализу процеса преноса слике кроз анализу пренетих, изгубљених и убачених информација.

Други аспект у процену квалитета уводи мапе значаја којима се пондеришу локалне вредности очувања информација о градијенту. Показано је да се увођењем мапа значаја могу побољшати перформансе мере очувања информација о градијенту.

Адитивни модели комбиновања очувања информација, праћени квалитетом вођеном интеграцијом су анализирани у трећем делу поглавља. Показано је да су адитивни модели очувања информација погоднији за процену квалитета од стандардних мултипликативних модела.

Сва три аспекта су илустрована кроз евалуацију бројних слика са различитим типовима и нивоима дисторзије.

4.1. Информациони аспект очувања информација о градијенту

Постојеће мере за објективну процену квалитета слика представљају квалитет тест слике графички (дводимензионална мапа квалитета) или нумерички (скаларна вредност) [50]. Као резултат, представа квалитета (било на један или други начин) даје делимичну слику о трансферу информација и њиховој репрезентацији.

Процена квалитета слике је сложен процес трансфера информација и њихове репрезентације и у неким апликацијама су поједини аспекти овог процеса критични и захтевају додатна разматрања. На пример, губитак битног обележја током процеса преноса медицинских слика може довести до давања погрешне дијагнозе и много озбиљнијих последица. У војним применама артефакти који се јављају у компресији слике могу довести до тога да се бенигни објекат класификује као претња односно валидни циљ. Са циљем да се размотре апликациони аспекти квалитета слике на робустан начин, потребна је много свестранија и исцрпнија евалуација квалитета слике [231, 232].

Како у практичним применама референтна слика на пријемној страни није доступна, за евалуацију квалитета потребне су NR мере. Већина техника за процену квалитета без референцирања на тест слику су технике које су прилагођене специфичним типовима дисторзије. Примери оваквих (NR) мера могу се наћи у [157-159], где се процена квалитета слике спроводи кроз евалуацију блоковског ефекта, и [160, 161] где се за процену квалитета мере замрљање слике и ефекат одјека – прстена. У скорије време се појављују генералне мере за процену квалитета без референцирања, као што су BIQI индекс (Blind Image Quality Assessment Index) [164] и BLIINDS индекс (Blind Image Integrity Notator using DCT Statistics) [165]. Ови приступи не подразумевају специфичан тип дисторзије слике и засновани су на статистици слика природних сцена (NSS) и статистици коефицијената DCT.

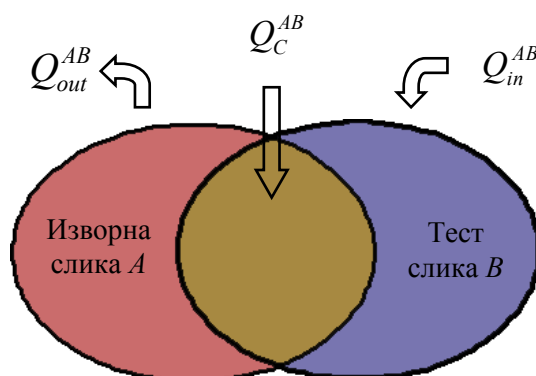
У овом делу дисертације објективна процена квалитета слике се посматра кроз исцрпну анализу засновану на квантификацији битних информационих компоненти укључених у трансфер слике: заједничке (пренете) информације, изгубљене информације (еквивокација) и унете информације (ирелеванција). Предложени приступ користи објективну процену квалитета слике засновану на очувању информација о градијенту, а приступ је илустрован на LIVE бази слика [223].

Објективна процена Q^{AB} дата релацијом (33) је у основи повезана са визуелно значајним информацијама – ивицама односно информацијама о градијенту [222]. Идеја од које се кренуло је да је пренос (линија везе) којим се очува више информација о градијенту у тест слици (слици на пријемној страни) бољи. У првом

кораку се издвајају параметри ивица у референтној и тест слици и то на свим позицијама. Након тога се коришћењем мера очувања информација одређује степен репрезентације параметара ивица изворне слике у тест слици. Како би се одредио квалитет комплетне тест слике, локалне вредности (вредности на нивоу пиксела) очувања информација се усредњавају и одређује крајња вредност квалитета слике са дисторзијом.

4.1.1. Пренете (заједничке) информације

Процес преноса информација се може посматрати као проблем у коме се скуп информација пресликава у нови скуп, а који може садржати све информације оригиналног скупа. Ово је графички илустровано коришћењем једноставног Веновог дијаграма на Сл. 4.1.



Сл. 4.1 Једноставни Венов дијаграма процеса преноса слике.

У случају преноса слике, информације се могу повезати са просторним локацијама где свака тачка представља индивидуални информациони скуп. Пренета информација, приказана на Сл. 4.1 као пресек између изворне слике A и тест слике B , је управо оно што се евалуира применом мере Q^{AB} (у односу на величину A). У овом делу дисертације мера Q^{AB} ће се обележавати са Q_C^{AB} , ради асоцијације на заједничке информације (common information).

Мера очувања информација о градијенту, Q_C^{AB} , представља количину информација пренетих од слике A до слике B , и у њу су интегрисане локалне

процене заједничких информација свих пиксела, $Q_C^{AB}(n,m)$. Заједничке информације, Q_C^{AB} , су обележене на Сл. 4.1.

4.1.2. Изгубљене информације

Информациони губитак (information loss), Q_{out}^{AB} , је мера информација изгубљених током преноса слике. Он представља информације које су присутне у референтној слици али их нема у слици са деградацијом. Овај губитак је обележен на Сл. 4.1 са Q_{out}^{AB} . Губитак информација се поистовећује са позицијама на којима је Q_C^{AB} мање од један. Међутим, како би се правилно евалуирао губитак информација, он се мора раздвојити од артефаката који се јављају током преноса, а који ће такође резултирати са $Q_C^{AB} < 1$. У предложеном приступу се за њихово раздвајање користи амплитуда градијента, (27). Наиме, на свакој позицији се користи следећа класификација: уколико је амплитуда градијента тест слике B , g_B , већа од одговарајуће амплитуде изворне слике, g_A , слика B садржи артефакт; насупрот томе, ослабљена амплитуда градијента слике B је показатељ губитка информација. Укупни губитак информација се добија интеграцијом (усредњавањем) локалних губитака, датих са $1 - Q_C^{AB}(n,m)$, на позицијама где је амплитуда градијента референтне слике већа од одговарајуће амплитуде тест слике, тј. на позицијама где је показатељ $r(n,m)=1$:

$$Q_{out}^{AB} = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M r(n,m) (1 - Q_C^{AB}(n,m)) \quad (34)$$

$$r(n,m) = \begin{cases} 1, & g_A(n,m) > g_B(n,m) \\ 0, & g_A(n,m) \leq g_B(n,m) \end{cases} \quad (35)$$

4.1.3. Убачене (лажне) информације

Информациони артефакти (лажне информације) представљају визуелне информације унете у тест слику током преноса, али које немају подударна обележја у референтној слици. То су информације које значајно смањују употребљивост тест слике и у неким применама могу довести до озбиљних последица. Артефакти који се

јављају током преноса су на Сл. 4.1 обележени са Q_{in}^{AB} . У оквиру предложеног приступа артефакти се могу евалуирати кроз информације о градијенту које постоје у тест слици B , али их нема у изворној слици A .

Локалне процене артефаката, $Q_{in}^{AB}(n,m)$, се евалуирају као губици информација на позицијама где су амплитуде градијента тест слике веће од одговарајућих амплитуда изворне слике. Збирна вредност артефаката у процесу преноса слике се одређује кроз интеграцију (усредњавање) локалних процена на свим позицијама тест слике:

$$Q_{in}^{AB} = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M (1-r(n,m))(1-Q_c^{AB}(n,m)) \quad (36)$$

Заједничке информације, Q_c^{AB} , изгубљене информације, Q_{out}^{AB} , и информациони артефакти, Q_{in}^{AB} , дају укупне информације:

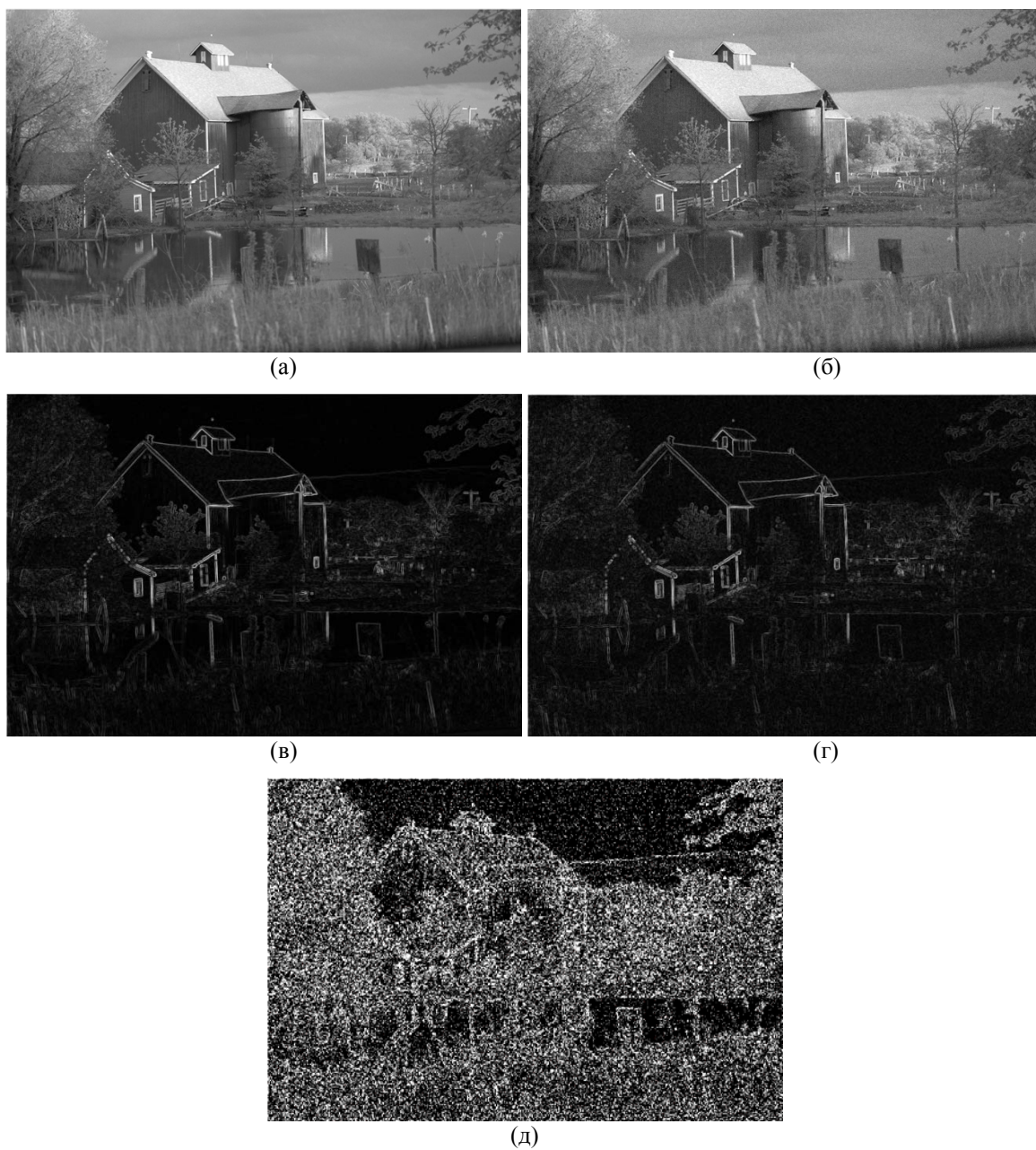
$$Q_c^{AB} + Q_{out}^{AB} + Q_{in}^{AB} = 1 \quad (37)$$

и у оквиру овог дела дисертације биће коришћене као три објективне мере за процену квалитета слике, тј. три различита аспекта преноса слике.

На Сл. 4.2-Сл. 4.4 приказане су изворне слике, њихове модификације настале додавањем Гаусовог шума, JPEG компресијом и замрљањем, амплитуде градијента изворних и тест слика и матрице класификације пиксела тест слика на основу амплитуде градијента (35). Бели пиксели матрица класификације представљају пикселе код којих је амплитуда градијента изворне слике већа од амплитуде градијента тест слике, тј. постоји информациони губитак. Насупрот њима, црни пиксели одговарају позицијама на којима постоје информациони артефакти, тј. лажне информације.

Присуство убачених информација на Сл. 4.2, настало због адитивног Гаусовог шума, нарочито је уочљиво у глатким регионима тест слике (црни пиксели). Додавањем шума у овим регионима долази до повећања амплитуде градијента. Овај ефекат није тако изражен у регионима са богатом текстуром (ефекат зависи и од снаге шума). У примеру са Сл. 4.2 највећи број пиксела тест слике сврстан је у категорију пиксела код којих постоје лажне информације – 72% пиксела, што се одражава и на вредност мере Q_{in}^{AB} . У датом примеру вредности заједничких, изгубљених и убачених информација су $[Q_c^{AB} \ Q_{out}^{AB} \ Q_{in}^{AB}] = [0.4858 \ 0.1088 \ 0.4054]$.

4. Посебни аспекти процене квалитета слика

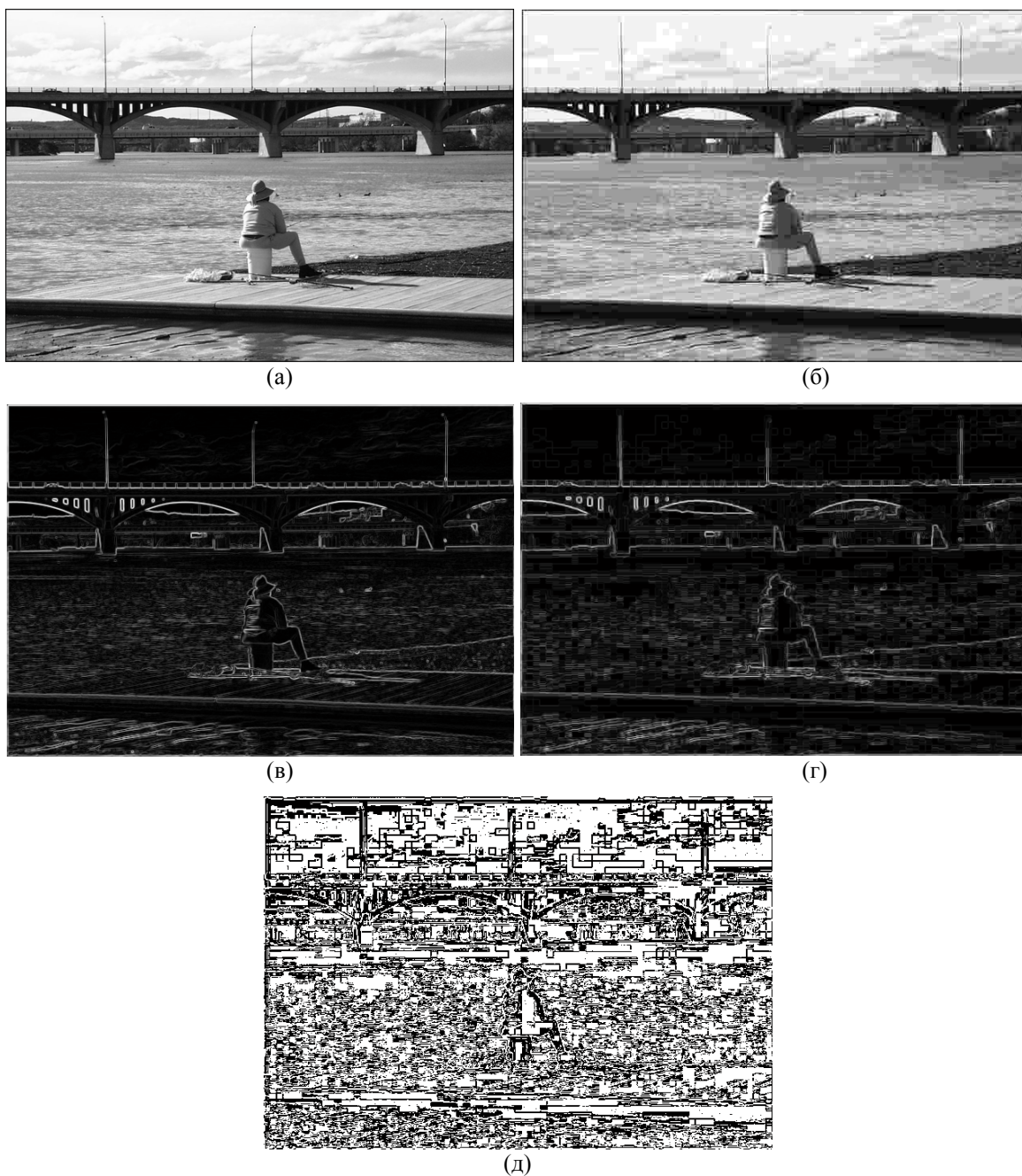


Сл. 4.2 (а) изворна слика, (б) тест слика настала додавањем Гаусовог шума изворној слици, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).

Деградација Сл. 4.3(а) настала JPEG компресијом довела је до губитка високофреквенцијских компоненти унутар блокова, као и до карактеристичних блоковских ефеката на граници суседних блокова. Губитак високофреквенцијских компоненти довео је до смањења амплитуде градијента, док су блоковски ефекти довели до повећања амплитуде градијента. У примеру са Сл. 4.3 58% пиксела тест

4. Посебни аспекти процене квалитета слика

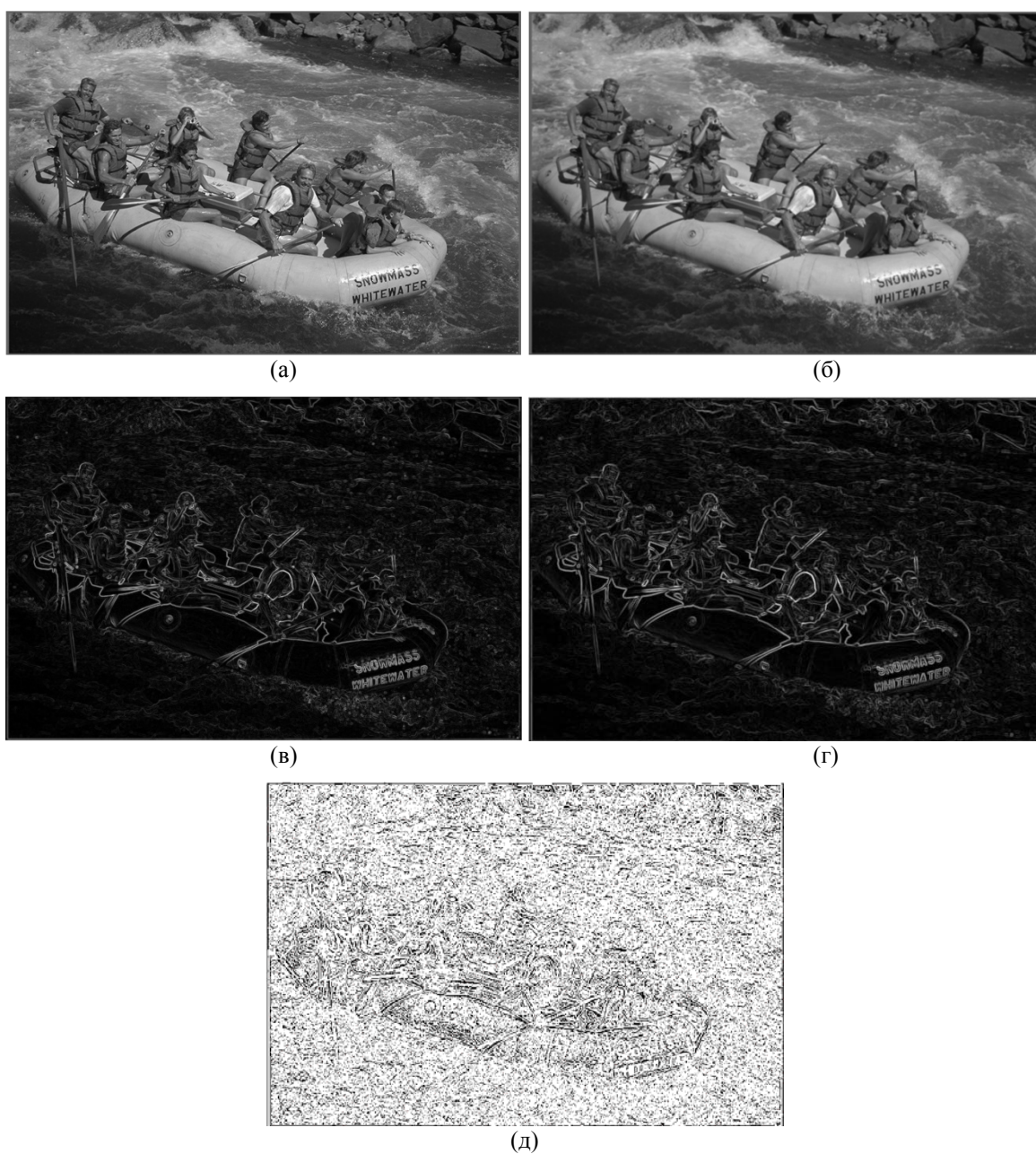
слике сврстано је у категорију пиксела код којих је дошло до губитка информација, док је 42% пиксела сврстано у категорију пиксела код којих постоје лажне информације. Оваква сразмера се одражава и на вредности заједничких, изгубљених и убачених информација, и за дати пример је $[Q_C^{AB} \ Q_{out}^{AB} \ Q_{in}^{AB}] = [0.3519 \ 0.3958 \ 0.2523]$.



Сл. 4.3 (а) изворна слика, (б) тест слика настала ЈРЕГ компресијом изворне слике, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).

4. Посебни аспекти процене квалитета слика

Насупрот примеру са Сл. 4.2, у коме је доминантан ефекат лажних информација, у примеру са Сл. 4.4 највећи број пиксела (80%) је према класификацији датој са (35) сврстан у класу пиксела са изгубљеним информацијама. Лажне информације су присутне у околини ивица. Ефекат губитка информација се уочава и кроз однос заједничких, изгубљених и убачених информација, где су добијене вредности $[Q_C^{AB} \ Q_{out}^{AB} \ Q_{in}^{AB}] = [0.6349 \ 0.2894 \ 0.0757]$.



Сл. 4.4 (а) изворна слика, (б) тест слика настала замрљањем изворне слике, (в) амплитуде градијента изворне слике, (г) амплитуде градијента тест слике и (д) класификација пиксела тест слике на основу (35).

4.1.4. Резултати

Аспекти трансфера слике су презентовани кроз евалуацију на LIVE бази слика [223]. Перформансе три аспекта су евалуиране на основу њихових способности естимације субјективних импресија квалитета слике [5], и то кроз монотоност предикције (SROCC) и тачност предикције (LCC). Поред перформанси предложена три аспекта дате су перформансе PSNR, SSIM и мере за процену квалитета без референцирања BIQI [164]. За нелинеарно пресликавање субјективних и објективних скорова квалитета користи се логистик крива са четири параметра.

Перформансе објективних мера процене на подскуповима и комплетној LIVE бази дате су у Таб. 4.1 и Таб. 4.2. Најбоља два резултата од анализирана три аспекта квалитета су осенчена.

Таб. 4.1 Корелација рангова (SROCC) субјективних и објективних скорова на подскуповима и комплетној LIVE бази.

<i>Тип деградације</i>	Q_{in}^{AB}	Q_{out}^{AB}	Q_C^{AB}	<i>PSNR</i>	<i>SSIM</i>	<i>BIQI</i>
JPEG2000	0.6206	0.8724	0.9503	0.8954	0.9355	0.7995
JPEG	0.2648	0.9439	0.9650	0.8809	0.9449	0.8914
White noise	0.8976	0.0653	0.9298	0.9854	0.9629	0.9510
Gaussian blur	0.7258	0.9377	0.9548	0.7823	0.8944	0.8463
Fast fading	0.4631	0.9294	0.9639	0.8907	0.9413	0.7067
All data	0.4399	0.5338	0.9385	0.8756	0.9104	0.8195

Таб. 4.2 Коefицијент линеарне корелације (LCC) субјективних и објективних скорова на подскуповима и комплетној LIVE бази.

<i>Тип деградације</i>	Q_{in}^{AB}	Q_{out}^{AB}	Q_C^{AB}	<i>PSNR</i>	<i>SSIM</i>	<i>BIQI</i>
JPEG2000	0.6701	0.8767	0.9560	0.8996	0.9410	0.8086
JPEG	0.2953	0.9445	0.9653	0.8879	0.9504	0.9011
White noise	0.9215	0.1093	0.9466	0.9857	0.9695	0.9538
Gaussian blur	0.6809	0.9235	0.9429	0.7835	0.8743	0.8293
Fast fading	0.3940	0.9434	0.9655	0.8895	0.9428	0.7328
All data	0.4305	0.6470	0.9378	0.8701	0.9014	0.8205

Напомена: Резултати у Таб. 4.1 и Таб. 4.2 за меру *BIQI* су преузети из [164].

У Таб. 4.2 дате су вредности коefицијента линеарне корелације за све типове деградације и збирно. Перформансе мере Q_C^{AB} и по тачности предикције (LCC) превазилазе перформансе Q_{in}^{AB} и Q_{out}^{AB} мера (на подскуповима и збирно). Овакав

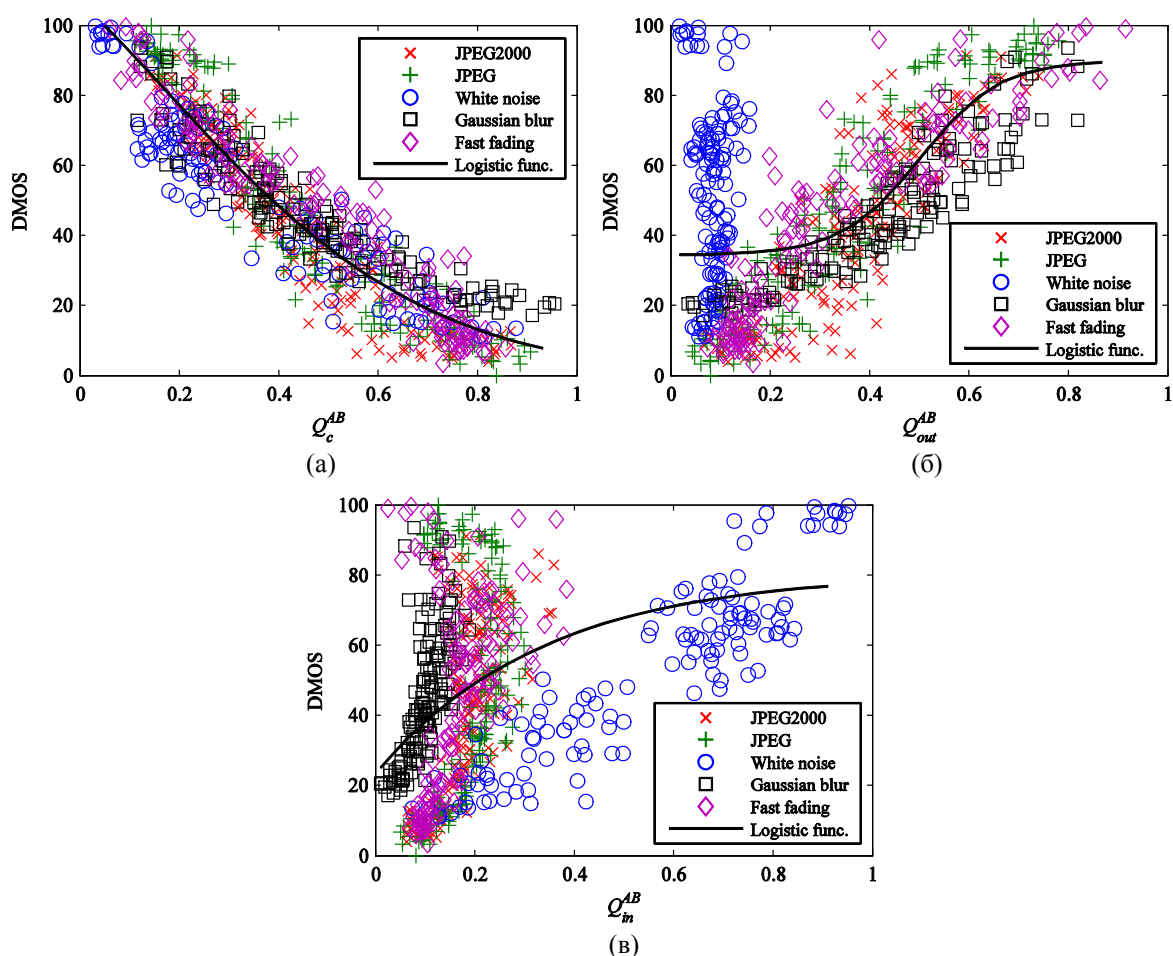
резултат је очекиван јер Q_{in}^{AB} и Q_{out}^{AB} аспекти користе информације које су присутне само у изворној слици или слици са деградацијом док Q_C^{AB} мера (аспект) користи заједничке информације.

Такође, док Q_{in}^{AB} има већи степен слагања са субјективним импресијама квалитета на White noise подскупу, Q_{out}^{AB} има већу корелацију на осталим подскуповима. И овакав резултат је очекиван јер су информациони артефакти који су евалуирани кроз Q_{in}^{AB} аспект доминантан вид дисторзије код слика са додатим шумом. Губитак информација који је заједнички за све технике са компресијом као и за слике са замрљањем је правилно евалуиран са Q_{out}^{AB} мером. Перформансе Q_{in}^{AB} мере на подскупу слика са адитивним шумом и перформансе Q_{out}^{AB} мере на осталим подскуповима су незнатно лошије од перформанси Q_C^{AB} мере.

Ако се упореди однос мера за објективну процену квалитета са потпуним референцирањем (Q_C^{AB} и SSIM) и мере BIQI за процену квалитета без референцирања уочава се да је предност на страни потпуног референцирања. Перформансе мера су упоредиве једино на класи слика са адитивним шумом. Перформансе мере BIQI су на глобалном плану и на три подскупа слика лошије и од конвенционалног PSNR приступа. Због тога би перформансе NR мере BIQI требало побољшати увођењем додатног аспекта који би на бољи начин обухватио изгубљене информације.

Визуелна илустрација односа субјективних података и објективних предикција за комплетну LIVE базу слика приказана је кроз дијаграме расипања на Сл. 4.5.

Са Сл. 4.5 се уочава линеарна веза субјективних скорова и скорова добијених одређивањем заједничких информација изворних и тест слика, Q_C^{AB} . Расипање скорова је константно у комплетном опсегу квалитета. Додатно, дијаграми расипања Q_{in}^{AB} и Q_{out}^{AB} мера илуструју зашто процена квалитета није једноставна – постоје значајне разлике у трендовима губитака информација и информационих артефаката за различите типове дисторзија. Може се рећи да се објективни скорови налазе између трендова слика са адитивним шумом (где је доминантно уношење лажних информација) и слика са замрљањем (где је доминантан губитак информација).



Сл. 4.5 Дијаграми расипања субјективних импресија (DMOS) и уведених аспеката на LIVE бази слика: (а) заједничке информације, (б) изгубљене информације и (в) убачене информације.

Индиректно се може закључити да коришћење једне стратегије у процени квалитета (која одговара евалуацији изгубљених или убачених информација) се може користити у процени квалитета слика са специфичним типом дисторзије, али да то није довољно за меру која би се користила на глобалном плану. Овакав закључак потврђује претпоставке да визуелни систем човека користи сложене механизме у процени квалитета слике. Комбинација различитих аспеката у процени квалитета може дати добре резултате субјективне предикције [57].

4.2. Примена мапе субјективног значаја у евалуацији квалитета

Евалуација субјективног значаја визуелних информација није једноставан задатак и она представља део моделовања визуелног система човека. Користи се у

апликацијама обраде и компресије слике и предмет је бројних истраживања. У контексту објективне процене квалитета слике, евалуација значаја идентификује информације на слици које значајно утичу на посматрача када формира утисак о квалитету. То омогућава да се повећа утицај битних информација на коначни резултат процене. Другим речима, значајније информације морају имати већи значај у укупној процени.

Процена квалитета слике најчешће се спроводи у две фазе. У првој фази квалитет се евалуира на локалном нивоу (у блоковима $k \times k$ пиксела или у регионима слике) и добија се мапа квалитета/дисторзије у простору слике (на нивоу пиксела). У другој фази, примењује се алгоритам интеграције локалних скорова квалитета/дисторзије како би се од мапе квалитета добила једна вредност која представља квалитет комплетне слике. И док је велики труд уложен у развој прве фазе процене квалитета, оптимална стратегија интеграције локалних скорова квалитета је предмет бројних истраживања [124-148]. Увођење мапа визуелног значаја у евалуацији квалитета и технике интеграције локалних скорова квалитета су описани у делу дисертације 2.7.

Уколико је позната мапа субјективног значаја, крајња мера квалитета се може добити као нормализована сума мапе квалитета тест слике модулисана са мапом субјективног значаја w . То значи да се мера квалитета тест слике B може добити као:

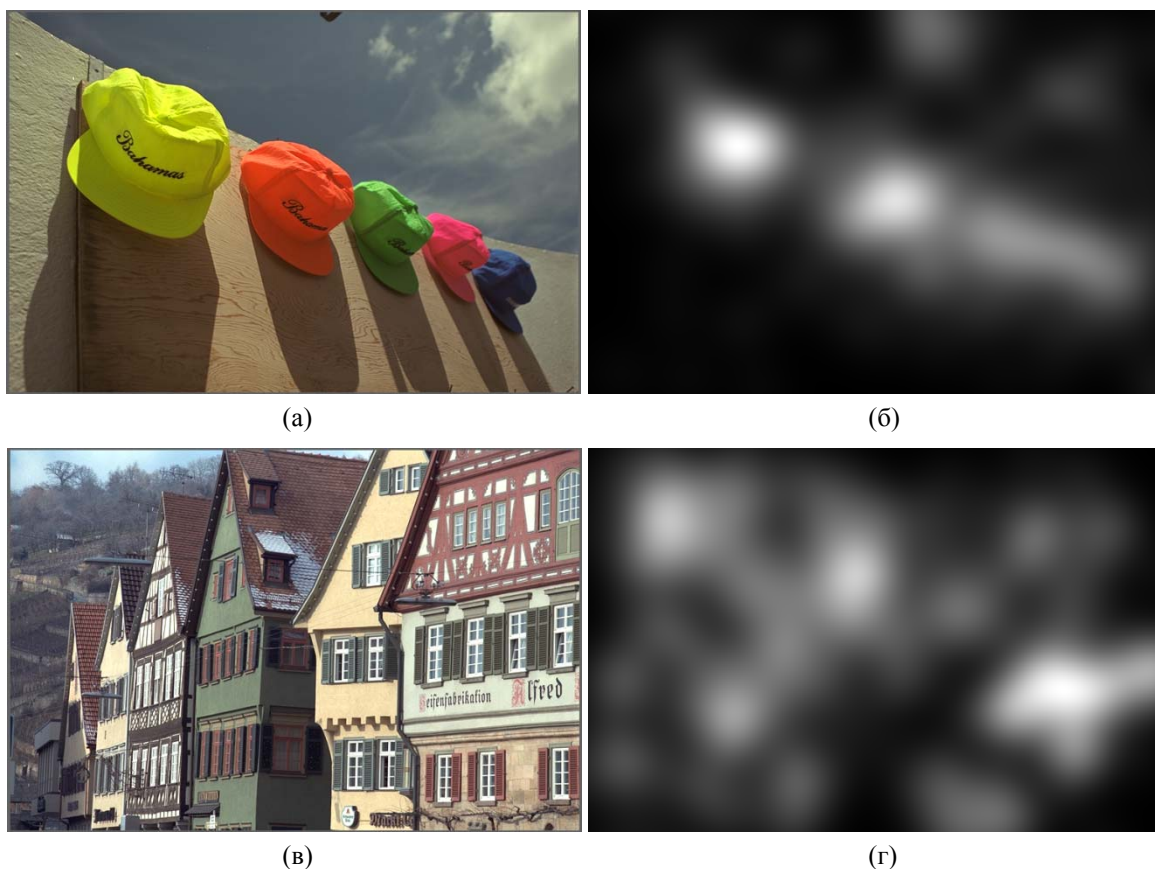
$$Q_w^{AB} = \frac{\sum_{n=1}^N \sum_{m=1}^M Q^{AB}(n,m)w(n,m)}{\sum_{n=1}^N \sum_{m=1}^M w(n,m)}, \quad (38)$$

где $Q^{AB}(n,m)$ представља вредност очувања информација о градијенту на позицији (n,m) . Уместо мере очувања информација о градијенту, Q^{AB} , може се користити било која друга мера која у евалуацији квалитета полази од анализе очувања информација или структурне сличности на нивоу пиксела.

Из (38) се уочава да је евалуација значаја потребна у истој резолуцији као што су резолуције тест и изворне слике и њен циљ је формирање мапе субјективног значаја, $w(n,m)$, која одсликава субјективну важност сваког пиксела слике. На Сл. 4.6

4. Посебни аспекти процене квалитета слика

приказане су две изворне слике и њихове мапе визуелног значаја (светлије вредности имају већи визуелни значај). Изворне слике су из LIVE базе слика [223], док су мапе визуелног значаја изведене из реалних експеримената са 20 посматрача [233, 234]. За разлику од мапе визуелног значаја изворне Сл. 4.6(а), где су посматрачи фокусирани на објекте у предњем плану, визуелни значај изворне Сл. 4.6(в) је равномерније распоређен.



Сл. 4.6 (а) изворна слика, (б) мапа визуелног значаја Сл. 4.6(а), (в) изворна слика и (г) мапа визуелног значаја Сл. 4.6(в).

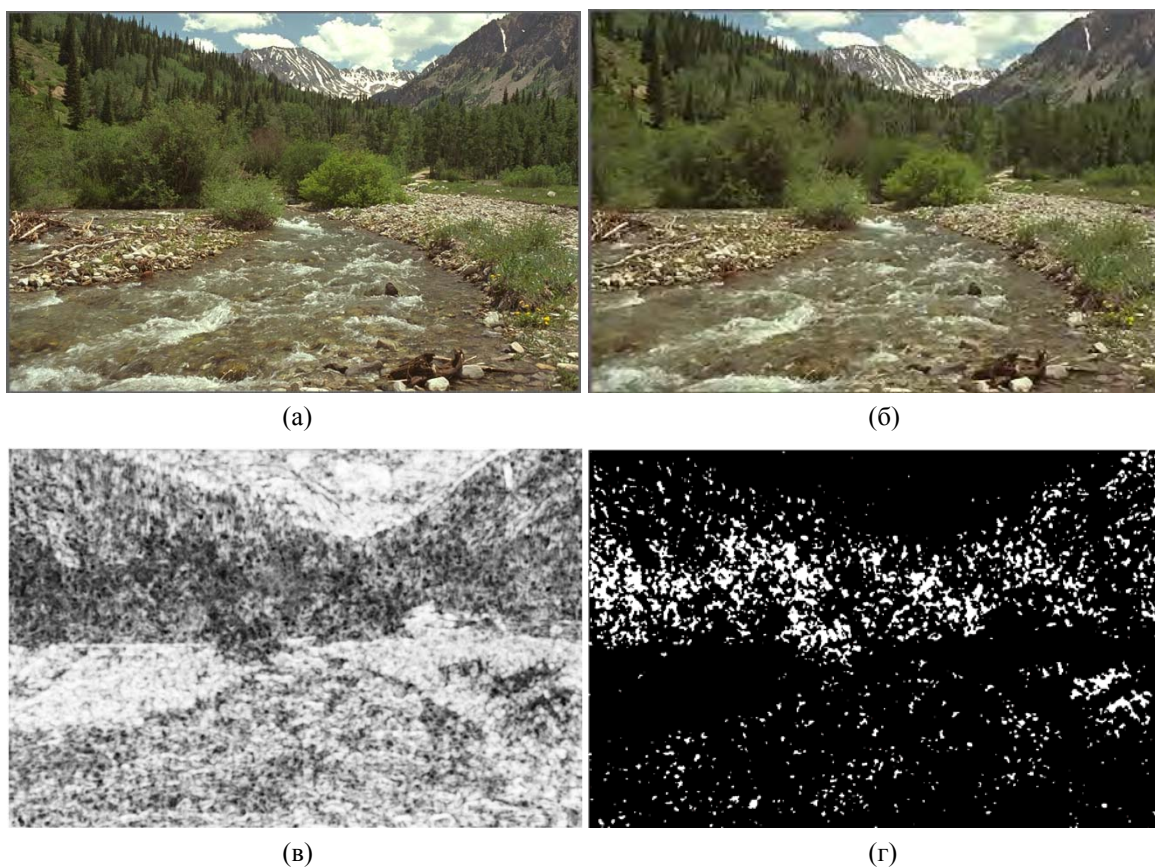
Дистрибуција субјективног значаја на сцени зависи од бројних фактора [220, 235]: контраста региона, димензија региона, облика региона, позиције региона, припадности предњем плану/позадини, припадности ентитетима вишег реда (објектима као што су путеви, реке, возила и сл.), боје, контекста, ...

У [220] субјективни значај пиксела одређен је као производ три фактора. Први фактор је амплитуда градијента, односно контраст. Амплитуда градијента је можда најбитнији фактор који утиче на субјективни значај детектоване ивице у сигналу слике. Израженије (снажније) ивице ће привући више пажње и због тога им се мора дати већи значај него слабијим ивицама. Други фактор је позиција пиксела на слици а трећи фактор припадност ентитетима вишег реда (објектима као што су путеви, реке, возила и сл.). Трећи фактор је у раду изведен из оријентације градијента пиксела слике.

У [236] предложен је трокомпонентни структурни индекс сличности (3-SSIM) модулисан са мапом субјективног значаја. Значај SSIM вредности се одређују након поделе пиксела слике у три класе (ивице, текстуре и глатки региони). Пикселима сврстаним у класу ивица дат је дупло већи значај него пикселима преостале две класе. С обзиром да су резултати процене добијени коришћењем 3-SSIM бољи од SSIM и MS-SSIM аутори су у [236] наговорили да је значај ивица можда већи од мултивеличинске анализе.

У раду [237] су истражене две хипотезе просторног удруживања за популарну SSIM меру. Прва је визуелна пажња (visual attention) и gaze direction – тј. где човек гледа. Побољшања су уочена нарочито за слике са замрљањем и федингом које теже да наруше структуру субјективно значајних карактеристика. Друга хипотеза је да посматрач тежи да “лоше” регионе слике обради са већом пажњом – тачношћу него “добре” регионе и на тај начин да утисак о слици формира давањем већег значаја “лошим” регионима слике. Због тога је разуман приступ коришћење $p\%$ најнижих локалних скорова квалитета који одговарају деловима слике са лошим квалитетом и давање већег значаја овим скоровима. Предложено је коришћење хипотезе о $p\%$ најнижих локалних скорова квалитета и једноставније мере SSIM као алтернативе MS-SSIM ($p=6\%$ је дало задовољавајуће резултате).

На Сл. 4.7 приказане су изворна и тест слика, мапа квалитета/дисторзије тест слике добијена применом SSIM мере и 6% пиксела са најнижим локалним скоровима квалитета (бели пиксели). Коришћењем најнижих скорова квалитета детектовани су региони код којих је дошло до потпуног губитка информација о структури.



Сл. 4.7 (а) изворна слика, (б) тест слика, (в) мапа квалитета добијена помоћу SSIM и (г) пиксели са најнижим локалним SSIM скоровима квалитета (бели пиксели).

У раду [238] мапе квалитета/дисторзије су одређене на два начина: као апсолутне разлике између изворне и тест слике и коришћењем индекса структурне сличности. Мапа значаја формирана је на два начина. Први начин је на основу броја реалних фиксација на свакој позицији пиксела. Фиксацијама је у овом случају дат исти значај. Други начин је на основу трајања реалних фиксација на свакој позицији пиксела. Међутим, увођењем мапа субјективног значаја нису остварена значајна побољшања у односу на просторно усредњавање вредности квалитета/дисторзије.

У [239] се користи неуниформни метод просторне интеграције локалних SSIM вредности, тј. уведена је мапа субјективног значаја која зависи од локалних варијанси нивоа сивог (variance-weighted pooling method):

$$w = g(\mathbf{x}, \mathbf{y}) = \sigma_x^2 + \sigma_y^2 + C_2, \quad (39)$$

где су σ_x и σ_y стандардне девијације сигнала x и y , тим редом, а константа C_2 представља основну – минималну тежину. Оправданост коришћења (39) лежи у чињеници да делови слике са већом енергијом (већа варијанса) садрже више информација. Уколико је примарни циљ визуелног система издвајање корисних информација са сцене, тада ће региони са већом енергијом привући већу пажњу посматрача, па им се због тога мора дати и већи значај.

У [124] су истражене три технике просторне интеграције: Минковски интеграција, интеграција на основу локалног квалитета (local quality/distortion-weighted pooling) и интеграција на основу информационог садржаја (information content-weighted pooling). Експерименти су показали да сва три метода могу побољшати предикцију мера за процену квалитета слика, али је трећи метод показао највећи потенцијал да би се прихватио као генералан и робустан. Функција тежине дефинисана на основу информационог садржаја је дата са:

$$w = g(\mathbf{x}, \mathbf{y}) = \log \left[\left(1 + \frac{\sigma_x^2}{C_2} \right) \left(1 + \frac{\sigma_y^2}{C_2} \right) \right], \quad (40)$$

где је C_2 константа, а σ_x и σ_y стандардне девијације вектора x и y . Уведени су информациони садржаји изворне и слике са дисторзијом како би алгоритам испуњавао услов симетричности. Међутим, из (40) се види да уколико су варијансе једнаке нули (глатки региони) њихов субјективни значај је такође једнак нули. И глатким регионима потребно је дати мали коначан значај w_{min} јер одсуство сигнала носи информацију о томе да нема промена нивоа сивог.

У раду [129] спроведена је широка студија на шест јавно доступних база слика којом је потврђено да се коришћењем информационог садржаја слике у пондерисању објективних скорова квалитета може остварити добитак.

Све ово указује да се увођењем локалног значаја може извршити побољшање објективних мера процене квалитета.

4.2.1. Одређивање мапе субјективног значаја

Људски визуелни систем користи бројне интерне механизме у одређивању субјективног значаја презентованих информација. Ови механизми обухватају више

нивое обраде као што су детекција и препознавање објеката, спознаја ситуације и сл. Дизајн предложене објективне мере има за циљ да искључиво користи ниже нивое обраде како би се смањила рачунарска комплексност па је одређивање мапе субјективног значаја рађено на нивоу пиксела.

У дисертацији је испитана техника одређивања визуелног значаја на основу информационог садржаја. Субјективни значај је одређен на основу локалног садржаја слике (изворне или обе слике), а не на основу локалног квалитета/дисторзије.

Основна претпоставка теорије информација је да се генерисање информације може моделовати као пробабилистички процес који се може мерити на начин који је у складу са интуицијом. У складу са овом претпоставком, за случајни догађај E који се јавља са вероватноћом $P(E)$ кажемо да садржи:

$$I(E) = \log_2 \frac{1}{P(E)} = -\log_2 P(E), \quad (41)$$

јединица информације. Величина $I(E)$ се често назива сопствена информација догађаја E . Генерално говорећи, количина информација придружена догађају E је обрнуто сразмерна са вероватноћом јављања догађаја E . Ако је $P(E)=1$ (тј. догађај E је сигуран догађај), пренета количина информација је $I(E)=0$. То је због тога што нема неизвесности догађаја E . Међутим, ако је $P(E)=0.99$, податак да се десио догађај E носи са собом малу количину информација.

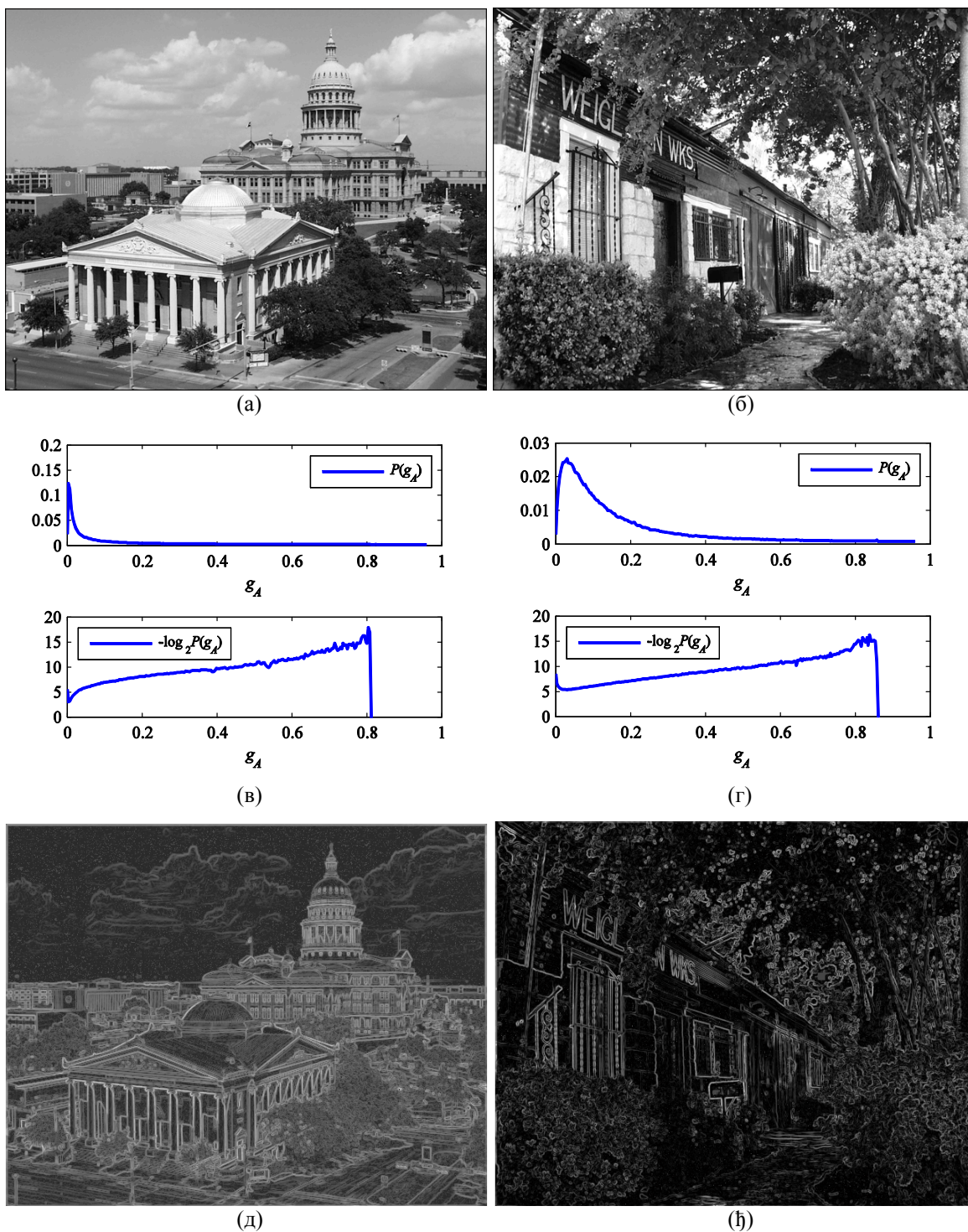
Субјективни значај у раду је одређен на два начина. Први начин је на основу локалног информационог садржаја изворне слике (слике A):

$$w(g_A) = -\log_2 P(g_A), \quad (42)$$

где је $P(g_A)$ функција густине вероватноће амплитуде градијента изворне слике. Из (42) се види да значај пиксела зависи од расподеле вредности градијента. Значај пиксела на позицији (n,m) се може тумачити као број бита потребних за репрезентацију вредности његовог градијента $g_A(n,m)$.

На Сл. 4.8 приказане су две изворне слике, хистограми градијента изворних слика, функције значаја пиксела и мапе значаја.

4. Посебни аспекти процене квалитета слика



Сл. 4.8 (а) (б) изворне слике, (в) (г) хистограми градијента изворних слика и функције значаја и (д) (ж) мале значаја.

Са Сл. 4.8(в) и Сл. 4.8(г) види се да хистограми градијента изворних слика личе на Рејлијеву расподелу. У извођењу оптималне вредности прага детекције

ивице често се полази од претпоставке да је шум у слици адитивни, бели Гаусов шум. Ако се модуо градијента g_A израчунава као квадранти корен из суме квадрата хоризонталног и вертикалног градијента, што је случај у овом раду, могуће је одредити расподелу густине вероватноће модула градијента $P(g_A)$ који потиче од шума у слици. За наведену претпоставку, модуо градијента који потиче од Гаусовог шума у слици има Рејлијеву (Rayleigh) расподелу, док је расподела угла градијента униформна [218]. Међутим, када се слика састоји од реалне сцене са суперпонираним шумом, хистограм модула градијента неће имати тачно Рејлијеву расподелу. Стога је у [240] извршена анализа понашања хистограма градијента под реалним условима. Анализа је показала да у реалним сликама модули градијената који потичу од стварних ивица имају знатно веће вредности од модула градијената који потичу од шума.

За претпоставку о Рејлијевој расподели градијента функција значаја (42) ће бити нелинеарна функција, што је у складу са нелинеарном природом визуелног система човека. Нелинеарне функције значаја се виде на сликама Сл. 4.8(в) и Сл. 4.8(г). Додатно се на Сл. 4.8(д) и Сл. 4.8(ђ) уочава да се оваквим приступом већи значај даје пикселима који носе информације о промени нивоа сивог (светлији пиксели), тј. о ивицама изворне слике.

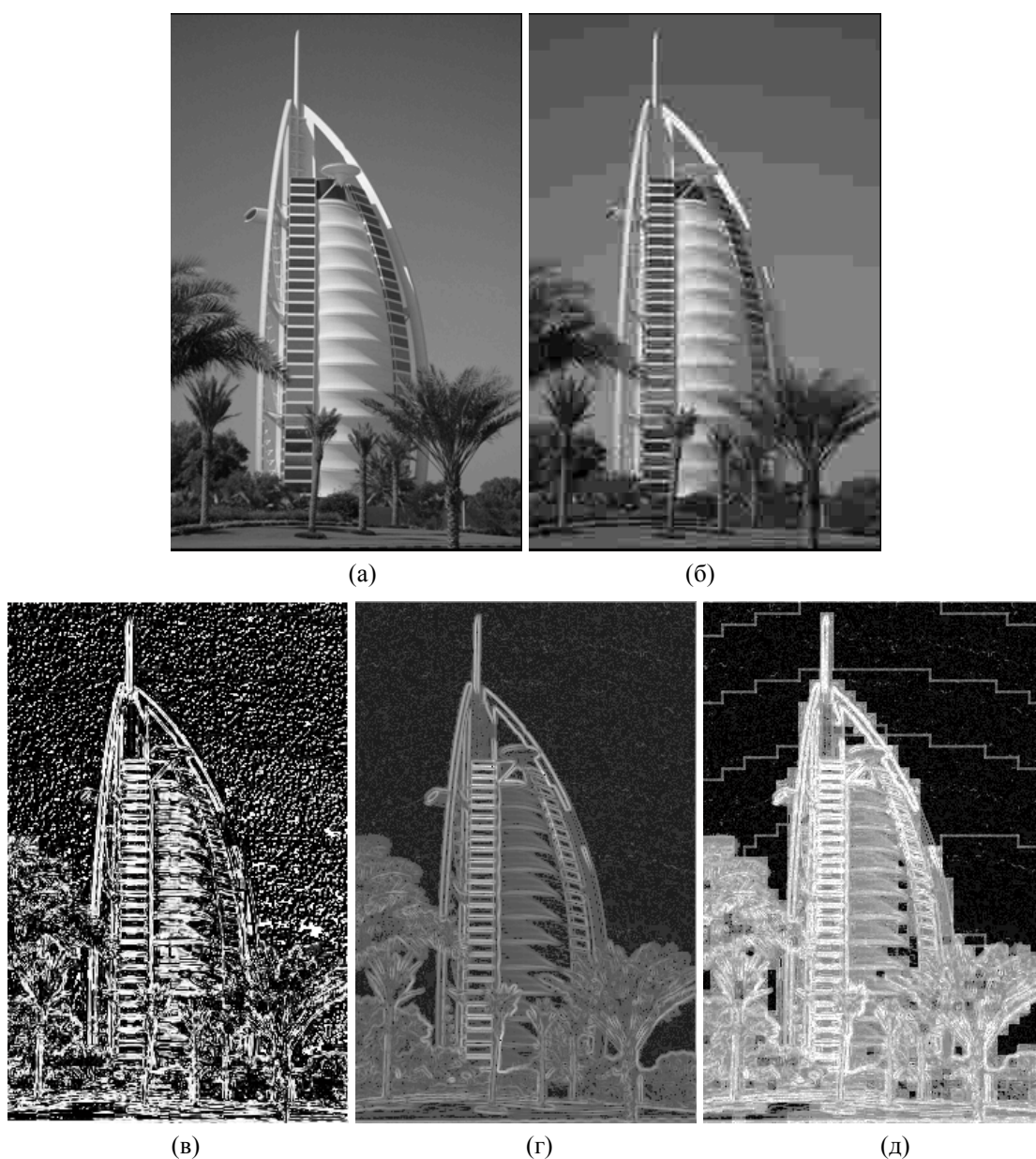
Како би се у обзир узео и локални информациони садржај тест слике уведена је и друга функција значаја пиксела:

$$w(g_A, g_B) = -\log_2 P(g_A, g_B), \quad (43)$$

где је $P(g_A, g_B)$ здружена расподела градијента изворне и тест слике. На овај начин збирна процена (38) ће испуњавати услов симетрије ($Q_w^{AB} = Q_w^{BA}$).

Увођењем функција значаја (42) и (43) избегава се хеуристичко одређивање значаја пиксела са нултим градијентом (константе C_2 из [239] – релација (39), односно w_{min} из [220]), јер значај зависи од фреквенције појаве нултог градијента, тј. одређује се из реалних података а не произвољним избором.

На Сл. 4.9 приказане су изворна и тест слика, мапа очувања информација о градијенту и мапе значаја добијене на предложене начине (светлије вредности имају већи значај).



Сл. 4.9 (а) изворна слика, (б) тест слика, (в) мапа очувања параметара градијента, (г) (д) мапе субјективног значаја добијене коришћењем (42) и (43), тим редом.

Са Сл. 4.9(г) и Сл. 4.9(д) види се да објекти у предњем плану имају већи субјективни значај од позадине (небо). Такође, са Сл. 4.9(д) види се да пиксели код којих постоје деградације добијају на значају – пиксели позадине код којих се уочава блоковски ефекат. Ови пиксели добијају на значају јер се у формирању мапе значаја поред изворне користи и тест слика (на којој су присутне деградације).

4.2.2. Примена мапе субјективног значаја у процени квалитета слике

Мапе субјективног значаја које су добијене на основу (42) и (43) интегрисане су са мапама очувања информација о градијенту (Сл. 3.1, односно релација (38)) и добијене су вредности квалитета Q_{w1}^{AB} и Q_{w2}^{AB} , тим редом. Мапа визуелног значаја која се користи код мере Q_{w1}^{AB} изведена је из амплитуде градијента изворне слике, док се код мере Q_{w2}^{AB} користи мапа визуелног значаја изведена из амплитуда градијента изворне и тест слике. Допринос употребе мапа субјективног значаја у процени квалитета анализиран је на базама слика коришћеним у претходном поглављу рада (LIVE, Тоуама, IVC и CSIQ). Резултати спроведене анализе дати су у Таб. 4.3–Таб. 4.11. У табелама су презентоване перформансе стандардне мере очувања ивица, Q^{AB} (без пондерисања), пондерисаних мера, Q_{w1}^{AB} и Q_{w2}^{AB} , и мера које су се у претходној анализи показале поузданим – MS-SSIM и VIF. Перформансе су презентоване кроз степен слагања субјективних и објективних скорова, како на нивоу база тако и на подскуповима слика у оквиру база. Подскупови слика и базе на којима је пондерисање са информационим садржајем донело побољшање су обележене знаком (+).

У Таб. 4.3 и Таб. 4.4 дате су перформансе објективних мера процене квалитета на LIVE бази слика. Уочава се да је пондерисање са мапама субјективног значаја на свим подскуповима као и на комплетној бази донело побољшање. Такође, може се рећи да не постоји нека значајна разлика у перформансама пондерисаних мера Q_{w1}^{AB} и Q_{w2}^{AB} . Увођењем пондерисања објективна процена на основу очувања информација о градијенту се по перформансама приближава VIF мери.

Таб. 4.3 Коefицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
MS-SSIM	0.9680	0.9804	0.9707	0.9486	0.9255	0.9382
VIF	0.9772	0.9864	0.9839	0.9743	0.9617	0.9597
Q^{AB}	0.9560	0.9653	0.9466	0.9429	0.9655	0.9378
Q_{w1}^{AB}	0.9645 (+)	0.9764 (+)	0.9664 (+)	0.9562 (+)	0.9655	0.9502 (+)
Q_{w2}^{AB}	0.9650 (+)	0.9758 (+)	0.9581 (+)	0.9563 (+)	0.9679 (+)	0.9461 (+)

4. Посебни аспекти процене квалитета слика

Таб. 4.4 Корелација рангова (SROCC) субјективних и објективних скорова на LIVE бази слика.

Мера	JPEG2000 images	JPEG images	White noise images	Gaussian blur images	Fast fading images	All images
MS-SSIM	0.9638	0.9814	0.9781	0.9542	0.9354	0.9527
VIF	0.9696	0.9846	0.9858	0.9728	0.9650	0.9636
Q^{AB}	0.9503	0.9650	0.9298	0.9548	0.9639	0.9385
Q_{w1}^{AB}	0.9565 (+)	0.9741 (+)	0.9528 (+)	0.9631 (+)	0.9686 (+)	0.9520 (+)
Q_{w2}^{AB}	0.9579 (+)	0.9739 (+)	0.9439 (+)	0.9639 (+)	0.9702 (+)	0.9488 (+)

Увођење мапа субјективног значаја у процени квалитета комплетне Тоуата базе није довело до значајног добитка, Таб. 4.5–Таб. 4.7. Перформансе пондерисаних мера на глобалном плану су у нивоу перформанси изворне, Q^{AB} процене. У оквиру базе дошло је до побољшања перформанси на JPEG2000 подскупу и до пропорционалног погоршања перформанси на JPEG подскупу. Иако је дошло до погоршања перформанси на JPEG подскупу, перформансе пондерисаних објективних мера су и даље боље од перформанси MS-SSIM и VIF мера.

Таб. 4.5 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.8948	0.8910	0.4305	0.5588	7.1429
VIF	0.9137	0.9077	0.4032	0.5087	5.3571
Q^{AB}	0.9322	0.9276	0.3482	0.4529	2.9762
Q_{w1}^{AB}	0.9352 (+)	0.9289 (+)	0.3335 (+)	0.4433 (+)	2.9762
Q_{w2}^{AB}	0.9341 (+)	0.9271	0.3355 (+)	0.4469 (+)	2.9762

Таб. 4.6 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG подскуп.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.8076	0.8659	1.0657	1.1998	41.6667
VIF	0.9000	0.9061	0.4430	0.5390	3.5714
Q^{AB}	0.9603	0.9543	0.2698	0.3450	2.3810
Q_{w1}^{AB}	0.9437	0.9380	0.3083	0.4092	2.3810
Q_{w2}^{AB}	0.9419	0.9342	0.3145	0.4155	2.3810

Таб. 4.7 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика – JPEG2000 подскуп.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.9416	0.9392	0.3237	0.4253	4.7619
VIF	0.9622	0.9559	0.2724	0.3440	3.5714
Q^{AB}	0.9234	0.9191	0.3515	0.4847	3.5714
Q_{w1}^{AB}	0.9405 (+)	0.9336 (+)	0.2984 (+)	0.4291 (+)	3.5714
Q_{w2}^{AB}	0.9391 (+)	0.9321 (+)	0.3020 (+)	0.4339 (+)	3.5714

4. Посебни аспекти процене квалитета слика

Пондерисање са мапама субјективног значаја визуелних информација довело је до побољшања перформанси ивичне објективне мере на три подскупа и комплетној IVC бази, Таб. 4.8 и Таб. 4.9. Занимљиво је уочити да је до побољшања перформанси дошло и на JPEG класи слика, тј. на класи слика за коју је дошло до погоршања перформанси у оквиру Тоуама базе.

Таб. 4.8 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>MS-SSIM</i>	0.9239	0.9467	0.9580	0.8927	0.8647	0.9108
<i>VIF</i>	0.9347	0.9387	0.9887	0.9149	0.9055	0.9026
Q^{AB}	0.8629	0.9090	0.9781	0.9332	0.8998	0.8813
Q_{w1}^{AB}	0.8964 (+)	0.9324 (+)	0.9759	0.9134	0.9190 (+)	0.8846 (+)
Q_{w2}^{AB}	0.8912 (+)	0.9241 (+)	0.9771	0.9141	0.9123 (+)	0.8864 (+)

Таб. 4.9 Корелација рангова (SROCC) субјективних и објективних скорова на IVC бази слика.

<i>Мера</i>	<i>JPEG2000 images</i>	<i>JPEG images</i>	<i>Gaussian blur images</i>	<i>JPEG lumichr images</i>	<i>LAR coded images</i>	<i>All images</i>
<i>MS-SSIM</i>	0.9264	0.9233	0.9458	0.8598	0.8587	0.8971
<i>VIF</i>	0.9356	0.9240	0.9729	0.8775	0.8878	0.8964
Q^{AB}	0.8606	0.8905	0.9669	0.9083	0.8460	0.8699
Q_{w1}^{AB}	0.8932 (+)	0.9153 (+)	0.9616	0.8879	0.8700 (+)	0.8759 (+)
Q_{w2}^{AB}	0.8905 (+)	0.8982 (+)	0.9646	0.8918	0.8652 (+)	0.8766 (+)

Пондерисање са мапама субјективног значаја доноси добитак на подскуповима CSIQ базе и бази у целини, Таб. 4.10 и Таб. 4.11. Овај добитак се нарочито уочава код слика са шумом. И у овом случају постоји побољшање перформанси пондерисаних мера над JPEG модификованим сликама.

Таб. 4.10 Коефицијент линеарне корелације (LCC) субјективних и објективних скорова (након нелинеарне регресије) на CSIQ бази слика.

<i>Мера</i>	<i>White noise images</i>	<i>JPEG images</i>	<i>JPEG2000 images</i>	<i>Pink noise images</i>	<i>Gaussian blur images</i>	<i>Contrast decrements images</i>	<i>All images</i>
<i>MS-SSIM</i>	0.9430	0.9834	0.9788	0.9387	0.9620	0.9453	0.8666
<i>VIF</i>	0.9606	0.9883	0.9776	0.9568	0.9736	0.9416	0.9252
Q^{AB}	0.8433	0.9527	0.9396	0.8222	0.9297	0.9422	0.8556
Q_{w1}^{AB}	0.8870 (+)	0.9785 (+)	0.9602 (+)	0.8752 (+)	0.9404 (+)	0.9496 (+)	0.8794 (+)
Q_{w2}^{AB}	0.8731 (+)	0.9777 (+)	0.9580 (+)	0.8616 (+)	0.9392 (+)	0.9482 (+)	0.8769 (+)

Таб. 4.11 Корелација рангова (SROCC) субјективних и објективних скорова на CSIQ бази слика.

Мера	White noise images	JPEG images	JPEG2000 images	Pink noise images	Gaussian blur images	Contrast decrements images	All images
<i>MS-SSIM</i>	0.9437	0.9643	0.9724	0.9327	0.9747	0.9448	0.8774
<i>VIF</i>	0.9575	0.9703	0.9671	0.9509	0.9744	0.9345	0.9194
<i>Q^{AB}</i>	0.8399	0.9378	0.9255	0.8170	0.9438	0.9437	0.8520
<i>Q_{w1}^{AB}</i>	0.8846 (+)	0.9598 (+)	0.9475 (+)	0.8655 (+)	0.9519 (+)	0.9511 (+)	0.8754 (+)
<i>Q_{w2}^{AB}</i>	0.8688 (+)	0.9602 (+)	0.9449 (+)	0.8512 (+)	0.9510 (+)	0.9491 (+)	0.8741 (+)

Генерално се може закључити да је на глобалном плану увођењем мапа субјективног значаја остварен добитак. Добитак зависи од базе над којом се ради. Утицај предложених мапа субјективног значаја на подскуповима слика из различитих база није једнозначан. На две базе слика (LIVE, CSIQ) је дошло до побољшања перформанси ивичне процене на свим подскуповима. На једној бази слика (IVC) дошло је до побољшања перформанси на три класе слика док су на две класе слика перформансе остале у нивоу перформанси стандардног приступа. Код Тоуата базе дошло је до побољшања перформанси на једној од две расположиве класе слика. На овој бази је дошло до погоршања перформанси пондерисаних мера на JPEG подскупу слика. Истовремено, на JPEG подскуповима LIVE, IVC и CSIQ база је дошло до побољшања перформанси. Овакав резултат може да буде последица реалног динамичког опсега квалитета слика са дисторзијом, где је расположиви опсег квалитета слика у Тоуата бази значајно ужи од динамичког опсега LIVE и CSIQ база. На свим базама слика разлике у перформансама између два предложена приступа са пондерисањем није значајна.

4.3. Сједињавање информација у процени квалитета слике

Индекс структурне сличности (SSIM) [48] је приступ за који се може рећи да је постао стандард у објективној процени квалитета слике, како због једноставности тако и због робусности. Како су слике природних сцена веома структуриране и како је HVS прилагођен издвајању оваквих информација, у основи SSIM мере је идеја да се кроз локална поређења осветљаја, контраста и структуре пореде структурне

информације оригиналне и слике са деградацијом [47, 48]. Три поређења се комбинују у један скор квалитета (SSIM) користећи мултипликативни модел.

Мултипликативна интеграција индивидуалних модела сличности дефинисана са (16) представља константно обележје SSIM и из ње изведених мера. Руководећи се искуством у раду са пробабилистичким системима где једна мала вредност значајно може утицати на глобални скор добијен кроз множење, анализиран је једноставнији, адитивни модел SSIM.

Ивична објективна процена комбинује у мултипликативном смислу информације из два канала (очување информација о амплитуди и оријентацији градијента), (32), тако да је примена адитивног модела сједињавања информација истражена и код процене квалитета кроз очување информација о градијенту.

Нови поглед на SSIM и ивичну објективну процену Q^{AB} , осим адитивног модела интеграције, користи и квалитетом вођену интеграцију локалних скорова индивидуалних канала очувања информација о структури/градијенту. Кроз обимну евалуацију над више од 1600 субјективно евалуираних слика из четири јавно доступне базе показано је да нови приступ даје доста боље слагање са субјективним импресијама квалитета за бројне типове и нивое деградације. Добијени резултати потврђују и теорију да се оцене посматрача о квалитету слике базирају на деловима посматране слике са најлошијим квалитетом.

4.3.1. Сједињавање информација о очувању структуре

Од момента публикавања, SSIM је постао основа за бројне метрике објективне процене квалитета сигнала [229]. У оквиру мноштва радова из ове области у дисертацији је представљен нови поглед на индекс структурне сличности кроз раздвајање компоненти основног модела, предлагање новог начина интеграције (адитивни модел – AM) и приступа просторне интеграције којима се раздвајају индивидуална структурна поређења и користи квалитетом вођена интеграција. Нови приступ је примењен на основни – интензитетски SSIM [48] и градијентни SSIM (GSSIM) [241]. Код градијентног индекса структурне сличности, GSSIM, користе се

ивице слике као носилац структурних информација. Коришћењем предложеног приступа показано је да се добијају поузданије мере, са скоро линеарном везом са субјективним импресијама квалитета.

4.3.1.1. Индекс структурне сличности

Различите технике структурног поређења између оригиналног и изведеног сигнала (сигнала са деградацијом) су у употреби последњих двадесетак година за процену квалитета слике и видеа [177, 178] и мултисензорског сједињавања слика [179]. Оне су шире афирмисане са формулацијом индекса структурне сличности (SSIM) за процену квалитета слике [47, 48]. Под претпоставком да је HVS адаптиран за издвајање структурних информација са сцене, процена квалитета на основу деградације структурних информација између два ненегативна сигнала x и y се може добити кроз директна поређења њихових осветљаја (l), контраста (c) и структуре (s), дефинисаних са:

$$l_{x,y} = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (44)$$

$$c_{x,y} = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (45)$$

$$s_{x,y} = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (46)$$

где су μ_x и μ_y средње вредности, σ_x^2 и σ_y^2 варијансе и σ_{xy} корелација између сигнала x и y . Константе C_1 , C_2 и C_3 су уведене како би се избегле ситуације у којима је именилац приближно једнак нули. Сва три поређења теже јединици када се разлика сигнала смањује (оне су једнаке јединици када је $x = y$), односно смањују се са порастом разлике између одбирака сигнала. На тај начин они одражавају квалитет репрезентације сигнала x у сигналу y и обратно (за поређења важи особина симетрије). Поређења се интегришу у једну меру структурне сличности користећи мултипликативни SSIM модел [48]:

$$SSIM_{x,y} = l_{x,y}^\alpha \cdot c_{x,y}^\beta \cdot s_{x,y}^\gamma \quad (47)$$

где константе α , β и γ дефинишу значај појединачних компоненти, при чему је у уобичајеној форми $\alpha=\beta=\gamma=1$. Како минимална вредност $s_{x,y}$ износи -1 (за инвертоване сигнале – негатив слике), динамички опсег $SSIM_{x,y}$ је од -1 до 1 ($x = y$).

За примену у обради сигнала слике SSIM се евалуира на свакој позицији (за сваки пиксел) слике користећи 11x11 Гаусов кернел кроз који се евалуирају локалне статистике (средња вредност μ и стандардна девијација σ) [48]. Универзални индекс квалитета слике, (UIQI) [47] користи правоугаони прозор уместо Гаусовог. На овај начин се добија скуп локалних естимација структурне сличности између две слике, $SSIM_{x,y}$ за $\forall x,y$. Глобална структурна сличност између две слике се добија као средња вредност локалних SSIM естимација.

Како се SSIM вредности одређују на локалном нивоу, један од начина да се мера унапреди је кроз бољи начин комбиновања локалних скорова квалитета [124]. Предложене су различите стратегије комбиновања локалних естимација: коришћење визуелних фиксација и локалног квалитета [237], пондерисањем на нивоу региона [242], пондерисањем на основу информационог садржаја [129] и коришћењем стварних визуелних информација битних за посматрача (добијених снимањем током процене квалитета) [238].

4.3.1.2. Квалитетом вођена процена квалитета

У дисертацији је предложен нови поглед на индекс структурне сличности код кога се врши декомпозиција основног модела интеграције из (47), напушта класично усредњавање локалних естимација квалитета и уводи квалитетом вођена интеграција индивидуалних структурних поређења.

У ранијим радовима је показано да на субјективну естимацију квалитета више утичу региони лошијег квалитета [237]. Ова чињеница је искоришћена у приступу интеграције, где се уводи квалитетом вођена интеграција – интеграција најнижих скорова квалитета, тј. у поступку интеграције се користи део ($p\%$) најнижих локалних скорова квалитета како би се извео глобални скор. За разлику од приступа из [237], код кога се користе локалне естимације SSIM, у раду је испитано

раздвајање индивидуалних структурних поређења, (44)-(46), како би се извршила њихова интеграција, индивидуално, у паровима и збирно. Циљ је да се анализира како независна поређења локалне структуре, контраста и осветљаја утичу на глобални скор квалитета.

Процес интеграције је спроведен на следећи начин. Иницијално су одређене три индивидуалне мапе сличности, $l_{x,y}$, $c_{x,y}$ и $s_{x,y}$, користећи мере поређења дефинисане са (44)-(46). Након одређивања мапа сличности извршено је сортирање њихових вредности у растућем редоследу амплитуда. Глобални скор квалитета добијен је кроз усредњавање најнижих $p\%$ скорова квалитета (48), чиме се добијају вредности $SSIM^{\{l,c,s\}}$. Вредности које су изван ових $p\%$ скорова су одбачене – нису узете у обзир. Поред анализе на индивидуалним мапама сличности (44)-(46), усредњавање најнижих скорова анализирано је и за њихове комбинације, тј. парове или тројку комбинација, кроз мултипликативни модел на нивоу пиксела ($l \times c$, $l \times s$, $c \times s$ и $l \times c \times s$).

$$SSIM^k = \frac{100}{pNM} \sum_{\forall k(n,m) \text{ у најнижих } p\%} k(n,m), \quad k \in \{l, c, s, l \times c, l \times s, c \times s, l \times c \times s\} \quad (48)$$

За валидацију предложеног приступа и процену оптималних вредности параметра p предложена просторне интеграције је примењена на основној SSIM [48] техници и градијентој GSSIM мери [241]. У фази обучавања коришћен је подскуп слика из LIVE базе [223]. LIVE база слика садржи 29 изворних слика различитих садржаја модификованих са различитим типовима и нивоима дисторзије, са доступним субјективним проценама. Из базе је на случајан начин одабран подскуп од 15 референтних слика и њихових модификација (374 слике) за обучавање, док је остатак (405 слика) слика коришћен у фази тестирања. У Таб. 4.12 и Таб. 4.13 дате су оптималне вредности параметра p (у заградама) и перформансе SSIM модела за различите стратегије просторне интеграције (најбољи скорови су подебљани). Анализа је спроведена на подскупу слика за обучавање из LIVE базе. Перформансе мера представљене су кроз коефицијент линеарне корелације (LCC) између субјективних и објективних скорова, Спирманову корелацију рангова (SROCC), средњу апсолутну грешку (MAE), корен средње квадратне грешке (RMSE) и одступање скорова изван интервала поузданости од 95% (OR). Оптималне

4. Посебни аспекти процене квалитета слика

вредности параметра p дате су по свих пет критеријума поређења. Мера поређења осветљаја из стандардног SSIM приступа је резултовала јако узаним динамичким опсегом квалитета што је довело до лошијих перформанси и због чега није узета у разматрање. Код GSSIM приступа овој мери одговара поређење средњих вредности локалних информација о градијенту (што је у ствари контраст). Коришћење овог поређења доприноси повећању коефицијента линеарне корелације са субјективним скоровима за 3% код стандардног GSSIM приступа (упоредити перформансе $c \times s$ и GSSIM у Таб. 4.13).

Таб. 4.12 Перформансе различитих SSIM мера на подскупу слика за обучавање из LIVE базе.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
контраст, c	0.8735	0.8877	10.8529	13.2526	10.1604
c (најнижих $p\%$)	0.9349 (2%)	0.9436 (2%)	7.7530 (2%)	9.6650 (2%)	2.6738 (2%)
структура, s	0.8726	0.8827	10.5205	13.2951	9.8930
s (најнижих $p\%$)	0.9343 (4%)	0.9373 (4%)	7.9105 (4%)	9.7017 (4%)	2.6738 (4-6%)
$I \times c$	0.8765	0.8901	10.7449	13.1077	9.8930
$I \times c$ (најнижих $p\%$)	0.9355 (2%)	0.9444 (2%)	7.7085 (2%)	9.6216 (2%)	2.4064 (2%)
$I \times s$	0.8759	0.8858	10.4078	13.1358	9.6257
$I \times s$ (најнижих $p\%$)	0.9343 (4%)	0.9376 (4%)	7.9185 (4%)	9.7029 (4%)	2.4064 (4%)
$c \times s$	0.9018	0.9091	9.2729	11.7676	7.4866
$c \times s$ (најнижих $p\%$)	0.9423 (2%)	0.9490 (2%)	7.0988 (2%)	9.1125 (2%)	2.4064 (4%)
SSIM (стандардни)	0.9029	0.9101	9.2163	11.7026	6.9519
SSIM (најнижих $p\%$)	0.9418 (2%)	0.9488 (2%)	7.1294 (2%)	9.1486 (2%)	2.4064 (4-6%)

Таб. 4.13 Перформансе различитих GSSIM мера на подскупу слика за обучавање из LIVE базе.

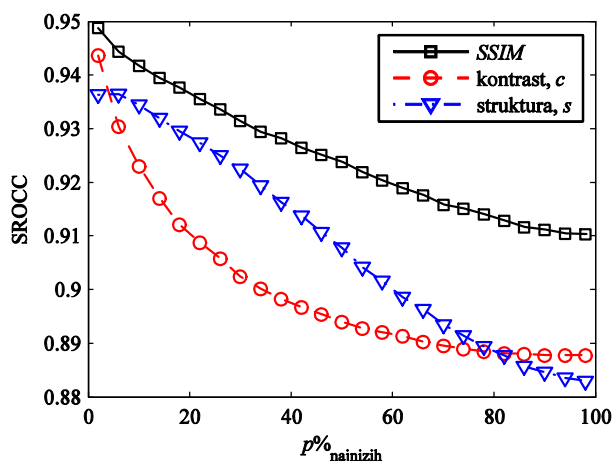
Мера	LCC	SROCC	MAE	RMSE	OR (%)
осветљај, I	0.8986	0.9034	9.6882	11.9451	5.6150
I (најнижих $p\%$)	0.9226 (4%)	0.9271 (4%)	8.4543 (4%)	10.5038 (4%)	2.4064 (10%)
контраст, c	0.8737	0.8922	10.5444	13.2424	10.9626
c (најнижих $p\%$)	0.9260 (2%)	0.9422 (2%)	8.4165 (2%)	10.2784 (2%)	4.5455 (2-4%)
структура, s	0.8490	0.8591	10.8794	14.3839	9.6257
s (најнижих $p\%$)	0.9025 (2%)	0.9102 (2%)	8.6794 (2%)	11.7270 (2%)	5.3476 (14%)
$I \times c$	0.9000	0.9078	9.6552	11.8656	6.9519
$I \times c$ (најнижих $p\%$)	0.9338 (2%)	0.9507 (2%)	7.9907 (2%)	9.7371 (2%)	2.4064 (2-4%)
$I \times s$	0.9306	0.9351	7.8056	9.9659	3.7433
$I \times s$ (најнижих $p\%$)	0.9466 (14%)	0.9497 (16%)	7.1571 (12%)	8.7746 (14%)	1.3369 (14-28%)
$c \times s$	0.9058	0.9127	8.9494	11.5348	6.4171
$c \times s$ (најнижих $p\%$)	0.9441 (12%)	0.9485 (16%)	7.1803 (8%)	8.9765 (12%)	1.6043 (16-24%)
GSSIM (стандардни)	0.9307	0.9348	7.7364	9.9595	3.2086
GSSIM (најнижих $p\%$)	0.9469 (10%)	0.9551 (14%)	6.9640 (4%)	8.7535 (10%)	1.0695 (10%)

Перформансе дате у Таб. 4.12 и Таб. 4.13 показују да се користећи део најнижих скорова квалитета (оптимално је користити 2% најнижих SSIM вредности и 14% најнижих GSSIM вредности – по критеријуму SROCC) значајно могу побољшати субјективно-објективна слагања и перформансе мера објективне процене. Ово важи, како за индивидуалне канале сличности, тако за њихове комбинације па и за комплетне SSIM и GSSIM мере. Закључак из [243] да се поређење осветљаја може игнорисати је потврђен у Таб. 4.12 (упоредити перформансе $c \times s$ и SSIM). Додатно се може уочити да се увођењем градијента у процену квалитета слике добија побољшање у односу на стандардну SSIM имплементацију на основу интензитета ($LCC_{GSSIM} > LCC_{SSIM}$).

На Сл. 4.10 приказана је корелација рангова субјективних и објективних скорова квалитета (SROCC) у функцији параметра p на подскупу слика за обучавање из LIVE базе и то за три мере процене – контраст c , структуру s и збирни SSIM ($lxcxs$) индекс. Јасно се уочава да се перформансе побољшавају са смањењем вредности параметра p , чиме су потврђени закључци из [237] (прикази из [237] и са Сл. 4.10 имају инвертоване $x(p)$ осе – p опада и p расте). У раду [237], где је такође анализиран утицај параметра p , сви пиксели учествују у одређивању глобалног SSIM скорa квалитета. Пикселима који се налазе унутар $p\%$ најнижих скорова придружен је значај r_1 док је остатку дат значај $r_2=1$ ($r_1 > r_2$). Због оваквог приступа у пондерисању мера из [237] је дефинисана за пун опсег вредности p (0-100%), и у перформансама презентованим кроз SROCC или LCC се појављује значајан скок код вредности $p=0\%$. Интеграција која је презентована овде узима у обзир само $p\%$ најнижих скорова игноришући остатак и тако није дефинисана за $p=0\%$. Ако се изузме ова екстремна вредност и скок који она прави, прикази из [237] и са Сл. 4.10 су веома слични. Оба наводе на закључак да се разматрањем дела пиксела са најнижим скоровима може доћи до бољег субјективно-објективног слагања. У овом случају то је 2% најнижих док је то по анализи из [237] 6% најнижих скорова. Разлика је највероватније последица различитих приступа при третирању пиксела – пондерисање или одбацивање.

Сл. 4.10 такође потврђује да се комбиновањем мера сличности у збирну SSIM меру добијају боље перформансе од индивидуалних мера поређења. Истраживањима

из [243] утврђено је да локална поређења структуре кроз меру сличности s имају већи значај од поређења контраста c (закључци су изведени након филтрације и децимације изворних слика). Са Сл. 4.10 се види да је закључак тачан уколико се занемаре SROCC скорови за најниже и највише p вредности.



Сл. 4.10 Корелација рангова (SROCC) у функцији параметра p за SSIM (47), контраст c (45) и структуру s (46) на LIVE подскупу слика за обучавање.

4.3.1.3. Адитивни индекс структурне сличности

Мултипликативна интеграција индивидуалних модела сличности дефинисана у (47) представља константно обележје SSIM и из ње изведених мера. Руководећи се искуством у раду са пробабилистичким системима где једна мала вредност значајно може утицати на глобални скор добијен кроз множење, у раду је анализиран једноставнији адитивни модел за SSIM и GSSIM ((49) и (50)). Робустности модела доприноси и расподела значаја различитих мера сличности у глобалном скору која је линеарна у адитивном моделу, за разлику од нелинеарне везе између SSIM и α , β и γ из (47).

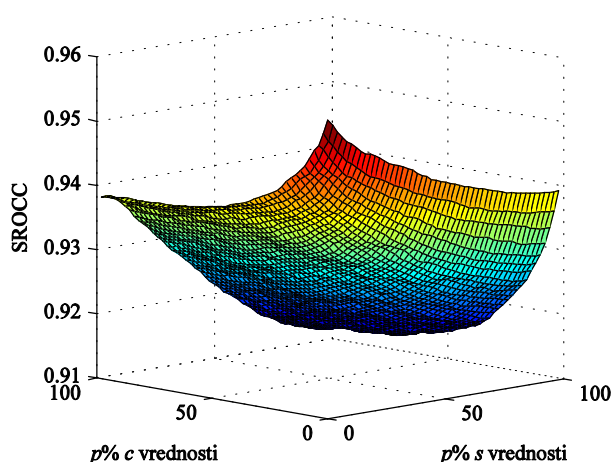
Тако су у раду дефинисани адитивни SSIM (AM-SSIM) и његов еквивалент у градијентном приступу (AM-GSSIM), релације (49) и (50), комбиновани са квалитетом вођеним усредњавањем најнижих скорова квалитета описаним у претходном делу рада:

$$AM-SSIM = \frac{100}{NM} \left(\sum_{k \in c, s} \frac{w_k}{p_k} \sum_{\forall k(n, m) \text{ у најнижих } p_k \%} k(n, m) \right) \quad (49)$$

$$AM-GSSIM = \frac{100}{NM} \left(\sum_{k \in l, c, s} \frac{w_k}{p_k} \sum_{\forall k(n, m) \text{ у најнижих } p_k \%} k(n, m) \right), l, c \text{ и } s \text{ су из [241]} \quad (50)$$

где w_l , w_c и w_s одређују релативни значај осветљаја, контраста и структуре ($w_l + w_c + w_s = 1$). Код SSIM приступа намерно је занемарено поређење осветљаја – луминанси ($w_l = 0$).

Како би се одредиле оптималне вредности параметара p и w , на скупу тренинг слика LIVE базе извршено је обучавање (тренинг) адитивних модела (AM). У процесу обучавања као објективна функција коришћена је корелација рангова (SROCC). На Сл. 4.11 приказан је p_c - p_s исечак 3D оптимизационе површине за AM-SSIM где је $w_c = 0.55$ ($w_s = 0.45$). Са слике су уочава да су оптималне вредности $p_c = 2\%$ и $p_s = 2\%$, за које је SROCC = 0.956. Поређења ради, на истом скупу тренинг слика LIVE базе стандардна SSIM мера има корелацију рангова од SROCC = 0.9107.

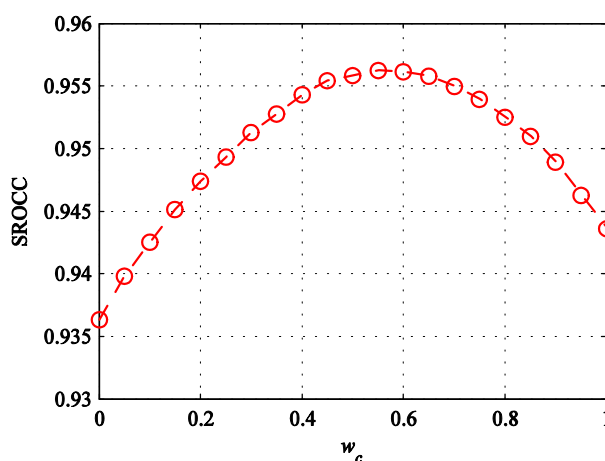


Сл. 4.11 SROCC 3D оптимизациона површина тренинг скупа LIVE базе у функцији SSIM p_c и p_s вредности из (49) ($w_c = 0.55$).

Ниске вредности параметра p и у овом случају обезбеђују најрелевантнија мерења квалитета али се робусне перформансе уочавају у комплетном опсегу p вредности (Сл. 4.11).

Ефекат расподеле значаја између канала сличности SSIM мере илустрован је на Сл. 4.12, где је приказана зависност корелације рангова AM-SSIM модела у функцији w_c ($w_s = 1 - w_c$), где је усвојено да су обе p вредности једнаке 2%. Уочава се да

је мера поређења контраста, c , битнија од мере поређења структуре, s (однос је 0.55 према 0.45), иако прелиминарна истраживања указују да расподела значаја зависи од скале/величине у којој се ради [126]. Еквивалентном процедуром оптимизације за AM-GSSIM добија се скуп оптималних вредности $[p_l, p_c, p_s, w_l, w_c, w_s] = [4, 2, 2, 0.4, 0.4, 0.2]$. У даљем раду се за анализу перформанси AM-SSIM и AM-GSSIM мера користе добијене оптималне вредности значаја канала и процента најнижих скорова квалитета, изведене на половини LIVE базе слика.



Сл. 4.12 Корелација рангова AM-SSIM тренинг скупа у функцији w_c ($w_s=1-w_c$).

4.3.1.4. Перформансе адитивних SSIM мера

Предложене мере структурне сличности AM-SSIM и AM-GSSIM су евалуиране на четири јавно доступне базе слика са субјективним импресијама квалитета: LIVE (подскуп слика који није коришћен за обучавање) [223], CSIQ [226], IVC [225] и Тоуата [224], односно укупно на 1624 тест слике које у неким случајевима значајно одступају од слика коришћених у фази обучавања. Предложене мере су упоређене са PSNR, полазним SSIM и GSSIM мерама (у стандардном облику и са квалитетом вођеном интеграцијом) и са актуелним мерама процене квалитета MS-SSIM [215], адаптивним (препорученим) SSIM [244], VSNR [59] и VIF мерама [58].

Перформансе објективних мера процене евалуиране су кроз три аспекта предикције перцепције квалитета слике, описана у делу рада 2.11. За нелинеарно пресликавање између објективних процена и субјективних скорова користи се интерполациона (логистик) функција са четири параметра [5].

Перформансе објективних мера процене на четири скупа слика дате су у Таб. 4.14, где су најбољи скорови и најбољи скорови SSIM техника (уколико се разликују) подебљани. Базични SSIM скорови су добијени коришћењем оригиналне мере (Матлаб функција `ssim_index.m` са сајта [244]) и они одговарају опису из [48]. Адаптивна (препоручена) верзија SSIM, која се може наћи у [244], представља SSIM технику код које се пре одређивања локалних скорова врши филтрација и децимација изворне и тест слике (функција `ssim.m` са сајта [244]), односно користи оптимална скала у евалуацији. Како је ово валидна верзија SSIM, она је узета у обзир при поређењу перформанси и у раду је обележена са SSIM ZW [244].

Из Таб. 4.14 се уочава да се не може издвојити универзално најбоља мера процене. Квалитетом вођено усредњавање код $SSIM_{\text{најнижих } 2\%}$ и $GSSIM_{\text{најнижих } 14\%}$, унапређују перформансе базичних мера на свим базама, где побољшање иде и до 5%. Перформансе базичних мера су додатно унапређене коришћењем адитивних модела интеграције, AM-SSIM и AM-GSSIM. То указује да је предложена измена додатно повећала робусност структурних објективних евалуација.

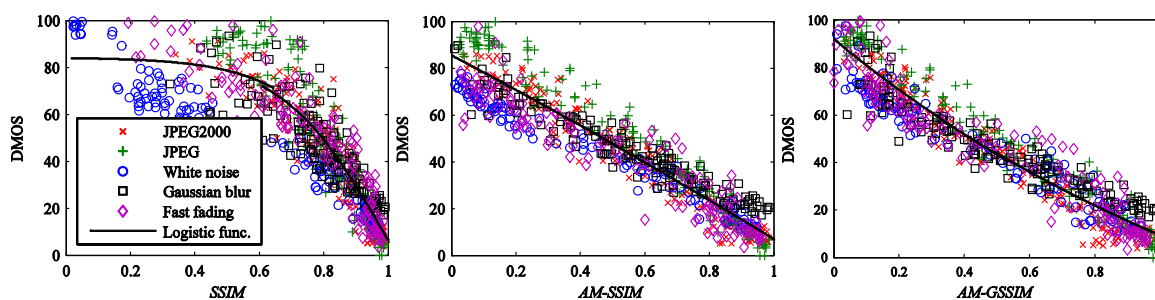
AM-GSSIM модел се боље понаша од осталих мера структурне сличности (најбољи на три од четири базе), бољи је од VIF мере на Тоуама бази и може се рећи да има перформансе које су у рангу перформанси VIF мере. То указује да ивице које се користе у AM-GSSIM обухватају структурне информације боље од интензитета који се користе код AM-SSIM.

На Сл. 4.13 приказане су релације између субјективних (MOS/DMOS) скорова и објективних предикција квалитета за све четири базе слика. Од објективних мера користе се базични SSIM и предложени адитивни модели, AM-SSIM и AM-GSSIM. Адитивни модели праћени одвојеном и квалитетом вођеном интеграцијом значајно побољшавају однос субјективних и објективних скорова. Интерполационе криве свих скупова слика показују да долази до линеаризације у простору субјективно-објективних скорова. Одзив AM мера је доста бољи у поређењу са озбиљним нелинеарностима које постоје код базичног SSIM, при чему се на значајно бољи начин користи расположиви динамички опсег објективног квалитета.

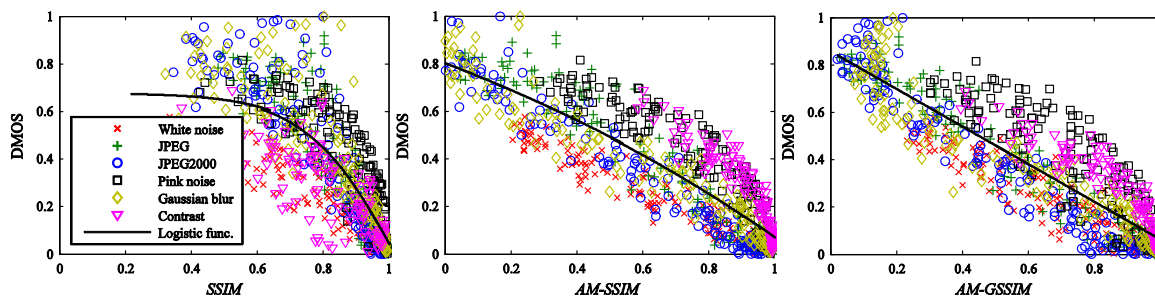
4. Посебни аспекти процене квалитета слика

Таб. 4.14 Перформансе објективних мера процене на LIVE, CSIQ, IVC и Тоуата база слика.

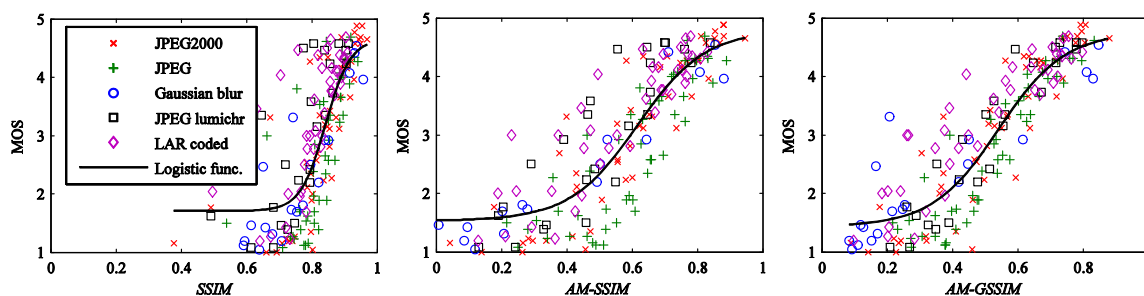
Мера процене	LCC	SROCC	MAE	RMSE	OR (%)
LIVE скуп тест слика					
<i>SSIM</i>	0.9008	0.9107	9.2729	11.8967	7.6543
<i>SSIM</i> _{најнижих 2%}	0.9410	0.9496	7.3143	9.2722	2.9630
<i>GSSIM</i>	0.9305	0.9358	7.6840	10.0348	3.2099
<i>GSSIM</i> _{најнижих 14%}	0.9490	0.9581	6.9916	8.6384	1.2346
<i>SSIM ZW</i> [244]	0.9411	0.9514	7.4006	9.2661	1.4815
<i>MS-SSIM</i>	0.9443	0.9596	7.2996	9.0185	2.7160
<i>AM-SSIM</i>	0.9492	0.9550	6.8169	8.6241	1.7284
<i>AM-GSSIM</i>	0.9547	0.9615	6.5811	8.1560	1.9753
<i>PSNR</i>	0.8784	0.8852	10.1182	13.0942	9.8765
<i>VSNR</i>	0.9265	0.9320	7.9889	10.3109	4.1975
<i>VIF</i>	0.9623	0.9662	6.0976	7.4502	0.2469
CSIQ база слика					
<i>SSIM</i>	0.8151	0.8368	0.1161	0.1521	33.4873
<i>SSIM</i> _{најнижих 2%}	0.8286	0.8378	0.1112	0.1470	32.5635
<i>GSSIM</i>	0.8650	0.8824	0.0973	0.1317	27.7136
<i>GSSIM</i> _{најнижих 14%}	0.8851	0.8895	0.0938	0.1222	27.4827
<i>SSIM ZW</i> [244]	0.8590	0.8749	0.1011	0.1344	29.3303
<i>MS-SSIM</i>	0.8666	0.8774	0.0972	0.1310	27.7136
<i>AM-SSIM</i>	0.8865	0.8781	0.1004	0.1215	29.7921
<i>AM-GSSIM</i>	0.8922	0.8837	0.0953	0.1186	28.0600
<i>PSNR</i>	0.7999	0.8057	0.1195	0.1576	34.2956
<i>VSNR</i>	0.8018	0.8132	0.1152	0.1569	30.1386
<i>VIF</i>	0.9252	0.9194	0.0753	0.0996	22.7483
IVC база слика					
<i>SSIM</i>	0.7924	0.7789	0.5547	0.7431	17.8378
<i>SSIM</i> _{најнижих 2%}	0.8226	0.8161	0.5197	0.6928	15.6757
<i>GSSIM</i>	0.8826	0.8686	0.4180	0.5727	8.1081
<i>GSSIM</i> _{најнижих 14%}	0.8829	0.8757	0.4415	0.5720	8.6486
<i>SSIM ZW</i> [244]	0.9119	0.9021	0.3774	0.5001	5.4054
<i>MS-SSIM</i>	0.9108	0.8971	0.3818	0.5031	5.9459
<i>AM-SSIM</i>	0.8622	0.8524	0.4674	0.6171	10.8108
<i>AM-GSSIM</i>	0.9024	0.8937	0.3986	0.5250	7.5676
<i>PSNR</i>	0.7192	0.6884	0.6688	0.8465	20.5405
<i>VSNR</i>	0.8027	0.7993	0.5585	0.7265	15.6757
<i>VIF</i>	0.9026	0.8964	0.4096	0.5244	7.5676
Тоуата база слика					
<i>SSIM</i>	0.7979	0.7870	0.5891	0.7545	14.2857
<i>SSIM</i> _{најнижих 2%}	0.8472	0.8391	0.5214	0.6648	12.5000
<i>GSSIM</i>	0.8976	0.8937	0.4353	0.5517	5.9524
<i>GSSIM</i> _{најнижих 14%}	0.8926	0.8870	0.4293	0.5642	7.1429
<i>SSIM ZW</i> [244]	0.8893	0.8822	0.4412	0.5724	8.3333
<i>MS-SSIM</i>	0.8948	0.8910	0.4305	0.5588	7.1429
<i>AM-SSIM</i>	0.8627	0.8531	0.4959	0.6328	10.7143
<i>AM-GSSIM</i>	0.9169	0.9078	0.3819	0.4995	7.1429
<i>PSNR</i>	0.6356	0.6132	0.7833	0.9662	22.0238
<i>VSNR</i>	0.8705	0.8608	0.4651	0.6160	9.5238
<i>VIF</i>	0.9137	0.9077	0.4032	0.5087	5.3571



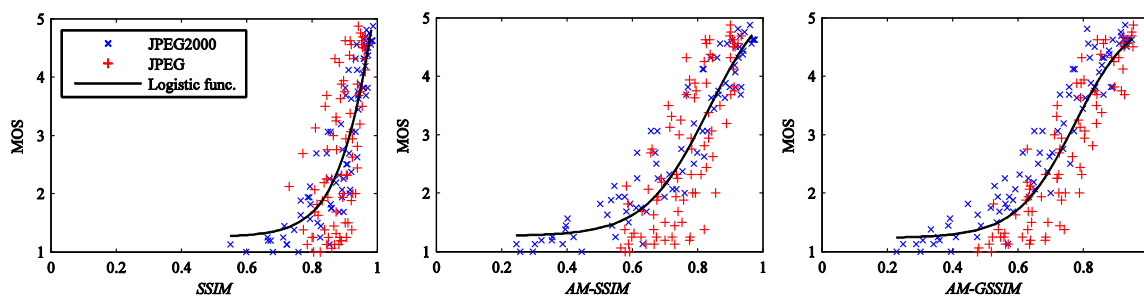
(a) LIVE база слика (тренинг и тест слике)



(б) CSIQ база слика



(в) IVC база слика



(г) Тоуата база слика

Сл. 4.13 Расипање субјективних DMOS/MOS и објективних SSIM/AM-SSIM/AM-GSSIM скорова квалитета. Свака тачка одговара једној тест слици. Интерполационе криве су дате пуном линијом.

4.3.1.5. Закључак

У раду је представљен нови поглед на примену индекса структурне сличности у процени квалитета слика. Нови приступ је базиран на робусном и стабилном

адитивном моделу интеграције и делотворном и сепарабилном квалитетом вођеном просторном интеграцијом локалних скорова квалитета. Новоформулисани модел је примењен на стандардни, интензитетски SSIM и градијентни SSIM, чиме су добијене AM-SSIM и AM-GSSIM мере. Оптимално слагање нових мера квалитета са субјективним импресијама квалитета добија се уколико се за формирање глобалног скорa користи део најнижих објективних скорова квалитета. Овим је подржана теорија да посматрачи одлуку о квалитету доносе на основу региона посматраног сигнала са најлошијим квалитетом [237]. Додатно је показано да очување контраста и структуре у концепту структурне сличности има скоро подједнак значај.

Експериментални резултати на четири јавно доступне базе слика показују да перформансе AM-GSSIM превазилазе остале SSIM базиране приступе, укључујући и сложену формулацију као што је више-резулациона MS-SSIM. Коришћењем адитивног приступа сједињавања информација и интеграције локалних скорова на основу вредности квалитета перформансе SSIM могу достићи (па и превазићи) перформансе много софистициранијих објективних мера као што је VIF [58].

Иако је исцрпном евалуацијом показано да предложени приступ достиже висок степен слагања са субјективним скоровима на широком асортиману слика, очекује се да мултивеличинска анализа и прави избор резолуције у којој се ради може додатно допринети побољшању перформанси, а што ће бити предмет даљих истраживања.

4.3.2. Примена адитивног модела сједињавања информација код објективне процене засноване на очувању градијента

Како је адитивни модел очувања структурних информација довео до побољшања у слагању субјективних и SSIM објективних скорова квалитета, његова примена је испитана и на објективној процени заснованој на очувању информација о градијенту, Q^{AB} . Објективна процена квалитета Q^{AB} користи два канала очувања информација – очување амплитуда градијента и очување оријентација. Очување амплитуда је еквивалентно очувању контраста код SSIM, док је очување оријентација еквивалентно очувању структуре код SSIM. Због тога је при

дефинисању адитивног ивичног/градијентног модела пресликан приступ који је коришћен код адитивног структурног приступа.

Предложена ивична објективна процена квалитета слика (33) користи неколико параметара од којих зависи слагање субјективних и објективних скорова. То су коефицијенти нелинеарних сигмоидних функција – k и σ у релацији (31). У раду су за ове коефицијенте преузете вредности које су добијене у истраживањима за потребе сједињавања слика [214, 221] и за које је утврђено да су субјективно значајне, $[k_g, \sigma_g, k_\alpha, \sigma_\alpha] = [-11, 0.7, -24, 0.8]$.

Сигмоидне функције имају за циљ да моделују одзив биолошких система који је нелинеаран. Због коришћења нелинеарних сигмоидних функција мера процене квалитета слика (33) се може назвати нелинеарном. Алтернативни приступ је приступ у коме се анализира промена сигнала и третира као вредност квалитета. Коришћењем овог, линеарног приступа, избегавају се коефицијенти нелинеарних сигмоидних функција (којих има укупно четири), и вредност мере квалитета слике се може дефинисати као:

$$\Delta^{AB} = \sqrt{\Delta_g^{AB} \cdot \Delta_\alpha^{AB}} \quad (51)$$

где Δ_g^{AB} и Δ_α^{AB} говоре о збирној промени сигнала – промени амплитуда градијента (29) и промени оријентације градијента (30):

$$\Delta_g^{AB} = \frac{1}{NM} \sum_{\forall n,m} \Delta_g^{AB}(n,m) \quad (52)$$

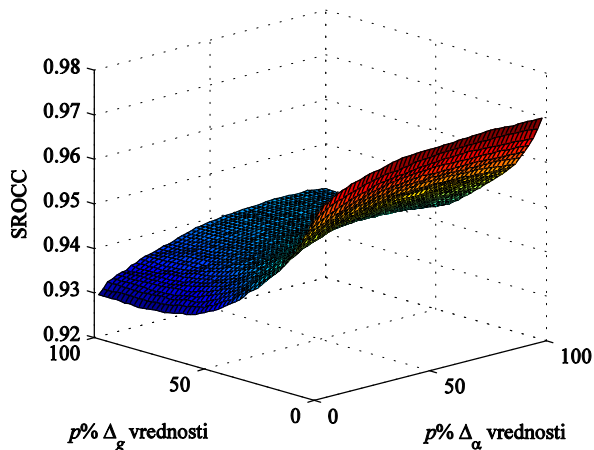
$$\Delta_\alpha^{AB} = \frac{1}{NM} \sum_{\forall n,m} \Delta_\alpha^{AB}(n,m) \quad (53)$$

Адитивни модел очувања информација је додатно комбинован са квалитетом вођеном интеграцијом индивидуалних канала (усредњавање најнижих скорова квалитета):

$$AM-\Delta^{AB} = \frac{100}{NM} \left(\sum_{k \in g, \alpha} \frac{w_k}{p_k} \sum_{\forall \Delta_k^{AB}(n,m) \text{ у најнижих } p_k \%} \Delta_k^{AB}(n,m) \right) \quad (54)$$

где је за адитивни приступ коришћен линеарни модел очувања информација о градијенту. То је урађено како би се избегла мултидимензиона оптимизација која би обухватила и четири параметра сигмоидних функција. У оптимизацији је тако потребно одредити вредности три параметра: w_g ($w_\alpha=1-w_g$), p_g и p_α . Параметар w_g одређује релативни значај очувања амплитуда градијента док параметри p_g и p_α говоре о томе колики проценат најнижих скорова из индивидуалних канала ће се узети у разматрање за глобалну процену квалитета.

Како би се одредиле оптималне вредности параметара w_g , p_g и p_α , на скупу тренинг слика LIVE базе извршено је обучавање (тренинг) адитивног модела (AM- Δ^{AB}). У процесу обучавања као објективна функција коришћена је корелација рангова (SROCC) а скуп тренинг слика је исти као у процесу обучавања SSIM адитивних мера. На Сл. 4.14 приказан је p_g - p_α исечак 3D оптимизационе површине за AM- Δ^{AB} где је $w_g=0.7$. Са слике су уочава да су оптималне вредности $p_g=2\%$ и $p_\alpha=78\%$.

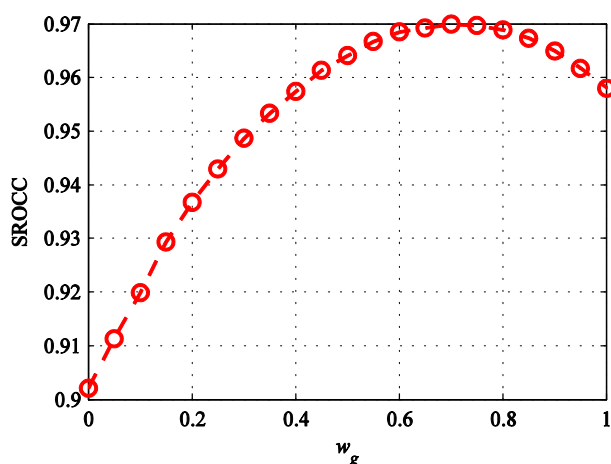


Сл. 4.14 SROCC 3D оптимизациона површина тренинг скупа LIVE базе у функцији p_g и p_α из (54) ($w_g=0.7$).

Ниске вредности параметра p_g и у овом случају обезбеђују релевантна мерења квалитета али се робусне перформансе уочавају у комплетном опсегу p_α вредности (Сл. 4.14). Узимање најнижих скорова квалитета много више утиче на очување информација о амплитуди градијента него на очување информација о оријентацији. Поређењем SROCC оптимизационих површина са Сл. 4.11 и Сл. 4.14 уочавају се

различити трендови корелације рангова у функцији најнижих скорова структурних информација/оријентација. Са повећањем p_α корелација рангова SSIM опада, док код ивичне процене благо расте након чега благо пада.

Ефекат расподеле значаја између канала сличности Δ^{AB} мере илустрован је на Сл. 4.15, где је приказана зависност корелације рангова AM- Δ^{AB} модела у функцији w_g , где је усвојено да су обе p вредности оптималне ($p_g=2\%$, $p_\alpha=78\%$). Уочава се да је мера поређења амплитуда ивица, Δ_g^{AB} , битнија од мере поређења оријентација, Δ_α^{AB} (однос је 0.7 према 0.3). Ово запажање потврђује однос значаја који је оптималан код адитивног SSIM мере (0.55 према 0.45).



Сл. 4.15 Корелација рангова AM- Δ^{AB} тренинг скупа у функцији w_g ($w_\alpha=1-w_g$).

Предложени адитивни AM- Δ^{AB} модел је евалуиран на истом скупу слика као и адитивни SSIM модели (при чему су усвојене оптималне вредности процента најнижих скорова квалитета и значаја канала добијене на тренинг скупу LIVE базе) – четири јавно доступне базе слика са субјективним импресијама квалитета. Осим адитивног модела презентоване су перформансе MS-SSIM и VIF мера, перформансе нелинеарне Q^{AB} мере и њених модификација добијених увођењем мапа значаја и презентоване су перформансе индивидуалних канала очувања информација о градијенту и мултипликативног модела. Перформансе објективних мера евалуиране су кроз тачност, монотоност и конзистентност предикције, а у пресликавању између објективних процена и субјективних скорова користи се логистик функција са четири параметра [5].

Перформансе објективних мера процене на четири скупа слика дате су у Таб. 4.15-Таб. 4.18, где су најбољи скорови подебљани. Адитивни ивични модел се

4. Посебни аспекти процене квалитета слика

најбоље понаша на LIVE бази слика где коефицијент корелације достиже 97%. На овој бази перформансе адитивног модела су боље од перформанси до сада најуспешније VIF мере. Добитак адитивног модела у односу на мултипликативни (линеарни и нелинеарни) је око 3%. Осим на LIVE бази, адитивни модел сједињавања информација је у односу на стандарни Q^{AB} приступ дао боље резултате на CSIQ и IVC базама слика.

Таб. 4.15 Поређење перформанси мера процене квалитета на LIVE бази слика (подскуп слика за тестирање).

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.9443	0.9596	7.2996	9.0185	2.7160
VIF	0.9623	0.9662	6.0976	7.4502	0.2469
Q^{AB}	0.9405	0.9418	7.5332	9.3083	2.9630
Q_{w1}^{AB}	0.9544	0.9570	6.7465	8.1752	1.7284
Q_{w2}^{AB}	0.9490	0.9523	7.0749	8.6398	1.9753
Δ_g^{AB}	0.9190	0.9235	8.3077	10.8030	4.6914
Δ_α^{AB}	0.9235	0.9150	8.4850	10.5109	3.2099
Δ^{AB}	0.9421	0.9447	7.3220	9.1905	2.4691
AM- Δ^{AB}	0.9692	0.9709	5.4455	6.7419	0.2469

Таб. 4.16 Поређење перформанси мера процене квалитета на CSIQ бази слика.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.8666	0.8774	0.0972	0.1310	27.7136
VIF	0.9252	0.9194	0.0753	0.0996	22.7483
Q^{AB}	0.8556	0.8520	0.1039	0.1359	31.1778
Q_{w1}^{AB}	0.8794	0.8754	0.0937	0.1250	28.6374
Q_{w2}^{AB}	0.8769	0.8741	0.0955	0.1262	29.5612
Δ_g^{AB}	0.8459	0.8690	0.1052	0.1400	30.1386
Δ_α^{AB}	0.7792	0.7147	0.1332	0.1646	40.9931
Δ^{AB}	0.8606	0.8593	0.1019	0.1337	29.6767
AM- Δ^{AB}	0.8847	0.8616	0.0986	0.1224	29.9076

Таб. 4.17 Поређење перформанси објективних мера процене квалитета на IVC бази слика.

Мера	LCC	SROCC	MAE	RMSE	OR (%)
MS-SSIM	0.9108	0.8971	0.3818	0.5031	5.9459
VIF	0.9026	0.8964	0.4096	0.5244	7.5676
Q^{AB}	0.8813	0.8699	0.4444	0.5757	5.4054
Q_{w1}^{AB}	0.8846	0.8759	0.4392	0.5682	5.9459
Q_{w2}^{AB}	0.8864	0.8766	0.4346	0.5640	5.9459
Δ_g^{AB}	0.8806	0.8720	0.4247	0.5773	7.5676
Δ_α^{AB}	0.8320	0.8253	0.5198	0.6758	11.3514
Δ^{AB}	0.8997	0.8895	0.4065	0.5318	5.4054
AM- Δ^{AB}	0.8901	0.8841	0.4228	0.5553	7.0270

4. Посебни аспекти процене квалитета слика

Таб. 4.18 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.

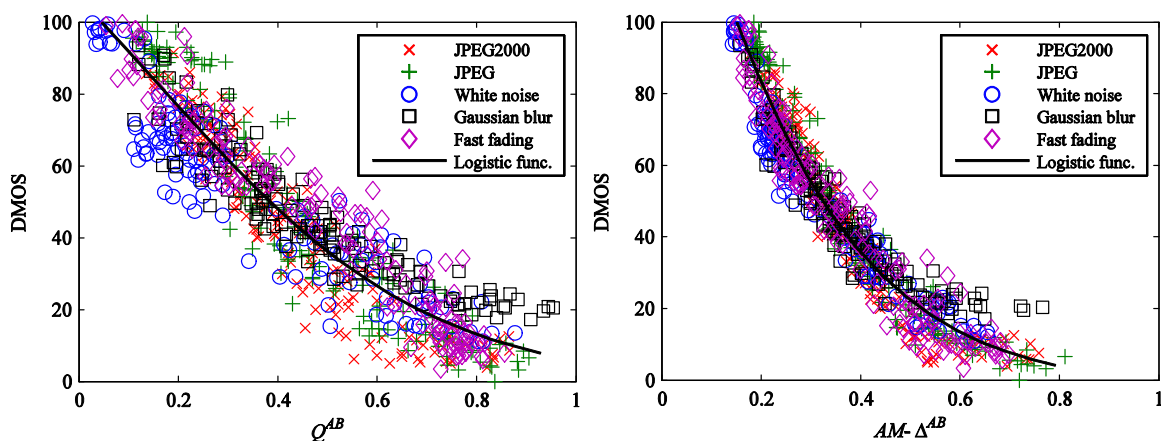
<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>MS-SSIM</i>	0.8948	0.8910	0.4305	0.5588	7.1429
<i>VIF</i>	0.9137	0.9077	0.4032	0.5087	5.3571
Q^{AB}	0.9322	0.9276	0.3482	0.4529	2.9762
Q_{w1}^{AB}	0.9352	0.9289	0.3335	0.4433	2.9762
Q_{w2}^{AB}	0.9341	0.9271	0.3355	0.4469	2.9762
Δ_g^{AB}	0.8921	0.8912	0.4408	0.5656	6.5476
Δ_α^{AB}	0.8427	0.8437	0.5300	0.6738	8.9286
Δ^{AB}	0.9462	0.9431	0.3099	0.4051	2.9762
$AM-\Delta^{AB}$	0.9115	0.9099	0.3994	0.5147	5.3571

Визуелна илустрација односа субјективних и објективних скорова квалитета дата је на Сл. 4.16 кроз дијаграме расипања у простору субјективно-објективних скорова. Представљени су резултати нелинеарног Q^{AB} и адитивног $AM-\Delta^{AB}$ приступа сједињавања информација. Код LIVE и CSIQ база слика уочава се знатно мање расипање око логистик криве уколико се користи адитивни приступ сједињавања информација. Додатно се за CSIQ субјективно-објективни простор расипања може уочити груписање скорова по класама слика. Коришћење дела најнижих локалних скорова квалитета додатно је на Тоуата бази слика сузило динамички опсег објективних скорова, тако да је у односу на стандардни приступ дошло до повећања стрмине логистик криве у простору субјективно-објективних скорова. Ефекат преоптерећења стрмине је вероватно разлог због кога је дошло до погоршања перформанси адитивног ивичног у односу на стандардни приступ (са 0.93 на 0.91). Међутим, и поред тога перформансе се могу сматрати задовољавајућим јер су боље од перформанси MS-SSIM и у нивоу перформанси VIF мере. Сужавање динамичког опсега и повећање нагиба логистик криве уочава се и код IVC базе слика.

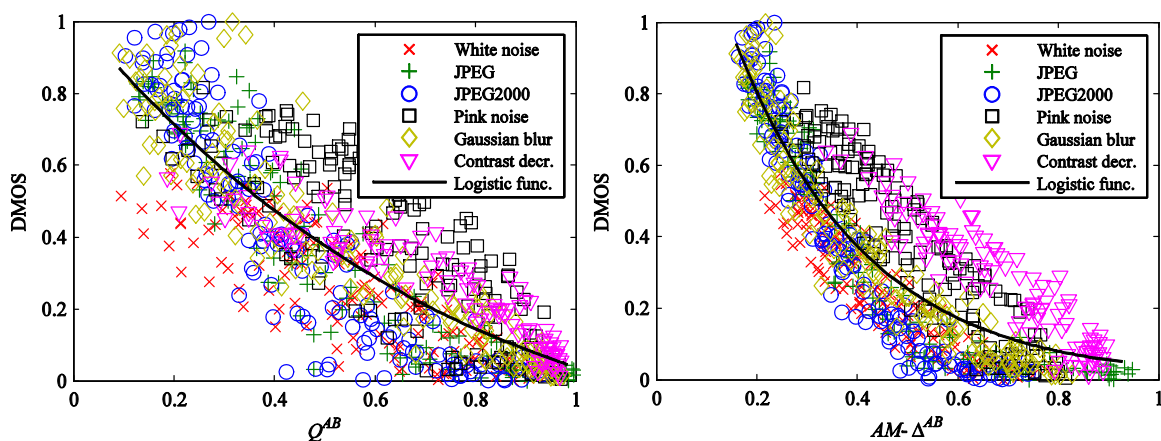
Ради лакшег поређења различитих мера очувања информација о градијенту, на Сл. 4.17 приказане су њихове перформансе на четири анализиране базе. Као критеријуми поређења користе се коефицијент линеарне корелације, LCC, и корелација рангова, SROCC, субјективних и објективних скорова. Поређењем перформанси индивидуалних канала очувања информација о градијенту уочава се да је предност на страни мере очувања информација о амплитуди градијента. Осим LIVE базе где су перформансе оба канала приближно исте, на преостале три базе слика постоји значајна разлика у перформансама. Сједињавањем информација о

4. Посебни аспекти процене квалитета слика

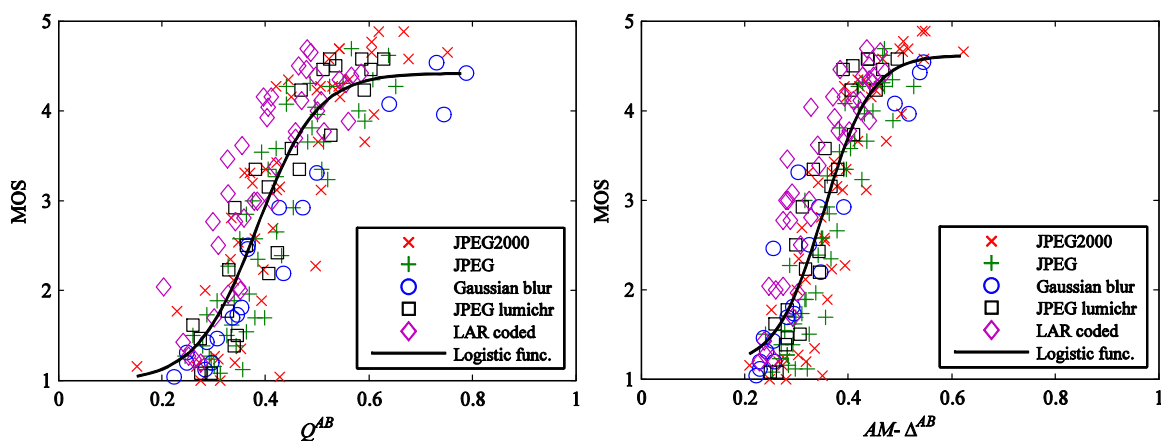
очувању ивичних информација, било на адитивни или мултипликативни начин, доводи до побољшања у предикцији субјективних импресија квалитета. Збирне мере које не користе сигмоидне функције (Δ^{AB} и $AM-\Delta^{AB}$) у потпуности могу заменити сложенији, нелинеарни Q^{AB} модел. Алтернативно, може се користити само канал очувања информација о амплитуди градијента.



(a) LIVE база слика (тренинг и тест слике)

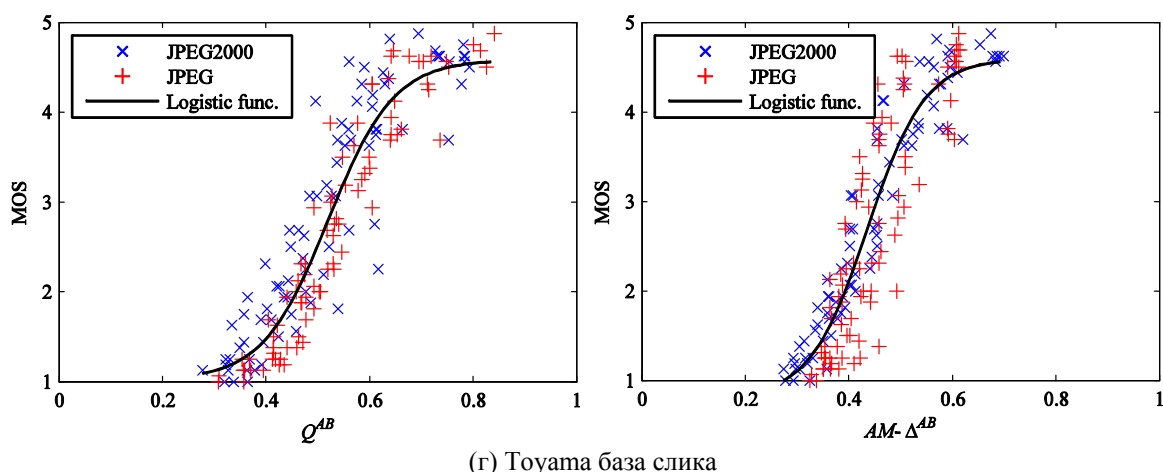


(б) CSIQ база слика

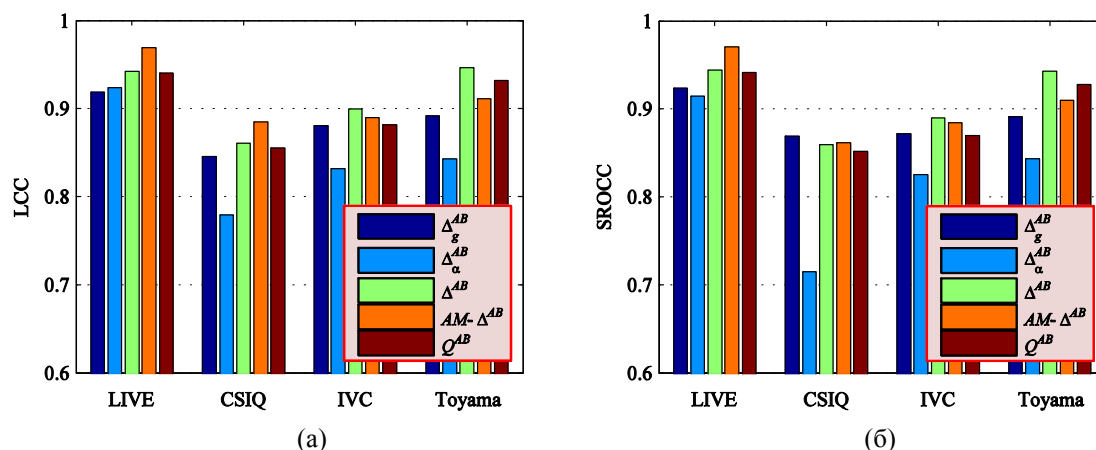


(в) IVC база слика

4. Посебни аспекти процене квалитета слика



Сл. 4.16 Расипање субјективних DMOS/MOS и објективних (Q^{AB} и $AM-\Delta^{AB}$) скорова квалитета.



Сл. 4.17 Поређење перформанси објективних градијентних процена на четири базе слика: (а) коефицијент линеарне корелације (LCC) субјективних и објективних скорова и (б) корелација рангова (SROCC) субјективних и објективних скорова.

Оптимизациони дијаграм адитивног приступа у очувању ивичних информација са Сл. 4.14 показује благо повећање перформанси са удаљавањем од координатног почетка и повећањем процента најнижих скорова очувања информација о оријентацији градијента. Перформансе након тога благо опадају како се у процени користи више локалних процена очувања информација о оријентацији. Због тога је у раду истражен адитивни модел код кога се комбинују најнижи и највиши скорови квалитета из оба канала. Оптимизациони поступак је додатно спроведен и за две преостале комбинације најнижих и највиших скорова квалитета. У процесу обучавања коришћен је LIVE подскуп слика који је већ коришћен у

4. Посебни аспекти процене квалитета слика

фазама обучавања. Оптималне вредности параметара адитивног модела дате су у Таб. 4.19, а као квантитативна мера коришћена је корелација рангова, SROCC.

Таб. 4.19 Оптималне вредности параметара адитивног модела очувања информација о градијенту.

		Δ_g^{AB}	
		најлошији	најбољи
Δ_α^{AB}	најлошији	$p_g=2\%$ $p_\alpha=78\%$ $w_g=0.7$ SROCC=0.9699	$p_g=100\%$ $p_\alpha=96\%$ $w_g=0.6$ SROCC=0.9410
	најбољи	$p_g=2\%$ $p_\alpha=84\%$ $w_g=0.65$ SROCC=0.9703	$p_g=100\%$ $p_\alpha=52\%$ $w_g=0.35$ SROCC=0.9444

Поређењем вредности SROCC из Таб. 4.19 уочава се да је у збирној процени квалитета најбоље узимати део најлошијих скорова очувања информација о амплитуди градијента. При томе је сасвим свеједно да ли ће се ти скорови комбиновати са најбољим или најлошијим скоровима очувања информација о оријентацији ивица. У збирној мери боље је већи значај дати каналу очувања информација о амплитуди градијента (0.65 односно 0.7). Адитивни модел комбиновања најлошијих скорова очувања амплитуда и најбољих скорова очувања оријентација (са оптималним параметрима из Таб. 4.19) је тестиран на четири доступне базе слика. Његове перформансе су упоређене са адитивним моделом који користи најлошије скорове из оба канала и дате су у Таб. 4.20–Таб. 4.23. Уочава се еквиваленција перформанси једног и другог начина комбиновања (најбољи скорови су подебљани). Овакав резултат могао се и наслутити због облика 3D оптимизационе површине са Сл. 4.14.

Таб. 4.20 Поређење перформанси мера процене квалитета на LIVE бази слика (подскуп слика за тестирање).

Мера	LCC	SROCC	MAE	RMSE	OR (%)
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_\alpha=78\%$ (лоших), $w_g=0.70$)	0.9692	0.9709	5.4455	6.7419	0.2469
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_\alpha=84\%$ (добрих), $w_g=0.65$)	0.9705	0.9722	5.3118	6.6080	0.2469

Таб. 4.21 Поређење перформанси објективних мера процене квалитета на CSIQ бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR</i>
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=78\%$ (лоших), $w_g=0.70$)	0.8847	0.8616	0.0986	0.1224	29.9076
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=84\%$ (добрих), $w_g=0.65$)	0.8891	0.8686	0.0970	0.1201	29.5612

Таб. 4.22 Поређење перформанси објективних мера процене квалитета на IVC бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=78\%$ (лоших), $w_g=0.70$)	0.8901	0.8841	0.4228	0.5553	7.0270
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=84\%$ (добрих), $w_g=0.65$)	0.8886	0.8840	0.4270	0.5588	6.4865

Таб. 4.23 Поређење перформанси објективних мера процене квалитета на Тоуата бази слика.

<i>Мера</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=78\%$ (лоших), $w_g=0.70$)	0.9115	0.9099	0.3994	0.5147	5.3571
$AM-\Delta^{AB}$ ($p_g=2\%$, $p_a=84\%$ (добрих), $w_g=0.65$)	0.9060	0.9044	0.4116	0.5296	5.9524

4.4. Закључак

У овом поглављу су анализирана три аспекта процене квалитета слике. Анализом трансфера слике са становишта теорије информација остварен је детаљнији увид у проблем преноса слике. Различити аспекти трансфера информација анализирани су коришћењем мере очувања информација о градијенту. Показано је да је приликом евалуације квалитета слике потребно користити више различитих аспеката који одговарају сложеним механизмима перцепције квалитета од стране посматрача. Приступ је илустрован кроз евалуацију бројних слика са различитим типовима и нивоима дисторзије. Од спроведене анализе очекује се да буде од великог значаја у развоју успешних мера за процену квалитета слике на глобалном плану.

У другом делу поглавља предложен је начин формирања мапа субјективног значаја. Мапе визуелног значаја изведене су из вероватноћа појављивања амплитуда градијента изворне и тест слике. Интеграцијом предложених мапа значаја са градијентном објективном проценом побољшавају се њене полазне перформансе. Добитак зависи од базе над којом се ради а утицај предложених мапа на подкуповима слика из различитих база није једнозначан.

У трећем делу поглавља представљени су адитивни модели сједињавања информација из индивидуалних канала евалуације квалитета. Адитивни модели комбиновани су са квалитетом вођеном интеграцијом. Показано је да се коришћењем адитивних SSIM мера побољшавају перформансе базичних SSIM и GSSIM мера. Добијени резултати у потпуности потврђују хипотезу да посматрачи своју одлуку о квалитету доносе на основу региона тест слике са најлошијим квалитетом.

Адитивни модел сједињавања информација о параметрима градијента је предложен као алтернатива стандардном, нелинеарном Q^{AB} моделу. У оквиру овог адитивног модела канал очувања информација о амплитуди градијента такође потврђује хипотезу да су најнижи скорови најбитнији у процени квалитета. Други канал, тј. канал очувања информација о оријентацији ивица се не уклапа у потпуности у наведену хипотезу. То може да буде због тога што се оријентација ивица рачуна применом Собелове маске димензија 3x3 пиксела и утицаја шума, који је израженији него код SSIM где се локалне процене квалитета спроводе унутар региона 11x11 пиксела.

Добијени резултати указују да је у процени квалитета слике и извођењу битних закључака веома тешко извести генералне (универзалне) закључке због бројних аспеката као што су: тест узорак, типови деградација, динамички опсег квалитета, објективна мера и њене појединачне компоненте, параметри објективних мера, субјективни тестови, итд. Тако је на неким базама слика битније поређење контраста док је на другим битније поређење структуре. Адитивни модел на неким базама је показао значајно побољшање, док је на неким остао у рангу стандардног приступа. Све то указује да је процена квалитета слике веома сложена.

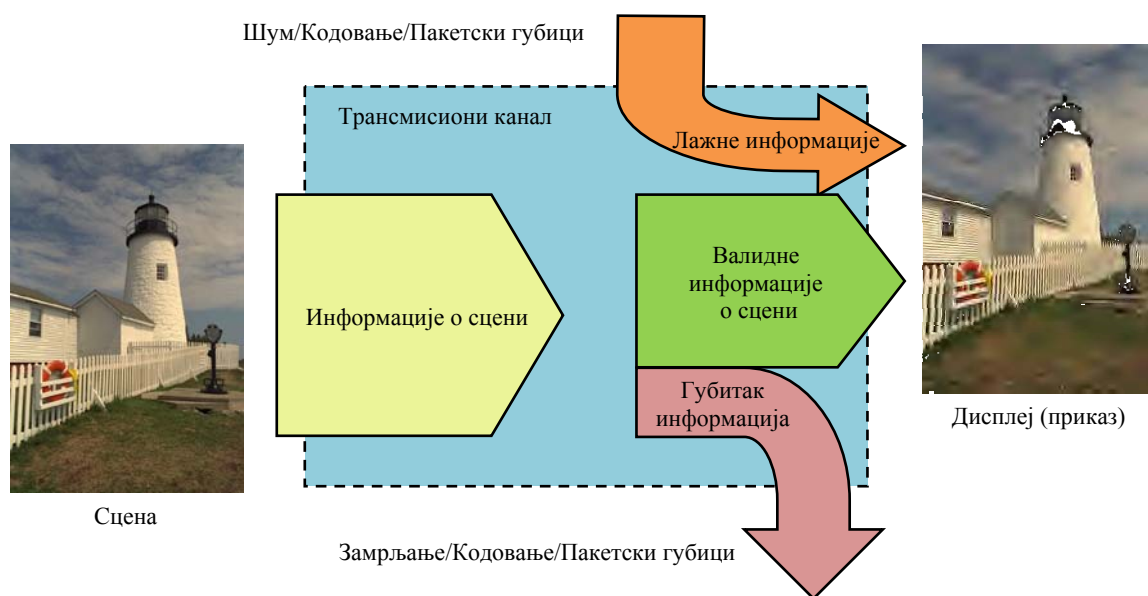
5. Процена квалитета видеа заснована на очувању информација о градијенту

5.1. Увод

Дигитални видео подаци који се преносе кроз комуникациони канал ограниченог капацитета изложени су различитим врстама дисторзије током аквизиције, компресије, обраде, преноса и репродукције. Тако се техникама компресије са губицима деградира квалитет како би се прилагодило капацитету канала за пренос. Такође, пакетски губици током преноса додатно нарушавају квалитет видеа. Због тога је од великог значаја уколико систем за пренос података има могућност да квантитативно опише деградацију која се јавља, са циљем да се контролише (прати) и уколико је могуће побољша квалитет видео сигнала. Робусна мера процене квалитета видеа је веома значајна за ову сврху.

Оваква мера се може користити за поређење различитих алгоритама компресије и обраде сигнала, оптимизацију перформанси алгоритама компресије и обраде под било којим ограничењима (капацитет канала, пакетски губици, однос сигнал шум, ...), поређење различитих система преноса и сл. [212]. Међутим одређивање квалитета није једноставан задатак. На Сл. 5.1 је илустрована поједностављена шема трансфера података разматрана у оквиру процене квалитета видеа.

Представа сцене је дата у форми дигиталне слике која се добија на излазу сензора. Секвенца слика се компримује и преноси кроз трансмисиони канал на пријемну страну, где се декомпримује и приказује посматрачима. Компресија и губици који се јављају током преноса деградирају стварне информације о сцени. Додатно, процеси током преноса могу дати карактеристичне лажне информације у примљеној видео секвенци, а који се могу третирати као стварне информације. Анализа процеса трансфера информација и одређивање квалитета реалног видеа на страни посматрача су предмет разматрања објективне мере предложене у овом поглављу.



Сл. 5.1 Пренос видео са становишта очувања и процене квалитета.

Циљ процеса преноса информација илустрованог кроз претходну слику јесте верна репрезентација визуелних информација из оригиналне секвенце (на страни сензора) у деградираној – приказаној секвенци (на страни посматрача). Ова формулација ће бити прихваћена у евалуацији квалитета видео.

Како би се мерило очување изворних информација у деградираној секвенци користи се потпуно референцирање, тј. информације изворне секвенце се пореде са информацијама деградиране секвенце. Сл. 5.2 илуструје овај приступ са више детаља.

Циљ приступа са Сл. 5.2 је издвајање и репрезентација различитих извора визуелних информација у изворној и деградираној секвенци. Због тога се врши разлагање видео сигнала као извора информација на:

1. просторне информације, којима одговарају структуре на сцени, облици објеката, просторне везе објеката, ...
2. временске информације, којима одговара кретање објеката, оптички ток, промене у структурама и облицима током времена, и
3. хроматске информације – боја и спектрална рефлектанса објеката на сцени.

Информације се издвајају из оригиналне и деградиране секвенце и пореде користећи специјализоване моделе очувања информација. Моделима се евалуира верност репрезентације (губитак) изворних визуелних информација у деградираној секвенци. Очување информација се анализира просторно и временски и након тога комбинује у једну вредност која представља квалитет приказаног видео.

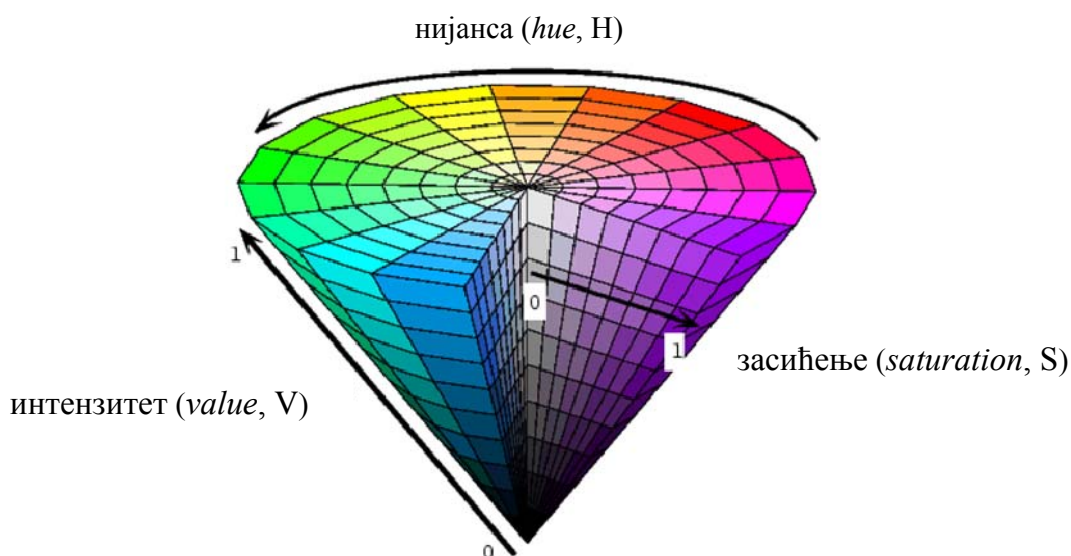


Сл. 5.2 Евалуација квалитета деградиране секвенце.

5.2. Разлагање визуелних информација

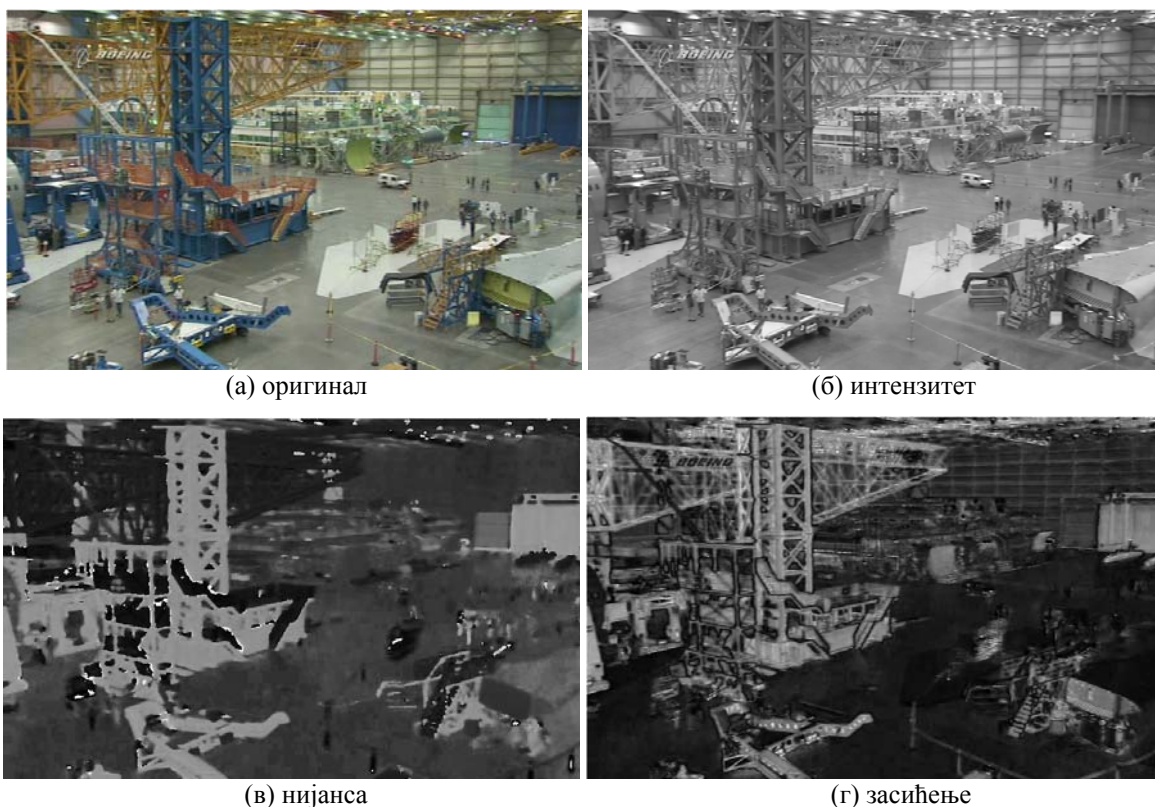
Како би се извршило разлагање извора информација, стандардни RGB видео сигнал се конвертује у погоднији колор систем. Постоје бројни колор координатни системи али се у овој дисертацији због једноставне имплементације и разлагања боје

у одговарајуће компоненте користи HSV (Hue-Saturation-Value) колор систем. Овај колор систем, илустрован на Сл. 5.3, раздваја оригинални RGB сигнал у три канала – компоненте: (1) *hue* (нијанса или врста боје) – који презентује таласну дужину (основну боју) сигнала на скали од црвене, преко зелене, плаве, до поново црвене, (2) *saturation* (засићење или чистоћа боја) – који говори о доминантности *hue* канала у сигналу, од сиве до пуне боје и (3) *value* или интензитет који говори о осветљају сигнала, или интензитету конвенционалног монохроматског сигнала [245].



Сл. 5.3 Илустрација HSV колор простора.

Трансформација из RGB у HSV колор систем је једноставна али није тривијална [246]. Коришћењем трансформације RGB видео секвенца се кадар по кадар конвертује у секвенцу HSV вредности. Хроматске информације се налазе у H и S (*hue* и *saturation*) равнима, док се информације о интензитету налазе у V (*value/intensity*) равни. Пример трансформације је дат на Сл. 5.4, где се кадар из RGB секвенце конвертује у HSV колор простор. Канал интензитета приказује просторне информације и структуре на сцени боље од остала два канала. Канал нијансе (H) приказује да постоји више примарних боја, док се из канала засићења види да боја није униформно расподељена на сцени.



Сл. 5.4 Трансформација RGB слике у HSV колор простор.

5.3. Очување просторних информација

Евалуација очувања просторних информација је спроведена кроз приступ који је успешно примењен у евалуацији квалитета сједињавања монохроматских слика, тј. кроз репрезентацију градијента [179, 214]. Евалуација очувања градијента је заснована на идеји да је *систем преноса који пренесе више информација о изворном градијенту бољи*. Просторне информације се издвајају из канала интензитета изворних и деградираних секвенци, A и B , у форми амплитуде градијента g и оријентације α (релације (27) и (28) у делу рада 3.2). Параметри градијента се извајају на свакој позицији (n,m) кадра током времена (t) , користећи Собелов оператор. Параметри градијента се користе како би се дефинисале релативне промене – фактори G и A , између оригиналне и деградираних секвенци (Сл. 3.1):

$$G_{n,m,t}^{AB} = \exp \left(- \left| \log \left(\frac{g_{n,m,t}^A}{g_{n,m,t}^B} \right) \right| \right) \quad (55)$$

$$A_{n,m,t}^{AB} = \left| \frac{|\alpha_{n,m,t}^A - \alpha_{n,m,t}^B|}{\pi} - 1 \right| \quad (56)$$

Оба параметра, (55) и (56), се налазе у опсегу $[0,1]$, где 0 одговара потпуном губитку информација док 1 говори о идеалном преносу. Параметар очувања амплитуде градијента G , је у опсегу $[0,1]$ независно од тога где је амплитуда градијента већа – у изворној или деградираној секвенци. Било каква промена амплитуде градијента (позитивна или негативна) се третира као дисторзија која се јавља током процеса преноса и даје вредности које су мање од 1.

Мера очувања просторних информација, Q_s , моделује субјективни губитак информација користећи нелинеарне сигмоидне функције и дефинисана је константама Γ , k_g , σ_g , k_α , и σ_α :

$$Q_{s_{n,m,t}} = \sqrt{Q_g \cdot Q_\alpha} = \sqrt{\frac{\Gamma_g}{1 + e^{k_g(G_{n,m,t}^{AB} - \sigma_g)}}} \cdot \sqrt{\frac{\Gamma_\alpha}{1 + e^{k_\alpha(A_{n,m,t}^{AB} - \sigma_\alpha)}}} \quad (57)$$

односно:

$$Q_{s_{n,m,t}} = \frac{\Gamma_s}{\sqrt{\left(1 + e^{k_g(G_{n,m,t}^{AB} - \sigma_g)}\right) \left(1 + e^{k_\alpha(A_{n,m,t}^{AB} - \sigma_\alpha)}\right)}} \quad (58)$$

Овај модел даје квантитативни губитак визуелних информација кроз промене параметара градијента, а које одговарају променама контраста (амплитуде градијента) и облика (оријентација градијента). Облици нелинеарних сигмоидних функција који се користе у релацији (57) су приказани на Сл. 3.5, за вредности које се користе у евалуацији квалитета сједињавања [179, 214] и које су прихваћене у дисертацији.

Крајња мера очувања просторних информација у секвенци, Q_s^{AB} , се добија као сума очувања вредности параметара градијента на свим позицијама и у свим кадровима секвенце, пондерисана са фактором локалног значаја, w_s , узетог као максимална локална вредност градијента изворне и деградиране секвенце:

$$w_{s_{n,m,t}} = \max\left(g_{s_{n,m,t}}^A, g_{s_{n,m,t}}^B\right) \quad (59)$$

$$Q_s^{AB} = \frac{\sum_{\forall n,m,t} Q_{s,n,m,t} w_{s,n,m,t}}{\sum_{\forall n,m,t} w_{s,n,m,t}} \quad (60)$$

Вредности мере Q_s^{AB} се такође налазе у опсегу $[0,1]$, где 0 одговара комплетном губитку просторних информација док је вредност 1 пренос без деградације. Оптималне вредности параметара нелинеарних функција које се користе у моделу за субјективни губитак информација могу се одредити кроз оптимизацију резултата објективних скорова и резултата субјективних тестова, као што је дато у [221]. Вредности параметара Γ_s , Γ_g и Γ_α се налазе аналитички, тако да вредности Q_s , Q_g и Q_α , тим редом, буду једнаке 1 када су G и A такође 1.

Једноставан модел очувања информација о просторном градијенту на нивоу пиксела, по узору на модел који се користи у процени квалитета сједињавања се може дефинисати и као:

$$S_{n,m,t} = \sqrt{A_{n,m,t}^{AB} \cdot G_{n,m,t}^{AB}} \quad (61)$$

Усредњавање очувања просторних информација на свим позицијама у кадру ($N \times M$) и у свим кадровима (којих је N_{FR}) даје:

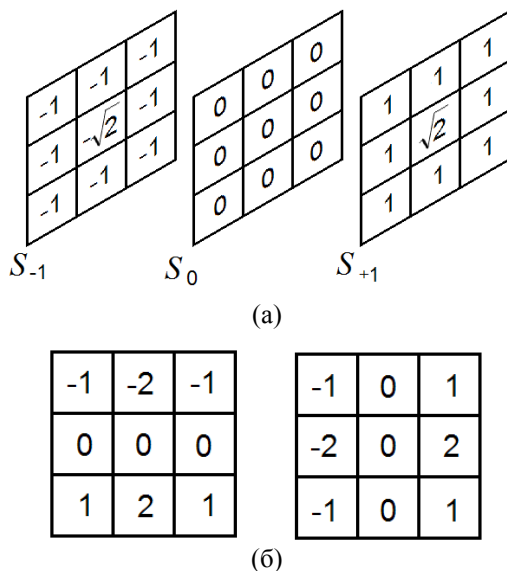
$$S^{AB} = \frac{\sum_{\forall n,m,t} S_{n,m,t}}{N \cdot M \cdot N_{FR}} \quad (62)$$

Ова мера квалитета се може тумачити као верност очувања изворних информација у секвенци са деградацијом (проста репрезентација информација у секвенци са деградацијом). Она не садржи модел визуелног система човека па се може назвати линеарном мером. Додатно, мера S^{AB} не узима у обзир значај региона на слици.

5.4. Очување временских информација

За евалуацију очувања временских информација користе се временски градијенти који се такође издвајају из канала интензитета. За ту сврху се користи временски градијентни оператор дат на Сл. 5.5(а). Овај оператор је аналоган

стандардном Собеловом оператору за издвајање ивица а који се користи за очување просторних информација (приказан на Сл. 5.5(б)).



Сл. 5.5 Оператори коришћени за издвајање: (а) временског градијента и (б) просторног градијента.

Временски градијент gt се добија кроз анализу промене интензитета у времену. За текући кадар, t , добија се конволуцијом (филтрирањем) претходног $t-1$ и наредног $t+1$ кадра секвенце са S_{-1} и S_{+1} операторима датим на Сл. 5.5(а):

$$gt_t = \left| (I_{t-1} * S_{-1}) + (I_{t+1} * S_{+1}) \right| \quad (63)$$

Слично као у анализи очувања просторног градијента, очување временског градијента gt се посматра између добијених вредности на свим локацијама и у свим кадровима изворне и деградиране секвенце. За разлику од модела коришћеног у анализи сједињавања [247], где се посматра очување амплитуде временског градијента, у процени квалитета секвенци (када постоји један извор – изворна секвенца) битан је и смер промене временског градијента. Промена временског градијента, T , се дефинише слично као очување амплитуде просторног градијента, али уз очување знака (смера градијента):

$$T_{n,m,t} = \text{sgn}\left(\frac{gt_{n,m,t}^A}{gt_{n,m,t}^B}\right) \exp\left(-\left|\log\left(\frac{gt_{n,m,t}^A}{gt_{n,m,t}^B}\right)\right|\right) \quad (64)$$

Субјективни губитак временских информација се добија пресликавањем промене T користећи нелинеарни модел очувања информација:

$$Q_{t,n,m,t} = \frac{\Gamma_t}{1 + e^{k_t(T_{n,m,t} - \sigma_t)}} \quad (65)$$

Очување градијента Q_t се налази у интервалу $[0,1]$, где 0 одговара потпуном губитку изворних (временских) информација а 1 је пренос без деградације.

Аналогно процени очувања просторних информација изворне секвенце, квалитет видеа са становишта очувања временских информација у деградираној секвенци, Q_t^{AB} , се добија као сума очувања временских информација на свим локацијама и у свим кадровима, пондерисана са локалним субјективним значајем, w_t , датим као максимум временског градијента изворне и деградиране секвенце:

$$w_{t,n,m,t} = \max(g_{t,n,m,t}^A, g_{t,n,m,t}^B) \quad (66)$$

$$Q_t^{AB} = \frac{\sum_{\forall n,m,t} Q_{t,n,m,t} w_{t,n,m,t}}{\sum_{\forall n,m,t} w_{t,n,m,t}} \quad (67)$$

Вредности мере Q_t^{AB} су у опсегу $[0,1]$, где 0 говори о потпуном губитку информација док $Q_t^{AB}=1$ говори да се ради о преносу без деградације. Параметри нелинеарне сигмоидне функције из (65) могу се одредити кроз оптимизацију резултата објективне мере и резултата субјективних тестова [221]. Вредност параметра Γ_t се налази аналитички, тако да вредност Q_t буде једнака 1 када је T такође 1, (65).

Алтернативно, временске и просторне информације се могу издвајати истовремено у форми просторно-временских градијената. На овај начин се унифицира издвајање информација из канала интензитета и поједностављује субјективни модел очувања информација (потребно одређивање мањег броја параметара нелинеарних сигмоидних функција). Међутим, на овај начин није могуће раздвојити временске информације (покрети на сцени) од информација о структури.

Уколико је циљ да се евалуира проста репрезентација временских информација из изворне секвенце у секвенци са деградацијом, то се може учини кроз линеарну меру:

$$T^{AB} = \frac{\sum_{\forall n,m,t, T_{n,m,t} > 0} T_{n,m,t}}{N \cdot M \cdot N_{FR}} \quad (68)$$

Оваквом дефиницијом очувања информација о временском градијенту уколико постоји промена смера временског градијента очување информација једнако је нули.

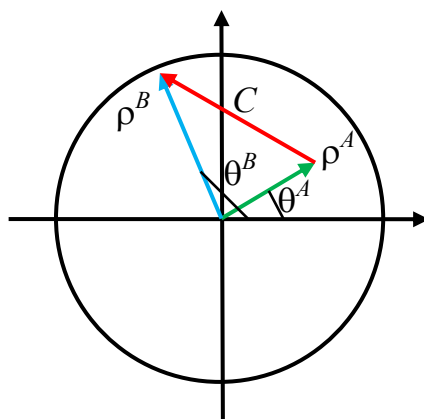
5.5. Очување хроматских информација

Хроматске информације се након преласка у HSV колор координатни систем, налазе у H и S каналима изворне и деградиране секвенце. Захваљујући природи овог система, вредности у овим каналима су у опсегу [0,1]. Аналогно анализи очувања просторних и временских информација, промене параметара боје се иницијално евалуирају на свим локацијама и у свим кадровима између изворне и деградиране секвенце.

Захваљујући природи HSV система, хроматске информације се анализирају кроз поларни модел очувања информација. Овај модел третира H и S вредности као поларне координате (ρ, θ) , што је илустровано на Сл. 5.3. У овом моделу, H канал је угао ротације, што одговара параметру θ ($\theta \in [0, 2\pi]$), а засићење S говори о удаљености од координатног почетка, што одговара параметру ρ ($\rho \in [0, 1]$). На овај начин се здружени хроматски вектори налазе унутар јединичног круга. Очување боје се дефинише као Еуклидско растојање хроматских вектора изворне и деградиране секвенце:

$$C_{n,m,t} = 1 - \frac{1}{2} \left\| (\rho_{n,m,t}^B, \theta_{n,m,t}^B) - (\rho_{n,m,t}^A, \theta_{n,m,t}^A) \right\| \quad (69)$$

Евалуација очувања хроматских информација је илустрована на Сл. 5.6. Хроматски простор се налази унутар јединичног круга јер $\rho \in [0, 1]$. Разлика боје између изворне и деградиране секвенце постоји због промене у засићењу, када се приближава или удаљава од координатног почетка и/или промене врсте боје – оријентације θ .



Сл. 5.6 Евалуација очувања хроматских информација.

Субјективни губитак/очување боје се добија коришћењем нелинеарне сигмоидне функције, дефинисане параметрима k_c и σ_c :

$$Qc_{n,m,t} = \frac{\Gamma_c}{1 + e^{k_c(C_{n,m,t} - \sigma_c)}} \quad (70)$$

Локалне евалуације боје се сумирају на свим позицијама и у свим кадровима и пондеришу користећи мапу субјективног значаја w_c . Како хроматске информације човек обрађује на начин који се разликује од обраде информација које одговарају кретању и структурама (временске и просторне информације) и како се очување боје може дати посматрањем комплетног видног поља, мапа значаја w_c је униформна у комплетној секвенци ($w_c=1$). Глобални фактор очувања боје се рачуна као:

$$Q_c^{AB} = \frac{\sum_{\forall n,m,t} Qc_{n,m,t}}{N \cdot M \cdot N_{FR}} \quad (71)$$

и такође се налази у опсегу $[0,1]$, који одговара комплетном губитку колор информација или преносу без деградације боје.

Алтернативни приступ очувања боје би био приступ у коме би се одвојено разматрало очување информација у Н и S каналима. Код овог приступа би било потребно одредити два пара вредности k и σ , један пар за један а други пар за други канал.

Очување боје из изворне секвенце у секвенци са деградацијом се може евалуирати кроз линеарну меру:

$$C^{AB} = \frac{\sum_{\forall n,m,t} C_{n,m,t}}{N \cdot M \cdot N_{FR}} \quad (72)$$

чиме се избегава коришћење нелинеарне сигмоидне функције.

5.6. Крајња мера квалитета видеа

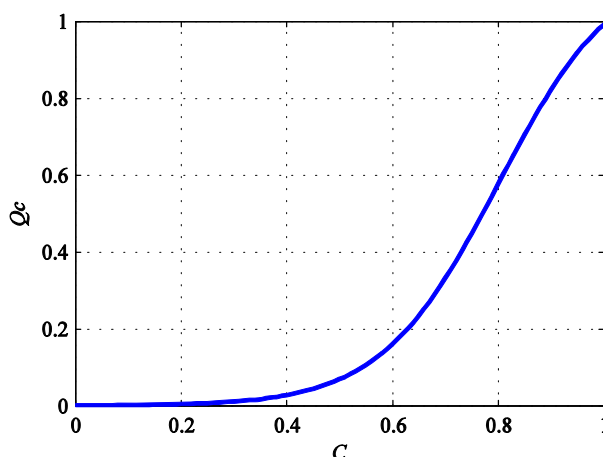
Појединачне естимације очувања информација изворне у деградираној секвенци, тј. очување просторних информација Q_s^{AB} , временских информација Q_t^{AB} и хроматских информација Q_c^{AB} се могу комбиновати у једну вредност која представља меру квалитета видео секвенце:

$$VQ^{AB} = k_s Q_s^{AB} + k_t Q_t^{AB} + k_c Q_c^{AB} \quad (73)$$

Коефицијенти значаја k_s , k_t и k_c ($k_s+k_t+k_c=1$) могу бити одређени кроз поступак оптимизације објективних скорова уз доступне субјективне скорове. Њихова вредност даје корисне информације о ефектима сваког од извора информација на људски визуелни систем у процени квалитета видеа. Динамички опсег индекса квалитета $VQ^{AB} \in [0,1]$ је очуван. За деградирану секвенцу са већом вредношћу VQ^{AB} каже се да је бољег квалитета.

Предложена мера за процену квалитета видеа (73) користи неколико параметара од којих зависи слагање објективних и субјективних скорова. Већину параметара чине коефицијенти нелинеарних сигмоидних функција – k и σ у релацијама (57), (65) и (70), а који утичу на субјективни губитак информација. Уместо да се спроведе мултидимензиони оптимизациони поступак, у раду су за ове коефицијенте преузете вредности које су добијене у истраживањима за потребе сједињавања слика [214, 221] и за које је утврђено да су субјективно значајне, $[k_g, \sigma_g, k_\alpha, \sigma_\alpha, k_t, \sigma_t] = [-11, 0.7, -24, 0.8, -11, 0.7]$. Како ранија истраживања нису обухватила модел очувања хроматских информација, урађена је делимична оптимизација, кроз коју је на изабраном скупу тест секвенци утврђено да су оптималне вредности параметара субјективног губитка хроматских информација $[k_c, \sigma_c] = [-9, 0.8]$. Добијене вредности одговарају скоро линеарној вези деградације хроматских информација и

субјективног утиска квалитета – Сл. 5.7. Ово је у складу са запажањем да у хроматске информације визуелни систем човека има мање поверења него у контраст који је евалуиран кроз просторне информације.



Сл. 5.7. Нелинеарна сигмоидна функција која се користи у очувању хроматских информација.

Због коришћења нелинеарних сигмоидних функција мера процене квалитета видеа (73) ће се у даљем раду називати нелинеарном. Алтернативни приступ је приступ у коме се анализира промена сигнала и третира као вредност квалитета. Коришћењем овог, линеарног приступа, избегавају се коефицијенти нелинеарних сигмоидних функција (којих има укупно осам), и вредност мере квалитета видеа се може дефинисати као:

$$VQ_{lin}^{AB} = k_s S_{lin}^{AB} + k_t T_{lin}^{AB} + k_c C_{lin}^{AB} \quad (74)$$

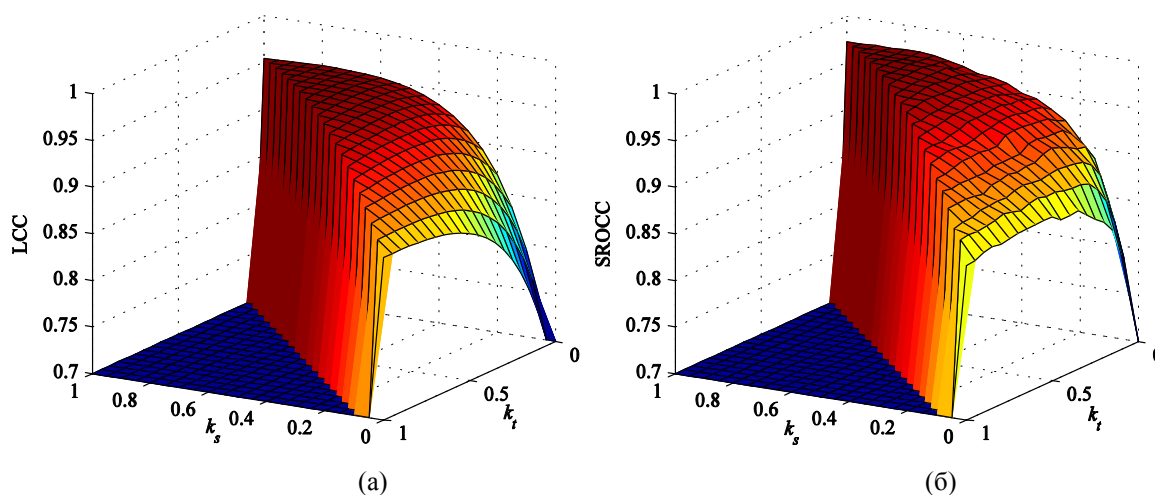
где S_{lin}^{AB} , T_{lin}^{AB} и C_{lin}^{AB} говоре о просторној, временској и хроматској промени сигнала (пондерисано са одговарајућим мапама значаја):

$$S_{lin}^{AB} = \frac{\sum_{\forall n,m,t} S_{n,m,t} w_{sn,m,t}}{\sum_{\forall n,m,t} w_{sn,m,t}} \quad (75)$$

$$T_{lin}^{AB} = \frac{\sum_{\forall n,m,t} T_{n,m,t} w_{tn,m,t}}{\sum_{\forall n,m,t} w_{tn,m,t}} \quad (76)$$

$$C_{lin}^{AB} = \frac{\sum_{\forall n,m,t} C_{n,m,t} W_{cn,m,t}}{\sum_{\forall n,m,t} W_{cn,m,t}} \quad (77)$$

Како би се одредиле оптималне вредности значаја различитих извора информација, (73), спроведена је детаљна оптимизација на скупу тренинг секвенци из једног од доступних сетова видео секвенци са субјективним импресијама квалитета (UoM (University of Manchester) база – биће описана у поглављу 5.9.1). Тренинг скуп UoM базе обухвата половину случајно одабраних тест секвенци (120 секвенци). Простор значаја канала је у основи дводимензионалан јер је сума све три тежине једнака 1, тј. $k_s+k_t+k_c=1$. На Сл. 5.8 приказане су оптимизационе површине за оптимизацију спроведену са кораком релативног значаја од 0.05 према коефицијенту линеарне корелације, LCC, и корелацији рангова, SROCC, између субјективних и објективних скорова. Плави троугао са леве стране приказа се односи на ситуације када је сума значаја већа од 1. Ако се посматра валидни део приказа, горњи леви угао одговара максималном значају просторног канала, горњи десни угао максималном значају хроматских информација и доњи десни угао максималном значају временских информација.



Сл. 5.8 Оптимизационе површине значаја изворних информација у објективној процени квалитета: (а) према коефицијенту линеарне корелације, LCC, (б) према корелацији рангова, SROCC.

Са оптимизационе површине се види да просторни канал даје најпоузданије естимације квалитета, након чега следи временски па хроматски канал. Оптималне

вредности значаја канала су $[k_s, k_t, k_c]=[0.8, 0.15, 0.05]$. Сама оптимизациона површина је заравњена па сличан вектор значаја такође даје меру квалитета са сличним перформансама. Због тога мера није претерано осетљива на ове параметре, што је предност саме мере.

5.7. Примери примене мере квалитета VQ^{AB}

На Сл. 5.9 и Сл. 5.10 дати су примери очувања изворних информација секвенце из сета UoM. Изворна секвенца је деградирана поступком компресије (H.264), за две битске брзине – 96 kbps (Сл. 5.9) и 416 kbps (Сл. 5.10). Приказана су по три суседна кадра изворне и деградиране секвенце, приказане су мапе очувања просторних, временских и колор информација, одговарајуће мапе субјективног очувања/губитка информација и мапе локалног значаја које се користе као пондери у процени квалитета (максималне амплитуде просторног/временског градијента). Светлије вредности са Сл. 5.9 и Сл. 5.10 одговарају вернијем трансферу информација – преносу без деградације, док пиксели црне боје говоре о потпуном губитку информација.

У оба анализирана случаја, тј. за обе битске брзине, скоро да постоји идеалан трансфер информација о боји. Временске информације су сачуване са мање успеха од просторних информација – на мапама очувања временских информација јављају се велике групе (региони) црних пиксела. Ово је очекивано јер се поступком компресије врши потискивање случајних и малих промена – покрета, и ресурси додељују на значајна кретања на сцени – у овом случају ток саобраћаја у десном делу слика. Додатно се на мапама очувања временских информација уочавају квадратне структуре због коришћења блоковске компресије. Очување просторних информација је верније за ивичне пикселе него за пикселе који се налазе у регионима који имају униформан ниво сивог. Такође, повећањем битског протока мапе очувања информација изворне секвенце постају светлије. Мапе значаја просторних информација одговарају ивицама док мапе значаја временских информација одговарају оптичком току. У Таб. 5.1 су за два анализирана примера дате вредности субјективне импресије квалитета (у форми MOS) и објективних мера. Вредности

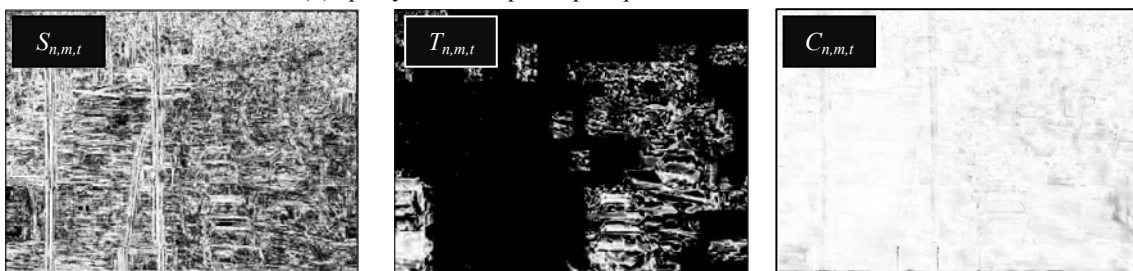
објективних мера Frame PSNR и Frame SSIM добијене су усредњавањем вредности PSNR/SSIM срачунатих на нивоу кадра.



(а) три суседна кадра оригиналне секвенце



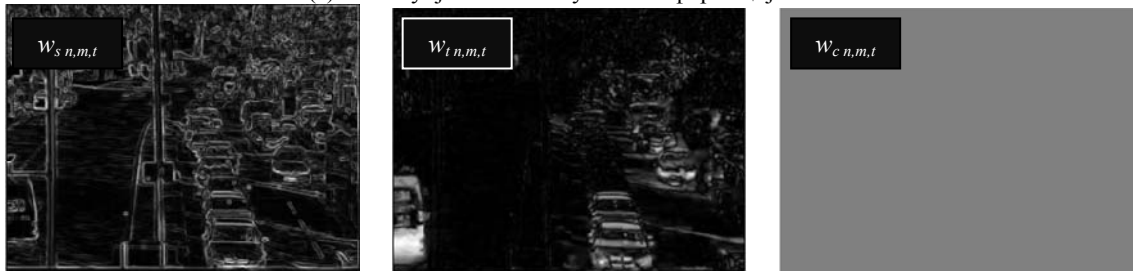
(б) три суседна кадра деградиране секвенце



(в) мапе очувања информација у три канала



(г) мапе субјективног очувања информација



(д) мапе субјективног значаја

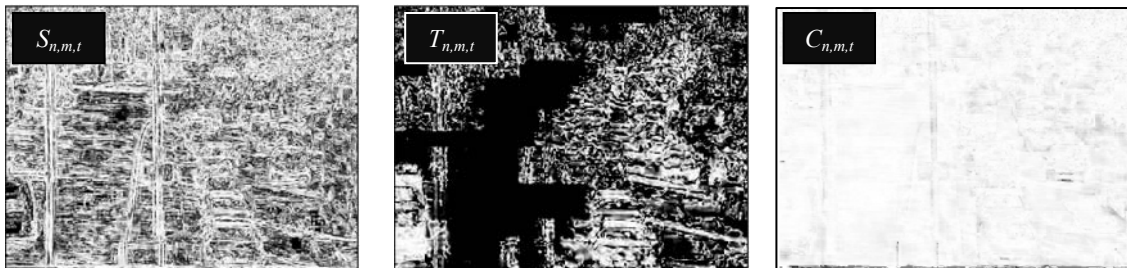
Сл. 5.9 Пример очувања информација изворне секвенце за битски проток 96 kbps.



(а) три суседна кадра оригиналне секвенце



(б) три суседна кадра деградиране секвенце



(в) мапе очувања информација у три канала



(г) мапе субјективног очувања информација



(д) мапе субјективног значаја

Сл. 5.10 Пример очувања информација изворне секвенце за битски проток 416 kbps.

Таб. 5.1 Вредности субјективних и објективних процена за секвенце анализирани на Сл. 5.9 и Сл. 5.10.

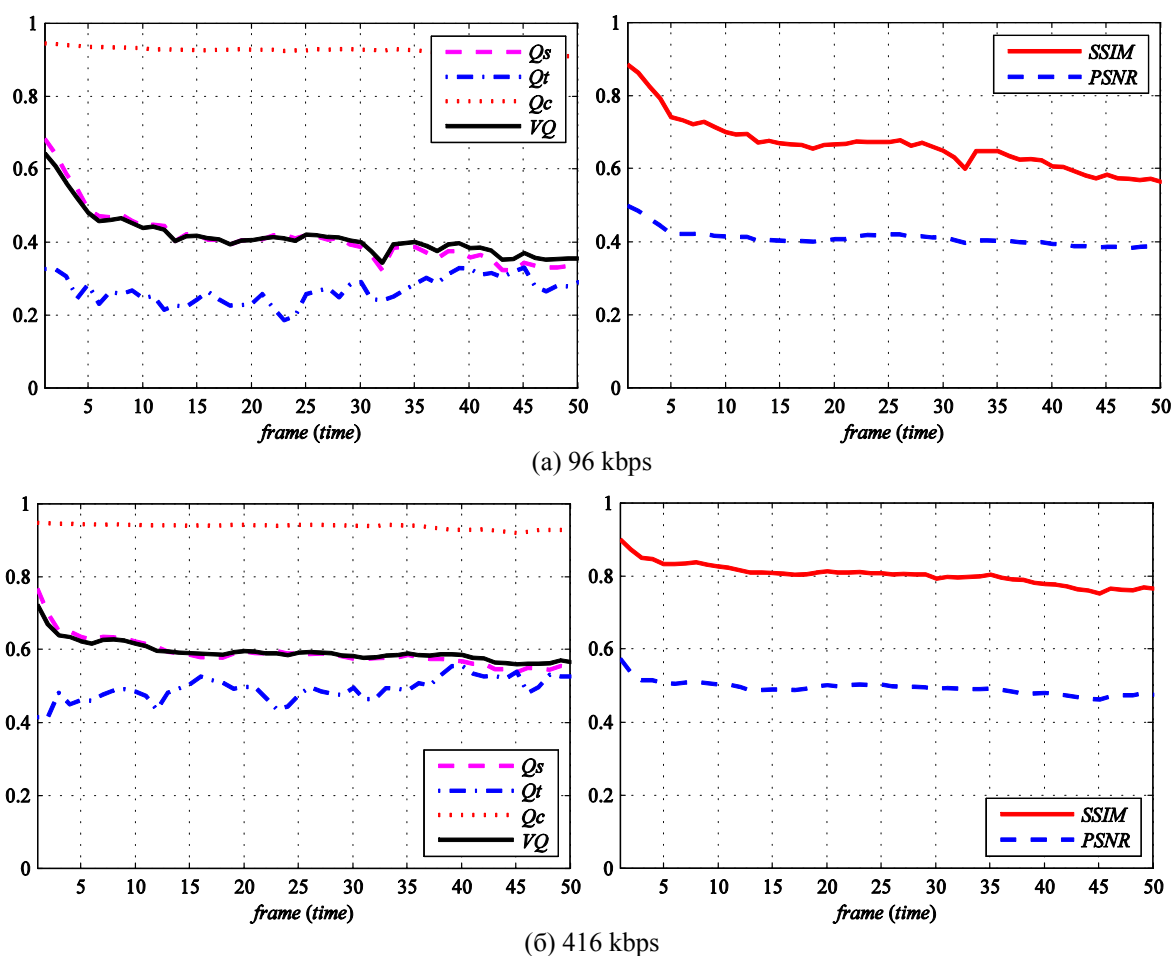
	<i>MOS</i>	<i>Frame PSNR</i>	<i>Frame SSIM</i>	Q_s^{AB}	Q_t^{AB}	Q_c^{AB}	VQ^{AB}
Сл. 5.9 (96 kbps)	2.2381	20.4000	0.6611	0.4098	0.2686	0.9227	0.4143
Сл. 5.10 (416 kbps)	3.6875	24.6124	0.8027	0.5893	0.4919	0.9358	0.5920

Из Таб. 5.1 се види да је по свим објективним мерама квалитет секвенце са битским протоком од 416 kbps бољи, а што је у складу са субјективном проценом. Такође се уочава да је разлика вредности квалитета боје, Q_c^{AB} , мања од 1.5% иако је разлика субјективних скорова значајна, тј. субјективни утисак квалитета различит.

Поред једне – глобалне вредности која говори о квалитету комплетне секвенце, предложена мера за процену квалитета се може користити и за добијање вредности квалитета на нивоу кадра. Полазећи од поређења сигнала на нивоу пиксела локалне вредности се помоћу мапе значаја w интеграле у вредност квалитета појединачних кадрова (на нивоу комплетне мере или кроз евалуацију у појединачним каналима). Приказ квалитета током времена може дати додатне информације о кретању квалитета видеа. Очување информација током времена за две секвенце анализирани у примерима са Сл. 5.9 и Сл. 5.10 дато је на Сл. 5.11. Поред очувања информација током времена за анализирана три канала и предложену меру, дате су и промене индекса структурне сличности, SSIM, и вршног односа сигнал шум, PSNR, (који је овде сведен на опсег [0,1] ради лакшег поређења).

Са Сл. 5.11 се види да су информације о боји боље очуване у комплетној секвенци од просторних и временских информација. Тренд промене вредности очувања просторних информација и вредности збирне мере VQ је сличан тренду промене индекса структурне просторне сличности, SSIM. Од овог тренда одступа очување временских информација. Иако се ради о секвенци релативно кратког трајања, очување информација се мења током времена и до 50%. Може се закључити да током времена долази до пада квалитета тест секвенце по све три објективне мере, VQ, SSIM и PSNR, а што је последица повећања оптичког тока. Додатно, улазак једног од возила у видно поље камере довео је до локалног пада вредности мера VQ и SSIM при битском протоку од 96 kbps (32 кадар, Сл. 5.11(a)). Овај улазак је довео до замрљања у тест секвенци, значајно већег од стварних димензија возила.

Једна од практичних примена анализе квалитета на нивоу кадрова је процена погодности кодека или система за приказ различитим садржајима са сцене. Како садржај сцене може утицати на квалитет видеа, систем који има могућност оптимизације се може користити да би се кодек или параметри система прилагодили новом садржају.



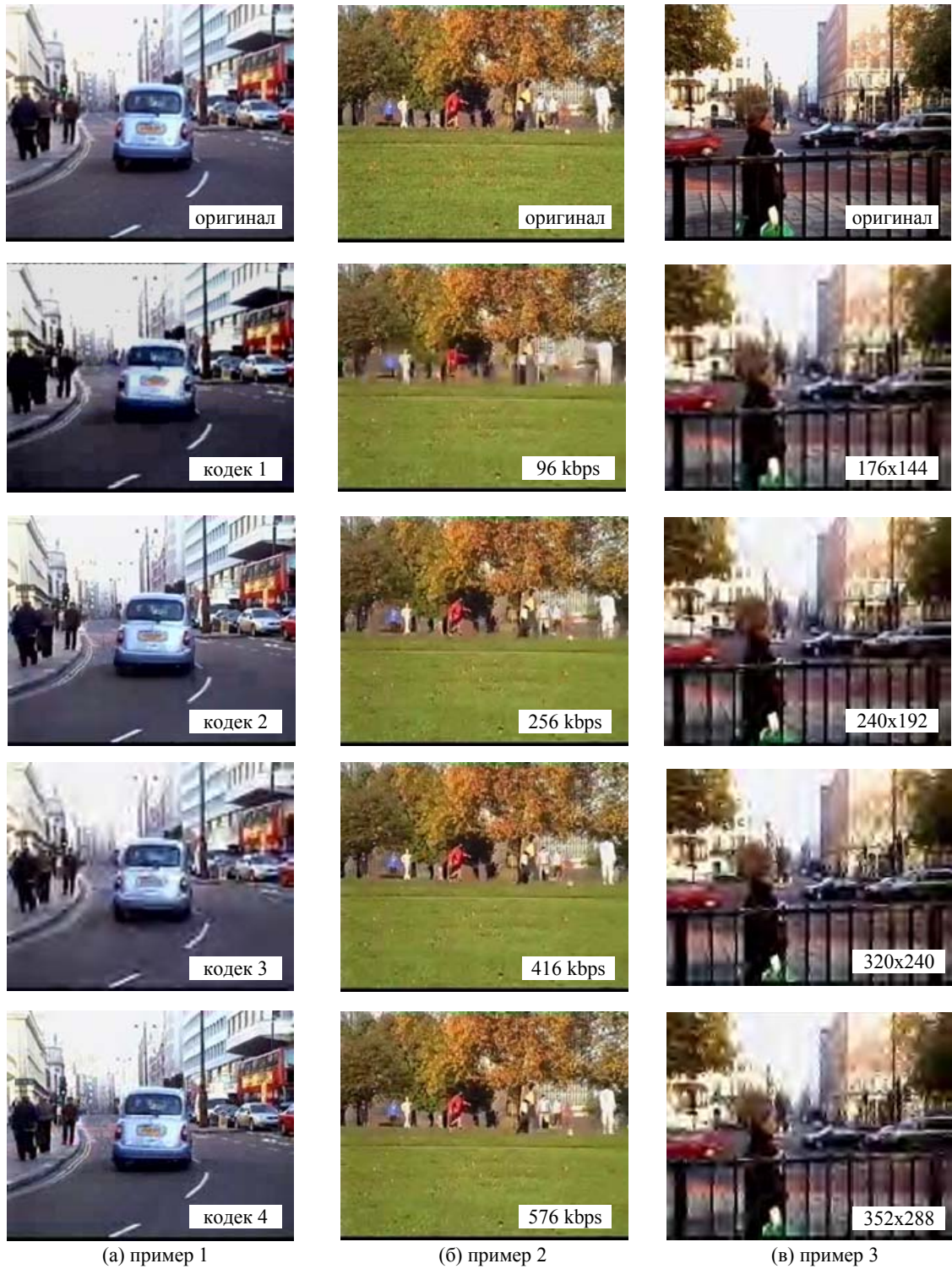
Сл. 5.11 Очување информација током времена за примере анализирани на Сл. 5.9 и Сл. 5.10.

На Сл. 5.12 приказани су кадрови изворних и деградираних секвенци за различите параметре система преноса, тј. промену кодека (пример 1), промену битске брзине (пример 2) и промену димензије кадра (пример 3). У анализираним примерима један од параметара се мењао док су остали параметри система преноса били фиксни. Као изворни сигнали коришћене су секвенце исте резолуције као и секвенце са деградацијом, тј. нису коришћене изворне секвенце највеће доступне резолуције. Иако се у примерима не може видети очување/трансфер информација о кретању објеката на сцени, из примера се види да квалитет зависи од избора параметара система. Додатне информације о анализираним примерима су дате у Таб. 5.2.

У примеру 1 се за кодек 1 уочава блоковски ефекат, док је за кодек 3 карактеристична појава замрљања. Ове ефекте посматрачи су уочили (вероватно третирали као губитак информација или лажне информације) и самим тим дали ниже

5. Процена квалитета видеа заснована на очувању информација о градијенту

вредности квалитета. Ранг субјективних скорова се у овом случају слаже са рангом све три објективне мере из Таб. 5.2.



Сл. 5.12 Кадрови изворних и деградираних секвенци за различите параметре система преноса: (а) промена кодека, (б) промена битске брзине и (в) промена димензије кадра.

У примеру 2 је за кодек 2 и димензију кадра од 320x240 пиксела анализиран утицај промене битске брзине. Са Сл. 5.12 се види да се за битску брзину од 96 kbps садржај на сцени не може најбоље описати, чак је и фудбалска лопта за овај битски проток нестала са сцене. Иако се за приказане кадрове за преостале три битске брзине разлике не могу уочити, посматрачи су секвенцама дали различите вредности квалитета, а које расту са порастом битске брзине. Ова разлика је вероватно последица очувања информација о кретању на сцени а које се не види кроз приказ појединачних кадрова. Ранг субјективних скорова се у овом случају слаже са рангом све три објективне мере из Таб. 5.2.

Таб. 5.2 Субјективни и објективни скорови за примере са Сл. 5.12.

	<i>MOS</i>	<i>Frame PSNR</i>	<i>Frame SSIM</i>	<i>VQ^{AB}</i>
<u>пример 1 (Сл. 5.12(а))</u>				
димензија кадра: 240x192, битски проток: 576 kbps				
кодек 1	2.5150	14.4410	0.6268	0.5368
кодек 2	4.1250	24.5747	0.8602	0.7676
кодек 3	2.7727	20.4448	0.7037	0.6254
кодек 4	4.4167	24.6773	0.8775	0.8014
<u>пример 2 (Сл. 5.12(б))</u>				
кодек 2, димензија кадра: 320x240				
96 kbps	2.2381	21.6647	0.6725	0.5984
256 kbps	3.3	22.6267	0.7628	0.6689
416 kbps	3.6875	23.9958	0.8095	0.7167
576 kbps	4.1429	25.0567	0.8171	0.7372
<u>пример 3 (Сл. 5.12(в))</u>				
кодек 3, битски проток: 576 kbps				
176x144	3	19.9364	0.7212	0.6573
240x192	2.7727	20.4448	0.7037	0.6254
320x240	2	21.2558	0.5766	0.5466
352x288	1.85	20.6788	0.4800	0.4952

Кроз пример 3 је анализиран утицај промене димензије кадра, при чему је коришћен кодек 3 уз битску брзину од 576 kbps (и даље је уочљив ефекат замрљања). Кроз приказ појединачних кадрова се не могу уочити разлике у квалитету кадрова за промену димензије кадра. Међутим, посматрачи, који су поред просторне димензије видели и временску димензију секвенци, су мишљења да се у секвенцама у нижој резолуцији пренело више изворних информација. Генерално, квалитет секвенци

добитених кодеком 3 је лошији од секвенци добитених коришћењем кодекса 2 и 4, тако да је и за битски проток од 576 kbps квалитет секвенци низак – у датом примеру посматрачи су дали максималну оцену 3 за секвенцу у најнижој резолуцији. За дати пример са смањењем резолуције расте субјективна импресија квалитета. Субјективна импресија квалитета се слаже са вредностима индекса структурне сличности, Frame SSIM, и предложене мере, VQ^{AB} , док од ње одступају вредности добитене коришћењем вршног односа сигнал шум, Frame PSNR.

5.8. Квалитетом вођена VQ^{AB} мера

Основна VQ^{AB} мера описана је релацијом (73) и код ње се користе вредности значаја канала $[k_s, k_t, k_c]=[0.8, 0.15, 0.05]$. Имајући у виду резултате добитене у претходном поглављу дисертације којима је наглашено да на субјективну импресију квалитета слике значајно утичу најниже локалне вредности квалитета, у објективној процени квалитета видеа искоришћена је слична логика. Наиме, за одређивање крајње вредности квалитета користи се део кадрова са најнижим скоровима квалитета. Након одређивања VQ^{AB} вредности квалитета на нивоу свих кадрова, врши се сортирање добитених вредности по растућој амплитуди. Крајња вредност квалитета се добија као средња вредност 20% најнижих VQ^{AB} скорова кадрова. Овакав приступ је вођен хипотезом да кадрови лошег квалитета могу имати доминантну улогу у субјективној импресији квалитета [237, 248]. Утицај најнижих скорова квалитета кадрова на слагање субјективних и објективних скорова биће додатно анализиран у једном од наредних делова дисертације (5.10). У даљем раду, уколико се другачије не нагласи, ознака VQ^{AB} односи се на градијентну процену квалитета видеа са редукованим бројем кадрова (20% кадрова са најнижим скоровима квалитета). За примере са скоровима квалитета кадрова са Сл. 5.11 секвенце су трајања 50 кадрова, па би се крајња вредност квалитета одређивала на основу квалитета 10 кадрова (у случају секвенце са битским протоком од 416 kbps, Сл. 5.11(б), то би било последњих 10 кадрова).

5.9. Анализа перформанси предложене мере

Како би се извршила евалуација предложене мере видео квалитета извршено је прикупљање података, тј. видео секвенци са доступним субјективним проценама. Секвенце су форматиране – преведене у сирови *avi* формат. Позивање на секвенце је урађено коришћењем стандардизованих текстуалних фајлова, прављених са циљем да се на исти начин приступа видео секвенцама свих база (референтним и деградираним) и да се аутоматски читавају субјективне импресије квалитета. Предобрада видео секвенци је омогућила да се избегну трансформације из домена у коме су секвенце добијене (qnt, YUV420p, YUV420i, HSI, ...) у домен који се користи за евалуацију квалитета (RGB). Овакав приступ је практичан и њиме се значајно скраћује време за евалуацију.

Доступне базе за процену квалитета омогућавају истраживачима да евалуирају перформансе алгоритама за објективно одређивање квалитета, као и да достигну основни циљ објективне процене – сагласност са субјективном импресијом квалитета. Прикупљено је неколико таквих база. Већина алгоритама за објективну процену квалитета је евалуирана на јавно доступној VQEG FRTV Phase I бази [249], која је застарела и која има недостатке као што су видео секвенце са проредом (interlaced) и лоша перцептуална сепарација видео. VQEG група није учинила потпуно доступним видео садржаје из каснијих студија – фаза истраживања (укључујући FRTV Phase II и Multimedia study), што ограничава истраживаче у поређењу нових алгоритама објективне процене квалитета са алгоритмима из VQEG извештаја.

Као додаток на LIVE Image Quality базу [223], LIVE група је развила две базе – LIVE Video и LIVE Wireless базе [250, 251]. Две академске институције Ecole Polytechnique Federale de Lausanne – Switzerland (EPFL) и Politecnico di Milano – Italy (PoliMI) су презентовали резултате субјективних тестова на 156 видео секвенци кодованих са H.264/AVC кодом уз симулацију пакетских губитака [252]. Доступно је 78 видео секвенци у CIF просторној резолуцији (352x288) и 78 секвенци у 4CIF просторној резолуцији (704x576). Након упућивања захтева ауторима рада [112] добијено је 60 секвенци QCIF формата (176x144), кодованих H.264 енкодером. Осим секвенци доступни су и резултати субјективних тестова.

Кроз сарадњу са Универзитетом у Манчестеру формиране су две базе видео секвенци са субјективним евалуацијама квалитета (UoM база и STrial база). UoM база је формирана током 2009. године док су сетови STrial базе прикупљени током 2012. године. Истраживања из 2009. године имала су за циљ предлагање мере објективне евалуације квалитета док су истраживања из 2012. године била усмерена на анализу перформанси предложене мере у процени квалитета видео секвенци са малим битским брзинама. Због тога су резултати истраживања из 2009. године одвојени од резултата добијених на STrial сетовима, односно у раду су детаљно описани резултати субјективних тестова на STrial сетовима заједно са анализом перформанси VQ^{AB} мере.

Поред перформанси предложене мере VQ^{AB} , у раду су представљене и перформансе конвенционалне PSNR мере, као и мера које су у истраживачкој заједници прихваћене као стандардне (SSIM [48], MOVIE [102], MAD [103]).

5.9.1. Анализа резултата на UoM бази секвенци

UoM база секвенци садржи 257 секвенци, 20 оригиналних (пет изворних секвенци у четири просторне резолуције) и 237 секвенци насталих компресијом – деградацијом). Деградиране секвенце су настале променом битске брзине, просторне резолуције и начина компресије. Вредности коришћених параметара су дате у Таб. 5.3 док су илустративни примери кадрова оригиналних секвенци дати на Сл. 5.13.

Таб. 5.3 Параметри коришћени при формирању UoM базе секвенци.

<i>Начин компресије</i>	<i>Битска брзина</i>	<i>Просторна резолуција</i>
<i>EV ≡ кодек 1</i>	96 kbps	176x144
<i>H.264 ≡ кодек 2</i>	256 kbps	240x192
<i>MJ2 ≡ кодек 3</i>	416 kbps	320x240
<i>MP4 ≡ кодек 4</i>	576 kbps	352x288

Како се користи пет изворних секвенци, а из табеле се види да су коришћена четири кодека, четири битске брзине и четири резолуције (могуће 5x64=320 секвенци), може се закључити да за сваку секвенцу нису коришћене све могуће комбинације.



Сл. 5.13 Приказ кадра изворних секвенци УоМ базе.

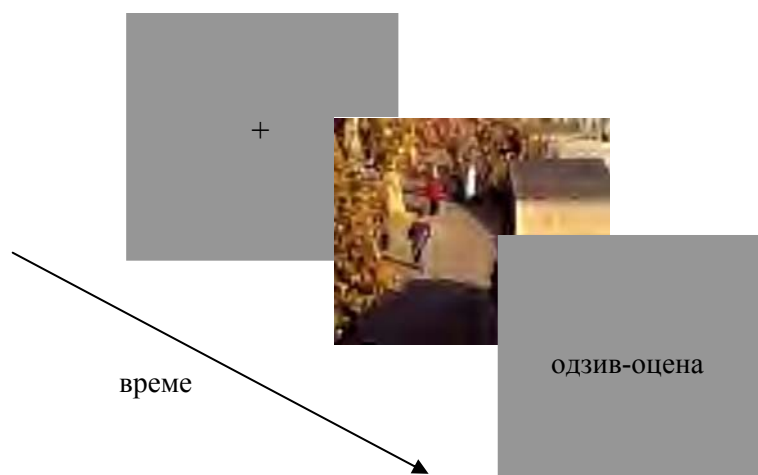
Прикупљање резултата субјективних тестова спроведено је кроз две студије, у амбијенталним условима прописаним од стране ITU [27, 28]. Прва студија имала је за циљ прикупљање субјективних импресија квалитета и у њој је учествовало 30 посматрача. Друга студија имала је за циљ да се прикупе подаци о томе где посматрачи гледају током евалуације квалитета (процене квалитета прикупљене у овој студији нису коришћене у евалуацијама). Пре спровођења субјективних тестова извршена је провера вида посматрача (оштрина вида и опажање боје).

Редослед приказивања видео секвенци је био случајан, како би се елиминисао ефекат учења и одржала пажња посматрача. Редослед посматрања видео секвенци и субјективне евалуације квалитета су аутоматски генерисани и архивирани у “csv” (comma-separated-value) фајлове.

Евалуиране видео секвенце су просторно скалиране тако да су посматрачима сви видео сигнали приказани у највишој доступној резолуцији (352x288), чиме сваки евалуирани видео заузима исти процентуални део од димензија монитора за приказ а посматрачи немају потребу за подешавањем видног угла. За скалирање је коришћен програмски пакет Матлаб и интерполација помоћу најближег суседа (nearest neighbour).

Видео сигнали су презентовани помоћу специјализованог програмског пакета намењеног за рад у Матлабу – Psychtoolbox (Psychophysical Toolbox for Matlab™) [253, 254]. Видео секвенце су презентоване на мониторима Sony Trinitron G250 20 inch CRT резолуције 800x600. Видео секвенце су на овај начин виђене под хоризонталним видним углом од 16.6° и вертикалним видним углом од 14.3°, за удаљеност посматрача 0.55m од монитора.

Током евалуације квалитета субјектима су прво приказане оригиналне видео секвенце у највишој резолуцији. То је урађено како би се посматрачи упознали са секвенцама и како би се смањио ефекат учења. Након тога су субјектима приказиване тест видео секвенце у случајном редоследу. Презентација једне секвенце се састоји од: фиксационе тачке на средини екрана, праћене приказом видео секвенце, као на Сл. 5.14. Сви видео записи су током приказа центрирани на монитору.



Сл. 5.14 Шематски приказ презентације видео садржаја.

Након приказа посматрачи су одговарали на питање: Можете ли да опишете шта се дешава у видео секвенци? Посматрачима је понуђен мени са пет могућих опција:

- опција 1 – сигурно не могу
- опција 2 – мислим да не могу (али нисам сигуран)
- опција 3 – не користи се
- опција 4 – мислим да могу
- опција 5 – сигуран сам да могу

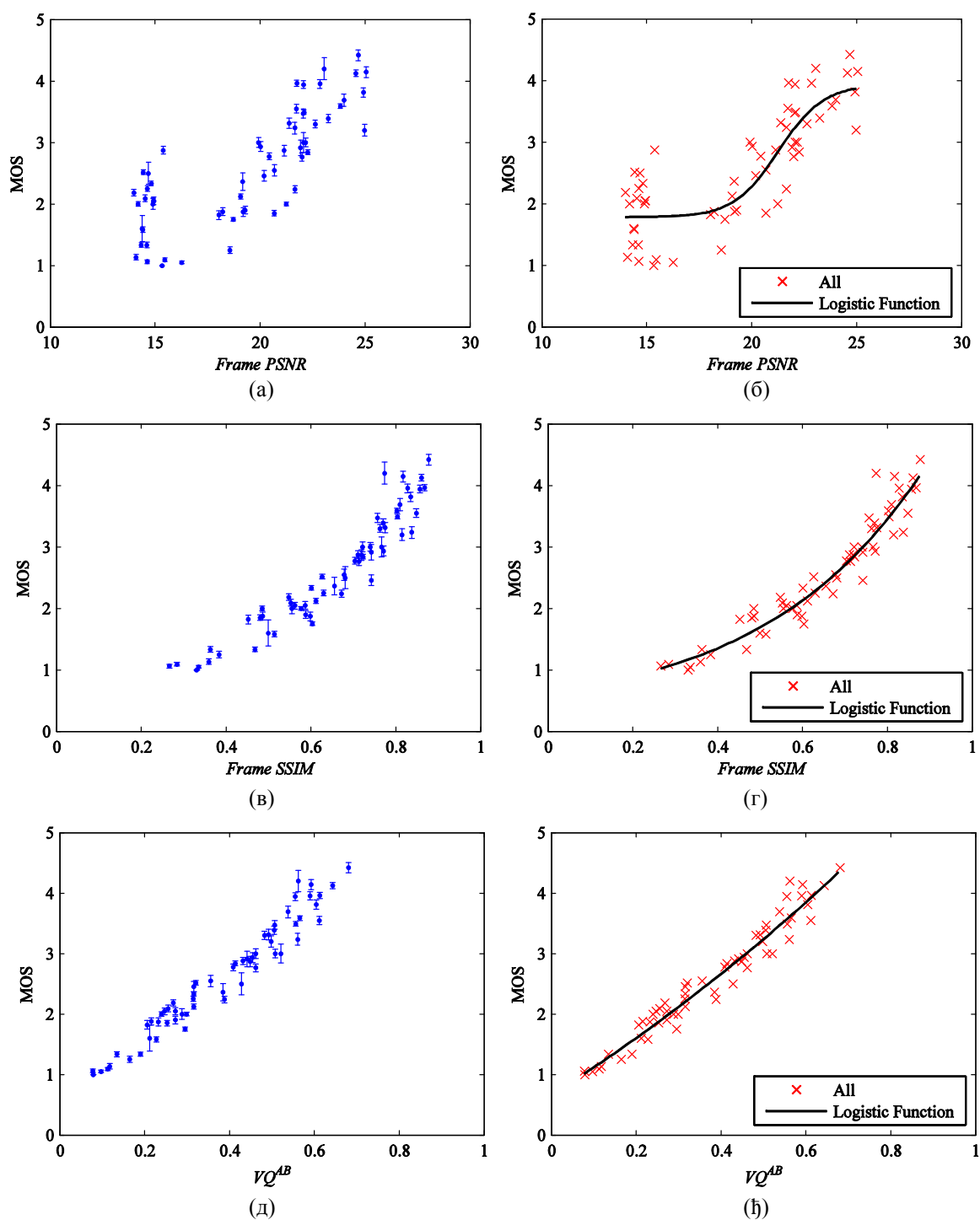
Током давања одговора не врши се приказ секвенци.

На комплетном UoM сету (237 секвенци) упоређене су перформансе три објективне мере процене – предложене VQ^{AB} , конвенционалне PSNR (Frame PSNR) и индекса структурне сличности (Frame SSIM). Резултати евалуације су дати у Таб. 5.4, при чему је коришћена интерполациона функција четвртог реда. Изворне и тест секвенце су просторно и временски регистроване.

Таб. 5.4 Перформансе три објективне мере на комплетном UoM сету.

<i>Мера процене</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>Frame PSNR</i>	0.8504	0.8042	0.4086	0.4828	81.6667
<i>Frame SSIM</i>	0.9627	0.9627	0.1812	0.2481	53.3333
<i>VQ^{AB}</i>	0.9765	0.9777	0.1554	0.1979	55.0000

Перформансе предложене мере су боље од перформанси друге две мере. При томе су перформансе мера VQ^{AB} и Frame SSIM значајно боље од перформанси вршног односа сигнал/шум, Frame PSNR. Перформансе предложене мере скоро да достижу теоретски максимум једнак 1 за LCC и SROCC мере. Ово је такође потврђено на дијаграмима расипања у простору субјективно објективних скорова, Сл. 5.15. Свака тачка у простору субјективно–објективних скорова на Сл. 5.15 представља усредњену вредност више субјективних-објективних скорова, тј. свака тачка представља субјективне-објективне скорове више секвенци са истом компресијом, истом битском брзином и резолуцијом а са различитим садржајима (237 секвенци је пресликовано на 60 подскупова – 60 тачака на дијаграмима). Дијаграми показују да иако је вршни однос достигао релативно добар степен корелације са субјективним скоровима, одзив мере је у великој мери нелинеаран а што није случај са преостале две мере. Наравно, по питању линеарности у овом простору најбоља је VQ^{AB} мера чији су параметри оптимизовани на половини UoM сета, а резултати дати на комплетном UoM скупу. Релативно мали интервали поузданости говоре о високом степену поверења у начин спровођења тестова, где се већина посматрача слаже у рангирању/евалуацији деградираних секвенци.



Сл. 5.15 Дијаграми расипања објективних и субјективних скорова за секвенце у УоМ сету: (а) (в) (д) прикази са интервалима поузданости субјективних скорова, (б) (г) и (ђ) прикази са оптималним интерполационим кривама.

Перформансе мера презентоване у Таб. 5.4 и илустроване на Сл. 5.15 говоре о понашању мера на глобалном нивоу – на нивоу комплетне базе. Некада је потребно

5. Процена квалитета видеа заснована на очувању информација о градијенту

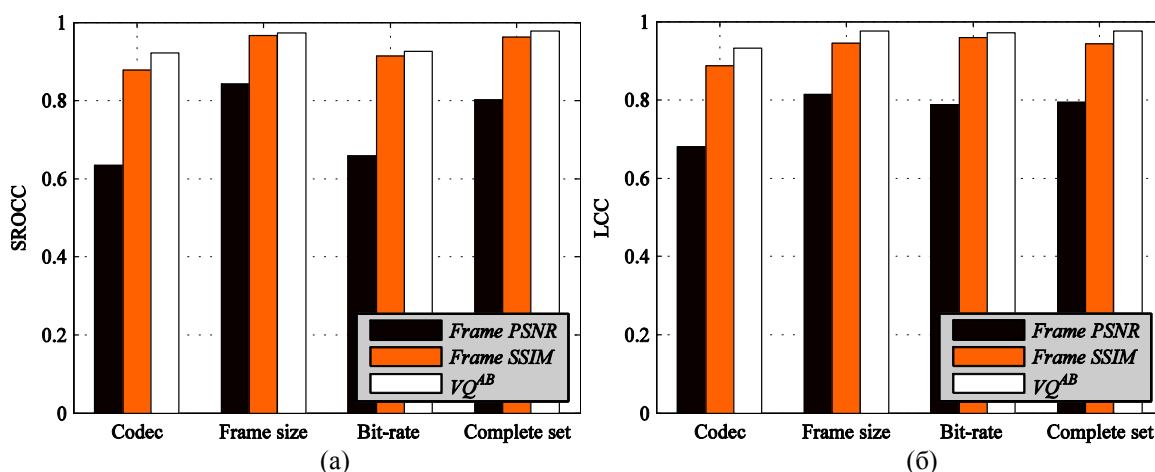
анализирати перформансе и на нивоу подскупова. У овом случају то су перформансе мера према типу компресије, битској брзини и димензији кадра. Тако су у Таб. 5.5 дати корелација рангова и коефицијент корелације (без интерполације помоћу логистик функција) на подскуповима и УоМ бази у целини. Приказани резултати омогућавају бољу анализу понашања објективних мера. Интерполација није рађена јер подскуп секвенци за исти тип кодека, исту димензију кадра или исту битску брзину има свега 15 елемената (60 тачака је подељено на четири класе са по 15 елемената). Због одсуства интерполације резултати корелације LCC на комплетном УоМ сету (последњи ред из Таб. 5.5) се разликују од резултата датих у Таб. 5.4, где је рађена интерполација.

Таб. 5.5 Анализа перформанси објективних мера на подскуповима и комплетној УоМ бази секвенци.

		SROCC			LCC		
		Frame PSNR	Frame SSIM	VQ ^{AB}	Frame PSNR	Frame SSIM	VQ ^{AB}
Кодек	кодек 1	0.45	0.91176	0.94118	0.53841	0.93477	0.93774
	кодек 2	0.74126	0.88112	0.95105	0.76487	0.90967	0.96254
	кодек 3	0.77941	0.89412	0.93529	0.78563	0.93482	0.97206
	кодек 4	0.56176	0.82353	0.85882	0.63207	0.76837	0.85725
	средња вредност	0.63311	0.87763	0.92159	0.68025	0.88691	0.9324
Димензија кадра	176x144	0.79021	0.98601	0.99301	0.72631	0.94254	0.96865
	240x192	0.89412	0.97059	0.98824	0.86799	0.94147	0.98514
	320x240	0.82059	0.93824	0.94706	0.84583	0.95478	0.97773
	352x288	0.86471	0.97059	0.96471	0.81601	0.93918	0.97178
	средња вредност	0.84241	0.96636	0.97325	0.81404	0.94449	0.97582
Битски проток	96 kbps	0.63214	0.93571	0.95357	0.91993	0.97086	0.98205
	256 kbps	0.625	0.86429	0.88929	0.73388	0.94957	0.96498
	416 kbps	0.68929	0.93571	0.93571	0.7789	0.9706	0.95488
	576 kbps	0.68571	0.91786	0.92857	0.71985	0.94255	0.98069
	средња вредност	0.65804	0.91339	0.92679	0.78814	0.9584	0.97065
Комплетан сет		0.80122	0.96188	0.97672	0.79354	0.94283	0.97549

Резултати приказани у Таб. 5.5 говоре о конзистентности (доследности) објективних VQ^{AB} и SSIM скорова и субјективних процена, независно од типа

компресије, димензије кадра или битског протока – корелација објективних и субјективних скорова је висока (благо одступање постоји код кодека 4). Насупрот овим мерама, корелација PSNR вредности и субјективних скорова значајно флукутира – зависи од подкупа секвенци. Ово запажање је илустровано на Сл. 5.16, где су приказане средње вредности SROCC и LLC скорова из Таб. 5.5.

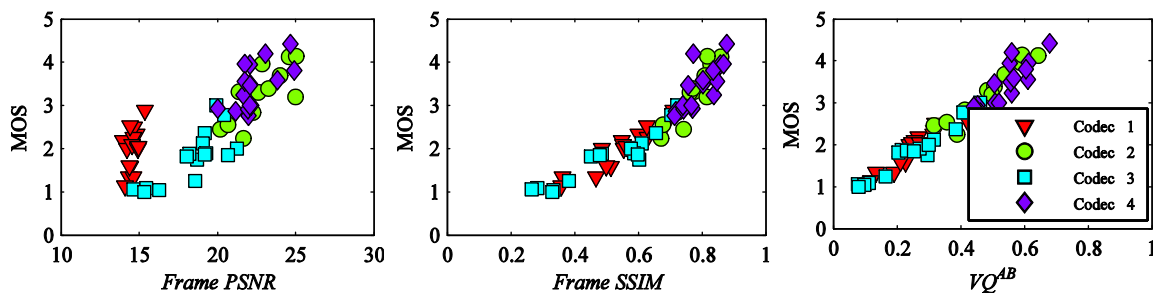


Сл. 5.16 Приказ резултата објективних мера из Таб. 5.5 (средње вредности по подкуповима и комплетан сет): (а) корелација рангова, SROCC, и (б) коефицијент линеарне корелације, LLC.

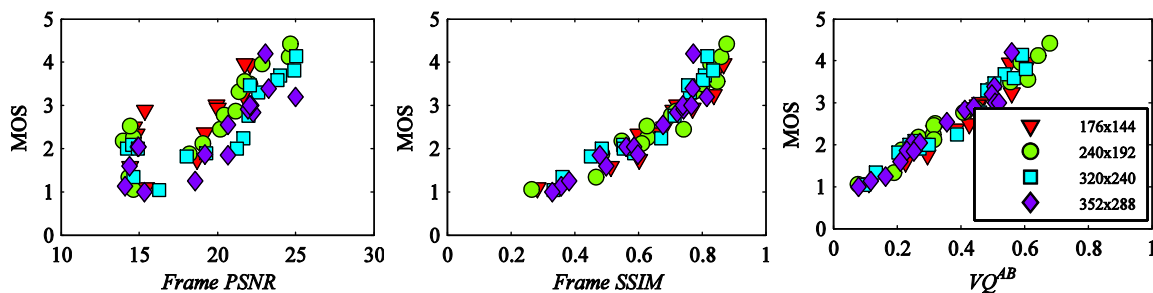
Резултати приказани у Таб. 5.5 су додатно илустровани у простору субјективних и објективних скорова на Сл. 5.17. Са слике се уочава да је коришћењем PSNR мере дошло до значајног груписања објективних скорова за кодек 1 (око 14 dB), иако се субјективни скорови налазе на интервалу од 1 (лош квалитет) до 3 (средњи квалитет). То је један од разлога лоших перформанси мере на глобалном нивоу. Груписање објективних скорова постоји и код кодека 4 (MP4 кодек), и то код све три мере, због чега су перформансе објективних мера Frame PSNR и VQ^{AB} лошије од перформанси на осталим кодецима (Таб. 5.5).

Један од разлога увођења објективних мера процене је да замене посматрача у евалуацији квалитета и да се одабере најбољи начин компресије уз ограничења која намеће канал за пренос (капацитет канала). На овој бази то би значило да се од четири кодека изврши њихово рангирање и одабере најбољи по питању квалитета видеа који даје на свом излазу. Ако се посматрају средње вредности субјективних

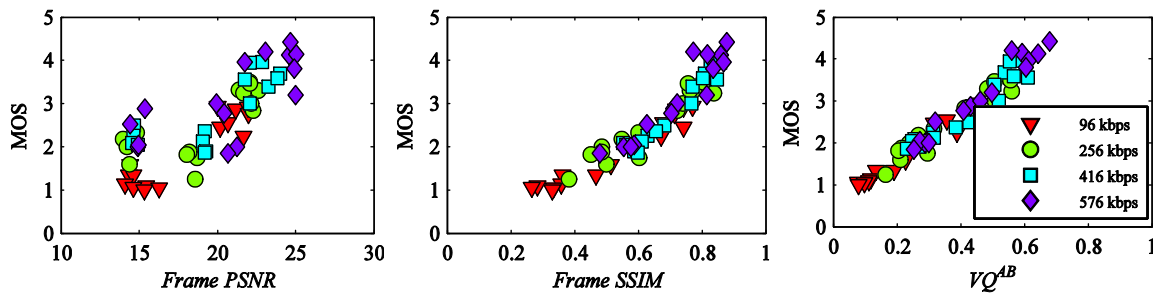
скорова, посматрачи су извршили рангирање у следећем редоследу: кодек 4 – кодек 2 – кодек 1 – кодек 3, Сл. 5.18(а). Овај ранг је очуван коришћењем објективних VQ^{AB} и SSIM скорова, док се значајно разликује од ранга добијеног коришћењем PSNR: кодек 2 – кодек 4 – кодек 3 – кодек 1.



(а) илустрација субјективних и објективних скорова према типу кодовања



(б) илустрација субјективних и објективних скорова према димензији кадра



(в) илустрација субјективних и објективних скорова према битском протоку

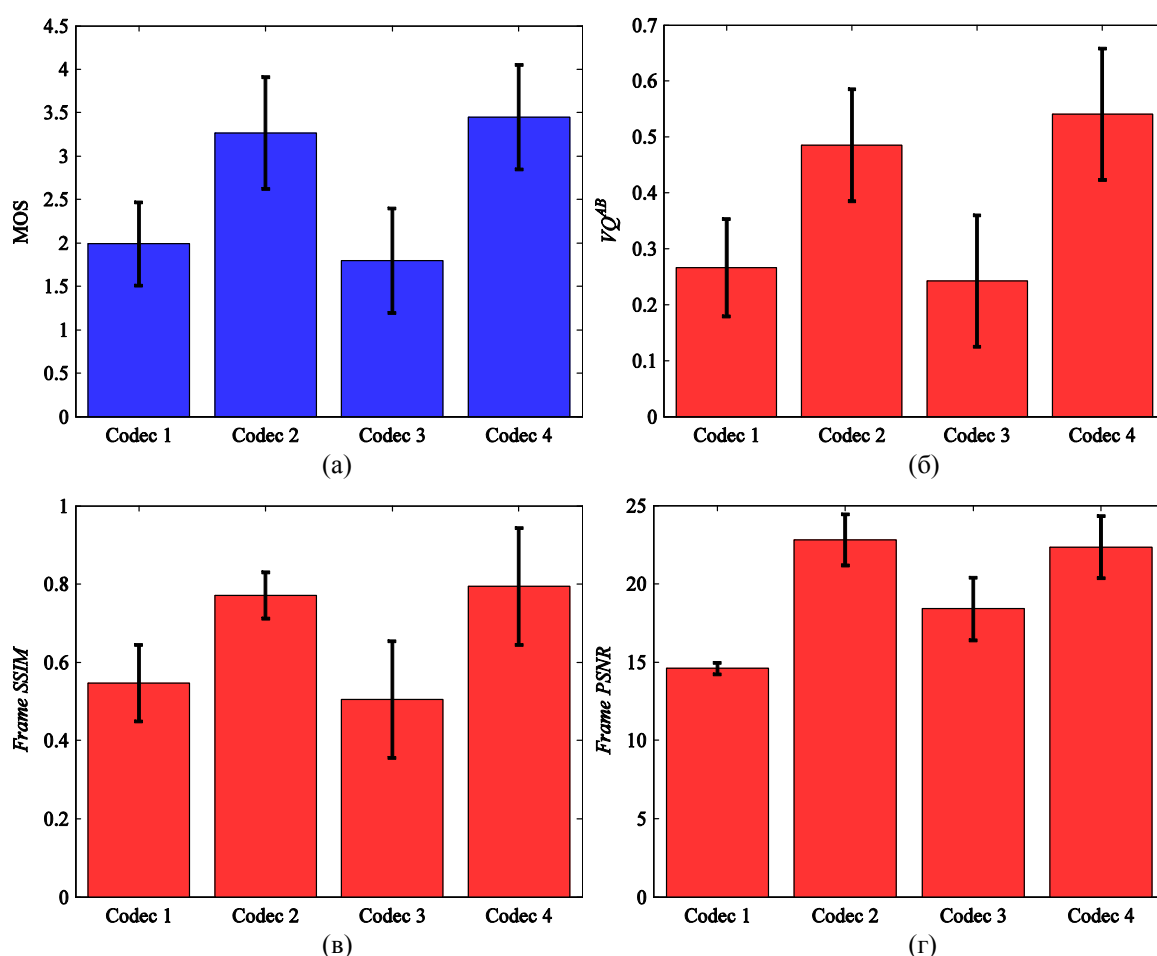
Сл. 5.17 Дијаграми расипања објективних и субјективних скорова за секвенце у UoM сету према различитим улазним параметрима: (а) типу компресије, (б) димензији кадра и (в) битском протоку.

5.9.2. Анализа резултата на FUB бази секвенци

У овој бази се користи шест оригиналних (стандардних) секвенци дужине 250 кадра: 'Coastguard', 'Container', 'Foreman', 'Japan football league', 'News' и 'Tempete' [112]. Илустративни кадрови оригиналних секвенци су дати на Сл. 5.19. Секвенце су

5. Процена квалитета видеа заснована на очувању информација о градијенту

компримоване користећи H.263 и H.264 енкодере, за битске брзине од 24 kbps до 384 kbps. Коришћене су две димензије кадра – QCIF (176x144) и CIF (352x288), при чему је број кадра мењан од 7.5 до 30 кадра у секунди. Користећи један тип енкодера, променом битске брзине, димензије кадра и броја кадра у секунди, добија се 15 тест секвенци, што се види из Таб. 5.6. Ова база није јавно доступна и након упућивања захтева ауторима из [112], добијен је део базе – 60 секвенци QCIF формата, кодovаних H.264 енкодером (шест изворних секвенци и по девет њихових модификација као што се види из Таб. 5.6).



Сл. 5.18 Средње вредности и стандардне девијације: (а) субјективних скорова, (б) објективних VQ^{AB} скорова, (в) објективних SSIM скорова и (г) објективних PSNR скорова, по кодовима коришћеним у UoM сету.

Таб. 5.6 Параметри коришћени при формирању FUB базе секвенци.

Димензија кадра	Број кадра	7.5 fps	15 fps	30 fps
CIF		64, 128 kbps	64, 128 kbps	128, 384 kbps
QCIF		24, 48, 64 kbps	24, 48, 64 kbps	48, 64, 128 kbps



Сл. 5.19 Приказ илустративних кадра изворних секвенци FUB сета.

Током субјективних тестова, све видео секвенце су приказане у највишој просторној и временској резолуцији, тј. у CIF формату и са 30 кадрава у секунди. Секвенце ниже резолуције су у одговарајући формат преведене кроз интерполацију и понављање кадрава. Аутори су користили скалу квалитета са пет нивоа, где вредности 1, 2, 3, 4 и 5 одговарају недовољном, довољном, добром, врло добром и одличном квалитету, тим редом. У субјективним тестовима је учествовало 20 посматрача (студенти универзитета, 10 момака и 10 девојака). За прикупљање субјективних скорова коришћен је DSIS II метод (Double Stimulus Impairment Scale, Variant II) [27, 28], при чему се прво приказује референтна а након ње тест секвенца. Пре давања оцене, посматрачима се у истом редоследу још једном приказују секвенце. Услови у којима су спроведени субјективни експерименти су усаглашени са препорукама ITU [27, 28].

На овој бази су анализирани перформансе три објективне мере, тј. перформансе MOVIE [102], MAD [103] и предложене мере VQ^{AB} . Анализирана је предикција скорова кроз тачност, монотоност и доследност [5].

5. Процена квалитета видеа заснована на очувању информација о градијенту

Скорови објективних мера на FUB бази секвенци су дати у Таб. 5.7 – Таб. 5.11. У табелама су подебљани и осенчени најбољи резултати две анализиране мере. Перформансе су дате за комплетан сет видео секвенци (укупно 54 секвенце) и подскупове секвенци са истим бројем кадрова у јединици времена (3 подскупа са по 18 секвенци).

Таб. 5.7 Коефицијент линеарне корелације (LCC) објективних и субјективних скорова на FUB бази секвенци.

<i>Мера</i>		<i>7.5 fps</i>	<i>15 fps</i>	<i>30 fps</i>	<i>All data</i>
MAD	Spatial	0.5115	0.7251	0.6207	0.5597
	Temporal	0.4935	0.4196	0.5218	0.3809
	Spatio-temporal	0.4935	0.4533	0.5410	0.4207
MOVIE	Spatial	0.8705	0.8742	0.8903	0.8474
	Temporal	0.3171	0.5263	0.7714	0.2905
	Spatio-temporal	0.6034	0.7268	0.8351	0.7345
VQ^{AB}		0.9045	0.9092	0.9495	0.8836

Таб. 5.8 Корелација рангова (SROCC) објективних и субјективних скорова на FUB бази секвенци.

<i>Мера</i>		<i>7.5 fps</i>	<i>15 fps</i>	<i>30 fps</i>	<i>All data</i>
MAD	Spatial	0.5933	0.6687	0.6322	0.5436
	Temporal	0.5246	0.2649	0.5227	0.3456
	Spatio-temporal	0.5439	0.3611	0.5475	0.3932
MOVIE	Spatial	0.8858	0.8588	0.8946	0.8626
	Temporal	0.2284	0.4524	0.6801	0.2838
	Spatio-temporal	0.6677	0.7305	0.8306	0.7434
VQ^{AB}		0.8879	0.8670	0.9287	0.8641

Таб. 5.9 Средња апсолутна грешка (MAE) између субјективних и објективних скорова на FUB бази секвенци.

<i>Мера</i>		<i>7.5 fps</i>	<i>15 fps</i>	<i>30 fps</i>	<i>All data</i>
MAD	Spatial	0.5550	0.4575	0.6915	0.6166
	Temporal	0.4844	0.6457	0.7215	0.6504
	Spatio-temporal	0.4844	0.6371	0.7008	0.6358
MOVIE	Spatial	0.2829	0.3483	0.3690	0.3820
	Temporal	0.5550	0.6910	0.5337	0.7099
	Spatio-temporal	0.4798	0.4921	0.4496	0.4925
VQ^{AB}		0.2631	0.2744	0.2337	0.3211

5. Процена квалитета видеа заснована на очувању информација о градијенту

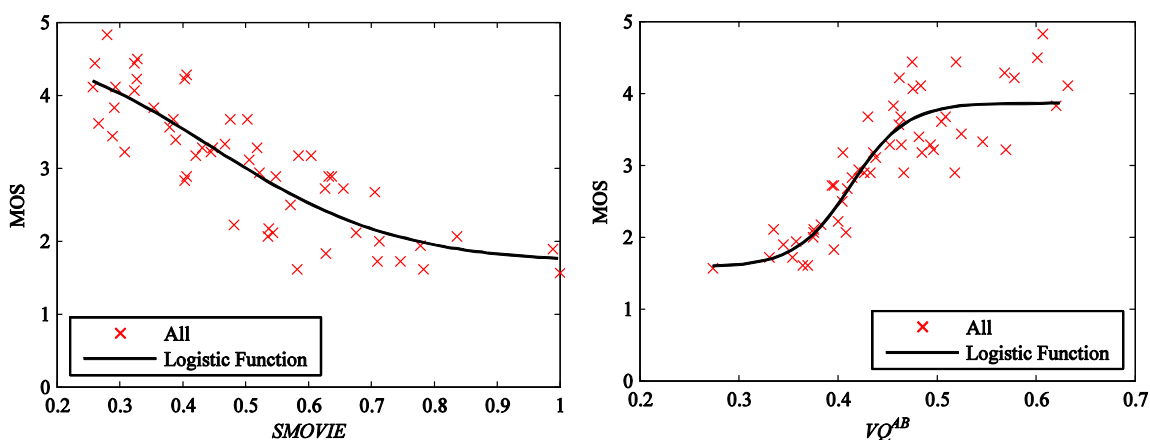
Таб. 5.10 Корен средње квадратне грешке (RMSE) између субјективних и објективних скорова на FUB бази секвенци.

<i>Мера</i>		<i>7.5 fps</i>	<i>15 fps</i>	<i>30 fps</i>	<i>All data</i>
MAD	Spatial	0.6843	0.6062	0.7960	0.7247
	Temporal	0.5952	0.7991	0.8661	0.8086
	Spatio-temporal	0.5952	0.7847	0.8539	0.7934
MOVIE	Spatial	0.3368	0.4275	0.4624	0.4643
	Temporal	0.6843	0.7509	0.6460	0.8368
	Spatio-temporal	0.5457	0.6047	0.5585	0.5935
VQ^{AB}		0.2918	0.3666	0.3187	0.4095

Таб. 5.11 Конзистентност (OR) објективних скорова на FUB бази секвенци.

<i>Мера</i>		<i>7.5 fps</i>	<i>15 fps</i>	<i>30 fps</i>	<i>All data</i>
MAD	Spatial	44.4444	38.8889	66.6667	55.5556
	Temporal	38.8889	55.5556	66.6667	50
	Spatio-temporal	38.8889	61.1111	61.1111	51.8519
MOVIE	Spatial	16.6667	33.3333	38.8889	31.4815
	Temporal	44.4444	66.6667	38.8889	59.2593
	Spatio-temporal	38.8889	44.4444	33.3333	50
VQ^{AB}		11.1111	27.7778	11.1111	22.2222

Из табела се види да су два најбоља објективна модела предложени VQ^{AB} модел и просторни MOVIE индекс, како на нивоу комплетне базе, тако и по сетовима секвенци са истим бројем кадрова у секунди. При томе се може рећи да је VQ^{AB} модел незнатно бољи од MOVIE индекса. На Сл. 5.20 је илустрована веза између субјективних података и објективних предикција за ове две мере, тј. приказани су дијаграми расипања субјективних и објективних скорова (MOVIE индекс има обрнуту скалу од VQ^{AB} мере, тј. код њега ниже вредности одговарају бољем квалитету).



Сл. 5.20 Дијаграми расипања субјективних и објективних (SMOVIE и VQ^{AB}) скорова (свака тачка одговара једној видео секвенци).

Пре провлачења интерполационе криве, вредности просторног MOVIE (SMOVIE) индекса су скалиране тако да максимална вредност буде једнака 1 (одговара лошем квалитету). На комплетном сету од 54 секвенце стварне вредности SMOVIE индекса се налазе у интервалу од 0.0068 до 0.0265, што је око 2 % од расположивог динамичког – теоретског опсега [0,1]. Сами аутори MOVIE мере су навели да добијене вредности генерално упадају у узан интервал због нормализације и коришћења нелинеарног модела и препоручили скалирање добијених вредности пре провлачења логистик криве. За исте секвенце вредности VQ^{AB} мере се налазе у опсегу од 0.2739 до 0.6324. Графици са Сл. 5.20 показују да код SMOVIE индекса постоји константно расипање објективних скорова око логистик криве, док је код VQ^{AB} мере расипање значајно мање код секвенци лошег и средњег квалитета.

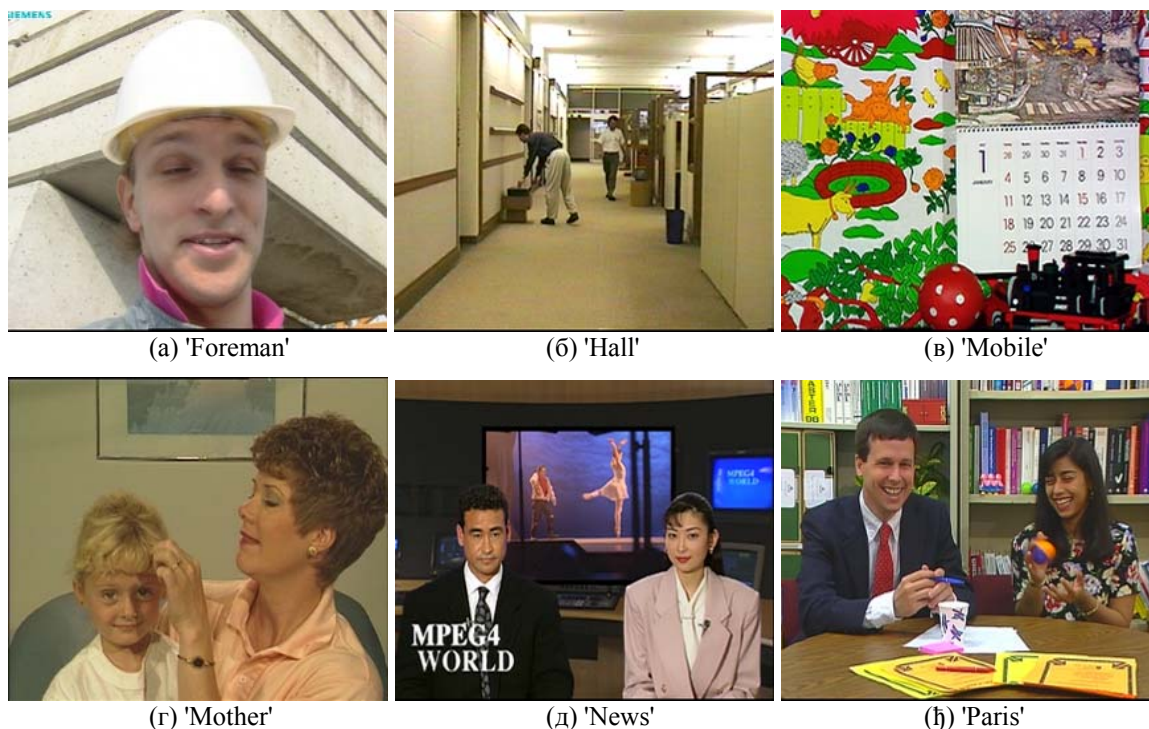
5.9.3. Анализа резултата на EPFL-PoliMI бази секвенци

Коришћење IP мрежа за испоруку мултимедијалних садржаја од стране провајдера ка бројним корисницима постаје све популарније. У случају видеа, пакетске мреже се на пример користе за дистрибуцију програма у IPTV апликацијама. У оваквим мрежама не постоји гаранција да ће садржај крајњем кориснику бити испоручен без грешака (гарантује се најбоља намера да се испоручи садржај). Са друге стране, интерес провајдера је да понуди сервисе који задовољавају очекивања корисника са становишта квалитета [255, 256].

Две академске институције – Политехнички институт у Милану (Politecnico di Milano – Italy) и Политехничка школа у Лозани (Ecole Polytechnique Fédérale de Lausanne – Switzerland) су учиниле доступном базу од 156 секвенци, кодованих са H.264/AVC енкодером и деградираних са пакетским губицима [252, 256]. Осим секвенци доступни су и резултати субјективних тестова рађених на обе академске институције. У субјективним тестовима учествовало је 40 посматрача. У овој дисертацији је коришћено 78 секвенци CIF просторне резолуције (352x288).

У субјективним експериментима разматрано је шест стандардних изворних видео секвенци CIF просторне резолуције које се користе у области обраде видеа: 'Foreman', 'Hall', 'Mobile', 'Mother', 'News' и 'Paris'. Изворне секвенце су доступне у

прогресивном формату са 30 кадрова у секунди. Секвенце су одабране тако да имају различите нивое просторне и временске комплексности. На Сл. 5.21 су приказани илустративни кадрови изворних секвенци.



Сл. 5.21 Приказ илустративних кадрова изворних секвенци EPFL-PoliMI сета (CIF подскуп).

За сваку од шест оригиналних H.264/AVC поворки бита, генерисане су деградиране поворке кроз одбацавање пакета. Одабрано је шест нивоа пакетских губитака (PLR), процентуално 0.1%, 0.4%, 1%, 3%, 5% и 10%. За сваки од нивоа пакетских губитака изабране су по две реализације настале одбацавањем различитих пакета, тако да се укупно добија 12 реализација по оригиналном садржају. На овај начин је добијен униформан опсег дисторзија, тј. субјективног утиска квалитета, при чему је број секвенци у бази задовољавајући. Параметри квантизације су пажљиво одабрани тако да се за сваку изворну секвенцу добије одличан визуелни квалитет при одсуству пакетских губитака. Перформансе оригиналних кодованих поворки бита су дате у Таб. 5.12. Укратко, параметри квантизације за сваку од секвенци су подешени тако да се не пређе битски проток од 600 kbps. Ова граница се може сматрати горњом границом за пренос видео садржаја CIF формата преко IP мрежа.

Таб. 5.12 Перформансе оригиналних кодованих поворки бита из EPFL-PoliMI базе секвенци.

<i>Име секвенце</i>	<i>Битски проток [kbps]</i>	<i>PSNR [dB]</i>
'Foreman'	353	34.35
'News'	283	37.27
'Mobile'	532	28.29
'Mother'	150	37.03
'Hall'	216	36.16
'Paris'	480	33.64

На Сл. 5.22 илустровано је како различите вредности пакетских губитака утичу на визуелни квалитет. Слика приказује кадрове секвенце 'Hall' за различите вредности PLR. Може се уочити да дисторзија слике нужно није повезана са нивоом пакетских губитака, тј. утисак квалитета значајно зависи од локације изгубљених пакета на слици (погледати у упоредити кадрове при $PLR = 3\%$ и $PLR = 5\%$).

У току субјективних тестова коришћен је SS (Single Stimulus) метод, при чему се презентује само тест секвенца без приказа оригинала. Субјективни тест обухвата приказ 83 видео секвенце: 5 секвенци за обучавање посматрача (5 реализација које одговарају различитим нивоима квалитета), 6 изворних секвенци (без пакетских губитака) и 72 деградираних секвенце (6 изворних x 12 тест секвенци).

Како је база формирана са циљем анализе пакетских губитака у IP мрежама, деградације су углавном локализоване у појединачним кадровима. На овој бази евалуиране су три објективне мере – PSNR, SSIM и предложена мера VQ^{AB} , а резултати евалуације су дати у Таб. 5.13.

Таб. 5.13 Перформансе објективних мера на комплетном EPFL-PoliMI сету.

<i>Мера процене</i>	<i>LCC</i>	<i>SROCC</i>	<i>MAE</i>	<i>RMSE</i>	<i>OR (%)</i>
<i>Frame PSNR</i>	0.6682	0.6619	0.8220	1.0034	65.3846
<i>Frame SSIM</i>	0.7735	0.7089	0.7019	0.8548	53.8462
<i>VQ^{AB}</i>	0.7909	0.7995	0.6804	0.8253	61.5385

Резултати из Таб. 5.13 показују да је предложена мера боља од PSNR, што је било и за очекивати. Тачност предложене мере (LCC, MAE, RMSE) је незнатно боља од SSIM, док је монотоност предикције значајно боља (SROCC).

Дијаграми расипања објективних, VQ^{AB} , и субјективних скорова приказани су на Сл. 5.23. Дата су два дијаграма расипања, први на коме је приказана оптимална

5. Процена квалитета видеа заснована на очувању информација о градијенту

интерполациона функција и други на коме су обележене вредности које потичу од различитих изворних секвенци.



(а) оригинал



(б) компресија без пакетских губитака, PLR = 0 %



(в) PLR = 0.1%



(г) PLR = 0.4%



(д) PLR = 1%



(е) PLR = 3%

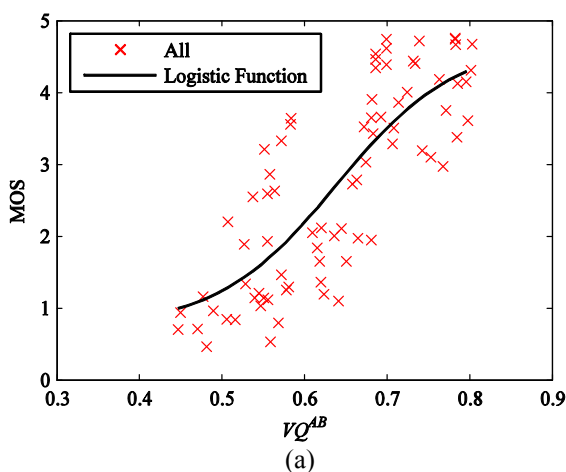


(e) PLP = 5%

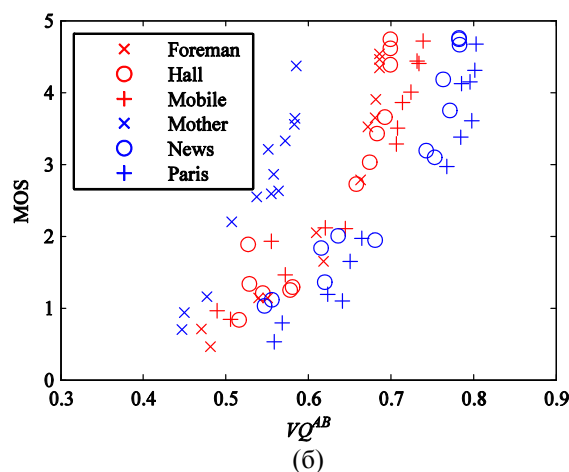


(ж) PLR = 10%

Сл. 5.22 Дисторзије добијене у преносу компримоване секвенце кроз канал са различитим нивоима пакетских губитака.



(a)



(б)

Сл. 5.23 Дијаграми расипања субјективних и објективних VQ^{AB} скорова (свака тачка одговара једној видео секвенци).

Са Сл. 5.23 уочава се линеарна веза између субјективних и објективних скорова као и релативно добро раздвајање објективних скорова по различитим изворним садржајима. Предложена мера јасно прати тренд квалитета процењеног од стране посматрача, што се уочава кроз рангирање тест секвенци које потичу од истог оригинала – раздвајање тачака које потичу од исте изворне секвенце доводи до правилног рангирања, Сл. 5.23(б). Ако се погледају вредности субјективних скорова, може се закључити да је истраживање исправно спроведено јер су субјективни скорови униформни у комплетном опсегу квалитета. Субјективне импресије квалитета покривају скоро комплетан опсег вредности $[0,5]$, док је код објективних мера присутно груписања објективних скорова: $VQ^{AB} \in [0.45, 0.80]$, $SSIM \in [0.73, 0.94]$ и $PSNR \in [19.50, 34.15]$.

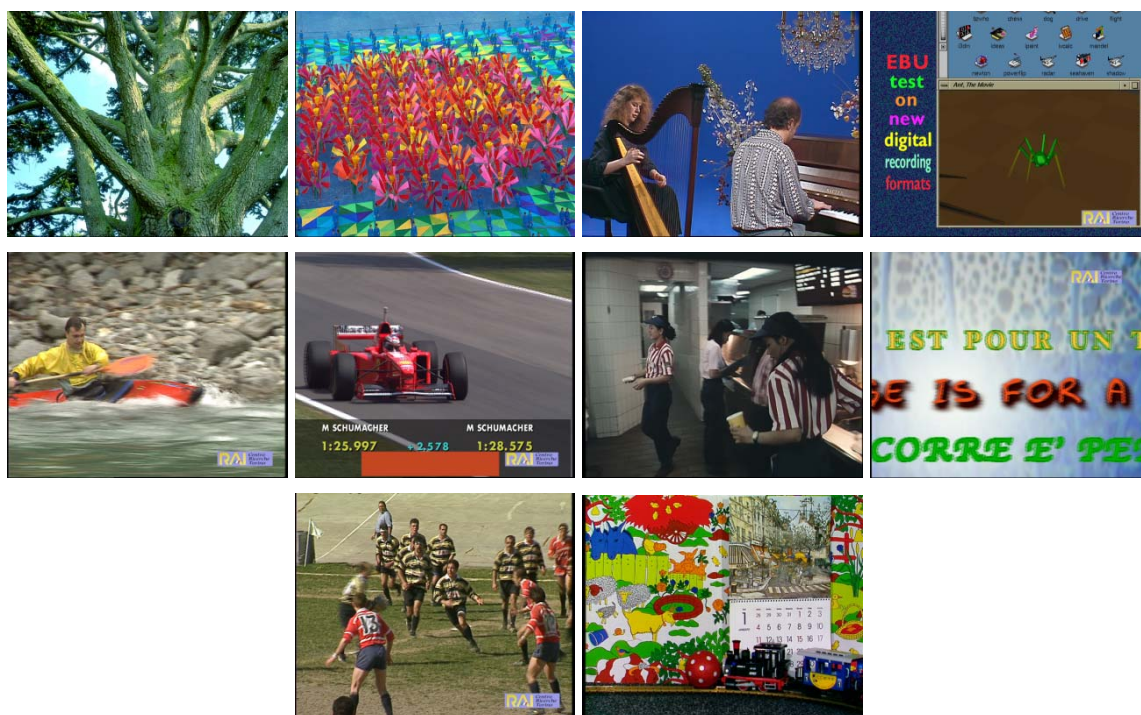
5.9.4. Анализа резултата на VQEG бази секвенци

У октобру 1997. године формирана је VQEG (Video Quality Experts Group) група са циљем да се врши евалуација нових мера за процену квалитета видеа и како би се помогло различитим ITU групама у писању препорука у овој области [213]. Након једног од састанака 1998. године (National Institute of Standards and Technology, Gaithersburg, USA), група је дефинисала оквирни план и процедуре за један опширни тест за евалуацију објективних мера. Циљ планираних активности је био да се изврши евалуација перформанси мера које пореде изворне и тест видео сигнале, тј. да се анализирају перформансе мера са потпуним референцирањем (“full-reference” или “double-ended” мере). FR мере, уз потпуно познате изворне секвенце, могу достићи високу корелацију са субјективним скоровима, који су прикупљени помоћу DSCQS поступка, описаног у [27].

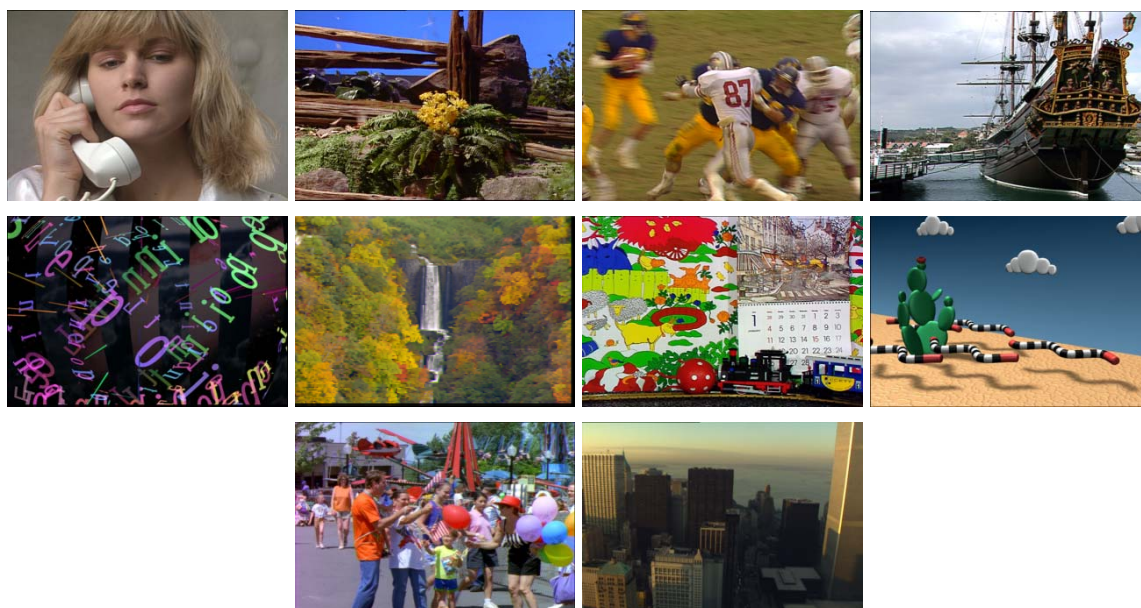
Десет учесника је поднело захтев за евалуацију својих мера објективне процене. Осим ових мера, VQEG група је укључила и PSNR као референтну објективну меру. Сваки од учесника је предао по једну меру за одређивање објективних скорова и поређење са субјективним импресијама на великом броју видео секвенци. Прикупљено је преко 26000 субјективних мишљења на 20 различитих изворних секвенци, обрађиваних на 16 различитих видео система и евалуираних у 8 независних лабораторија широм света. Субјективни тестови су подељени у четири сесије: 50 Hz/високи квалитет, 50 Hz/ниски квалитет, 60 Hz/високи квалитет и 60 Hz/ниски квалитет. Високи квалитет у овом контексту се односи на пренос у TV дифузији (broadcast) док ниски квалитет одговара квалитету у дистрибуцији (distribution). Битске брзине секвенци из сесија високог квалитета се налазе у опсегу од 3 Mb/s до 50 Mb/s, док су битске брзине секвенци сесија ниског квалитета од 768 kb/s до 4.5 Mb/s.

На Сл. 5.24 и Сл. 5.25 приказани су илустративни кадрови изворних секвенци VQEG сета. Из приказаних кадрова се јасно уочава да је овај сет примарно намењен апликацијама у TV дифузији, како по садржајима које приказује и димензијама слика, тако и по крајњем квалитету.

5. Процена kvaliteta videa zasnovana na ocuvanju informacija o gradijentu



Сл. 5.24 Примери кадрова изворних секвенци из VQEG Phase I сета – 625 линија/50 Hz.

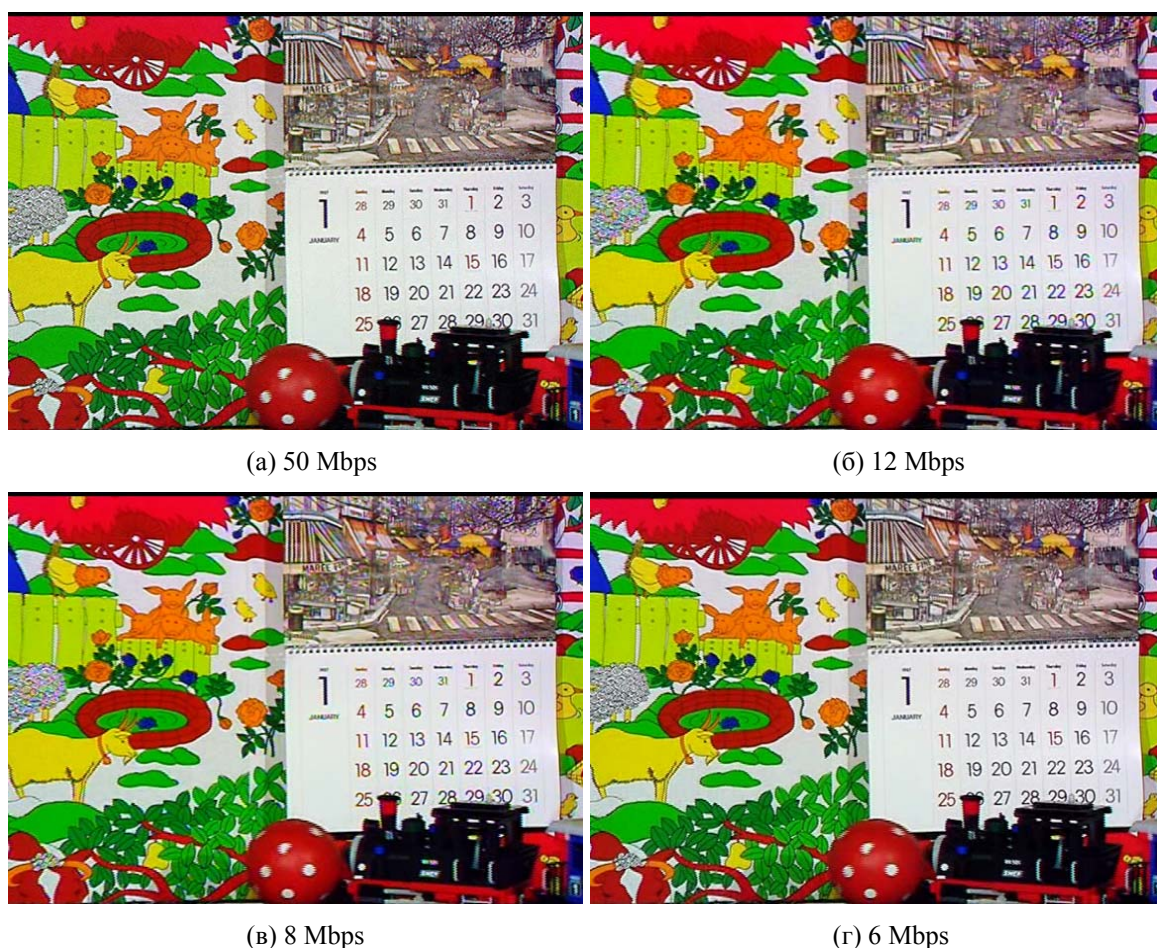


Сл. 5.25 Примери кадрова изворних секвенци из VQEG Phase I сета – 525 линија/60 Hz.

Изворне секвенце су деградиране кроз компресију и прилагођење преносу са различитим битским брзинама. На Сл. 5.26 и Сл. 5.27 приказани су кадрови једне од изворних секвенци из различитих верзија видеа – видеа високог и ниског квалитета.

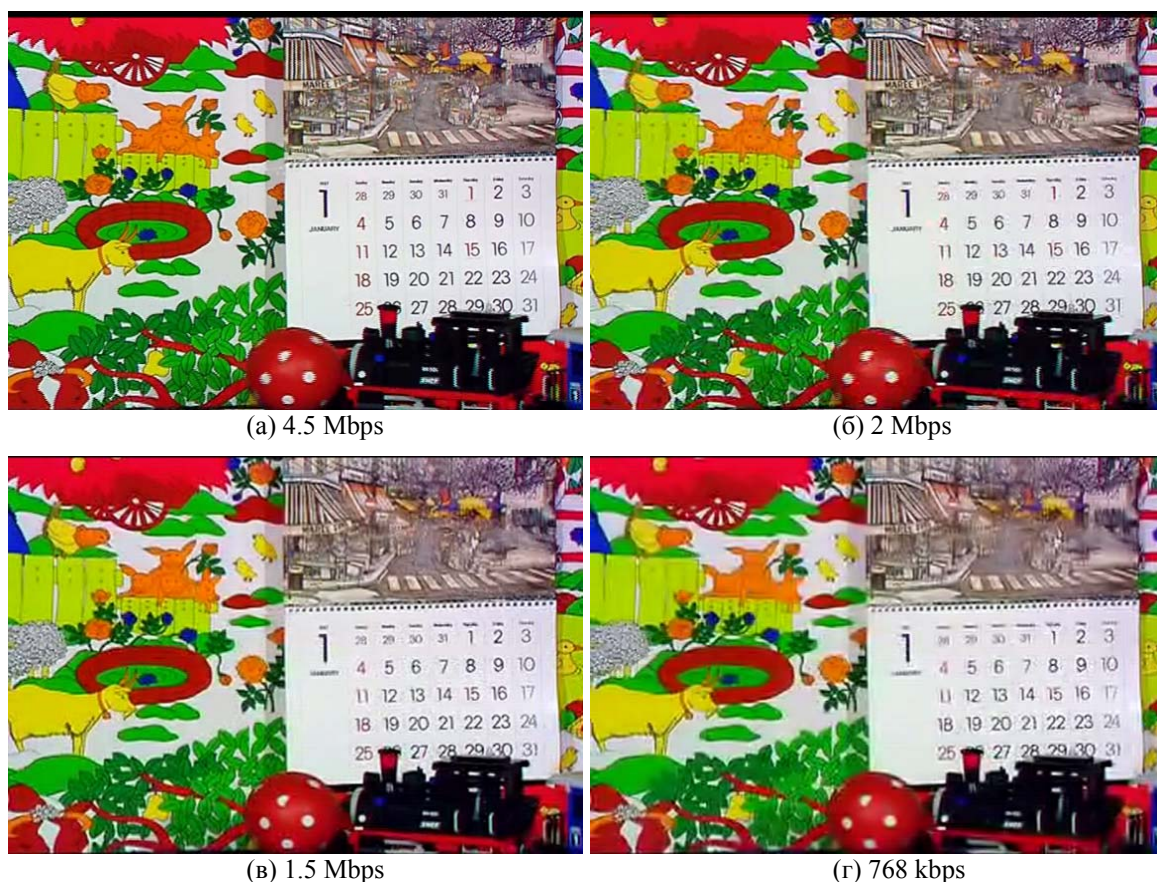
5. Процена квалитета видеа заснована на очувању информација о градијенту

Код секвенци високог квалитета приказани су кадрови секвенци са битским протоцима од 50Mbps до 6 Mbps, док су за секвенце ниског квалитета приказани кадрови секвенци са битским протоцима од 4.5 Mbps до 768 kbps.



Сл. 5.26 Примери кадрова компримованих тест секвенци са различитим битским брзинама – сесија са секвенцама високог квалитета.

Са Сл. 5.26 се јасно уочава лоша перцептуална сепарација видеа са различитим битским брзинама. Веома тешко је детектовати и идентификовати било какву деградацију у било којој од секвенци, иако се битске брзине значајно разликују. Слично запажање се може извести и за Сл. 5.27, где је такође тешко уочити постојање деградације код кадра секвенце са највишим битским протоком – 4.5 Mbps.



Сл. 5.27 Примери кадрова компримованих тест секвенци са различитим битским брзинама – сесија са секвенцама ниског квалитета.

Током прве фазе реализованих истраживања утврђено је да осам објективних мера има перформансе које су статистички еквивалентне. Перформансе ових мера су, што је било изненађујуће, еквивалентне перформансама PSNR [257, 258]. Могући разлог за овај резултат могу бити непоуздани субјективни скорови због лоше перцептуалне сепарације видеа, а што је показано кроз два претходна примера (Сл. 5.26 и Сл. 5.27).

У Таб. 5.14 дате су перформансе мера објективне процене на VQEG Phase I бази секвенци. Узете су у обзир секвенце из све четири сесије, а са *P0-P9* означене су мере које су анализирани у VQEG извештајима, из којих су и преузете дате вредности. Из Таб. 5.14 се види да се PSNR мера понаша скоро исто као већина мера, укључујући и овде предложену VQ^{AB} меру. Ово је резултат лоше перцептуалне сепарације видео записа у овој бази а што је илустровано раније. Генерално,

резултати VQ^{AB} мере су у нивоу резултата бољих мера чије су перформансе статистички еквивалентне.

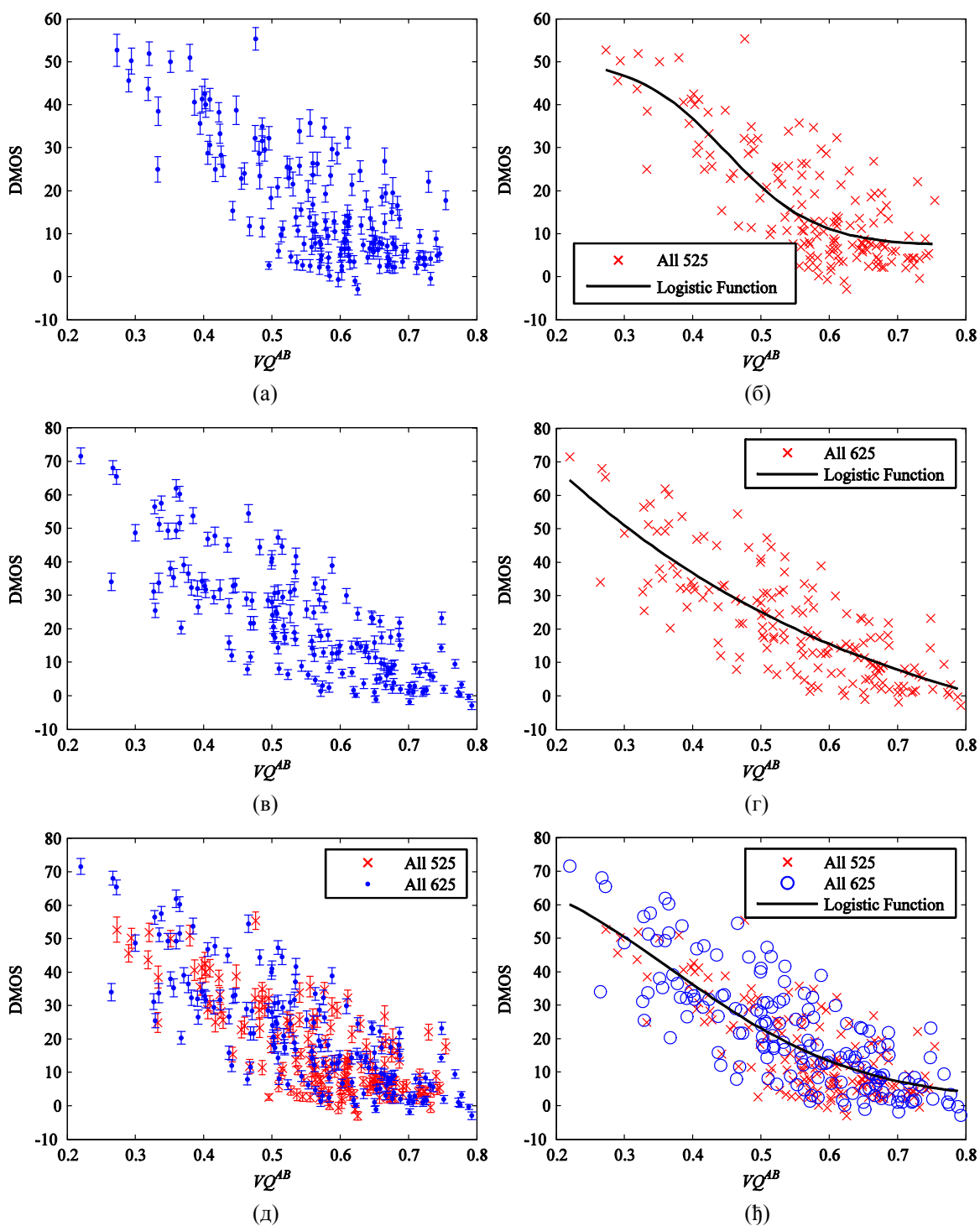
Таб. 5.14 Поређење перформанси мера објективне процене на VQEG Phase I сету (узете су у обзир секвенце из све четири сесије).

<i>Мера процене</i>	<i>LCC</i>	<i>SROCC</i>	<i>OR</i>
<i>P0 (PSNR)</i>	0.779	0.786	0.678
<i>P1 (CPqD)</i>	0.794	0.781	0.650
<i>P2 (Tektronix/Sarnoff)</i>	0.805	0.792	0.656
<i>P3 (NHK/Mitsubishi)</i>	0.751	0.718	0.725
<i>P4 (KDD)</i>	0.624	0.645	0.703
<i>P5 (EPFL)</i>	0.777	0.784	0.611
<i>P6 (TAPESTRIES)</i>	0.310	0.248	0.844
<i>P7 (NASA)</i>	0.770	0.786	0.636
<i>P8 (KPN/Swisscom CT)</i>	0.827	0.803	0.578
<i>P9 (NTIA)</i>	0.782	0.775	0.711
<i>VQ^{AB}</i>	0.792	0.729	0.662

На Сл. 5.28 илустровани су резултати објективне VQ^{AB} мере на подскуповима и комплетном VQEG Phase I сету. Дијаграми расипања приказују расподеле субјективних и објективних скорова. Са леве стране дати су интервали поузданости субјективних скорова (стандардна грешка), а са десне стране су приказане логистик криве у простору субјективно-објективних скорова. Глатке криве логистик функција говоре о доброј вези субјективних и објективних скорова.

5.9.5. Анализа резултата на LIVE Video бази секвенци

Са циљем да се превазиђу проблеми који су уочени код VQEG базе (анализа слике са проредом (у полусликама), лоша перцептуална сепарација видеа и коришћење енкодера старије генерације), LIVE група је направила базу видео секвенци – LIVE Video Quality Database [250] и учинила је јавно доступном. Ова база, између осталих, обухвата компримоване видео секвенце поступком H.264, као и видео секвенце добијене преносом H.264 видео секвенци кроз комуникациони канал са губицима.

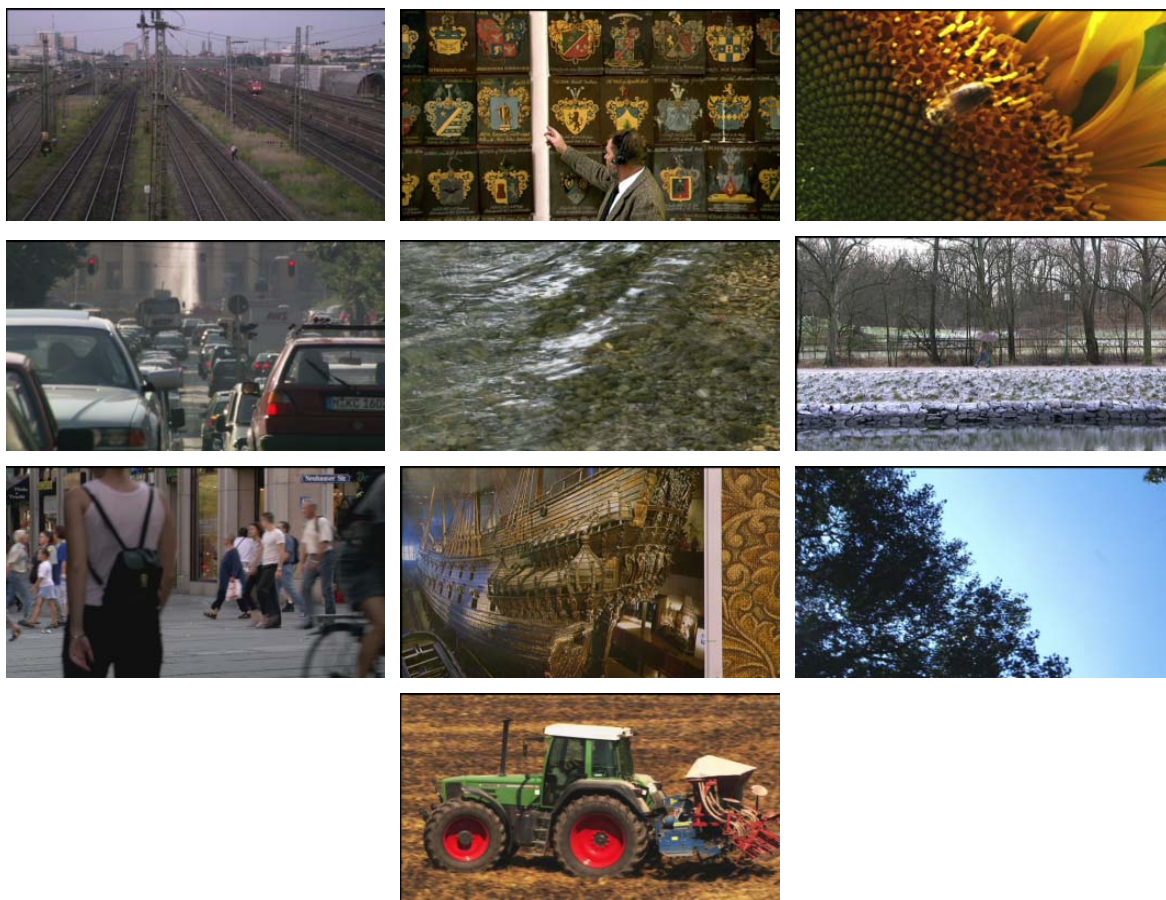


Сл. 5.28 Графички приказ резултата предложене VQ^{AB} мере на VQEG Phase I сету - дијаграми расипања за: (а) 525/60Hz сет, (в) 625/50Hz сет и (д) комплетну базу, и прикази одговарајућих логистик функција (б), (г) и (е).

LIVE база секвенци је формирана полазећи од 10 некомпримованих секвенци високог квалитета са различитим садржајима. Формиран је сет од 150

5. Процена квалитета видео заснована на очувању информација о градијенту

модификованих оригиналних секвенци (15 модификованих секвенци по свакој изворној) користећи четири типа деградације – MPEG-2 компресија (MPEG-2 сет), H.264 компресија (H.264 сет) и симулација трансфера H.264 компримованих поворки бита кроз IP (IP сет) и бежичне мреже (Wireless сет). Степен деградације је тако подешаван да се осигура перцептуално раздвајање видео записа. У субјективним тестовима учествовало је 38 посматрача кроз SS (Single Stimulus) поступак евалуације. Квалитет видео је даван на континуалној скали. Посматрачима су приказиване и референтне секвенце како би се субјективни скорови превели у форму диференцијалног MOS скорa (DMOS) (елиминација скривене референце). Оцене непоузданих посматрача нису узете у обзир, тј. поступило се по поступку описаном у [27]. Субјективне импресије (средња вредност и варијанса DMOS скорова) су доступне заједно са изворним и деградираним секвенцама, а детаљан опис базе се може наћи у [259]. Примери кадрова изворних секвенци су дати на Сл. 5.29.



Сл. 5.29 Примери кадрова изворних секвенци LIVE Video сета.

Видео записи из ове базе су у прогресивном формату (слика по слика) и омогућавају истраживачима да се фокусирају на развој алгоритама објективне процене квалитета видеа. База покрива доста широк опсег квалитета – секвенце ниског квалитета су изабране да буду сличног квалитета као при преносу видеа на интернету (YouTube, бежични видео, пренос видеа кроз канал ограниченог капацитета уживо и сл.).

Резолуција видео записа је 768x432 пиксела. Ова резолуција је изабрана како би се очувао аспектни однос HD резолуције и минимизирале визуелне деградације. Додатно, ова резолуција обезбеђује да је број редова и колона умножак броја 16, што често захтевају системи са компресијом као што је MPEG-2.

Системи са компресијом као што су MPEG-2 и H.264 углавном дају униформну дисторзију/квалитет у видеу, како просторно тако и временски. Мрежни губици, са друге стране, условљавају краткотрајне дисторзије у видеу, и просторно и временски. Видео записи који се добијају симулацијама преноса кроз бежичну мрежу са губицима садрже грешке које су ограничене на мале регионе у оквиру кадра, док видео записи који се добијају симулацијом пакетског преноса са губицима обухватају веће регионе у оквиру кадра. Грешке у бежичним и IP мрежама су и временски краткотрајне и појављују се као треперења у видеу. На Сл. 5.30 приказани су кадрови секвенце 'Pedestrian area' деградирани са сва четири типа деградације коришћених у LIVE Video бази.

Са слике се види да су визуелни ефекти дисторзија различити за сваки од кадрова. Видео секвенце са компресијом имају типичне компресионе артефакте, као што су блоковски ефекат, замрљање, ефекат прстена и лоша компензација покрета око ивица објеката. Занимљиво је такође и уочити разлику деградација насталих MPEG-2 и H.264 компресијама, где је блоковски ефекат у великој мери редукован код H.264 секвенце. Грешке у бежичним и IP мрежама су временски краткотрајне и појављују се као изненадне појаве у видеу. Скоро половина видео секвенци из ове базе садржи просторно-временски локализоване дисторзије, док VQEG Phase I база садржи углавном компримоване видео секвенце и неколицину видеа са локализованим просторно-временским дисторзијама.



Сл. 5.30 Примери деградација унетих у изворну видео секвенцу LIVE базе.

На LIVE Video бази секвенци анализиране су перформансе предложене мере и упоређене са резултатима јавно доступних објективних модела. Коришћене су две метрике – Спирманова корелација рангова (SROCC), кроз коју се мери монотоност предикција објективног модела према субјективним скоровима и коефицијент линеарне корелације (LCC) кроз који се мери тачност предикција. Из рада [259] су преузети резултати евалуације 10 објективних мера и презентовани су заједно са резултатима евалуације предложене мере у Таб. 5.15 и Таб. 5.16 (перформансе алгоритма који дао најбоље резултате су подебљане). Дати су резултати евалуације мера које су већ описане (коришћене) – PSNR, SSIM, MS-SSIM, VSNR, VQM и MOVIE мере. Изузимајући VQM и MOVIE индекс, за остале набројане мере процене квалитета врши се по кадровима у каналу интензитета, а након чега се крајња мера добија усредњавањем по анализираним кадровима. Додатно су дате и перформансе мера:

- *Speed SSIM* – мера процене квалитета која користи једновеличински SSIM индекс узезан са статистичким моделима визуелне перцепције [259].

5. Процена квалитета видеа заснована на очувању информација о градијенту

- *V-VIF* – објективна мера процене квалитета која представља проширење VIF мере дефинисане за појединачне слике, прилагођене за потребе процене квалитета видеа коришћењем временских градијената [259].

Таб. 5.15 Поређење перформанси алгоритама објективне процене видеа на LIVE Video бази – анализа SROCC за сваку од деградација и базу у целини.

Мера процене	Wireless	IP	H.264	MPEG-2	Комплетан сет
<i>Frame PSNR</i>	0.4334	0.3206	0.4296	0.3588	0.3684
<i>Frame SSIM</i>	0.5233	0.4550	0.6514	0.5545	0.5257
<i>MS-SSIM</i>	0.7285	0.6534	0.7051	0.6617	0.7361
<i>Speed SSIM</i>	0.5630	0.4727	0.7086	0.6185	0.5849
<i>VSNR</i>	0.7019	0.6894	0.6460	0.5915	0.6755
<i>VQM</i>	0.7214	0.6383	0.6520	0.7810	0.7026
<i>V-VIF</i>	0.5507	0.4736	0.6807	0.6116	0.5710
<i>Spatial MOVIE</i>	0.7934	0.6828	0.7023	0.6967	0.7262
<i>Temporal MOVIE</i>	0.6475	0.6000	0.7257	0.8043	0.7029
<i>MOVIE</i>	0.7931	0.6712	0.7158	0.7442	0.7397
<i>VQ^{AB}</i>	0.6734	0.5942	0.7696	0.8223	0.6330

Таб. 5.16 Поређење перформанси алгоритама објективне процене видеа на LIVE Video бази – анализа LCC за сваку од деградација и базу у целини.

Мера процене	Wireless	IP	H.264	MPEG-2	Комплетан сет
<i>Frame PSNR</i>	0.4675	0.4108	0.4385	0.3856	0.4035
<i>Frame SSIM</i>	0.5401	0.5119	0.6656	0.5491	0.5444
<i>MS-SSIM</i>	0.7170	0.7219	0.6919	0.6604	0.7441
<i>Speed SSIM</i>	0.5867	0.5587	0.7206	0.6270	0.5962
<i>VSNR</i>	0.6992	0.7341	0.6216	0.5980	0.6896
<i>VQM</i>	0.7325	0.6480	0.6459	0.7860	0.7236
<i>V-VIF</i>	0.5488	0.5102	0.6911	0.6145	0.5756
<i>Spatial MOVIE</i>	0.7876	0.7372	0.7146	0.6604	0.7435
<i>Temporal MOVIE</i>	0.6564	0.6070	0.7556	0.8016	0.7100
<i>MOVIE</i>	0.8075	0.7223	0.7459	0.7028	0.7606
<i>VQ^{AB}</i>	0.6863	0.6700	0.7611	0.8260	0.6378

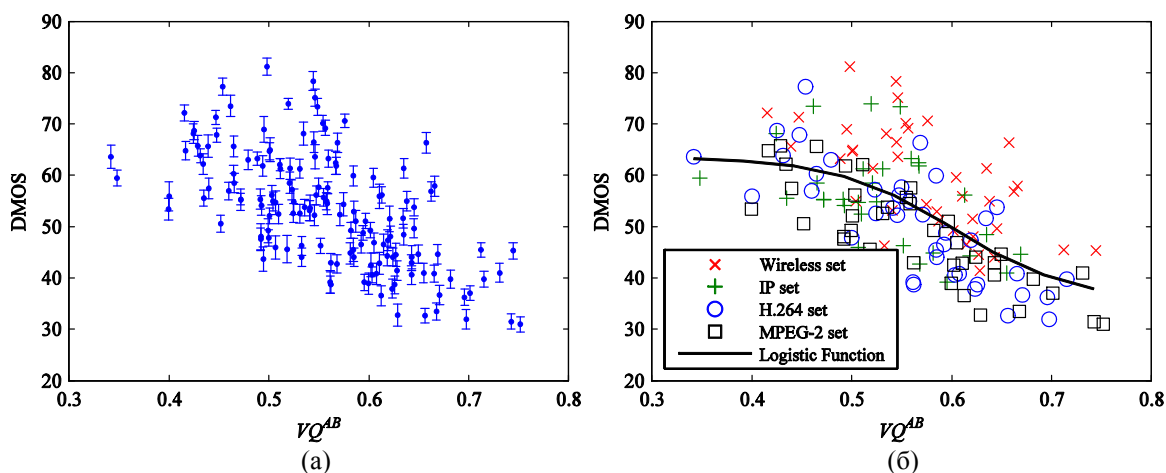
Из Таб. 5.15 и Таб. 5.16 треба приметити да скорови по типовима деградација нису усредњени како би се добио крајњи скор на бази. Уместо усредњавања, LCC и SROCC су оцењени на комплетном сету секвенци, што захтева рангирање знатно већег броја видео записа него код индивидуалних деградација.

Захваљујући slabим резултатима на деградацијама насталим током бежичног и IP преноса, перформансе предложене мере на комплетном сету нису најбоље (и по ранговима видео секвенци и по предикцији тачности). Међутим, предложена мера даје најбоље резултате на подацима добијеним компресијом (како по H.264 тако и по

MPEG-2). Овакав раскорак резултата на појединачним сетовима и бази у целини указује да је потребно кроз побољшање перформанси појединачних деградација побољшати перформансе на глобалном нивоу.

Дијаграми расипања субјективно-објективних скорова за VQ^{AB} меру на комплетном LIVE Video сету, са интервалима поузданости субјективних скорова и логистик кривом, су дати на Сл. 5.31. Логистик крива објективних VQ^{AB} вредности је глатка и скоро линеарна, са значајним расипањем око логистик криве. Расипање скорова је нарочито изражено код Wireless и IP сетова.

Може се закључити да предложена мера добро евалуира просторно хомогене дисторзије, а има проблема са локализованим дисторзијама унутар кадрова, што може бити правац даљих унапређења.



Сл. 5.31 Дијаграми расипања објективних VQ^{AB} и субјективних скорова за све видео секвенце у LIVE Video бази: (а) приказ са интервалима поузданости субјективних скорова, (б) приказ са оптималном логистик кривом.

Резултати такође показују да се правилним избором секвенци у бази могу изложити ограничења PSNR као мере објективног квалитета видеа, а што није уочено током анализе секвенци из VQEG базе ($LCC_{PSNR}=0.4035$, $SROCC_{PSNR}=0.3684$). Перформансе свих тестираних алгоритама су боље од перформанси PSNR. На изабраном скупу тестираних објективних мера, и по LCC и SROCC, се издваја MOVIE индекс. Овај индекс, који је комбинација просторног и временског MOVIE индекса, се понаша доследно на свим типовима дисторзије. Интересантно је да је просторни MOVIE индекс бољи код бежичног и IP преноса, док је временски MOVIE индекс бољи код преостале две категорије (H.264 и MPEG-2). Такође, из ове

анализе се може потврдити да коришћење информација о кретањима на сцени може побољшати перформансе алгоритма процене квалитета видеа. На пример, Speed SSIM има боље перформансе од простог, једновеличинског индекса структурне сличности рачунатог на нивоу појединачних кадрова (Frame SSIM).

5.9.6. Анализа резултата на LIVE Wireless бази секвенци

Како би се евалуирале перформансе алгоритама објективне процене квалитета видеа за специфичан циљ – пренос H.264/AVC компримованог видеа преко бежичних мрежа, LIVE група је спровела обимно истраживање у коме су прикупљени субјективни подаци и анализирали објективне мере процене користећи H.264 компримоване видео секвенце и симулацију грешака током преноса. У субјективним тестовима коришћено је 160 тест секвенци, које су тако одабране да имају добру перцептуалну сепарацију. H.264 компресија се широко користи у технологијама бежичног преноса.

Изворне видео секвенце су у некомпримованом прогресивном YUV420 формату, са резолуцијом од 768x480 пиксела и са 30 кадрова у секунди. Секвенце су добијене од компаније *Boeing*. Од опширне колекције добијених видеа, одабране су референтне секвенце које имају различита кретања објеката и људи. Неки од видеа су добијени у ноћним условима. Узето је 10 изворних секвенци трајања по 10 s, односно са по 300 кадрова [260]. Примери кадрова изворних секвенци LIVE Wireless базе приказани су на Сл. 5.32.

Свакој од изворних секвенци одговара по 16 модификованих, које су добијене комбинацијама 4 различите битске брзине и 4 различита нивоа пакетских губитака. Коришћене су битске брзине од 500 kbps, 1 Mbps, 1.5 Mbps и 2 Mbps. Пренос компримованих видео секвенци је симулиран за пакетске губитке од 0.5%, 2%, 5% и 17%. На Сл. 5.33 приказани су кадрови секвенце компримоване за пренос од 2 Mbps и деградирани са различитим нивоима пакетских губитака.

Дисторзије које се јављају у секвенцама ове базе су реалне и добро одсликавају природу реалних канала за пренос информација. Међутим, LIVE Wireless база је повучена од стране аутора, тј. више није јавно доступна. Аутори су

се на то одлучили из више разлога: (1) прављења веће и репрезентативније базе са већим бројем различитих изворних садржаја, (2) садржаји које приказују секвенце су веома слични (спора и једноставна кретања објеката у и око хангара за смештај авиона) и (3) због једноставних и сличних садржаја корелација објективних и субјективних скорова се креће око 95-96%, па није могуће на прави начин анализирати перформансе објективних мера. Као алтернативу аутори су предложили коришћење LIVE Video базе.



Сл. 5.32 Примери кадрова изворних секвенци LIVE Wireless базе.

На LIVE Wireless бази анализиране су перформансе предложеног алгоритма објективне процене квалитета и упоређене са перформансама водећих алгоритама објективне процене, чије перформансе су преузете из [260]. Перформансе мера су изражене кроз мерење тачности предикције субјективних скорова (LCC и RMSE) и

њихову монотоност (SROCC). Перформансе анализираних мера су дате у Таб. 5.17, где су и најбољи резултати подебљани. Већина техника је иста као код анализе спроведене у претходном делу дисертације где је коришћена LIVE Video база секвенци. Додатно су на овој бази анализирани и следеће мере:

- *Speed-weighted SSIM (SW-SSIM)* – Уведено је пондерисање SSIM скорова по кадровима на основу кретања присутног у секвенци. Пондерисани скорови су усредњени по кадровима, а у процени се користи само канал интензитета [260].
- *P-SS-SSIM* – Модификовани SSIM индекс, где се уместо усредњавања локалних скорова квалитета у процени користи $p\%$ најнижих SSIM вредности по кадровима. Крајња мера квалитета добија се усредњавањем добијених скорова по кадровима, а у процени се користи само канал интензитета [260].



Сл. 5.33 Деградиције које се јављају у компримованој видео секвенци са 2 Mbps пренетој кроз канал са различитим нивоима пакетских губитака.

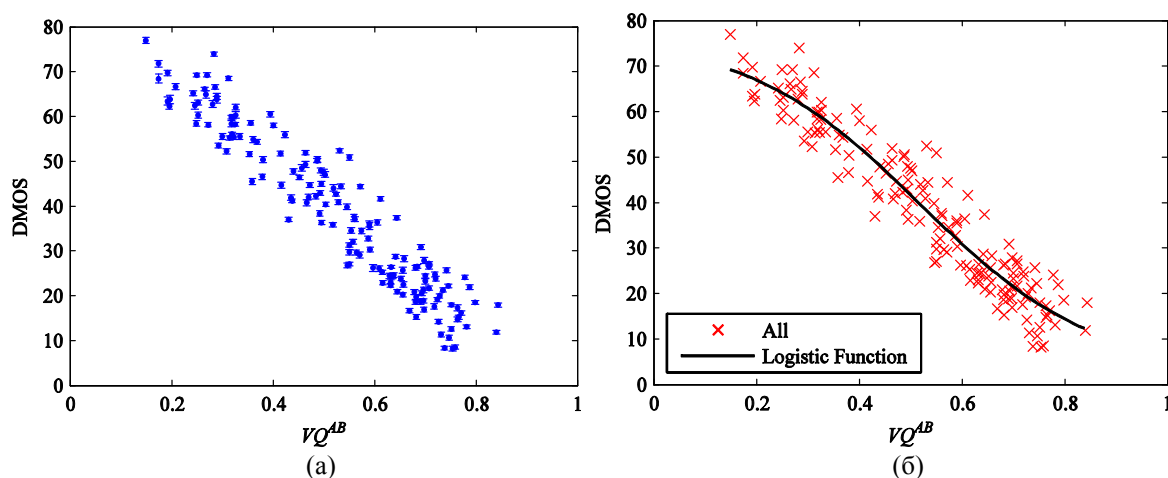
Коефицијент корелације објективних и субјективних скорова је значајно већи од корелације на VQEG и LIVE Video сетовима, што се може објаснити

униформношћу садржаја и униформношћу дисторзије. Генерално, најбољи резултати добијају се коришћењем VQM мере, али је већина алгоритама постигла висок ниво сагласности објективних и субјективни процена, где спада и мера VQ^{AB} ($LCC_{VQ}=0.9548$).

Таб. 5.17 Перформансе мера објективне процене на LIVE Wireless бази секвенци.

Мера процене	LCC	SROCC	RMSE
Frame PSNR	0.8639	0.8615	8.8997
Frame SSIM	0.8875	0.8967	8.1448
MS-SSIM	0.9588	0.9608	5.0196
VQM	0.9711	0.9721	4.2172
VSNR	0.9484	0.9418	5.6028
SW-SSIM	0.9617	0.9599	4.8450
P-SS-SSIM	0.9637	0.9628	4.7180
V-VIF	0.9524	0.9470	5.3854
VQ^{AB}	0.9548	0.9498	5.2516

Дијаграми расипања субјективно-објективних скорова за VQ^{AB} меру на комплетном LIVE Wireless сету, са интервалима поузданости субјективних скорова и оптималном логистик кривом, су дати на Сл. 5.34. Интервали поузданости указују на висок степен сагласности субјектата у оцењивању и поузданост резултата субјективних тестова. Дијаграми такође показују да је логистик крива објективних VQ^{AB} вредности глатка и линеарна.



Сл. 5.34 Дијаграми расипања објективних VQ^{AB} и субјективних скорова за све видео секвенце у LIVE Wireless бази: (а) приказ са интервалима поузданости субјективних скорова, (б) приказ са оптималном логистик кривом.

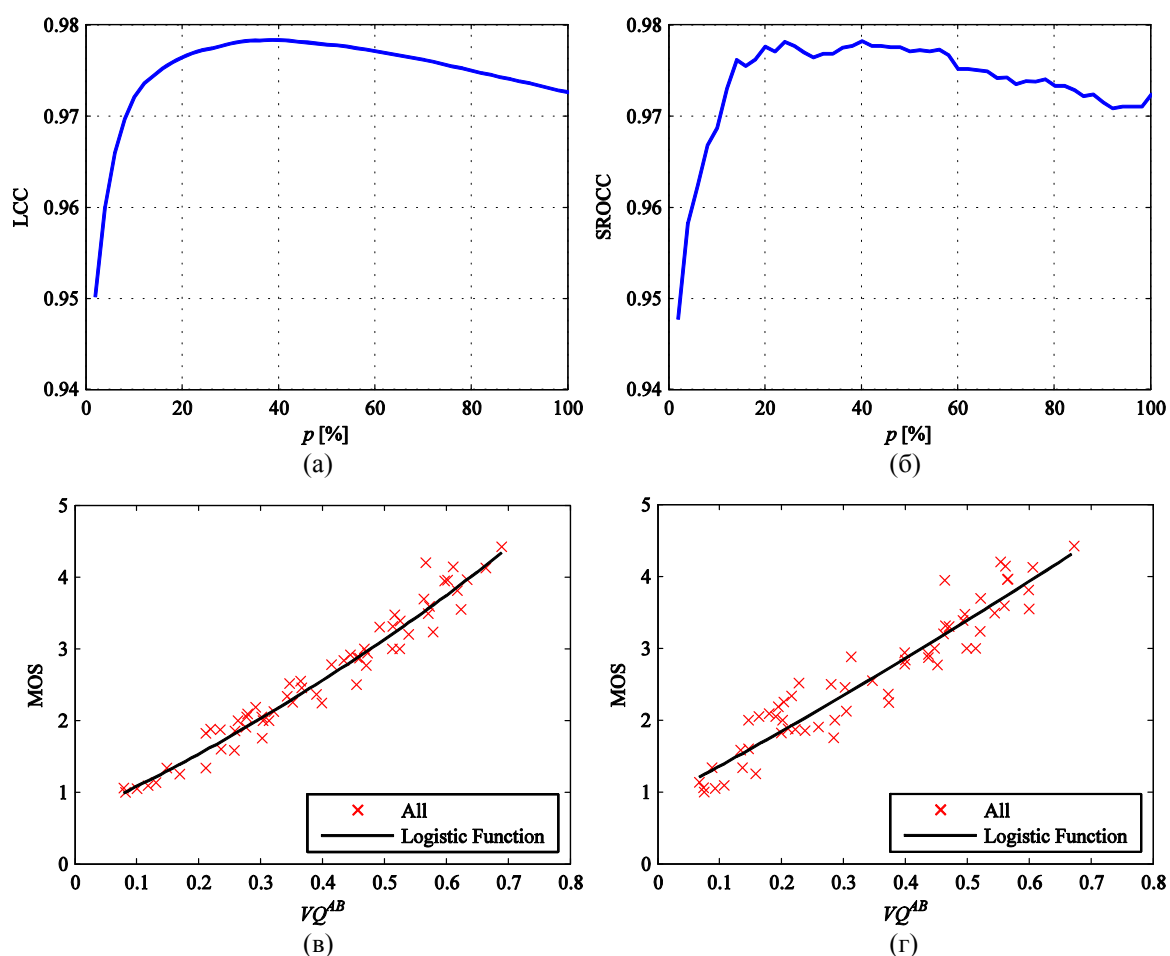
5.10. Увођење значаја објективних скорова кадрова

У дисертацији је анализирана могућност побољшања перформанси VQ^{AB} мере увођењем визуелног значаја VQ^{AB} скорова појединачних кадрова. Интуитивно је јасно да кадрови тест секвенце немају исти значај за посматрача. Коришћењем хипотезе да су за процену квалитета неки кадрови тест секвенце битнији од других, селекција скорова квалитета током трајања секвенце $VQ^{AB}(t)$ је једна од могућности за побољшање перформанси објективне мере. У радовима се обично наводе две хипотезе које могу утицати на људску перцепцију квалитета слике/видеа. Прва хипотеза је да на перцепцију утиче то који регион слике се посматра. Друга хипотеза је да људи обраћају више пажње на кадрове лошег квалитета него на кадрове доброг квалитета и самим тим ти кадрови имају већи значај у процени квалитета [237]. Тако је у раду разматрана селекција VQ^{AB} скорова кроз увођење визуелног значаја кадра, представљеног кроз анализу најнижих објективних скорова по кадровима.

У даљој дискусији подразумева се да су доступне вредности квалитета појединачних кадрова добијене коришћењем VQ^{AB} мере. Ове вредности су сортиране у растућем редоследу. Усвојена је хипотеза да кадрови лошег квалитета могу имати пресудан значај на субјективну перцепцију квалитета. Због тога је разуман приступ да се већи значај да кадровима лошег квалитета (или деловима слике лошег квалитета). Постоје бројни начини да се уведе значај кадра. У дисертацији је размотрен приступ коришћења одређеног процента ($p\%$) кадрова лошег квалитета (мереног кроз објективну меру), али је остало питање коликог процента и колики значај кадра увести. Како би се одредила оптимална вредност параметра p , спроведена је детаљна оптимизација полазећи од $p=2\%$ до $p=100\%$ са инкрементом од 2% кадрова. Уместо да се уводи произвољна монотона функција квалитета (рецимо опадајућа експоненцијална функција), коришћен је статистички принцип увођења већег значаја екстремним вредностима – у овом случају нижим скоровима. Најнижих $p\%$ VQ^{AB} скорова су једнако пондерисани, а остатак објективних вредности је игнорисан, тј. њихов значај је једнак 0. Коришћењем средње вредности најнижих објективних скорова добија се модификована верзија процене VQ^{AB} .

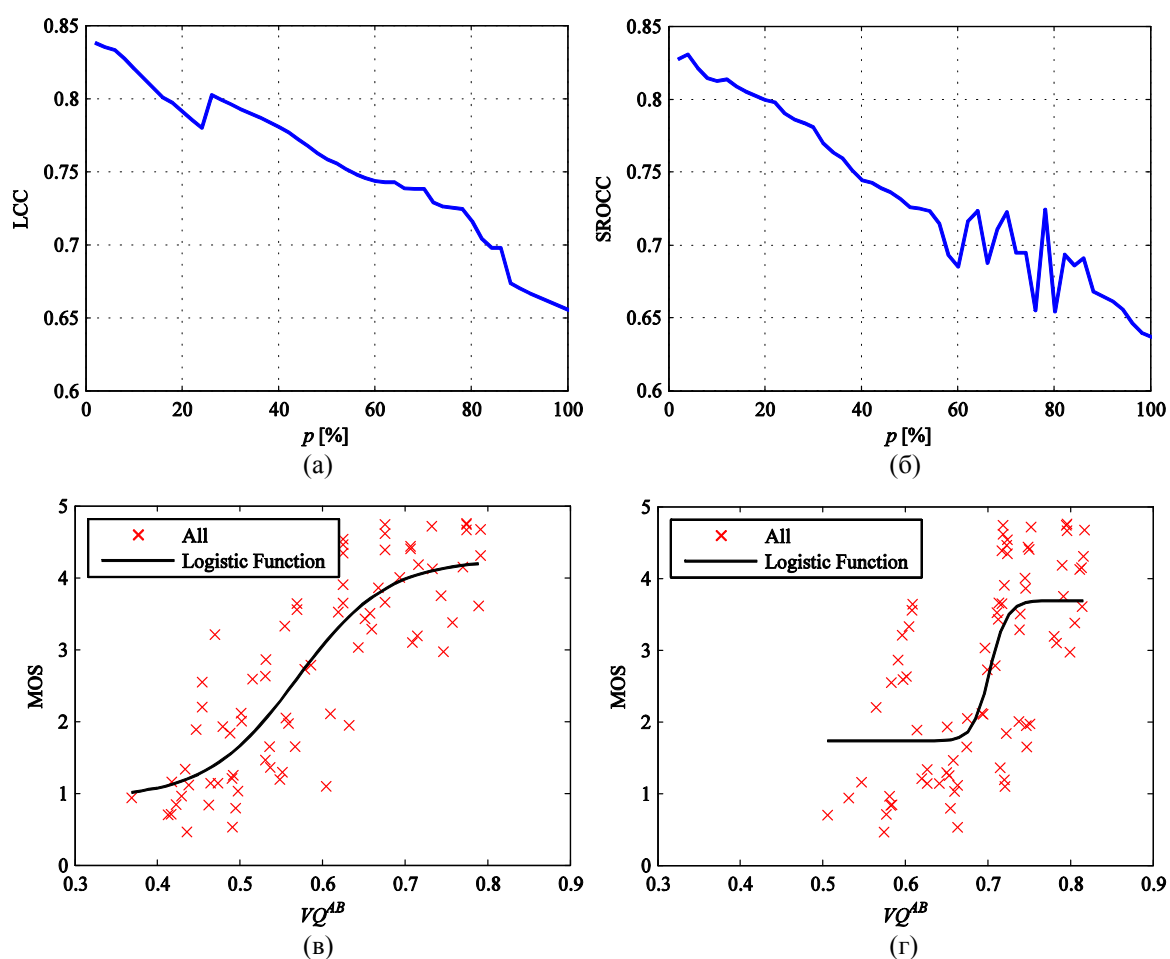
У валидацији предложеног приступа и оптимизацији VQ^{AB} мере коришћено је пет база секвенци (оптимизација је рађена за сваку базу понаособ). Перформансе мере су евалуиране кроз поређење са субјективним DMOS/MOS скоровима користећи две конвенционалне мере процене – LCC и SROCC. За пресликавање скорова коришћена је логистик функција са четири параметра. Добијени резултати (зависност перформанси VQ^{AB} мере од селекције релевантних $p\%$ кадрова и дијаграми расипања) су илустровани на Сл. 5.35 – Сл. 5.40.

На Сл. 5.35 приказани су дијаграми оптимизације добијени за UoM базу секвенци и то по две мере које показују потпуно слагање по питању избора значајних кадрова. На слици су такође приказане логистик криве за најбољи и најлошији случај избора сортираних објективних скорова кадрова.



Сл. 5.35 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на UoM сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.9784, $p=38\%$) и (г) најлошији случај (LCC=0.9502, $p=2\%$).

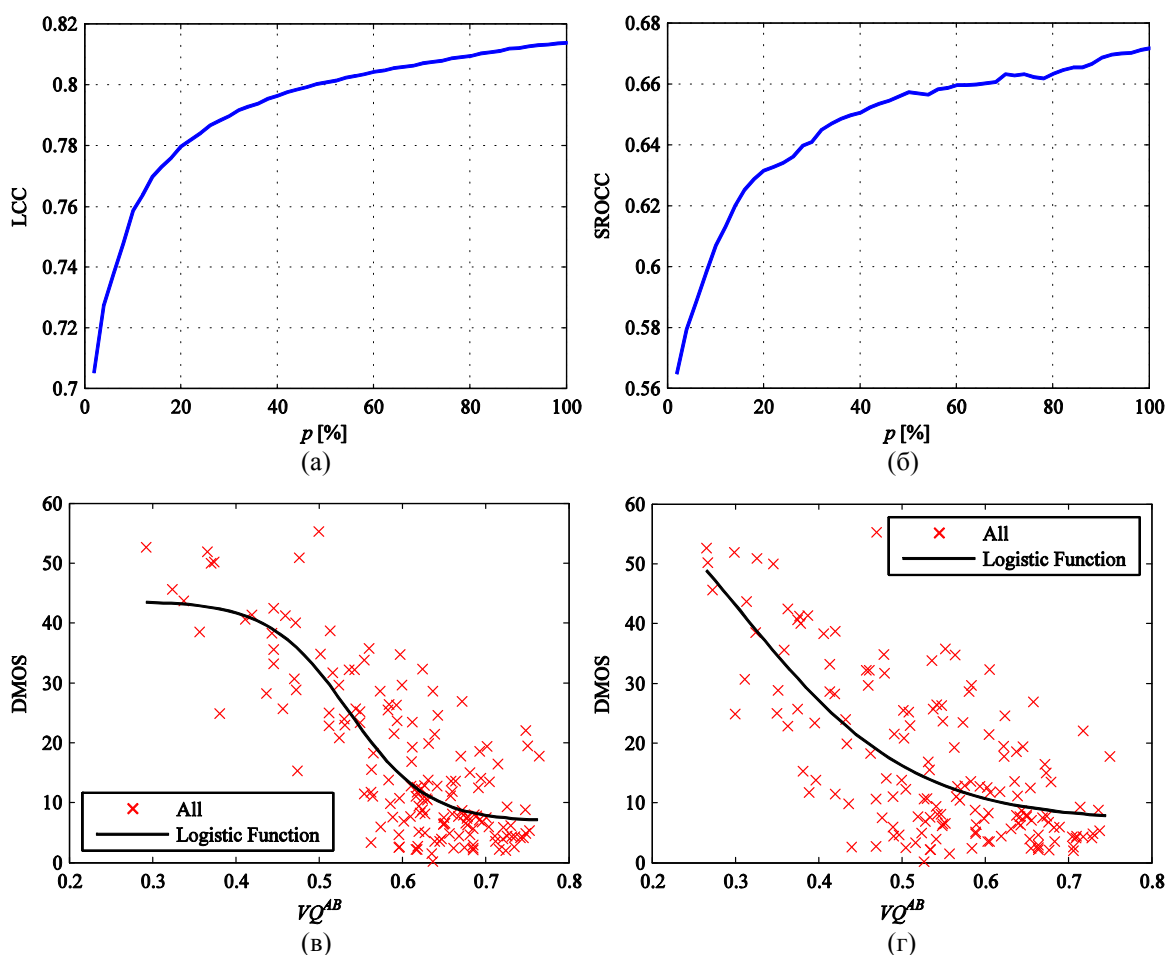
Резултати за EPFL-PoliMI сет видео секвенци су приказани на Сл. 5.36. Оптимална вредност параметра p је близу нуле, тј. $p=2\%$. Ово није неочекивано узимајући у обзир да ова база садржи деградације настале пакетским губицима насталим током преноса кроз мрежу. Видљиве деградације су у овом случају просторно локализоване и уочљиве само у неким кадровима – кадровима лошег квалитета.



Сл. 5.36 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на EPFL-PoliMI сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај ($LCC=0.8380$, $p=2\%$) и (г) најлошији случај ($LCC=0.6542$, $p=100\%$).

Резултати оптимизације за VQEG 525/60Hz сет видео секвенци приказани су на Сл. 5.37. Уочава се потпуно супротан тренд оптимизационе криве од криве код EPFL-PoliMI сета, тј. најбоље је у анализи узети свих $p=100\%$ кадрова. Исти тренд је добијен за други сет – 625/50Hz ове базе, Сл. 5.38, где је оптимално узети $p=94\%$

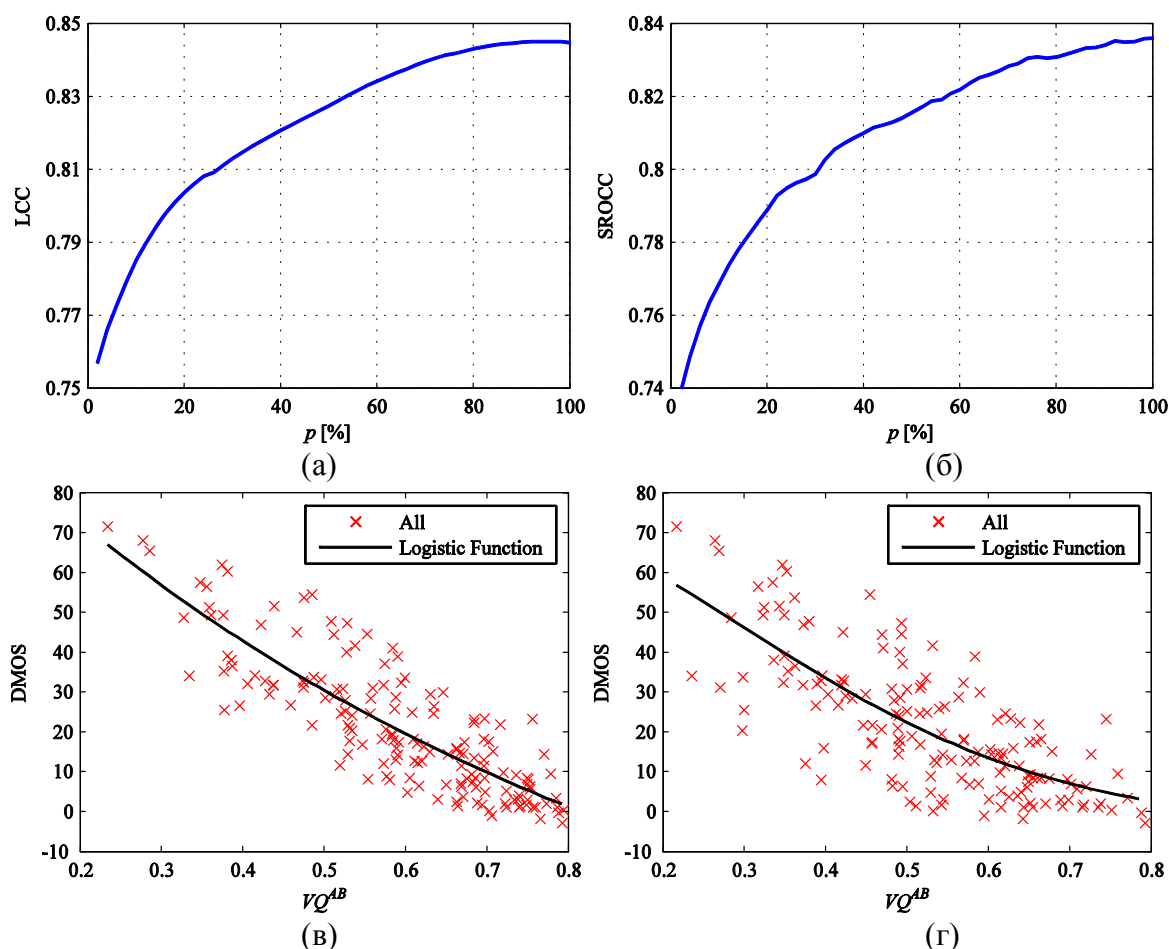
кадрова са најнижим скоровима. У овом случају нешто је већа корелација субјективних и објективних скорова, али је динамички опсег корелације приближно исти.



Сл. 5.37 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на VQEG 525/60Hz сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.8144, p =100%) и (г) најлошији случај (LCC=0.7053, p =2%).

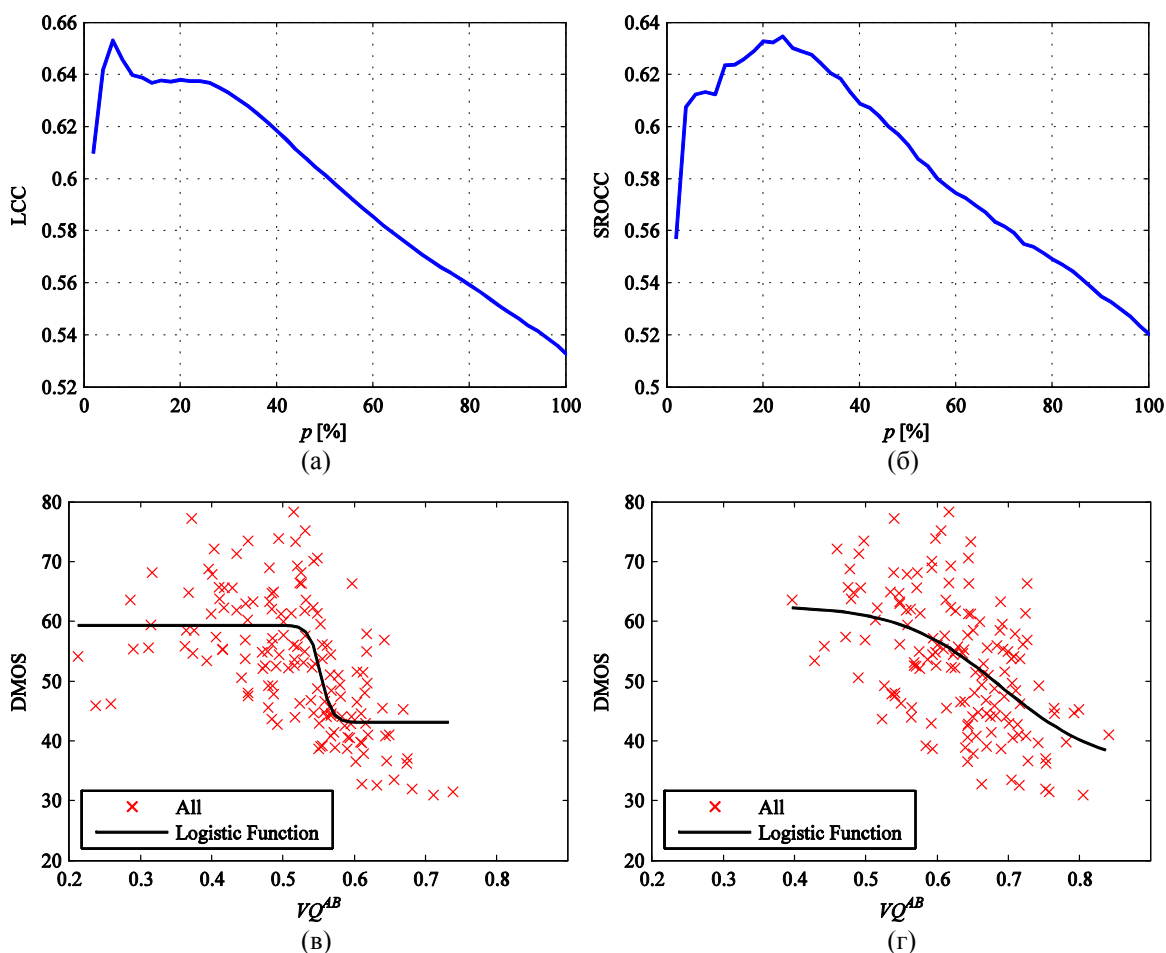
Еквивалентни оптимизациони дијаграм за LIVE Video базу секвенци приказан је на Сл. 5.39. Облик криве се у великој мери разликује од облика VQEG сетова при чему су перформансе објективне мере лошије у комплетном опсегу вредности параметра p . У овом случају оптимална вредност параметра p је ближа 0, па се избором релативно малог броја кадрова лошег квалитета добијају оптималне перформансе процене квалитета. Овај ефекат се такође може објаснити природом дисторзија присутних у видео секвенцама са деградацијом. За разлику од VQEG базе

која садржи деградације настале компресијом, LIVE Video база садржи и деградације настале пакетским губицима, симулираним за IP и бежичне мреже. То значи да је деградација квалитета више локализована у оквиру секвенце бита и вероватно присутна само у неким кадровима. Ови кадрови су од виталног значаја за субјективну процену квалитета, а самим тим и за објективну процену.



Сл. 5.38 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова на VQEG 625/50Hz сету: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.8453, $p=94\%$) и (г) најлошији случај (LCC=0.7567, $p=2\%$).

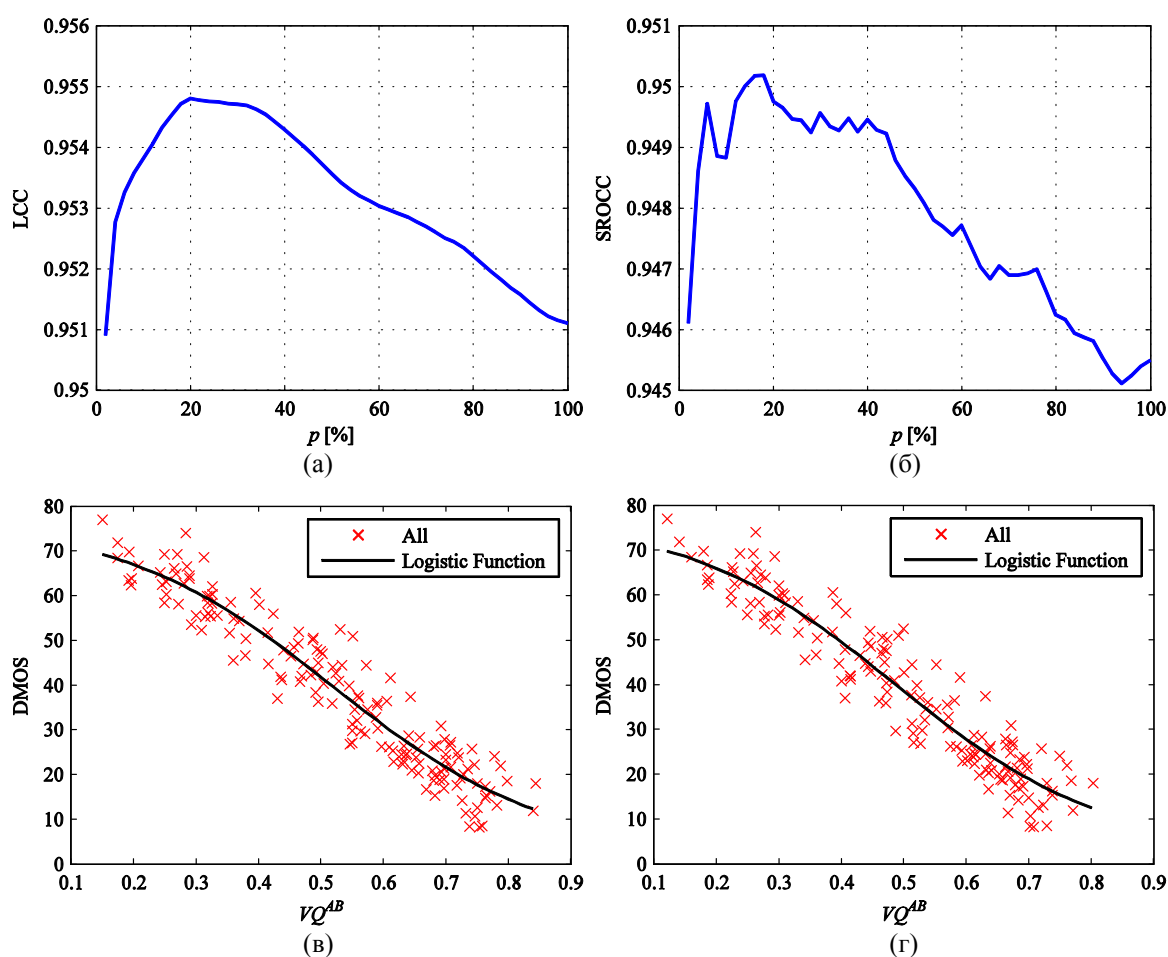
Сличан тренд је добијен кроз оптимизацију на LIVE Wireless бази секвенци, Сл. 5.40, која садржи видео секвенце код којих пакетски губици доминирају над другим формама деградације. Сада је максимум корелације за $p=20\%$, при чему је динамички опсег вредности корелације јако мали – мањи од 1% од расположивог опсега. Тиме се добија сличан тренд логистик крива за најбољи и најгори сценарио избора вредности параметра p .



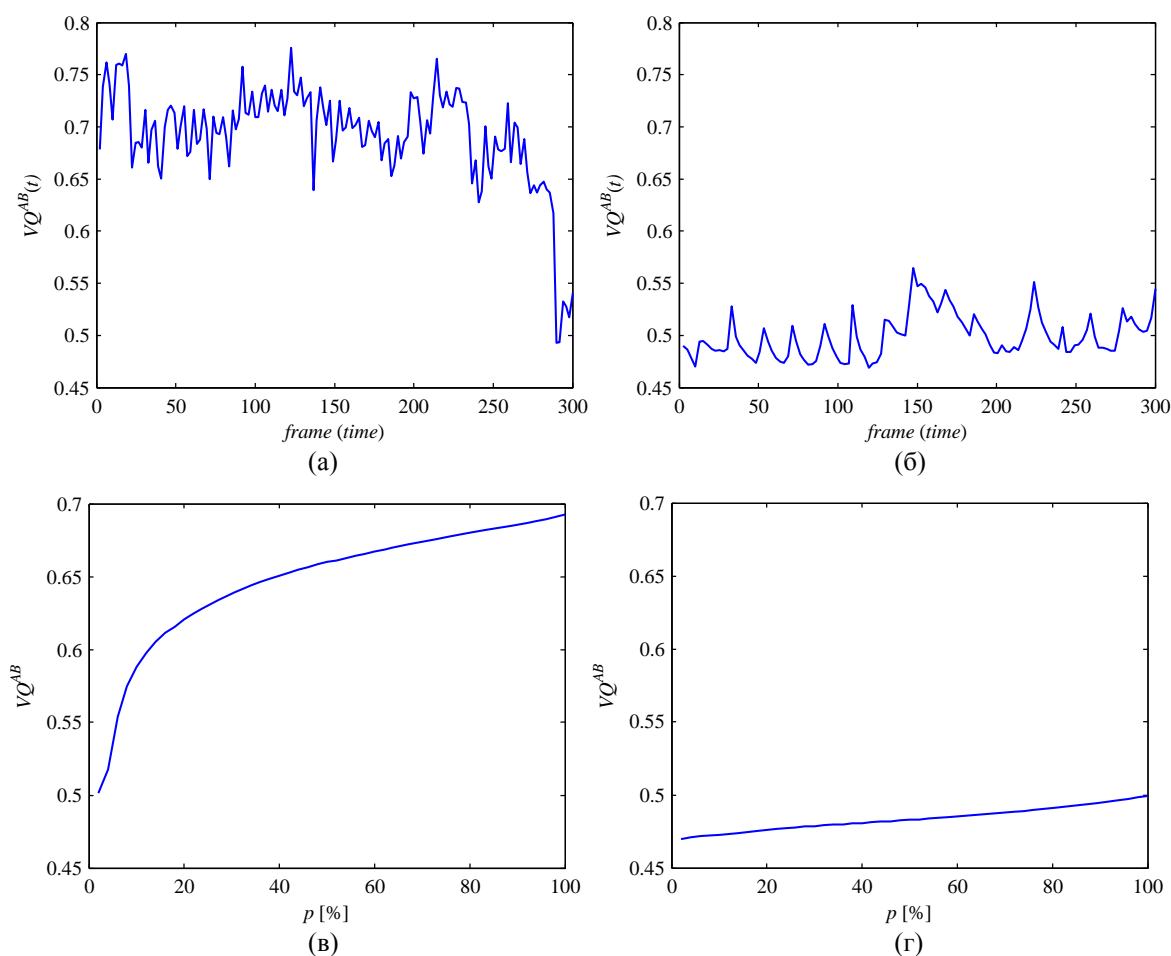
Сл. 5.39 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова за LIVE Video базу секвенци: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.6532, $p=6\%$) и (г) најлошији случај (LCC=0.5323, $p=100\%$).

Из спроведене анализе је јасно да оптимална вредност параметра p којим се врши ограничење броја кадрова у крајњој естимацији квалитета зависи од типа деградације видео сигнала. Мера за процену квалитета је релативно независна од избора вредности параметра p када је деградација доминантно настала компресијом, при чему су резултати поуздани у широком опсегу вредности параметра p . У случају када пакетски губици доминирају, пожељно је у евалуацији користити мањи број кадрова. Овај закључак је додатно потврђен илустрацијама датим на Сл. 5.41 и Сл. 5.42. На Сл. 5.41(а) и Сл. 5.41(б) приказана је процена квалитета кадрова тест секвенци из две анализирани базе – EPFL PoliMI и VQEG. Тест секвенце су настале од исте изворне секвенце – секвенца Mobile, а за процену квалитета користи се предложена мера VQ^{AB} . Код тест секвенце из EPFL PoliMI базе постоје значајне

варијације квалитета. Нарочито је уочљив пад квалитета при крају секвенце кадрова. Овај пад је последица пакетских губитака (PLR=3%). Варијације квалитета постоје и у тест секвенци из VQEG базе (последица компресије), али је динамички опсег квалитета значајно ужи од динамичког опсега који је последица пакетских губитака. На Сл. 5.41(в) и Сл. 5.41(г) је за анализиране тест секвенце приказана зависност предложене мере од процента релевантних кадрова са најнижим скоровима квалитета. Због малих варијација и суженог динамичког опсега мера VQ^{AB} мало зависи од процента релевантних кадрова VQEG базе. Са друге стране од изабране вредности параметра p значајно зависи вредност квалитета VQ^{AB} код секвенце из EPFL PoliMI базе.

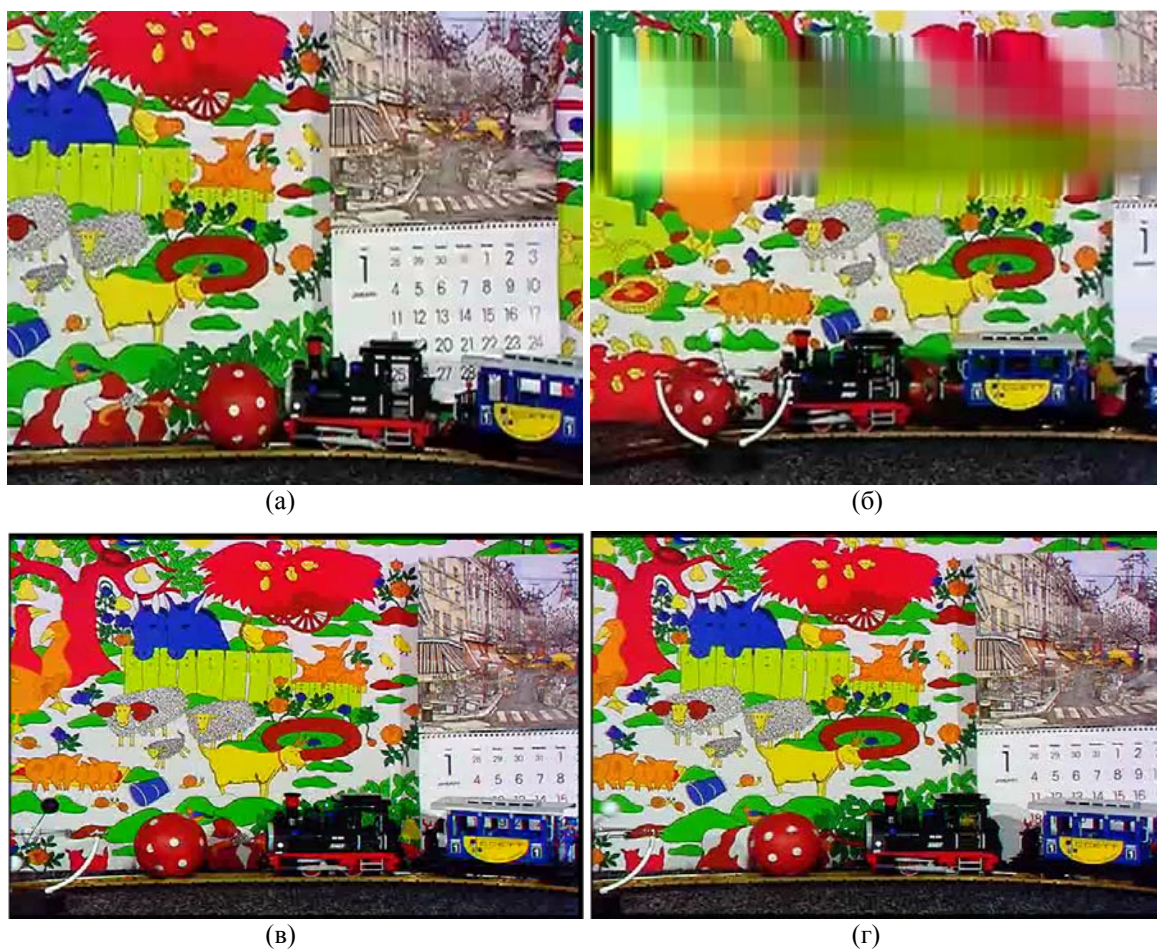


Сл. 5.40 Зависност перформанси предложене VQ^{AB} мере од процента релевантних кадрова за LIVE Wireless базу секвенци: (а) LCC, (б) SROCC и логистик криве за (в) најбољи случај (LCC=0.9548, $p=20\%$) и (г) најлошији случај (LCC=0.9509, $p=2\%$).



Сл. 5.41 (а) процена квалитета кадрова тест секвенце из EPFL-PoliMI базе, (б) процена квалитета кадрова тест секвенце из VQEG базе, (в) зависност мере VQ^{AB} од процента релевантних кадрова секвенце из EPFL-PoliMI базе и (г) зависност мере VQ^{AB} од процента релевантних кадрова секвенце из VQEG базе. Тест секвенце су настале од исте изворне секвенце – секвенце *Mobile*.

За анализиране тест секвенце су на Сл. 5.42 приказани кадрови најбољег и најлошијег квалитета – квалитет је мерен кроз VQ^{AB} меру. Кадрови су сведени на исте димензије али се њихове оригиналне димензије значајно разликују (352x288 пиксела у EPFL-PoliMI бази и 720x486 пиксела у VQEG тест секвенци). Уочавају се значајне разлике у квалитету кадрова из EPFL-PoliMI базе. Због пакетских губитака у кадру најлошијег квалитета дошло је до значајног нарушавања структура из оригиналне секвенце, а што има пресудну улогу у субјективној импресији квалитета. Нарушавање квалитета је присутно у горњој половини приказаног кадра. Код кадрова тест секвенце из VQEG базе не може се уочити значајна разлика у квалитету кадрова.



Сл. 5.42 Кадрови анализираних тест секвенци: (а) кадар најбољег квалитета из EPFL-PoliMI базе, (б) кадар најлошијег квалитета из EPFL-PoliMI базе, (в) кадар најбољег квалитета из VQEG базе и (г) кадар најлошијег квалитета из VQEG базе.

Добијени резултати где се оптимална вредност параметра p на анализираним базама креће у интервалу од $p=2\%$ до $p=100\%$ могу се тумачити и на други начин. Наиме, овакве вредности се могу објаснити са различитим садржајима присутним у секвенцама и начином (циљем) спровођења субјективних тестова. При формирању база и прикупљању субјективних скорова, посматрачима су постављени различити задаци: (1) опис дешавања на сцени (UoM), (2) одређивање квалитета секвенце (VQEG, LIVE Video) или (3) процена степена деградације (LIVE Wireless, EPFL-PoliMI). Генерално се пред посматраче у спровођењу субјективних тестова постављају два задатка. Први је процена квалитета а други процена корисности посматраних секвенци, а што се може добити из различитог броја кадрова. На пример, посматрач који је учествовао у прикупљању субјективних скорова за EPFL-

Polimi базу може одговорити на питање: Да ли су у секвенци присутна велика изобличења? (тј. појаве које деградирају структуре присутне на сцени или формирају нове облике), на основу релативно малог броја кадрова из комплетне секвенце. Тиме оптимална вредност параметра p тежи нули ($p=2\%$). Посматрачима који су имали за циљ да евалуирају квалитет секвенце или да опишу дешавања на сцени потребан је већи број кадрова за поуздану одлуку, као што се десило у UoM и VQEG сетовима. Због тога оптималну вредност параметра p у реалним апликацијама треба одредити уважавајући контекст у коме се процена спроводи као и одговор који се очекује од посматрача.

Стандардна имплементација VQ^{AB} мере у процени квалитета користи $p=20\%$ кадрова са најнижим скоровима квалитета. Усвојена вредност процента кадрова која се користи у евалуацији квалитета представља компромисно решење којим се не улази у контекст у коме се процена спроводи или циљ са којим се врши прикупљање субјективних импресија квалитета.

5.11. Поређење линеарног и нелинеарног модела очувања информација о градијенту

Нелинеарни модел очувања информација (73) користи осам параметара за одређивање информационог губитка (сличности) између улаза A и B , а који се користе за дефинисање сигмоидних функција описаних константама $k_g, \sigma_g, k_a, \sigma_a, k_b, \sigma_b, k_c$ и σ_c . Разлог увођења ових константи био је да се моделује одзив биолошких система који је нелинеаран. Алтернативни приступ (74) за процену квалитета видеа посматра промену изворног сигнала кроз очување информација о просторном и временском градијенту и очување боје. У овом приступу иста релативна промена сигнала даје пропорционално исту релативну промену квалитета (очувања информација). Код оба приступа уведене су додатно мапе значаја (w_s, w_t), значај изворних информација (k_s, k_t, k_c) и избор релевантних кадрова ($p=20\%$).

Поређење линеарног и нелинеарног модела је спроведено на пет база описаних у претходним деловима рада. Резултати анализе приказани су у Таб. 5.18. Из приказаних резултата се не може јасно издвојити оптимални приступ.

5. Процена квалитета видеа заснована на очувању информација о градијенту

Нелинеарни модел је бољи на VQEG сетовима, док је линеарни модел – модел верности информација, бољи на три сета (EPFL-PoliMI, LIVE Video и LIVE Wireless). Код UoM сета благу предност има нелинеарни модел. Генерално гледано, перформансе предложених приступа се могу значајно разликовати.

Таб. 5.18 Перформансе предложене мере објективне процене квалитета кроз поређење линеарног и нелинеарног модела очувања информација.

База секвенци	модел	LCC	SROCC	MAE	RMSE	OR (%)
VQEG 625/50 Hz	линеарни	0.7065	0.6983	8.9422	11.7924	67.5000
	нелинеарни	0.8035	0.7888	7.7997	9.9193	69.3750
VQEG 525/60 Hz	линеарни	0.6999	0.5388	7.6233	9.6745	63.7500
	нелинеарни	0.7799	0.6316	6.6324	8.4781	60.6250
LIVE Video	линеарни	0.6984	0.7034	6.3588	7.8569	2.6667
	нелинеарни	0.6378	0.6330	6.7417	8.4543	4
LIVE Wireless	линеарни	0.9650	0.9599	3.8901	4.6337	85
	нелинеарни	0.9548	0.9498	4.2846	5.2516	81.2500
EPFL-PoliMI	линеарни	0.8269	0.8219	0.6183	0.7585	56.4103
	нелинеарни	0.7909	0.7995	0.6804	0.8253	61.5385
UoM	линеарни	0.9698	0.9650	0.1768	0.2238	61.6667
	нелинеарни	0.9765	0.9777	0.1554	0.1979	55

5.12. Опис STrial 0 и STrial 1 сетова намењених процени квалитета видео секвенци са малим битским брзинама

У ранијим фазама истраживања предложена је мера објективне процене квалитета видео секвенци – VQ^{AB} . Мера је оптимизирана на половини UoM сета видео секвенци и на комплетном сету (од анализираних сетова) постигла највећи степен сагласности са субјективним скоровима – 97%. За UoM би се могло рећи да у потпуности не одговара ономе што се од једне базе намењене процени квалитета захтева: (1) база садржи свега пет изворних секвенци, (2) у прикупљању информација о томе где гледају посматрачи током процене квалитета учествовао је мали број посматрача, (3) време које су посматрачи трошили на анализу квалитета је било знатно веће од времена препорученог од стране ITU, (4) субјективни скорови нису дати по појединачним секвенцама већ одсликавају квалитет више секвенци са истим параметрима преноса. Од осталих, јавно доступних база секвенци са субјективним проценама, по битским брзинама које се могу сматрати прихватљивим за оно што је циљ истраживања у тези могу се навести FUB и EPFL-PoliMI базе

секвенци. FUB база је некомплетна, садржи један тип деградације изворних секвенци и свега шест изворних секвенци. Са друге стране EPFL-PoliMI база секвенци је циљно прављена како би се анализирали пакетски губици који настају током преноса. Остале базе генерално садрже секвенце високе резолуције што излази из оквира онога што је циљ израде дисертације.

Због тога је, како би се презентовале преформансе VQ^{AB} мере у складу са стварним циљем објективне евалуације – евалуација видео секвенци компримованих на различите начине за потребе видео надзора, спроведена евалуација на новом сету репрезентативних видео секвенци. Спроведени су субјективни тестови са следећим циљевима:

- потврђивање научног кредибилитета објективне мере процене видео квалитета VQ^{AB} кроз субјективне тестове и прикупљање субјективних импресија у контролисаним условима;
- израда софтвера потребног за спровођење субјективних тестова – приказ секвенци, навигација током спровођења тестова и прикупљање скорова унутар једноставног окружења;
- евалуација/валидација мере објективне процене користећи прикупљене субјективне податке.

Постављени циљеви су остварени кроз рад у четири фазе:

- формирање базе секвенци: прикупљање изворних видео секвенци, избор енкодера и форматирање секвенци;
- израда софтвера за приказ, навигацију и прикупљање субјективних скорова;
- спровођење субјективних тестова (на Универзитету у Манчестеру) у контролисаним условима и са репрезентативним бројем посматрача;
- евалуација/валидација мера објективне процене.

Избор изворних видео секвенци је спроведен тако да се покрију типичне ситуације видео надзора, са различитим сценама и условима окружења. Деградације изворних видео секвенци су добијене кроз различите начине компресије за релевантне битске брзине. Спроведена су два субјективна теста у којима су

посматрачи имали прилике да виде шест (2 енкодера x 3 битске брзине) односно девет (3 енкодера x 3 битске брзине) модификованих изворних секвенци.

Софтвер за прикупљање субјективних података (VAL-STAT – Visual Automation Ltd. Subjective Trial Automation Tool) који је намењен аутоматизацији процедуре субјективне процене и прикупљања субјективних скорова је имплементиран у Microsoft Windows окружењу. Софтвер омогућава унос података о посматрачу (године, корекција вида, искуство) и истовремени приказ видео записа у случајном редоследу. Кориснички интерфејс је веома једноставан и омогућава фокусирање посматрача на презентовани видео садржај. Софтвер је тестиран на стандардној хардверској конфигурацији како би се осигурало да деградација квалитета не настаје због софтверске имплементације или карактеристика хардвера.

Додатно је прављен софтвер за анализу података прикупљених током субјективних тестова. Овај софтвер је развијен у Matlab[®] окружењу и омогућава учитавање, сумирање и анализу субјективних података и као излаз даје субјективне импресије у формату потребном за валидацију објективних скорова. Овај пакет такође даје бројне показатеље поузданости субјективних тестова, статистике оцена посматрача и поузданост и сигурност добијених субјективних резултата.

Субјективни тестови су спроведени током септембра месеца 2012. године. Први субјективни тест је трајао три дана и послужио је како би се процедура процене и софтвер проверили и додатно побољшали. Други субјективни тест је трајао четири дана. Тестови су спроведени у различитим, контролисаним условима, при чему се на услове спровођења тестова – окружење, осветљење, шум, процедура процене, много више обраћала пажња у другом субјективном тесту. У тестовима је евалуирано 129 секвенци са деградацијама (48 и 81, тим редом) насталих од 17 различитих изворних секвенци.

Субјекти који су учествовали у тестовима су навели да нису имали потешкоћа у комплетирању тестова што је довело до конзистентних резултата погодних за валидацију објективних мера. Други субјективни тест је спроведен са две различите групе и на две локације.

5.12.1.Опис референтних видео секвенци

Видео секвенце су преузете из EYELIDS базе CCTV секвенци видео надзора и из јавно доступних материјала академских институција (University of Ohio, Dublin University, The University of Reading). Примери кадрова изворних секвенци приказани су на Сл. 5.43 док је опис секвенци дат у Таб. 5.19.



Сл. 5.43 Примери кадрова изворних секвенци STrial сетова.

5.12.2.Опис коришћених кодека

Ефекти деградације настале током преноса кроз канал ограниченог капацитета су симулирани компресијом изворних података са различитим енкодерима и различитим степеном компресије. Степен компресије је повезан са ограничењима реалних комуникационих инфраструктура.

5. Процена квалитета видеа заснована на очувању информација о градијенту

Таб. 5.19 Опис изворних секвенци коришћених у Strial сетовима.

Име секвенце	fps	Трајање	Опис сцене	Активност на сцени
20060708_01_140x440 Сл. 5.43(а)	15	20 s	Маскирани људи седе на неплодном терену. Окружени су стенама и дрвећем.	Скоро да је нема. Слабе варијације због температуре и скоро неприметног кретања једног од субјеката.
abandoned_luggage Сл. 5.43(б)	25	59 s	Пешачка стаза на железничкој станици.	Људи ходају и трче. Воз долази, напушта станицу и стиже следећи воз.
crowding_detection (црно-бела видео секвенца) Сл. 5.43(в)	0.5	16 min 42 s	Подземна железничка станица.	Људи се шетају и групишу се на сцени. Воз долази, путници се укрцавају и воз одлази.
db3_2b_30fps Сл. 5.43(г)	30	20 s	Двориште универзитетског насеља снимљено са веће висине.	Људи се крећу на различитим даљинама од камере. Сенка облака прелази преко сцене.
ME_Training Сл. 5.43(д)	30	12 min 14 s	Прометна раскрсница снимљена са веће висине.	Кретање саобраћајног тока од камере и ка камери, са лева удесно и обратно. Наизменично заустављање саобраћајног тока и прелазак пешака преко улице.
parked_vehicle_day Сл. 5.43(ђ)	25	1 min 47 s	Пут са умереним саобраћајем снимљен са веће висине.	Наизменично кретање саобраћајног тока (од камере и ка камери) и преласка пешака преко улице.
Parked_Vehicle_night Сл. 5.43(е)	25	1 min 40 s	Пут и пешачка стаза осветљене улице снимљене ноћу.	Возило и пешак се крећу ка камери. Једно возило се креће од камере, зауставља и наставља са кретањем.
Balcony3_Vis Сл. 5.43(ж)	25	7 min 54 s	Двориште универзитетског насеља у Даблину снимљено са висине.	Возила у позадини пролазе са лева у десно и обратно. Пешаци се крећу од камере и ка камери као и са лева у десно и са десна у лево.
S1-T1-C-2 Сл. 5.43(з)	25	2 min	Чекаоница железничке станице.	Кретање великог броја пешака са лева у десно, са десна у лево, од камере и ка камери.

Изабрана су три кодека којима је могуће достићи мале битске брзине:

- MPEG-4 Part 2

Кодек са променљивом брзином преноса заснован на дискретној косинусној трансформацији и последњи пут побољшан у 2009. години. MPEG-4 Part 2 је био веома популаран кодек за видео и као такав је имплементиран у DivX, Xvid и QuickTime формате.

- H.264/AVC

Кодек са променљивом битском брзином заснован на компензацији блоковских кретања. Обухвата бројне напредне технике за побољшање визуелног квалитета и захтева два пута мању битску брзину за исти квалитет преноса у односу на MPEG-4 Part 2. Широко је прихваћен од стране видео дистрибутера, део је Blue-Ray спецификације и има подршку од већине интернет претраживача у апликацијама видео преноса.

- RealVideo 10

Дизајниран као кодек за трансфер видео сигнала, има динамичко подешавање броја кадрова и резолуције кодованог сигнала како би се задовољили тачни захтеви по питању битске брзине. Нестандардна стратегија кодовања и технолошки захтеви нису дозволили да се овај кодек прихвати у широкој употреби, али се и даље користи у трансферу информација у реалном времену кроз мреже ограниченог капацитета.

5.12.3. Просторна резолуција видео секвенци

Неколико фактора је утицало на резолуцију видео секвенци из базе:

- већина оригиналних изворних секвенци је доступна у формату који је близак PAL резолуцији – углавном 360x288 пиксела;
- методологија спровођења субјективних тестова је захтевала да су изворна видео секвенца и њене две/три верзије настале деградацијом истовремено видљиве на екрану за приказ;

- видео секвенце ове природе (видео надзор) ће често бити анализирани на ручним уређајима са ограниченом просторном резолуцијом.

Тако је за секвенце усвојена просторна резолуција од 360x288 пиксела. Током припреме већине видео материјала уочено је да се циљна битска брзина од 10 kbps не може постићи за резолуцију од 360x288 пиксела ако се користи MPEG-4 кодовање. Због тога је резолуција ручно смањивана док се не достигне циљна битска брзина, након чега се враћало назад у жељену резолуцију. Овај процес је потребно спровести и у реалној примени.

5.12.4. Битске брзине

Валидација објективних мера процене је извршена кроз поређење са резултатима субјективних тестова репрезентативних реалних видео апликација и њихових операционих ограничења. Тако су одабране битске брзине из стварних сценарија:

- 1) 10 kbps
- 2) 64 kbps
- 3) 128 kbps

Потребно је напоменути да је за овако мале битске протоке у процесу кодовања прихватљиво кашњење (латенција – latency). MPEG-4 Part 2 и H.264 кодеци раде тако што се за кључне кадрове преноси огромна количина података иза чега следи серија много мањих трансфера. Овакав поступак кодовања/декодовања захтева складиштење како би кључни кадрови били пренети у реалном времену кроз мрежу ограниченог капацитета.

Треба рећи и да су MPEG-4 Part 2 и H.264 намењени преносу у системима са релативно великим пропусним опсегом чиме се захтеви по питању резолуције и броја кадрова у секунди постижу на рачун губитка информација о пикселима. Са фиксним бројем кадрова у секунди и фиксном резолуцијом постоји минимална битска брзина коју могу постићи MPEG-4 и H.264 кодоване секвенце (зависно од динамике на сцени).

5.12.5. Трајање секвенци

Конзистенција оцена видео квалитета од стране посматрача је боља уколико се у процени користе краће секвенце које се могу посматрати више пута и поредити. Краће секвенце посматрач лакше и боље памти, пореди и оцењује. Због тога су посматрачима презентоване секвенце трајања по 10 s.

Од изворног материјала који обухвата неколико десетина минута видео, изабрани су делови који репрезентују различита кретања на сцени и имају различите нивое активности. Коришћен је различит број кадрова у секунди који утиче на субјективну перцепцију континуитета кретања као и на перформансе кодека. Генерално, у оквиру трајања субјективних тестова презентован је максималан број различитих видео записа.

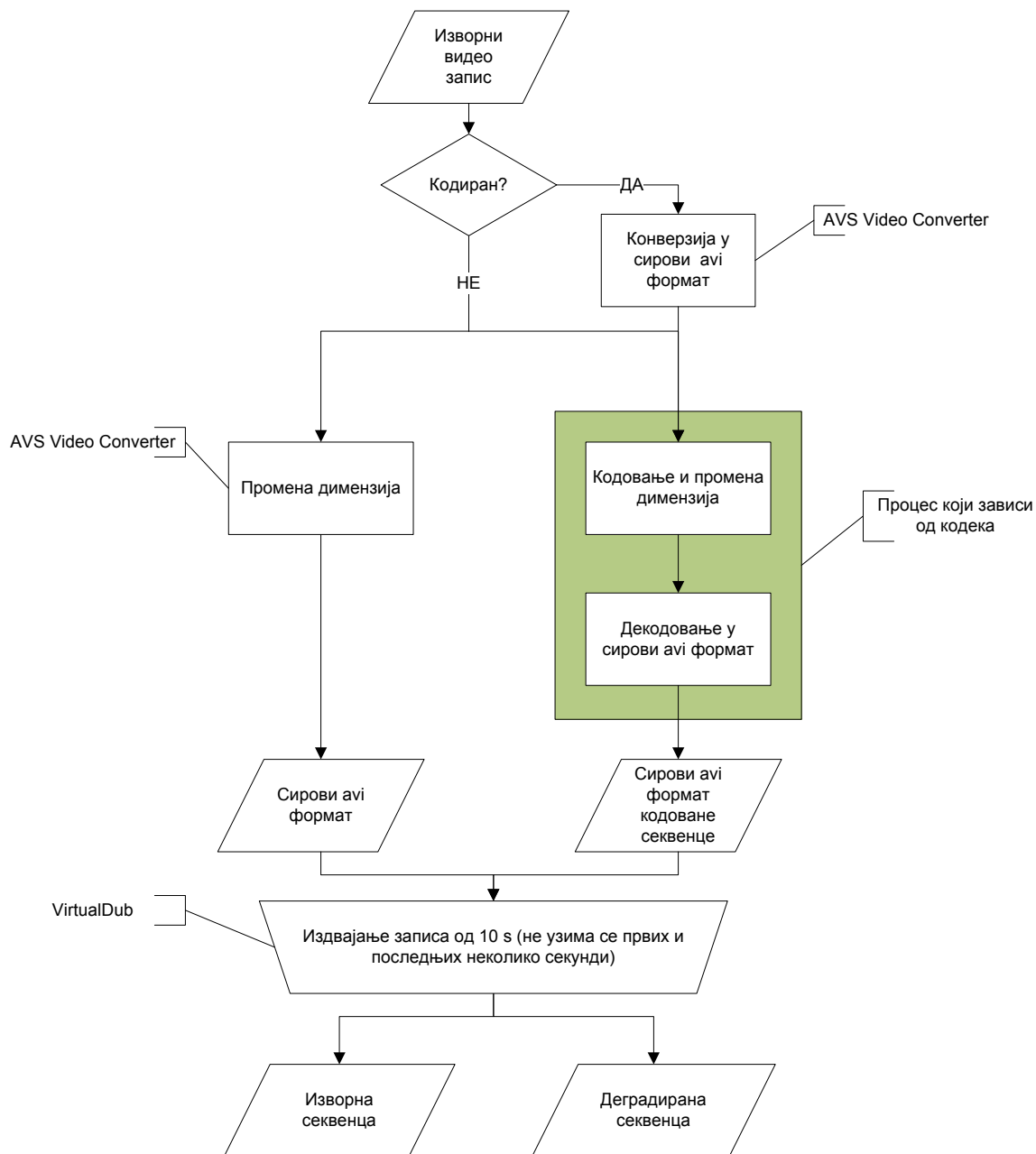
5.12.6. Генерисање видео записа

Полазећи од изворног материјала у доступној резолуцији и расположивом квалитету, коришћењем комерцијално доступних софтвера за кодовање/декодовање и обраду видео извршена је модификација како би се постигао захтевани ниво деградације. Тако је коришћен следећи софтвер:

- AVS Video Converter (MPEG-4 Part 2 и H.264/AVC кодовање; MPEG-4 Part 2, H.264/AVC и Windows Media Video декодовање)
- Bitrate Viewer (анализа битског протока MPEG-4 Part 2 и H.264/AVC)
- RealProducer (RealVideo 10 кодовање)
- RealPlayer Converter (конверзија RealVideo 10 у Windows Media Video)
- VirtualDub (издвајање репрезентативних секвенци по 10 s)

Поједностављена блок шема поступка којим се од изворног материјала долази до крајњег резултата – изворне секвенце и њених девет модификација трајања по

10s, приказана је на Сл. 5.44. Ова блок шема је разрађена у више детаља по изабраној врсти кодовања и приказана са изгледима менија током рада на Сл. 5.45 и Сл. 5.46.

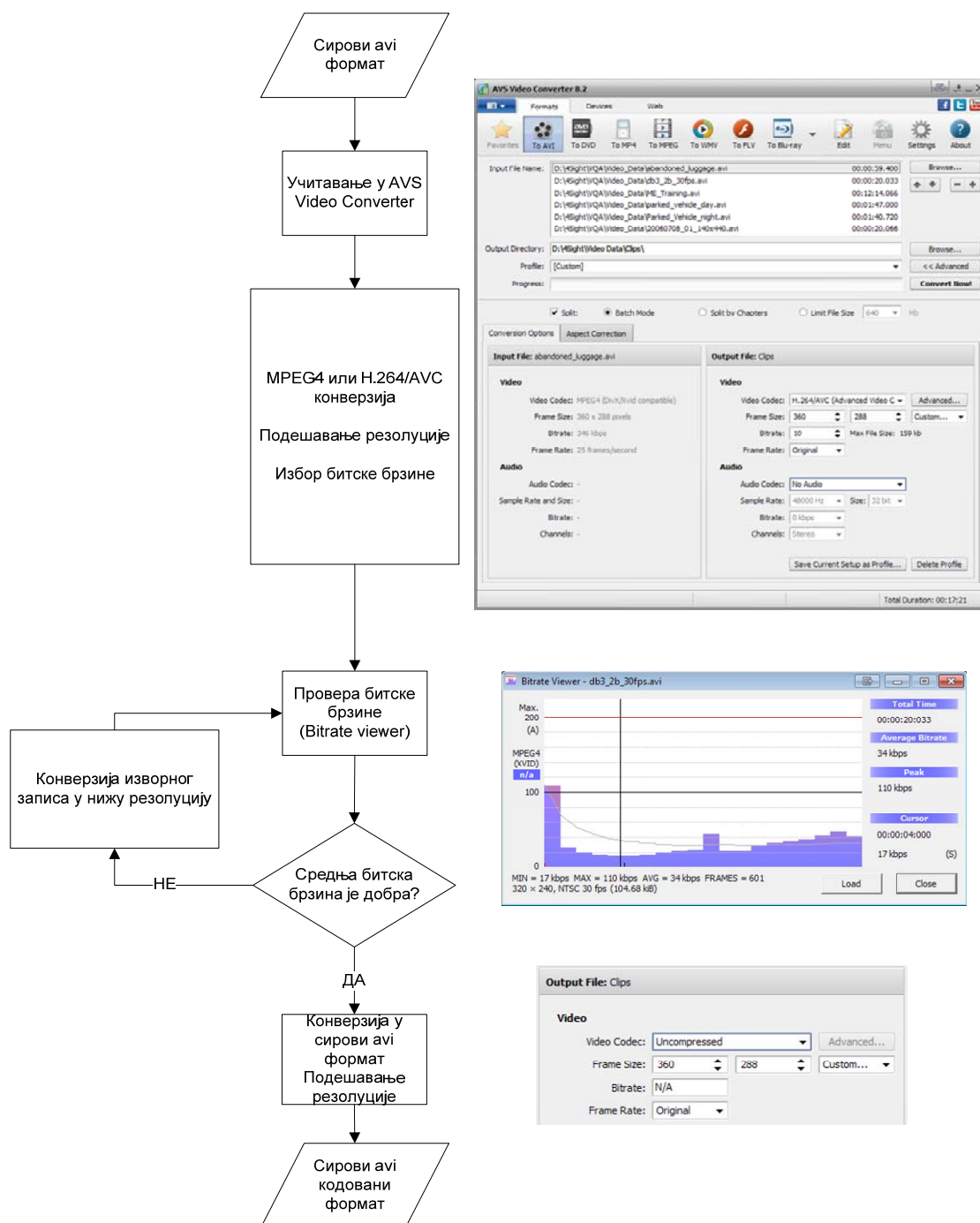


Сл. 5.44 Поједностављена блок шема поступка добијања видео материјала прилагођеног захтевима субјективног теста.

Излазни фајлови поступака илустрованих на сликама Сл. 5.45 и Сл. 5.46 су некомпримовани фајлови настали поступком кодовања и идентичног су трајања као и изворни материјал (неколико десетина секунди до неколико минута). За издавање

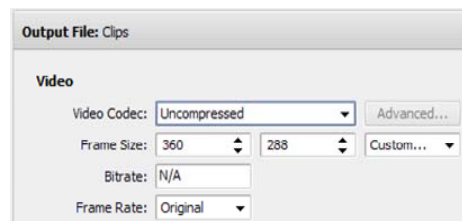
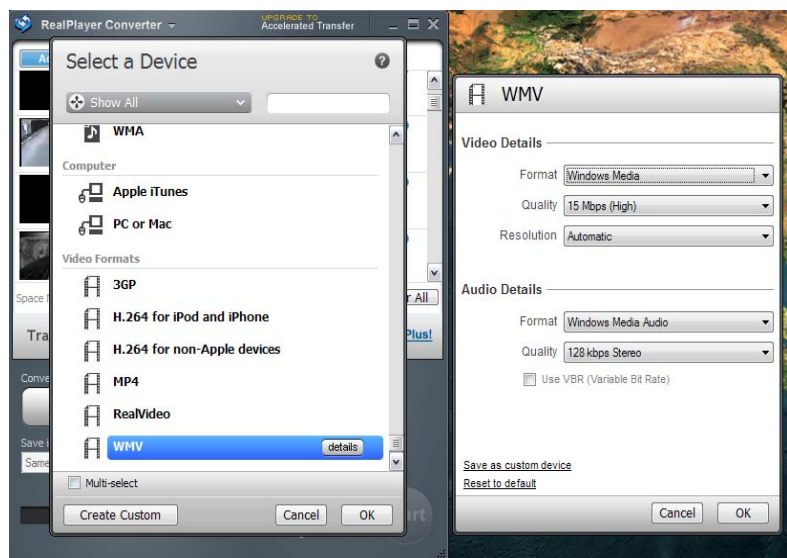
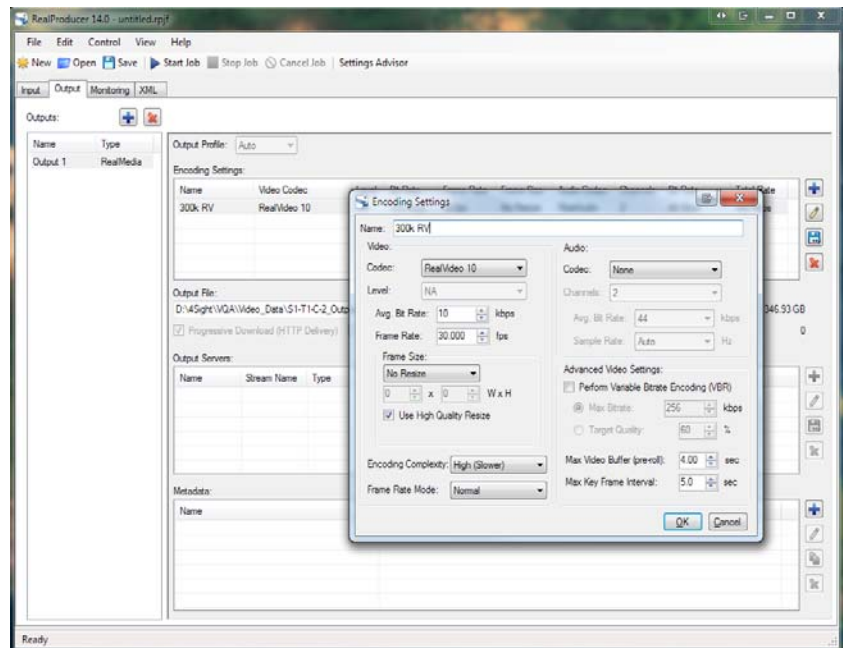
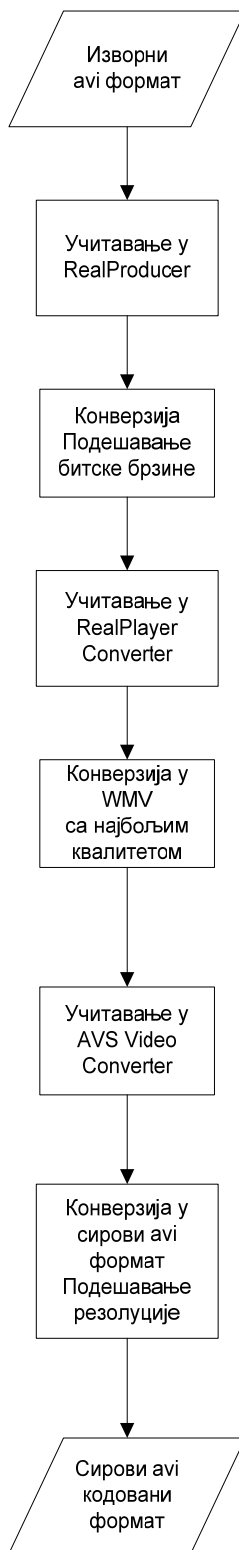
5. Процена квалитета видео заснована на очувању информација о градијенту

видео записа од по 10 s коришћен је софтвер VirtualDub. За сваки видео запис одбачене су прве и последње две секунде видео материјала. На овај начин су елиминисане појаве које се јављају на почетку и крају кодовања, тј. појаве у прелазним процесима током којих се врши прилагођење кодека самој секвенци. Блок шема овог поступка са изгледима менија током рада приказана је на Сл. 5.47.



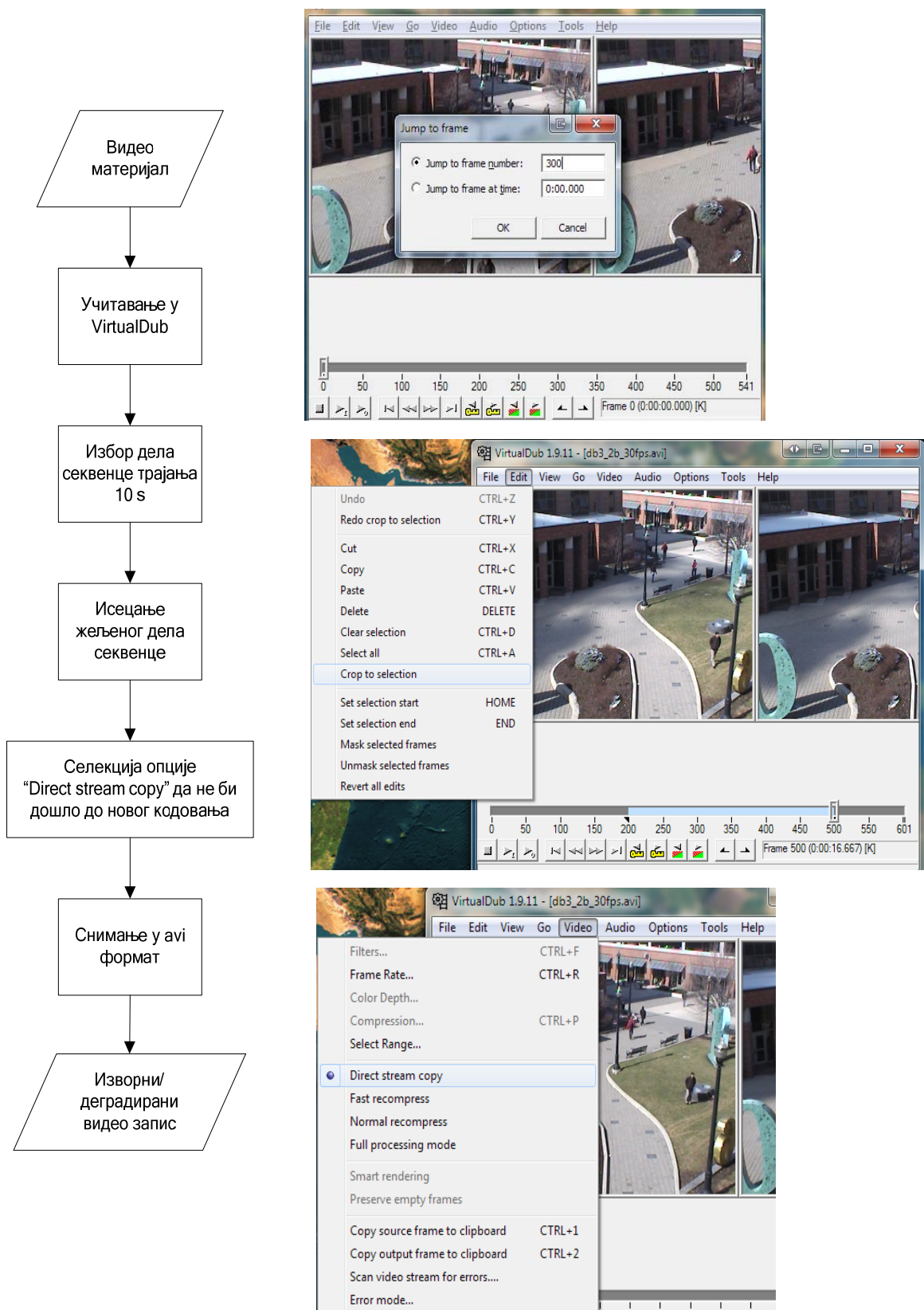
Сл. 5.45 Поједностављена блок шема MPEG-4 Part2 и H.264/AVC поступка кодовања са одговарајућим приказима екрана током кодовања.

5. Процена квалитета видео заснована на очувању информација о градијенту



Сл. 5.46 Поједностављена блок шема RealVideo 10 поступка кодовања са одговарајућим приказима екрана током кодовања.

5. Процена квалитета видео заснована на очувању информација о градијенту



Сл. 5.47 Поједностављена блок шема поступка издвајања репрезентативног дела секвенце са одговарајућим приказима екрана током рада.

5.12.7.Опрема за испитивање

Хардвер

Четири (три) некомпримована видео су истовремено приказивана на свакој страници током тестирања (оригинал и три његове деградације у STrial 1 сету, односно оригинал и две његове деградације у STrial 0 сету), Сл. 5.48. Због тога је хардвер изабран тако да се прилагоди великој количини података и да има довољно меморије за складиштење у RAM.

Изабрани хардвер се може прилагодити и сличним тестовима за неке будуће потребе. На пример, CPU и GPU су довољно јаки да подрже декодовање више видео у HD резолуцији уколико то буде потребно а систем за приказ се може користити у конфигурацији са више монитора.



Сл. 5.48 Изглед странице за приказ и навигацију (STrial 1 сет).

Систем за приказ

За приказ су изабрани IPS монитори због добре репродукције боје и природног контраста. Резолуција монитора за приказ је током субјективних тестова смањена са максималних 1920x1200 на 1280x800 пиксела са фиксним аспектним

скалирањем. Ово је урађено како би се прилагодило димензијама секвенци и осигурало да субјекти могу уочити детаље у сваком од видео записа. Осветљеност у просторији у којој су спроведени тестови су подешени да одговарају реалним ситуацијама. Тако је осветљај монитора подешен на 200 cd/m^2 а ниво осветљаја позадине на 20 lux.

Систем за приказ и хардверску конфигурацију чине:

Систем за приказ

- Dell U2412M IPS 24” Monitor – резолуције 1920x1200

PC

- Intel i5 3450 IvyBridge Quad Core CPU – 3.1 GHz
- 12 GB, 1600 MHz, DDR3 меморија
- два 240 GB SATA 6 Gb/s SSD диска у RAID 1 конфигурацији
- nVidia GTX 560Ti графички адаптер.

Калибрација

Spyder4 Elite screen calibrator.

5.12.8. Софтвер за тестирање

Софтвер за приказ, навигацију и прикупљање субјективних скорова – VQATrial, омогућава вишеструко понављање видео фајлова са захтеваним бројем кадрова у секунди, са добром синхронизацијом и са континуитетом у приказу. Софтвер је тестиран у STrial 0 сету, након чега су направљене ситне измене пре спровођења крајњег (меродавног), STrial 1 сета.

5.12.9. Процедура тестирања

Колекција видео секвенци је подељена на подскупове, где се подскупови секвенци приказују посматрачу у оквиру једне стране, као што је илустровано на Сл. 5.48. Један подскуп чине изворна секвенца и њене три модификације (од три кодека)

којима посматрач оцењује квалитет у опсегу од 0 до 5. Оцена 0 говори да модификована секвенца нема сличности са изворном, док оцена 5 говори да се не може уочити разлика изворне и деградиране секвенце.

Посматрачи су у оцењивању учествовали на добровољној основи а пре спровођења теста су им дате инструкције о самом тесту. Процедура тестирања своди се на:

- упознавање субјеката са циљевима тестирања, задацима које треба да ураде и информацијама које ће се прикупити током тестирања,
- упознавање субјеката да је тестирање на добровољној основи,
- субјектима се пружа помоћ у уношењу иницијалних података и током анализе прве стране теста, док се навикну на софтвер и процедуру тестирања,
- тест се ради динамиком која одговара посматрачу.

Планирано време за анализу видео записа је 30 минута, иако је време спровођења теста значајно варијало од посматрача до посматрача. Материјал који се приказује посматрачима даје се у случајном редоследу како би се елиминисао било какав вид учења.

5.12.10. Анализа прикупљених података

Процедура прикупљања субјективних импресија је аутоматизована. За сваког посматрача одзиви се складиште на жељену локацију. На крају тестирања у једном директоријуму се налазе импресије од свих учесника. На Сл. 5.49 приказан је садржај директоријума са листингом фајлова субјективних импресија за сет STrial 1.

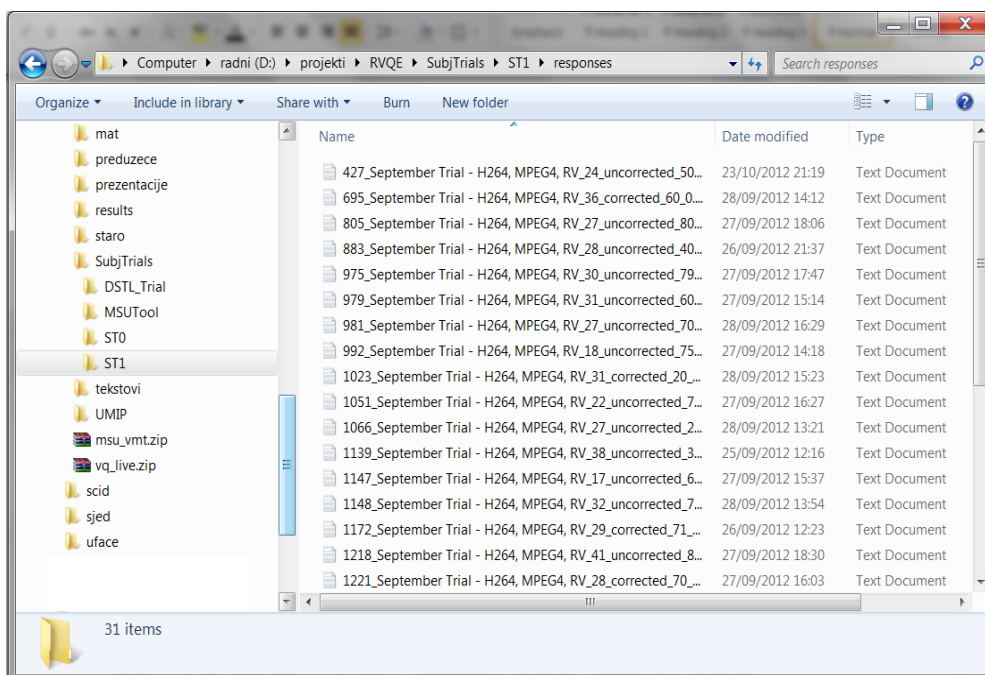
Конвенција назива текстуалних фајлова са субјективним импресијама је следећа:

- случајно генерисан идентификациони број (ID) посматрача, нпр. “427” за првог посматрача,
- предефинисани низ карактера који идентификује сет који се користи, за STrial 1 је то низ “September Trial - H264, MPEG4, RV”,

5. Процена квалитета видеа заснована на очувању информација о градијенту

- број година посматрача, рецимо 24 за првог посматрача,
- стање вида посматрача (без корекције или са корекцијом),
- степен искуства посматрача у процени квалитета видеа и субјективним тестовима, за првог посматрача је 50%.

На овај начин прикупљене су основне информације о посматрачу, а који је идентификован кроз ID број.



Сл. 5.49 Листинг фајлова са субјективним импресијама на STrial 1 сету.

Током спровођења субјективних тестова на STrial 0 сету коришћена је једноставнија конвенција текстуалних датотека са субјективним импресијама, тј. коришћени су број година, стање вида и искуство посматрача.

Као што се види са Сл. 5.49 субјективне импресије једног посматрача снимају се у текстуалну датотеку (.txt) и то према унапред договореном формату. За један сет приказаних секвенци (референтна + три секвенце са деградацијом) формат записа се поједностављено може представити у форми:

RefVidFname, No. replays, StartTime, EndTime, DegVid1, DegVid1_Score, DegVid2, DegVid2_Score, DegVid3, DegVid3_Score

где су уведена следећа поља:

- *RefVidFname* – име оригиналне (изворне) секвенце,
- *No. replays* – укупан број прегледа видео записа из једног сета пре преласка на следећи сет,
- *StartTime* – датум и време када је видео сет први пут приказан на екрану,
- *EndTime* – датум и време када је посматрач прешао на оцењивање следећег сета,
- *DegVidX* где је $X \in \{1, 2, 3\}$ – име секвенце са деградацијом,
- *DegVidX_Score* где је $X \in \{1, 2, 3\}$ – нумеричка представа квалитета коју је посматрач дао секвенци са деградацијом.

Садржај текстуалне датотеке за једног од учесника STrial 1 сета дат је на Сл. 5.50.

<p>Balcony3_Vis_653-703.avi, 4 replays, 27/09/2012 13:58:13, 27/09/2012 13:59:53, Balcony3_Vis_mpeg4_64kbps_653-703.avi, 3.5, Balcony3_Vis_RV_64kbps_653-703.avi, 2, Balcony3_Vis_h264_64kbps_653-703.avi, 4 Balcony3_Vis_653-703.avi, 2 replays, 27/09/2012 13:59:54, 27/09/2012 14:00:45, Balcony3_Vis_RV_128kbps_653-703.avi, 2, Balcony3_Vis_h264_128kbps_653-703.avi, 4, Balcony3_Vis_mpeg4_128kbps_653-703.avi, 3.5 Balcony3_Vis_653-703.avi, 3 replays, 27/09/2012 14:00:45, 27/09/2012 14:02:00, Balcony3_Vis_RV_10kbps_653-703.avi, 3, Balcony3_Vis_mpeg4_10kbps_653-703.avi, 1.5, Balcony3_Vis_h264_10kbps_653-703.avi, 2.5</p> <p>S1-T1-C-2_126-136.avi, 1 replays, 27/09/2012 14:02:00, 27/09/2012 14:02:27, S1-T1-C- 2_mpeg4_128kbps_126-136.avi, 3.5, S1-T1-C-2_h264_128kbps_126-136.avi, 4, S1-T1-C- 2_RV_128kbps_126-136.avi, 2 S1-T1-C-2_126-136.avi, 1 replays, 27/09/2012 14:02:27, 27/09/2012 14:02:49, S1-T1-C- 2_RV_64kbps_126-136.avi, 3, S1-T1-C-2_h264_64kbps_126-136.avi, 4, S1-T1-C- 2_mpeg4_64kbps_126-136.avi, 2.5 S1-T1-C-2_126-136.avi, 2 replays, 27/09/2012 14:02:49, 27/09/2012 14:03:18, S1-T1-C- 2_RV_10kbps_126-136.avi, 2.5, S1-T1-C-2_mpeg4_10kbps_126-136.avi, 1, S1-T1-C- 2_h264_10kbps_126-136.avi, 3</p>

Сл. 5.50 Садржај текстуалне датотеке за једног од учесника STrial 1 сета.

5.12.11. Опис првог субјективног теста – тест STrial 0

Иницијални тест је послужио да се тестирају протокол и софтвер, прикупе подаци и изврши њихова анализа. Генерално, тест је успео и дао је употребљиве

податке као и бројне смернице које су послужиле за побољшање софтвера и протокола у наредним тестовима.

Тест је спроведен у канцеларији Универзитета у Манчестеру. У соби нису постојали прозори тако да су услови спровођења теста били постојани. Ниво осветљености се није мењао и остао је на нивоу осветљаја канцеларије. Просторија је на таквој локацији да су субјекти били изоловани од спољних утицаја током теста.

Посматрачи су седели у стандардној канцеларијској столици а опрема за тестирање се налазила на канцеларијском столу. Испитаници су били слободни да опрему и могуће услове подешавају према својим афинитетима.

У тесту је учествовало 15 волонтера. Њихово професионално искуство се кретало од студената у области обраде слике, предавача и истраживача па до административних радника.

У току теста коришћено је осам изворних секвенци и секвенце са деградацијама насталих коришћењем два начина кодовања, MPEG-4 Part 2 и H.264/AVC, за три битске брзине: 128, 64 и 10 kbps. Тако је укупно евалуирано 54 секвенце где су на екрану истовремено приказиване по 3 секвенце (оригинал + 2 модификације).

Посматрачима је у просеку требало 20 минута да заврше тест.

Резолуција приказиваних видеа је фиксирана на 360x288 пиксела.

Додатне информације о референтним видео секвенцама могу се наћи у Таб. 5.20.

5.12.12. Опис другог субјективног теста – тест STrial 1

Субјективни STrial 1 тест је спроведен са циљем да се прикупе субјективни подаци који ће се користити у валидацији објективних мера процене. Како би се прикупило што више субјективних резултата тестови су рађени на две локације.

5. Процена квалитета видеа заснована на очувању информација о градијенту

Таб. 5.20 Додатне информације о изворним секвенцама STrial 0 сета – назив секвенце, временски исечак секвенце коришћен у анализи, жељена и достигнута средња битска брзина.

abandoned_luggage 00:00:12 to 00:00:22	MPEG-4 Part 2	H.264/AVC
128 kbps	129 kbps	127 kbps
64 kbps	65 kbps	63 kbps
10 kbps	10 kbps	10 kbps
crowding_detection 00:12:30 to 00:12:40	MPEG-4 Part 2	H.264/AVC
128 kbps	107 kbps	128 kbps
64 kbps	63 kbps	64 kbps
10 kbps	7 kbps	10 kbps
db3_2b_30fps 00:00:10 to 00:00:20	MPEG-4 Part 2	H.264/AVC
128 kbps	135 kbps	134 kbps
64 kbps	71 kbps	69 kbps
10 kbps	11 kbps	12 kbps
parked_vehicle_day 00:01:33 to 00:01:43	MPEG-4 Part 2	H.264/AVC
128 kbps	130 kbps	130 kbps
64 kbps	66 kbps	65 kbps
10 kbps	10 kbps	12 kbps
Parked_Vehicle_night 00:01:02 to 00:01:15	MPEG-4 Part 2	H.264/AVC
128 kbps	130 kbps	130 kbps
64 kbps	65 kbps	66 kbps
10 kbps	10 kbps	11 kbps
20060708_01_140x440 00:00:10 to 00:00:20	MPEG-4 Part 2	H.264/AVC
128 kbps	129 kbps	131 kbps
64 kbps	67 kbps	69 kbps
10 kbps	11 kbps	15 kbps
ME_Training 00:06:18 to 00:06:28	MPEG-4 Part 2	H.264/AVC
128 kbps	130 kbps	129 kbps
64 kbps	67 kbps	64 kbps
10 kbps	10 kbps	13 kbps
S1-T1-C-2 00:00:02 to 00:00:12	MPEG-4 Part 2	H.264/AVC
128 kbps	132 kbps	131 kbps
64 kbps	66 kbps	67 kbps
10 kbps	11 kbps	11 kbps

Прва локација је на Универзитету у Манчестеру – Imaging Science Department (ISBE) а друга локација је у научном парку Универзитета у Манчестеру – Cubic Motion.

ISBE окружење

Тестови на овој локацији су спроведени у конференцијској сали. У сали нису постојали прозори а ниво осветљаја је контролисан помоћу лампе на 20 lux. Просторија је изолована од спољних утицаја.

Посматрачи су седели у стандардној канцеларијској столици а опрема за тестирање се налазила на канцеларијском столу. Испитаници су били слободни да опрему, угао и даљину посматрања подешавају према својим афинитетима.

Cubic Motion окружење

Тест је спроведен у пространој и тихој канцеларији Cubic Motion научног парка. Просторија је осветљена флуоресцентном светлошћу и у њој се налазио један прозор у доњем делу једног од зидова али без директног излагања сунчевим зрацима. На тај начин ниво осветљености позадине се није могао контролисати.

Субјекти

У субјективном тесту на ISBE локацији учествовало је 18 студената и особља управе. Њихово професионално искуство се кретало од студената у области обраде слике, предавача и истраживача па до административних радника.

Додатно је у Cubic Motion парку у субјективном тесту учествовало 13 волонтера. Како је Cubic Motion компанија која се бави графиком и анимацијом, сваки од учесника на овој локацији је био веома искусан у уређивању и евалуацији графике и видео детекције.

Видео подаци

У току теста коришћено је девет изворних секвенци и секвенце са деградацијама насталих коришћењем три начина кодовања, MPEG-4 Part 2,

5. Процена kvaliteta videa zasnovana na ocuvanju informacija o gradijentu

H.264/AVC и RealVideo 10, за три битске брзине: 128, 64 и 10 kbps. Тако је укупно евалуирана 81 секвенца где су на екрану истовремено приказиване по 4 секвенце (оригинал + 3 модификације).

Посматрачима је у просеку требало 30 минута да заврше тест.

Резолуција приказиваних videa је фиксирана на 360x288 пиксела.

Додатне информације о референтним видео секвенцама могу се наћи у Таб. 5.21.

Таб. 5.21 Додатне информације о изворним секвенцама STrial 1 сета – назив секвенце, временски исечак секвенце коришћен у анализи, жељена и достигнута средња битска брзина.

abandoned_luggage 00:00:36 to 00:00:46	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	129 kbps	127 kbps	128 kbps
64 kbps	65 kbps	63 kbps	64 kbps
10 kbps	10 kbps	10 kbps	10 kbps
crowding_detection 00:15:34 to 00:15:44	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	107 kbps	128 kbps	128 kbps
64 kbps	63 kbps	64 kbps	64 kbps
10 kbps	7 kbps	10 kbps	10 kbps
db3_2b_30fps 00:00:02 to 00:00:12	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	135 kbps	134 kbps	128 kbps
64 kbps	71 kbps	69 kbps	64 kbps
10 kbps	11 kbps	12 kbps	10 kbps
parked_vehicle_day 00:01:18 to 00:01:28	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	130 kbps	130 kbps	128 kbps
64 kbps	66 kbps	65 kbps	64 kbps
10 kbps	10 kbps	12 kbps	10 kbps
Parked_Vehicle_night 00:01:21 to 00:01:22	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	130 kbps	130 kbps	128 kbps
64 kbps	65 kbps	66 kbps	64 kbps
10 kbps	10 kbps	11 kbps	10 kbps
20060708_01_140x440 00:00:10 to 00:00:20	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	129 kbps	131 kbps	128 kbps
64 kbps	67 kbps	69 kbps	64 kbps
10 kbps	11 kbps	15 kbps	10 kbps

5. Процена квалитета видеа заснована на очувању информација о градијенту

ME_Training 00:02:25 to 00:02:35	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	130 kbps	129 kbps	128 kbps
64 kbps	67 kbps	64 kbps	64 kbps
10 kbps	10 kbps	13 kbps	10 kbps

S1-T1-C-2 00:01:26 to 00:01:36	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	132 kbps	131 kbps	128 kbps
64 kbps	66 kbps	67 kbps	64 kbps
10 kbps	11 kbps	11 kbps	10 kbps

Balcony3_Vis 00:06:53 to 00:07:03	MPEG-4 Part 2	H.264/AVC	RealVideo 10
128 kbps	133 kbps	129 kbps	128 kbps
64 kbps	68 kbps	65 kbps	64 kbps
10 kbps	11 kbps	12 kbps	10 kbps

5.12.13. Статистика субјективних скорова STrial 0 теста

Субјективни одзиви који су прикупљени и архивирани у један директоријум су анализирани и њихове основне статистике су евалуиране на нивоу појединачних видеа који су приказивани током теста. У Таб. 5.22 дате су статистике субјективних скорова за сваки од видеа STrial 0 сета.

Оно што је очигледно из Таб. 5.22 јесте да је већина евалуираних видео секвенци добила доста широк опсег вредности квалитета. На пример, секвенца која је обележена са 43 добила је оцене од 1.6 до 4.8. Ово показује да људи на различите начине примећују/схватају апсолутне нивое квалитета.

Таб. 5.22 Статистике субјективних скорова видео секвенци STrial 0 сета.

р.бр. видеа	Минимални скор	Максимални скор	Средња вредност	Медијана	Стандардна девијација	Стандардна грешка
1	0	2.8	1.93333	2	0.716805	0.185078
2	0	1.2	0.553333	0.5	0.396172	0.102291
3	0.4	4.8	3.82667	4.1	1.08395	0.279875
4	0.6	4.9	2.66	3	1.14005	0.29436
5	0.5	5	3.98	4	1.14966	0.296841
6	0.6	5	3.91333	4	1.14322	0.295178
7	1.4	5	4.15333	4.5	0.841654	0.217314
8	0.9	4.5	3.47333	3.7	0.909055	0.234717
9	0	2.7	0.826667	0.5	0.772257	0.199396

5. Процена квалитета видеа заснована на очувању информација о градијенту

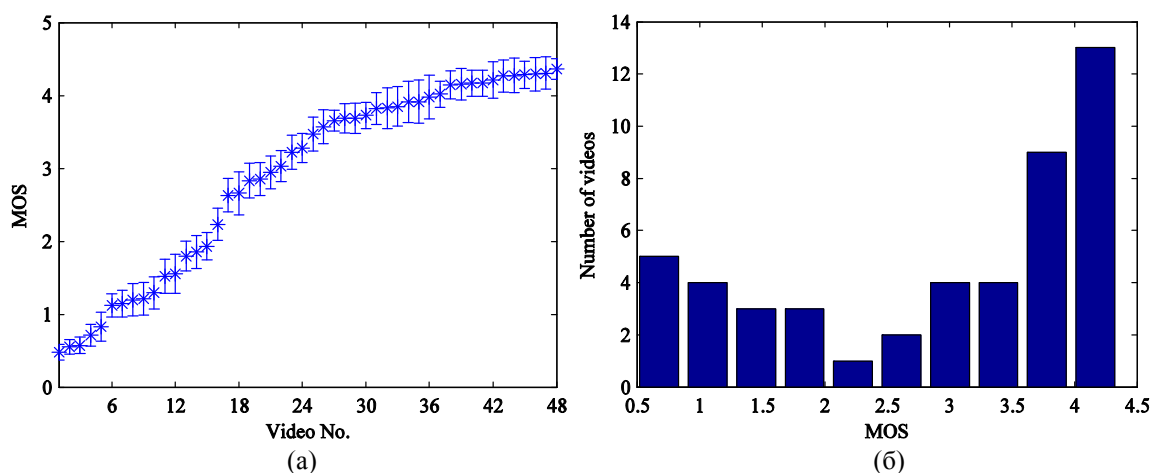
р.бр. видеа	Минимални скор	Максимални скор	Средња вредност	Медијана	Стандардна девијација	Стандардна грешка
10	1.1	4.2	3.22	3.3	0.915111	0.236281
11	1.2	4.8	3.82	4	0.844478	0.218043
12	0	3.1	1.2	1	0.85105	0.21974
13	0.6	3.5	1.85333	2	0.879827	0.22717
14	2.4	5	4.28667	4.5	0.72493	0.187176
15	1.1	4.4	3.28	3.5	0.774781	0.200048
16	0	1.5	0.573333	0.4	0.443149	0.114421
17	1.3	5	4.27333	4.5	0.916099	0.236536
18	1.7	4.8	3.72667	4	0.696385	0.179806
19	2.2	4.9	4.16667	4.5	0.70068	0.180915
20	0.1	3.4	1.29333	1	0.85479	0.220706
21	2.5	4.9	4.16667	4.5	0.684175	0.176653
22	1.9	4.7	3.68667	4	0.803445	0.207449
23	0.4	3.5	1.52	1.5	0.900952	0.232625
24	2	5	4.29333	4.5	0.8908	0.230003
25	2.7	4.5	3.65333	3.6	0.548852	0.141713
26	0	1.7	0.713333	0.5	0.568038	0.146667
27	1.1	5	4.21333	4.5	0.973849	0.251447
28	0	4.3	1.55333	1.4	1.04462	0.269721
29	1	5	3.91333	4.2	1.10897	0.286334
30	0.9	5	3.84667	3.8	1.04667	0.27025
31	0.7	4.3	2.94667	3	0.867564	0.224004
32	1.5	5	4.30667	4.5	0.863933	0.223067
33	0.5	3.2	1.8	1.8	0.806226	0.208167
34	0.8	3.7	2.23333	2.2	0.854958	0.220749
35	0.8	4.2	3.03333	3	0.819988	0.21172
36	0	1.3	0.48	0.4	0.414384	0.106994
37	0.2	2.4	1.12	1	0.624729	0.161304
38	1.3	4.6	2.85333	3	0.869209	0.224429
39	1.7	4.8	4.14667	4.5	0.754857	0.194903
40	0.3	3	1.14667	0.9	0.720978	0.186156
41	0.9	4.7	3.57333	3.8	0.901163	0.232679
42	0.9	4.3	2.83333	2.7	0.922471	0.238181
43	1.6	4.8	3.68667	4	0.780903	0.201628
44	0.9	3.8	2.63333	2.5	0.88533	0.228591
45	2	5	4.26667	4.5	0.864099	0.223109
46	2.2	4.7	4.02	4.2	0.693026	0.178939
47	0	2.5	1.21333	1.2	0.86674	0.223791
48	3	4.9	4.36	4.6	0.552656	0.142695

Просечна субјективна оцена (средња вредност из табеле) има мању варијацију од медијане скорова, што може бити реалније у ситуацијама као што је анализа

videа са редним бројем 43, где је расподела субјективних скорова веома несиметрична.

Стандардна девијација скорова је релативно висока и креће се од 0.4 до 1.15. Нормализацијом са бројем субјеката који су учествовали у тесту (15 субјеката) долази се до стандардне грешке која показује неодређеност скорова од $\pm 10\%$.

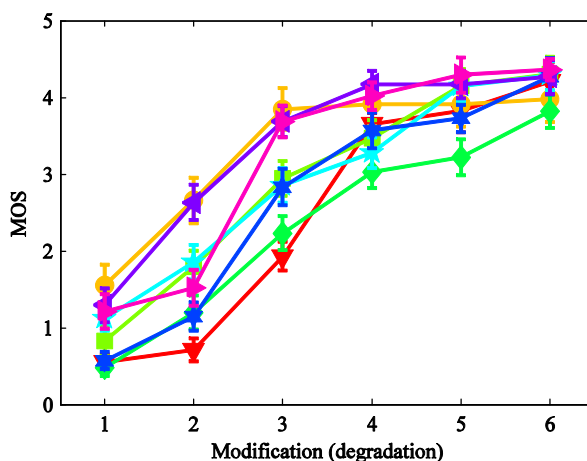
Додатне илустрације анализе субјективних скорова приказане су на Сл. 5.51, где су дате сортиране вредности субјективних скорова у растућем редоследу као и расподела скорова у облику хистограма са 10 подопсега.



Сл. 5.51 Анализа резултата STrial 0 сега: (а) сортиране MOS вредности са интервалима поузданости ($MOS \pm SE$) и (б) хистограм MOS вредности.

Може се уочити скоро линеаран тренд сортираних вредности са добрим распоном квалитета видеа јер MOS вредности покривају око 80% расположивог динамичког опсега (вредности се крећу од 0.5 до 4.5). Такође, са хистограма се уочава да се трећина вредности субјективних скорова налази испод вредности 2 (лош квалитет) што је највероватније због секвенци са битским протоком од 10 kbps (16 од 48 секвенци). Преостале секвенце (битски проток од 64 и 128 kbps) су сврстане у секвенце средњег и доброг квалитета, а због максимума хистограма који се помера ка високим MOS вредностима може се закључити да битски проток од 64 kbps може дати секвенце доброг квалитета.

Један од начина анализе рангирања вредности квалитета добијених за STrial 0 сет приказан је на Сл. 5.52. Субјективни скорови су анализирани према референтним секвенцама, тј. свака линија се односи на видео секвенце које потичу од истог изворног записа.



Сл. 5.52 Расподела субјективних скорова за различите референтне секвенце у STrial 0 сету.

Сл. 5.52 илуструје добар распон видео квалитета. То указује да је тест добро уравнотежен/усклађен (нпр. зелена и љубичаста линија показују добру расподелу скорова од ниског до високог квалитета) тако да садржи комплексне видео секвенце које у потпуности упошљавају енкодер као и мање изазовне секвенце за које кодеци веома брзо достижу висок квалитет који не зависи од избора кодека или битске брзине (наранџаста линија која има засићење са десне стране приказа – ова линија одговара модификацијама “Crowding detection” видео секвенце са екстремно ниским темпом од 2 fps).

Поузданост субјективних скорова на појединачним изворним секвенцама је мерена кроз стандардну девијацију субјективних скорова придружених модификацијама једне изворне секвенце, што је дато у Таб. 5.23. Најмању стандардну девијацију субјективних скорова има сет настао од секвенце “Crowding detection”. Ово може бити последица малог темпа (2 fps) који омогућава да посматрачи имају довољно времена да анализирају просторни квалитет сваког од кадрова чиме су скорови поузданији а тиме и рангирање различитих кодека и битских брзина.

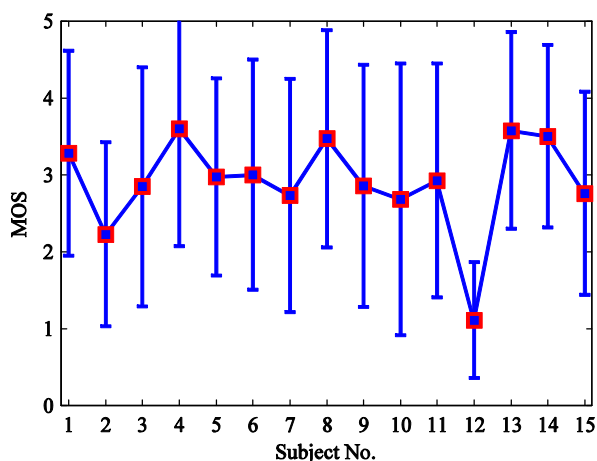
Таб. 5.23 Стандардне девијације субјективних скорова деградираних STrial 0 видео секвенци насталих од исте изворне секвенце.

Референтна видео секвенца	Стандардна девијација
20060708_01_140x440_10-20.avi	1.2557
ME_Training_618-628.avi	1.2818
Parked_Vehicle_night_102-112.avi	1.1857
S1-T1-C-2_2-12.avi	1.4295
abandoned_luggage_12-22.avi	1.6326
crowding_detection_1230-1240.avi	0.9973
db3_2b_30fps_10-20.avi	1.3702
parked_vehicle_day_133-143.avi	1.4996

Један од индикатора успешног субјективног теста јесте поузданост посматрача. Овде смо се бавили анализом конзистентности њихових одговора. Наравно, заинтересовани смо за тестове са високим нивоом поузданости где се субјекти у потпуности (100%) слажу. У стварности ово се никада не дешава тако да извесна несигурност постоји како у апсолутном квалитету тако и у рангирању квалитета.

Основне статистике субјективних скорова појединачних посматрача приказане су на Сл. 5.53, где су дате средње вредности и стандардне девијације њихових скорова ($MOS \pm \sigma$). Уочава се да постоје релативно мале варијације средњих вредности субјективних скорова индивидуалних посматрача.

Статистика једног од посматрача, посматрача са редним бројем 12, због коришћеног опсега квалитета од 0.1 до 3 одступа од статистика осталих испитаника, како по средњој вредности тако и по стандардној девијацији. Неке од студија видео квалитета и квалитета слике би искључиле оваквог посматрача из збирне анализе, то у овом случају није урађено јер иако је посматрач користио различит динамички опсег скорова, по питању рангова квалитета његови одговори се нису разликовали од одговора осталих субјеката.



Сл. 5.53 Средње вредности и стандардне девијације субјективних скорова ($MOS \pm \sigma$ опсер) за сваког од посматрача STrial 0 сета.

5.12.14. Статистика субјективних скорова STrial 1 теста

Субјективни тест STrial 1 је спроведен у другој половини септембра месеца 2012. године, укључио је 31 субјекта који су евалуирали девет различитих видео сетова (референтних видеа) са 3x3 деградационом матрицом (9 деградираних видео секвенци по свакој од референтних, 3 кодека за 3 битске брзине).

Субјективни одзиви који су прикупљени и архивирани у један директоријум су анализирани и њихове основне статистике су евалуиране на нивоу појединачних видеа који су приказивани током теста. У Таб. 5.24 дате су статистике субјективних скорова за сваки од видеа STrial 1 сета.

Таб. 5.24 Статистике субјективних скорова видео секвенци STrial 1 сета.

р.бр. видеа	Минимални скор	Максимални скор	Средња вредност	Медијана	Стандардна девијација	Стандардна грешка
1	0.1	3.8	1.7323	1.3	1.1571	0.20782
2	2.3	5	4.0774	4.2	0.64276	0.11544
3	2.8	5	4.5097	4.7	0.55699	0.10004
4	0.4	3.4	1.6677	1.7	0.78119	0.14031
5	2.5	5	4.1806	4.4	0.71945	0.12922
6	2.7	5	4.3516	4.5	0.63239	0.11358
7	0.3	4.8	2.3677	2.5	1.1022	0.19797
8	3	5	4.3806	4.5	0.55343	0.099398
9	3.4	5	4.5774	4.8	0.49512	0.088925

5. Процена квалитета видеа заснована на очувању информација о градијенту

р.бр. видеа	Минимални скор	Максимални скор	Средња вредност	Медијана	Стандардна девијација	Стандардна грешка
10	0	5	3.9742	4.2	1.0185	0.18293
11	2.5	5	4.3452	4.7	0.79366	0.14255
12	2.1	5	4.4161	4.8	0.81857	0.14702
13	0.1	3.9	1.8452	2	0.93339	0.16764
14	2.2	5	3.9548	4	0.69514	0.12485
15	3.2	5	4.3645	4.5	0.56128	0.10081
16	0	2.8	1.1806	1.1	0.674	0.12105
17	2.5	4.8	3.7742	4	0.58136	0.10442
18	3	5	4.271	4.5	0.57746	0.10372
19	0.1	3	1.2774	1.2	0.71214	0.1279
20	2.4	5	4.0839	4.1	0.69814	0.12539
21	2.3	5	4.4387	4.7	0.74192	0.13325
22	0.4	3.8	2.7	3	0.87254	0.15671
23	0.7	5	4.1581	4.4	0.83177	0.14939
24	2.4	5	4.3516	4.5	0.63711	0.11443
25	0.8	4	2.6613	2.7	0.83254	0.14953
26	3.6	5	4.5903	4.7	0.40526	0.072787
27	0	5	4.4548	4.8	0.92117	0.16545
28	0	3.9	1.1387	1	0.88982	0.15982
29	1	4.5	2.9484	3	0.94335	0.16943
30	0.2	4.7	2.8226	3	1.0372	0.18629
31	0	2.9	1.0419	0.9	0.75929	0.13637
32	2.3	4.8	3.6871	3.9	0.71028	0.12757
33	3	5	4.3742	4.5	0.57269	0.10286
34	0.2	3.3	1.5258	1.5	0.82542	0.14825
35	1.5	5	3.9226	4.1	0.78728	0.1414
36	3.1	5	4.2323	4.5	0.60409	0.1085
37	0	3	1.4903	1.5	0.78032	0.14015
38	2.5	5	4.2097	4.6	0.7884	0.1416
39	2.8	5	4.2645	4.6	0.79101	0.14207
40	0	2.6	0.88065	0.8	0.64157	0.11523
41	1.4	4.7	3.2645	3.3	0.79647	0.14305
42	2.6	4.8	3.8548	4	0.61635	0.1107
43	0	1.8	0.61935	0.5	0.48677	0.087427
44	1.2	4.1	2.6355	2.5	0.71719	0.12881
45	2.1	4.8	3.5323	3.5	0.66803	0.11998
46	0	1.8	0.5871	0.5	0.53338	0.095798
47	0.5	4.2	2.7	2.9	0.93131	0.16727
48	1.5	4.8	3.5161	3.8	0.81326	0.14607
49	0	2.9	1.029	1	0.73493	0.132

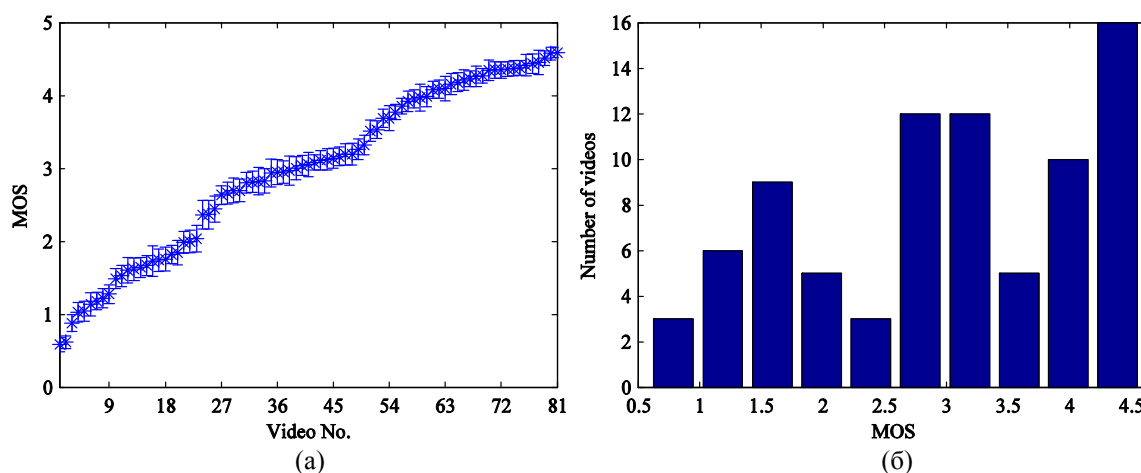
5. Процена квалитета видеа заснована на очувању информација о градијенту

р.бр. видеа	Минимални скор	Максимални скор	Средња вредност	Медијана	Стандардна девијација	Стандардна грешка
50	0.6	5	3.6903	4	0.94458	0.16965
51	2.1	5	3.9839	4	0.75458	0.13553
52	0.1	3	1.2194	1.1	0.75473	0.13555
53	1	4.6	3.3226	3.4	0.73016	0.13114
54	0	5	4.0968	4.4	0.92141	0.16549
55	0	3.6	1.6032	1.6	0.96695	0.17367
56	0.7	4.2	2.3645	2.4	1.0837	0.19463
57	0.5	4.2	2.4419	2.5	0.99992	0.17959
58	0.5	3.6	1.9968	2	0.86197	0.15481
59	1.5	4.4	2.8161	3	0.80627	0.14481
60	1.4	4.4	3.1645	3.2	0.71113	0.12772
61	0	3.9	2.0387	2	1.023	0.18373
62	1.1	4.3	2.9516	3	0.87135	0.1565
63	1.5	4.4	3.1226	3.1	0.81351	0.14611
64	0	3.7	1.9871	2	0.84921	0.15252
65	0	4.5	2.9355	3	1.06	0.19039
66	0.7	4.5	2.971	3.3	1.1175	0.20071
67	0.1	3.1	1.7419	1.7	0.84883	0.15245
68	1.8	4.6	3.0387	3	0.76362	0.13715
69	1.7	4.4	3.1387	3.2	0.75924	0.13636
70	0	3	1.6194	1.6	0.853	0.1532
71	1	4.4	2.8032	2.9	0.80062	0.1438
72	1.5	4.5	3.1968	3.1	0.83566	0.15009
73	0.1	3.5	1.7548	1.6	0.90805	0.16309
74	1	4.3	2.8323	2.8	0.92678	0.16645
75	1.2	4.5	3.0516	3	0.92335	0.16584
76	0.3	3.2	1.8194	1.9	0.70447	0.12653
77	0.2	4.1	3.0065	3	0.88541	0.15902
78	1.2	4.4	3.1935	3.5	0.81769	0.14686
79	0.2	3.3	1.6419	1.6	0.77965	0.14003
80	2.2	4.3	3.0806	3	0.58333	0.10477
81	1.5	4.4	3.1129	3	0.74955	0.13462

Из приказаних резултата је поново уочљиво да је већина деградираних секвенци добила широк опсег вредности скорова, као на пример секвенца обележена са редним бројем 73, где су субјективне оцене из распона од 0.1 до 3.5. Ово је наравно последица субјективног искуства, јер неки од посматрача виде секвенце на

различит начин од осталих. Стандардне девијације скорова су већином мање од 1, док је стандардна грешка мања од 10% од вредности укупног опсега квалитета. Пад вредности стандардне грешке у односу на грешку STrial 0 сета је настао из два разлога: (1) већег броја испитаника у субјективном тесту, чиме се смањује неодређеност мерења перцептуалног квалитета и (2) додатног искуства испитаника који су већ учествовали у STrial 0 тесту.

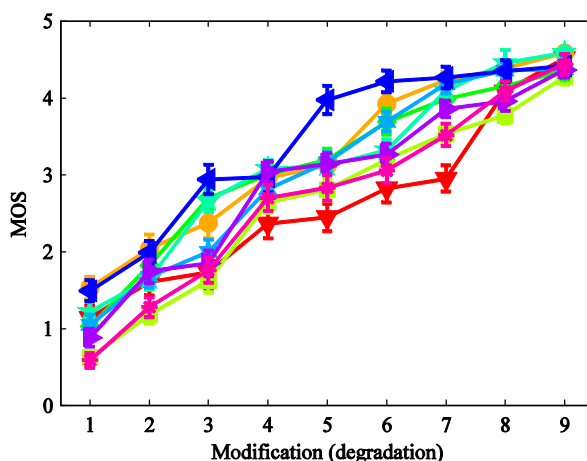
Додатне илустрације анализе субјективних скорова приказане су на Сл. 5.54, где су дате сортиране вредности субјективних скорова у растућем редоследу као и расподела скорова у облику хистограма са 10 подопсега.



Сл. 5.54 Анализа резултата STrial 1 сета: (а) сортиране MOS вредности са интервалима поузданости ($MOS \pm SE$) и (б) хистограм MOS вредности.

Поново се може уочити скоро линеаран тренд сортираних вредности са добрим распоном квалитета видеа јер MOS вредности покривају око 80% расположивог динамичког опсега (вредности се крећу од 0.5 до 4.5). Такође, са хистограма се уочавају три локална максимума који највероватније одговарају трима битским брзинама коришћеним за трансфер података (10, 64 и 128 kbps).

Рангирање субјективних вредности добијених за STrial 1 сет приказано је на Сл. 5.55. Субјективни скорови су анализирани према референтним секвенцама, тј. свака линија се односи на видео секвенце које потичу од истог изворног записа.



Сл. 5.55 Распореда субјективних скорова за различите референтне секвенце у STrial 1 сету.

Сл. 5.55 илуструје добар распон видео квалитета. То указује да је тест добро уравнотежен/усклађен (нпр. црвена и љубичаста линија показују добру расподелу скорова од ниског до високог квалитета) тако да садржи комплексне видео секвенце које у потпуности упошљавају енкодер као и мање изазовне секвенце за које кодеци веома брзо достижу висок квалитет који не зависи од избора кодека или битске брзине (плава линија која има засићење са десне стране приказа – ова линија одговара модификацијама “Crowding detection” видео секвенце са екстремно ниским темпом од 2 fps). Имајући у виду овакву расподелу субјективних скорова може се рећи да су тест секвенце перцептуално сепарабилне, па је за очекивати добро раздвајање и објективних скорова квалитета односно добро раздвајање скорова на дијаграмима расипања.

Поузданост субјективних скорова на појединачним изворним секвенцама је мерена кроз стандардно одступање субјективних скорова придружених модификацијама једне изворне секвенце, што је дато у Таб. 5.25. Може се говорити о конзистентности стандардних одступања субјективних скорова по изворним секвенцама. Генерално, стандардно одступање субјективних скорова се смањило у односу на стандардно одступање добијено за STrial 0 сет, јер се у овом случају по изворној секвенци анализира 9 за разлику од 6 модификација у претходном сету.

5. Процена квалитета видео заснована на очувању информација о градијенту

Таб. 5.25 Стандардне девијације субјективних скорова деградираних STrial 1 видео секвенци насталих од исте изворне секвенце.

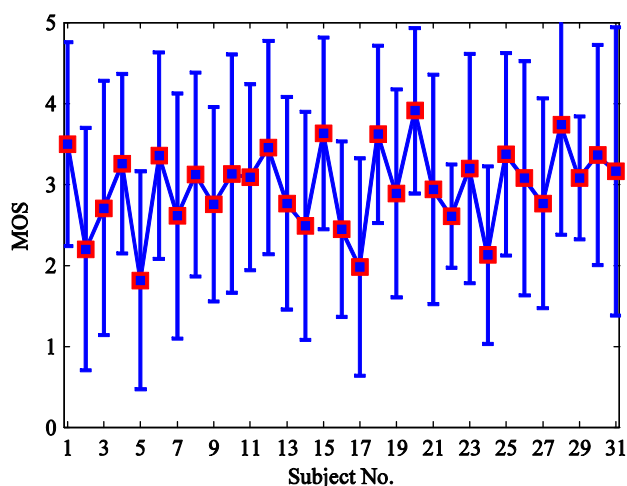
Референтна видео секвенца	Стандардна девијација
20060708_01_140x440_10-20.avi	1.1156
Balcony3_Vis_653-703.avi	1.2416
ME_Training_225-235.avi	1.2374
Parked_Vehicle_night_121-131.avi	1.1678
S1-T1-C-2_126-136.avi	1.2809
abandoned_luggage_36-46.avi	1.1056
crowding_detection_1534-1544.avi	1.1137
db3_2b_30fps_2-12.avi	1.1698
parked_vehicle_day_118-128.avi	1.1012

У Таб. 5.26 дате су вредности сагласности субјеката у STrial 1 тесту. Сагласност је мерена за сваки деградациони сет, тј. по свакој референтној секвенци и то кроз средњу вредност стандардне грешке MOS вредности и средње вредности корелације рангова (SROCC) и линеарне корелације (LCC) скорова за сваки пар посматрача. Средња вредност стандардне грешке говори о томе колико су се субјекти слагали по апсолутној вредности квалитета. Са друге стране средње вредности SROCC и LCC говоре о томе колико се разликују рангови и њихов релативни квалитет од субјекта до субјекта. За две изворне секвенце са малом динамиком ('20060708_01_140x440_10-20' и 'crowding_detection_1534-1544'), где је већина модификованих секвенци високог квалитета, скорови имају случајан карактер јер посматрачи нису у могућности да на прави начин раздвоје видео секвенце по питању квалитета (између енкодера односно битских брзина). За остале изворне секвенце слагање субјеката у евалуацији квалитета је веће од 80%.

Таб. 5.26 Сагласност субјеката у STrial 1 сету.

Референтна видео секвенца	Средња стандардна грешка	Средња вредност корелације рангова (SROCC) између субјеката	Средња вредност линеарне корелације (LCC) између субјеката
20060708_01_140x440_10-20.avi	0.16519	0.75276	0.78763
Balcony3_Vis_653-703.avi	0.1412	0.86344	0.86982
ME_Training_225-235.avi	0.12361	0.82099	0.86109
Parked_Vehicle_night_121-131.avi	0.14335	0.87079	0.87651
S1-T1-C-2_126-136.avi	0.13326	0.80756	0.81607
abandoned_luggage_36-46.avi	0.13081	0.88179	0.88516
crowding_detection_1534-1544.avi	0.15999	0.76899	0.81084
db3_2b_30fps_2-12.avi	0.13203	0.87374	0.86504
parked_vehicle_day_118-128.avi	0.14345	0.88588	0.88158

Основне статистике субјективних скорова појединачних посматрача приказане су на Сл. 5.56, где су дате средње вредности и стандардне девијације њихових скорова ($MOS \pm \sigma$). Уочава се да постоје варијације средњих вредности субјективних скорова индивидуалних посматрача. Иако варијације средњих вредности постоје, овај пут оне нису такве да би се неки од посматрача посебно издвојио од већине као што је био случај у STrial 0 сету.



Сл. 5.56 Средње вредности и стандардне девијације субјективних скорова ($MOS \pm \sigma$ опсер) за сваког од посматрача STrial 1 сета.

5.12.15. Резултати субјективних тестова – коментари

Иницијални резултати субјективних тестова спроведених на Универзитету у Манчестеру су дати у форми MOS, који одсликава квалитет видеа. Спроведени тестови дају могућност да се евалуирају карактеристике посматрача са становишта опсервације видеа. Валидација мера објективне процене квалитета може уз додатну анализу прикупљених резултата дати бољи увид у начин одлучивања и на уплив различитих утицаја на одлуке које посматрачи доносе по питању приказиваних видео записа.

Први значајан резултат је укључивање резултата субјективног STrial 0 теста за потребе даље анализе и валидације. Иницијални план за овај тест је био да то буде припрема за опширнији тест, STrial 1, из кога би се резултати користили у

валидацији мера објективне процене. Очекивало се да ће STrial 0 тест дати непоуздане резултате и естимације квалитета са малим поверењем и то из бројних разлога: недостатка искуства посматрача, почетних верзија софтвера, малог броја деградираних секвенци по изворној секвенци, услова у којима је спроведен тест (који нису под контролом) и слично.

Међутим, поузданост субјективних импресија квалитета STrial 0 теста је практично иста као и у STrial 1 тесту. У Таб. 5.27 дате су средње вредности стандардних девијација ($E[\sigma_i]$) и стандардних грешака ($E[SE_i]$) субјективних скорова добијених за комплетне сетове (за све видео секвенце из сетова). Мања стандардна грешка у STrial 1 сету се може објаснити тиме што је у експерименту учествовао већи број посматрача (два пута више, чиме је однос стандардних грешака $1/\sqrt{2}$). Мала разлика у средњој вредности стандардних девијација се може објаснити повећаним искуством у евалуацији видео квалитета другог сета. Наравно, дошло је до смањења средње вредности, што је и очекивано.

Таб. 5.27 Средње вредности стандардних девијација и стандардних грешака субјективних скорова на комплетним STrial сетовима.

Видео сет	$E[\sigma_i]$	$E[SE_i]$
STrial 0	0.820577	0.211872
STrial 1	0.787459	0.141432

Мале разлике у поузданости скорова спроведених тестова такође указују на релативно слаб ефекат спољашњих услова на вредности субјективних скорова. Први тест је спроведен у канцеларијским условима и амбијенталним светлом док је други тест спроведен у замраченим, скоро потпуно контролисаним условима осветљаја.

Уочен је још један ефекат током припреме објашњења самих тестова. Неки од посматрача су били забринути да можда неће испунити стандардне захтеве (очекивања) тражене у тестовима. Иако не постоји стандард и иако се индивидуалне перформансе не мере, посматрачи су некада забринути да ће направити грешку у новом окружењу. Због тога им се мора експлицитно нагласити да не постоје тачни

или погрешни одговори како би се умирили, повећали поузданост и смањили стрес под којим се налазе.

Последње запажање из спроведених тестова односи се на темпо. У оба сета се налазе секвенце са slabим и високом темпом (fps). Постоји разлика у апсолутном квалитету између секвенци са 2 fps и 30 fps. Ово се може објаснити са повећањем квалитета слике (просторних детаља у кадру) за секвенце са slabим темпом, јер се са смањењем броја кадрова може користити представа сваког од пиксела са већим бројем бита. Међутим, динамика се негативно одразила на перцепцију квалитета. Блиско са овим запажањем, потребно је правити компромис између очувања просторних информација и континуитета кретања на слици (смањењем темпа секвенце смањити битску брзину без великог губитка у квалитету).

5.13. Анализа резултата објективних мера процене на STrial сетовима

У овом делу дисертације дати су резултати субјективне валидације објективних мера процене кроз поређење објективних и субјективних скорова добијених на STrial сетовима видео секвенци. Спроведена је и валидација овде предложене VQ^{AB} мере за потребе због којих се и развијала – процена квалитета компримованих видео секвенци коришћењем различитих енкодера а за потребе надзора.

За валидацију, односно поређење субјективних и објективних скорова квалитета развијен је софтвер у програмском пакету Matlab[®]. Софтвер омогућава учитавање, поређење и квантитативну и квалитативну анализу скорова (субјективно-објективне мере корелације).

Објективна VQ^{AB} мера је разматрана кроз ширу анализу – анализу на нивоу приказаних сетова, анализу на нивоу референтних секвенци, анализу по начинима компресије, анализу по битским брзинама па до глобалне анализе, тј. анализе на нивоу комплетних сетова. Конвенционална PSNR мера, као и неколико тренутно

најзаступљених мера, SSIM, MAD и MOVIE мере, су анализирани на глобалном нивоу и њихове перформансе су упоређене са перформансама VQ^{AB} мере.

5.13.1. Мере и софтвер за евалуацију објективних мера процене

Субјективне импресије квалитета прикупљене током STrial тестова су преликани у MOS, тј. једну вредност која одсликава субјективни утисак квалитета сваке од деградираних секвенци. Са друге стране, објективне мере процене дају објективну, нумеричку вредност која одсликава објективну естимацију квалитета деградираних видеа добијену кроз поређење са изворном секвенцом. Најдиректнији начин одређивање њихове корелације је кроз обраду субјективних и објективних естимација квалитета на комплетном сету (на свим видео секвенцама из базе), тј. кроз одређивање корелације два независна узорка. На овај начин се добијају перформансе објективне мере на глобалном нивоу.

Додатно се могу дефинисати алтернативе овакве евалуације користећи одговарајуће подскупе субјективних и објективних естимација квалитета. Субјективни тестови су спроведени тако што су посматрачима приказиване групе деградираних видео секвенци са њиховим оригиналом. Групни приказ (приказ једног сета) потиче од исте референтне секвенце. Референтна секвенца је такође коришћена да би се направило више различитих сетова за приказ. Добро је познато да људи као посматрачи могу без много напора да одлуче о релативном квалитету сигнала са истим садржајем. Ово указује да ће посматрачи поуздано рангирати видео записе који потичу од исте референтне секвенце. Због тога су перформансе објективних мера процене квалитета анализирани и на нивоу исте референце (reference wise).

У евалуацији на нивоу референтних секвенци субјективни и објективни скорови квалитета су раздвојени на групе које потичу од истог референтног сигнала. Тако су подаци из STrial 0 сета подељени на осам група (због осам референтних секвенци) од по шест видео записа. Субјективно-објективна корелација се након тога одређивала за сваки сет након чега је одређена средња вредност над осам група као глобални показатељ слагања.

На најнижем нивоу, посматрачима су приказиване по две (STrial 0 сет) или три (STrial 1 сет) деградираних секвенци истовремено. Због тога је можда најпоузданија естимација релативног квалитета видеа једног групног приказа. Због тога се евалуација објективних мера може одредити и кроз предикцију релативног субјективног квалитета само за видео записе који припадају једном групном сету (set wise). Аналогно анализи на нивоу референтних секвенци, естимације квалитета су подељене на групе које одговарају приказима (24 групе од по 2 видеа из STrial 0 сета и 27 група од по 3 видеа из STrial 1 сета). Показатељи перформанси, корелација рангова (SROCC) и коефицијент линеарне корелације (LCC), су након тога одређени за сваку групу посебно. Средња вредност сваког од показатеља је након тога узета као индикатор перформансе на глобалном нивоу – нивоу сета.

Како би се евалуација спровела на нивоу групе, референце или глобално, формиране су текстуалне датотеке са потребним подацима. Део текстуалне датотеке која описује STrial 1 сет дат је на Сл. 5.57. Овај запис се аутоматски учитава како би се спровела потребна евалуација. Сваки ред у овом фајлу односи се на једну секвенцу са деградацијом и садржи више поља:

- **Display Set ID:** идентификациони број који представља редни број приказаног сета (у оквиру сета приказано је више секвенци које су посматрачи евалуирали).
- **Degradation Set ID:** идентификациони број референтне секвенце (редни број којим су секвенци које потичу од истог извора приказиване).
- **Име референтне секвенце**
- **Име деградираних секвенци**
- **MOS:** естимација субјективног квалитета деградираних секвенци
- **Стандардна грешка MOS**

5. Процена квалитета видео заснована на очувању информација о градијенту

1	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x140_h264_10kbps_10-20.avi	,	1.732258	,	0.207825
1	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_mpeg4_10kbps_10-20.avi	,	1.138710	,	0.159817
1	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_RV_10kbps_10-20.avi	,	1.603226	,	0.173669
2	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x140_h264_64kbps_10-20.avi	,	4.077419	,	0.115443
2	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_mpeg4_64kbps_10-20.avi	,	2.948387	,	0.169431
2	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_RV_64kbps_10-20.avi	,	2.364516	,	0.194635
3	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x140_h264_128kbps_10-20.avi	,	4.509677	,	0.100038
3	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_mpeg4_128kbps_10-20.avi	,	2.822581	,	0.186289
3	,	1	,	20060708_01_140x440_10-20.avi	,	20060708_01_140x440_RV_128kbps_10-20.avi	,	2.441935	,	0.179592
4	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_h264_10kbps_36-46.avi	,	1.667742	,	0.140306
4	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_mpeg4_10kbps_36-46.avi	,	1.041935	,	0.136372
4	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_RV_10kbps_36-46.avi	,	1.996774	,	0.154814
5	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_h264_64kbps_36-46.avi	,	4.180645	,	0.129218
5	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_mpeg4_64kbps_36-46.avi	,	3.687097	,	0.127570
5	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_RV_64kbps_36-46.avi	,	2.816129	,	0.144810
6	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_h264_128kbps_36-46.avi	,	4.351613	,	0.113580
6	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_mpeg4_128kbps_36-46.avi	,	4.374194	,	0.102859
6	,	6	,	abandoned_luggage_36-46.avi	,	abandoned_luggage_RV_128kbps_36-46.avi	,	3.164516	,	0.127722

Сл. 5.57 Део листинга субјективних скорова STrial 1 сета.

Можда најбитнији аспект перформанси објективне мере јесте њена способност поуздане евалуације и рангирања различитих система за видео пренос/компресију (у контексту евалуације која ће се овде спровести то би била анализа према различитим кодецима или комбинацијама кодека и битских брзина), а што се може постићи и кроз скупе субјективне тестове. Субјективна евалуација различитих сетова (кодека и битских брзина) је доступна из субјективних тестова и може се директно поредити са објективним скоровима за специфичне системе преноса. Методолошки, ово укључује издвајање скорова квалитета деградираних видео секвенци добијених за специфичне системе преноса и њихово збрајање у један скор који одсликава тај систем преноса (рецимо $MOS_{H.264}$). Ово је урађено за субјективне и објективне скорове за све системе преноса. Субјективно-објективна корелација резултујућих скорова квалитета је након тога мерена кришћењем стандардних мера процене перформанси.

Од објективне мере се очекује да поуздано процени субјективни квалитет и ранг различитих система преноса на исти начин како би то урадили посматрачи.

5.13.2. Мере за процену перформанси

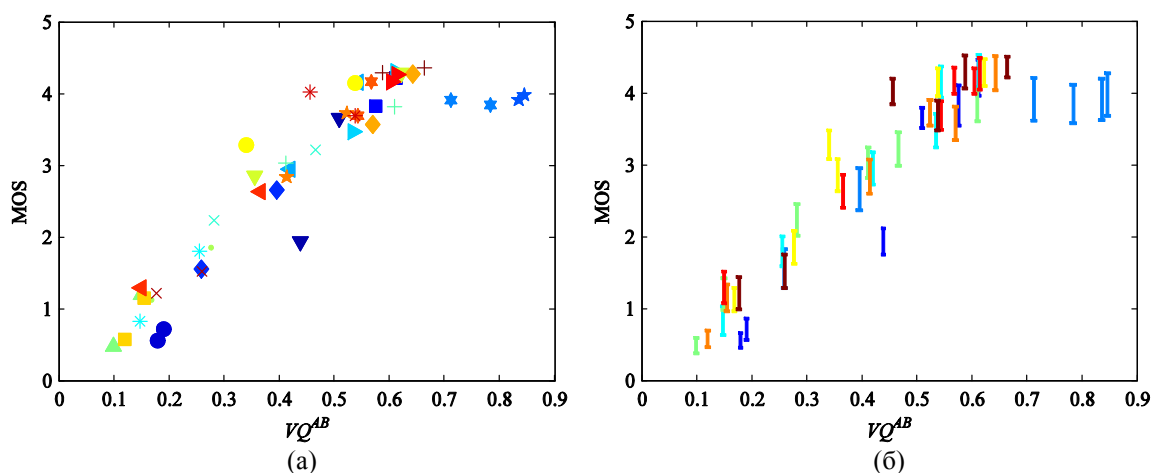
Софтвер за анализу перформанси објективних мера омогућава анализу на два начина. Први начин обухвата одређивање линеарне корелације и корелације рангова

подскупова субјективних и објективних скорова (према сетовима, референцама, битским брзинама, кодецима и њиховим комбинацијама) и глобално. Код ове анализе нису се користиле логистик функције мапирања како би се линеаризовале релације субјективних и објективних скорова квалитета, а које утичу на вредности корелације. У неким ситуацијама то би било и несврхисходно јер се у подскуповима може наћи мали узорак. Други начин обухвата одређивање вредности показатеља предложених од стране ITU [5] (тачност, монотоност и конзистентност предикције), али на глобалном нивоу – нивоу STrial сетова. У овом случају се користе логистик криве са четири параметра.

Показатељи перформанси предложене VQ^{AB} мере су одређени на свим наведеним нивоима, док су перформансе осталих објективних мера дате на глобалном нивоу.

5.13.3. Анализа перформанси објективних мера процене на STrial 0 сету

На Сл. 5.58 приказани су дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета и то према приказаним сетовима и референтним секвенцама. Скорови који припадају истом скупу (сету или потичу од исте референтне секвенце) су приказани различито обојеним симболима. Додатно су на једном од дијаграма расипања дати и интервали поузданости.



Сл. 5.58 Дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета: (а) према приказаним сетовима и (б) према референтним секвенцама.

На оба дијаграма се јасно уочава добро слагање субјективних и објективних скорова квалитета при чему постоји скоро линеарна веза између MOS и VQ^{AB} скорова. Већина естимација квалитета се налази дуж дијагоналне линије, иако постоји неколико изузетака од којих је већина са десне стране главне корелационе линије. Ово указује да су посматрачи овим секвенцама дали ниже вредности квалитета у односу на објективне скорове, или постоји негативни тренд у објективној естимацији квалитета. Такође, може се уочити и регион сатурације у опсегу високих вредности VQ^{AB} мере, где неколико видеа има скоро исте MOS вредности али различите VQ^{AB} вредности.

У Таб. 5.28 сумиране су перформансе VQ^{AB} мере на STrial 0 субјективним подацима. Може се говорити о добром учинку објективне мере, где су њене перформансе у нивоу перформанси UoM сета, на коме је мера оптимизована. Средња вредност корелације рангова по референтним секвенцама је изнад 0.96, док је корелација рангова на комплетном скупу од 48 расположивих секвенци изнад 0.91. Треба напоменути да је по питању рангирања у оквиру приказаних сетова, у 22 од 24 сета извршено рангирање које је у складу са субјективним рангом, односно у две ситуације је дошло до инверзије рангирања ($SROCC=-1$) што је значајно умањило средњу вредност по приказаним сетовима на 0.83.

Таб. 5.28 Квантитативни показатељи перформанси VQ^{AB} мере на STrial 0 сету.

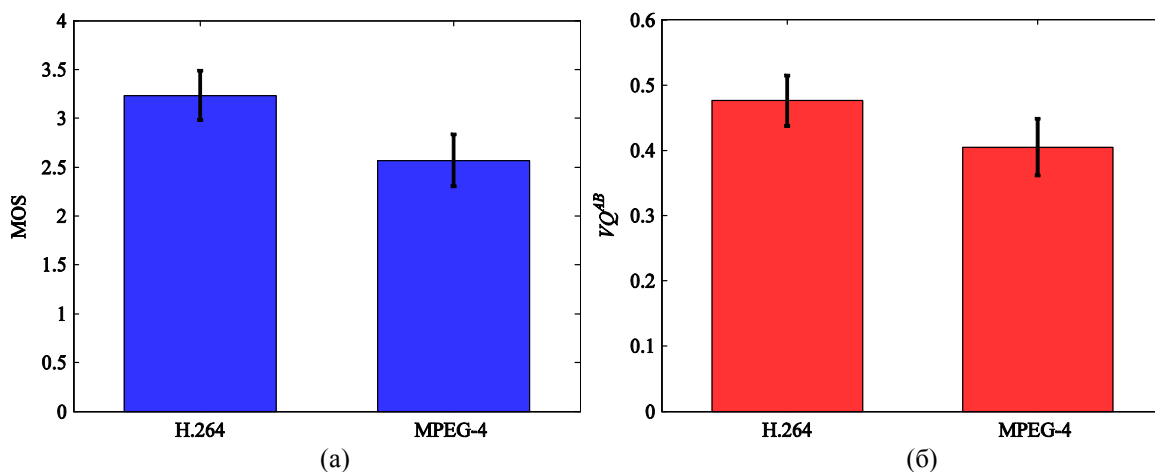
Мера	SROCC	LCC
Средња вредност мере по сетовима	0.83333	0.83333
Средња вредност мере по референцама	0.96429	0.9758
Комплетан сет	0.91413	0.91625

Перформансе објективне мере према референтним секвенцама су дате у Таб. 5.29. Овакав начин евалуације узима у обзир људске способности да се на лак начин рангирају секвенце које потичу од истог референтног сигнала. Види се да објективна VQ^{AB} мера има одличне перформансе рангирања, где се рангирање у потпуности подудара са субјективним ранговима за три од осам деградационих сетова. Коефицијент линеарне корелације субјективних и објективних скорова квалитета је изнад 0.95, за сваки референти видео.

Таб. 5.29 Квантитативни показатељи перформанси према референтним секвенцама.

Изворни видео	SROCC	LCC
abandoned_luggage_12-22.avi	1	0.97242
crowding_detection_1230-1240.avi	0.94286	0.96822
db3_2b_30fps_10-20.avi	1	0.989
ME_Training_618-628.avi	1	0.98163
20060708_01_140x440_10-20.avi	0.94286	0.95092
parked_vehicle_day_133-143.avi	0.94286	0.99222
Parked_Vehicle_night_102-112.avi	0.94286	0.99583
S1-T1-C-2_2-12.avi	0.94286	0.95617

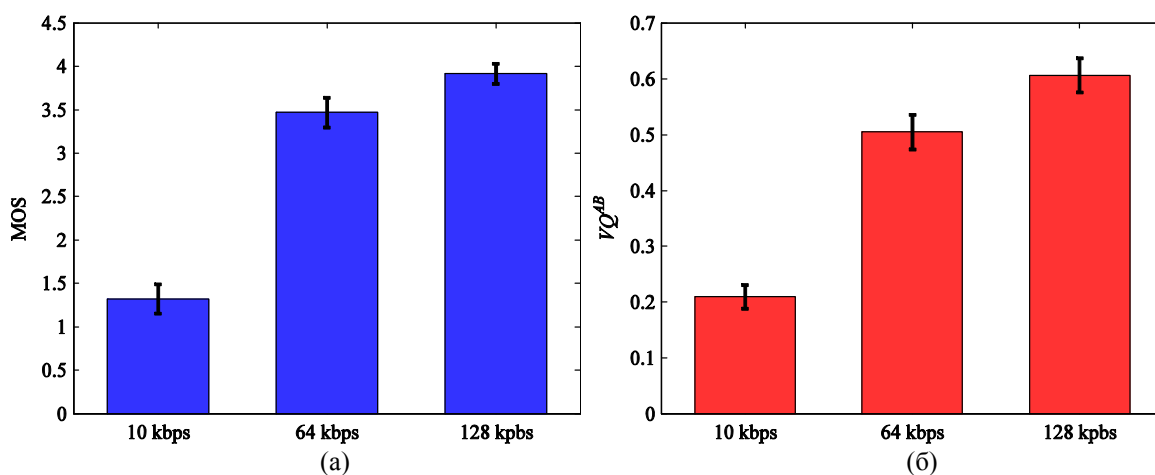
Објективна VQ^{AB} мера је искоришћена за евалуацију видео квалитета различитих система преноса. Естимације квалитета за два кодека и три битске брзине су издвојене и усредњене како би се добила једна субјективна и једна објективна вредност за одређену битску брзину, кодек или и конкретан кодек и битску брзину. На Сл. 5.59(а) и (б) приказане су субјективне и објективне естимације квалитета за H.264 и MPEG-4 системе преноса евалуиране током STrial 0 теста. Са вертикалних приказа се уочава да објективна мера даје добру апсолутну и релативну предикцију квалитета оба кодека.



Сл. 5.59 (а) субјективни и (б) објективни видео квалитет добијен за различите кодеке STrial 0 сета.

На Сл. 5.60 приказане су субјективне и објективне естимације квалитета секвенци са различитим битским брзинама STrial 0 сета – 10, 64 и 128 kbps. Поново

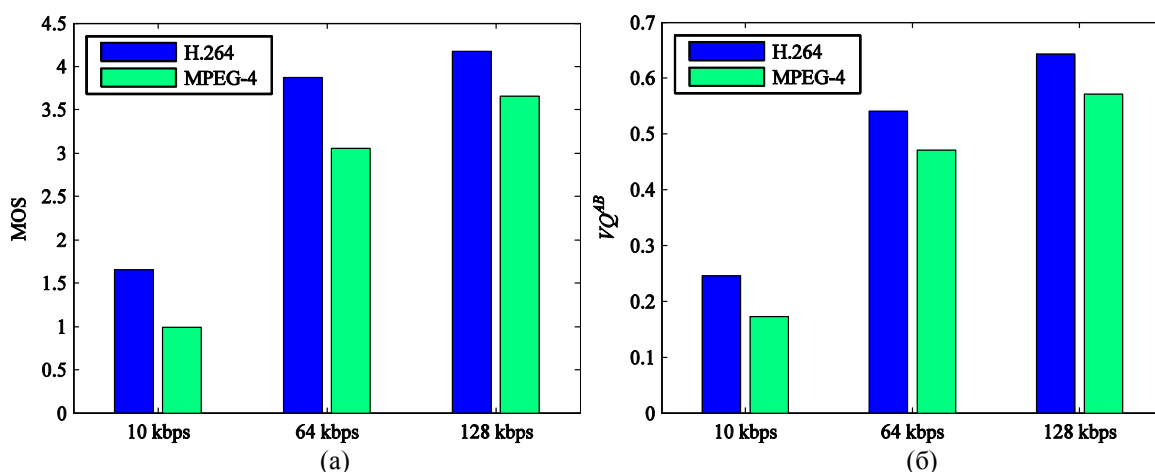
се јасно уочава да објективна мера има добру предикцију како апсолутног тако и релативног субјективног квалитета за све три битске брзине. Објективна мера такође правилно евалуира субјективну разлику која одговара анализираним битским брзинама – мања је разлика квалитета између 64 и 128 kbps од разлике квалитета која постоји између 10 и 64 kbps система преноса. Коефицијент линеарне корелације између субјективних и објективних естимација квалитета секвенци са различитим битским брзинама је 0.996 (наравно корелација рангова је максимална).



Сл. 5.60 (а) субјективни и (б) објективни квалитет видео секвенци са различитим битским брзинама STrial 0 сета.

Осим одвојених анализа по кодецима и битским брзинама, понашање система преноса се може анализирати и по њиховим комбинацијама (кодек + битска брзина). На Сл. 5.61 дате су субјективне и објективне вредности естимација квалитета секвенци из различитих система преноса (различитих кодера) а по конкретним битским брзинама у оквиру STrial 0 сета. Може се уочити да су посматрачи за све битске брзине дали предност систему преноса са H.264 кодовањем, а што је испраћено и по објективним проценама. Коефицијент линеарне корелације процењених субјективних и објективних вредности је у овој анализи већи од 0.99, док је корелација рангова око 0.95 (јер је $MOS_{H.264+64kbps} > MOS_{MPEG4+128kbps}$ а $VQ_{H.264+64kbps} < VQ_{MPEG4+128kbps}$).

5. Процена квалитета видео заснована на очувању информација о градијенту

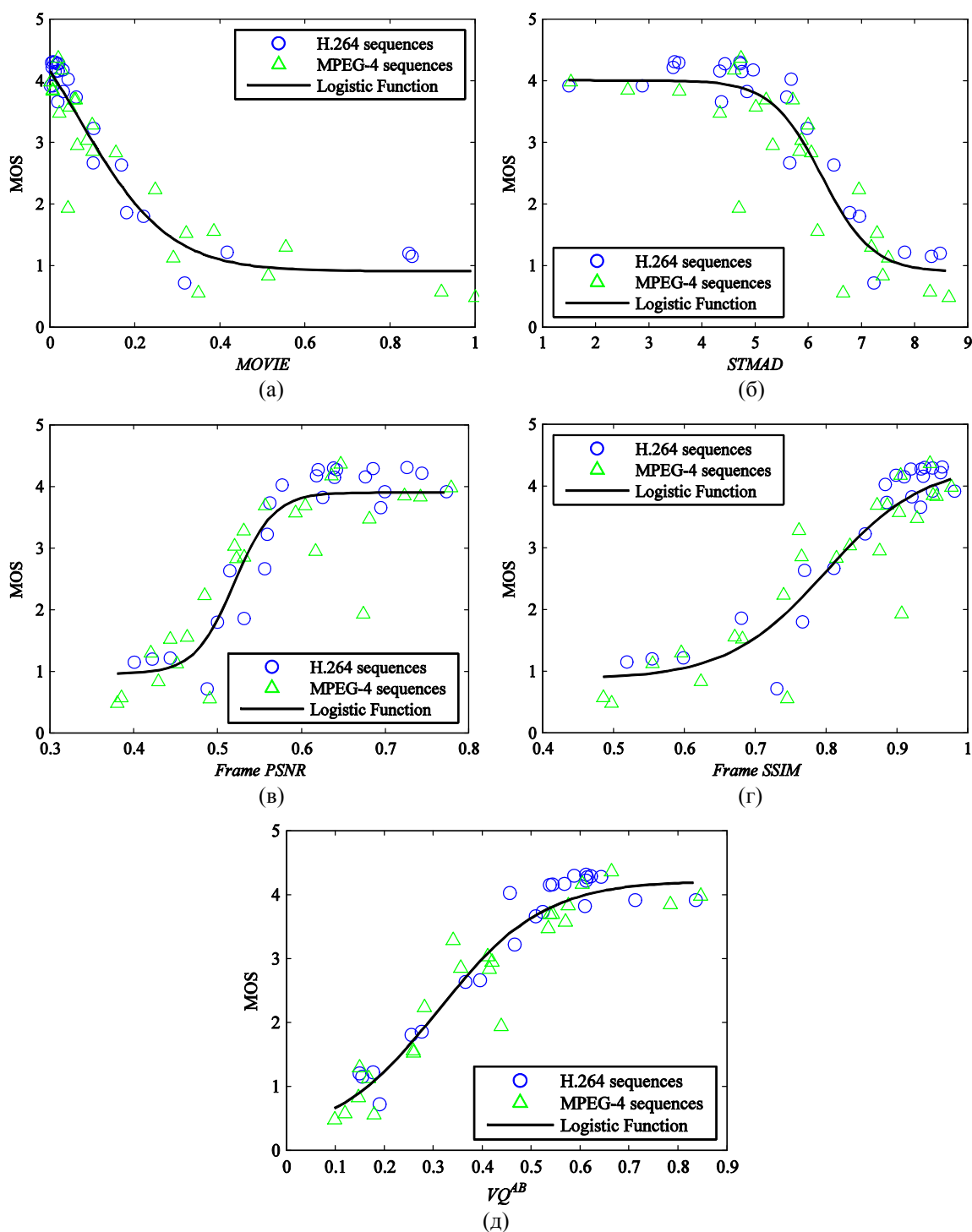


Сл. 5.61 (а) субјективни и (б) објективни видео квалитет кодовања са различитим енкодерима при различитим битским брзинама STrial 0 сета.

Перформансе предложене мере и неколико објективних мера анализиране су на глобалном нивоу (нивоу комплетног сета STrial 0) по мерама предложеним од стране VQEG групе, тј. након интерполације помоћу нелинеарних логистик функција (логистик криве са четири параметра). Перформансе објективних мера процене су дате у Таб. 5.30. Уочава се да је по свим параметрима предикција субјективних скорова најбоља уколико се користи предложена VQ^{AB} објективна мера. На другом месту по слагању субјективних и објективних скорова је MOVIE мера. Додатно су на Сл. 5.62 приказани дијаграми расипања субјективних и објективних скорова на STrial 0 сету, заједно са кривама интерполације.

Таб. 5.30 Перформансе објективних мера на STrial 0 сету.

	LCC	SROCC	MAE	RMSE	OR
	MAD				
<i>Spatial MAD</i>	0.9273	0.9048	0.3659	0.4782	31.2500
<i>Temporal MAD</i>	0.9014	0.8069	0.4109	0.5531	39.5833
<i>Spatio-temporal MAD</i>	0.9190	0.8451	0.3709	0.5037	33.3333
	MOVIE				
<i>Spatial MOVIE</i>	0.9374	0.8869	0.3386	0.4450	27.0833
<i>Temporal MOVIE</i>	0.9435	0.8866	0.3226	0.4235	25
<i>MOVIE</i>	0.9503	0.8998	0.3101	0.3978	18.7500
	Конвенционалне мере – PSNR и SSIM				
<i>Frame PSNR</i>	0.9178	0.8353	0.3758	0.5074	27.0833
<i>Frame SSIM</i>	0.9217	0.8775	0.3539	0.4955	29.1667
	Предложена VQ мера				
VQ^{AB}	0.9631	0.9143	0.2652	0.3436	10.4167

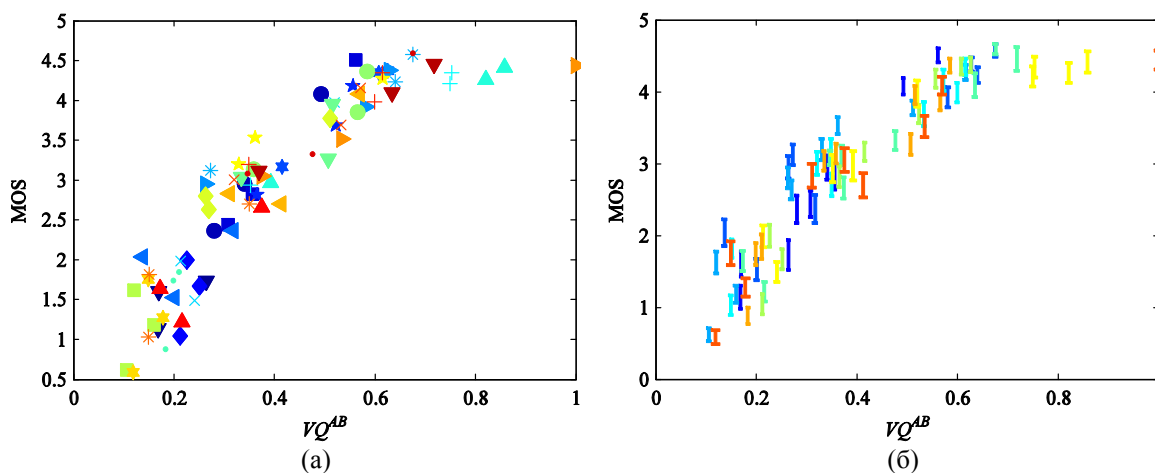


Сл. 5.62 Дијаграми расипања субјективних и објективних скорова на STrial 0 сету.

5.13.4.Анализа перформанси објективних мера процене на STrial 1 сету

На Сл. 5.63 приказани су дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета и то према приказаним сетовима и референтним

секвенцама. Скорови који припадају истом скупу (сету или потичу од исте референтне секвенце) су приказани различито обојеним симболима. Додатно су на једном од дијаграма расипања дати и интервали поузданости.



Сл. 5.63 Дијаграми расипања субјективних и објективних VQ^{AB} скорова квалитета: (а) према приказаним сетовима и (б) према референтним секвенцама.

На оба дијаграма се јасно уочава да је корелација субјективних и објективних скорова веома висока. Њихова веза је скоро линеарна, са малим регионом сатурације на горњој граници објективних скорова.

У Таб. 5.31 сумиране су перформансе VQ^{AB} мере на STrial 1 субјективним подацима. Може се говорити о добром учинку објективне мере, где су њене перформансе у нивоу перформанси UoM сета, на коме је мера оптимизирана, односно у нивоу перформанси STrial 0 сета. Средња вредност корелације рангова по референтним секвенцама је изнад 0.95, док је корелација рангова на комплетном скупу од 81 расположиве секвенце изнад 0.94. Одговарајуће линеарне корелације су нешто ниже због региона сатурације видљивог на претходним дијаграмима. Од расположивих 27 сетова са по три секвенце, рангирање субјективних и објективних скорова је исто у 19 приказа (корелација рангова је нешто већа од корелације рангова на STrial 0 сету).

Таб. 5.31 Квантитативни показатељи перформанси VQ^{AB} мере на STrial 1 сету.

Мера	SROCC	LCC
Средња вредност мере по сетовима	0.85185	0.87575
Средња вредност мере по референцама	0.9537	0.94619
Комплетан сет	0.94329	0.91007

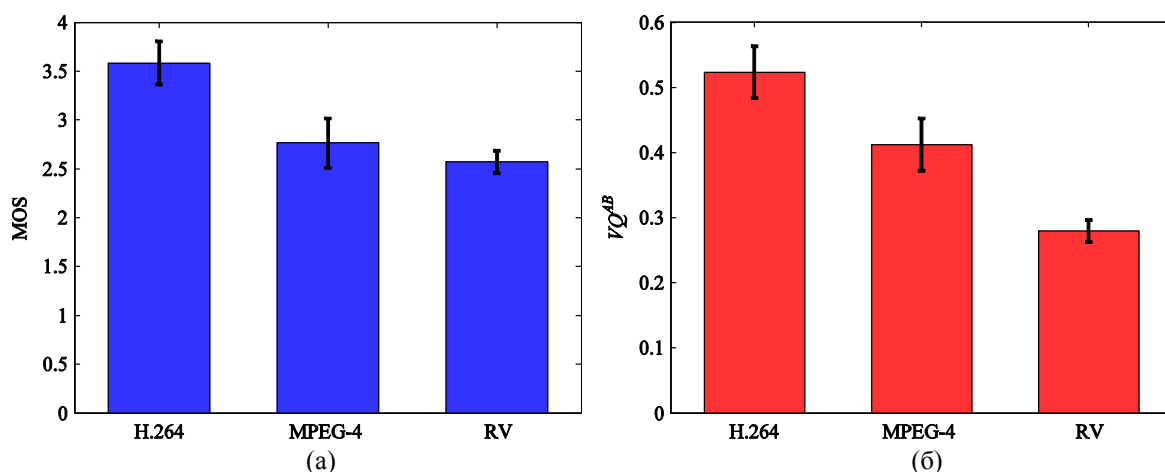
Перформансе објективне мере према референтним секвенцама су дате у Таб. 5.32. Овакав начин евалуације узима у обзир људске способности да се на лак начин рангирају секвенце које потичу од истог референтног сигнала. Види се да објективна VQ^{AB} мера има добре перформансе рангирања, где је линеарна корелација скорова код осам изворних секвенци већа од 0.92. Коефицијент корелације рангова субјективних и објективних скорова квалитета је изнад 0.91, за сваки референтни видео.

Таб. 5.32 Квантитативни показатељи перформанси према референтним секвенцама.

Изворни видео	SROCC	LCC
20060708_01_140x440_10-20.avi	0.98333	0.98383
Balcony3_Vis_653-703.avi	0.91667	0.93641
ME_Training_225-235.avi	0.96667	0.92897
Parked_Vehicle_night_121-131.avi	0.93333	0.9581
S1-T1-C-2_126-136.avi	0.91667	0.95823
abandoned_luggage_36-46.avi	0.98333	0.97906
crowding_detection_1534-1544.avi	0.96667	0.93616
db3_2b_30fps_2-12.avi	0.98333	0.94421
parked_vehicle_day_118-128.avi	0.93333	0.89075

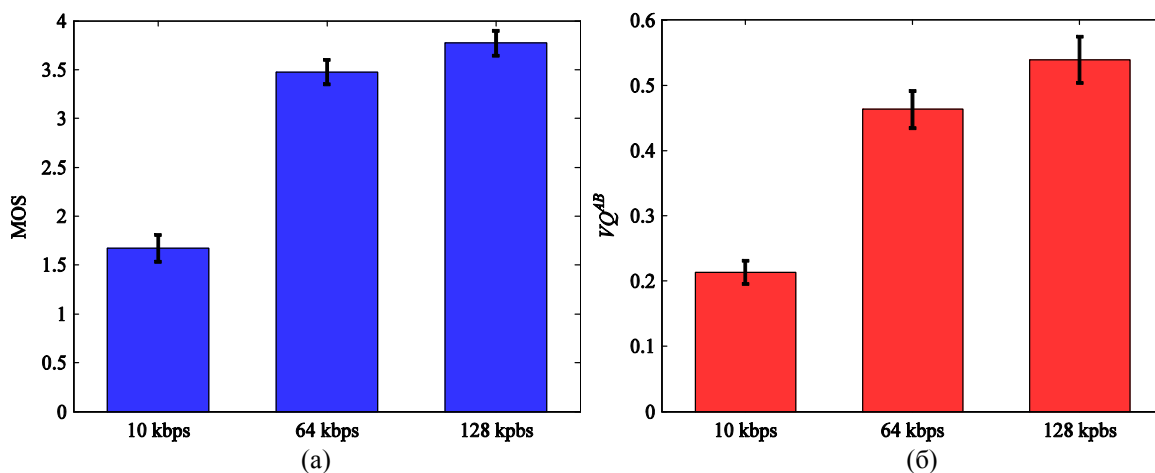
Објективна VQ^{AB} мера је искоришћена и за евалуацију видео квалитета различитих система преноса. Овај пут естимације квалитета за три кодека и три битске брзине су издвојене и усредњене како би се добила једна субјективна и једна објективна вредност за одређену битску брзину, кодек или и конкретан кодек и битску брзину. На Сл. 5.64 приказане су субјективне и објективне естимације квалитета за H.264, MPEG-4 и RV кодеке (системе преноса) евалуиране током STrial 1 теста. Са вертикалних приказа се уочава да је објективно рангирање система преноса идентично субјективном рангирању. Такође, уочава се да се релативни субјективни однос естимација квалитета MPEG-4 и RV система преноса разликује од одговарајућег објективног односа (због чега је коефицијент линеарне корелације добијених средњих вредности субјективних и објективних скорова нешто изнад 0.92).

На Сл. 5.65 приказане су субјективне и објективне естимације квалитета секвенци са различитим битским брзинама STrial 1 сета – 10, 64 и 128 kbps.



Сл. 5.64 (а) субјективни и (б) објективни видео квалитет добијен за различите кодеке STrial 1 сета.

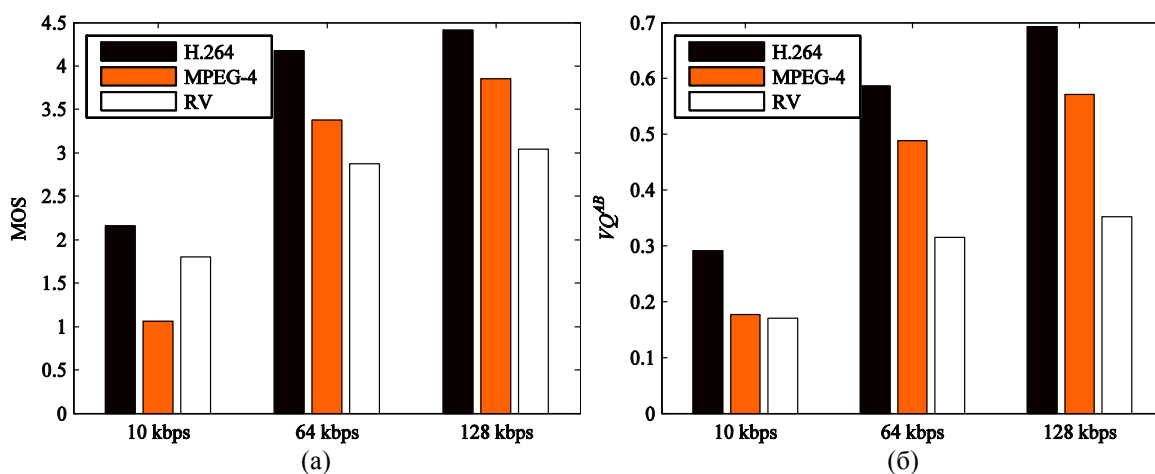
Поново се јасно уочава да објективна мера има добру предикцију како апсолутног тако и релативног субјективног квалитета за све три битске брзине. Објективна мера такође правилно евалуира субјективну разлику која одговара анализираним битским брзинама – мања је разлика квалитета између 64 и 128 kbps од разлике квалитета која постоји између 10 и 64 kbps система преноса. Коефицијент линеарне корелације између субјективних и објективних естимација квалитета секвенци са различитим битским брзинама је 0.996 док је корелација рангова максимална.



Сл. 5.65 (а) субјективни и (б) објективни квалитет видео секвенци са различитим битским брзинама STrial 1 сета.

Осим одвојених анализа по кодецима и битским брзинама, понашање система је анализирано и по њиховим комбинацијама (кодек + битска брзина). На Сл. 5.66 дате су субјективне и објективне вредности естимација квалитета секвенци из различитих система преноса (различитих кодера) а по конкретним битским брзинама у оквиру STrial 1 сета. Може се уочити да су посматрачи за све битске брзине дали предност систему преноса са H.264 кодовањем, а што је испраћено и по објективним проценама. Коефицијент линеарне корелације естимираних субјективних и објективних вредности у овој анализи је 0.96, док је корелација рангова већа од 0.98 (јер је $MOS_{RV+10kbps} > MOS_{MPEG4+10kbps}$ а $VQ_{RV+10kbps} < VQ_{MPEG4+10kbps}$).

Субјективни и објективни рангови су различити само у једном случају – за битску брзину од 10 kbps. На овој битској брзини посматрачи су дали предност квалитету RV система преноса над MPEG-4 системом преноса, док је по VQ^{AB} мери блага предност на страни MPEG-4 система преноса. Разлика у рангирању је највероватније последица алгорита достизања циљне битске брзине од 10 kbps. Наиме, код MPEG-4 компресије овако мала битска брзина је у неким ситуацијама достигнута смањењем резолуције изворних секвенци. Након тога је рађена интерполација на циљану резолуцију од 360x288 пиксела. Смањењем резолуције неповратно је изгубљен део изворних информација, а што су посматрачи уочили и дали предност RV систему преноса. Поред тога, у MPEG-4 секвенцама има више локалних ефеката насталих блоковском компресијом, за разлику од RV секвенци које су равномерно деградиране.



Сл. 5.66 (а) субјективни и (б) објективни квалитет видео кодовања са различитим енкодерима при различитим битским брзинама STrial 1 сета.

Перформансе предложене мере и неколико објективних мера анализирани су на глобалном нивоу (нивоу комплетног сета STrial 1), и по мерама предложеним од стране VQEG групе, тј. након интерполације помоћу нелинеарних логистик функција (логистик криве са четири параметра). Перформансе објективних мера процене су дате у Таб. 5.33.

Из Таб. 5.33 уочава се да је по свим показатељима најбоље слагање субјективних и објективних VQ^{AB} скорова. На другом месту по слагању би били скорови добијени MOVIE техником, при чему је сада већа разлика у перформансама VQ^{AB} и MOVIE техника него на STrial 0 сету. Перформансе MAD и Frame SSIM техника су у нивоу перформанси добијених на STrial 0 сету, а перформансе Frame PSNR технике на овом сету са три кодека су значајно лошије од перформанси на сету са два кодека. Додатно су на Сл. 5.67 приказани дијаграми расипања субјективних и објективних скорова на STrial 1 сету, заједно са кривама интерполације.

Таб. 5.33 Перформансе објективних мера на STrial 1 сету.

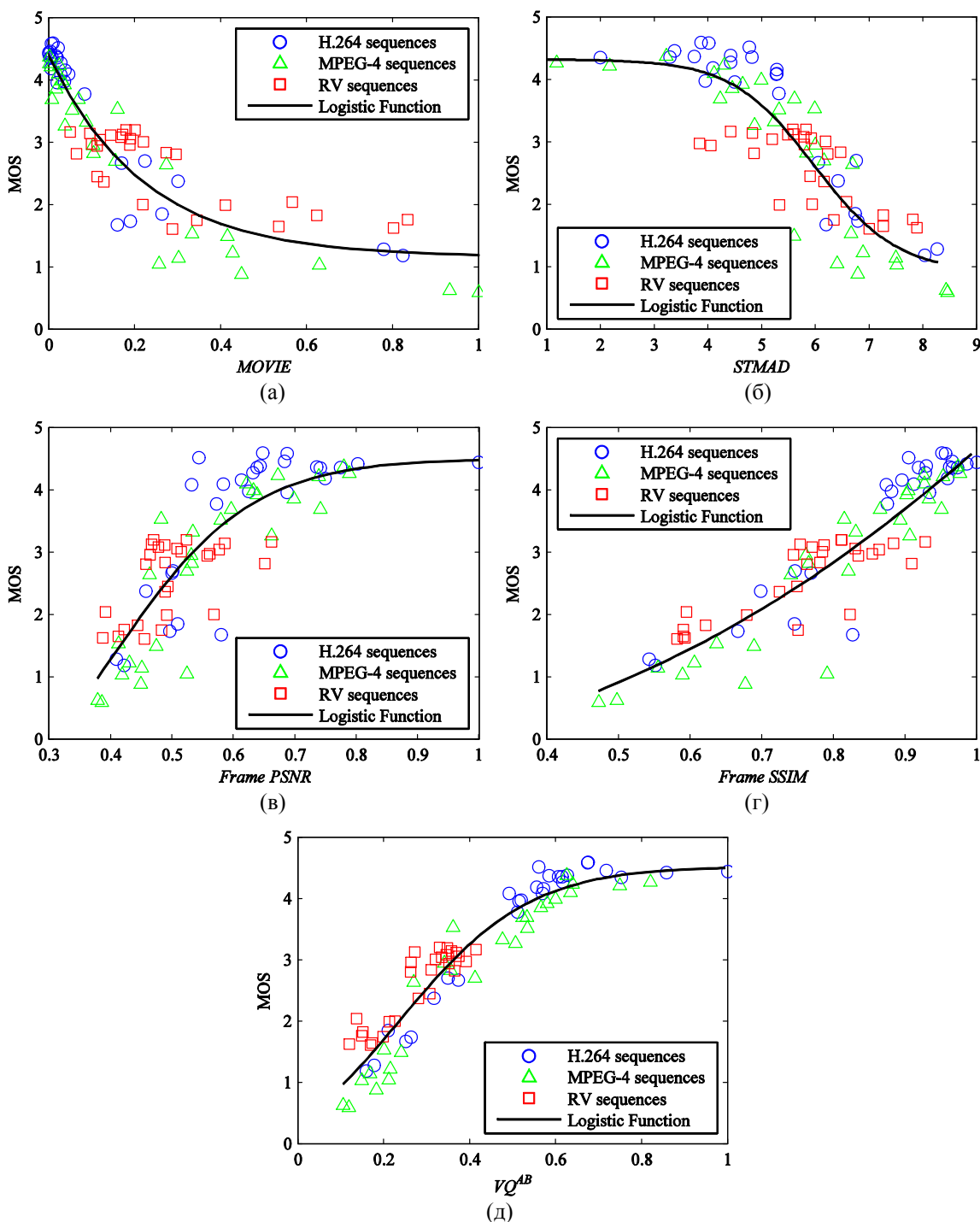
	LCC	SROCC	MAE	RMSE	OR
	MAD				
<i>Spatial MAD</i>	0.8999	0.9037	0.3713	0.4955	53.0864
<i>Temporal MAD</i>	0.8958	0.8943	0.3969	0.5048	54.3210
<i>Spatio-temporal MAD</i>	0.8867	0.8859	0.4002	0.5252	54.3210
	MOVIE				
<i>Spatial MOVIE</i>	0.8797	0.8767	0.4342	0.5402	59.2593
<i>Temporal MOVIE</i>	0.9218	0.9100	0.3443	0.4404	45.6790
<i>MOVIE</i>	0.9230	0.9212	0.3474	0.4371	54.3210
	Конвенционалне мере – PSNR и SSIM				
<i>Frame PSNR</i>	0.8410	0.8361	0.4712	0.6145	60.4938
<i>Frame SSIM</i>	0.9153	0.9122	0.3437	0.4575	43.2099
	Предложена VQ мера				
VQ^{AB}	0.9535	0.9432	0.2759	0.3423	35.8025

5.14. Закључак

У овом делу дисертације предложена је објективна мера VQ^{AB} за процену квалитета видео секвенци са потпуним референцирањем. Мера је настала проширењем мере процене квалитета појединачних слика. Наиме, поред анализе очувања просторних информација на нивоу кадрова (спроведене по узору на меру коришћену у процени квалитета слика), анализирано је очување информација о кретању и очување информација о боји. Квалитет видео секвенце добија се

5. Процена квалитета видеа заснована на очувању информација о градијенту

узимањем у обзир очувања информација по сва три канала – просторном, временском и каналу боје. Највећи значај у крајњој мери квалитета има очување просторних информација (80%), а најмањи значај очување информација о боји (5%). Оваква расподела значаја добијена је кроз поступак оптимизације на једној од доступних база (UoM база секвенци).



Сл. 5.67 Дијаграми расипања субјективних и објективних скорова на STrial 1 сету.

Валидација предложене мере је спроведена кроз поређење VQ^{AB} и субјективних скорова квалитета на неколико доступних база видео секвенци. Иако је објективна мера развијана за потребе процене квалитета секвенци са малим битским брзинама (до 576 kbps), перформансе мере су представљене и за базе са значајно већим битским протоком (реда десетина Mbps). Од анализираних објективних мера процене квалитета, предложена мера се показала најбољом на базама секвенци са малим битским протоком (UoM, FUB, EPFL-PoliMI и STrial). Степен слагања предложене мере са субјективним импресијама квалитета (мерен кроз коефицијент линеарне корелације) на овим базама секвенци се значајно разликује. Најнижи степен слагања је на EPFL-PoliMI бази секвенци – око 80%, док је на осталим базама секвенци већи од 90%.

На новоформираним базама секвенци Универзитета у Манчестеру (STrial 0 и STrial 1), формираним за битски проток до 128 kbps, објективна VQ^{AB} мера достигла је висок степен корелације са субјективним проценама квалитета. Корелација рангова (мерена Спирмановим коефицијентом корелације) ове мере на комплетном STrial 0 сету је изнад 0.91, док је средња вредност корелација по оригиналним секвенцама изнад 0.96 (од максимума који је 1). Коефицијент линеарне корелације је такође висок, изнад 0.91 на глобалном плану и гледано према оригиналним секвенцама изнад 0.97. Рангирање различитих система преноса по овој мери се у потпуности слаже са ранговима добијеним кроз субјективне процене.

Слични резултати су добијени на STrial 1 сету. Могућност исправног рангирања видеа према перцептуалном квалитету (мерено кроз Спирманову корелацију рангова) за ову меру на комплетном STrial 1 сету је изнад 0.94 док је одговарајућа вредност, гледано по изворним секвенцама, изнад 0.95. Коефицијент линеарне корелације субјективних и објективних скорова је такође висок, тј. преко 0.94 и 0.91 тим редом (од максимума који је једнак 1). По питању евалуације различитих система преноса према битским брзинама односно начинима кодовања, објективни VQ^{AB} и субјективни скорови се у потпуности слажу.

Перформансе мере на STrial сетовима су сразмерне са субјективно-објективном корелацијом постигнутом на UoM сету, на којем је мера и

оптимизована. Тиме су потврђене могућности VQ^{AB} мере у поузданој евалуацији субјективног квалитета видеа са малим битским протоком.

Валидација предложене мере на VQEG и LIVE Wireless базама секвенци са већим битским протоком показала је да су перформансе предложене мере у рангу најбољих мера објективне процене квалитета. Перформансе предложене мере су нешто ниже од најбољих мера процене квалитета на LIVE Video бази секвенци. Корелација са субјективним импресијама квалитета предложене мере на глобалном плану је 0.63, док је корелација најбоље, MOVIE мере, 0.73. Интересантно је да су на нивоу две класе секвенци са компресијом из LIVE Video базе (H.264 и MPEG-2) перформансе предложене мере биле боље од осталих анализираних објективних мера процене квалитета.

Стандардна имплементација VQ^{AB} мере користи сигмоидне функције за моделовање нелинеарног одзива визуелног система човека. За моделовање се користи осам параметара нелинеарних сигмоидних функција. Поред овог, нелинеарног модела, на неколико база секвенци испитан је и модел који не користи сигмоидне функције, тј. посматра се сличност референтног и тест сигнала кроз очување информација о структури, о кретању и боји. На VQEG базама секвенци бољим се показао нелинеарни модел (линеарна корелација је 0.80 насупрот 0.71 и 0.78 насупрот 0.70), док се на LIVE Video, LIVE Wireless и EPFL PoliMi базама бољим показао линеарни модел. Нарочито је значајан допринос линеарног модела на LIVE Video бази секвенци где се његовим коришћењем по перформансама на глобалном плану приближава перформансама најбољој MOVIE мери (коэффициент линеарне корелације субјективних и објективних скорова је порастао са 0.64 на 0.70). Разлика у перформансама линеарног и нелинеарног модела на UoM бази секвенци није значајна. Имајући у виду да је VQEG база застарела и да су бројни истраживачи нагласили да секвенце из ове базе нису перцептуално сепарабилне, тренутни резултати указују да је предност на страни линеарног модела очувања информација, а који је рачунарски једноставнији.

У оквиру дисертације анализиран је и значај квалитета кадрова на крајњу меру квалитета и слагање са субјективним импресијама квалитета. Показано је да се поуздана процена квалитета може добити коришћењем дела кадрова са најнижим

скоровима квалитета. Предложено је да се у одређивању крајње вредности квалитета видео секвенце узме у обзир 20% кадрова са најнижим вредностима квалитета. Ова вредност представља добар компромис између ситуација када у секвенцама постоје краткотрајне уочљиве сметње настале пакетским губицима и када се поуздана процена квалитета добија коришћењем веома малог броја кадрова са најнижим скоровима квалитета (рецимо 2%), и ситуација у којима су деградације присутне у свим кадровима и при којима је за поуздану процену потребно узети у обзир већи број кадрова (више од 20%, а у неким ситуацијама и свих 100%).

Оптималну вредност параметра p у реалним апликацијама треба одредити уважавајући контекст у коме се процена спроводи као и одговор који се очекује од посматрача. Посматрачима који имају за циљ да евалуирају квалитет секвенце или да опишу дешавања на сцени потребан је већи број кадрова за поуздану одлуку, као што је случај у UoM и VQEG сетовима. Насупрот њима, посматрачи који имају за циљ да процене степен деградације (LIVE Wireless и EPFL-PoliMI сетови), то могу учинити на основу релативно малог броја кадрова комплетне секвенце.

На три базе секвенци са малим битским протоком (FUB, STrial 0 и STrial 1) уочено је да постоји регион сатурације на врху опсега квалитета VQ^{AB} мере – за секвенце "доброг" квалитета. Како би се достигле циљне битске брзине које не прелазе 128 kbps, секвенце из ових база су са редукованом резолуцијом (176x144 у оквиру FUB базе и 360x288 у оквиру STrial сетова). Даљина посматрања, смањена резолуција и истовремени приказ више секвенци на монитору ограничили су могућности уочавања разлика код секвенци са битским протоком од 128 kbps, што је довело до засићења у субјективним импресијама квалитета. Истовремено, субјективно неуочљиве разлике између тест и изворних секвенци довеле су до размицања објективних скорова квалитета и до појаве сатурације. Ова појава би се могла елиминисати кроз предобраду видео сигнала ниског квалитета, тј. кроз филтрацију изворне и тест секвенце коришћењем филтра пропусника ниских учестаности.

Коначно, приказани резултати указују да се поуздана процена квалитета видео може добити кроз очување информација о градијенту.

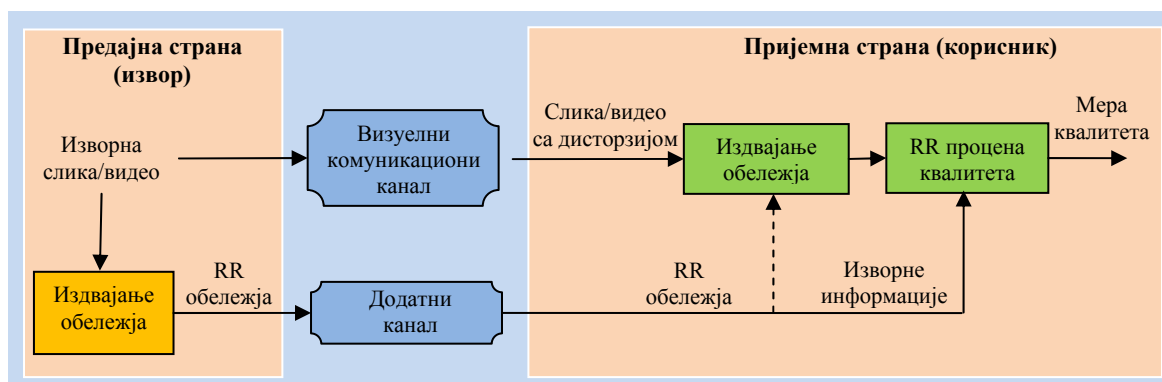
6. Процена квалитета видеа у условима редукованог референцирања

6.1. Увод

Технике редукованог референцирања се налазе између техника са потпуним референцирањем и техника без референцирања и дизајниране су да дају практична решења у процени квалитета док се задржава тачност процене. За разлику од система са потпуним референцирањем где је њихова практична корист ограничена и техника без референцирања где је њихова поузданост дискутабилна, технике са делимичним референцирањем представљају компромисно решење, тј. праве баланс између тачности процене квалитета и потребне количине референтних података [6]. Иако је концепт процене квалитета са делимичним референцирањем осмишљен деведесетих година прошлог века као прагматичан приступ надзору видео квалитета у оквиру мултимедијалних комуникационих мрежа (у реалном времену), имајући у виду данашњи брзи развој апликација мрежних видео сервиса напредак у области процене квалитета уз делимично референцирање (као и без референцирања) је јако спор [6, 212].

На Сл. 6.1 илустрован је генерални концепт рада система са делимичним референцирањем на изворни сигнал (слику или видео). На страни извора (предаје) врши се издвајање обележја референтног визуелног сигнала. Издвојена обележја се преносе на страну посматрача кроз додатни канал. Најчешће се подразумева да је трансфер информација у овом каналу без грешака. На страни посматрача се такође користи издвајање обележја из сигнала са деградацијом. Поступак издвајања информација из сигнала са деградацијом може бити исти као и на предајној страни, или се може прилагодити доступним информацијама са предајне стране. У завршној фази врши се поређење издвојених обележја из изворног сигнала и сигнала са деградацијом како би се добила вредност која одсликава квалитет сигнала са деградацијом.

6. Процена квалитета видеа у условима редукованог референцирања



Сл. 6.1 Генерални приступ при процени квалитета слике/видеа са делимичним референцирањем.

У основи RR мера је статистичка анализа очувања обележја изворног сигнала у сигналу са деградацијом. Код ових техника могу се користити делови изворног сигнала од интереса или карактеристике изведене из самог сигнала. Наравно, обележја која се преносе морају имати много мању битску брзину од некомпримоване поворке бита слике/видеа да би мере биле практичне. Ова битска брзина је одређена капацитетом додатног канала којим се врши трансфер података или меморијом за њихово складиштење [261].

У [262] се у анализи процене квалитета слика користи референцирање дела изворног сигнала. За процену квалитета коришћене су PSNR и SSIM технике. Са изворних слика изабран је сет репрезентативних пиксела (за одређивање позиције пиксела од интереса коришћен је генератор псеудо-случајних бројева са униформном расподелом). Аутори су посматрали одступање процене квалитета са референцирањем на део изворног сигнала, од процене квалитета која се добија када је комплетан изворни сигнал доступан. Аутори су закључили да се SSIM процене квалитета разликују за мање од 1% уколико се са слика димензија 512x512 пиксела узме свега 50 одбирака (делова слика), док се при истом степену редукције изворних информација PSNR вредности разликују чак и за 18%.

Случајан одабир позиција пиксела од интереса повлачи са собом и пренос позиција случајно изабраних пиксела чиме се повећава количина потребних података које је потребно пренети на страну пријемника. Такође, оваквим избором обележја се врши оскудан опис референтног сигнала и не врши адекватно сажимање

слике/видеа [6]. Додатно, дисторзије могу погодити пикселе који нису одабрани и на тај начин се неће детектовати битне промене. Због тога је један од највећих изазова за истраживаче у области процене квалитета са редукованим референцирањем избор сета обележја уз ограничену битску брзину додатне линије везе или канала за пренос (избор битних али довољних обележја).

Добар сет обележја потребно је да на ефикасан начин сумира садржај референтног сигнала, да је осетљив на специфичне дисторзије или шири спектар дисторзија и наравно да обухвата аспект субјективитета, односно човека као посматрача визуелних сигнала [6]. Тако су предложени алгоритми за RR процену квалитета где се за издвајање обележја користе модели статистике природних сцена и маргинална расподела вејвлет коефицијената слике [263], хибридни вејвлети и дирекционе банке филтара [264], ивице делова изворне слике [192], разлика ентропија вејвлет коефицијената референтног сигнала и сигнала са дисторзијом [265], губитак просторних информација и статистика хистограма разлике суседних кадрова [266], средње вредности нивоа сивог и случајне пројекције макроблокова у слици [267], коефицијенти бројних фреквенцијских трансформација [268], структурне разлике настале у каналу преноса [151], структурне разлике праћене различитим аспектима људске перцепције слике [82] и сл.

Приступ из [263] користи свега 162 бита по слици. Приступом из [192] се добијају добре перформансе ако се користи 1/1290 део изворног некомпимованог сигнала. У раду [265] је предложена фамилија алгоритама за процену квалитета са редукованим референцирањем, где се алгоритми разликују по количини података изворног сигнала на основу које се врши процена квалитета. Када су аутори у процени квалитета користили 2.5% изворног сигнала, перформансе су биле у рангу перформанси најбољих техника са потпуним референцирањем. Када се за процену квалитета користи само један скалар изведен из референтног сигнала, перформансе на нивоу појединачних дисторзија изворних слика су у рангу најбољих техника за процену квалитета са потпуним референцирањем, док је постизање добрих перформанси на глобалном нивоу – нивоу сета слика са различитим деградацијама предмет даљих истраживања. У раду [266] је имплементирана RR техника за процену квалитета у реалном систему извор-корисник. Предложена имплементација

за пренос обележја од интереса користи интернет комуникацију, при чему је битска брзина RR обележја издвојених из изворног сигнала мања од 10 kbit/s. Вођени добрим перформансама индекса структурне сличности, аутори из [267] су предложили алгоритам којим се коришћењем RR обележја апроксимира FR SSIM вредност (SSIM добијен потпуним познавањем изворног сигнала). Аутори су навели да за њихов приступ потребан капацитета додатног канала од 38 kbit/s. Кроз мултивеличинску геометријску анализу и аспекте визуелног система човека, аутори из [268] су кроз бројне декомпозиције сигнала и издвајање обележја постигли висок степен слагања субјективних и објективних скорова. У зависности од типа декомпозиције сигнала, слика је представљена са веома малим бројем обележја (од 10 до 25 карактеристика).

Једно од питања код RR техника је како пренети RR карактеристике на страну пријемника. Ово обично захтева додатни, поуздани канал који може бити скуп или недоступан. Један интересантан метод за праћење квалитета слике без додатног канала коришћењем модерних техника воденог жига је предложен у [269]. Код ове технике издвојена обележја утискују се у слику са деградацијом у облику порука које нису видљиве посматрачу. Овакве технике не захтевају додатни канал, а не утичу на визуелни квалитет преношене слике. Сличан приступ је предложен за RR процену квалитета видеа [270], где се издвојена обележја утискују у преношени видео коришћењем 3D дискретне косинусне трансформације. Додатно је у систем преноса уграђен и блок за детекцију грешака током преноса како би се повећала робусност система.

Обележја изворних сигнала која се користе за процену квалитета сигнала на страни посматрача омогућавају крајњим корисницима и да на основу њих изврше поправку квалитета пријемног сигнала [269].

Генерално, у бројним радовима из области процене квалитета са делимичним референцирањем показано је да се са веома малом битском брзином додатног канала могу постићи предикције квалитета које су веома близу перформанси техника са потпуним референцирањем, а неки од приступа су и тестирани на реалним системима преноса и праћења квалитета видеа коришћењем додатног канала за пренос информација [271].

Након тестова спроведених од стране VQEG групе и позитивног мишљења о техникама делимичног референцирања при процени квалитета, ITU унија је донела неколико препорука из ове области [35-37, 40]. У наведеним препорукама описује се мера EPSNR (Edge Peak Signal to Noise Ratio), код које се за делимично референцирање користи деградација ивица изворног сигнала. Мера је заснована на запажању да је визуелни систем човека осетљив на деградације у околини ивица. Капацитет додатног канала који се користи за пренос информација о изворним видео секвенцама за сваку од наведених препорука дат је у Таб. 6.1. Види се да капацитет додатног канала зависи од резолуције видео формата, при чему не прелази 256 kbit/s.

Таб. 6.1 Типичне битске брзине за додатни канал делимичног референцирања.

Препорука	Видео формат	Капацитет додатног канала
ITU-T J.246 (2008)	QCIF (176x144 пиксела)	1 kbit/s, 10 kbit/s
	CIF (352x288 пиксела)	10 kbit/s, 64 kbit/s
	VGA (640x480 пиксела)	10 kbit/s, 64 kbit/s, 128 kbit/s
ITU-T J.249 (2010)	525 линија (720x486 пиксела)	15 kbit/s, 80 kbit/s, 256 kbit/s
	625 линија (720x576 пиксела)	15 kbit/s, 80 kbit/s, 256 kbit/s
ITU-R BT.1867 (2010)	QCIF (176x144 пиксела)	1 kbit/s, 10 kbit/s
	CIF (352x288 пиксела)	10 kbit/s, 64 kbit/s
	VGA (640x480 пиксела)	10 kbit/s, 64 kbit/s, 128 kbit/s
ITU-T J.342 (2011)	1080i 60 Hz (29.97 fps) 1080p (29.97 fps) (1920x1080 пиксела)	56 kbit/s, 128 kbit/s, 256 kbit/s
	1080i 50 Hz (25 fps) 1080p (25 fps) (1920x1080 пиксела)	56 kbit/s, 128 kbit/s, 256 kbit/s

У блиској вези са истраживањима у области процене квалитета са делимичним референцирањем су истраживања у којима се анализира одзив посматрача на различите визуелне стимулације. Оваква истраживања могу помоћи при одабиру значајних обележја сигнала побуде за редуковано референцирање.

Чуло вида је доминантан канал перцепције преко кога човек остварује комуникацију са информационим и комуникационим системима. Међутим, главна препрека људске визуелне комуникације је то да човек у једном тренутку може обрађати пажњу на ограничен број догађаја односно појава. Ово има велики утицај на визуелну комуникацију, јер оно што представља успешну комуникацију зависи од бројних механизма у мозгу којима се деле расположиви ресурси и одређује где ће човек гледати и чему ће посветити пажњу [273].

Значај визуелних информација је битна компонента визуелног система човека (HVS) и он је у доброј мери анализиран кроз увезивање са проценом квалитета слике и видеа. Интуитивно се може очекивати да увођење значаја визуелних информација треба да поправи перформансе мере за објективну евалуацију квалитета сигнала. На пример, артефакт (деградација) у региону од интереса је много битнији од деградације која постоји у неупадљивом региону. Због тога се често за побољшање перформанси мера објективне процене користи увођење значаја визуелних информација. Генерално, бројне мере процене квалитета слике и видеа са потпуним референцирањем су имплементирани кроз две фазе. У првој фази се локално се евалуира дисторзија слике/видеа и као резултат добија мапа квалитета/дисторзије. У другој фази се користе механизми просторног и/или временског комбиновања локалних скорова из мапе квалитета у једну вредност која одсликава квалитет посматране слике/видеа [238].

У литератури се могу наћи покушаји да се значај визуелних информација користи за побољшање предикције квалитета објективних мера [124-148]. Визуелно значајни подаци се могу прикупити кроз субјективне експерименте (помоћу система који служе за прикупљање података о кретању зенице људског ока), или се могу моделовати. На први поглед, коришћење реалних података је прикладније, јер резултати не зависе од поузданости модела визуелног значаја. Тако су истраживачи у радовима [273, 274] показали да се подаци о кретању зенице људског ока прикупљени током евалуације квалитета JPEG компримованих слика разликују од података прикупљених током посматрања истог сета слика али без неког унапред задатог циља (free looking). Ови резултати указују да артефакти (деградације) могу одвратити пажњу од региона који по статистици припадају природним сценама (тамо где би човек у просеку гледао и на шта се ослањају модели људског визуелног система), и на тај начин утицати на субјективну процену квалитета. Деградација неких региона може бити мање уочљива због ефеката маскирања осветљаја или текстуре. Додатно се применом JPEG компресије могу добити артефакти који су уочљивији у позадини него у предњем плану слике. На овај начин се одступа од статистике значаја изведене из посматрања слика природних сцена [275].

Последњих година истраживачи су се фокусирали на анализу утицаја дисторзија слике на кретања ока и фиксације на слици. Са друге стране, веома мало

радова је посвећено утицају дисторзија током посматрања видео секвенци, где су дисторзије настале компресијом генерално много јасније него на сликама.

Утицаји циља субјективне евалуације квалитета и садржаја сцене на импресију субјективног квалитета компримованог видео дати су у [276]. Заједно са субјективном импресијом квалитета снимани су и резултати кретања зенице посматрача. Резултати указују да на субјективну импресију квалитета велики утицај има присуство/одсуство деградације људских фигура у предњем плану сцене. Такође, потврђено је да на субјективну импресију квалитета утичу и инструкције које се дају посматрачима пре спровођења теста (на који део слике да се фокусирају – предњи план или позадину).

Аутори из [277] су истражили утицај кретања на сцени на субјективни значај региона. Њихови експерименти су показали да је кретање битно обележје које утиче за субјективни значај пиксела/региона. Слична анализа је спроведена у [278], где су проучавана кретања зенице људског ока при различитим кретањима камере – при промени азимута, елевације и зума.

Аутори из [279] су анализирали како дисторзије компримованог видео утичу на кретања ока током посматрања видео записа. Анализирано је на који начин циљ са којим се спроводи експеримент (процена квалитета видео или опис дешавања на сцени) утиче на кретања ока. Такође, анализирано је на који начин степен дисторзије утиче на кретања ока током експеримената спровођених са различитим циљевима. Као и у неким претходним студијама, аутори су закључили да циљ који је презентован посматрачима значајно утиче на кретања људског ока током субјективних тестова. Аутори су дошли до закључка да у тесту где су посматрачи добили задатак да посматрају видео секвенце са циљем да опишу дешавања на сцени, постоје статистичке разлике кретања ока приликом посматрања видео без дисторзије и видео са дисторзијом. Међутим, у субјективним експериментима спроведеним са циљем да се евалуира квалитет видео, постоји већи степен слагања кретања људског ока приликом посматрања видео без дисторзије и видео са дисторзијом. Из овога се може закључити да се стратегија претраживања коју посматрачи користе у евалуацији квалитета видео са деградацијом значајно не мења (стратегија зависи од садржаја видео и типа деградације). Једна од јавно доступних

база фиксација током гледања видео секвенци описана је у [280]. Како су посматрачи два пута гледали исти сет видео секвенци, аутори су након анализе закључили да се мапе фиксација од исте секвенце разликују. Данас се може наћи преко 20 јавно доступних база фиксација на сликама и видео секвенцама [132, 281].

Кроз бројна експериментална истраживања показано је да се увођењем реалних мапа субјективног значаја могу побољшати перформансе објективних мера процене квалитета слика. Добитак зависи од садржаја слике, избора објективне мере и типа дисторзије изворних слика. Како је највећи број радова обухватио интеграцију реалних мапа субјективног значаја са мерама процене квалитета слика, у дисертацији је анализирана примена реалних мапа субјективног значаја у евалуацији квалитета видео сигнала.

У даљем делу овог поглавља биће описан начин формирања мапа субјективног значаја изведених из реалних (и случајних) фиксација и начин интеграције ових мапа са објективним мерама које дају евалуацију квалитета на нивоу пиксела. Перформансе овако пондерисаних објективних мера за процену квалитета видеа биће дате у другом делу поглавља, заједно са најбитнијим запажањима и правцима даљег рада.

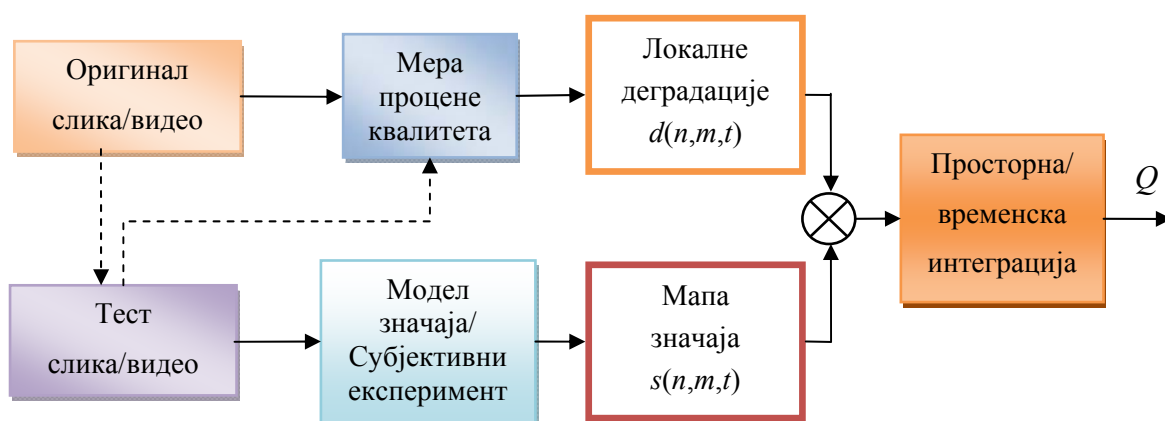
6.2. Увођење мапе субјективног значаја

Интуитивно је јасно да су за посматрача неки региони на слици битнији (атрактивнији) него суседни и тиме привлаче већу пажњу. Већа пажња се може објаснити тиме да локалне карактеристике слике стимулишу процесе који се одвијају у визуелном систему човека, као што су локални контраст, присуство ивица, оријентација, боја, кретање и слично. Осим ових једноставних обележја, на пажњу може утицати и искуство посматрача, који на слици од мноштва објеката тражи специфичне облике (рецимо људско лице, возило и сл.).

Мапа визуелног значаја идентификује (представља) битне регионе на слици. То су региони који ће највероватније привући пажњу посматрача. Она се може посматрати као дводимензионални запис локација свих фиксација, било да су у питању индивидуални или групни записи. Оваква мапа врши топографско

енкодовање упадљивих локалних региона на слици, и може се упоредити са мапом неког предела где узвишење представља регион од интереса за посматрача [282].

Генерални модел увођења значаја визуелних информација у објективној процени квалитета илустрован је на Сл. 6.2. Значај визуелних информација се може добити коришћењем резултата субјективних тестова или увођењем различитих модела изведених на основу субјективних тестова или разумних претпоставки. У нашем случају се користе резултати субјективних тестова у форми снимљених путања кретања људског ока, чиме се уводи субјективни значај за сваку од локација (пиксела) у тест секвенци/слици. Из фиксација се добија мапа значаја којом се пондеришу локалне вредности квалитета, и коначно се добијене вредности комбинују (интеграле) у једну вредност која представља меру квалитета секвенце која се приказује посматрачу.



Сл. 6.2 Уобичајени приступ интеграције значаја визуелних информација у процес процене квалитета.

Визуелни значај код човека је уско повезан са кретањем његовог ока. Под претпоставком да на пажњу утиче значај различитих делова сцене, снимљене локације фиксација се могу искористити за добијање мапе расподеле визуелног значаја (подједнако битна за одређивање значаја и процену квалитета је и идентификација региона које посматрач није гледао). Мапа значаја се из снимљених кретања људског ока у раду одредила у два корака. У првом кораку, кретања ока посматрача су преведена у значајне тачке, тј. битне локације на слици, а након тога су ове локације преведене у континуалне мапе значаја односно пажње.

6.3. Експериментални подаци о кретању зенице људског ока

Након прве студије у којој су прикупљене субјективне импресије квалитета на УоМ сету (студија је описана у претходном поглављу дисертације, део 5.9.1), на Универзитету у Манчестеру спроведена је и друга студија са циљем да се прикупе подаци о томе где посматрачи гледају током евалуације квалитета. У другој студији су учествовала четири посматрача, који су евалуацију радила у оквиру четири сесије.

Током прикупљања субјективних импресија о евалуираним секвенцама извршено је и прикупљање података о покретима ока посматрача. За аквизицију података коришћен је ASL 500 Eye Tracker систем. То је систем који се ставља на главу (Сл. 6.3), користи инфрацрвени извор и камеру за прикупљање података. Тачност система је боља од 0.5° а подаци о покретима ока се узимају са учестаношћу од 50 Hz. Покрети главе посматрача су ограничени помоћу држача за браду (Сл. 6.3).

За контролу презентације видеа и команде тригерована за прикупљање података са ASL Eye Tracker система коришћен је Psychtoolbox пакет [253, 254].



Сл. 6.3 (а) ASL 500 Eye Tracker систем и (б) експериментални услови друге студије у којима је спроведено прикупљање података.

Субјективно мишљење и трајекторије зенице ока се аутоматски архивирају. За архивирање покрета ока, од расположивих 257 секвенци коришћене су 192 (20 оригиналних и 172 секвенце са деградацијом).

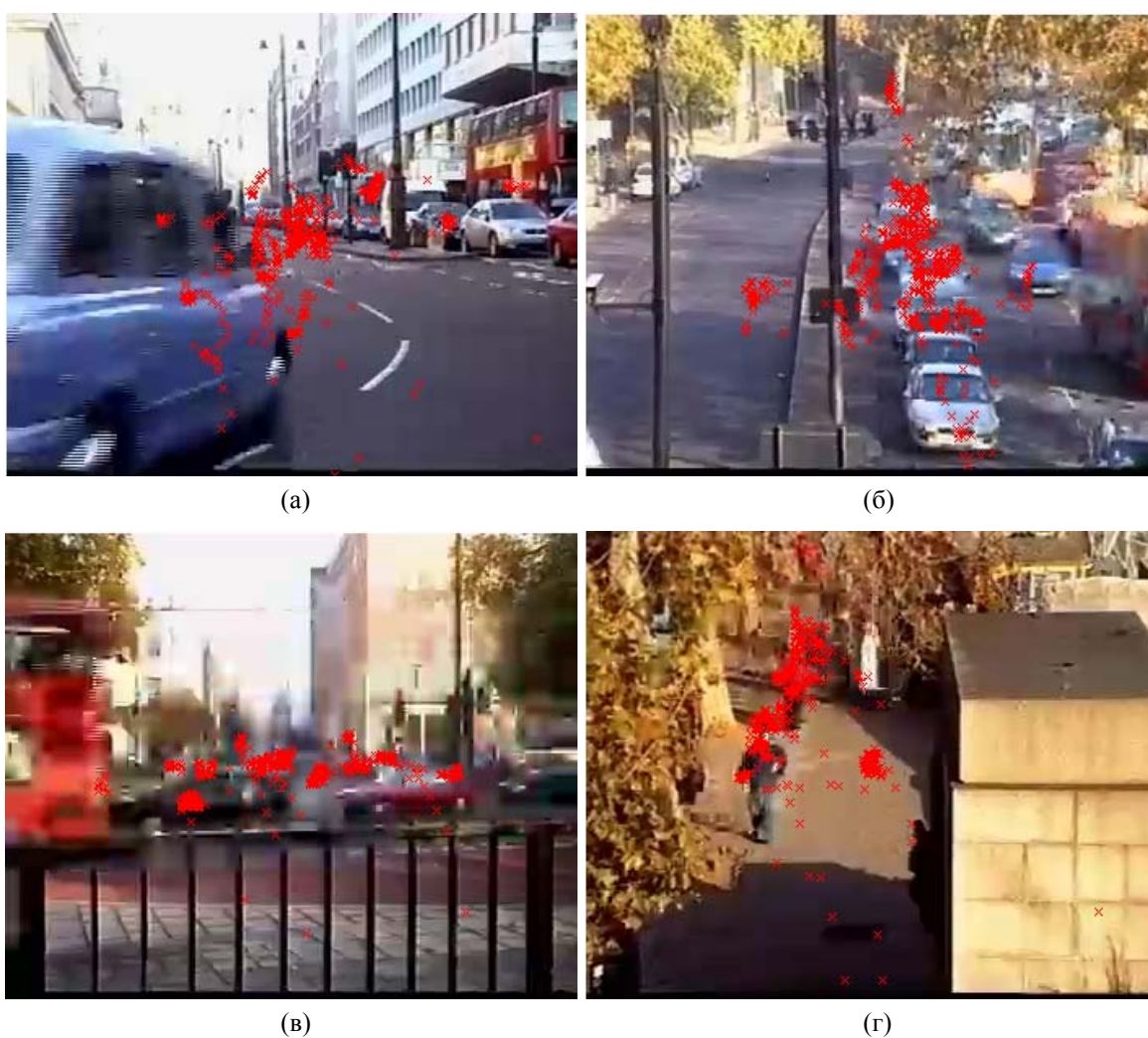
Колекције записа кретања зенице ока које су добијене од Универзитета у Манчестеру било је потребно обрадити и анализирати, тј. превести у формат погодан за издвајање корисних информација. Ово превођење обухвата упаривање снимљених фиксација и оригиналних видеа и усаглашавање информација са видео сетовима коришћеним у субјективним тестовима и евалуацији објективних мера – снимци кретања ока сваког посматрача сведени су на формат који одговара видном углу секвенце из базе. Након упаривања, подаци о фиксацијама се могу директно користити у процесу пондерисања локалних вредности квалитета анализираних мера. Уколико снимљене фиксације одговарају регионима на основу којих су посматрачи дали мишљење о квалитету секвенце, за очекивати је да се перформансе објективних мера побољшају. Другим речима, циљ овог дела дисертације је да се анализирају субјективно пондерисане објективне евалуације квалитета и да се покуша одговорити на питање: Могу ли се коришћењем стварних (мерених) субјективно битних информација у поступку евалуације квалитета побољшати перформансе мере објективне процене квалитета видеа?

Нормализација скале и подешавање резолуције су спроведени како би се усагласиле разлике које постоје у видео секвенцама, доступним у 4 резолуције, и посматраних секвенци које су приказиване једино у највећој резолуцији (352x288 пиксела). Кретања зенице ока су снимљена у приказаној (највећој) резолуцији и било их је потребно скалирати на видни угао сваке од секвенци.

На Сл. 6.4 илустровани су прикупљени подаци на четири видео секвенце UoM сета. Фиксације су приказане преко 25.-тог кадра секвенци, који нужно не представља сва дешавања која се могу видети у комплетној секвенци. У сва четири случаја фиксације прате покрете (кретања) у секвенци: кретање пешака и возила горе лево, кретање колоне возила горе десно, кретање возила и пешака доле лево и кретање пешака доле десно.

Један од недостатака доступног сета је то што садржи свега пет различитих сцена. На тај начин је било тешко избећи ефекте који утичу на резултате – замор посматрача и досаду. Свака од сесија има 80 видео записа од по 10 секунди, и узимањем у обзир време припреме посматрача, укупно време спровођења

субјективног теста је било дуже од пола часа, што представља велики изазов за очување пажње посматрача. Као резултат, део фиксација је био ван приказане слике на екрану.



Сл. 6.4 Приказ фиксација посматрача преко кадрова тест секвенци.

Међутим, из приказаних фиксација са Сл. 6.4 може се рећи да локације фиксација нису случајне и да се могу користити за одређивање региона који ће посматрачу привући више пажње а тиме и пондерисати са резултатима објективних мера процене. У следећем делу дисертације ће бити описан начин на који су ове информације коришћене заједно са мерама објективне процене квалитета.

6.4. Формирање мапе визуелног значаја из реалних узорака

Како би се формирала мапа визуелног значаја за сваки од видео записа у УоМ бази, подаци о кретањима ока посматрача су искоришћени за формирање мапе значаја за сваког од посматрача – у сваком кадру видео секвенце и за комплетан видео. Мапа значаја је базирана на фиксацијама за сваку од просторних позиција (n_j, m_j) , тако да се мапа значаја SM_k за k -тог посматрача добија као:

$$SM_k(n, m) = \sum_{j=1}^{N_{FP}} \delta(n - n_j, m - m_j) \quad (78)$$

где је N_{FP} број фиксација прикупљених током теста а δ Диракова делта функција. Свака фиксација има исту тежину, тј. вредност 1. На овај начин се добија бинарна мапа значаја, димензија као и оригинална слика и која има вредност 0 на свим позицијама осим на позицијама интересантним за посматрача.

Како би се генерализовала мапа значаја за сваку од секвенци и у суштини одредило који су визуелно најбитнији региони, бинарне мапе значаја свих посматрача су сведене на усредњене мапе, SM , које се добијају као:

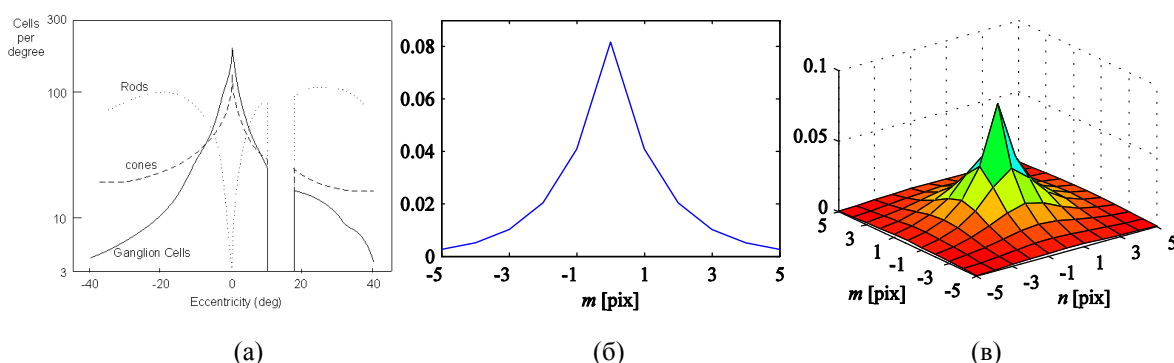
$$SM(n, m) = \frac{1}{K} \sum_{k=1}^K SM_k(n, m) \quad (79)$$

где је K укупан број посматрача (за УоМ базу видео секвенци $K=4$).

Пример суперпозиције фиксација приказан на Сл. 6.4 генерално указује на централне локације од интереса за посматрача. Регион који визуелни систем човека истражује а чији је центар на датој локацији има коначну површину која је одређена централним видним углом визуелног система. То указује да је потребно размотрити регион на слици око тачке од интереса како би се коректно моделовало овакво понашање. Због тога је полазећи од бинарне, дводимензионалне мапе фиксација SM потребно креирати континуалну мапу значаја којом ће се пондерисати локалне процене квалитета.

Фоторецептори (чепићи (cones) и штапићи (rods)) и ћелије ганглија нису униформно расподељени у мрежњачи људског ока, што је илустровано на Сл. 6.5(а). Штапићи су врло осетљиви фоторецептори који омогућавају вид при врло ниским

нивоима осветљености али нису способни да разликују боје. Заузимају већи део мрежњаче и омогућавају да видно поље сваког ока буде преко 180 степени. Највише су концентрисани око центра мрежњаче, тј. у опсегу од 10 до 20 угаоних степени око места где оптичка оса система пролази кроз мрежњачу (овај део мрежњаче назива се фовеа – жута мрља). Другу групу фоторецептора чине чепићи. Налазе се у уском појасу од око 5 угаоних степени у односу на место где оптичка оса система пролази кроз мрежњачу. Чепићи су задужени за fine детаље у слици и они су једине ћелије које су осетљиве на боју. Густина чепића у жутој мрљи је око 120 чепића по лучном степену видног поља, па је угаона резолуција ока око 0.5' [245]. Када се човек фокусира на неку тачку, слика сцене променљиве резолуције се преноси са почетка визуелног канала ка вишим нивоима процесирања у људском мозгу. Регион око фиксационе тачке се пројектује на жуту мрљу са највећом резолуцијом и самим тим се прима са највећом осетљивошћу.



Сл. 6.5 (а) густина фоторецептора и ћелија ганглија на мрежњачи, (б) експоненцијални филтар који се користи за моделовање одзива ћелија мрежњаче и (в) 3-D приказ филтра који се користи у формирању континуалне мапе значаја.

Узимајући у обзир одзив ћелија мрежњаче приказан на Сл. 6.5(а) може се доћи до закључка да значај унутар региона од интереса опада са повећањем удаљености од позиције фиксације. Овај ефекат се може моделовати са Гаусовим или експоненцијалним филтром [238]. Пример експоненцијалног филтра дат је на Сл. 6.5(б), где се јасно уочава да облик филтра прати промену густине фоторецептора задужених за прикупљање информација о финим детаљима са сцене.

У раду се за моделовање ове појаве (значаја пиксела) користила кружно-симетрична 2-D експоненцијална функција:

$$e_s(d) = 0.5^d \quad (80)$$

где је d удаљеност од централног пиксела.

Овакво понашање је идентично дуж обе осе (x и y), што се може видети на Сл. 6.5(в). Како веома мали број локација око централног пиксела садржи значајне коефицијенте, у раду су узете димензије филтра 11x11 пиксела, тј. централни пиксел и по пет пиксела на сваку страну. Добијени филтар се користи за филтрацију локација фиксација како би се добила континуална мапа значаја.

Континуалне мапе субјективног значаја се добијају филтрирањем мапа фиксација SM са 2-D експоненцијалним филтром e_s . Како је фреквенција узимања фиксација значајно већа од фреквенције прикупљања кадрова (fps), сваки кадар секвенце садржи више фиксација. Свака од ових локација на појединачним кадровима се користи како би се креирала мапа значајних локација кадра, тј. ова мапа се филтрира како би се добила континуална мапа субјективног значаја. Као резултат се добија 2-D мапа субјективног значаја $s(n,m,t)$ за сваки од пиксела слике (n,m) , односно секвенце:

$$s(n,m,t) = SM(n,m,t) * e_s(n,m) \quad (81)$$

Резултујућа мапа субјективног значаја $s(n,m,t)$ одсликава најинтересантније делове кадра за групу посматрача, тј. представља просечан одзив посматрача.

На Сл. 6.6 и Сл. 6.7 приказани су кадрови једне од изворних секвенци УоМ базе и одговарајући кадрови секвенце са деградацијом (темпо = 5 fps; трајање секвенце = 50 кадрова). Као и код осталих секвенци из ове база камера је стационарна а на сцени постоје покретни објекти, што су у овом случају возила саобраћајног тока која се приближавају или удаљавају од сензора.

На слици Сл. 6.7 се уочавају ефекти грубе квантизације EV алгоритма компресије којим се губе фини детаљи региона на слици/секвенци, са блоковским ефектом који је веома уочљив на компримованим сликама.

6. Процена квалитета видео у условима редукованог референцирања

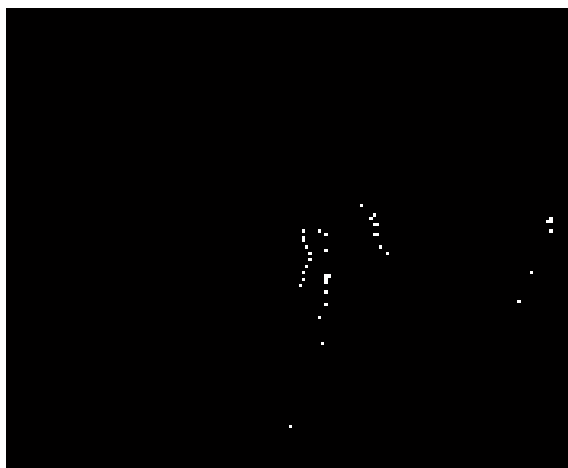


Сл. 6.6 Кадрови изворне секвенце 'Original_176x144_Road_50.avi'.

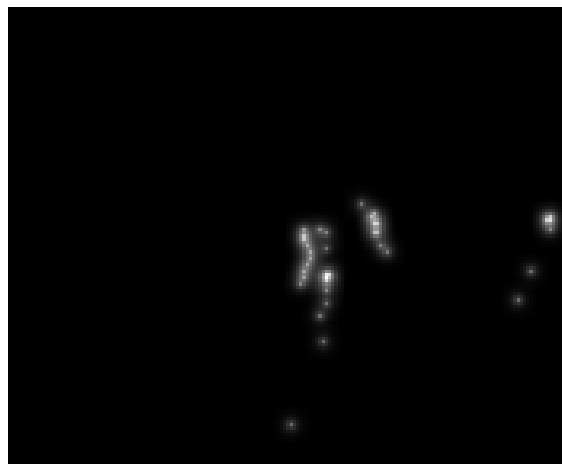


Сл. 6.7 Кадрови секвенце са деградацијом – EV компресијом 'EV_176x144_576_S3_014.avi'.

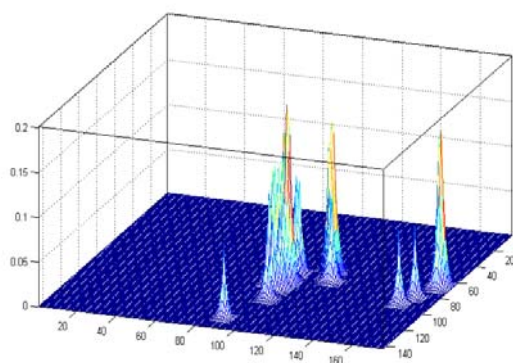
На Сл. 6.8 приказане су позиције фиксација у првом кадру секвенце са деградацијом чији су кадрови дати на Сл. 6.7. Позиције фиксација су на Сл. 6.8(а) приказане белим тачкама на црној позадини. Резултујућа мапа значаја након филтрације је приказана као сива слика на Сл. 6.8(б). Ова мапа значаја је на много живописнији начин представљена у 3-D простору на Сл. 6.8(г), где висина сваке тачке одређује њен субјективни значај. Уочава се да региони фиксација обухватају веома мали део комплетне сцене и да су груписани око неколико облика или ивица. То се боље уочава на Сл. 6.8(д), где су на самом кадру црвеном бојом оивичени региони који су визуелно значајни, тј. региони чији је значај већи од нуле. Сл. 6.8 показује да посматрачи, као што би неко и очекивао, углавном посвећују пажњу покретним објектима, у овом случају саобраћајном току.



(а)



(б)



(в)



(г)

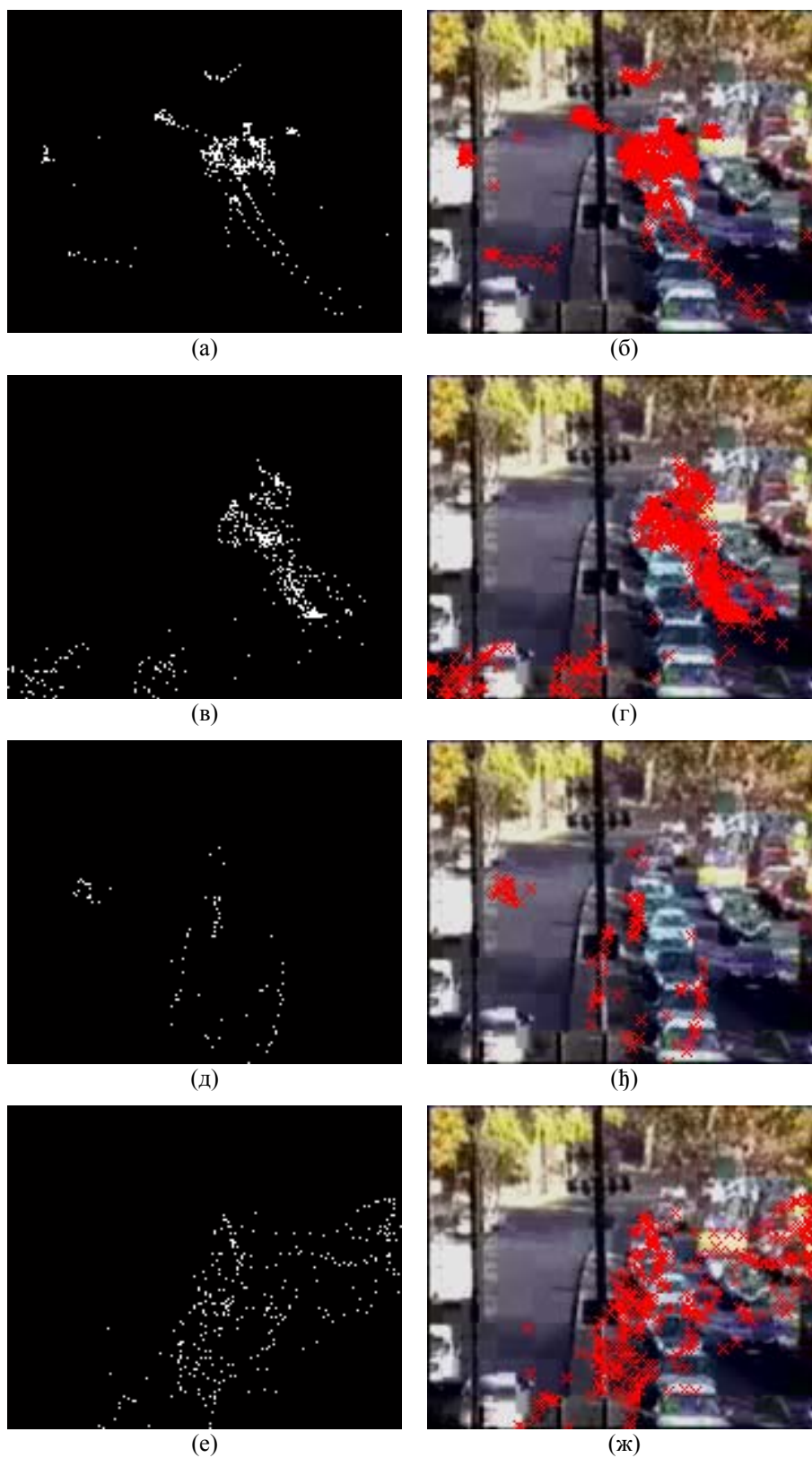
Сл. 6.8 Региони фиксација за први кадар секвенце са деградацијом анализане на Сл. 6.7: (а) бинарна мапа фиксација, (б) финална мапа значаја, (в) мапа значаја приказана у 3-D простору и (г) региони од значаја у анализираном кадру са деградацијом (црвеном бојом су оивичени региони за које је субјективни значај већи од нуле).

Како је резолуција анализираног кадра 176x144 пиксела и за њега је детектовано 37 фиксација (од максимално могућих 40), филтрирана мапа значаја би обухватила највише 17.7% од оригиналног сигнала који се користи у евалуацији квалитета (уколико не постоји преклапање региона што генерално није случај). Овако добијена мапа значаја је директно уведена у објективне мере процене квалитета кроз пондерисање локалних евалуација квалитета добијених коришћењем PSNR, SSIM и предложене мере VQ^{AB}.

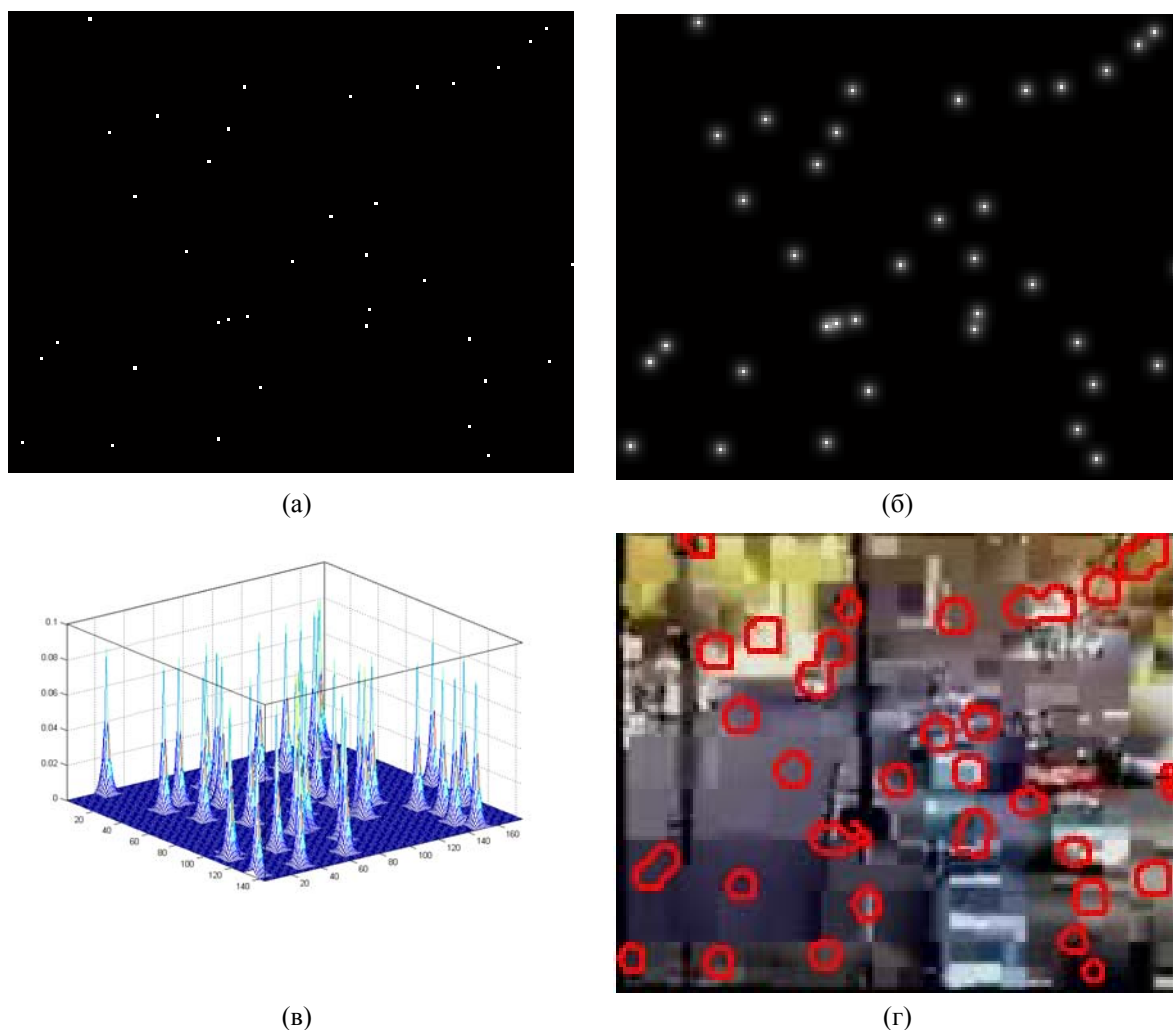
На Сл. 6.9 приказане су локације фиксација за комплетну секвенцу 'EV_176x144_576_S3_014.avi' дужине 50 кадрова и то за сва четири посматрача. Фиксације су приказане преко 36.-тог кадра тест секвенце. Генерално се може говорити о високом степену слагања субјективно битних региона у секвенци, односно о регионима који посматрачима привлаче већу пажњу. Расподела пажње за сва четири посматрача је условљена кретањима на сцени. Оно што је интересантно је да иако постоји релативно споро кретање саобраћајног тока ка посматрачу (десни део кадра), оно је привукло више пажње од возила која се брзо крећу у супротном смеру (леви део кадра).

Иако се може говорити о слагању битних региона за различите посматраче уочава се да посматрачи имају различите технике претраживања. Такође, може се уочити и да посматрач 3 није био фокусиран на секвенцу током свих 10 секунди њеног приказивања. Број снимљених фиксација за овог посматрача је значајно мањи од броја фиксација осталих посматрача, односно већина снимљених фиксација излази изван оквира приказиване секвенце. Пример са Сл. 6.9 је дат ради поређења интересантних региона различитих посматрача током комплетне видео секвенце. За потребе процене квалитета видео сигнала, фиксације су временски раздвојене и коришћене за формирање мапе визуелног значаја сваког кадра понаособ.

У оквиру рада спроведен је контролни експеримент у коме су реалне фиксације замењене са случајно расподељеним фиксацијама у видном пољу – број случајних фиксација у сваком од кадрова једнак је броју снимљених фиксација. Пример случајно расподељених фиксација приказан је на Сл. 6.10(a), док су региони случајних фиксација за претходно анализирани кадар са Сл. 6.8 приказани на Сл. 6.10(г). За сваки кадар прављена је друга мапа случајних фиксација.



Сл. 6.9 Позиције фиксација за сваког од посматрача током посматрања комплетне тест секвенце (позиције су приказане преко 36.-тог кадра анализиране секвенце): (а) (б) посматрач 1 (сесија 4, 37-а секвенца), (в) (г) посматрач 2 (сесија 1, 46-та секвенца), (д) (ж) посматрач 3 (сесија 3, 7-а секвенца) и (е) (ж) посматрач 4 (сесија 1, 31-а секвенца).



Сл. 6.10 Региони случајних фиксација за први кадар секвенце са деградацијом анализане на Сл. 6.7: (а) бинарна мапа фиксација, (б) финална мапа значаја, (в) мапа значаја приказана у 3-D простору и (г) случајни региони од значаја у анализираном кадру (оивчени црвеном бојом).

Ефектом увођења случајних локација од интереса се улази у домен мера са редукованим референцирањем. У претходно анализираном примеру, уколико би број фиксација био максималан (40 фиксација) у објективној евалуацији квалитета користило би се 19.1% информација од изворног сигнала. То би захтевало пренос додатних 19.1% информација оригиналног сигнала (односно 4.7% за максималну резолуцију анализирану у раду, 352x288 пиксела) како би се вршила евалуација квалитета на пријемној страни – страни посматрача. Иако је за некога пренос додатних 19.1% информација изворног сигнала мала цена којом се добија поуздана евалуација квалитета, ова количина информација је значајно већа од очекиване

величине компримованог низа информација и значајно већа од онога што би било очекивано у практичним мерама (системима) са редукованим референцирањем.

6.5. Реално пондерисане објективне мере процене квалитета

Реално пондерисане објективне мере су евалуиране на УоМ бази секвенци, за коју су прикупљене трајекторије кретања зенице људског ока. Прикупљени подаци су коришћени како би се добиле мапе субјективног значаја (поступак је описан у делу дисертације 6.4), које су потом интегрисане са резултатима објективних мера како би се одредио њихов допринос у процени квалитета. Уведене мапе субјективног значаја/пажње су интегрисане са три објективне мере: PSNR, SSIM и предложеном мером VQ^{AB} (као и њеним саставним елементима). Наведене објективне мере врше процену квалитета локално, дајући квантитативне вредности дисторзије/квалитета.

Мапе субјективног значаја су укључене у PSNR, SSIM и VQ^{AB} мере кроз пондерисање локалних скорова квалитета. Иако постоје бројне стратегије комбиновања мапа значаја са резултатима објективних мера [124, 237], у раду се користила једноставна стратегија – множење и сумирање. Увођењем мапа субјективног значаја добијају се пондерисане објективне мере које ће у раду бити обележене индексом W испред стандардних назива ($WPSNR$, $WSSIM$, WVQ^{AB} и сл.).

6.5.1. Пондерисани вршни однос сигнал шум - WPSNR

Пондерисани вршни однос сигнал шум се добија као:

$$WPSNR = 10 \cdot \log_{10} \left(\frac{L^2}{WMSE} \right) \quad (82)$$

где је:

$$WMSE = \frac{\sum_{n=1}^N \sum_{m=1}^M (R(n, m) - D(n, m))^2 \cdot s(n, m)}{\sum_{n=1}^N \sum_{m=1}^M s(n, m)} \quad (83)$$

У (83) је са D означена слика са дисторзијом док R представља изворну слику. Димензије слика су $N \times M$ а $s(n, m)$ је одговарајућа мапа значаја изведена из субјективних експеримената и описана у претходном делу дисертације. Вредност L

из (82) представља динамички опсег слике, који је типично [0,255]. Код процене квалитета видео секвенци WPSNR се одређује за сваки кадар и усредњава како би се добила вредност квалитета секвенце – Frame WPSNR (користи се само канал интензитета).

6.5.2. Пондерисани индекс структурне сличности - WSSIM

Увођење мапе субјективног значаја код индекса структурне сличности врши се на сличан начин као и код вршног односа сигнал/шум, пондерисањем локалних скорова квалитета и њиховим усредњавањем:

$$WSSIM = \frac{\sum_{n=1}^N \sum_{m=1}^M SSIM(n, m) \cdot s(n, m)}{\sum_{n=1}^N \sum_{m=1}^M s(n, m)} \quad (84)$$

где се SSIM матрица добија коришћењем SSIM мере из [48] над сликом са дисторзијом, D , и оригиналом, R . Код процене квалитета видео секвенци WSSIM се одређује за сваки кадар понаособ и усредњава како би се добила вредност квалитета секвенце – Frame WSSIM (користи се само канал интензитета).

6.5.3. Пондерисана мера VQ^{AB} – WVQ^{AB}

Предложена мера VQ^{AB} (73) врши декомпозицију информација у три канала (просторни, временски и колор) и за сваки од канала уводи различите тежине ($[k_s, k_t, k_c]=[0.8, 0.15, 0.05]$). Додатно се након добијања вредности мере VQ^{AB} за сваки од кадрова секвенце са деградацијом користи варијанта која узима у обзир кадрове у којима је квалитет најлошији, тј. кренуло се од претпоставке да кадрови лошег квалитета могу имати пресудну улогу на субјективну импресију квалитета. Тако се за предложену меру прво врши анализа квалитета сваког од кадрова, након чега се добијене вредности квалитета сортирају и у укупној процени се користи $p\%$ кадрова са најмањим VQ^{AB} скоровима, тј. са најлошијим квалитетом (у анализи спроведеној у претходним деловима рада предложена је вредност $p=20\%$). Како су за тест секвенце фиксације доступне за сваки од кадрова, како би компаративна анализа резултата била валидна, у овом делу рада су за добијање вредности мере квалитета

VQ^{AB}/WVQ^{AB} коришћени сви кадрови тест секвенце, тј. извршено је усредњавање пондерисаних локалних скорова квалитета (средња вредност квалитета по кадровима).

Увођење мапе субјективног значаја код VQ^{AB} мере врши се на сличан начин као и код вршног односа сигнал/шум и индекса структурне сличности, пондерисањем локалних скорова квалитета и њиховим усредњавањем:

$$WVQ^{AB} = \frac{\sum_{\forall n,m,t} VQ_{n,m,t}^{AB} \cdot s_{n,m,t}}{\sum_{\forall n,m,t} s_{n,m,t}} \quad (85)$$

6.6. Анализа перформанси пондерисаних објективних мера на УоМ сету

У Таб. 6.2 дате су перформансе објективних мера процене (PSNR, SSIM и VQ^{AB}) након увођења мапа субјективног значаја s (перформансе две мере са најбољим резултатима су обележене масним словима). Коришћене су реалне мапе субјективног значаја изведене из снимљених фиксација и њима одговара ознака W (од weighted). Такође, уведене су и мапе значаја добијене на основу случајних позиција фиксација. Број фиксација унутар кадра одговара броју снимљених фиксација. Резултати који се добијају применом оваквих мапа значаја су обележени са W_{rand} (од random weighted).

Таб. 6.2 Перформансе објективних мера на УоМ сету.

Мера процене	LCC	SROCC	MAE	RMSE	OR [%]	
PSNR	Frame PSNR	0.8610	0.8121	0.3973	0.4666	85
	Frame WPSN	0.9057	0.8946	0.3184	0.3890	73.3333
	Frame W_{rand} PSNR	0.9090	0.8752	0.3031	0.3824	70
SSIM	Frame SSIM	0.9665	0.9689	0.1761	0.2353	58.3333
	Frame WSSIM	0.9685	0.9650	0.1868	0.2286	58.3333
	Frame W_{rand} SSIM	0.9666	0.9675	0.1716	0.2351	60
VQ^{AB}	VQ^{AB}	0.9743	0.9755	0.1574	0.2068	48.3333
	WVQ^{AB}	0.9684	0.9631	0.1689	0.2290	58.3333
	$W_{rand}VQ^{AB}$	0.9722	0.9762	0.1613	0.2148	53.3333

Разлике у перформансама стандардних техника из Таб. 6.2 (PSNR, SSIM и VQ^{AB}) и перформанси датих у Таб. 5.4 постоје због броја секвенци коришћених у

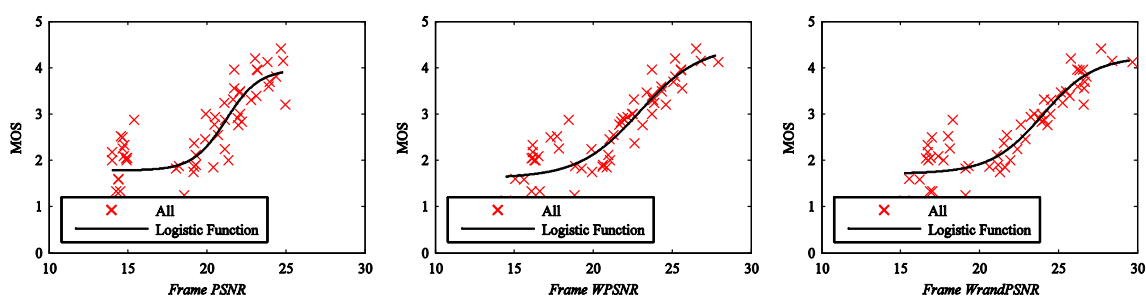
анализи (237 секвенци у претходном поглављу и 172 секвенце у овом поглављу). Додатно, резултати предложене мере VQ^{AB} из Таб. 5.4 су добијене коришћењем 20% најнижих скорова по кадровима, а овде се разматрају вредности квалитета свих кадрова.

Из Таб. 6.2 се уочава да се увођењем мапа субјективног значаја (реалних и случајних) значајно поправљају перформансе PSNR по свих пет параметара – око 5% се повећава корелација субјективних и објективних скорова. Како се перформансе побољшавају и увођењем случајних фиксација, може се закључити да би у процени квалитета требало користити локално усредњене PSNR вредности. Сличан закључак је изведен у [283] где је побољшање перформанси PSNR мере у процени квалитета видеа добијено усредњавањем PSNR вредности квалитета али по кадровима. Увођење мапа значаја у објективну процену квалитета додатно је илустровано дијаграмима расипања субјективних и објективних скорова на Сл. 6.11. Увођење локализованих реалних и случајних мапа значаја доводи до линеаризације PSNR вредности. Значајно се побољшавају резултати стандардне PSNR технике које су груписане са леве стране дијаграма расипања, где малим вредностима PSNR (око 14 dB) одговарају субјективне оцене које се крећу лошег до средњег квалитета (приближно од 1 до 3). Са Сл. 6.12 на којој су дати дијаграми расипања по типу кодека види се да те вредности одговарају једном од кодека (EV кодек). Пондерисањем локалних вредности квалитета ове вредности се приближавају тренду осталих кодека.

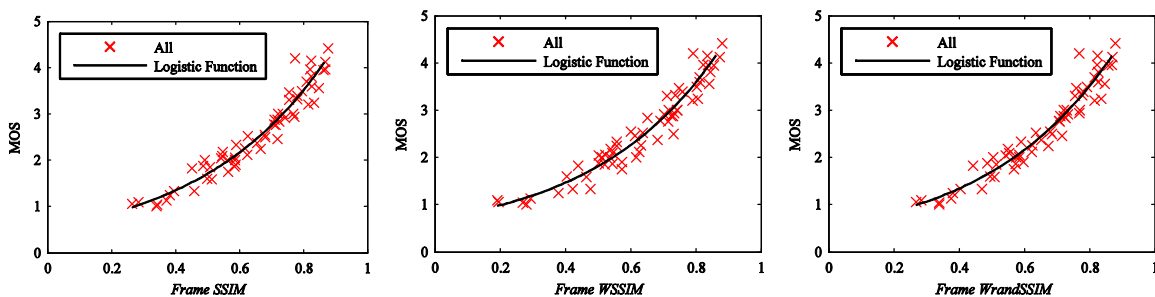
Увођење реалних мапа субјективног значаја код SSIM и VQ^{AB} мера не доводи до побољшања перформанси у односу на конвенционалне приступе (код VQ^{AB} дошло је до благог пада у перформансама – коефицијент линеарне корелације је за око 1% лошији). Имајући у виду јако висок степен слагања субјективних и конвенционалних објективних скорова (око 97%), то је можда било и за очекивати. Изненађујући је резултат да се случајним фиксацијама (односно редукованим референцирањем са случајним фиксацијама) добијају одлични резултати слагања субјективних и објективних скорова, како по SSIM тако и по VQ^{AB} мери. Овакав резултат се може тумачити квалитетом SSIM и VQ^{AB} мера. Наиме, добра објективна мера даје и на локалном нивоу (пиксел или група пиксела) добру процену квалитета.

6. Процена квалитета видеа у условима редукованог референцирања

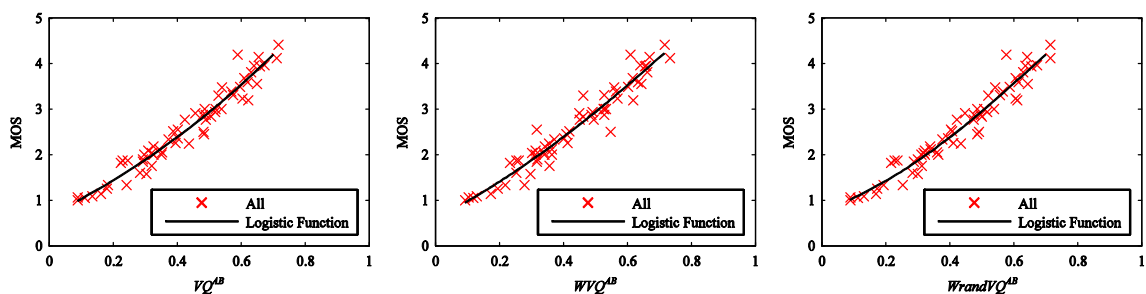
Стога, ако мера процене даје добре естимације квалитета у свим деловима слике/видеа, резултат објективне процене не би требало значајно да зависи од избора дела/делова изворног сигнала који се користе у процени – независно од тога да ли су у питању делови сигнала који одговарају реалним или случајним фиксацијама. Наравно, овакво тумачење важи под условом да модификација слике погађа комплетну слику, а што је у компресији случај. Како се коришћењем реалних и случајних фиксација перформансе SSIM и VQ^{AB} мера статистички не мењају, може се закључити да је то због поузданих процена квалитета на локалном нивоу. Дијаграми расипања објективних SSIM и VQ^{AB} вредности, дати на Сл. 6.11 и Сл. 6.12, приказују скоро линеарну везу између објективних и субјективних скорова, која је очувана увођењем како реалних тако и случајних фиксација.



(а) дијаграми расипања PSNR вредности



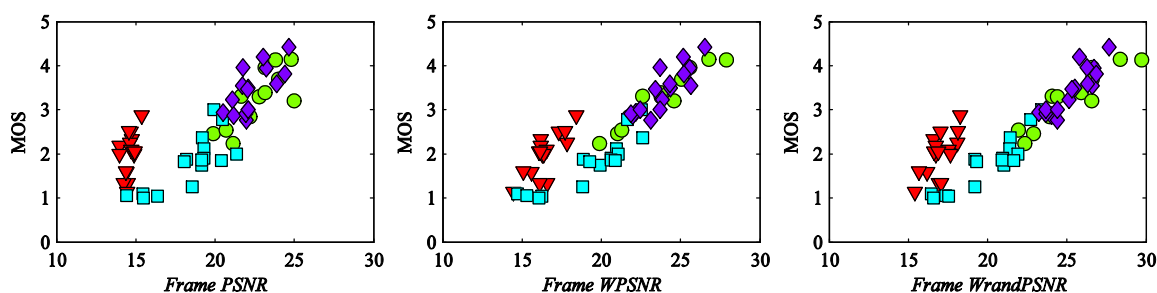
(б) дијаграми расипања SSIM вредности



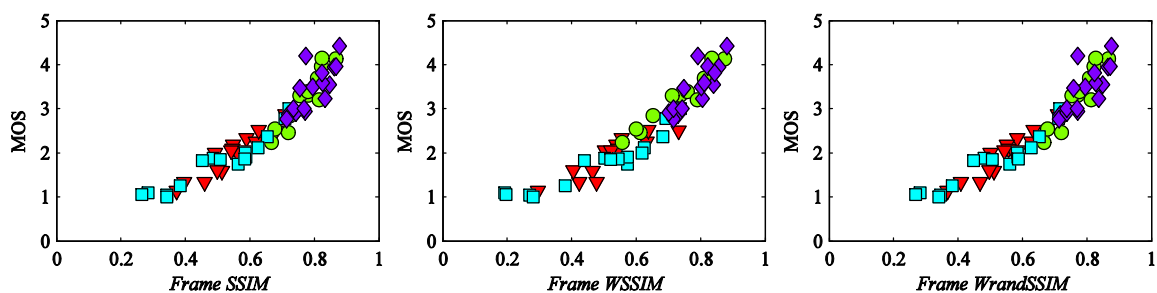
(в) дијаграми расипања VQ^{AB} вредности

Сл. 6.11 Дијаграми расипања објективних, пондерисаних објективних и субјективних скорова за секвенце у UoM сету: (а) PSNR, (б) SSIM и (в) VQ^{AB} .

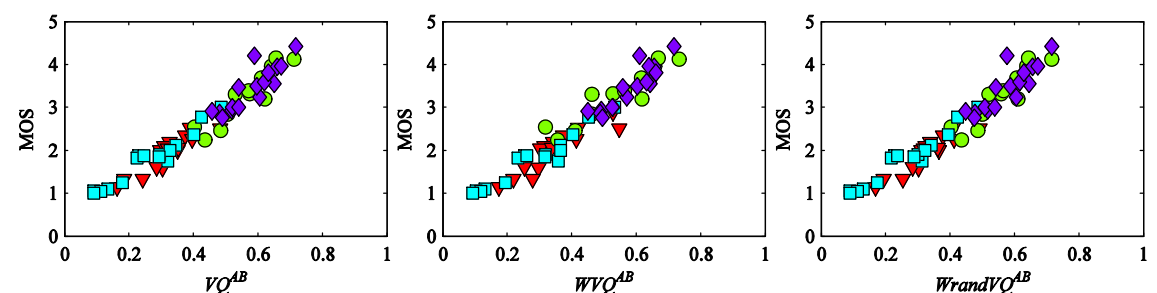
Такође, треба нагласити да изворно предложена мера за објективну процену квалитета видеа коришћењем SSIM [98] подразумева и неку врсту редукованог референцирања где се предлаже да се из комплетног кадра користи мање од 100 узорака (блокова) како би се радило у реалном времену али без значајног утицаја на перформансе саме мере.



(а) дијаграми расипања PSNR вредности



(б) дијаграми расипања SSIM вредности

(в) дијаграми расипања VQ^{AB} вредности

Сл. 6.12 Дијаграми расипања објективних, пондерисаних објективних и субјективних скорова за секвенце у UoM сету: (а) PSNR, (б) SSIM и (в) VQ^{AB} (анализа према типу кодека).

Генерално, добитак који се очекује од увођења реалних фиксација у објективну процену квалитета зависи од мере процене квалитета са којом се фиксације интегришу, односно фиксације се не могу одвојено посматрати од мере процене квалитета. Како се коришћењем дела изворног сигнала добијеним из

случајних фиксација (референцирање на део изворног сигнала) добија јако добро слагање субјективних и објективних скорова ($LCC_{\{SSIM, VQ\}} > 0.97$), може се закључити да је редуковано референцирање могуће и да се њиме могу достићи перформансе мера процене са потпуним референцирањем.

Увођењем мапе субјективног значаја s код објективне процене VQ^{AB} постоји двоструко пондерисање. Мера VQ^{AB} садржи нелинеарне сигмоидне функције и пондерисање са мапама субјективног значаја изведеним из вредности амплитуде градијента (просторни и временски градијент). Сигмоидне функције и мапе субјективног значаја $w_{\{s,t,c\}}$ су уведене како би се моделовао визуелни систем човека. Додатним увођењем мапа субјективног значаја s , које у себи садрже одзив (модел) просечног посматрача, врши се дуплирање утицаја посматрача. Због тога су у раду дати резултати процене квалитета (очувања информација) за сваки од канала који учествује у коначној процени и то кроз:

- репрезентацију изворних информација у сигналу са деградацијом: S^{AB} (линеарни модел очувања просторних информација (62)), T^{AB} (линеарни модел очувања временских информација (68)) и C^{AB} (линеарни модел очувања информација о боји (72)) мере;
- нелинеарно пресликавање репрезентације изворних информација у мере процене: Q_s (нелинеарни модел очувања просторних информација (58)), Q_t (нелинеарни модел очувања временских информација (65)) и Q_c (нелинеарни модел очувања информација о боји (70)) мере;
- репрезентацију изворних информација након нелинеарног пресликавања и пондерисања са мапама значаја w : Q_s^{AB} (нелинеарни модел очувања просторних информација са пондерисањем (60)), Q_t^{AB} (нелинеарни модел очувања временских информација са пондерисањем (67)) и Q_c^{AB} (нелинеарни модел очувања информација о боји са пондерисањем (70)) мере.

Анализа перформанси набројаних мера процене обухвата и пондерисање са мапама субјективног значаја изведеним из реалних експеримената и из случајних фиксација. Пондерисање се спроводи на локалном нивоу, тј. за сваку од мера се одређује локални квалитет, спроводи пондерисање са мапом значаја и врши интеграција локалних скорова у глобални скор квалитета. Пондерисање и

6. Процена квалитета видеа у условима редукованог референцирања

интеграција се спроводе по узору на (85). Додатно су процене очувања информација из сваког од канала искоришћене за одређивање вредности локалног квалитета VQ (са значајем евалуације из канала $[k_s, k_t, k_c]=[0.8, 0.15, 0.05]$). Локални квалитет срачунат је за сва три модела – линеарни модел, нелинеарни модел и нелинеарни модел са пондерисањем. Добијене локалне вредности скорова су пондерисане са мапом значаја и тако добијени производ је искоришћен за одређивање квалитета комплетне секвенце.

Перформансе мера процене квалитета у појединачним каналима и збирне мере, без пондерисања и са пондерисањем, дате су у Таб. 6.3, Таб. 6.4 и Таб. 6.5 (перформансе мере са најбољим резултатима су обележене масним словима). Додатно су резултати анализе из табела представљени графички на Сл. 6.13 и Сл. 6.14 кроз вредности линеарне корелације и корелације рангова. Код мера процене из Таб. 6.4 визуелни систем човека укључен је у разматрање два пута – први пут кроз нелинеарне сигмоидне функције објективне мере а други пут кроз мапу значаја визуелних информација. Слична ситуација је и код мера из Таб. 6.5 – визуелни систем човека је обухваћен нелинеарним сигмоидним функцијама и локалним значајем изведеним из амплитуде просторног/временског градијента и кроз мапу значаја визуелних информација. Из Таб. 6.4 и Таб. 6.5 се уочавају идентичне перформансе за Q_c и Q_c^{AB} мере јер се у оба случаја користе нелинеарне сигмоидне функције а мапа значаја локалних информација о очувању боје је униформна.

Таб. 6.3 Перформансе мера репрезентације изворних информација у секвенцама са деградацијом UoM сега.

Мера процене	LCC	SROCC	MAE	RMSE	OR [%]	
S^{AB}	S^{AB}	0.9199	0.9249	0.2659	0.3597	66.6667
	WS^{AB}	0.9274	0.9272	0.2331	0.3431	61.6667
	$WrandS^{AB}$	0.9025	0.8974	0.2966	0.3951	73.3333
T^{AB}	T^{AB}	0.1189	0.0069	0.7750	0.9109	90
	WT^{AB}	0.5533	0.5396	0.6184	0.7642	81.6667
	$WrandT^{AB}$	0.1083	-0.0248	0.7751	0.9120	90
C^{AB}	C^{AB}	0.8091	0.8023	0.4340	0.5392	80
	WC^{AB}	0.7632	0.7478	0.4927	0.5928	81.6667
	$WrandC^{AB}$	0.7988	0.7831	0.4398	0.5520	75
VQ	VQ	0.9078	0.8991	0.2907	0.3847	73.3333
	WVQ	0.9194	0.9183	0.2524	0.3607	60
	$WrandVQ$	0.8902	0.8833	0.3168	0.4179	76.6667

$$VQ = 0.8S^{AB} + 0.15T^{AB} + 0.05C^{AB}$$

6. Процена квалитета видеа у условима редукованог референцирања

Таб. 6.4 Перформансе мера репрезентације изворних информација након примене нелинеарних сигмоидних функција у секвенцама са деградацијом UoM сега.

Мера процене		LCC	SROCC	MAE	RMSE	OR [%]
Q _s	Q _s	0.9572	0.9590	0.2035	0.2654	65
	WQ _s	0.9480	0.9491	0.1910	0.2921	53.3333
	WrandQ _s	0.9502	0.9537	0.2167	0.2858	70
Q _t	Q _t	0.4223	0.3895	0.7191	0.8316	93.3333
	WQ _t	0.6931	0.6756	0.5546	0.6613	86.6667
	WrandQ _t	0.4100	0.3776	0.7294	0.8368	95
Q _c	Q _c	0.8124	0.8032	0.4320	0.5350	78.3333
	WQ _c	0.7639	0.7506	0.4923	0.5920	78.3333
	WrandQ _c	0.8020	0.7864	0.4380	0.5480	76.6667
VQ [*]	VQ [*]	0.9612	0.9630	0.1935	0.2532	68.3333
	WVQ [*]	0.9503	0.9461	0.1897	0.2856	53.3333
	WrandVQ [*]	0.9549	0.9573	0.2057	0.2723	71.6667

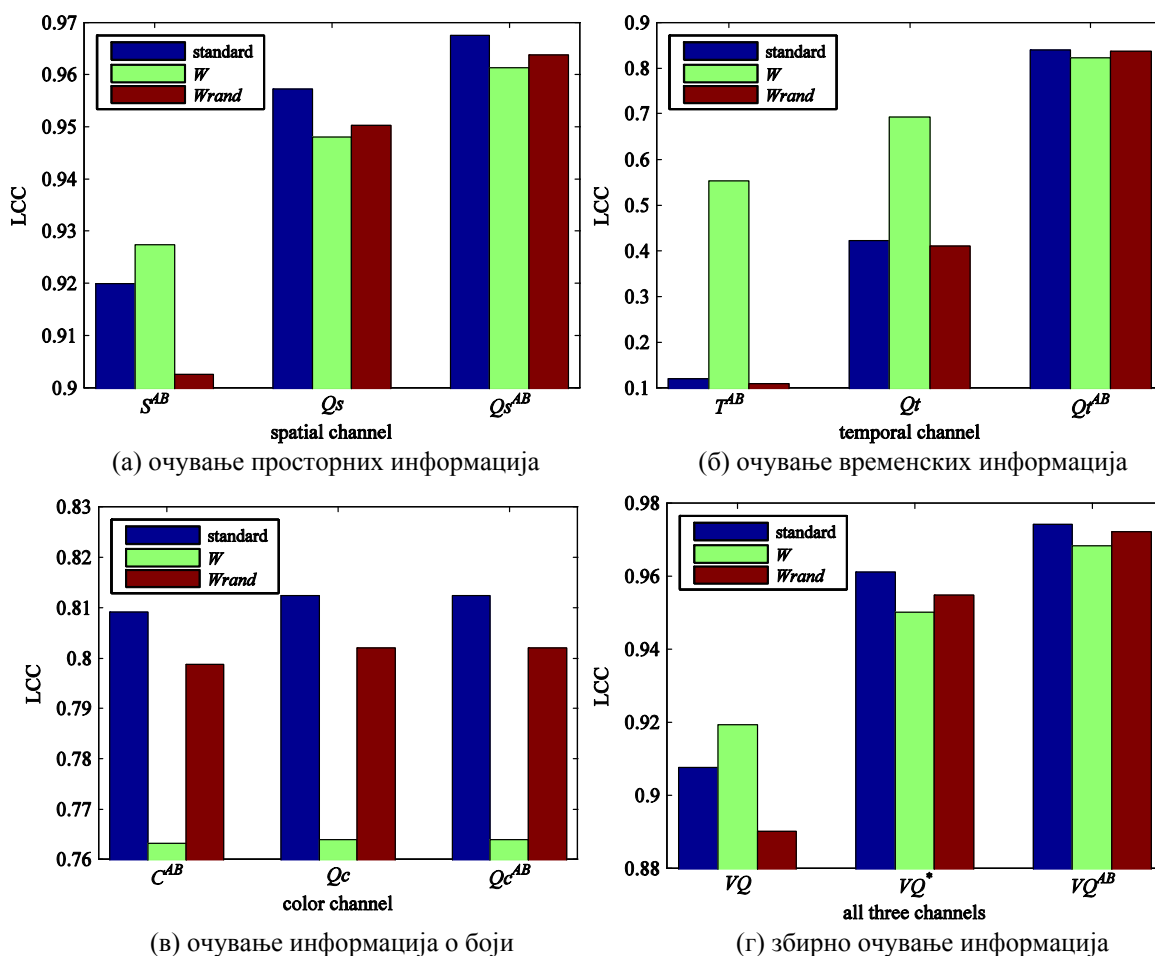
$$VQ^* = 0.8Q_s + 0.15Q_t + 0.05Q_c$$

Таб. 6.5 Перформансе мера очувања изворних информација након примене нелинеарних сигмоидних функција и увођења значаја изведеног из амплитуде градијента у секвенцама са деградацијом UoM сега.

Мера процене		LCC	SROCC	MAE	RMSE	OR [%]
Q _s ^{AB}	Q _s ^{AB}	0.9675	0.9712	0.1775	0.2321	61.6667
	WQ _s ^{AB}	0.9613	0.9631	0.1777	0.2528	50
	WrandQ _s ^{AB}	0.9638	0.9689	0.1876	0.2447	63.3333
Q _t ^{AB}	Q _t ^{AB}	0.8398	0.8309	0.4157	0.4981	85
	WQ _t ^{AB}	0.8222	0.8112	0.4313	0.5222	83.3333
	WrandQ _t ^{AB}	0.8361	0.8228	0.4259	0.5032	85
Q _c ^{AB}	Q _c ^{AB}	0.8124	0.8032	0.4320	0.5350	78.3333
	WQ _c ^{AB}	0.7639	0.7506	0.4923	0.5920	78.3333
	WrandQ _c ^{AB}	0.8020	0.7864	0.4380	0.5480	76.6667
VQ ^{AB}	VQ ^{AB}	0.9743	0.9755	0.1574	0.2068	48.3333
	WVQ ^{AB}	0.9684	0.9631	0.1689	0.2290	58.3333
	WrandVQ ^{AB}	0.9722	0.9762	0.1613	0.2148	53.3333

$$VQ^{AB} = 0.8Q_s^{AB} + 0.15Q_t^{AB} + 0.05Q_c^{AB}$$

Уколико се посматрају перформансе мера које говоре о очувању просторних информација може се закључити да се увођењем сигмоидне функције и мапе значаја w_s значајно поправља слагање субјективних и објективних скорова (коэффициент корелације се повећава за око 4%). Тиме се може закључити да су параметри сигмоидне функције и значај пиксела адекватно одабрани, јер су потврђени субјективним експериментима. Коришћењем малог дела изворног сигнала (кроз реалне или случајне фиксације) слагање субјективних и објективних скорова је практично исто.

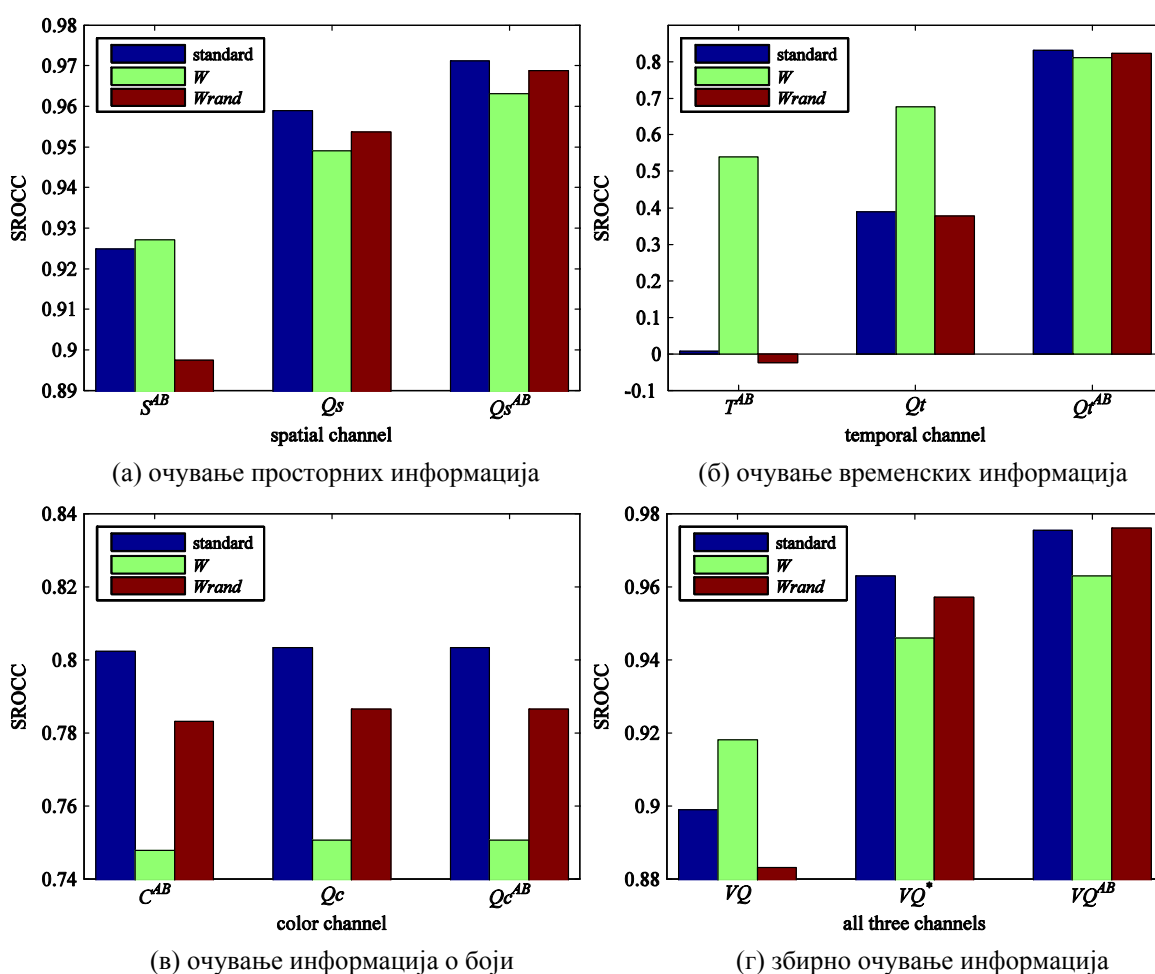


Сл. 6.13 Перформансе објективних мера процене квалитета (коэффициент линеарне корелације LCC) кроз: (а) очување просторних информација, (б) очување временских информација, (в) очување информација о боји и (г) очување информација у сва три канала.

Најзанимљивији закључци су везани за очување временских информација. Како је изворно репрезентација изворних временских информација дефинисана у (64) тако да је смер очувања временског градијента битан, овде је узето да ако постоји промена знака временског градијента, T^{AB} , очување информација је једнако 0. Резултати добијени у овом делу рада говоре да то можда и није најбоље решење за репрезентацију временских информација. Може се уочити да се увођењем сигмоидне функције и мапе значаја пиксела w_t значајно поправља слагање субјективних и објективних скорова, тј. од нечега што није корелисано ($LCC=0.11$), долази корелације од преко 80%, па се може закључити да су параметри нелинеарне сигмоидне функције и мапа значаја адекватно одабрани. Коришћењем реалних фиксација такође се значајно поправља степен слагања мере која говори о

6. Процена kvaliteta videa у условима редукованог референцирања

репрезентацији изворних информација, T^{AB} (упоредити вредности корелације за T^{AB} и WT^{AB}). Овакав резултат говори да су посматрачи гледали тамо где мера добро ради (региони у којима постоји кретање на сцени), тј. на локалном нивоу правилно евалуира очување изворних информација. Увођењем случајних фиксација перформансе мера остају на нивоу перформанси изворних објективних мера, T^{AB} , Q_t и Q_t^{AB} . Овакви резултати потврђују закључак од раније да добитак који се очекује од увођења мапа значаја изведених из реалних експеримената значајно зависи од избора објективне мере.



Сл. 6.14 Перформансе објективних мера процене квалитета (корелација рангова SROCC) кроз: (а) очување просторних информација, (б) очување временских информација, (в) очување информација о боји и (г) очување информација у сва три канала.

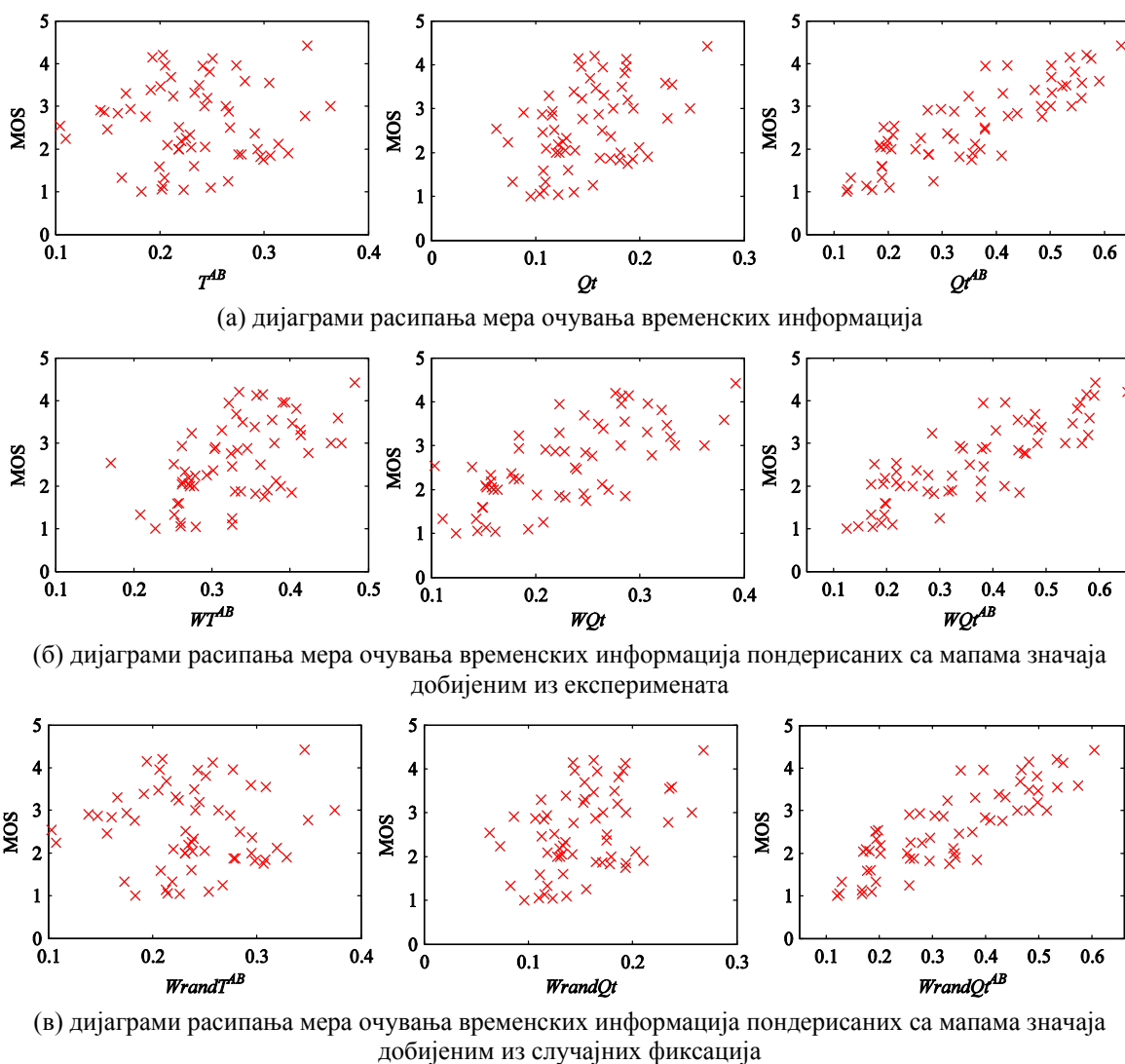
Увођење сигмоидне функције у очување информација о боји не поправља перформансе. У овом каналу сви пиксели на слици имају исти значај па је $Q_c = Q_c^{AB}$.

6. Процена квалитета видеа у условима редукованог референцирања

Коришћењем дела изворног сигнала добијеног из реалних или случајних фиксација опада степен слагања субјективних и објективних скорова.

Збирно очување изворних информација које се добија кроз пондерисање очувања информација из појединачних канала у великој мери прати тренд који постоји за просторни канал (а који у збирној мери има и највећи утицај – 80%).

С обзиром на запажања добијена квантитативном анализом очувања временских информација, на Сл. 6.15 приказани су дијаграми расипања мера очувања временских информација (без пондерисања и са пондерисањем).



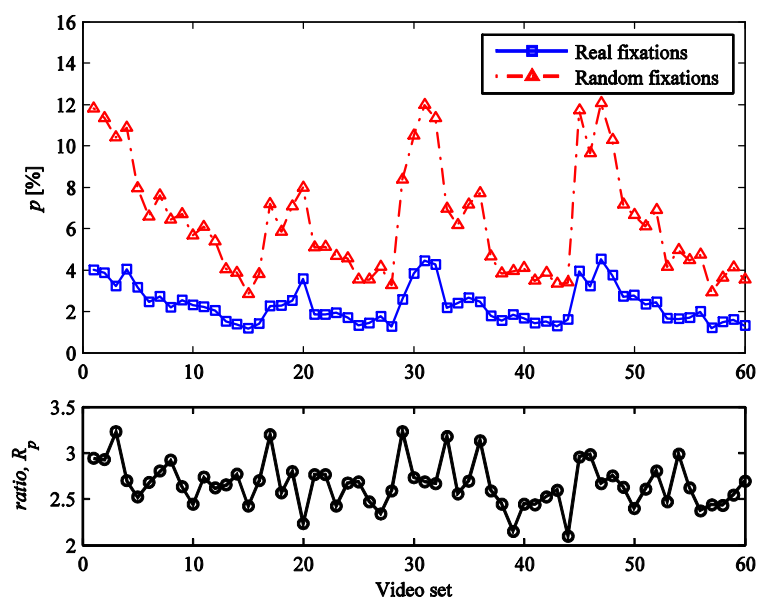
Сл. 6.15 Дијаграми расипања субјективних и објективних скорова – мера очувања временских информација.

Ако се посматрају дијаграми расипања из првог реда са Сл. 6.15, уочава се како се од изотропног облака субјективно објективних скорова кроз увођење нелинеарне сигмоидне функције и мапе значаја пиксела w_i долази до прилично уређеног низа објективних скорова који прате тренд субјективних импресија квалитета. Са дијаграма расипања из прве колоне са Сл. 6.15 уочава се боље груписање објективних процена квалитета уколико се за пондерисање локалних вредности очувања временских информација T^{AB} користе реалне фиксације (други ред прве колоне). Коришћењем случајних фиксација и даље постоји значајно расипање у домену субјективно-објективних импресија квалитета. Са дијаграма расипања из последње колоне са Сл. 6.15 може се закључити да би на мери процени квалитета очувања временских информација потребно и даље радити. Корелација субјективних и објективних скорова је око 0.83, па је због тога и процентуално учешће у укупној процени квалитета свега 15%. Сличан закључак важи и за меру очувања информација о боји где је корелација субјективних и објективних скорова нешто нижа (због чега је и учешће у збирној процени квалитета 5%).

Увођењем реалних и случајних фиксација на претходно описани начин у процени квалитета видео секвенци користи се део изворног сигнала. Како су реалне фиксације груписане на малом делу кадра а случајне фиксације разбацане по комплетном кадру, анализирано је процентуално учешће потребних изворних информација у RR евалуацији квалитета за сваки од 60 скупова UoM базе (Сл. 6.16). Осим процентуалног дела изворног сигнала који се користи у процени квалитета на један односно други начин (помоћу случајних или стварних фиксација), дат је и њихов однос, R_p ($R_p = p_{random}/p_{real}$).

Са Сл. 6.16 се види да се количник R_p креће око вредности 3. Имајући у виду да не постоје значајне статистичке разлике између перформанси добијених увођењем реалних и случајних фиксација у објективну евалуацију квалитета (по све три анализирани мере – PSNR, SSIM и VQ^{AB} (Таб. 6.2)), може се закључити да је за редуковано референцирање увођењем случајних фиксација потребна већа количина изворних информација како би перформансе биле у нивоу перформанси редукованог референцирања уз коришћење стварних фиксација (у овом случају три пута више изворних информација). Додатно, овакав резултат потврђује да су подаци кретања

зенице ока током евалуације квалитета валидни. Процентуално учешће дела изворног сигнала у процени квалитета увођењем реалних фиксације се креће од 1.17% до 4.53%, док се формирањем мапа значаја изведених из случајних фиксација користи од 2.84% до 12.07% изворног сигнала. Овакав динамички опсег је последица коришћења изворних секвенци са различитим димензијама кадрова и коришћења 2D експоненцијалног филтра чије су димензије 11x11 пиксела за све резолуције анализираних секвенци. Додатно на процентуално учешће утиче и број реалних фиксација палих на посматрану видео секвенцу током евалуације квалитета. Како се за формирање мапа значаја користи експоненцијални филтар за који значај пиксела брзо опада са удаљавањем од централног пиксела, додатним смањењем димензија филтра не би требало значајно да се утиче на објективну евалуацију квалитета а додатно се редукује количина изворних информација у евалуацији квалитета са редукованим референцирањем.



Сл. 6.16 Процентуално учешће изворних информација коришћених у процени квалитета са редукованим референцирањем UoM сета.

6.7. Закључак и правци даљег рада

Резултати кратке студије о објективним мерама са референцирањем на део изворног сигнала подстичу на даљи рад и сугеришу да се правилним избором

региона слике могу добити веома корисне мере са редукованим референцирањем. У спроведеним експериментима коришћено је око 4% информација изворног сигнала а за практичне мере редукованог референцирања потребно је додатно смањити количину пренесених изворних информација. Коначна количина информација изворног сигнала коју је потребно пренети зависи и од низа додатних фактора, укључујући и компресију (без губитака или уз минималне губитке) података пре слања, тако да би наведени проценат пао, али је опет пар процената велико када се узме у обзир да ће компримована тест секвенца бити неколико десетина пута мања од оригинала (нпр. 1.5% од оригинала).

Потребна је и темељна анализа правила селекције региона од значаја као и избор кадрова који ће се користити у евалуацији квалитета. Додатна истраживања се могу усмерити и ка једноставнијим моделима поређења података (мерама губитка информација) како би се додатно смањиле информације по пикселу које се преносе на страну пријемника.

Како би се у потпуности разумели механизми визулног значаја у процени квалитета слика/видеа потребно је спровести додатне студије. Раздвајање снимљених позиција од интереса на фиксације (посматрач је фокусиран на део слике) и сакаде (пут између две фиксације) може имати утицај на анализу сигнала на вишим нивоима. Идентификација стварних фиксација је згодан начин да се минимизира комплексност података о кретањима зенице ока (и количина података) а да се са друге стране оствари циљ истраживања – разумевање когнитивних и визуелних процеса. Избацивање сакада из разматрања у евалуацији квалитета је разумно пошто се током трајања сакаде процесира (или уопште не процесира) јако мало визуелних информација [284], а и актуелни пут током сакаде је неважан за објективну процену квалитета.

У анализи је коришћено свега пет сцена евалуираних од стране четири посматрача. Самим тим како би се потврдили добијени резултати потребно је новом анализом обухватити већи број сцена са већим бројем субјеката.

Интеграција субјективног значаја и мапа дисторзије коришћена у раду је једноставна – пондерисање локалних скорова квалитета и њихова интеграција. За

очекивати је да се напредним техникама сједињавања може остварити добитак у процени квалитета.

Приказани резултати иду у прилог тврдњи да се увођењем мапа значаја изведених из реалних фиксација могу побољшати резултати објективних мера. Добитак (или губитак) зависи од објективне мере са којом се врши интеграција.

7. Закључак

У овом поглављу дисертације се сумирају резултати истраживања о процени квалитета слике и видеа кроз очување информација о градијенту. Биће истакнуте најбитније појединости теоријских и експерименталних истраживања спроведених током израде дисертације, као и закључци који се могу извести из добијених резултата. Резултати представљени у дисертацији проистичу из десетогодишњих истраживања из области објективне процене квалитета слике и видеа.

Процена квалитета слике и видеа представљена у дисертацији дата је кроз развој алгоритама/мера за објективну процену квалитета слике/видеа и кроз субјективну процену квалитета видео секвенци са малим битским брзинама. Широки спектар истраживања овог типа представљен је кроз преглед литературе у другом поглављу дисертације. Анализирани су најбитнији аспекти процене квалитета слике и видеа. Описан је значај који за развој алгоритама субјективне и објективне евалуације квалитета има VQEG група, а дати су и текући пројекти ове групе. Преглед најбитнијих мера за објективну евалуацију квалитета слике и видеа дат је по неколико аспеката: једноставне мере процене квалитета, хибридне технике процене квалитета, процена квалитета коришћењем колор модела, примена модела визуелног система човека у евалуацији квалитета, увођење мапа визуелног значаја у процени квалитета, очување градијента у процени квалитета, процена квалитета са делимичним референцирањем и процена квалитета без референцирања на изворни сигнал. Коначно, указано је на потребу за поузданом и ефикасном евалуацијом видео секвенци са малим битским брзинама, при чему је дат преглед релативно малог броја радова из ове области.

Предмет истраживања представљеног у овој дисертацији јесу алгоритми за потпуну и делимично референцирану објективну процену квалитета сигнала слике и видеа, модификованих компресијом, губицима у преносу и другом обрадом. Развијене су робусне, рачунарски ефикасне објективне мере за аутоматску процену квалитета слике и видеа. Мере дају нумерички скор квалитета који је у доброј

корелацији са просечном субјективном проценом квалитета сигнала различитих типова и широког опсега деградације. У основи мера је поређење локалних параметара градијентних поља изворног и тест сигнала (слике/видеа).

У трећем делу дисертације предложена је објективна мера за процену квалитета слика. Разлике параметара градијента (амплитуда и оријентација) пресликане су у локалне процене квалитета коришћењем биолошки инспирисаних модела. Интеграцијом (усредњавањем) локалних процена квалитета изводи се глобални скор квалитета који је у опсегу од нула (лош квалитет) до један (идеалан квалитет). Перформансе предложене мере представљене су на четири јавно доступне базе слика (LIVE, CSIQ, Тоуама, IVC) са доступним субјективним скоровима. Перформансе предложене мере су анализирани на нивоу сваке од база, као и на бројним подскуповима слика унутар база. Предложена мера се по перформансама налази међу прве три од седам анализираних објективних мера процене. Степен слагања објективних процена добијених по предложеној мери и субјективних скорова квалитета (мерен кроз линеарну корелацију) зависи од базе слика са којом се ради и креће се на глобалном плану (нивоу комплетне базе) од 85% (CSIQ база слика) до 94% (LIVE база слика). Овако значајне разлике у перформансама су последица различитих типова деградација присутних у оквиру база, али је током анализе уочено да перформансе објективних мера значајно зависе од начина припреме тест слика и од коришћеног динамичког опсега квалитета. Разлике у перформансама на различитим базама постоје и за остале анализирани објективне мере.

У четвртом поглављу дисертације су на бројним типовима и нивоима дисторзије разматрана три аспекта процене квалитета. Кроз први аспект је спроведена исцрпна анализа процеса преноса слике кроз анализу заједничких, изгубљених и убачених информација. Предложена објективна мера процене квалитета одсликава степен заједничких информација изворне и тест слике. Експерименталним резултатима је показано да се евалуација изгубљених или убачених информација може користити у процени квалитета слика са специфичним типом дисторзије, али да то није довољно за меру која би се користила на глобалном плану. Овакав закључак потврђује претпоставке да визуелни систем човека користи

сложене механизме у процени квалитета слике. Једно од отворених питања јесте на који начин комбиновати добијене нумеричке вредности које представљају степен заједничких, изгубљених и убачених информација. Коришћење убачених информација се може користити за поуздану процену квалитета слика са адитивним шумом, док се изгубљене информације могу користити за поуздану процену квалитета слика са замрљањем и компресијом.

Кроз други аспект процене квалитета се врши побољшање предложене објективне мере. Увођење мапа субјективног значаја и њихова интеграција са локалним скоровима објективног квалитета довела је до побољшања перформанси од 0.5% до 2% на доступним базама слика. Међутим, утицај предложених мапа субјективног значаја на перформансе мере у оквиру подскупова слика из различитих база није увек доводио до побољшања перформанси.

Кроз трећи аспект процене квалитета потврђена је хипотеза да су у субјективној евалуацији квалитета најбитнији региони са највећим степеном деградације. Коришћењем најнижих скорова квалитета (квалитетом вођену интеграцију) и адитивног модела сједињавања информација уведена је алтернатива стандардном мултипликативном приступу који се користи у процени квалитета. Код већине анализираних објективних мера за поуздану процену квалитета довољно је користити од 2% до 6% најнижих локалних скорова квалитета. Од тренда већине објективних мера одступа предложена мера очувања оријентација градијента (највероватније због утицаја шума), где је потребно користити значајно већи део локалних скорова квалитета (више од 70%).

Како су приликом коришћења биолошки инспирисаних модела преузети параметри нелинеарних сигмоидних функција коришћени за потребе процене квалитета сједињавања слика, у даљем раду могу се тражити оптимални параметри за потребе процене квалитета специфичних типова дисторзија слика или оптимални параметри на глобалном плану. Иако оптимизација параметара сигмоидних функција није рађена, усвојене вредности параметара дају поуздане процене квалитета.

Одвојене анализе примене очувања амплитуда градијента (које одговарају контрасту на слици) и оријентација градијента (које одговарају структурама на

слици) у процени квалитета слике, показале су да је предност на страни очувања амплитуда градијента. Меру очувања оријентације градијента могуће је даље надограђивати. Тренутно се оријентација градијента прорачунава унутар региона 3x3 пиксела коришћењем Собеловог филтра. Почетна истраживања указују да поузданост мере очувања оријентација зависи од димензија региона унутар кога се прорачунава. Такође, оријентације градијента могуће је поредити на позицијама које припадају ивицама слике, тј. на позицијама које носе информацију о структури на слици. На један и други начин би се делимично елиминисао утицај шума на одређивање оријентације градијента.

Предложена објективна мера процене квалитета слика је проширена за процену квалитета видеа и презентована у петом делу дисертације. Осим очувања амплитуда и оријентација градијента на нивоу кадрова (просторном нивоу), посматра се очување амплитуда временског градијента и очување информација о боји. За очување информација о боји дефинисани су посебни модели за мерење перцептуалних разлика, док су за очување временског градијента преузети модели коришћени за потребе процене квалитета сједињавања слика. Оптимизацијом је утврђено да поузданој процени квалитета највише доприноси предложена мера очувања просторних информација (80%), након чега је битно очување временских информација (15%), а најмањи значај (5%) носи очување информација о боји. У овом делу дисертације описани су и субјективни тестови спроведени у сарадњи са Универзитетом у Манчестеру. Субјективним тестовима извршена је евалуација више од 300 релевантних секвенци добијених при компресији изворних секвенци за мале битске брзине. Перформансе предложене објективне мере процене квалитета видеа анализирани су на седам база видео секвенци са доступним субјективним импресијама квалитета. При томе се предложена мера показала веома погодном за процену квалитета секвенци са малим битским брзинама – до 576 kbps. Од анализираних објективних мера процене квалитета, предложена мера је дала најбоље резултате на четири базе. Перформансе предложене мере (али и других објективних мера) и у овом случају значајно зависе од базе секвенци. Степен слагања субјективних и објективних скорова (кофицијент корелације) је најнижи на LIVE Video бази секвенци (63%), а највиши на UoM и STrial базама (преко 97%). Анализа

по класама секвенци је показала да предложеној мери више погодују секвенце са компресијом од секвенци са присутним пакетским губицима. Предложено је да се у одређивању крајње вредности квалитета користи 20% кадрова са најнижим скоровима квалитета. Ова вредност представља компромис између ситуација када у секвенцама постоје краткотрајне уочљиве сметње настале пакетским губицима и када се поуздана процена квалитета добија коришћењем веома малог броја кадрова са најнижим скоровима квалитета и ситуација у којима су деградације присутне у свим кадровима и када је за поуздану процену потребно разматрати очување информација у скоро свим кадровима.

Резултати представљени у шестом поглављу рада представљају почетак истраживања у области процене квалитета видеа са редукованим референцирањем. Како је током спровођења једног од субјективних тестова на Универзитету у Манчестеру вршено снимање кретања зенице ока посматрача, снимци су преведени у координатни систем видеа и искоришћени за пондерисање објективних скорова квалитета и процену квалитета видеа са редукованим референцирањем. Показано је да се референцирањем на мали део изворног сигнала (око 4% изворног некомпримованог сигнала) може добити поуздана евалуација квалитета. Како су за редукацију изворних информација коришћени реални снимци, предмет даљих истраживање може да буде одређивање субјективно битних региона из изворног сигнала. Осим реалних снимака кретања зенице ока у евалуацију квалитета убачене су и случајне фиксације. Показано је да се са три пута већом количином изворних информација изведених из случајних фиксација могу достићи перформансе објективне мере која користи реалне фиксације. Снимци кретања зенице ока посматрача се у даљем раду могу искористити за анализу понашања визуелног система човека када гледа оригинале и њихове модификације, анализу стратегије оцењивања током времена (учење током времена) и сл. Дијаграми расипања субјективних и објективних скорова добијених коришћењем мере очувања временских информација, а представљени су у овом делу дисертације, указују да је на мери потребно даље радити. Наиме, постоји значајно расипање у простору субјективно-објективних скорова квалитета, због чега је значај очувања временских информација у укупном скору квалитета свега 15%.

Добијени резултати указују да је у процени квалитета слике/видеа и извођењу битних закључака веома тешко извести генералне (универзалне) закључке због бројних аспеката као што су: тест узорак, типови деградација, динамички опсег квалитета, објективна мера и њене појединачне компоненте, параметри објективних мера, субјективни тестови, итд. Тако је на неким базама слика битније поређење контраста док је на другим битније поређење структуре. Негде је у евалуацији квалитета пожељно узети мали број кадрова са најнижим скоровима квалитета, а негде посматрати све скорове квалитета кадрова. Увођење мапа субјективног значаја негде доноси значајан добитак а у неким ситуацијама може довести и до погоршања перформанси. Све ово заједно говори да је објективна евалуација квалитета веома комплексан проблем, али га је могуће решити и доћи до високих перформанси коришћењем приступа који су предложени за евалуацију квалитета слика и видеа.

8. Литература

- [1] A.C. Bovik, Automatic prediction of perceptual image and video quality, Proceedings of the IEEE, Vol. 101, No. 9, pp. 2008-2024, 2013.
- [2] A.C. Bovik, Perceptual video processing: Seeing the future [Point of view], Proceedings of the IEEE, Vol. 98, No. 11, pp. 1799-1803, 2010.
- [3] Z. Wang, A.C. Bovik, L. Lu, Why is image quality assessment so difficult?, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Proceedings of papers, Vol. IV, pp. 3313-3316, 2002.
- [4] S. Winkler, Video quality measurement standards – Current status and trends, 7th International Conference on Information, Communications and Signal Processing (ICICS), Proceedings of papers, pp. 1-5, 2009.
- [5] ITU TUTORIAL, Objective perceptual assessment of video quality: Full reference television, ITU-T Telecommunication Standardization Bureau, 2004.
- [6] Z. Wang, A.C. Bovik, Reduced- and no-reference image quality assessment, IEEE Signal Processing Magazine, Vol. 28, No. 6, pp. 29-40, 2011.
- [7] I. Avcibas, N. Memon, B. Sankur, Steganalysis using image quality metrics, IEEE Transactions on Image Processing, Vol. 12, No. 2, pp. 221-229, 2003.
- [8] F. Zhang, W. Liu, W. Lin, K.N. Ngan, Spread spectrum image watermarking based on perceptual quality metric, IEEE Transactions on Image Processing, Vol. 20, No. 11, pp. 3207-3218, 2011.
- [9] G. Piella, H. Heijmans, A new quality metric for image fusion, 11th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, Vol. III, pp. 173-176, 2003.
- [10] N. Cvejic, A. Loza, D. Bull, N. Canagarajah, A similarity metric for assessment of image fusion algorithms, International Journal of Signal Processing, Vol. 2, No. 3, pp. 178-182, 2005.
- [11] A. Rehman, Z. Wang, SSIM-based non-local means image denoising, 18th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 217-220, 2011.
- [12] K. Zeng, Z. Wang, Perceptual evaluation of image denoising algorithms, 47th Asilomar Conference on Signals, Systems and Computers, Proceedings of papers, pp. 1351-1355, 2013.
- [13] S.-D. Chen, A new image quality measure for assessment of histogram equalization-based contrast enhancement techniques, Digital Signal Processing, Vol. 22, No. 4, pp. 640-647, 2012.
- [14] H. Yeganeh, M. Rostami, Z. Wang, Objective quality assessment for image super-resolution: A natural scene statistics approach, 19th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 1481-1484, 2012.
- [15] A.M. Demirtas, A.R. Reibman, H. Jafarkhani, Image quality estimation for different spatial resolutions, 20th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 379-382, 2013.

- [16] H. Yeganeh, Z. Wang, Objective quality assessment of tone-mapped images, *IEEE Transactions on Image Processing*, Vol. 22, No. 2, pp. 657-667, 2013.
- [17] Y. Fang, K. Zeng, Z. Wang, W. Lin, Z. Fang, C.-W. Lin, Objective quality assessment for image retargeting based on structural similarity, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, Vol. 4, No. 1, pp. 95-105, 2014.
- [18] F. Zhang, L. Ma, S. Li, K.N. Ngan, Practical image quality metric applied to image coding, *IEEE Transactions on Multimedia*, Vol. 13, No. 14, pp. 615-624, 2011.
- [19] S. Wang, A. Rehman, Z. Wang, S. Ma, W. Gao, Perceptual video coding based on SSIM-inspired divisive normalization, *IEEE Transactions on Image Processing*, Vol. 22, No. 4, pp. 1418-1429, 2013.
- [20] E. Christophe, D. Léger, C. Mailhes, Quality criteria benchmark for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 9, pp. 2103-2114, 2005.
- [21] L. Snidaro, G.L. Foresti, A multi-camera approach to sensor evaluation in video surveillance, 15th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, Vol. I, pp. 1101-1104, 2005.
- [22] H. Chang, J. Zhang, New metrics for clutter affecting human target acquisition, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 42, No. 1, pp. 361-368, 2006.
- [23] A. Łoza, L. Mihaylova, D. Bull, N. Canagarajah, Structural similarity-based object tracking in multimodality surveillance videos, *Machine Vision and Applications*, Vol. 20, No. 2, pp. 71-83, 2009.
- [24] A. Rehman, Y. Gao, J. Wang, Z. Wang, Image classification based on complex wavelet structural similarity, *Signal Processing: Image Communication*, Vol. 28, No. 8, pp. 984-992, 2013.
- [25] L. Atzori, A. Floris, G. Ginesu, D.D. Giusto, Quality perception when streaming video on tablet devices, *Journal of Visual Communication and Image Representation*, Vol. 25, No. 3, pp. 586-595, 2014.
- [26] S. Winkler, Analysis of public image and video databases for quality assessment, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6, No. 6, pp. 616-625, 2012.
- [27] ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union, Geneva, Switzerland, 2002.
- [28] ITU-T Recommendation P.910, Subjective video quality assessment methods for multimedia applications, International Telecommunication Union, Geneva, Switzerland, 2008.
- [29] U. Engelke, A. Maeder, H.-J. Zepernick, Human observer confidence in image quality assessment, *Signal Processing: Image Communication*, Vol. 27, No. 9, pp. 935-947, 2012.
- [30] ITU-T Recommendation J.144, Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference, International Telecommunication Union, Geneva, Switzerland, 2004.

- [31] ITU-R Recommendation BT. 1683, Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference, International Telecommunication Union, Geneva, Switzerland, 2004.
- [32] K. Brunnstrom, G. Cermak, D. Hands, M. Pinson, F. Speranza, A. Webster, Final report from the Video Quality Experts Group on the validation of objective models of multimedia quality assessment, Phase I, VQEG, Available at: <http://www.vqeg.org>, 2008.
- [33] ITU-T Recommendation J.247, Objective perceptual multimedia video quality measurement in the presence of a full reference, International Telecommunication Union, Geneva, Switzerland, 2008.
- [34] ITU-R Recommendation BT.1866, Objective perceptual visual quality measurement techniques for broadcasting applications using low definition television in the presence of a full reference signal, International Telecommunication Union, Geneva, Switzerland, 2010.
- [35] ITU-T Recommendation J.246, Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference, International Telecommunication Union, Geneva, Switzerland, 2008.
- [36] ITU-R Recommendation BT.1867, Objective perceptual visual quality measurement techniques for broadcasting applications using low definition television in the presence of a reduced bandwidth reference, International Telecommunication Union, Geneva, Switzerland, 2010.
- [37] ITU-T Recommendation J.249, Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference, International Telecommunication Union, Geneva, Switzerland, 2010.
- [38] ITU-T Recommendation J.340, Reference algorithm for computing peak signal to noise ratio (PSNR) of a processed video sequence with constant spatial shifts and a constant delay, International Telecommunication Union, Geneva, Switzerland, 2010.
- [39] ITU-T Recommendation J.341, Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference, International Telecommunication Union, Geneva, Switzerland, 2011.
- [40] ITU-T Recommendation J.342, Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal, International Telecommunication Union, Geneva, Switzerland, 2011.
- [41] ITU-T Recommendation J.244, Calibration methods for constant misalignment of spatial and temporal domains with constant gain and offset, International Telecommunication Union, Geneva, Switzerland, 2008.
- [42] R.C. Streijl, S. Winkler, D.S. Hands, Perceptual quality measurement – Towards a more efficient process for validating objective models [Standards in a nutshell], IEEE Signal Processing Magazine, Vol. 27, No. 4, pp. 136-140, 2010.
- [43] ATIS Standard 0800008, QoS metrics for linear broadcast IPTV, Alliance for Telecommunications Industry Solutions, Washington, DC, USA, 2007.

- [44] T. Rahrer, R. Fiandra, S. Wright, Triple-play services quality of experience (QoE) requirements, DSL Forum, Architecture and Transport Working Group, Technical Report, TR-126, 2006.
- [45] Video Services Forum, Recommended video over IP metrics, VSF Test and Measurements Activity Group, Technical report, 2006.
- [46] Z. Wang, A.C. Bovik, Mean squared error: Love it or leave it? A new look at signal fidelity measures, *IEEE Signal Processing Magazine*, Vol. 26, No. 1, pp. 98-117, 2009.
- [47] Z. Wang, A.C. Bovik, A universal image quality index, *IEEE Signal Processing Letters*, Vol. 9, No. 3, pp. 81-84, 2002.
- [48] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600-612, 2004.
- [49] D. Van der Weken, M. Nachtgaele, E.E. Kerre, Using similarity measures and homogeneity for the comparison of images, *Image and Vision Computing*, Vol. 22, No. 9, pp. 695-702, 2004.
- [50] A. Shnayderman, A. Gusev, A.M. Eskicioglu, An SVD-based grayscale image quality measure for local and global assessment, *IEEE Transactions on Image Processing*, Vol. 15, No. 2, pp. 422-429, 2006.
- [51] C.-Y. Wee, R. Paramesran, R. Mukundan, X. Jiang, Image quality assessment by discrete orthogonal moments, *Pattern Recognition*, Vol. 43, No. 12, pp. 4055-4068, 2010.
- [52] J. Qian, D. Wu, L. Li, D. Cheng, X. Wang, Image quality assessment based on multi-scale representation of structure, *Digital Signal Processing*, Vol. 33, pp. 125-133, 2014.
- [53] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, A.C. Bovik, Image quality assessment based on a degradation model, *IEEE Transactions on Image Processing*, Vol. 9, No. 4, pp. 636-650, 2000.
- [54] Z. Liu, R. Laganière, Phase congruence measurement for image similarity assessment, *Pattern Recognition Letters*, Vol. 28, No. 1, pp. 166-172, 2007.
- [55] L. Zhang, H. Li, SR-SIM: A fast and high performance IQA index based on spectral residual, 19th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 1473-1476, 2012.
- [56] M. Narwaria, W. Lin, I.V. McLoughlin, S. Emmanuel, L.-T. Chia, Fourier transform-based scalable image quality measure, *IEEE Transactions on Image Processing*, Vol. 21, No. 8, pp. 3364-3377, 2012.
- [57] E.C. Larson, D.M. Chandler, Most apparent distortion: Full-reference image quality assessment and the role of strategy, *Journal of Electronic Imaging*, Vol. 19, Paper No. 011006, 2010.
- [58] H.R. Sheikh, A.C. Bovik, Image information and visual quality, *IEEE Transactions on Image Processing*, Vol. 15, No. 2, pp. 430-444, 2006.
- [59] D.M. Chandler, S.S. Hemami, VSNR: A wavelet-based visual signal-to-noise ratio for natural images, *IEEE Transactions on Image Processing*, Vol. 16, No. 9, pp. 2284-2298, 2007

-
- [60] S. Li, F. Zhang, L. Ma, K.N. Ngan, Image quality assessment by separately evaluating detail losses and additive impairments, *IEEE Transactions on Multimedia*, Vol. 13, No. 5, pp. 935-949, 2011.
- [61] K.-H. Thung, R. Paramesran, C.-L. Lim, Content-based image quality metric using similarity measure of moment vectors, *Pattern Recognition*, Vol. 45, No. 6, pp. 2193-2204, 2012.
- [62] L. Zhang, D. Zhang, X. Mou, RFSIM: A feature based image quality assessment metric using Riesz transforms, *17th IEEE International Conference on Image Processing (ICIP)*, Proceedings of papers, pp. 321-324, 2010.
- [63] L. Zhang, D. Zhang, X. Mou, FSIM: A feature similarity index for image quality assessment, *IEEE Transactions on Image Processing*, Vol. 20, No. 8, pp. 2378-2386, 2011.
- [64] P. Singh, D.M. Chandler, F-MAD: A feature-based extension of the most apparent distortion algorithm for image quality assessment, *Proc. SPIE 8653, Image Quality and System Performance X*, Paper No. 86530I, 2013.
- [65] M. Narwaria, W. Lin, Objective image quality assessment based on support vector regression, *IEEE Transactions on Neural Networks*, Vol. 21, No. 3, pp. 515-519, 2010.
- [66] M. Narwaria, W. Lin, A.E. Cetin, Scalable image quality assessment with 2D mel-cepstrum and machine learning approach, *Pattern Recognition*, Vol. 45, No. 1, pp. 299-313, 2012.
- [67] C. Charrier, O. Lézoray, G. Lebrun, Machine learning to design full-reference image quality assessment algorithm, *Signal Processing: Image Communication*, Vol. 27, No. 3, pp. 209-219, 2012.
- [68] T.-J. Liu, W. Lin, C.-C. Kuo, Image quality assessment using multi-method fusion, *IEEE Transactions on Image Processing*, Vol. 22, No. 5, pp. 1793-1807, 2013.
- [69] S. Bianco, G. Ciocca, F. Marini, R. Schettini, Image quality assessment by preprocessing and full reference model combination, *Proc. SPIE 7242, IS&T/SPIE Electronic Imaging*, Paper No. 72420O, 2009.
- [70] L. Ma, S. Li, K.N. Ngan, Visual horizontal effect for image quality assessment, *IEEE Signal Processing Letters*, Vol. 17, No. 7, pp. 627-630, 2010.
- [71] K. Okarma, Quality assessment of images with multiple distortions using combined metrics, *Elektronika ir Elektrotechnika*, Vol. 20, No. 6, pp. 128-131, 2014.
- [72] A. Chetouani, A. Beghdadi, M. Deriche, A hybrid system for distortion classification and image quality evaluation, *Signal Processing: Image Communication*, Vol. 27, No. 9, pp. 948-960, 2012.
- [73] Q. Zhang, Y. Han, Y. Cai, Novel full-reference image quality assessment metric based on entropy fusion, *Optik-International Journal for Light and Electron Optics*, Vol. 124, No. 21, pp. 5149-5153, 2013.
- [74] K. Gu, G. Zhai, X. Yang, W. Zhang, A new psychovisual paradigm for image quality assessment: From differentiating distortion types to discriminating quality conditions, *Signal, Image and Video Processing (SIViP)*, Vol. 7, No. 3, pp. 423-436, 2013.
-

- [75] L. Capodiferro, G. Jacovitti, E.D. Di Claudio, Two-dimensional approach to full-reference image quality assessment based on positional structural information, *IEEE Transactions on Image Processing*, Vol. 21, No. 2, pp. 505-516, 2012.
- [76] A. Barri, A. Doms, B. Jansen, P. Schelkens, A locally adaptive system for the fusion of objective quality measures, *IEEE Transactions on Image Processing*, Vol. 23, No. 6, pp. 2446-2458, 2014.
- [77] J. Shen, Q. Li, G. Erlebacher, Hybrid no-reference natural image quality assessment of noisy, blurry, JPEG2000 and JPEG images, *IEEE Transactions on Image Processing*, Vol. 20, No. 8, pp. 2089-2098, 2011.
- [78] K. Okarma, Application of the CIE color spaces for the digital image quality assessment, *Annales UMCS Informatica AI*, Vol. 10, No. 2, pp. 69-77, 2010.
- [79] X. Zhang, D.A. Silverstein, J.E. Farrell, B.A. Wandell, Color image quality metric S-CIELAB and its application on halftone texture visibility, in *Comcon '97, Proceedings of papers*, pp. 44-48, 1997.
- [80] L. He, X. Gao, W. Lu, X. Li, D. Tao, Image quality assessment based on S-CIELAB model, *Signal, Image and Video Processing (SIVIP)*, Vol. 5, No. 3, pp. 283-290, 2011.
- [81] P. Le Callet, D. Barba, A robust quality metric for color image quality assessment, *11th IEEE International Conference on Image Processing (ICIP)*, Proceedings of papers, Vol. I, pp. 437-440, 2003.
- [82] M. Carnec, P. Le Callet, D. Barba, Objective quality assessment of color images based on a generic perceptual reduced reference, *Signal Processing: Image Communication*, Vol. 23, No. 4, pp. 239-256, 2008.
- [83] A. Toet, M.P. Lucassen, A new universal colour image fidelity metric, *Displays*, Vol. 24, No. 4, pp. 197-207, 2003.
- [84] A. Kolaman, O. Yadid-Pecht, Quaternion structural similarity: A new quality index for color images, *IEEE Transactions on Image Processing*, Vol. 21, No. 4, pp. 1526-1536, 2012.
- [85] H. Kikuchi, S. Kataoka, S. Muramatsu, H. Huttunen, Color-tone similarity of digital images, *20th IEEE International Conference on Image Processing (ICIP)*, Proceedings of papers, pp. 393-397, 2013.
- [86] S.J. Daly, Visible differences predictor: An algorithm for the assessment of image fidelity, *Proc. SPIE 1616, Human Vision, Visual Processing and Digital Display III*, pp. 2-15, 1992.
- [87] C.C. Taylor, Z. Pizlo, J.P. Allebach, C.A. Bouman, Image quality assessment with a Gabor pyramid model of the Human Visual System, *Proc. SPIE 3016, International Symposium on Electronic Imaging, Science and Technology*, pp. 58-69, 1997.
- [88] A.P. Bradley, A wavelet visible difference predictor, *IEEE Transactions on Image Processing*, Vol. 8, No. 5, pp. 717-730, 1999.
- [89] F. Zhang, D.R. Bull, Quality assessment methods for perceptual video compression, *20th IEEE International Conference on Image Processing (ICIP)*, Proceedings of papers, pp. 39-43, 2013.

-
- [90] X. Yang, W. Ling, Z. Lu, E.P. Ong, S. Yao, Just noticeable distortion model and its applications in video coding, *Signal Processing: Image Communication*, Vol. 20, No. 7, pp. 662-680, 2005.
- [91] Y. Jia, W. Lin, A.A. Kassim, Estimating just-noticeable distortion for video, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, No. 7, pp. 820-829, 2006.
- [92] H. Liu, I. Heynderickx, A simplified human vision model applied to a blocking artifact metric, *Computer Analysis of Images and Patterns (CAIP)*, LNCS 4673, Springer, pp. 334-341, 2007.
- [93] A.C. Brooks, X. Zhao, T.N. Pappas, Structural similarity quality metrics in a coding context: Exploring the space of realistic distortions, *IEEE Transactions on Image Processing*, Vol. 17, No. 8, pp. 1261-1273, 2008.
- [94] X. Fei, L. Xiao, Y. Sun, Z. Wei, Perceptual image quality assessment based on structural similarity and visual masking, *Signal Processing: Image Communication*, Vol. 27, No. 7, pp. 772-783, 2012.
- [95] R. Brémond, J.-P. Tarel, E. Dumont, N. Hautiere, Vision models for image quality assessment: One is not enough, *Journal of Electronic Imaging*, Vol. 19, No. 4, Paper No. 043004, 2010.
- [96] F. Zhang, S. Li, L. Ma, K.N. Ngan, Limitation and challenges of image quality measurement, *Proc. SPIE 7744, Visual Communications and Image Processing*, Paper No. 774402, 2010.
- [97] A. Bhat, I. Richardson, S. Kannangara, A new perceptual quality metric for compressed video based on mean squared error, *Signal Processing: Image Communication*, Vol. 25, No. 8, pp. 588-596, 2010.
- [98] Z. Wang, L. Lu, A.C. Bovik, Video quality assessment based on structural distortion measurement, *Signal Processing: Image Communication*, Vol. 19, No. 2, pp. 121-132, 2004.
- [99] A.K. Moorthy, A.C. Bovik, Efficient video quality assessment along temporal trajectories, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 11, pp. 1653-1658, 2010.
- [100] C. Vu, S. Deshpande, ViMSSIM: From image to video quality assessment, 4th Workshop on Mobile Video (MoVid), *Proceedings of papers*, pp. 1-6, 2012.
- [101] K. Zeng, Z. Wang, 3D-SSIM for video quality assessment, 19th IEEE International Conference on Image Processing (ICIP), *Proceedings of papers*, pp. 621-624, 2012.
- [102] K. Seshadrinathan, A.C. Bovik, Motion tuned spatio-temporal quality assessment of natural videos, *IEEE Transactions on Image Processing*, Vol. 19, No. 2, pp. 335-350, 2010.
- [103] P.V. Vu, C.T. Vu, D.M. Chandler, A spatiotemporal most-apparent-distortion model for video quality assessment, 18th IEEE International Conference on Image Processing (ICIP), *Proceedings of papers*, pp. 2505-2508, 2011.
- [104] S. Li, L. Ma, K.N. Ngan, Full-reference video quality assessment by decoupling detail losses and additive impairments, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 22, No. 7, pp. 1100-1112, 2012.
-

-
- [105] E.P. Ong, X. Yang, W. Lin, Z. Lu, S. Yao, X. Lin, S. Rahardja, B.C. Seng, Perceptual quality and objective quality measurements of compressed videos, *Journal of Visual Communication and Image Representation*, Vol. 17, No. 4, pp. 717-737, 2006.
- [106] J. Xia, Y. Shi, K. Teunissen, I. Heynderickx, Perceivable artifacts in compressed video and their relation to video quality, *Signal Processing: Image Communication*, Vol. 24, No. 7, pp. 548-556, 2009.
- [107] S. Rimac-Drlje, M. Vranješ, D. Žagar, Foveated mean squared error – A novel video quality metric, *Multimedia Tools and Applications*, Vol. 49, No. 3, pp. 425-445, 2010.
- [108] A.K. Moorthy, L.K. Choi, A.C. Bovik, G. De Veciana, Video quality assessment on mobile devices: Subjective, behavioral and objective studies, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6, No. 6, pp. 652-671, 2012.
- [109] A. Besson, F. De Simone, T. Ebrahimi, Objective quality metrics for video scalability, 20th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 59-63, 2013.
- [110] P. Brun, G. Hauske, T. Stockhammer, Subjective assessment of H. 264-AVC video for low-bitrate multimedia messaging services, 12th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 1145-1148, 2004.
- [111] M.A. Masry, S.S. Hemami, A metric for continuous quality evaluation of compressed video with severe distortions, *Signal processing: Image communication*, Vol. 19, No. 2, pp. 133-146, 2004.
- [112] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, M. Etoh, Cross-dimensional perceptual quality assessment for low bit-rate videos, *IEEE Transactions on Multimedia*, Vol. 10, No. 7, pp. 1316-1324, 2008.
- [113] A. Khan, L. Sun, J. Fajardo, I. Taboada, F. Liberal, E. Ifeakor, Impact of end devices on subjective video quality assessment for QCIF video sequences, 3rd IEEE International Workshop on Quality of Multimedia Experience (QoMEX), Proceedings of papers, pp. 177-182, 2011.
- [114] Q. Huynh-Thu, M. Ghanbari, Modelling of spatio-temporal interaction for video quality assessment, *Signal Processing: Image Communication*, Vol. 25, No. 7, pp. 535-546, 2010.
- [115] C. Yim, A.C. Bovik, Evaluation of temporal variation of video quality in packet loss networks, *Signal Processing: Image Communication*, Vol. 26, No. 1, pp. 24-38, 2011.
- [116] M. Narwaria, W. Lin, A. Liu, Low-complexity video quality assessment using temporal quality variations, *IEEE Transactions on Multimedia*, Vol. 14, No. 3, pp. 525-535, 2012.
- [117] A. Ninassi, O. Le Meur, P. Le Callet, D. Barba, Considering temporal variations of spatial visual distortions in video quality assessment, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 3, No. 2, pp. 253-265, 2009.
- [118] A. Rehman, Z. Wang, Perceptual experience of time-varying video quality, 5th IEEE International Workshop on Quality of Multimedia Experience (QoMEX), Proceedings of papers, pp. 218-223, 2013.
-

-
- [119] Y.-F. Ou, Y. Xue, Y. Wang, Q-STAR: A perceptual video quality model considering impact of spatial, temporal and amplitude resolutions, *IEEE Transactions on Image Processing*, Vol. 23, No. 6, pp. 2473-2486, 2014.
- [120] F. De Simone, M. Tagliasacchi, M. Naccari, S. Tubaro, T. Ebrahimi, A H.264/AVC video database for the evaluation of quality metrics, *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Proceedings of papers, pp. 2430-2433, 2010.
- [121] K. Zeng, A. Rehman, J. Wang, Z. Wang, From H.264 to HEVC: Coding gain predicted by objective video quality assessment models, *International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, 2013.
- [122] D. Grois, D. Marpe, A. Mulayoff, B. Itzhaky, O. Hadar, Performance comparison of H.265/MPEG-HEVC, VP9 and H.264/MPEG-AVC encoders, *Picture Coding Symposium (PCS)*, Proceedings of papers, pp. 8-11, 2013.
- [123] P. Hanhart, M. Rerabek, F. De Simone, T. Ebrahimi, Subjective quality evaluation of the upcoming HEVC video compression standard, *Proc. SPIE 8499, Application to Digital Image Processing XXXV*, Paper No. 84990V, 2012.
- [124] Z. Wang, X. Shang, Spatial pooling strategies for perceptual image quality assessment, *14th IEEE International Conference on Image Processing*, Proceedings of papers, pp. 2945-2948, 2006.
- [125] M. Gong, M. Pedersen, Spatial pooling for measuring color printing quality attributes, *Journal of Visual Communication and Image Representation*, Vol. 23, No. 5, pp. 685-696, 2012.
- [126] B.P. Bondzulic, V.S. Petrovic, S.T. Mitrovic, B.Z. Pavlovic, M.S. Andric, Visual attention pooling and understanding the structural similarity index in multi-scale analysis, *Optica Applicata*, Vol. 44, No. 2, pp. 267-283, 2014.
- [127] V. Petrović, V. Dimitrijević, Focused pooling for image fusion evaluation, *Information Fusion*, Vol. 22, pp. 119-126, 2015.
- [128] J. Park, K. Seshadrinathan, S. Lee, A.C. Bovik, Video quality pooling adaptive to perceptual distortion severity, *IEEE Transactions on Image Processing*, Vol. 22, No. 2, pp. 610-620, 2013.
- [129] Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment, *IEEE Transactions on Image Processing*, Vol. 20, No. 5, pp. 1185-1198, 2011.
- [130] H. Liu, I. Heynderickx, Visual attention in objective image quality assessment: Based on eye-tracking data, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 21, No. 7, pp. 971-982, 2011.
- [131] H. Liu, U. Engelke, J. Wang, P. Le Callet, I. Heynderickx, How does image content affect the added value of visual attention in objective image quality assessment?, *IEEE Signal Processing Letters*, Vol. 20, No. 4, pp. 355-358, 2013.
- [132] S. Winkler, S. Ramanathan, Overview of eye tracking datasets, *5th International Workshop on Quality of Multimedia Experience (QoMEX)*, Proceedings of papers, pp. 212-217, 2013.
-

-
- [133] V. Kadiyala, S. Pinneli, E.C. Larson, D.M. Chandler, Quantifying the perceived interest of objects in images: Effects of size, location, blur, and contrast, Proc. SPIE 6806, Human Vision and Electronic Imaging XIII, Paper No. 68060S, 2008.
- [134] G. Abdollahian, Z. Pizlo, E.J. Delp, A study on the effect of camera motion on human visual attention, 15th IEEE International Conference on Image Processing, Proceedings of papers, pp. 693-696, 2008.
- [135] O. Le Meur, A. Ninassi, P. Le Callet, D. Barba, Overt visual attention for free-viewing and quality assessment tasks: Impact of the regions of interest on a video quality metric, Signal Processing: Image Communication, Vol. 25, No. 7, pp. 547-558, 2010.
- [136] A. Mittal, A.K. Moorthy, W.S. Geisler, A.C. Bovik, Task dependence of visual attention on compressed videos: Point of gaze statistics and analysis, Proc. SPIE 7865, Human Vision and Electronic Imaging XVI, Paper No. 78650T, 2011.
- [137] K. Seshadrinathan, A.C. Bovik, Temporal hysteresis model of time varying subjective video quality, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Proceedings of papers, pp. 1153-1156, 2011.
- [138] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, 1998.
- [139] U. Rajashekar, I. Van Der Linde, A.C. Bovik, L.K. Cormack, GAFFE: A gaze-attentive fixation finding engine, IEEE Transactions on Image Processing, Vol. 17, No. 4, pp. 564-573, 2008.
- [140] M.C. Farias, W.Y. Akamine, On performance of image quality metrics enhanced with visual attention computational models, Electronics Letters, Vol. 48, No. 11, pp. 631-633, 2012.
- [141] W.Y. Akamine, M.C. Farias, Incorporating visual attention models into video quality metrics, Proc. SPIE 9016, Image Quality and System Performance XI, Paper No. 90160O, 2014.
- [142] X. Hou, L. Zhang, Saliency detection: A spectral residual approach, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Proceedings of papers, pp. 1-8, 2007.
- [143] Q. Ma, L. Zhang, Image quality assessment with visual attention, 19th International Conference on Pattern Recognition (ICPR), Proceedings of papers, pp. 1-4, 2008.
- [144] L. Ma, S. Li, K.N. Ngan, Motion trajectory based visual saliency for video quality assessment, 18th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 233-236, 2011.
- [145] H. Liu, I. Heynderickx, Towards an efficient model of visual saliency for objective image quality assessment, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Proceedings of papers, pp. 1153-1156, 2012.
- [146] X. Feng, T. Liu, D. Yang, Y. Wang, Saliency inspired full-reference quality metrics for packet-loss-impaired video, IEEE Transactions on Broadcasting, Vol. 57, No. 1, pp. 81-88, 2011.
-

- [147] L. Shen, Z. Liu, Z. Zhang, A novel H.264 rate control algorithm with consideration of visual attention, *Multimedia Tools and Applications*, Vol. 63, No. 3, pp. 709-727, 2013.
- [148] J.-S. Lee, F. De Simone, T. Ebrahimi, Efficient video coding based on audio-visual focus of attention, *Journal of Visual Communication and Image Representation*, Vol. 22, No. 8, pp. 704-711, 2011.
- [149] H. Wu, M. Yuen, A generalized block-edge impairment metric for video coding, *IEEE Signal Processing Letters*, Vol. 4, No. 11, pp. 317-320, 1997.
- [150] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, A no-reference perceptual blur metric, 10th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, Vol. III, pp. 57-60, 2002.
- [151] U. Engelke, M. Kusuma, H.-J. Zepernick, M. Caldera, Reduced-reference metric design for objective perceptual quality assessment in wireless imaging, *Signal Processing: Image Communication*, Vol. 24, No. 7, pp. 525-547, 2009.
- [152] G. Cheng, L. Cheng, Reduced reference image quality assessment based on dual derivative priors, *Electronics Letters*, Vol. 45, No. 18, pp. 937-939, 2009.
- [153] G. Cheng, J. Huang, Z. Liu, C. Lizhi, Image quality assessment using natural image statistics in gradient domain, *AEU-International Journal of Electronics and Communications*, Vol. 65, No. 5, pp. 392-397, 2011.
- [154] A. Rehman, Z. Wang, Reduced-reference image quality assessment by structural similarity estimation, *IEEE Transactions on Image Processing*, Vol. 21, No. 8, pp. 3378-3389, 2012.
- [155] L. Ma, S. Li, K.N. Ngan, Reduced-reference image quality assessment in reorganized DCT domain, *Signal Processing: Image Communication*, Vol. 28, No. 8, pp. 884-902, 2013.
- [156] L. Atzori, A. Floris, G. Ginesu, D. Giusto, Streaming video over wireless channels: Exploiting reduced-reference quality estimation at the user-side, *Signal Processing: Image Communication*, Vol. 27, No. 10, pp. 1049-1065, 2012.
- [157] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, X. Yang, A locally adaptive algorithm for measuring blocking artifacts in images and videos, *Signal Processing: Image Communication*, Vol. 19, No. 6, pp. 499-506, 2004.
- [158] G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, No-reference noticeable blockiness estimation in images, *Signal Processing: Image Communication*, Vol. 23, No. 6, pp. 417-432, 2008.
- [159] S. Lee, S.J. Park, A new image quality assessment method to detect and measure strength of blocking artifacts, *Signal Processing: Image Communication*, Vol. 27, No. 1, pp. 31-38, 2012.
- [160] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, Perceptual blur and ringing metrics: Application to JPEG2000, *Signal Processing: Image Communication*, Vol. 19, No. 2, pp. 163-172, 2004.
- [161] J. Zhang, S.H. Ong, T.M. Le, Kurtosis-based no-reference quality assessment of JPEG2000 images, *Signal Processing: Image Communication*, Vol. 26, No. 1, pp. 13-23, 2011.

- [162] G. Callico, S. Lopez, S. Gabarda, E. Gil, G. Cristobal, J. Lopez, R. Sarmiento, Anisotropic quality measurement applied to H.264 video compression, Proc. SPIE 7363, VLSI Circuits and Systems IV, Paper No. 73630I, 2009.
- [163] T. Brandão, M.P. Queluz, No-reference quality assessment of H.264/AVC encoded video, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 20, No. 11, pp. 1437-1447, 2010.
- [164] A.K. Moorthy, A.C. Bovik, A two-step framework for constructing blind image quality indices, IEEE Signal Processing Letters, Vol. 17, No. 5, pp. 513-516, 2010.
- [165] M.A. Saad, A.C. Bovik, C. Charrier, A DCT statistics-based blind image quality index, IEEE Signal Processing Letters, Vol. 17, No. 6, pp. 583-586, 2010.
- [166] L. Liu, H. Dong, H. Huang, A.C. Bovik, No-reference image quality assessment in curvelet domain, Signal Processing: Image Communication, Vol. 29, No. 4, pp. 494-505, 2014.
- [167] L. Liu, B. Liu, H. Huang, A.C. Bovik, No-reference image quality assessment based on spatial and spectral entropies, Signal Processing: Image Communication, Vol. 29, No. 8, pp. 856-863, 2014.
- [168] C. Li, Y. Ju, A.C. Bovik, X. Wu, Q. Sang, No-training, no-reference image quality index using perceptual features, Optical Engineering, Vol. 52, No. 5, Paper No. 057003, 2013.
- [169] A. Mittal, R. Soundararajan, A.C. Bovik, Making a “completely blind” image quality analyzer, IEEE Signal Processing Letters, Vol. 20, No. 3, pp. 209-212, 2013.
- [170] C. Li, A.C. Bovik, X. Wu, Blind image quality assessment using a general regression neural network, IEEE Transactions on Neural Networks, Vol. 22, No. 5, pp. 793-799, 2011.
- [171] A. Mittal, A.K. Moorthy, A.C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Transactions on Image Processing, Vol. 21, No. 12, pp. 4695-4708, 2012.
- [172] F. Yang, S. Wan, Y. Chang, H.R. Wu, A novel objective no-reference metric for digital video quality assessment, IEEE Signal Processing Letters, Vol. 12, No. 10, pp. 685-688, 2005.
- [173] M.A. Saad, A.C. Bovik, C. Charrier, Blind prediction of natural video quality, IEEE Transactions on Image Processing, Vol. 23, No. 3, pp. 1352-1365, 2014.
- [174] M.C. Farias, M. Carli, S.K. Mitra, Objective video quality metric based on data hiding, IEEE Transactions on Consumer Electronics, Vol. 51, No. 3, pp. 983-992, 2005.
- [175] D. Culibrk, M. Mirkovic, V. Zlokolica, M. Pokric, V. Crnojevic, D. Kukolj, Salient motion features for video quality assessment, IEEE Transactions on Image Processing, Vol. 20, No. 4, pp. 948-958, 2011.
- [176] X. Lin, H. Ma, L. Luo, Y. Chen, No-reference video quality assessment in the compressed domain, IEEE Transactions on Consumer Electronics, Vol. 58, No. 2, pp. 505-512, 2012.

- [177] A.A. Webster, C.T. Jones, M.H. Pinson, S.D. Voran, S. Wolf, Objective video quality assessment system based on human perception, Proc. SPIE 1913, Human Vision, Visual Processing and Digital Display IV, pp. 15-26, 1993.
- [178] P. Fränti, Blockwise distortion measure for statistical and structural errors in digital images, Signal Processing: Image Communication, Vol. 13, No. 2, pp. 89-98, 1998.
- [179] C. Xydeas, V. Petrović, Objective image fusion performance measure, Electronics Letters, Vol. 36, No. 4, pp. 308-309, 2000.
- [180] M.H. Pinson, S. Wolf, A new standardized method for objectively measuring video quality, IEEE Transactions on Broadcasting, Vol. 50, No. 3, pp. 312-322, 2004.
- [181] F. Pan, X. Lin, S. Rahardja, E. Ong, W. Lin, Using edge direction information for measuring blocking artifacts of images, Multidimensional Systems and Signal Processing, 18 (2007) 297-308.
- [182] D.-O. Kim, H.-S. Han, R.-H. Park, Gradient information-based image quality metric, IEEE Transactions on Consumer Electronics, Vol. 56, No. 2, pp. 930-936, 2010.
- [183] D.-O. Kim, R.-H. Park, Image quality assessment using the amplitude/phase quantization code, IEEE Transactions on Consumer Electronics, Vol. 56, No. 4, pp. 2756-2762, 2010.
- [184] M. Zhang, X. Mou, D. Zhang, Non-shift edge based ratio (NSER): An image quality assessment metric based on early vision features, IEEE Signal Processing Letters, Vol. 18, No. 5, pp. 315-318, 2011.
- [185] Z.A. Seghir, F. Hachouf, Edge-region information measure based on deformed and displaced pixel for image quality assessment, Signal Processing: Image Communication, Vol. 26, No. 8, pp. 534-549, 2011.
- [186] J.-F. Zhao, H.-J. Feng, Z.-H. Xu, Q. Li, An image quality assessment based on regional weight, Optik-International Journal for Light and Electron Optics, Vol. 123, No. 6, pp. 494-497, 2012.
- [187] A. Liu, W. Lin, M. Narwaria, Image quality assessment based on gradient similarity, IEEE Transactions on Image Processing, Vol. 21, No. 4, pp. 1500-1512, 2012.
- [188] J. Zhu, N. Wang, Image quality assessment by visual gradient similarity, Image IEEE Transactions on Processing, Vol. 21, No. 3, pp. 919-933, 2012.
- [189] X. Zhang, X. Feng, W. Wang, W. Xue, Edge strength similarity for image quality assessment, IEEE Signal Processing Letters, Vol. 20, No. 4, pp. 319-322, 2013.
- [190] H. Zhang, Q. Zhu, C. Fan, D. Deng, Image quality assessment based on Prewitt magnitude, AEU-International Journal of Electronics and Communications, Vol. 67, No. 9, pp. 799-803, 2013.
- [191] W. Xue, L. Zhang, X. Mou, A.C. Bovik, Gradient magnitude similarity deviation: A highly efficient perceptual image quality index, IEEE Transactions on Image Processing, Vol. 23, No. 2, pp. 684-695, 2014.
- [192] M.G. Martini, B. Villarini, F. Fiorucci, A reduced-reference perceptual image and video quality metric based on edge preservation, EURASIP Journal on Advances in Signal Processing 2012, 2012:66 1-13.

- [193] Z.P. Sazzad, R. Sultana, H. Siddiqa, M.R. Dey, Objective image quality evaluation model for JPEG2000 coded images based on edge information measures, *Journal of Electrical and Electronic Engineering*, Vol. 1, No. 4, pp. 101-106, 2013.
- [194] I. Avcibas, B. Sankur, K. Sayood, Statistical evaluation of image quality measures, *Journal of Electronic Imaging*, Vol. 11, No. 2, pp. 206-223, 2002.
- [195] H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, *IEEE Transactions on Image Processing*, Vol. 15, No. 11, pp. 3440-3451, 2006.
- [196] M. Pedersen, J.Y. Hardeberg, Survey of full-reference image quality metrics, *Hogskolen I Gjoviks rapportserie*, Nr.5, 2009.
- [197] S.A. Monjas, Performance evaluation of quality metrics for video quality assessment, Master Thesis, Fakultat IV – Elektrotechnik und Informatik, 2010.
- [198] S. Chikkerur, V. Sundaram, M. Reisslein, L.J. Karam, Objective video quality assessment methods: A classification, review and performance comparison, *IEEE Transactions on Broadcasting*, Vol. 57, No. 2, pp. 165-182, 2011.
- [199] K. Seshadrinathan, A.C. Bovik, Automatic prediction of perceptual quality of multimedia signals – A survey, *Multimedia Tools and Applications*, Vol. 51, No. 1, pp. 163-186, 2011.
- [200] W. Lin, C.-C.J. Kuo, Perceptual visual quality metrics: A survey, *Journal of Visual Communication and Image Representation*, Vol. 22, No. 4, pp. 297-312, 2011.
- [201] A.K. Moorthy, A.C. Bovik, Visual quality assessment algorithms: What does the future hold?, *Multimedia Tools and Applications*, Vol. 51, No. 2, pp. 675-696, 2011.
- [202] L. Zhang, L. Zhang, X. Mou, D. Zhang, A comprehensive evaluation of full reference image quality assessment algorithms, 19th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 1477-1480, 2012.
- [203] A. Beghdadi, M.-C. Larabi, A. Bouzerdoum, K.M. Iftikharuddin, A survey of perceptual image processing methods, *Signal Processing: Image Communication*, Vol. 28, No. 8, pp. 811-831, 2013.
- [204] A. Lahoulou, A. Bouridane, E. Viennet, M. Haddadi, Full-reference image quality metrics performance evaluation over image quality databases, *Arabian Journal for Science and Engineering*, Vol. 38, No. 9, pp. 2327-2356, 2013.
- [205] M. Vranješ, S. Rimac-Drlje, K. Grgić, Review of objective video quality metrics and performance comparison using different databases, *Signal Processing: Image Communication*, Vol. 28, No. 1, pp. 1-19, 2013.
- [206] D.M. Chandler, Seven challenges in image quality assessment: Past, present and future research, *ISRN Signal Processing*, Hindawi Publishing Corporation, Article ID 905685, pp. 1-53, 2013.
- [207] M. Shahid, A. Rossholm, B. Lövsström, H.-J. Zepernick, No-reference image and video quality assessment: a classification and review of recent approaches, *EURASIP Journal on Image and Video Processing* 2014, 2014:40 1-32.
- [208] S. Winkler, *Digital video quality: Vision models and metrics*, John Wiley & Sons Ltd, 2005.

-
- [209] H.R. Wu, K.R. Rao, *Digital video image quality and perceptual coding*, Taylor & Francis Group, 2006.
- [210] Z. Wang, A.C. Bovik, *Modern image quality assessment*, Synthesis Lectures on Image, Video and Multimedia Processing, Morgan & Claypool Publishers, 2006.
- [211] L. Zhang, W. Lin, *Selective visual attention: Computational models and applications*, John Wiley & Sons Singapore Pte. Ltd., 2013
- [212] Z. Wang, Applications of objective image quality assessment methods [Applications corner], *IEEE Signal Processing Magazine*, Vol. 28, No. 6, pp. 137-142, 2011.
- [213] Video Quality Experts Group (VQEG), [Online] Available: www.its.bldrdoc.gov/vqeg/projects-home.aspx
- [214] V. Petrović, C. Xydeas, Objective evaluation of signal-level image fusion performance, *Optical Engineering*, Vol. 44, No. 8, Paper No. 087003, 2005.
- [215] Z. Wang, E.P. Simoncelli, A.C. Bovik, Multi-scale structural similarity for image quality assessment, *Conference Record of the 37th Asilomar Conference on Signals, Systems and Computers*, pp. 1398-1402, 2003.
- [216] W. Handee, P. Wells, *The perception of visual information*, Springer, New York, 1997.
- [217] R.C. Gonzales, R.E. Woods, *Digital image processing*, Second Edition, Prentice Hall, New Jersey, pp. 567-581, 2002.
- [218] W.K. Pratt, *Digital image processing*, Fourth Edition, John Wiley & Sons, Inc., pp. 462-522, 2007.
- [219] P. Barten, *Contrast sensitivity of the human eye and its effects on image quality*, HV Press, Knegsel, 1999.
- [220] V. Petrović, Multisensor pixel-level image fusion, PhD thesis, University of Manchester, 2001.
- [221] V. Petrović, Subjective tests for image fusion evaluation and objective metric validation, *Information Fusion*, Vol. 8, pp. 208-216, 2007.
- [222] B. Bondžulić, V. Petrović, Edge-based objective evaluation of image quality, 18th IEEE International Conference on Image Processing (ICIP), *Proceedings of papers*, pp. 3305-3308, 2011.
- [223] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, LIVE image quality assessment database release 2, [Online] Available: <http://live.ece.utexas.edu/research/quality/subjective.htm>
- [224] Y. Horita, K. Shibata, Y. Kawayoke, Z.M. Parvez Sazzad, MICT image quality evaluation database (Toyama database), [Online] Available: <http://mict.eng.u-toyama.ac.jp/mictdb.html>
- [225] P. Le Callet, F. Autrusseau, Subjective quality assessment IRCCyN/IVC database, [Online] Available: <http://www.irccyn.ec-nantes.fr/ivcdb/>
- [226] E.C. Larson, D.M. Chandler, The CSIQ image database, [Online] Available: <http://vision.okstate.edu/?loc=csiq>
- [227] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, J. Hakkinen, CID2013: A database for evaluating no-reference image quality assessment algorithms, *IEEE Transactions on Image Processing*, Vol. 24, No. 1, pp. 390-402, 2015.
-

-
- [228] S. Tourancheau, F. Atrousseau, Z.M. Parvez Sazzad, Y. Horita, Impact of subjective dataset on the performance of image quality metrics, 15th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 365-368, 2008.
- [229] B. Bondzulić, V. Petrović, Additive models and separable pooling, a new look at structural similarity, *Signal Processing*, Vol. 97, pp. 110-116, 2014.
- [230] K. Okarma, Colour image quality assessment using structural similarity index and singular value decomposition, *International Conference on Computer Vision and Graphics (ICCVG)*, LNCS 5337, Springer, pp. 55-65, 2009.
- [231] V. Petrović, C. Xydeas, Objective image fusion performance characterisation, 10th IEEE International Conference on Computer Vision (ICCV), Proceedings of papers, Vol. 2, pp. 1866-1871, 2005.
- [232] B. Bondžulić, V. Petrović, Objective image quality characterisation, 10th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS), Proceedings of papers, pp. 341-344, 2011.
- [233] H. Liu, I. Heynderickx, TUD image quality database: Eye-tracking release 1, [Online] Available: http://mmi.tudelft.nl/iqlab/eye_tracking_1.html
- [234] H. Liu, I. Heynderickx, Studying the added value of visual attention in objective image quality metrics based on eye movement data, 16th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 3097-3100, 2009.
- [235] W. Osberger, A.J. Maeder, Automatic identification of perceptually important regions in an image, 14th IEEE International Conference on Pattern Recognition (ICPR), Proceedings of papers, pp. 701-704, 1998.
- [236] C. Li, A.C. Bovik, Three-component weighted structural similarity index, *Proc. SPIE 7242, Image Quality and System Performance VI*, Paper No. 72420Q, 2009.
- [237] A.K. Moorthy, A.C. Bovik, Visual importance pooling for image quality assessment, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 3, No. 2, pp. 193-201, 2009.
- [238] A. Ninassi, O. Le Meur, P. Le Callet, D. Barba, Does where you gaze on an image affect your perception of quality? Applying visual attention to image quality metric, 14th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 169-172, 2007.
- [239] Z. Wang, E.P. Simoncelli, Stimulus synthesis for efficient evaluation and refinement of perceptual image quality metrics, *Proc. SPIE 5292, Human Vision and Electronic Imaging IX*, pp. 99-108, 2004.
- [240] N. Amodaj, M. Popović, Adaptive procedure for threshold selection in directional derivative edge detectors, *Proc. SPIE 1349, Applications of Digital Image Processing XIII*, pp. 128-138, 1990.
- [241] G.-H. Chen, C.-L. Yang, S.-L. Xie, Gradient-based structural similarity for image quality assessment, 13th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 2929-2932, 2006.
- [242] C. Li, A.C. Bovik, Content-partitioned structural similarity index for image quality assessment, *Signal Processing: Image Communication*, Vol. 25, No. 7, pp. 517-526, 2010.
-

- [243] D.M. Rouse, S.S. Hemami, Understanding and simplifying the structural similarity metric, 15th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 1188-1191, 2008.
- [244] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, The SSIM index for image quality assessment – suggested usage, [Online] Available: <https://ece.uwaterloo.ca/~z70wang/research/ssim/index.html>
- [245] M. Popović, *Digitalna obrada slike*, Akademska misao, Beograd, 2006.
- [246] M. Sonka, V. Hlavac, R. Boyle, *Image processing, analysis and machine vision*, 2nd edition, PWS Publishing, Pacific Grove, 1998.
- [247] V. Petrović, T. Cootes, R. Pavlović, Dynamic image fusion performance evaluation, 10th International Conference on Information Fusion, Proceedings of papers, pp. 1-7, 2007.
- [248] V. Petrović, B. Bondžulić, Objective assessment of surveillance video quality, Sensor Signal Processing for Defence Conference (SSPD), Proceedings of papers, pp. 1-5, 2012.
- [249] Video Quality Experts Group (VQEG), [Online] Available: <http://www.its.bldrdoc.gov/vqeg/downloads.aspx>
- [250] Laboratory for Image and Video Engineering, LIVE Video Quality Database, [Online]. Available: http://live.ece.utexas.edu/research/quality/live_video.html
- [251] Laboratory for Image and Video Engineering, LIVE Wireless Video Database, [Online]. Available: http://live.ece.utexas.edu/research/quality/live_wireless_video.html
- [252] EPFL-PoliMI Video Quality Assessment Database, [Online]. Available: <http://vqa.como.polimi.it/index.htm>
- [253] D.H. Brainard, The Psychophysics toolbox, *Spatial Vision*, Vol. 10, pp. 433-436, 1997.
- [254] D.G. Pelli, The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision*, Vol. 10, No. 4, pp. 437-442, 1997.
- [255] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, T. Ebrahimi, Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel, First International Workshop on Quality of Multimedia Experience (QoMEX), Proceedings of papers, pp. 204-209, 2009.
- [256] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, T. Ebrahimi, Subjective quality assessment of H.264/AVC video streaming with packet losses, *EURASIP Journal on Image and Video Processing*, Hindawi Publishing Corporation, Article ID 190431, pp. 1-12, 2011.
- [257] Video Quality Experts Group (VQEG), Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Technical report, 2000.
- [258] P. Coriveau, et al., Video Quality Experts Group: Current results and future directions, *Proc. SPIE 4067, Visual Communications and Image Processing*, 2000.
- [259] K. Seshadrinathan, R. Soundararajan, A.C. Bovik, L.K. Cormack, Study of subjective and objective quality assessment of video, *IEEE Transactions on Image Processing*, Vol. 19, No. 6, pp. 1427-1441, 2010.

-
- [260] A.K. Moorthy, A.C. Bovik, Wireless video quality assessment: A study of subjective scores and objective algorithms, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 4, pp. 587-599, 2010.
- [261] M.C.Q. Farias, Visual-quality estimation using objective metrics, *Journal of the Society for Information Display*, Vol. 19, No. 11, pp. 764-770, 2011.
- [262] K. Okarma, P. Lech, A statistical reduced-reference approach to digital image quality assessment, *Computer Vision and Graphics (CVG), LNCS 5337*, Springer, pp. 43-54, 2009.
- [263] Z. Wang, E.P. Simoncelli, Reduced-reference image quality assesment using a wavelet domain natural image statistic model, *Proc. SPIE 5666, Electronic Imaging X*, pp. 149-159, 2005.
- [264] X. Li, D. Tao, X. Gao, W. Lu, A natural image quality evaluation metric, *Signal Processing*, Vol. 89, No. 4, pp. 548-555, 2009.
- [265] R. Soundararajan, A.C. Bovik, RRED indices: Reduced reference entropic differencing for image quality assessment, *IEEE Transactions on Image Processing*, Vol. 21, No. 2, pp. 517-526, 2012.
- [266] L. Ma, S. Li, K.N. Ngan, Reduced-reference video quality assessment of compressed video sequences, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 22, No. 10, pp. 1441-1456, 2012.
- [267] M. Tagliasacchi, G. Valenzise, M. Naccari, S. Tubaro, A reduced-reference structural similarity approximation for videos corrupted by channel errors, *Multimedia Tools and Applications*, Vol. 48, No. 3, pp. 471-492, 2010.
- [268] X. Gao, W. Lu, D. Tao, X. Li, Image quality assessment based on multiscale geometric analysis, *IEEE Transactions on Image Processing*, Vol. 18, No. 7, pp. 1409-1423, 2009.
- [269] Z. Wang, G. Wu, H.R. Sheikh, E.P. Simoncelli, E.-H. Yang, A.C. Bovik, Quality-aware images, *IEEE Transactions on Image Processing*, Vol. 15, No. 6, pp. 1680-1689, 2006.
- [270] K. Zeng, Z. Wang, Quality-aware video based on robust embedding of intra- and inter-frame reduced-reference features, *17th IEEE International Conference on Image Processing (ICIP), Proceedings of papers*, pp. 3229-3232, 2010.
- [271] S. Wolf, M.H. Pinson, Low bandwidth reduced reference video quality monitoring system, *International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), Proceedings of papers*, pp. 23-25, 2005.
- [272] Gaze-based Communication Project, [Online]. Available: www.gazecom.eu
- [273] A. Ninassi, O. Le Meur, P. Le Callet, D. Barba, A. Tirel, Task impact on the visual attention in subjective image quality assessment, *European Signal Processing Conference (EUSIPCO), Invited paper*, 2006.
- [274] C. Vu, E.C. Larson, D.M. Chandler, Visual fixation patterns when judging image quality: Effects of distortion type, amount and subject experience, *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Proceedings of papers*, pp. 73-76, 2008.
- [275] J. Redi, H. Liu, P. Gastaldo, R. Zunino, I. Heynderickx, How to apply spatial saliency into objective metrics for JPEG compressed images? *16th IEEE*
-

- International Conference on Image Processing (ICIP), Proceedings of papers, pp. 961-964, 2009.
- [276] Y. Zhong, I. Richardson, A. Sahraie, P. McGeorge, Influence of task and scene content on subjective video quality, Image Analysis and Recognition, LNCS 3211, Springer, pp. 295-301, 2004.
- [277] D. Mahapatra, S. Winkler, S.-C. Yen, Motion saliency outweighs other low-level features while watching videos, Proc. SPIE 6806, Human Vision and Electronic Imaging XIII, Paper No. 6806P, 2008.
- [278] G. Abdollahian, Z. Pizlo, E.J. Delp, A study on the effect of camera motion on human visual attention, 15th IEEE International Conference on Image Processing (ICIP), Proceedings of papers, pp. 693-696, 2008.
- [279] A.K. Moorthy, W.S. Geisler, A.C. Bovik, Evaluating the task dependence of eye movements for compressed videos, 5th International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), Proceedings of papers, pp. 12-16, 2010.
- [280] H. Hadizadeh, M.J. Enriquez, I.V. Bajic, Eye-tracking database for a set of standard video sequences, IEEE Transactions on Image Processing, Vol. 21, No. 2, pp. 898-903, 2012.
- [281] Stefan Winkler, Image and video quality resources, [Online]. Available: <http://stefan.winkler.net/resources.html>
- [282] D.S. Wooding, Fixation maps: Quantifying eye-movement traces, Eye Tracking Research and Applications Symposium (ETRA), Proceedings of papers, pp. 31-36, 2002.
- [283] M. Vranješ, S. Rimac-Drlje, K. Grgić, Locally averaged PSNR as a simple objective video quality metric, 50th International Symposium ELMAR, Proceedings of papers, Vol. 1, pp. 17-20, 2008.
- [284] D.D. Salvucci, J.H. Goldberg, Identifying fixations and saccades in eye-tracking protocols, Eye Tracking Research and Applications Symposium (ETRA), Proceedings of papers, pp. 71-78, 2000.