



Универзитет у Новом Саду
Факултет техничких наука
Одсек за рачунарску технику
и рачунарске комуникације



Докторска дисертација:

МУЛТИ-РЕЗОЛУЦИОНА МЕРА
ЗА ОБЈЕКТИВНУ ОЦЕНУ КВАЛИТЕТА
СИНТЕТИЗОВАНИХ СЛИКА ФТВ ВИДЕО СИГНАЛА

Ментор:

Др Драган Кукољ

Кандидат:

Мр Драгана Сандић-Станковић

Нови Сад, 2016.

МУЛТИ-РЕЗОЛУЦИОНА МЕРА
ЗА ОБЈЕКТИВНУ ОЦЕНУ КВАЛИТЕТА
СИНТЕТИЗОВАНИХ СЛИКА
ФТВ ВИДЕО СИГНАЛА



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 Драгана Сандић-Станковић, дипл. инж.
Ментор, МН:	проф. др Драган Кукољ
Наслов рада, НР:	Мулти-резулациона мера за објективну оцену квалитета синтетизованих слика ФТВ видео сигнала
Језик публикације, ЈП:	Српски
Језик извода, ЈИ:	Српски
Земља публикавања, ЗП:	Република Србија
Уже географско подручје, УГП:	
Година, ГО:	2016
Издавач, ИЗ:	Ауторски репринт
Место и адреса, МА:	Нови Сад, Трг Доситеја Обрадовића 6
Физички опис рада, ФО: (поглавља/страна/	8 поглавља, 129 страница, 102 референце, 20 табела, 49 слика
Научна област, НО:	Електротехничко и рачунарско инжењерство
Научна дисциплина, НД:	Телекомуникације
Предметна одредница/Кључне речи, ПО:	обрада слике, процена квалитета ДИБР-синтетизоване слике, мултирезулациона метрика за оцену квалитета слике, декомпозиција слике морфолошком пирамидом и морфолошким таласићима
УДК	
Чува се, ЧУ:	У библиотеци Факултета техничких наука, Нови Сад
Важна напомена, ВН:	
Извод, ИЗ:	Основни допринос ове докторске дисертације је развој алгоритама за објективну процену визуелног квалитета слике синтетизоване применом ДИБР (Depth Image Based Rendering) техника које узрокују неуниформна изобличења у области ивица. Применом нелинеарних морфолошких филтара у мултирезулационој декомпозицији слика код израчунавања предложене метрике, важне геометријске информације као што су ивице су добро очуване без помака и замућења у сликама на различитим скалама мултирезулационе репрезентације. Израчунавањем МСЕ по подопсезима који садрже ивице, пиксел по пиксел, прецизно се мери разлика две мултирезулационе репрезентације. Тако се највећи значај у процени квалитета додељује области ивица. Процене предложене метрике се добро поклапају са субјективним оценама.
Датум прихватања теме, ДП:	22. фебруар 2016.

Датум одбране, ДО:			
Чланови комисије, КО:	Председник:	Др Миодраг Темеринац редовни професор	
	Члан:	Др Ирини Рељин редовни професор	
	Члан:	Др Мирослав Поповић редовни професор	Потпис
	Члан:	Др Никола Теслић редовни професор	
	Члан, ментор:	Др Драган Кукољ редовни професор	



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 :	Dragana Sandić-Stanković, Msc.
Mentor, MN :	Prof. Dragan Kukulj, PhD
Title, TI :	Multi-scale metric for objective synthesized image quality assessment for FTV
Language of text, LT :	Serbian
Language of abstract, LA :	Serbian
Country of publication, CP :	Republic of Serbia
Locality of publication, LP :	
Publication year, PY :	2016
Publisher, PB :	Author's reprint
Publication place, PP :	Novi Sad, Dositeja Obradovića sq. 6
Physical description, PD : (chapters/pages/ref./tables/pictures/graphs/app endixes)	8 chapters, 129 pages, 102 references, 20 tables, 49 pictures
Scientific field, SF :	Electrical and computer engineering
Scientific discipline, SD :	Telecommunications
Subject/Key words, S/KW :	image processing, DIBR synthesized view quality assessment, multi-scale IQA metric, multi-scale decomposition using morphological filters, morphological pyramid, morphological wavelets
UC	
Holding data, HD :	The Library of Faculty of Technical Sciences, NoviSad, Serbia
Note, N :	
Abstract, AB :	The main contribution of this doctoral thesis is the development of algorithms for objective DIBR-synthesized view quality assessment. DIBR algorithms introduce nonuniform geometric distortions affecting the edge coherency in the synthesized images. The non-linear morphological filters used in multi-scale image decompositions of the proposed metric maintain important geometric information such as edges across different resolution levels. Calculating MSE pixel-by-pixel through subbands in which the edges are extracted, the difference of the two multiresolution representations, the reference and the synthesized image, is precisely measured. In that way the importance of edge areas which are prone to synthesis artifacts is emphasized in the image quality assessment. The proposed metric has very good agreement with human judgment.

Accepted by the Scientific Board on, ASB:	22. February 2016.	
Defended on, DE:		
Defended Board, DB:	President:	Miodrag Temerinac, Phd.
	Member:	Irina Reljin, Phd
	Member:	Miroslav Popović, Phd
	Member:	Nikola Teslić, Phd
	Member, Mentor:	Dragan Kukulj, Phd
		Mentor's sign

Захвалница

Захваљујем се ментору професору др Драгану Кукољу на великој професионалној подршци, поверењу, разумевању и помоћи током израде докторске дисертације. Веома сам захвална ментору на помоћи у остваривању контаката са истраживачима у свету: за научну посету (Short Term Scientific Mission, STSM) истраживачкој групи IRCCyN/IVC Универзитета у Нанту (Ecole polytechnique de l'Universite de Nantes, IRCCyN Image & Video Communication Lab) у децембру 2013. године која је финансирана од стране e-COST (European Cooperation in Science and Technology) Action IC1003 QUALINET (European Network on Quality of Experience in Multimedia Systems and Services). Захвална сам професору Patrick Le Callet на гостопримству током STSM у истраживачкој групи IRCCyN/IVC Универзитета у Нанту (Ecole polytechnique de l'Universite de Nantes) и за праћење мог рада током израде дисертације. Током STSM користила сам базу слика коју је креирала група IRCCyN/IVC за тестирање метрике која је предложена у овом раду а касније сам користила и њихову базу видео секвенци. Захваљујем се и професору Marcus Barkowsky на сарадњи током STSM и објашњењима о нелинеарном мапирању резултата објективне метрике према субјективним оценама према препоруци групе VQEG (Video Quality Experts Group). Резултате остварене током STSM сам презентовала на састанку у организацији e-COST Action IC1003 QUALINET у Берлину у марту 2014.

Захвална сам ментору Кукољу и професору Le Callet за учествовање у летњој школи (Training School) са темом 3D AudioVisual Content Processing and Communications одржаној у Лисабону у јулу 2015 коју је организовао e-COST Action IC1105: 3D-ConTourNet и за учествовање на конференцији 3DTV која је одржана у истој недељи такође на Институту Универзитета у Лисабону (Iscte-Iul, Instituto Universitário De Lisboa). Присутствовала сам састанцима групе COST Action IC11053D-ConTourNet (3D Content Creation, Coding and Transmission over Future Media Networks) на Одсеку за инжењерство Универзитета у Риму (Engineering Department at University of Roma TRE) у фебруару 2016 и на Факултету за информационе и комуникационе системе Универзитета у Шведској (Dept. of Information and Communication Systems Mid-Sweden University, Sundsvall, Sweden) у марту 2016.

Хвала мојој породици, супругу Предрагу и синовима Милану и Душану.

Сажетак

ЗДВ (3Д видео) је друга фаза у развоју ФТВ (Free-viewpoint TV) стандарда коју спроводи MPEG (Moving Picture Experts Group). ФТВ, као најсавременији визуелни медиј, омогућава кориснику да интерактивно мења позицију посматрања 3Д сцене. Према ЗДВ стандарду мањи број снимљених видео секвенци исте сцене на различитим позицијама се шаље са предајне стране а на пријемној страни се генерише већи број видео секвенци на основу примљених снимљених видео секвенци и њима одговарајућих мапа дубине. Видео секвенце се генеришу применом ДИБР (Depth-Image-Based-Rendering) алгоритма синтезе за позиције посматрања сцене за које нема снимљених видео секвенци. ДИБР алгоритми за синтезу слика уносе нови тип изобличења у синтетизоване слике која се појављују углавном око области које су заклоњене на снимљеној слици а требало би да буду видљиве у синтетизованој слици на новој позицији. Изобличења слике због ДИБР синтезе нису униформно распоређена на слици као изобличења због компресије већ су уочљива у области ивица и неки објекти су померени. Типична изобличења због синтезе су црне празнине, замућење ивица, померај ивица или дупле ивице. ДИБР алгоритми примењују геометријске трансформације слика па синтетизоване слике садрже геометријске дисторзије у области ивица.

Квалитет синтетизованих слика је веома важан за успех 3Д видео апликација. Стандардне мере за оцену квалитета слика не дају добре резултате при оцењивању синтетизованих слика које садрже неуниформна изобличења ивица. Процене метрика које су посебно дизајниране за оцену квалитета слика синтетизованих ДИБР алгоритмима се боље поклапају са субјективним оценама него стандардне метрике при оцењивању синтетизованих слика, али и даље постоји потреба за метрикама са још бољим перформансама.

Да би добили метрику за оцену квалитета синтетизованих слика која се добро поклапа са субјективним оценама, у раду је предложена примена мултирезолуционе декомпозиције синтетизоване и референтне слике коришћењем морфолошких филтара у првој фази израчунавања метрике. Експериментално је потврђено да човек визуелни систем обрађује визуелне информације на мултирезолуциони начин и примена мултирезолуционе декомпозиције слика при израчунавању квалитета слике доприноси повећању поузданости

метрике. Применом нелинеарних морфолошких филтара у мултирезолуционој декомпозицији, важне геометријске информације као што су ивице су добро очуване без помака и замућења у сликама на различитим скалама мултирезолуционе репрезентације. Израчунавањем средње квадратне грешке као мере изобличења по подопсезима мултирезолуционе репрезентације који садрже ивице пиксел по пиксел, прецизно се мери разлика две мултирезолуционе репрезентације. Тако се највећи значај у оцени квалитета додељује области ивица које су склоне дисторзијама у синтетизованим сликама због процеса ДИБР синтезе. Комбиновањем средњих квадратних грешака подопсега мултирезолуционе репрезентације синтетизоване слике добија се резултујућа оцена метрике.

Анализиране су две верзије морфолошке мултирезолуционе метрике: MP-PSNR (Morphological Pyramid Peak Signal-to-Noise Ratio) која се заснива на морфолошкој пирамидалној декомпозицији и MW-PSNR (Morphological Wavelet Peak Signal-to-Noise Ratio) у којој је примењена декомпозиција слике морфолошким таласићима.

Предложена метрика је тестирана на три базе слика/видео секвенци које садрже слике/видео секвенце генерисане применом ДИБР техника синтезе. За тестирање метрике коришћене су: база слика која садржи слике генерисане применом седам различитих алгоритама ДИБР синтезе, IRCCyN/IVCDIBR, део базе стерео слика генерисаних применом четири различите технике синтезе, MCL-3D, и база видео секвенци са навигацијом кроз различите позиције, FVV, које садрже изобличења због процеса синтезе и због компресије мапа дубине. Обе верзије предложене метрике, MP-PSNR и MW-PSNR, показују значајно боље перформансе од стандардних метрика за оцену квалитета слика као што су PSNR, SSIM, MS-SSIM и значајно боље перформансе од метрике 3DswIM, посебно дизајниране за оцену квалитета слика синтетизованих ДИБР алгоритмима.

Показано је да се резултати највише коришћене објективне мере, вршне вредности сигнал/шум, PSNR, добро поклапају са субјективним оценама када се PSNR израчунава на подопсезима детаља виших нивоа морфолошке мултирезолуционе репрезентације синтетизоване слике. Предложене су редуковане верзије морфолошке мултирезолуционе метрике, MP-PSNR_r и MW-PSNR_r, које користе само подопсега виших нивоа

декомпозиције. Редуковане верзије метрика показују боље перформансе од основних верзија.

Морфолошки оператори коришћени у мултирезолуционој декомпозицији слика су једноставни за имплементацију јер користе само операције минимум, максимум и сабирање, раде са целим бројевима и као резултат дају целе бројеве. Израчунавање средње квадратне грешке која се користи као мера изобличења по подопсезима мултирезолуционе репрезентације слике је такође веома једноставно па је предложена метрика рачунски врло ефикасна.

Кључне речи: оцена квалитета ДИБР-синтетизованих слика/видеа, мултирезолуциона метрика за оцену квалитета слике, мултирезолуциона декомпозиција слике применом морфолошких филтара, морфолошка пирамида, морфолошки таласићи

Abstract

3DV (3D video) is the second phase in the development of FTV (Free-viewpoint TV) standard implemented by MPEG (Moving Picture Experts Group). FTV is the most immersive visual media that enables the user to view a three-dimensional scene by freely changing viewpoints. In the 3DV standard, smaller number of captured views is transmitted and greater number of views is synthesized at the receiver side. Video sequences are generated using DIBR (Depth-Image-Based-Rendering) techniques which introduce new types of artefacts mostly located around the disoccluded regions. As the DIBR algorithms involve geometric transformations, most of them introduce nonuniform geometric distortions affecting the edge coherency in the synthesized images. Typical artefacts due to synthesis process are black holes, blurred edges, shifted edges or ghost edges.

The synthesized image quality is of great importance for the success of the 3D video applications. The common image quality assessment metrics, which are primarily designed for other types of distortions, don't handle efficiently the synthesis distortion. The metrics dedicated to synthesized image quality assessment show better performances in the quality assessment of synthesized views, but there is still room for improvement.

In order to better deal with specific geometric distortions in the DIBR synthesized images, full-reference metric based on multi-scale image decompositions using morphological filters is proposed in this thesis. Psychophysics and physiological experiments have shown that multiscale transforms seem to appear in the visual cortex of mammals. Introducing multiresolution decomposition of the reference and the synthesized images in the image quality assessment, better correlation of the objective metric's results to subjective results is achieved. The non-linear morphological filters used in the multi-scale image decomposition maintain important geometric information such as edges on their true positions, neither drifted nor blurred, across different resolution levels. Morphological multiresolution decomposition can be interpreted as structural image decomposition tending to enhance image features such as edges which are segregated by scale at the various resolution levels. Edge distortion between appropriate subbands of the multi-scale representations of the reference image and the DIBR synthesized image is measured using Mean Squared Error (MSE) precisely pixel-by-pixel. In this way, areas around edges that are prone to synthesis artifacts are emphasized in the metric score. Subbands' mean squared errors

are combined into multi-scale mean squared error, which is transformed into multi-scale peak signal-to-noise ratio measure.

Two versions of morphological multi-scale metric have been analysed: Morphological Pyramid Peak Signal-to-Noise Ratio (MP-PSNR) based on morphological pyramid decomposition, and Morphological Wavelet Peak Signal-to-Noise Ratio (MW-PSNR) based on morphological wavelet decomposition.

The performances of the proposed metrics are tested using three image/video databases which contain DIBR-synthesis artefacts. The used databases are the IRCCyN/IVCDIBR database which contains images synthesized using seven DIBR algorithms, the part of the MCL-3D database which contains stereoscopic images synthesized using four DIBR algorithms, and the Free-Viewpoint video database, FVSV, which contains video sequences with distortions due to synthesis process and depth maps compression. Proposed metrics achieve significantly higher correlation with human judgment compared to the state-of-the-art image quality metrics such as PSNR, SSIM, MS-SSIM and compared to the tested metric dedicated to synthesis-related artifacts, 3DswIM.

We have demonstrated that PSNR has particularly good agreement with human judgment when it is calculated between images at higher scales of morphological multi-scale representations. Consequently, simplified and in essence reduced versions of multi-scale metrics, MP-PSNR_r and MW-PSNR_r, are proposed, taking into account only detailed images at higher decomposition scales. The performances of reduced versions of the morphological multi-scale measure are improved comparing to their full versions.

Morphological operators used in the image multiresolution decomposition are computationally efficient since they involve only integer numbers and simple computations like a min, max and sum. The calculation of MSE by subbands is also computationally simple and the proposed metric is of low computational complexity.

Keywords: DIBR-synthesized view quality assessment, Multi-scale image quality metric, Morphological multi-scale image decomposition, morphological pyramid, morphological wavelets

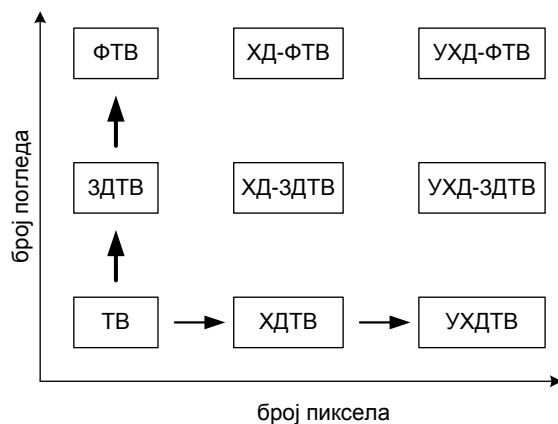
Садржај

1	Увод	15
1.1	Предмет истраживања	16
1.2	Циљ истраживања	22
1.3	Допринос истраживања	23
1.4	Организација дисертације	24
2	Преглед метрика за оцену квалитета ДИБР синтетизованих слика/видеа	26
3	Мултирезолуциона репрезентација слике	33
3.1	Преглед мултирезолуционих метрика за оцену квалитета слика	33
3.2	Опште шеме мултирезолуционе декомпозиције слике	35
3.2.1	Редундантна мултирезолуциона репрезентација слике	36
3.2.2	Нередундантна мултирезолуциона репрезентација слике	38
4	Предложена мулти-резолуциона метрика за оцену квалитета ДИБР-синтетизованих слика	43
4.1	Мултирезолуциона декомпозиција слике применом морфолошких оператора	44
4.1.1	Пирамидална декомпозиција слике применом морфолошких оператора	45
4.1.2	Декомпозиција слике морфолошким таласићима	48
4.2	Израчунавање дисторзије	59
4.2.1	Израчунавање MP-PSNR	60
4.2.2	Израчунавање MW-PSNR	61
4.3	Рачунска комплексност метрике	63
5	Резултати	67
5.1	Мере за поређење перформанси објективних метрика за оцену квалитета слика	67
5.2	Тест пример 1: база синтетизованих слика IRCCyN/IVCDIBR	69
5.2.1	Опис базе IRCCyN/IVCDIBR	69

5.2.2. Перформансе предложене метрике при оцењивању базе IRCCyN/IVCDIBR	72
5.3. Тест пример 2: база синтетизованих стерео слика MCL-3D	88
5.3.1. Опис базе MCL-3D	88
5.3.2. Перформансе предложене метрике при оцењивању базе MCL-3D	94
5.4. Тест пример 3: база синтетизованих видео секвенци FVSV	102
5.4.1. Опис базе FVSV	102
5.4.2. Перформансе предложене метрике при оцењивању базе FVSV	104
6 Закључак	110
7 Литература	112
8 Прилог	121
8.1 Списак слика	121
8.2 Списак табела	126
8.3 Списак скраћеница	128

1 УВОД

Визуелни медији као што су телевизија, филм и фотографија су дигитализовани системи засновани на пикселима. Један правац њиховог развоја је повећања квалитета слике повећањем броја пиксела. После телевизије стандардне резолуције, развијена је телевизија високе резолуције (ХДТВ) а затим телевизија ултра високе резолуције (УХДТВ, 4к/8к), слика 1.1. Сви ови системи омогућују кориснику посматрање сцене са једне позиције. Други правац развоја телевизије је повећање броја позиција посматрања сцене. Да би корисницима омогућили доживљај дубине као у реалном свету, после конвенционалне 2Д телевизије која омогућује кориснику посматрање сцене са једне позиције дошло је до развоја тродимензионе телевизије (ЗДТВ). Стерео ЗДТВ омогућује кориснику осећај дубине сцене. Стерео слика садржи две слике које одговарају позицијама које су мало размакнуте. Свака слика се посебно пројектује на мрежњачу ока и на основу њих човеков визуелни систем креира слику у мозгу која омогућује 3Д доживљај. Ауто-стерео ЗДТВ омогућује 3Д доживљај без ношења наочара на неколико позиција унутар ограничене зоне посматрања. Импресивну комуникацију са слободним избором позиције посматрања омогућује тек телевизија са слободним избором позиције посматрања, ФТВ (Free-viewpoint TV). ФТВ доноси епохалне промене у историји телевизије омогућујући бесконачно много погледа и позиција посматрања може бити било где[1]. ФТВ се сматра крајњим циљем ЗДТВ.



Слика 1.1. Категоризација телевизије

1.1 Предмет истраживања

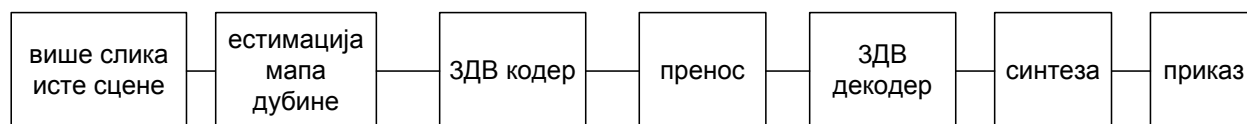
ФТВ као најсавременији визуелни медиј омогућава кориснику посматрање 3Д сцене без наочара и без замора очију уз интерактивну контролу позиције посматрања. За реализацију ФТВ, потребне су видео секвенце на великом броју позиција. Због ограничења пропусног опсега у преносу, ограничења у трошковима, за мањи број позиција видео секвенце се снимају, кодују и преносе а за већи број позиција видео секвенце се синтетизују на пријемној страни у реалном времену. За ефикасну репрезентацију 3Д података усвојен је формат MVD (Multiview Video plus Depth, MVD) [2]. MVD садржи више видео секвенци снимљених за исту сцену на различитим позицијама и њима одговарајућих секвенци мапа дубине (depth maps). Мапа дубине је монохроматска слика у скали сивог која садржи информације о геометрији 3Д сцене. Интензитет пиксела мапе дубине представља удаљеност оптичког центра камере од површине објекта 3Д сцене. Видео секвенце се на пријемној страни синтетизују за позиције за које нема снимљених видео секвенци на основу снимљених видео секвенци и снимљених или естимираних мапа дубине (depth maps) применом ДИБР (Depth Image Based Rendering) алгоритама [3].

Интернационалну стандардизацију ФТВ спроводи MPEG (Moving Picture Experts Group) од 2001. године кроз 3 фазе [4]. У првој фази која је трајала од марта 2004 до маја 2009 усвојен је стандард за ефикасно кодовање скупа видео секвенци [5], МВЦ (Multi-view Video Coding, MVC), као проширење стандарда за кодовање једне видео секвенце H.264/MPEG-4 AVC (Advanced Video Coding). Између видео секвенци снимљених са различитих позиција око исте сцене постоји корелација. Редундантност између видео секвенци се може отклонити предикцијом између видео секвенци компензацијом покрета као код конвенционалног видеа. МВЦ примењује компензацију покрета по времену и по погледима. МВЦ систем је приказан на слици 1.2. Систем са више синхронизованих камера распоређених у 3Д простору снима сцену на више позиција (multiview) и омогућује корисницима на пријемној страни посматрање сцене са фиксног броја позиција у широкој зони посматрања. Број позиција снимања сцене на предајној страни и број позиција посматрања сцене на пријемној страни је исти. МВЦ је примењен за Блу-реј 3Д.



Слика 1.2. МВЦ систем

Друга фаза је 3Д видео (ЗДВ) и трајала је у периоду од априла 2007. до 2014. године. Развијен је стандард за ефикасно кодовање скупа видео секвенци и њима одговарајућих мапа дубине у MVD (multiview plus depth) формату [2]. Кодовање садржаја у MVD формату се заснива на стандардима 3D-AVC [6] или 3D-HEVC (High Efficiency Video Coding) [7]. ЗДВ систем је приказан на слици 1.3. Код ЗДВ се скуп видео секвенци и мапа дубине у MVD формату компримује пре преноса а на пријемној страни се на основу декодованих података синтетизују нове видео секвенце за позиције за које нема снимљених видео секвенци применом ДИБР алгоритама. ДИБР технике користе геометрију сцене за синтезу слика на новим (виртуелним) позицијама. Метод синтезе бољих перформанси омогућује смањење броја камера за снимање, смањење пропусног опсега за пренос, смањење трошкова. Укупан број видео секвенци на пријемној страни је већи него на предајној. Циљ је прилагођавање садржаја за приказ на више дисплеја и за различите позиције. Аутостерео 3Д дисплеји омогућују посматрање 3Д садржаја из различитих углова без наочара. 3Д видео апликације имају примену у области забаве (3Д телевизија, 3Д биоскоп, ФТВ, 3Д игре) или као апликације за посебне намене (телеконференција, даљинско образовање, навигација робота, медицинске слике).



Слика 1.3. ЗДВ систем

Трећа фаза ФТВ је почела у августу 2013. и има за циљ “super multiview“ апликације и апликације са произвољним избором позиције посматрања (free navigation). ФТВ систем је приказан на слици 1.4. Излаз ФТВ-а је “super multiview“ за “super multiview“ дисплеје или један поглед за слободну навигацију. “Super multiview” дисплеји приказују стотине слика распоређених ултра-густо и тако омогућују пријатан 3Д доживљај без наочара у широком

углу посматрања, гладак прелаз између суседних позиција и доживљај шетње око објеката у предњем делу сцене. За приказ на “super multiview“ дисплејима, потребан је велики број камера за снимање сцене (типично 80 камера). За пренос великог броја снимљених видео секвенци потребна је велика ширина пропусног опсега. Синтеза видео секвенци се заснива на интерполацији снимљених слика. Коришћењем камера које снимају и дубину сцене може се повећати опсег избора позиција посматрања и смањити укупан број камера за снимање. Код апликација са слободном навигацијом корисник може имати осећај шетње или летења кроз сцену.



Слика 1.4. ФТВ систем

Комитет ФИФА (Међународна федерација фудбалских асоцијација) светског купа Јапана је планирао да се пренос ФИФА светског купа 2022. године оствари преко ФТВ. Јапан планира испоруку 3Д реплике фудбалског стадиона по целом свету преко ФТВ [8]. ФТВ ће наћи примену у области широкопојасног преноса, комуникација, забаве, реклама, изложби, образовања, медицине, уметности, архивирања, безбедности, надзора. Већ је реализован у реалном времену и ФТВ се може гледати на рачунару или мобилном телефону.

ДИБР синтеза слика

Пошто је број камера које снимају сцену ограничен, за позиције за које недостају снимљене слике ради се синтеза (интерполација) на основу снимљених слика са суседних позиција. Методе синтезе се могу поделити према томе да ли користе геометријске информације о сцени или не. Неке од најпознатијих метода које не користе геометријске информације и примењују интерполацију и филтрирање за синтезу слика су „light field rendering“ [9], „concentric mosaic“ [10] или „lumigraph“ [11]. Другу групу чине методе које користе геометријске информације садржане у мапама дубине (depth maps) за пројекцију пиксела из референтне слике на коректну позицију у синтетизованој слици (Depth Image-Based Rendering, DIBR).

ДИБР је техника генерисања слике на новој позицији на основу слике и мапе дубине на суседној позицији [12]. Први корак ДИБР синтезе је пројектовање (3D warping) колор слике (текстуре) и мапе дубине на виртуелну позицију. Пиксели референтне слике се пројектују у 3Д простор коришћењем информација из мапе дубине и информација о камерама. Затим се пиксели из 3Д простора пројектују на 2Д слику на виртуелној позицији. Овај корак има највећи утицај на квалитет синтетизоване слике. Процес синтезе не утиче на вредност (интензитет) пиксела у синтетизованој слици али мења позицију пиксела. Када се израчунава позиција пиксела у синтетизованој слици резултат може бити нецелобројан па се заокружује на целобројан или се ради интерполација. У оба случаја се прави грешка. Квалитет пројектоване мапе дубине зависи од броја ивица у предњем делу сцене јер се грешке концентришу на ивицама. Ограничење броја нивоа квантизације мапе дубине узрокује грешке при пројекцији. Пројекција може узроковати пукотине (cracks) у виртуелној слици. Да би квалитет синтетизоване слике био бољи, грешке пројекције треба минимизовати. То се може постићи генерисањем мапе дубине са више нивоа квантизације. Квалитет синтетизоване слике се може поправити и обрадом ивица објеката у предњем делу сцене. Већина алгоритама ДИБР синтезе користи 2 референтне слике са 2 позиције (лево и десно у односу на нову позицију) за генерисање слике на позицији која недостаје (виртуелна позиција). Слике настале пројекцијом 2 референтне слике (леве и десне) се комбинују (blend) и тако се значајно смањују дисоклузије а тиме и грешка синтетизоване слике. Овај корак значајно доприноси квалитету синтетизоване слике. У последњем кораку се попуњавају преостале дисоклузије применом алгорита за попуњавање (inpainting algorithms). Иако процентуално мали број пиксела слике припада области “дисоклузија”, тако се поправља субјективни квалитет слике јер непопуњене црне празнине значајно нарушавају визуелни квалитет слике. Анализа корака ДИБР алгорита је представљена у [13]. Угао између суседних камера је важан фактор за 3Д МВВ (Multi View Video). Мада је деградација квалитета по једном степену угла између виртуелне и референтне позиције мала, не препоручују се углови већи од 15 степени јер узрокују велике области “дисоклузија”.

Изобличења у ДИБР синтетизованим сликама

У 3Д видео системима процес ДИБР синтезе и грешке у мапама дубине су нови извори дисторзија у односу на оне који постоје у конвенционалним 2Д видео системима. Они узрокују изобличења ивица објеката (просторна дисторзија) у синтетизованим сликама и светлуцање при гледању видеа (временска дисторзија). За разлику од дисторзија због компресије које су раширене по целој слици, процес ДИБР синтезе уноси изобличења углавном дуж ивица објеката [14]. Униформне текстуре су без дисторзије а ивице објеката и комплексне текстуре садрже изобличења. Нису ни све контуре једнако изобличене. Геометријска изобличења у области ивица су веома уочљива. Код објеката који се крећу у предњем делу сцене изобличења због синтезе у области ивица између објеката и позадине постају још уочљивија. Најуочљивија временска дисторзија код синтетизованих видео секвенци је светлуцање (flickering) [15].

Како ДИБР алгоритми садрже геометријске трансформације они узрокују углавном геометријска изобличења слике у области ивица. Процес ДИБР синтезе уноси нови тип изобличења у синтетизовану слику, највише у области “дисклузија“. “Дисклузије“ су области сцене које се не виде са позиције са које се снима референтна слика (делови позадине заклоњени објектима у предњем делу сцене) али постају видљиве кад се промени позиција. Како за њихову синтезу нема података о њиховом садржају оне се манифестују као непопуњене области (празнине) у синтетизованој слици. Због помераја целог видног поља на виртуелној позицији у односу на оригиналну позицију дуж ивица синтетизоване слике се јавља “дисклузија“. “Дисклузија“ се јавља и дуж ивица објеката на слици. Када се слика синтетизује на основу снимљених слика са обе стране, леве и десне, “дисклузије“ су мање. Делови слике на местима где су “дисклузије“ се попуњавају применом алгоритама за попуњавање делова слике који недостају (inpainting) који користе технике екстраполације или интраполације. У зависности од примењеног метода екстраполације у алгоритму за попуњавање делова слике који недостају, региони у синтетизованој слици могу бити помакнути или промењене величине (resized). Због попуњавања “дисклузија“ ивице могу бити замућене (blurry) што је посебно видљиво на прелазима између позадине и објеката у предњем делу сцене.

Грешке у мапама дубине настају због ограничености уређаја за снимање или при естимацији мапе дубине и због квантизације мапе дубине. Ове грешке узрокују

геометријске дисторзије у синтетизованој слици јер доводе до грешака у мапирању пиксела из референтне слике у синтетизовану слику на погрешне позиције. Тако настаје померај објеката у синтетизованим сликама. Промена позиција грешака око региона високог контраста око ивица објеката у предњем делу сцене узрокују светлуцање приликом гледања видеа. Због несавршености мапа дубине могу се јавити пукотине у синтетизованој слици. Непоклапање ивица слике (текстуре) и мапе дубине може узроковати пресликавање пиксела који припадају објектима у предњем делу слике у позадину или обрнуто.



Слика1.5. Типична изобличења у сликама због процеса ДИБР синтезе. Делови оригиналних слика су у левој колони а делови синтетизованих слика које садрже изобличења у десној.

На слици 1.5 су приказани примери описаних типичних изобличења у сликама настали због процеса ДИБР синтезе.

1.2 Циљ истраживања

Оцена квалитета синтетизованих слика/видеа има значајну улогу код оцењивања перформанси различитих делова видео система као што су снимање 3Д садржаја, генерисање и обрада мапа дубине, компресија 3Д садржаја, расподела ресурса (алокација битског протока) између видео секвенци и мапа дубине при преносу и синтеза видео секвенци. Поуздана метода оцене квалитета слика/видеа је значајна за оптимизацију перформанси целог 3Д видео система у циљу побољшања квалитета доживљаја (Quality of Experience, QoE) крајњег корисника. Циљ функције оптимизације је минимизација дисторзије синтетизованих слика/видеа. Квалитет синтетизованих слика/видеа је најзначајнији критеријум код оцењивања целог 3Д видео система. За интерактивне видео сервисе, потребно је подешавање параметара компресије на основу квалитета видео сигнала. Мерење квалитета 3Д видео сигнала на пријему се може користити као повратна информација за фино подешавање параметара система тако да се остваре боље перформансе система. Могућност промене параметара компресије и преноса је веома важна за испоруку робусног висококвалитетног 2Д/3Д видео сервиса преко непоузданих комуникационих канала.

Субјективно оцењивање спроводе људи додељујући оцене садржају који посматрају. Субјективне методе оцењивања су временски захтевне, скупе и непрактичне у системима у којима је потребна оцена квалитета у реалном времену. Објективне методе имају за циљ да аутоматски оцене квалитет слике/видеа тако да се оцена што боље поклапа са субјективном оценом.

У овом раду је предложена метрика за оцену квалитета слика/видеа синтетизованих коришћењем ДИБР техника синтезе. ДИБР алгоритми садрже геометријске трансформације и узрокују углавном геометријска изобличења слике у области ивица. Стандардни алгоритми за оцену квалитета слике нису поуздани у оцењивању слика са изобличењима због ДИБР синтезе.

1.3 Допринос истраживања

За оцену квалитета ДИБР синтетизованих слика у овом раду је предложена метрика заснована на мултирезолуционој декомпозицији применом морфолошких филтара. Експериментално је доказано да човеков визуелни систем процесира информације на мултирезолуциони начин. Примена мултирезолуционе декомпозиције слике у метрици за оцењивања квалитета слике омогућује боље поклапање резултата метрике са субјективним оценама. Морфолошка мултирезолуциона декомпозиција је структурна декомпозиција слике којом се издвајају ивице различитих величина на различитим нивоима декомпозиције [16]. Постоји подударност између шеме морфолошке мултирезолуционе декомпозиције и човекове визуелне перцепције [16]. Применом нелинеарних морфолошких филтара у мултирезолуционој декомпозицији важне геометријске информације као што су ивице остају очуване на њиховим оригиналним позицијама и без замућења у сликама свих нивоа декомпозиције [17]. Израчунавањем средње квадратне грешке између одговарајућих подопсега мултирезолуционих репрезентација референтне и синтетизоване слике у којима су издвојене ивице пиксел по пиксел, прецизно се мери разлика између две репрезентације слике. Тако се у оцени синтетизоване слике највећи значај додељује дисторзији ивица која је карактеристична за синтетизоване слике а значајно нарушава визуелни квалитет слике. Средње квадратне грешке више подопсега мултирезолуционих репрезентације слике се комбинују у резултујућу оцену.

Анализиране су две верзије метрике засноване на мултирезолуционој декомпозицији слике применом морфолошких филтара: Morphological Pyramid Peak Signal-to-Noise Ratio, MP-PSNR [18], заснована на пирамидалној декомпозицији слике применом морфолошких оператора [19] и Morphological Wavelet Peak Signal-to-Noise Ratio, MW-PSNR [20] заснована на декомпозицији слике применом морфолошких таласића [21]. Испитан је утицај два типа морфолошких таласића у сепарабилној декомпозицији на поузданост метрике. Утицај несепарабилне декомпозиције применом морфолошких таласића са узорковањем на “quincunx“ решетки на поузданост метрике је такође испитан.

Обе верзије метрике, MP-PSNR и MW-PSNR, се добро поклапају са субјективним оценама људи, знатно боље од стандардних метрика за оцену квалитета слика и знатно боље од

других метрика намењених оцењивању квалитета синтетизованих слика. Како су морфолошки оператори који се користе у мултирезолуционој декомпозицији слике једноставни за израчунавање као и средња квадратна грешка која се користи као мера изобличења подопсега мултирезолуционе репрезентације слике, рачунска комплексност морфолошких мултирезолуционих мера није велика.

Анализиране су перформансе метрике PSNR по подопсезима мултирезолуционе декомпозиције применом морфолошких филтара. Показано је да се метрика PSNR добро поклапа са субјективним оценама људи када се израчуна за подопсеге синтетизоване слике виших нивоа морфолошке декомпозиције. Перформансе метрике PSNR израчунате за подопсеге виших нивоа декомпозиције су знатно боље од перформанси метрике PSNR израчунатих за подопсеге нижих нивоа декомпозиције. Због тога су предложене редуковане верзије мултирезолуционе метрике са морфолошким филтрима, MP-PSNR_r [22] и MW-PSNR_r [20], коришћењем само подопсега виших нивоа декомпозиције. Перформансе редукованих верзија мултирезолуционе метрике са морфолошким филтрима су боље од перформанси основних верзија.

Предложене верзије метрике су коришћене за оцену квалитета фрејмова синтетизованих видео секвенци као први корак у оцени квалитета 3Д садржаја и за оцену квалитета ФТВ видео секвенци. Метрике су тестиране на базама слика које садрже само изобличења због процеса синтезе као и на бази ФТВ видео секвенци са навигацијом кроз различите позиције које садрже изобличења и због компресије мапа дубине и због процеса синтезе [23]. Предложене метрике показују добро поклапање са субјективним оценама када се користе за оцену изобличења због ДИБР синтезе. Могу се користити за поређење алгоритама ДИБР синтезе.

1.4 Организација дисертације

Дисертација садржи осам поглавља. После увода, у другом поглављу је преглед метрика за оцену слика и видео секвенци синтетизованих применом алгоритама ДИБР синтезе. Треће поглавље садржи преглед мултирезолуционих метрика за оцену квалитета слика и преглед општих шема редундантне и нередундантне мултирезолуционе декомпозиције сигнала. У четвртом поглављу је описана предложена мултирезолуциона метрика за оцену

квалитета ДИБР синтетизованих слика. Прво су описане декомпозиције слике морфолошком пирамидом и морфолошким таласићима које су анализиране за примену у првој фази израчунавања предложене мултирезолуционе метрике. Затим је описано израчунавање дисторзије по подопсезима декомпоноване слике за оба типа декомпозиције. Анализирана је сложеност израчунавања метрике. У петом поглављу су описане три базе ДИБР синтетизованих слика/видео секвенци коришћене за тестирање предложене метрике и приказане су перформансе предложене метрике тестиране на описаним базама слика/видео секвенци. У шестом поглављу је закључак. Седмо поглавље садржи преглед коришћене литературе. Последње поглавље садржи списак слика, табела и скраћеница.

2 ПРЕГЛЕД МЕТРИКА ЗА ОЦЕНУ КВАЛИТЕТА ДИБР СИНТЕТИЗОВАНИХ СЛИКА/ВИДЕА

Већина стандардних метрика за оцену квалитета слика су намењене за други тип изобличења најчешће распрострањених по целој слици и нису погодне за оцену квалитета слика које садрже изобличења због ДИБР синтезе углавном у областима “дисклузија”. Поузданост стандардних 2Д метрика за оцену квалитета синтетизованих слика које садрже само изобличења због ДИБР синтезе је тестирана у [14]. За тестирање је коришћена база слика које садрже само изобличења због процеса ДИБР синтезе [24]. Резултати показују да стандардне 2Д метрике нису поуздане у оцењивању квалитета синтетизованих слика. Стандардне 2Д метрике су тестирание и за оцену квалитета синтетизованих видео секвенци које садрже само изобличења због ДИБР синтезе у [25]. Ни једна од тестираних метрика не достиже ни 50% корелације са субјективним оценама када се користе за оцену квалитета синтетизованих слика и видеа. Постоје два главна разлога због којих конвенционалне 2Д методе нису довољно добре у оцењивању квалитета синтетизованих слика/видеа [15]. Прво, конвенционалне објективне метрике потцењују доминантне дисторзије у синтетизованим сликама/видеу као што су дисторзије ивица, местимичан померај објеката и светлугање које су веома уочљиве људима. Конвенционалне метрике квалитета слике које третирају све регионе слике једнако потцењују ефекат оштећења ивица и недовољно га приказују у оцени. Код оцењивања квалитета синтетизованих слика оштећење ивица треба да има значајан утицај на оцену објективне метрике да би се она боље подударала са оценом људи. Друго, конвенционалне објективне метрике прецењују изобличења као што су мале геометријске дисторзије, сталан помак објеката, шум настао приликом снимања са више камера на различитим позицијама и разлике у осветљају (illumination) које су једва приметне људима. Ове дисторзије ће утицати на смањење вредности PSNR синтетизоване слике али неће смањити визуелни квалитет слике. Субјективни (визуелни) квалитет синтетизованих слика које садрже објекте који су померени је висок.

С обзиром да стандардне метрике нису довољно поуздане у оцењивању ДИБР синтетизованих слика потребне су нове метрике посебно дизајниране за оцену квалитета

ДИБР синтетизованих слика. Неколико метрика је дизајнирано за оцену квалитета синтетизованих слика.

Индикатор структурне деградације заснован на анализи ивица је предложен у [26]. Индикатор се израчунава на основу помераја контура у синтетизованој слици. За неколико синтетизованих слика је приказано добро поклапање индикатора са субјективним оценама.

VSQA (View Synthesis Quality Assessment) метрика је намењена детекцији изобличења у синтетизованим сликама [27]. Метрика израчунава 3 мапе које описују комплексност текстуре, разноликост оријентације градијента и присуство великог контраста. Метрика се може дефинисати тако да буде проширење било које метрике за оцену квалитета 2Д слика. У раду је приказана VSQA-SSIM као екстензија метрике SSIM. Предложена метрика тестирана на бази слика [24] показује корелисаност са субјективним оценама од 61.42% што је за 17.8% боље од корелисаности SSIM метрике.

3DSwIM (3D Synthesized view Image Quality Metric) метрика је намењена детекцији изобличења у синтетизованим сликама [28]. Након поделе на блокове, ради се регистрација која омогућује да се пореде блокови из две слике, референтне и синтетизоване, који се најбоље поклапају. На основу претпоставке да су изобличења још уочљивија уколико су објекти на слици људи, метрика садржи корак детекције коже који додељује веће тежине блоковима слике који садрже оштећења на лицу, врату, итд. Након првог нивоа декомпозиције Харовим таласићима, деградација слике се мери анализом статистичких варијација у подопсегу са хоризонталним детаљима. Дисторзија по блоковима се израчунава на основу хистограма осветљаја. Непопуњене области које одговарају “дисоклузијама“ су дуж вертикалних ивица на слици. Када се попуне те области садрже високофреквентне компоненте у хоризонталном правцу. У раду је приказано да метрика постиже корелисаност 76.17% са субјективним оценама када се користи за оцену квалитета слика из базе [24].

Према студији о оцењивању квалитета синтетизованих слика/видео у MVD формату квалитет синтетизованог садржаја се може оцењивати на 3 начина [29]. Први начин је израчунавање метрике између видео секвенце синтетизоване на основу некомпримованих

података и оригиналне видео секвенце. На тај начин се оцењују ефекти ДИБР синтезе. Други начин је израчунавање метрике између видео секвенци синтетизоване на основу некомпримоване видео секвенце и компримоване секвенце мапа дубине и видео секвенце синтетизоване на основу некомпримованих података. На тај начин се оцењује ефекат компресије мапе дубине. Трећи начин је израчунавање метрике између видео секвенце синтетизоване на основу некомпримоване видео секвенце и компримоване секвенце мапа дубине и оригиналне видео секвенце. Тако се оцењује ефекат и ДИБР синтезе и компресије мапе дубине комбиновано. За оцену квалитета синтетизованих видео секвенци је коришћена PSNR метрика. Показано је да процес синтезе уноси велике дисторзије у синтетизовану видео секвенцу које маскирају дисторзије због компресије мапе дубине. Закључак је да када треба оценити перформансе кодека, за референтну секвенцу треба користити секвенцу синтетизовану на основу некомпримованих података.

Поузданост стандардних 2Д метрика за оцену квалитета синтетизованих видео секвенци које садрже изобличења због компресије мапе дубине је тестирана у [30]. За тестирање је коришћена база видео секвенци које не садрже временске дисторзије јер сви фрејмови видео секвенце одговарају истом временском тренутку [31]. Видео секвенце симулирају кретање камере код слободне навигације ФТВ. Свака видео секвенца садржи фрејмове синтетизоване на 49 позиција између 2 позиције на којима су снимљене видео секвенце. Фрејмови су синтетизовани на основу некомпримованих текстура и компримованих мапа дубине. Резултати показују да стандардне 2Д метрике нису поуздане у оцењивању квалитета ФТВ видео секвенци синтетизованих слика.

Бољу поузданост од стандардних метрика у оцењивању квалитета дела FVSV базе [31] која садржи видео секвенце са изобличењима због компресије мапа дубине и ДИБР синтезе показује метрика предложена у [32]. Метрика садржи процедуру регистрације за поравнање између синтетизоване и оригиналне слике. Користећи процедуру детекције коже метрика додељује веће тежинске факторе блоковима који садрже лице и врат особа јер је дисторзија уочљивија на људима него на другим деловима сцене. Да би задржали просторну локализацију изобличења синтетизовани и референтни фрејм су декомпоновани таласићима (wavelets) по блоковима. Деградација се израчунава на основу хистограма осветљаја хоризонталних подопсега првог нивоа декомпозиције синтетизоване

и референтне слике. Метрика је израчуната на делу базе који садржи видео секвенце синтетизоване на основу некомпримованих видео секвенци (текстура) и компримоване секвенци мапа дубине. Фрејмови видео секвенце су синтетизовани применом мода мешање (blend) тако да се слике са обе позиције (лева и десна) користе за попуњавање празнина у слици. За компресију мапа дубине коришћена су 3 алгоритма кодовања: HEVC, H264 и JPEG2000.

Метрика за оцену квалитета синтетизованих видео секвенци које садрже изобличења због компресије мапе дубине и/или текстуре је представљена у [15]. Метрика се фокусира на временску дисторзију светлцање због компресије мапе дубине и процеса синтезе. Састоји се од две мере квалитета из просторног и временског домена. Једна мери временску дисторзију светлцање која настаје због компресије мапе дубине и процеса синтезе. Друга мери просторно-временске промене, замућење и блок ефекат у синтетизованој секвенци због компресије текстуре. Резултати показују да предложена метрика има добре перформансе при оцењивању базе ДИБР синтетизованих видео секвенци СИАТ која је представљена у истом раду, боље од перформанси стандардних метрика.

Метрика за оцену квалитета слика које садрже изобличења ДИБР синтезе због грешака у мапи дубине предложена је у [33]. Грешке у мапи дубине узрокују хоризонталне помаке објеката у синтетизованој слици. Стандардне метрике квалитета које израчунавају резултат пиксел по пиксел оцењују да је таква слика са помереним објектима смањеног квалитета док је визуелни квалитет слике који оцењују субјекти очуван. Предложена метрика најпре елиминише сталан помак унутар објеката слике а затим израчунава квалитет слике применом стандардне 2Д метрике SSIM. За одређивање помака користи се алгоритам поклапања блокова (block-matching) у правцу x -осе. Након компензације помака објеката, резултат 2Д метрике квалитета слике се боље поклапа са субјективним оценама. Други део мери несталан помак објеката дуж ивица и израчунава структурну оцену користећи Хаусдорф удаљеност карактеристичних тачака издвојених Канијевим детектором ивица по блоковима синтетизоване и референтне слике. У тестирању су коришћене MPEG MVD видео секвенце. Мапе дубине су изобличене тако да садрже изобличења због квантизације, Гаусовог шума, офсета. Креирана су 42 стерео пара тако да

је једна слика стерео пара оригинал а друга синтетизована. Предложена метрика се боље поклапа са субјективним оценама од стандардних метрика.

Метрика CSED (Color and Sharpness of Edge Distortion) [34] за оцену квалитета ДИБР синтетизованих слика се састоји из две компоненте. Једна компонента мери дисторзију осветљености у области “дисоклузија”. Друга компонента мери изобличења ивица око области “дисоклузија”. Метрика је намењена за оцену квалитета стерео 3Д видеа у формату SVD (Single View Video) који садржи једну видео секвенцу и секвенцу мапа дубине.

Зависност квалитета 3Д стерео видеа од квалитета мапа дубине које су коришћене у процесу ДИБР синтезе је анализирана у [35]. У експериментима су коришћене стандардне метрике за оцену квалитета слика и видео секвенце које су препоручене од MPEG. Стерео видео је формиран тако да је лева видео секвенца оригинална а десна видео секвенца синтетизована. За синтезу десне слике стерео пара коришћена је оригинална лева слика и мапа дубине са изобличењима због компресије и због грешака у преносу. За симулацију изобличења компресије коришћен је HEVC кодер. За симулацију изобличења због губитка пакета, коришћен је симулатор преноса кроз мрежу који користи H.264/AVC алгоритам кодовања. Резултати показују да квалитет 3Д видеа значајно зависи од квалитета мапе дубине али да само квалитет мапа дубине није довољан за предикцију квалитета 3Д видеа.

Однос квалитета синтетизоване слике и квалитета мапе дубине која се користи за синтезу је анализиран у раду [36]. Предложена је нова метрика за оцену квалитета мапа дубине заснована на моделу дисторзија који априксимира грешке синтезе због грешака пиксела у мапи дубине. Резултат предложене метрике је добро корелисан са оценом квалитета синтетизоване слике, боље него резултати стандардних метрика. Дата је илустрација примене предложене метрике уградњом у кодер у фазу избора мода кодовања (intra frame, inter frame). Тако се постиже уштеда до 30% битског протока у поређењу са традиционалном шемом избора мода кодовања која се заснива на суми квадрата грешака.

Оцена квалитета синтетизованих видео секвенци метрикама које су проширење стандардних метрика као што су PSNR и SSIM информацијама из мапе дубине је предложена у [37]. Метрика додељује веће тежинске факторе регионима слике који су

подложни изобличењима због процеса синтезе. Метрика узима у обзир и временску дисторзију видео секвенце. У тестовима су оцењиване видео секвенце синтетизоване на основу колор слика и мапа дубине које садрже изобличења. Перформансе предложене методе су мерене коришћењем корелације са VQM метриком која је високо корелисана са субјективним оценама. Предложена метода показује боље перформансе од PSNR и SSIM.

Метод за оцену квалитета стерео 3Д видеа (3D Video Quality Measure, 3VQM) генерисаног применом ДИБР је представљен у [38]. Метод се заснива на концепту идеалне мапе дубине која омогућује ДИБР синтезу слике без дисторзија. На основу естимиране идеалне мапе дубине и реалне мапе дубине изводе се две мере визуелних дисторзија стерео видеа: ТО (temporal outliers) и СО (spatial outliers) које оцењују временске и просторне варијације мапе дубине које узрокују несклад леве и десне слике, брзе промене разлика и геометријске дисторзије. Због грешака у реалној мапи дубине, пиксели (блокови) колор слике се пројектују на погрешне позиције у синтетизованој слици. Визуелно се те грешке манифестују просторно и временски. За геометријске дисторзије користи се мера СО која се израчунава као стандардна девијација разлике идеалне и реалне мапе дубине. Временска варијација разлике идеалне и реалне мапе дубине по фрејмовима је индикатор визуелне дисторзије. У области текстуре дисторзије се уочавају као значајне промене интензитета а око равних области као светлуцање. Мера ТО се израчунава као стандардна девијација промене разлика идеалне и реалне мапе дубине за два суседна фрејма. Брзе промене диспаритета се визуелно манифестују као светлуцање. За овај облик дисторзије је предложена мера ТИ (temporal inconsistencies) која се израчунава као стандардна девијација разлике реалних мапа дубине два суседна фрејма. Комбинацијом три мере, СО, ТО и ТИ, добија се 3VQM. Резултати показују да се предложена мера добро поклапа са субјективним оценама и има боље перформансе од метрика PSNR и SSIM.

Метрика за оцену квалитета 3Д видеа у СВД формату са редукованом референцом се заснива на чињеници да ивице/контуре мапе дубине могу представљати различите нивое дубине и могу се користити за мерење структурне деградације [39]. Пошто се ивице у мапи дубине поклапају са одговарајућим ивицама објеката колор слике (текстуре), ивице мапе дубине и ивице текстуре се пореде при одређивању индекса квалитета структурне деградације слика видео секвенце. Предложена метрика упоређује ивице оригиналне и

мапе дубине са дисторзијама и мери структурну деградацију мапе дубине. На основу информација о ивицама у мапи дубине могу се идентификовати објекти слике у одговарајућој слици (текстури). Бинарна мапа ивица се преноси са предајне на пријемну страну са мањим оптерећењем ресурса него да се преноси референтна колор слика. Преносе се и параметри везано за осветљење и контраст оригиналне и мапе дубине са изобличењем. SSIM метрика је коришћена за поређење структурних разлика коришћењем информација о ивицама. Предложена метода постиже добре резултате у поређењу са метриком која користи целу референцу.

3 МУЛТИРЕЗОЛУЦИОНА РЕПРЕЗЕНТАЦИЈА СЛИКЕ

Мултирезолуционе методе се примењују у случајевима када информације од интереса постоје на више скала. Анализом слике на само једној скали не би добили информације о слици које постоје на другим скалама. У таквим случајевима треба анализирати слику по свим скалама. Значај анализе слике кроз више скала (резолуција) потиче од природе самих слика [40]. Сцена садржи објекте разних величина а објекти садрже различите карактеристике. Објекти могу бити на различитим удаљеностима од посматрача. Мултирезолуционе методе се користе у видео праћењу, проблемима подударности слика као што је стерео визија, сегментацији, претраживању (налажењу) слика по садржају и другим претраживањима. Експериментима је утврђено да се мултирезолуциона трансформација одвија у кори великог мозга сисара [41].

Као и у другим областима обраде и анализе слике, и у области оцењивања квалитета слике постиже се побољшање перформанси метрике применом мултирезолуционе декомпозиције слике пре израчунавања изобличења/квалитета. У метрикама које су намењене за оцену квалитета слика које садрже изобличења због компресије и преноса користе се пирамидалне декомпозиције и декомпозиције таласићима.

3.1 ПРЕГЛЕД МУЛТИРЕЗОЛУЦИОНИХ МЕТРИКА ЗА ОЦЕНУ КВАЛИТЕТА СЛИКА

Мултирезолуциона метрика структурне сличности MS-SSIM (Multi-scale Structural Similarity) [42] садржи пирамидалну декомпозицију Гаусовог типа [43] референтне слике и слике са изобличењем применом линеарних филтара. Између одговарајућих слика пирамидалних репрезентација референтне и изобличене слике израчунавају се мере структурне сличности које се комбинују у резултујућу меру MS-SSIM. Перформансе MS-SSIM су боље од перформанси SSIM када се користи за оцењивање слика са изобличењима због компресије и преноса.

Пирамидална декомпозиција Гаусовог типа је коришћена и у метрици MS-PSNR [44]. Уместо Гаусових филтара примењени су једноставни линеарни филтри којима се израчунава средња вредност пиксела на прозору 2x2 који су коришћени и код метрике MS-SSIM. Након понављања филтрирања и узорковања кроз више нивоа декомпозиције, високофреквентне компоненте су углавном потиснуте и нискофреквентне копије

референтне и изобличене слике постају све сличније. Након декомпозиције референтне слике и слике са изобличењем на низ слика све нижих резолуција, израчунавају се средње квадратне грешке између слика две пирамиде исте резолуције са истог нивоа декомпозиције почев од другог нивоа декомпозиције. Комбиновањем MSE са различитих нивоа пирамиде добија се MS-MSE која се конвертује у MS-PSNR. Метрика је тестирана на бази стерео слика LIVE3D које садрже изобличења због компресије и преноса и замућење. Метрика остварује знатно боље перформансе од PSNR и SSIM а сличне перформансе као MS-SSIM од које је рачунски вишеструко ефикаснија.

Мултирезолуционе метрике са расподелом тежинских фактора према информационом садржају (Information content Weighted Structural Similarity, IW-SSIM, и Information content Weighted Peak Signal-to-Noise Ratio, IW-PSNR) [45] садрже Лапласову пирамидалну декомпозицију [43] референтне слике и слике са изобличењем. Између одговарајућих слика две пирамиде, израчунавају се мере структурне сличности за метрику IW-SSIM, односно средње квадратне грешке за метрику IW-PSNR. Комбинацијом мера по подопсезима израчунава се крајња мера. Метрике су тестиране на базама слика које садрже стандардне типове изобличења због компресије, преноса, замућење, шум итд. Остварено је мало побољшање перформанси увођењем тежинских фактора према информационом садржају (повећање корелације применом IW-SSIM у односу на MS-SSIM је 0.3% - 4.3 %). Велико побољшање перформанси је остварено применом Лапласове пирамидалне декомпозиције референтне и слике са изобличењем пре израчунавања PSNR по сликама пирамида и применом тежинских фактора према информационом садржају (IW-PSNR остварује повећање корелације 5.1% - 26.2% у односу на PSNR).

Мултирезолуциона метрика за оцену квалитета слике која користи декомпозицију референтне и оригиналне слике применом оператора ДоГ (difference of Gaussian, DoG) представљена је у [46]. На сваком нивоу резолуције, креира се скуп слика по више скала применом ДоГ оператора. Скуп слика ДоГ структуре се израчунава одузимањем слика добијених филтрирањем полазне слике Гаусовим кернелом са различитим (скалираним) стандардним девијацијама (скалама). Полазна слика следеће октаве се добија узорковањем слике последње скале претходне октаве. Оцена квалитета слике се креира поређењем ДоГ структуре референтне и изобличене слике различитих октава и скала. Код креирања оцене узимају се у обзир и тежински фактори који зависе од информационог садржаја слике. За

сваки пар ДоГ слика израчунава се мапа тежинских коефицијената. Ниже скале ДоГ репрезентације су осетљиве на дисторзије слике које утичу на високофреквентне компоненте као што је бели Гаусов шум. Више скале ДоГ репрезентације су осетљиве на дисторзије слике које утичу на нисокофреквентне компоненте. Зато се може рећи да је ДоГ флексибилна репрезентација структуре слике. Метода је тестирана на базама слика које садрже изобличења због компресије и преноса, замућење и шумове. Метод показује боље перформансе од осталих тестираних метрика.

Декомпозиција слике Харовим таласићима је примењена у метрици намењеној за оцену квалитета слика које садрже изобличења због компресије [47]. Експериментално је показано да декомпозиција Харовим филтрима успешно симулира декомпозицију човековог визуелног система.

Харови таласићи се користе у декомпозицији слике унутар метрике за оцену квалитета слике предложене у [48]. Харови филтри су изабрани због једноставности и због тога што омогућују тачније резултате метрике него други таласићи. Мапе грешака се израчунавају за сваки подопсег детаља на сваком нивоу декомпозиције и њиховим сумирањем се добија збирна мапа грешака. Између збирних мапа грешака оригиналне слике и слике са изобличењем израчунава се PSNR. PSNR се израчунава и између слика апроксимација оригиналне и слике са изобличењем на последњем нивоу декомпозиције. Линеарном комбинацијом ова два PSNR добија се крајња оцена. Метрика је тестирана на бази слика које садрже изобличења због компресије и преноса. Метрика израчунава квалитет слике тачније од класичне PSNR метрике и може се користити у апликацијама у реалном времену.

3.2 ОПШТЕ ШЕМЕ МУЛТИРЕЗОЛУЦИОНЕ ДЕКОМПОЗИЦИЈЕ СЛИКЕ

Дат је преглед два типа шема за мултирезолуциону декомпозицију слике: редувантне пирамидалне декомпозиције слике и нередувантне декомпозиције применом таласића (wavelets). Пирамиде и таласићи су најчешће коришћени алат за мултирезолуциону декомпозицију слика.

3.2.1 Редундантна мултирезолуциона репрезентација слике

Редундантна мултирезолуциона репрезентација слике садржи већи број пиксела од саме слике. Шема једног нивоа редундантне мултирезолуционе декомпозиције, слика 3.1.1.а, [19], садржи један оператор за анализу сигнала $\Psi_j^\uparrow : V_j \rightarrow V_{j+1}$, тако да важи $s_{j+1} = \Psi_j^\uparrow(s_j)$ и један оператора за синтезу сигнала $\Psi_j^\downarrow : V_{j+1} \rightarrow V_j$, тако да важи $\Psi_j^\downarrow(s_{j+1}) = \hat{s}_j$. V_j је простор сигнала на нивоу j , s_j је сигнал на нивоу декомпозиције j а \hat{s}_j је апроксимација сигнала. Применом оператора анализе на сигнал s_j отклањају се информације из сигнала које нису обавезно везане за фреквенцију. Геометријске информације могу бити отклоњене такође, у зависности од изабраног оператора анализе. Применом оператора анализе на сигнал а затим оператора синтезе на резултујући сигнал добија се сигнал апроксимације $\hat{s}_j = \Psi_j^\downarrow \Psi_j^\uparrow(s_j) \in \hat{V}_j$. Оператор синтезе се бира тако да је сигнал апроксимације што сличнији сигналу који се декомпонује. Сигнал грешке $d_j = s_j - \hat{s}_j \in Y_j$ садржи информације о сигналу s_j које нису присутне у сигналу апроксимације, \hat{s}_j , где је Y_j простор сигнала детаља на нивоу j . Оператори анализе и синтезе треба да задовоље пирамидални услов према коме применом оператора синтезе на сигнал а затим оператора анализе на резултујући сигнал поново се добија полазни сигнал, $\Psi_j^\uparrow \Psi_j^\downarrow = id$, где је id оператор идентитета. Рекурзивна шема анализе сигнала за редундантну репрезентацију сигнала са M нивоа декомпозиције је:

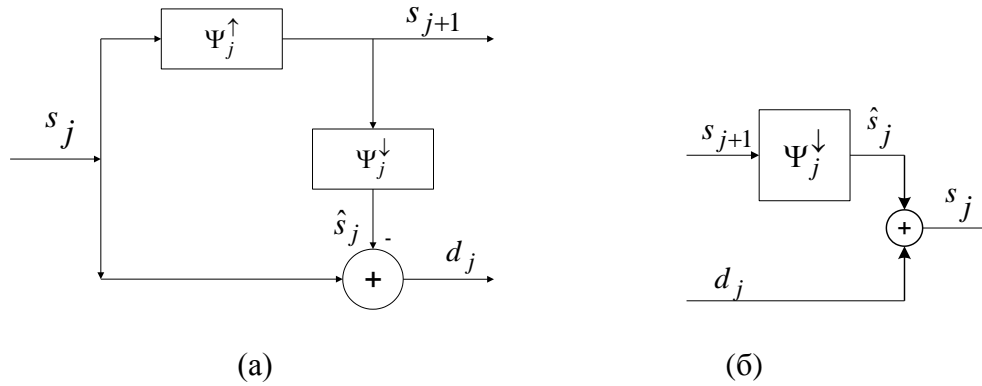
$$x \rightarrow (d_0, s_1) \rightarrow (d_0, d_1, s_2) \rightarrow \dots \rightarrow (d_0, d_1, \dots, d_{j-1}, s_j) \rightarrow \dots \rightarrow (d_0, d_1, \dots, d_{M-1}, s_M) \quad (3.1)$$

Применом шеме редундантне мултирезолуционе декомпозиције, сигнал се декомпонује на низ сигнала апроксимација $\{s_1, s_2, \dots, s_M\}$ и низ сигнала грешака (детаља)

$\{d_0, d_1, \dots, d_{M-1}\}$ све нижих резолуција.

$$s_0 = x \in V_0, \quad s_{j+1} = \Psi_j^\uparrow(s_j) \in V_{j+1}, \quad 0 \leq j \leq M-1 \quad (3.2)$$

$$d_j = s_j - \Psi_j^\downarrow(s_{j+1}) \in Y_j \quad (3.3)$$



Слика. 3.1.1 Редундантна мултирезолуциона репрезентација: (а) шема декомпозиције са једним оператором за анализу и једним оператором за синтезу; (б) шема реконструкције са једним оператором синтезе

Сигнал x се може перфектно реконструисати на основу низа сигнала детаља $d_j, j=0, 1, \dots, M-1$ и сигнала апроксимације најниже резолуције s_M применом рекурзије уназад. На слици 3.1.1.б приказана је шема једног нивоа реконструкције на основу сигнала апроксимације ниже резолуције s_{j+1} и сигнала детаља d_j . Реконструисани сигнал s_j се добија сабирањем сигнала детаља d_j и сигнала добијеног применом оператора синтезе на сигнал апроксимације следећег нивоа мултирезолуционе репрезентације ниже резолуције s_{j+1} :

$$s_j = \Psi_j^\downarrow(s_{j+1}) + d_j, \quad M-1 \geq j \geq 0, \quad x = s_0 \quad (3.4)$$

Могућност перфектне реконструкције слике на основу њене мултирезолуционе репрезентације је корисна особина за примену у оцењивању квалитета слике јер гарантује да нема губитка информација и ако се две слике разликују, разликују се и њихове мултирезолуционе репрезентације [49].

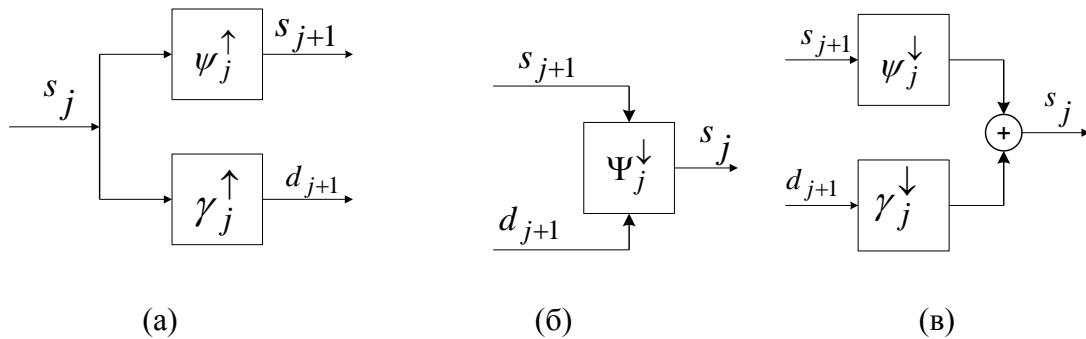
Најпознатији пример редундантне мултирезолуционе шеме је Лапласова пирамида [43]. Лапласова пирамида је погодна за анализу слике по скалама. Сlike детаља Лапласове пирамиде су копије оригиналне слике филтриране појасно пропусним филтром са истакнутим ивицама распоређеним по величини тако да су фини детаљи издвојени у сликама нижих нивоа декомпозиције а грубљи детаљи у вишим. Паралелно са

генерисањем Лапласове пирамиде, генерише се и Гаусова пирамида која садржи низ слика апроксимација све нижих резолуција. Сlike Гаусове пирамиде су копије оригиналне слике генерисане филтрирањем оригиналне слике нископропусним филтром и децимацијом након филтрирања по нивоима декомпозиције.

Пирамидална репрезентација има доста заједничког са начином на који људи посматрају свет, тј. визуелни систем примата је мултирезолуционог карактера [49]. Пирамидална репрезентација описује слику по више просторних резолуција уз очување локалне просторне структуре која нам омогућује да видимо слику на свакој скали [50].

3.2.2 Нередувантна мултирезолуциона репрезентација слике

Нередувантна мултирезолуциона репрезентација слике садржи једнак број пиксела као сама слика. Шема једног нивоа нередувантне мултирезолуционе декомпозиције [21] садржи два оператора анализе, слика 3.2.1.а.



Слика 3.2.1. Нередувантна мултирезолуциона репрезентација: (а) шема декомпозиције са два оператора за анализу (б) шема реконструкције са једним оператором синтезе (в) шема реконструкције са два оператора синтезе

Применом једног оператора анализе $\psi_j^\uparrow : V_j \rightarrow V_{j+1}$ (V_j је простор сигнала на нивоу j) на сигнал s_j добија се апроксимација сигнала ниже резолуције $s_{j+1} = \psi_j^\uparrow(s_j)$. Применом другог оператора анализе $\gamma_j^\uparrow : V_j \rightarrow Y_{j+1}$ (Y_j је простор детаља на нивоу j) на сигнал добија

се сигнал детаља ниже резолуције, $d_{j+1} = \gamma_j^\uparrow(s_j)$. Рекурзивна шема анализе сигнала за нередундантну репрезентацију сигнала са M нивоа декомпозиције је:

$$x \rightarrow (d_1, s_1) \rightarrow (d_1, d_2, s_2) \rightarrow \dots \rightarrow (d_1, d_2, \dots, d_j, s_j) \rightarrow \dots \rightarrow (d_1, d_2, \dots, d_M, s_M) \quad (3.5)$$

Нередундантна пирамидална декомпозиција се може представити:

$$s_0 = x \in V_0, \quad s_{j+1} = \psi_j^\uparrow(s_j) \in V_{j+1}, \quad 0 \leq j \leq M-1 \quad (3.6)$$

$$d_{j+1} = \gamma_j^\uparrow(s_j) \in Y_{j+1} \quad (3.7)$$

Сигнал се може реконструисати на основу нередундантне мултирезолуционе репрезентације применом шеме реконструкције са једним или са два оператора синтезе [21].

Шема реконструкције са једним оператором синтезе Ψ_j^\downarrow је приказана на слици 3.2.1.б. Да би била могућа перфектна реконструкција треба да важи:

$$\Psi_j^\downarrow(\psi_j^\uparrow(s), \gamma_j^\uparrow(s)) = s, \quad s \in V_j, \quad \Psi_j^\downarrow: V_{j+1} \times Y_{j+1} \rightarrow V_j \quad (3.8)$$

Да би се избегла редундантност, за оператор синтезе треба да важи:

$$\psi_j^\uparrow(\Psi_j^\downarrow(s, d)) = s, \quad s \in V_{j+1}, \quad d \in Y_{j+1} \quad (3.9)$$

$$\gamma_j^\uparrow(\Psi_j^\downarrow(s, d)) = d, \quad s \in V_{j+1}, \quad d \in Y_{j+1} \quad (3.10)$$

Сигнал x се може реконструисати рекурзијом уназад:

$$s_j = \Psi_j^\downarrow(s_{j+1}, d_{j+1}), \quad M-1 \geq j \geq 0, \quad x = s_0 \quad (3.11)$$

Шема реконструкције са два оператора синтезе, слика 3.2.1.в, је специјални случај шеме са једним оператором синтезе [21]. Шема садржи оператор синтезе сигнала $\psi_j^\downarrow: V_{j+1} \rightarrow V_j$,

$\psi_j^\downarrow(s_{j+1}) = \hat{s}_j$ и оператор синтезе детаља $\gamma_j^\downarrow: Y_{j+1} \rightarrow V_j$, $\gamma_j^\downarrow(d_{j+1}) = \hat{d}_j$ тако да важи:

$$\Psi_j^\downarrow(s_{j+1}, d_{j+1}) = \psi_j^\downarrow(s_{j+1}) + \gamma_j^\downarrow(d_{j+1}) \quad (3.12)$$

Да би била могућа перфектна реконструкција сигнала, за операторе анализе и синтезе треба да важи:

$$\psi_j^\downarrow \psi_j^\uparrow(s) + \gamma_j^\downarrow \gamma_j^\uparrow(s) = s, \quad \forall s \in V_j \quad (3.13)$$

Сигнал x се може реконструисати рекурзивним поступком уназад:

$$s_j = \psi_j^\downarrow(s_{j+1}) + \gamma_j^\downarrow(d_{j+1}), \quad M-1 \geq j \geq 0, \quad x = s_0 \quad (3.14)$$

Да би се избегла редундантност треба да важи:

$$\psi_j^\uparrow(\psi_j^\downarrow(s) + \gamma_j^\downarrow(d)) = s, \quad s \in V_{j+1}, d \in Y_{j+1} \quad (3.15)$$

$$\gamma_j^\uparrow(\psi_j^\downarrow(s) + \gamma_j^\downarrow(d)) = d, \quad s \in V_{j+1}, d \in Y_{j+1} \quad (3.16)$$

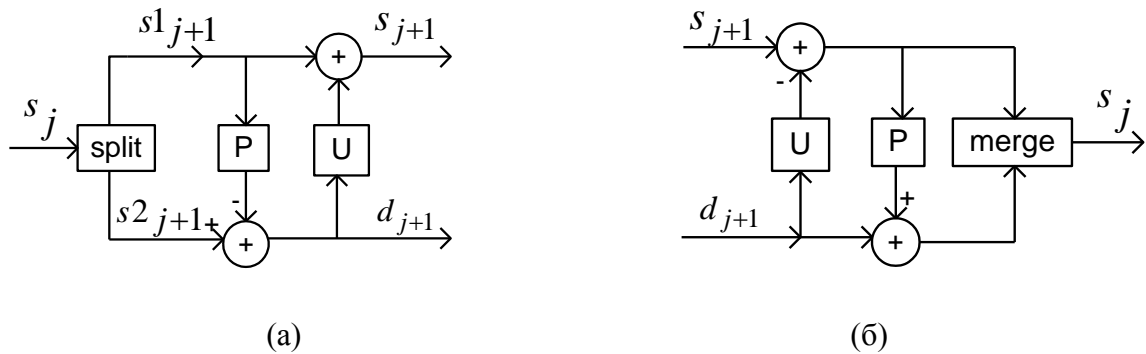
Декомпозиција сигнала $x \in V_0$ на низ сигнала $(d_1, d_2, \dots, d_M, s_M)$ је трансформација таласићима (wavelet transformation). Трансформацијом таласићима сигнал се разлаже на различите фреквенцијске компоненте.

Декомпозицијом слике таласићима добија се репрезентација слике која садржи исти број пиксела као оригинална слика. То је последица чињенице да се сигнал на нивоу j реконструише на основу сигнала апроксимације и сигнала детаља на нивоу $j+1$ ниже резолуције. Трансформацијом слике применом таласића добија се мултирезолуциона репрезентација слике која омогућује декорелацију у просторном и фреквентном домену. Користи се у многим областима обраде слике као што су компресија слике и отклањање шума.

Лифтинг шема

Основни алат за конструкцију таласића прве генерације је Фуријеова трансформација. Друга генерација таласића омогућује ефикасну трансформацију сигнала на основу лифтинг шеме [51] у просторном домену за разлику од традиционалног начина израчунавања у фреквентном домену. Свака трансформација сигнала таласићима се може

реализовати помоћу лифтинг шеме [52]. Једна од великих предности реализације трансформације таласићима помоћу лифтинг шеме је да се филтрирање реализује кроз веома једноставне кораке и са мањом рачунском сложености у поређењу са стандардним алгоритмима филтрирања [51]. У лифтинг шеми оператори анализе и синтезе су модификовани тако да су особине модификоване шеме боље од оригиналне. Лифтинг шема омогућује израчунавање трансформације таласићима без алокације додатне меморије. Лифтинг шема је увек инвертибилна и омогућује перфектну реконструкцију. Лифтинг шема омогућује конструкцију различитих декомпозиција линеарним и нелинеарним таласићима.



Слика 3.2.2 Трансформација сигнала таласићима применом лифтинг шеме садржи три корака: подела сигнала (split), процена (predict, P) и поправка сигнала (update, U). (а) декомпозиција (б) инверзна трансформација, синтеза.

Један ниво декомпозиције таласићима применом лифтинг шеме приказан на слици 3.2.2.а, састоји се из 3 дела: подела (split), процена (predict, P) и поправка (update, U). Улазни сигнал s_j се дели на два дисјунктна сигнала $s1_{j+1}$ и $s2_{j+1}$ ниже резолуције на нивоу декомпозиције $j+1$:

$$(s1_{j+1}, s2_{j+1}) = split(s_j) \tag{3.17}$$

Како су сигнали $s1_{j+1}$ и $s2_{j+1}$ високо корелисани, на основу једног од њих може се урадити процена (предикција) другог. Сигнал грешке (разлике) d_{j+1} представља грешку процене и израчунава се као разлика сигнала и његове процене:

$$d_{j+1} = s2_{j+1} - P(s1_{j+1}) \quad (3.18)$$

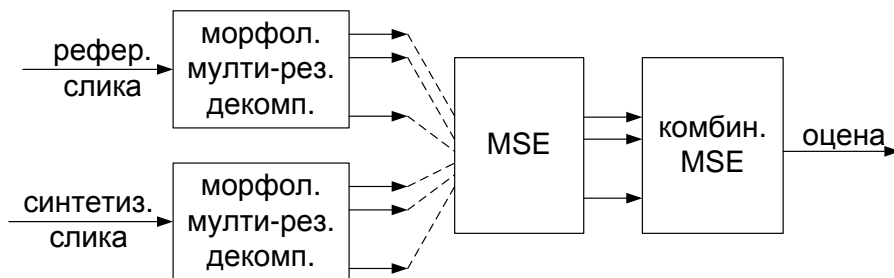
Сигнал разлике се користи за поправку, ажурирање (update) сигнала $s1_{j+1}$ и тако се добија апроксимација полазног сигнала ниже резолуције s_{j+1} :

$$s_{j+1} = s1_{j+1} + U(d_{j+1}) \quad (3.19)$$

Инверзна трансформација таласићима се израчунава применом операција у обрнутом редоследу и са променом знака операција, слика 3.2.2.б.

4 ПРЕДЛОЖЕНА МУЛТИ-РЕЗОЛУЦИОНА МЕТРИКА ЗА ОЦЕНУ КВАЛИТЕТА ДИБР-СИНТЕТИЗОВАНИХ СЛИКА

Израчунавање мултирезолуционе метрике за оцену квалитета слике се може описати као процес кроз три фазе. У првој фази референтна и слика са изобличењем се декомпонују на низ слика опадајућих резолуција применом мултирезолуционе декомпозиције. У другој фази се израчунавају мере квалитета/дисторзије по подопсезима декомпоноване слике. У трећој фази се резултати подопсега комбинују у резултујућу оцену.



Слика 4.1. Израчунавање мултирезолуционе метрике кроз три фазе

Мултирезолуциона декомпозиција слика у првој фази израчунавања метрике може имати значајан утицај на перформансе метрике па је потребно испитати различите врсте декомпозиција. Већина постојећих мултирезолуционих метрика за оцену квалитета слика користи линеарне филтре у мултирезолуционој декомпозицији. У овом раду је испитан утицај мултирезолуционих декомпозиција са нелинеарним морфолошким филтрима на перформансе метрике за оцену квалитета слика које садрже изобличења због процеса ДИБР синтезе. Испитан је утицај редундантне морфолошке пирамидалне декомпозиције са морфолошким операторима ерозија и дилатација на перформансе метрике. Утицај декомпозиције морфолошким таласићима је испитан такође. Два типа морфолошких таласића су примењена у сепарабилној декомпозицији и један тип морфолошких таласића је примењен у несепарабилној декомпозицији са узорковањем на “quincunx“ решетки. У другој фази израчунавања резултата метрике предложено је коришћење средње квадратне грешке као мера дисторзије подопсега мултирезолуционе репрезентације, слика 4.1. У трећој фази израчунавања се средње квадратне грешке подопсега комбинују у резултујућу оцену метрике.

4.1 МУЛТИРЕЗОЛУЦИОНА ДЕКОМПОЗИЦИЈА СЛИКЕ ПРИМЕНОМ МОРФОЛОШКИХ ОПЕРАТОРА

Мултирезолуциона репрезентација је одређена трансформацијом којом се из слике више резолуције добија слика ниже резолуције. У линеарним скала просторима оператор за промену скале је конволуција Гаусовим кернелом. Након конволуције Гаусовим кернелом, слика је униформно замућена па и посебно важни региони слике као што су ивице [54]. То је мана у анализи слике линеарним филтрима јер ивице одговарају физичким границама објеката. Многа истраживања су показала да је људско око врло осетљиво на изобличења ивица и контура у слици. Информације о ивицама и контурама су за човека најважније информације о структури слике за схватање сцене. У многим областима анализе слике пожељно је да ивице остану на својој позицији без померања при промени скале (резолуције). Да би се то остварило, предложене су нелинеарне мултирезолуционе декомпозиције са морфолошким операторима које омогућују очување позиције ивица кроз више скала.

Морфолошке операције

Математичка морфологија уноси квалитативну новину у процесирње слика. За разлику од других приступа који су аналитичке и статистичке природе, морфолошки приступ анализе слике је геометријски [55], [56]. Морфолошки приступ се користи у разним областима обраде слике као што су побољшање квалитета слике потискивањем шума и/или истицањем ивица, сегментација, детекција ивица, репрезентација облика, компресија, итд. Математичка морфологија је моћан алат за анализу слике јер омогућује нелинеарну анализу слике уз очување облика и геометрије слике. Према морфолошком приступу анализе слике, објекти у слици се третирају као геометријски облици. Морфолошке операције анализирају слику помоћу структурног елемента од чијих карактеристика (облика, величине, оријентације) зависи резултат анализе. Предност примене морфолошких оператора је отклањање детаља слике без нарушавања остатка структуре слике. Облик и величина отклоњених детаља зависи од облика и величине СЕ. Операције математичке морфологије могу се применити на бинарне слике, слике у скали сивог и колор слике. Као и слика, и СЕ може бити бинаран или у скали сивог. Најчешће се користи бинарни СЕ. Величина СЕ утиче на време обраде слике, тј. на број операција.

Основне морфолошке операције су дилатација и ерозија. Морфолошке операције за бинарне слике се дефинишу као геометријске трансформације скупова. Проширењем морфолошке теорије применом операција минимум и максимум дефинисане су морфолошке операције за слике у скали сивог (grayscale) [57]. За слику у скали сивог f и бинарни СЕ (SE) основне морфолошке операције се израчунавају као максимална односно минимална вредност у области дефинисаној структурним елементом:

$$D: \text{dilation}_{SE}(f)(x) = \max_{y \in SE} f(x - y) \quad (4.1)$$

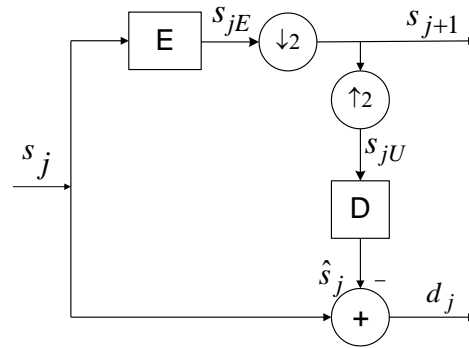
$$E: \text{erosion}_{SE}(f)(x) = \min_{y \in SE} f(x + y) \quad (4.2)$$

Применом оператора дилатација на слику у скали сивог, слика постаје светлија а тамни детаљи се смањују или отклањају зависно од величине СЕ. Применом оператора ерозија на слику у скали сивог, слика постаје тамнија а светли детаљи се смањују или отклањају. Дилатација и ерозија су транслационо инваријантне операције. Морфолошке операције су једноставне за имплементацију и ефикасне у израчунавању. Комбиновањем морфолошких оператора добијају се морфолошки филтри. Морфолошко филтрирање сигнала је нелинеарна трансформација сигнала којом се локално модификују геометријске карактеристике сигнала.

4.1.1 ПИРАМИДАЛНА ДЕКОМПОЗИЦИЈА СЛИКЕ ПРИМЕНОМ МОРФОЛОШКИХ ОПЕРАТОРА

Већина постојећих метода за оцену квалитета слике које се заснивају на пирамидалној декомпозицији слике, користе линеарне филтре. У овом раду је испитан утицај пирамидалне декомпозиције слике са морфолошким филтрима на перформансе метрике MP-PSNR за оцену квалитета ДИБР-синтетизованих слика. У литератури [16] постоји тврдња да се морфолошка пирамидална декомпозиција поклапа са човековом визуелном перцепцијом. Морфолошка пирамидална репрезентација се генерише применом шеме декомпозиције за редувантну репрезентацију слике, слика 3.1.1а, у којој се као оператори анализе и синтезе користе морфолошки оператори [19]: као оператор анализе Ψ_j^\uparrow користи се морфолошки оператор ерозија (E) и иза њега децимација а као оператор синтезе Ψ_j^\downarrow

морфолошки оператор дилатација (D) коме претходи интерполација. Тако се добија шема морфолошке пирамидалне декомпозиције МПД ЕД [19] приказана на слици 4.1.1. Пирамида МПД ЕД је коришћена код визуализације 3Д слика [58].



Слика 4.1.1. Шема морфолошке пирамидалне декомпозиције: Е ерозија, D дилатација.

Сигнал апроксимације ниже резолуције s_{j+1} добија се применом морфолошког оператора ерозија на слику апроксимације претходног нивоа пирамиде s_j а затим децимацијом са фактором 2 по свакој димензији слике δ^\downarrow . Када се користи структурни елемент демензија $(2r+1) \times (2r+1)$ за морфолошко филтрирање, слика апроксимације ниже резолуције s_{j+1} добија се према:

$$s_{jE}(m,n) = \min \{ s_j(m+k, n+l), \quad | -r \leq k, l \leq r \}$$

$$s_{j+1} = \delta^\downarrow(s_{jE}) \tag{4.3}$$

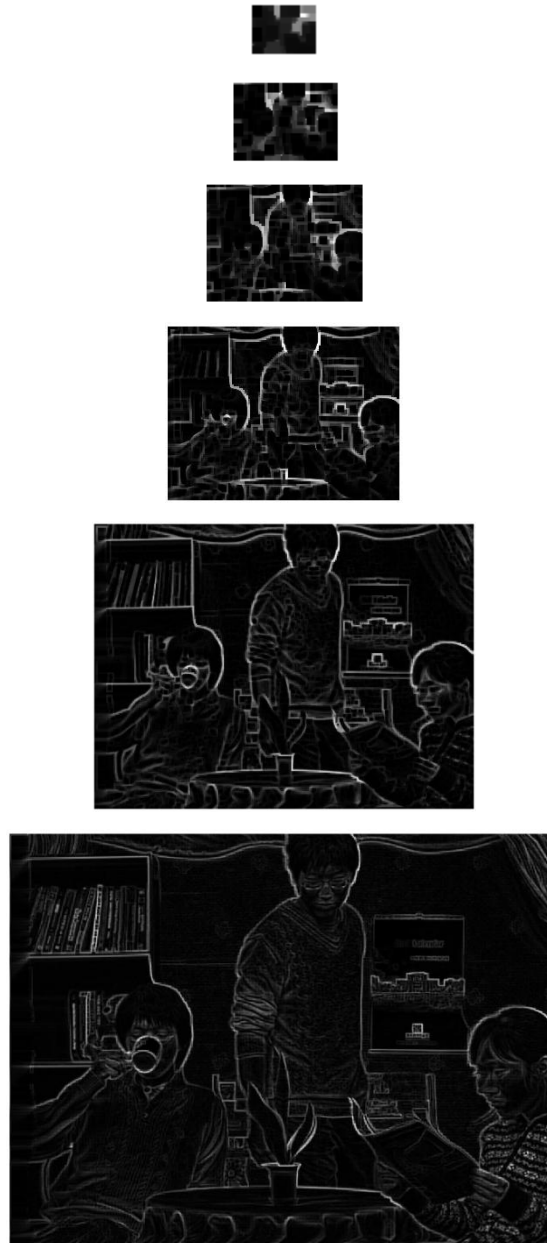
Када је СЕ у облику квадрата, морфолошко филтрирање слике се може урадити ефикасније сепарабилно по врстама и колонама. Слика детаља d_j се израчунава одузимањем филтриране интерполиране слике апроксимације следећег нивоа \hat{s}_j од слике апроксимације текућег нивоа s_j . Филтрирана интерполирана слика \hat{s}_j се добија интерполацијом δ^\uparrow слике следећег нивоа пирамиде s_{j+1} по свакој димензији слике са фактором 2 а затим филтрирањем применом оператора дилатација.

$$s_{jU} = \delta^{\uparrow}(s_{j+1})$$

$$\hat{s}_j(m,n) = \max \{ s_{jU}(m-k, n-l), \mid -r \leq k, l \leq r \}$$

$$d_j = s_j - \hat{s}_j$$

(4.4)



Слика 4.1.2. Морфолошка пирамида МПД ЕД генерисана на основу синтетизованог фрејма видео секвенце Newspaper. За морфолошке операције коришћен је СЕ димензија 7x7.

Морфолошка пирамида детаља МПД (Morphological Bandpass Pyramid, MBP) садржи слике детаља опадајуће резолуције $d_j, j=0, \dots, M-1$ и слику апроксимације најниже резолуције на врху пирамиде s_M [16]. МПД се може интерпретирати као структурна декомпозиција која истиче ивице у слици и издваја их по величини тако да су фини детаљи издвојени у сликама нижих нивоа декомпозиције које су веће резолуције а грубљи детаљи су у сликама виших нивоа декомпозиције који су ниже резолуције [19]. Слике детаља МПД ЕД садрже позитивне вредности. Максимална вредност пиксела у слици детаља вишег нивоа декомпозиције је мања од максималне вредности пиксела у слици детаља нижег нивоа декомпозиције. Минимална вредност пиксела у слици детаља вишег нивоа декомпозиције је већа од минималне вредности пиксела у слици детаља нижег нивоа декомпозиције. МПД испуњава пирамидални услов и могућа је перфектна реконструкција оригиналне слике. МПД ЕД пирамида у којој се користи СЕ димензија 2x2 је морфолошка Харова пирамида [19]. На слици 4.1.2 приказана је МПД ЕД пирамида генерисана на основу фрејма из видео секвенце Newspaper коришћењем СЕ димензија 7x7 за морфолошке операције.

4.1.2 ДЕКОМПОЗИЦИЈА СЛИКЕ МОРФОЛОШКИМ ТАЛАСИЋИМА

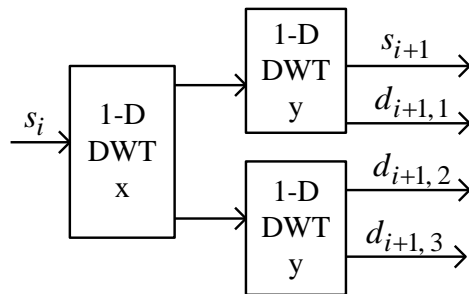
Већина постојећих метода за оцену квалитета слике које се заснивају на мултирезолуционој декомпозицији слике таласићима, користе линеарне таласиће. У овом раду је испитан утицај нелинеарних морфолошких таласића у декомпозицији слике на перформансе метрике MW-PSNR за оцену квалитета ДИБР-синтетизованих слика. Испитан је утицај сепарабилне декомпозиције са морфолошким таласићима minNaag и minLift инесепарабилне декомпозиције са узорковањем на “quincunx” решетки применом морфолошких таласића minLiftQ.

У оригиналном облику, таласићи су алат за линеарну анализу сигнала. Појава лифтинг шеме је подстакла развој нелинеарних таласића. Показано је да се свака трансформација слике таласићима може израчунати применом лифтинг шеме рачунски једноставније него применом стандардног алгоритма филтрирања [52]. Док се за конструкцију линеарних таласића може користити Фуријеова техника, лифтинг шема је једини општи метод за конструкцију морфолошких таласића [21]. Декомпозиција слике морфолошким

таласићима је нелинеарна трансформација слике којом ивице у слици остају очуване у подопсезима различитих резолуција [21]. Применом лифтинг шемеса морфолошким операторима може се реализовати и сепарабилна и несепарабилна декомпозиција слике морфолошким таласићима. Морфолошки таласићи су коришћени код кодовања видеа за кодовање фрејмова резидуала [59] и код скалабилног кодовања слика са применом за “digital watermarking“ [60].

Сепарабилна декомпозиција слике таласићима

Сепарабилна декомпозиција слике таласићима се реализује применом две једнодимензионе декомпозиције таласићима, у хоризонталном и у вертикалном правцу. Тако се добија три подопсега детаља $d_{i+1,1}$, $d_{i+1,2}$, $d_{i+1,3}$ и слика апроксимација s_{i+1} , слика 4.1.3. Подопсези детаља садрже вертикалне, хоризонталне и дијагоналне детаље. Сепарабилном декомпозицијом слике се добијају четири слике које су по обе димензије два пута мање од слике која је декомпонована. Слика апроксимације се даље декомпоњује у следећем нивоу декомпозиције (dyadic decomposition).



Слика 4.1.3. Сепарабилна декомпозиција слике таласићима применом две једнодимензионе декомпозиције таласићима у вертикалном и хоризонталном правцу.

Једнодимензиона декомпозиција слике морфолошким таласићима се може реализовати применом лифтинг шеме са слике 3.2.2а у којој су оператори предикције (prediction, P) и поправке (update, U) реализовани применом морфолошких оператора. Најједноставнија декомпозиција таласићима је тривијална декомпозиција (lazy transform) којом се једнодимензиони сигнал дели (split) на два сигнала, тако да један сигнал садржи парне узорке а други сигнал садржи непарне узорке.

Имплементирана је сепарабилна декомпозиција слике коришћењем два типа морфолошких таласића, *minNaar* и *minLift*, и испитан је њихов утицај на перформансе метрике MW-PSNR за оцену квалитета слике. Ради поређења са линеарним таласићима, имплементирана је и сепарабилна декомпозиција слике коришћењем Харовог таласића (*Naar wavelet*) и таласића *cdf* (2,2).

- *1-Д трансформација морфолошким Харовим таласићем (minNaar)*

Један од најједноставнијих нелинеарних морфолошких таласића је морфолошки Харов минимум таласић, *minNaar* [21]. Сличан је линеарном Харовом таласићу али се у кораку поправке (*update*, U) у лифтинг шеми примењује нелинеарни оператор минимум. Процена (предикција) сигнала је изузетно једноставна: сами парни узорци сигнала који се декомпонује се користе као процена непарних. Сигнал разлике се израчунава на исти начин и код линеарног Харовог таласића. Применом лифтинг шеме, сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} се израчунавају на следећи начин [61]:

$$d_{j+1}[n] = s_j[2n+1] - s_j[2n],$$

$$s_{j+1}[n] = s_j[2n] + \min(0, d_{j+1}[n]) \quad (4.5)$$

Декомпозицијом сигнала применом морфолошког Харовог таласића, ивице у сигналима детаља су боље очуване него у случају примене линеарног Харовог таласића [21]. Морфолошки Харов таласић је транслаторно инваријантан у просторном домену, инваријантан је у односу на помак нивоа сивог (*gray-shift invariant*) и инваријантан је у односу на скалирање нивоа сивог (*gray-multiplication invariant*) [61].

- *1Д Харов таласић (Naar)*

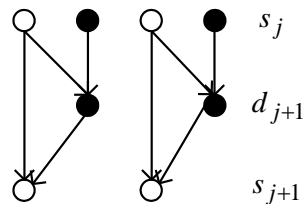
Најједноставнија линеарна декомпозиција сигнала таласићима је декомпозиција применом Харовог таласића. Применом лифтинг шеме [52], процена узорака сигнала који садржи непарне узорке сигнала који се декомпонује су сами парни узорци истог сигнала. Процена је тачна за сигнал који има константну вредност и у том случају је сигнал детаља једнак нули. Корак поправке (*update*, U) сигнала осигурава да је средња вредност сигнала апроксимације ниже резолуције једнака средњој вредности сигнала који се декомпонује.

Применом лифтинг шеме, сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} се израчунавају према:

$$d_{j+1}[n] = s_j[2n+1] - s_j[2n],$$

$$s_{j+1}[n] = s_j[2n] + \frac{1}{2}d_{j+1}[n] \quad (4.6)$$

Илустрација једног корака декомпозиције сигнала применом линеарног Харовог таласића и морфолошког Харовог таласића приказана је на слици 4.1.4. Парни узорци сигнала s_j су представљени белом бојом а непарни узорци црном. Узорак сигнала детаља d_{j+1} се израчунава на основу истих узорака (2 суседна узорака, парни и непарни) сигнала s_j и на исти начин (као њихова разлика) и код декомпозиције применом линеарног Харовог таласића и применом морфолошког Харовог таласића. Узорак сигнала апроксимације ниже резолуције се израчунава на основу истих узорака сигнала s_j и детаља d_{j+1} али применом различитих оператора код декомпозиције линеарним и морфолошким Харовим таласићима.



Слика 4.1.4. Декомпозиција сигнала s_j на сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} морфолошким и линеарним Харовим таласићима применом лифтинг шеме

- *1D декомпозиција морфолошким таласићем (minLift)*

Декомпозиција сигнала применом таласића minLift се реализује применом лифтинг шеме која садржи морфолошке операторе за израчунавање процене (P) код израчунавања сигнала детаља и за израчунавање поправке (U) сигнала апроксимације [62], [21]. Након поделе сигнала на два сигнала тако да један садржи парне узорке а други непарне,

израчунава се процена сигнала који садржи непарне узорке као минимална вредност два суседна узорка сигнала који садржи парне узорке. Корекција (update) је изабрана тако да се локални минимум сигнала пренесе у сигнал апроксимације ниже резолуције. Корекција узорка сигнала апроксимације се ради коришћењем два узорка сигнала детаља применом операције минимум.

$$d_{j+1}[n] = s_j[2n+1] - \min(s_j[2n], s_j[2n+2])$$

$$s_{j+1}[n] = s_j[2n] + \min(0, d_{j+1}[n-1], d_{j+1}[n]) \quad (4.7)$$



Слика 4.1.5. Подопсези првог нивоа декомпозиције фрејма из видео секвенце Newspaper морфолшким таласићем minLift садрже ивице по правцима: вертикалне, хоризонталне, угаоне.

Декомпозицијом сигнала применом таласића minLift чува се локална минимална вредност сигнала кроз сигнале апроксимације на свим нивоима декомпозиције а не генеришу се нове [21]. Декомпозиција слике таласићем minLift је инваријантна у односу на помак

нивоа сивог (gray-shift invariant) и инваријантна у односу на скалирање нивоа сивог (gray-multiplication invariant) [61].

Пример који илуструје декомпозицију сигнала применом таласића minLift приказан је у [62], [21]. Сигнал детаља има вредност око нуле у областима благе варијације нивоа сивог у слици која се декомпонује. Нагле промене интензитета слике у скали сивог се пресликавају у позитивне вредности сигнала детаља. Подопсеци првог нивоа декомпозиције фрејма из видео секвенце Newspaper морфолшким таласићем minLift су приказани на слици 4.1.5. Подопсеци садрже вертикалне детаље, хоризонталне детаље и детаље око углова (дијагоналне).

- *Линеарни таласић cdf(2,2)*

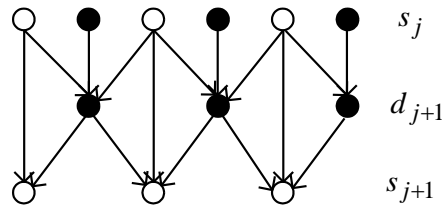
Линеарни биортогонални таласић назван према истраживачима који су га представили, Cohen-Daubechies-Feauveau, cdf(2,2), је имплементиран применом лифтинг шеме[52] због поређења са њему одговарајућим морфолошким таласићем minLift. Процена узорка сигнала који садржи непарне узорке се израчунава једноставно као средња вредност два суседна узорка сигнала који садржи парне узорке. Процена је тачна за линеарне сигнале и у том случају је сигнал детаља једнак нули. Сигнал детаља показује колико оригинални сигнал одступа од линеарног. Сигнал детаља садржи високофреквенцијске компоненте оригиналног сигнала. Поправка (update) сигнала који садржи парне узорке омогућује очување средње вредности сигнала. Декомпозицијом сигнала помоћу cdf(2,2) таласића применом лифтинг шеме узорак сигнала детаља d_{j+1} и сигнала апроксимације s_{j+1} се израчунавају према:

$$d_{j+1}[n] = s_j[2n+1] - \frac{1}{2}(s_j[2n] + s_j[2n+2])$$

$$s_{j+1}[n] = s_j[2n] + \frac{1}{4}(d_{j+1}[n-1] + d_{j+1}[n]) \quad (4.8)$$

Илустрација израчунавања узорака сигнала детаља и сигнала апроксимације код декомпозиције сигнала морфолошким таласићем minLift и линеарним таласићем cdf(2,2) применом лифтинг шеме је приказана на слици 4.1.6. Најпре се сигнал који се декомпонује s_j подели на два сигнала, сигнал са парним узорцима (бели) и сигнал са непарним

узорцима (црни). Код обе декомпозиције, морфолошке и код линеарне, узорак сигнала детаља d_{j+1} се израчунава на основу иста 3 узорка сигнала који се декомпонује, s_j , али применом различитих операција. Узорак сигнала апроксимације s_{j+1} се израчунава поправком узорка сигнала који се декомпонује s_j помоћу иста два узорка сигнала детаља d_{j+1} код обе декомпозиције, морфолошке и линеарне, али применом различитих операција.



Слика 4.1.6. Декомпозиција сигнала s_j на сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} применом морфолошког таласића minLift или линеарног таласића $\text{cdf}(2,2)$ коришћењем лифтинг шеме

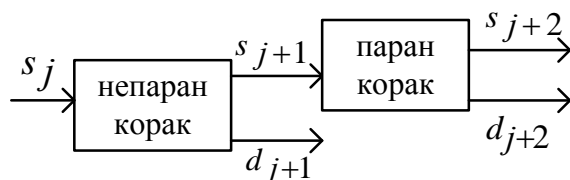
Несепарабилна декомпозиција таласићима на “quincunx” решетки

Слике као дводимензиони низови се могу трансформисати правом 2Д трансформацијом којом се слика трансформише као област а не по врстама и колонама [63]. Применом сепарабилне декомпозиције слике коришћењем 1Д трансформација таласићима по вертикалном и по хоризонталном правцу, ивице које се не поклапају са ова два правца не могу се издвојити у подопсезима. Применом несепарабилне декомпозиције, овај недостатак се може отклонити. Несепарабилно узорковање омогућује декомпозицију слике која се боље поклапа са човековим визуелним системом [64]. “Quincunx“ решетка (решетка као шаховско поље) је најједноставнија несепарабилна мултидимензиона структура узорковања. Због симетрије ”quincunx“ решетке, несепарабилна трансформација је неосетљива на правац ивица и оријентацију слике. Применом несепарабилне трансформације слика се декомпонује са мултирезолуционим фактором $\sqrt{2}$ по скали. Да би се остварио исти ниво децимације слике применом сепарабилне и несепарабилне

декомпозиције са “quincunx“ узорковањем, слику треба декомпоновати са душло већим бројем нивоа декомпозиције применом несепарабилне декомпозиције.

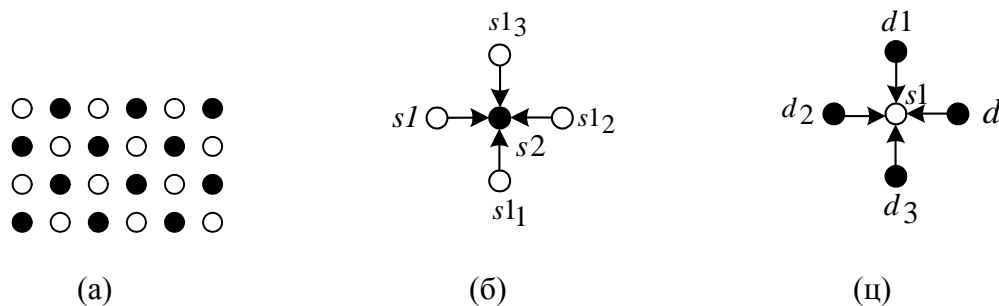
У овом раду је испитан утицај несепарабилне декомпозиције слике морфолошким таласићима на “quincunx“ решетки на поузданост мере MW-PSNR за оцену квалитета ДИБР-синтетизованих слика. Због поређења, имплементирана је и несепарабилна декомпозиција слике линеарним таласићима на “quincunx“ решетки коришћењем лифтинг шеме.

Несепарабилна 2Д трансформација таласићима на “quincunx“ решетки је реализована наизменичном применом непарног и парног корака, слика 4.1.7. У сваком кораку се генерише слика детаља и слика апроксимације. Слика апроксимације се може декомпоновати даље.



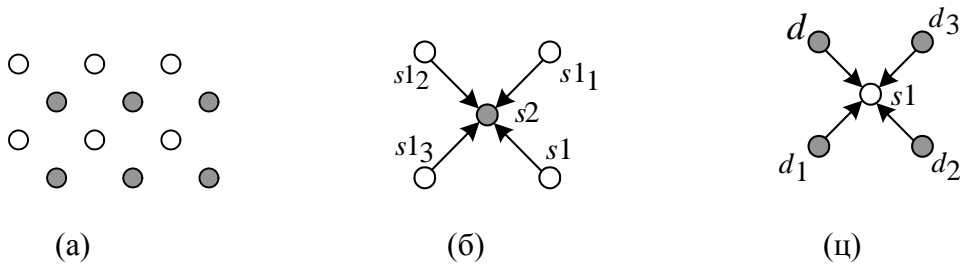
Слика 4.1.7. Несепарабилна декомпозиција слике таласићима коришћењем “quincunx“ решетке применом непарног и парног корака

И паран и непаран корак су имплементирани коришћењем лифтинг шеме са слике 3.2.2а.



Слика 4.1.8. Непарни корак несепарабилне декомпозиције слике применом лифтинг шеме
 а) слика која се декомпонује садржи узорке на Cartesian решетки најпре се дели на два сигнала, оба са пикселима на “quincunx“ решетки б) процена црног пиксела се израчунава на основу 4 бела пиксела који га окружују ц) бели пиксел се ажурира на на основу 4 пиксела детаља који га окружују

У непарном кораку декомпозиције, пиксели слике s_j се најпре деле (split) на 2 сигнала, $(s1_{j+1}, s2_{j+1}) = split(s_j)$ оба на “quincunx“ решетки, тако да сигнал $s1_{j+1}$ садржи беле пикселе а сигнал $s2_{j+1}$ црне пикселе, слика 4.1.8.а. За сваки црни пиксел израчунава се његова процена на основу 4 најближа бела пиксела која га окружују, слика 4.1.8б. Пиксел сигнала грешке d_{j+1} се израчунава као разлика црног пиксела и његове процене, $d = s2 - P(s1, s1_1, s1_2, s1_3)$. Затим се сваки бели пиксел сигнала $s1_{j+1}$ ажурира (поправља) на основу 4 пиксела сигнала грешке (детаља) који га окружују, $s = s1 + U(d, d_1, d_2, d_3)$ слика 4.1.8.ц. Сигнал апроксимације s_{j+1} се даље декомпонира у парном кораку декомпозиције.



Слика 4.1.9. Парни корак несепарабилне декомпозиције на “quincunx“ решетки применом лифтинг шеме а) сигнал апроксимације са пикселима на quincunx решетки се дели на два сигнала, оба на Cartesian решетки б) процена сивог пиксела на основу 4 бела пиксела који га окружују по дијагоналним правцима ц) бели пиксел се ажурира на на основу 4 пиксела детаља који га окружују по дијагоналним правцима

У парном кораку несепарабилне декомпозиције слике, сигнал апроксимације s_{j+1} генерисан у непарном кораку декомпозиције који садржи узорке на “quincunx“ решетки дели се на два сигнала, сигнал $s1_{j+2}$ представљен белим пикселима и $s2_{j+2}$ представљен сивим пикселима на слици 4.1.9.а, оба на Cartesian решетки. Процена пиксела сигнала $s2_{j+2}$ израчунава се на основу 4 бела пиксела сигнала $s1_{j+2}$ која га окружују по дијагоналним правцима, слика 4.1.9б. Пиксел сигнала грешке d_{j+2} се израчунава као разлика сивог пиксела и његове процене. Затим се сваки бели пиксел сигнала $s1_{j+2}$

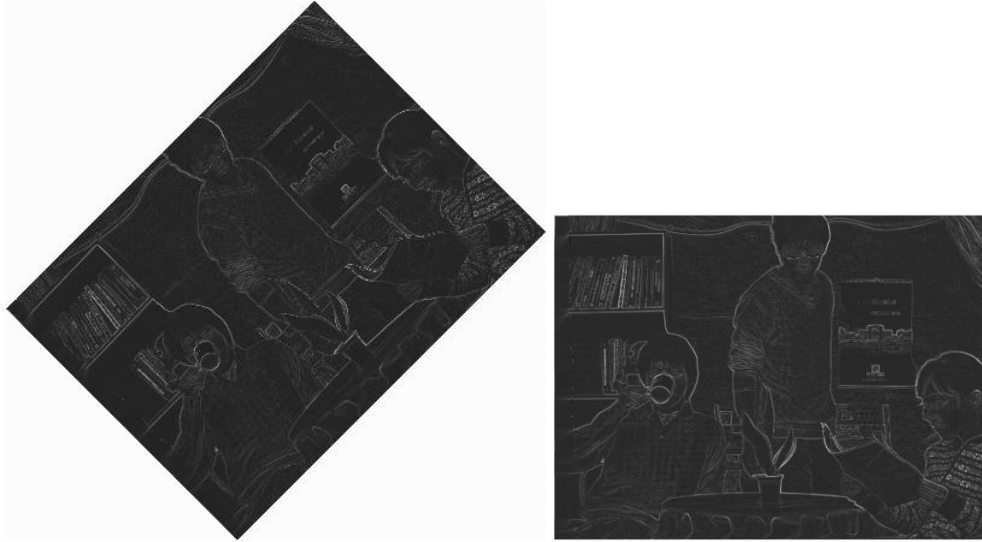
ажурира (поправља) на основу 4 пиксела сигнала детаља који га окружују по дијагоналним правцима, слика 4.1.9.ц, и тако се добија сигнал апроксимације s_{j+2} .

- *Несепарабилна нелинеарна декомпозиција слике са “quincunx“ узорковањем морфолошким таласићем minLiftQ*

Несепарабилна декомпозиција морфолошким таласићем minLiftQ на “quincunx“ решетки применом лифтинг шеме аналогна сепарабилној декомпозицији морфолошким таласићем minLift је представљена у [21], [65]. Прво се сигнал слике подели на два сигнала тако да један садржи црне пикселе а други беле, сваки на “quincunx“ решетки, слика 4.1.8а. У непарном кораку декомпозиције, процена пиксела сигнала који садржи црне пикселе на слици се израчунава као минимална вредност 4 пиксела који га окружују сигнала који садржи беле пикселе, слика 4.1.8.б. Сигнал разлике се израчунава као разлика сигнала који садржи црне пикселе и његове предикције (4.9). Пиксел сигнала апроксимације се ажурира додавањем минималне вредности 4 пиксела сигнала детаља који га окружују (4.10), слика 4.1.8.ц. Упарном кораку несепарабилне декомпозиције, након поделе сигнала апроксимације на два сигнала, оба на Cartesian решетки, израчунава се процена сигнала који садржи сиве пикселе. Процена сивог пиксела се израчунава као минимална вредност 4 пиксела која га окружују по дијагоналним правцима, слика 4.1.9б. Грешка процене се израчунава према (4.9). У непарном кораку, бели пиксел сигнала апроксимације се ажурира (поправља) на основу 4 пиксела детаља који га окружују по дијагоналним правцима, слика 4.1.9ц, коришћењем операције минимум. Илустрација непарног корака декомпозиције морфолошким таласићем minLiftQ на “quincunx“ решетки је приказана у [61].

$$d = s_2 - \min(s_1, s_{1_1}, s_{1_2}, s_{1_3}) \quad (4.9)$$

$$s = s_1 + \min(d, d_1, d_2, d_3, 0) \quad (4.10)$$



Слика 4.1.10. Подопсеци прва два нивоа несепарабилне декомпозиције фрејма из синтетизоване видео секвенце Newspaper са узорковањем према “quincunx“ решетки, непарног и парног, применом морфолошког таласића minLiftQ. Слика детаља непарног корака декомпозиције је ротирана за 45° .

Подопсеци првог нивоа несепарабилне декомпозиције фрејма из синтетизоване видео секвенце Newspaper на “quincunx“ решетки применом морфолошког таласића minLiftQ приказани су на слици 4.1.10. Слика детаља непарног корака декомпозиције је ротирана 45° пре приказа. Области велике промене нивоа сивог (ивице) су мапиране у бело. Области са малом променом нивоа сивог су мапиране у тамно.

- *Несепарабилна линеарна декомпозиција слике са “quincunx“ узорковањем таласићем $cdfQ(2,2)$*

Због поређења са нелинеарном несепарабилном декомпозицијом слике морфолошким таласићем, имплементирана је несепарабилна линеарна декомпозиција са “quincunx“ узорковањем таласићем $cdf(2,2)Q$ [66] који је аналоган таласићу $cdf(2,2)$ за сепарабилну декомпозицију. Применом лифтинг шеме, након поделе пиксела слике на 2 дела према слици 4.1.8а, у непарном кораку декомпозиције, израчунава се сигнал грешке као разлика сигнала који садржи црне пикселе и његове процене према (4.11). Процена црног пиксела се израчунава као средња вредност 4 најближа бела пиксела, слика 4.1.8.б. Четири пиксела сигнала грешке се користи за ажурирање пиксела сигнала апроксимације, слика 4.1.8.ц, према (4.12).

$$d = s_2 - \frac{1}{4}(s_1 + s_{11} + s_{12} + s_{13}) \quad (4.11)$$

$$s = s_1 + \frac{1}{8}(d + d_1 + d_2 + d_3) \quad (4.12)$$

У парном кораку декомпозиције, сигнал апроксимације генерисан у непарном кораку се дели на 2 дела према слици 4.1.9.а. Процена пиксела за израчунавање сигнала грешке се израчунава као средња вредност 4 околна пиксела, слика 4.1.9.б. Сигнал грешке се израчунава према (4.11). Пиксел сигнала апроксимације се ажурира на основу 4 околна пиксела сигнала детаља, слика 4.1.9.ц, према (4.12).

4.2 ИЗРАЧУНАВАЊЕ ДИСТОРЗИЈЕ

Средња квадратна грешка (Mean Squared Error), MSE, и вршна вредност сигнал-шум (Peak Signal-to-Noise Ratio), PSNR, су највише коришћене мере дисторзије/квалитета. То су најједноставнији начини за израчунавање сличности између две слике, референтне и изобличене. MSE има физичко значење – дефинише енергију сигнала грешке. Особина очувања енергије гарантује да је енергија изобличеног сигнала иста у трансформационом домену као и у домену сигнала. MSE је одлична метрика у контексту оптимизације. За поређење различитих алгоритама највише се користе MSE и PSNR [67].

У овом раду је коришћена MSE као мера дисторзије између две мултирезолуционе репрезентације, референтне слике и синтетизоване слике, у другој фази израчунавања мултирезолуционе метрике. Како су у подопсезима мултирезолуционе репрезентације слике издвојене ивице, израчунавањем MSE по подопсезима истиче се важност ивица у оцени мултирезолуционе метрике. Применом морфолошких филтара у мултирезолуционој декомпозицији ивице у подопсезима су издвојене без замућења и помераја. Израчунавањем MSE пиксел по пиксел прецизно се мери разлика у области ивица између слика две мултирезолуционе репрезентације, синтетизоване слике и референтне слике, на свим нивоима. Тако је добијена метрика чија се оцена добро подудара са субјективним оценама.

4.2.1 Израчунавање MP-PSNR

Израчунавање оцене метрике MP-PSNR која користи пирамидалну декомпозицију слика приказано је на слици 4.2.1. Између слика две пирамидалне репрезентације, референтне слике x_{R_j} и синтетизоване слике x_{D_j} , на сваком нивоу декомпозиције j , димензија $K_j N_j$, израчунава се средња квадратна грешка MSE_j :

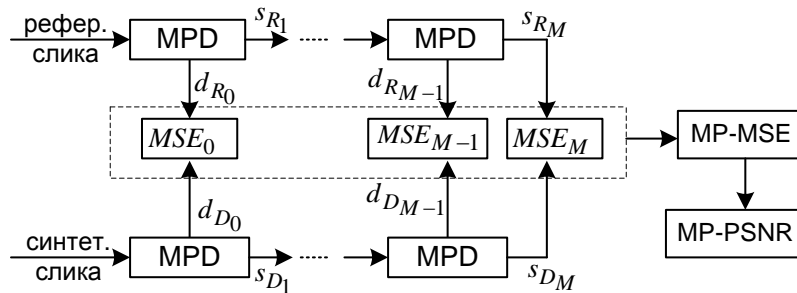
$$MSE_j = \frac{1}{N_j \cdot K_j} \sum_{k=1}^{K_j} \sum_{n=1}^{N_j} (x_{R_j}(k,n) - x_{D_j}(k,n))^2 \quad (4.13)$$

Средња квадратна грешка пирамиде MP-MSE се израчунава као производ средњих квадратних грешака свих слика пирамиде са тежинским фактором $\beta_j = \frac{1}{M+1}$, где је М број нивоа декомпозиције.

$$MP_MSE = \prod_{j=0}^M [MSE_j]^{\beta_j} \quad (4.14)$$

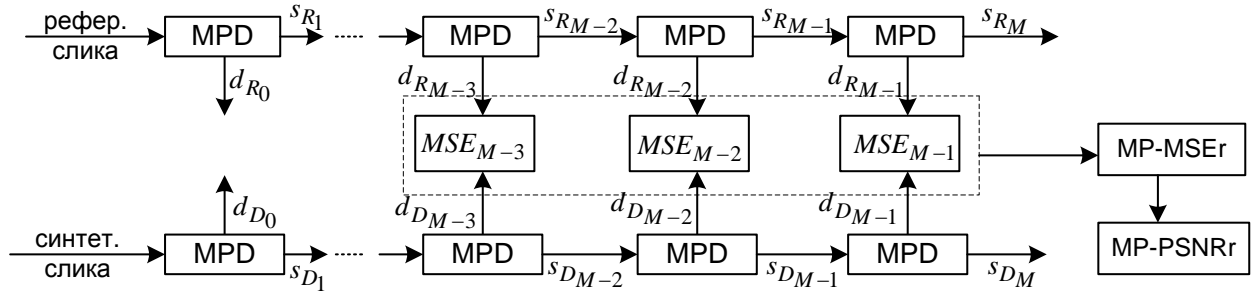
За слику максималног динамичког ранга R, MP-MSE се трансформише у MP-PSNR:

$$MP_PSNR = 10 \cdot \log_{10} \left(\frac{R^2}{MP_MSE} \right) \quad (4.15)$$



Слика 4.2.1. Израчунавање MP-PSNR на основу MSE_j свих слика пирамиде: MPD- један ниво морфолошке пирамидалне декомпозиције

Осим основне верзије метрике која користи све слике пирамиде за израчунавање дисторзије, анализирана је и редукована верзија метрике MP-PSNR_r која користи само слике детаља виших нивоа пирамидалне репрезентације, слика 4.2.2.



Слика 4.2.2. Израчунавање редуковане верзије MP-PSNRr на основу MSE_j слика детаља виших нивоа пирамиде: MPD-један ниво морфолошке пирамидалне декомпозиције

Средња квадратна грешка редуковане пирамиде MP-MSEr се израчунава као сума средњих квадратних грешака слика детаља виших нивоа пирамиде са тежинским фактором $\beta_j = \frac{1}{M_r}$, где је M_r је број слика пирамиде коришћен за израчунавање редуковане верзије метрике, слика 4.2.2.

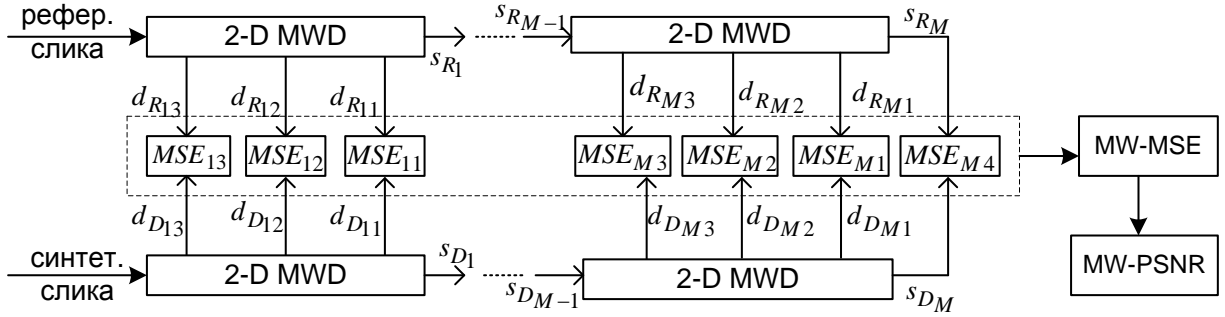
$$MP_MSEr = \sum_j \beta_j \cdot MSE_j \quad (4.16)$$

$$MP_PSNRr = 10 \cdot \log_{10} \left(\frac{R^2}{MP_MSEr} \right) \quad (4.17)$$

4.2.2 Израчунавање MW-PSNR

Израчунавање оцене метрике MW-PSNR која користи декомпозицију слика таласићима приказано је на слици 4.2.3. MSE_{ji} је коришћена за мерење дисторзије између подопсега две мултирезолуционе репрезентације, референтне и синтетизоване слике, након декомпозиције таласићима. Између одговарајућих подопсега референтне и синтетизоване слике свих нивоа декомпозиције израчунавају се средње квадратне грешке MSE_{ji} , за подопсег i на нивоу декомпозиције j .

$$MSE_{j,i} = \frac{1}{N_j \cdot K_j} \sum_{k=1}^{K_j} \sum_{n=1}^{N_j} (x_{R_{j,i}}(k,n) - x_{D_{j,i}}(k,n))^2 \quad (4.18)$$



Слика 4.2.3. Израчунавање MW-PSNR на основу MSE_{ji} свих подопсега i свих нивоа декомпозиције j добијених сепарабилном декомпозицијом 2-D MWD синтетизоване слике таласићима

Декомпозицијом слике таласићима добија се већи број подопсега. Човеков визуелни систем је комплексан и није потпуно познат тако да није познато на који начин је најбоље комбиновати подопсега код израчунавања метрике. Зато је изабран једноставан метод чијом применом су остварене добре перформансе метрике. Средња квадратна грешка мулти-резулационе репрезентације слике добијене декомпозицијом таласићима, MW-MSE, се израчунава као збир средњих квадратних грешака свих подопсега са једнаким тежинским факторима $\beta_{ji} = \frac{1}{M \cdot D + 1}$, M је број нивоа декомпозиције а D број подопсега детаља једног нивоа декомпозиције (код сепарабилне декомпозиције $D=3$, код несепарабилне $D=2$).

$$MW - MSE = MSE_{M,D+1} \cdot \beta_{M,D+1} + \sum_{j=1}^M \sum_{i=1}^D MSE_{j,i} \cdot \beta_{j,i} \quad (4.19)$$

За слику максималног динамичког ранга R , MW-MSE се трансформише у MW-PSNR:

$$MW - PSNR = 10 \cdot \log_{10} \left(\frac{R^2}{MW - MSE} \right) \quad (4.20)$$

Редукована верзија метрике, MW-PSNR_r, која примењује декомпозицију слика таласићима користи само подопсега виших нивоа декомпозиције. Сабирањем средњих квадратних грешака подопсега и нормализацијом добија се средња квадратна грешка за више подопсега:

$$MW - MSE_r = \sum_j \sum_i MSE_{ji} \cdot \beta_{ji} \quad (4.21)$$

где тежински фактор $\beta_{ji} = \frac{1}{M_r}$ зависи од броја подопсега M_r који се користе за израчунавање редуковане верзије метрике.

$$MW - PSNR_r = 10 \cdot \log_{10} \left(\frac{R^2}{MW - MSE_r} \right) \quad (4.22)$$

4.3 РАЧУНСКА КОМПЛЕКСНОСТ ПРЕДЛОЖЕНЕ МЕТРИКЕ

Морфолошки оператори који се користе за мултирезолуциону декомпозицију слика раде са целим бројевима и као резултат генеришу целе бројеве применом операција минимум, максимум и сабирање. Једноставни су за израчунавање. Израчунавање MSE је такође једноставно (потребно је извршити 1 одузимање, 1 множење и 1 сабирање по пикселу). Израчунавање морфолошких мултирезолуционих мера, MP-PSNR и MW-PSNR, је једноставно.

Морфолошко филтрирање пирамидалне декомпозиције је имплементирано рачунски једноставније у два корака, сепарабилно по колонама и по врстама применом структурног елемента величине $P \times 1$ пиксела по колонама и $1 \times P$ пиксела по врстама, $P=2, 3, 5, 7, 9, 11, 13$. За филтрирање слике димензија NK пиксела морфолошким оператором ерозија у пирамидалној декомпозицији потребно је $\frac{3}{2}(P-1)NK$ поређења (comparisons, C). Исти број операција поређења је потребан за интерполационо филтрирање морфолошким оператором дилатација пре одузимања (subtraction, S) код израчунавања слике разлике d_1 , слика. Број операција за израчунавање слике детаља на првом нивоу декомпозиције d_1 , је $(3 \cdot (P-1) \cdot C + S) \cdot NK$. Укупан број операција за израчунавање морфолошке пирамиде ED са L нивоа је:

$$(3 \cdot (P-1) \cdot C + S) \cdot NK \cdot \left(1 + \frac{1}{2^2} + \frac{1}{2^4} + \dots + \frac{1}{2^{2(L-1)}} \right) =$$

$$(3 \cdot (P-1) \cdot C + S) \cdot \frac{4}{3} NK, \quad L \rightarrow \infty \quad (4.23)$$

Број операција за израчунавање морфолошке пирамиде ED зависи од величине структурног елемента SE као што је приказано у табели 4.3.1. Са повећањем димензија структурног елемента, комплексност израчунавања расте.

Табела 4.3.1. Број операција за израчунавање морфолошке пирамиде ерозија/дилатација са L нивоа применом структурног елемента величине 1xP пиксела по врстама и Px1 пиксела по колонама. Коришћене су операцијеодузимање (S) и поређење (C).

P (SE)	L	број операција за декомп.
2	6	$(3 \cdot C + S) \cdot \frac{1365}{1024} NK$
3	5	$(6 \cdot C + S) \cdot \frac{341}{256} NK$
5	5	$(12 \cdot C + S) \cdot \frac{341}{256} NK$
7	5	$(18 \cdot C + S) \cdot \frac{341}{256} NK$
9	4	$(24 \cdot C + S) \cdot \frac{85}{64} NK$
11	4	$(30 \cdot C + S) \cdot \frac{85}{64} NK$
13	4	$(36 \cdot C + S) \cdot \frac{85}{64} NK$

Број операција за израчунавање MSE за основну верзију MP-PSNR и за редуковану верзију MP-PSNRr је приказан у табели 4.3.2. Рачунска сложеност израчунавања редуковане верзије MP-PSNRr је мања него рачунска сложеност израчунавања основне верзије MP-PSNR јер се MSE израчунавају само за слике детаља ниже резолуције док се код основне верзије MSE израчунава за све слике пирамиде.

Табела 4.3.2. Број операција за израчунавање MSE слика пирамиде за основну верзију MP-PSNR и за редуковану верзију MP-PSNRr. Морфолошка пирамидална декомпозиција са L нивоа применом SE димензија $P \times P$. MSE се израчунава коришћењем операција одузимање (S), множење (M) и сабирање (A).

P (SE)	L	број операција за MSE
MP-PSNR		
2	6	$(S + M + A) \frac{1365}{1024} NK$
3	5	$(S + M + A) \frac{341}{256} NK$
5	5	
7	5	
9	4	$(S + M + A) \frac{85}{64} NK$
11	4	
13	4	
MP-PSNRr		
2	4-6	$(S + M + A) \frac{21}{1024} NK$
3	3-5	$(S + M + A) \frac{21}{256} NK$
5	3-5	
7	3-5	
9	2-4	$(S + M + A) \frac{21}{64} NK$
11	2-4	
13	2-4	

Број операција за израчунавање декомпозиције морфолошким таласићима за слику димензија NK је приказан у табели 4.3.3.

Табела 4.3.3. Број операција за израчунавање декомпозиције слике таласићима (S одузимање, C поређење, A сабирање, M множење)

декомпозиција	таласићи	број операција
сепарабилна	minHaar	$(S + C + A) \cdot \frac{4}{3} NK$
	Haar	$(S + M + A) \cdot \frac{4}{3} NK$
	minLift	$(3C + S + A) \cdot \frac{4}{3} NK$
	cdf(2,2)	$(3A + 2M + S) \cdot \frac{4}{3} NK$
несепарабилна	minLiftQ	$(7C + S + A) \cdot 2NK$
	cdf(2,2)Q	$(7A + 2M + S) \cdot 2NK$

Сепарабилна декомпозиција морфолошким таласићима је рачунски једноставнија од несепарабилне. Једноставнија је и од морфолошке пирамидалне декомпозиције за исту дужину филтара. На пример, декомпозиција слике морфолошким таласићима minHaar је рачунски једноставнија од декомпозиције слике морфолошким Харовом пирамидом. Такође, декомпозиција слике морфолошким таласићима minLift је рачунски једноставнија од декомпозиције слике морфолошким пирамидом у којој је коришћен структурни елемент дужине 3 за сепарабилну пирамидалну декомпозицију. Декомпозиција слике морфолошким Харовим таласићима је рачунски најједноставнија а њеном применом у метрици MW-PSNR се постижу најбоље перформансе метрике.

Број операција за израчунавање MSE свих подопсега добијених декомпозицијом слике таласићима је приближно $(S + M + A)NK$. Укупан број пиксела у свим подопсезима је NK .

5 РЕЗУЛТАТИ

Предложена метрика је тестирана на три базе: на бази слика које садрже изобличења због процеса ДИБР синтезе, IRCCyN/IVCDIBR, на бази стерео слика синтетизованих различитим техникама синтезе, MCL-3D, и на бази видео секвенци са навигацијом кроз различите позиције, FVSV, које садрже изобличења због процеса синтезе и због компресије мапа дубине које се користе у процесу синтезе.

5.1 МЕРЕ ЗА ПОРЕЂЕЊЕ ПЕРФОРМАНСИ ОБЈЕКТИВНИХ МЕТРИКА ЗА ОЦЕНУ КВАЛИТЕТА СЛИКА

У овој секцији су описане мере за оцењивање перформанси објективних метрика за оцену квалитета слика/видео које су коришћене у овом раду. Анализа перформанси објективних метрика које користе референтне слике/видео секвенце за израчунавање оцене слике/видео секвенце са изобличењем се заснива на субјективним оценама DMOS (Differential Mean Opinion Score). На основу субјективне средње оцене (Mean Opinion Score) за референтну слику/секвенцу $MOS(r)$, и средње субјективне оцене за слику/секвенцу са изобличењем $MOS(d)$, израчунава се DMOS(d) за слику/секвенцу са изобличењем [69]:

$$DMOS(d) = MOS(d) - MOS(r) + 5 \quad (5.1)$$

Већа вредност DMOS означава бољи квалитет слике/секвенце. Доња граница за DMOS је 1 а горња може бити и већа од 5 ако је референтној слици додељена нижа субјективна оцена MOS него слици са изобличењем.

Субјективне оцене су обично компримоване на крајевима скале док објективне оцене нису. Због тога се примењује нелинеарно мапирање пре израчунавања перформанси објективне метрике [69]. Применом нелинеарне логистичке функције оцене објективне метрике Q се пресликавају у предикцију оцена субјективне метрике $DMOS_p$:

$$DMOS_p(Q) = aQ^3 + bQ^2 + cQ + d \quad (5.2)$$

где су a , b , c , d коефицијенти функције нелинеарног мапирања. Нелинеарним мапирањем оцена објективне метрике према субјективним оценама, максимизује се корелација између субјективних и објективних оцена.

Израчунавање DMOS на основу MOS оцена добијених субјективним оцењивањем и нелинеарно мапирање објективних према субјективним оценама је урађено према Плану тестирања за оцењивање модела квалитета видеа за ТВ високе резолуције који је предложила група VQEG HDTV[69].

За поређење перформанси мера за оцену квалитета слика коришћене су следеће метрике: корен средње квадратне грешке (Root Mean Squared Error, RMSE) између субјективних и нелинеарно мапираних објективних резултата, Пирсонов коефицијент корелације (Pearson's Correlation Coefficient, PCC) нелинеарно мапираних објективних према субјективним оценама и Спирманов коефицијент корелације ранга (Spearman's correlation coefficient, SCC).

Тачност објективне метрике се оцењује израчунавањем корена из средње квадратне грешке (Root Mean Squared Error, RMSE) између субјективних оцена DMOS и нелинеарно мапираних објективних оцена $DMOS_p$ на основу N слика за које су израчунате оцене објективне метрике из базе слика:

$$RMSE = \sqrt{\frac{1}{N-d} \sum_{i=1}^N (DMOS_{p_i} - DMOS_i)^2} \quad (5.3)$$

где је d број степена слободе нелинеарне функције мапирања, тј. број коефицијената нелинеарне функције ($d=4$ када се користи монотона полиномска функција трећег реда за нелинеарно мапирање објективних према субјективним оценама).

Пирсонов коефицијент корелације (PCC) се израчунава након нелинеарног мапирања између субјективних и објективних оцена. За базу која садржи N слика Пирсонов коефицијент корелације се израчунава између субјективних оцена DMOS и нелинеарно мапираних објективних оцена $DMOS_p$ свих слика базе:

$$PCC = \frac{\sum_{i=1}^N (DMOS_i - \overline{DMOS}) * (DMOS_{p_i} - \overline{DMOS}_p)}{\sqrt{\sum_{i=1}^N (DMOS_i - \overline{DMOS})^2} * \sqrt{\sum_{i=1}^N (DMOS_{p_i} - \overline{DMOS}_p)^2}} \quad (5.4)$$

Вредности Пирсоновог коефицијента корелације су у интервалу [-1, 1].

Спирманов коефицијент корелације (SCC) се израчунава на основу разлике ранга субјективне и објективне оцене сваке слике, d_i :

$$SCC = 1 - \frac{6 \sum d_i^2}{N(N^2 - 1)} \quad (5.5)$$

RMSE и PCC оцењују тачност предикције, а SCC монотоност предикције. Боље објективне мере имају виши PCC и SCC а нижи RMSE.

5.2 ТЕСТ ПРИМЕР 1: БАЗА СИНТЕТИЗОВАНИХ СЛИКА IRCCyN/IVCDIBR

После описа базе IRCCyN/IVCDIBR која садржи синтетизоване фрејмове видео секвенци, приказане су перформансе предложене метрике у оцењивању ове базе.

5.2.1 Опис базе IRCCyN/IVCDIBR

База слика IRCCyN/IVCDIBR [24], [14] садржи слике синтетизоване применом различитих алгорита ДИБР синтезе на основу некомпримованих слика и одговарајућих некомпримованих мапа дубине. Слике базе садрже изобличења због процеса ДИБР синтезе. IRCCyN/IVCDIBR база садржи фрејмове из 3 MVD видео секвенце, слика 5.2.1, описане у табели 5.2.1.



Слика 5.2.1. Фрејмови MVD видео секвенци базе IRCCyN/IVCDIBR: Book Arrival, Lovebird1 и Newspaper.

Табела 5.2.1. База IRCCyN/IVCDIBR садржи фрејмове три MVD секвенце

секвенца	резолюција	број камера	удаљеност камера
Book Arrival	1024x768	16	6.5 цм
Lovebird1		12	3.5 цм
Newspaper		9	5 цм

При креирању базе IRCCyN/IVCDIBR, свака слика је синтетизована на основу једног изабраног фрејма из снимљене видео секвенце и њему одговарајуће мапе дубине. Из сваке MVD секвенце изабране су 2 оригиналне видео секвенце на 2 позиције према табели 5.2.2. На основу изабраног фрејма једне видео секвенце и њему одговарајуће мапе дубине синтетизоване су 2 слике на 2 позиције које одговарају првој суседној позицији и следећој суседној позицији у односу на позицију видео секвенце из које је изабрани фрејм према табели 5.2.2.

Табела 5.2.2. База слика IRCCyN/IVCDIBR: 3 MVD садржаја, позиције оригиналне и синтетизоване слике

садржај	позиција оригиналне слике	позиција синтетизоване слике	број фрејма
BookArrival	10	8	60
	8	9	54
	10	9	60
	8	10	54
Lovebird	8	6	112
	6	7	104
	8	7	112
	6	8	104
Newspaper	6	4	104
	4	5	136
	6	5	104
	4	6	136

За синтезу слика коришћено је 7 алгоритама синтезе, A1-A7. Алгоритам A1 и A2 примењује метод [12] којим се мапа дубине филтрира 2Д нископропусним Гаусовим филтром чиме се губи оштрина ивица. На тај начин се избегава појава “дисоклузија” унутар синтетизоване слике али се јављају геометријске дисторзије које су посебно уочљиве на прелазу између позадине и предњег дела слике. Ипак, област дуж ивице слике остаје непопуњена. Применом алгоритма A1 та област се одсеца па се слика интерполира до оригиналне величине. Применом алгоритма A2 ивична област слике се попуњава применом алгоритма за попуњавање рупа [70]. Метод A1 је погодан само за стерео видео када се једна слика стерео пара преноси а друга се синтетизује на пријемној страни. За

више видео секвенци из MVD секвенце, алгоритам A1 није примењив јер би требало урадити одсецање у свим сликама свих MVD видео секвенци на свим позицијама да би се очувао стерео утисак што би довело до губитка информација и повећале би се дисторзије у сликама.

ДИБР алгоритми се разликују по начину попуњавања “дисоклузија” у синтетизованим сликама. Алгоритам A3 користи алгоритам за попуњавање [70] који уноси замућење у област дисоклузије. Метод [70] не успева да реконструише комплексне текстуре. Алгоритам A3 [71] је усвојен као референтни софтвер за MPEG експерименте стандардизације. Алгоритам A4 [72] попуњава празнине у синтетизованој слици линију по линију користећи суседне информације из позадине. Метод не успева да реконструише вертикалне и косе структуре и комплексне текстуре. Погодан је за реконструкцију једноставне глатке позадине без текстуре. У алгоритму A5 попуњавање рупа у синтетизованој слици се ради применом методе синтезе текстуре тако што се копира постојећи садржај слике у делове који недостају [73]. Применом алгоритма A5 праве ивице су добро реконструисане у синтетизованој слици. Применом алгоритма A1-A5 сваки фрејм видео секвенце се синтетизује независно од осталих. Алгоритам A6 користи садржај претходног фрејма за попуњавање рупа текућег фрејма [74]. Алгоритам A7 не попуњава области на синтетизованој слици које одговарају “дисоклузијама” – остају празнине у слици.

База садржи 84 синтетизоване слике (3 MVD x 4 фрејма x 7 алгоритама) и 12 оригиналних слика (3 MVD x 4 фрејма).

Слике синтетизоване применом алгоритма A1 су добиле највеће оцене приликом субјективног рестирања а најниже оцене применом свих објективних метрика. То је због тога што алгоритам A1 не примењује поступак за попуњавање (inpainting) “дисоклузија”. “Дисоклузија” дуж ивице слике је одсечена а затим је слика рескалирана. “Дисоклузије” унутар слике су избегнуте нископропусним филтрирањем мапа дубине. Мада тако синтетизована слика не садржи приметна изобличења, не може се користити за 3Д апликације јер геометрија сцене више не одговара референтној слици. Пошто анализирам методе синтезе за генерисање 3Д садржаја овај метод ће бити изостављен код тестирања.

Слике синтетизоване применом алгорита синтезе A2-A7 су приказане на слици 5.2.2.



Слика 5.2.2.Слике синтетизоване на основу фрејма видео секвенце Newspaper применом алгоритама A2-A7 с лева на десно, одозго на доле

За субјективно оцењивање слика базе изабрана је методологија ACR-HR (Absolute Categorical Rating with Hidden Reference Removal) према којој се субјектима приказује једна по једна слика по случајном редоследу и они је оцењују оценама 1-5. Референтне слике су укључене у оцењивање као и остале слике базе па отуд назив „скривена референца”. Током субјективног оцењивања слике су приказиване на екрану TVLogic LVM401W. У тестирању је учествовало четрдесеттри испитаника, углавном студената. Примењена је методологија за субјективну оцену квалитета телевизијских слика [76]. За сваку слику је израчуната средња оцена свих посматрача и добијена је MOC (Mean Opinion Score) оцена.

5.2.2 Перформансе предложене метрике при оцењивању базе IRCCyN/IVCDIBR

У овој секцији је приказана анализа и резултати предложене метрике при оцењивању слика базе IRCCyN/IVCDIBR. Анализирани су основне верзије метрике MP-PSNR и MW-PSNR и редуковане верзије метрике MP-PSNRr и MW-PSNRr. Анализирана је и метрика PSNR посебно по подопсезима декомпозиције.

Анализа перформанси метрике MP-PSNR

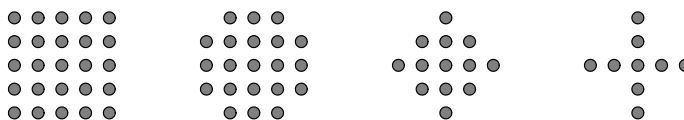
У овој секцији су анализирани перформансе метрике Морфолошка пирамидална вршна вредност сигнал/шум, MP-PSNR (Morphological Pyramid Peak Signal-to-Noise Ratio). Прва фаза у израчунавању метрике MP-PSNR је мултирезолуциона декомпозиција референтне и синтетизоване слике морфолошким пирамидом ерозија/дилатација, МПД ЕД. Испитан је утицај броја нивоа пирамидалне декомпозиције на перформансе метрике MP-PSNR. Анализиран је утицај облика и величине структурног елемента који се користи у морфолошким операцијама ерозија и дилатација морфолошке пирамидалне декомпозиције слика. За поређење су израчунате и перформансе мултирезолуционе метрике засноване на Лапласовој пирамидалној декомпозицији применом линеарних филтара, LP-PSNR.

Анализиране су перформансе метрике PSNR по нивоима декомпозиције израчунате између одговарајућих слика пирамидалне репрезентације референтне слике и пирамидалне репрезентације синтетизоване слике. Показано је да су перформансе метрике PSNR израчунате на вишим нивоима пирамидалне репрезентације знатно боље од перформанси метрике PSNR израчунате на нижим нивоима пирамиде. Због тога је предложена редукована метрика MP-PSNR_r која користи само слике пирамиде са виших нивоа декомпозиције. Анализиране су перформансе редуковане метрике MP-PSNR_r зависно од димензија структурног елемента који се користи у морфолошким операцијама при креирању морфолошке пирамидалне репрезентације слике.

Од облика и величине структурног елемента који се користи за морфолошко филтрирање зависи које геометријске карактеристике ће бити очуване у филтрираној слици а посебно правац повећања или сужавања објеката слике. Када се користи квадратни СЕ објекти слике се повећавају или сужавају подједнако по свим правцима. Квадратни СЕ је погодан за детекцију правих линија а кружни СЕ за детекцију циркуларних.

Анализиран је утицај структурних елемената различитих облика (квадратни, кружни, ромб, крстасти), слика 5.2.3, који се користе у морфолошкој операцији ерозија при креирању морфолошке пирамидалне репрезентације слике на перформансе метрике MP-PSNR. Применом квадратног и кружног СЕ остварене су боље перформансе MP-PSNR метрике него применом СЕ у облику ромба или крста. Квадратни СЕ је погодан за

имплементацију јер омогућује пирамидалну декомпозицију сепарабилно по врстама и колонама са децимацијом после сваког корака па је рачунска сложеност знатно мања. Због тога је изабран СЕ квадратног облика за морфолошке операције током израчунавања морфолошке пирамидалне декомпозиције.



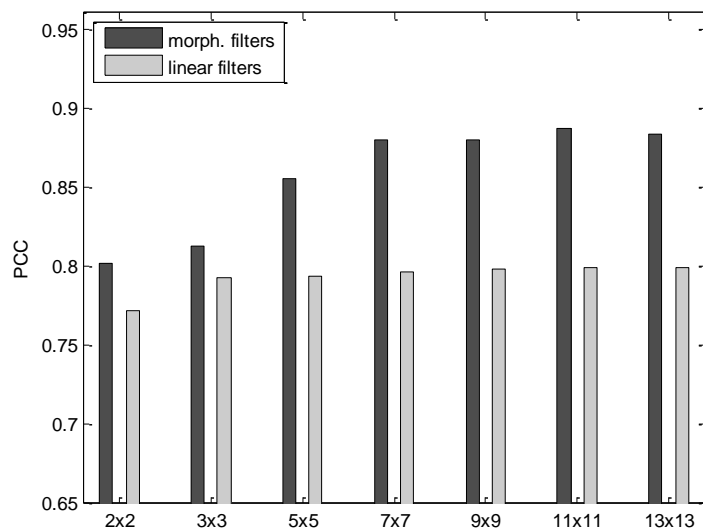
Слика 5.2.3. Структурни елементи различитих облика: квадратни, кружни, ромб, крстасти (с лева на десно)

Табела 5.2.3. Перформансе основне верзије MP-PSNR и редуковане верзије MP-PSNR_r

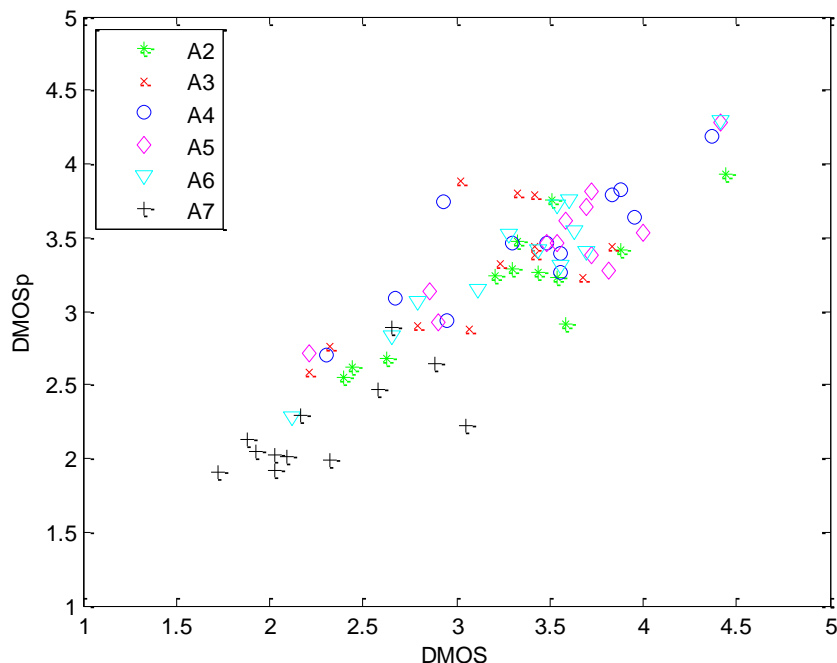
SE	нивоа	RMSE	PCC	SCC
MP-PSNR				
2x2	6	0.4101	0.8019	0.7083
3x3	5	0.3996	0.8131	0.7101
5x5	5	0.3561	0.8549	0.7759
7x7	5	0.3264	0.8796	0.8050
9x9	5	0.3263	0.8798	0.8015
11x11	4	0.3165	0.8874	0.8175
13x13	4	0.3221	0.8830	0.8021
MP-PSNR _r				
2x2	4-6	0.3660	0.8459	0.7775
3x3	3-5	0.3252	0.8806	0.8185
5x5	3-5	0.2936	0.9039	0.8634
7x7	3-5	0.2931	0.9042	0.8573
9x9	2-4	0.2997	0.8996	0.8614
11x11	2-4	0.2922	0.9048	0.8684
13x13	2-4	0.2920	0.9050	0.8684

Испитан је утицај величине СЕ и броја нивоа пирамидалне декомпозиције на перформансе метрике MP-PSNR. Примењени су квадратни структурни елементи димензија 2x2, 3x3,

5x5, ..., 13x13. Анализирane су перформансе MP-PSNR за број нивоа пирамидалне декомпозиције од 1 до 7. Изабрана је комбинација величине SE и броја нивоа декомпозиције која омогућује најбоље перформансе MP-PSNR што је приказано у горњем делу табеле 5.2.3. Када је примењен SE димензија 2x2, најбоље перформансе метрике MP-PSNR су постигнуте са 6 нивоа пирамидалне декомпозиције. Применом SE димензија 3x3, 5x5, 7x7 и 9x9 најбоље перформансе метрике MP-PSNR су остварене са 5 нивоа пирамидалне декомпозиције. За SE већих димензија најбоље перформансе метрике MP-PSNR су остварене са 4 нивоа пирамидалне декомпозиције. Што је већи SE више детаља се уклања из слике морфолошким филтрирањем. Повећањем димензија SE, перформансе метрике MP-PSNR су боље али је и дуже време израчунавања декомпозиције и резултата метрике. Пирсонов коефицијент корелације метрике MP-PSNR према DMOS-у за SE различитих димензија приказан је на слици 5.2.4. Применом SE димензија већих од 5x5, постижу се добре перформансе метрике MP-PSNR. Најбоље перформансе метрике MP-PSNR су остварене применом SE димензија 11x11: Пирсонов коефицијент корелације 0.887 и Спирманов коефицијент корелације 0.817. Зависност нелинеарно пресликаних вредности MP-PSNR према DMOS приказана је графички на слици 5.2.5. Свака тачка на графикону одговара једној слици из базе.



Слика 5.2.4. Пирсонови коефицијенти корелације (PCC) метрике која користи пирамидалну декомпозицију слика са морфолошким филтрима MP-PSNR и са линеарним филтрима LP-PSNR према DMOS за различите величине структурног елемента за морфолошке филтре и различите величине прозора за линеарне филтре, 2x2 до 13x13.



Слика 5.2.5.База IRCCyN/IVCDIBR: DMOSp (MP-PSNR) према DMOS

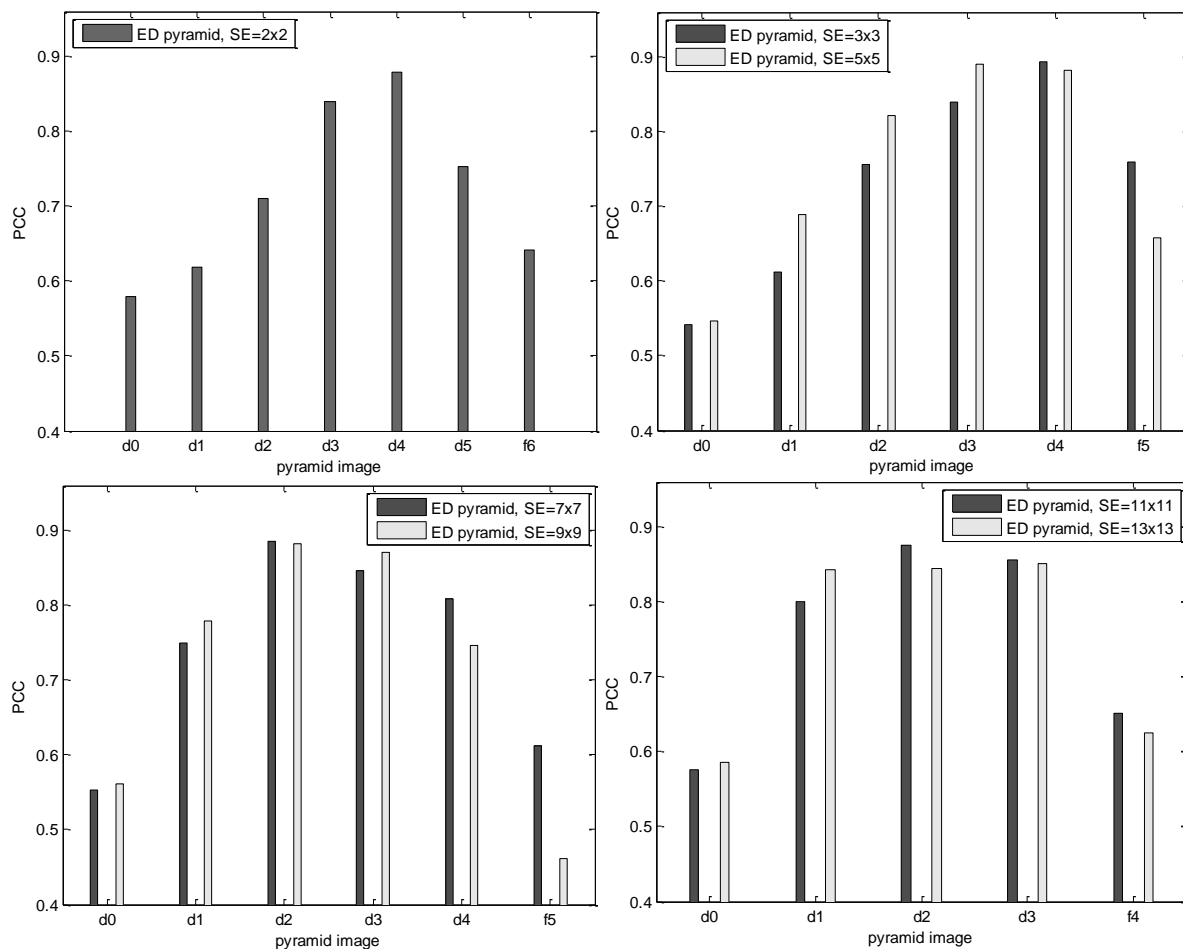
Због поређења је имплементирана и мултирезолуциона метрика заснована на Лапласовој пирамидалној декомпозицији са линеарним филтрима, LP-PSNR. Коришћени су једноставни и ефикасни биномни филтри као апроксимација Гаусових филтара[77]. Биномни коефицијенти филтара су из Паскаловог троугла, нормализовани њиховим сумама. Дводимензиони филтар је имплементиран као каскада једнодимензионих филтара сепарабилно по врстама и по колонама. Анализиран је утицај дужине филтра и броја нивоа декомпозиције на перформансе метрике LP-PSNR. За линеарне филтре дужине 2, 3, 5, ..., 13 најбоље перформансе метрике LP-PSNR су остварене пирамидалном декомпозицијом са 6 нивоа. Перформансе метрике LP-PSNR не зависе од дужине филтара што је приказано на слици 2. Пирсонов коефицијент корелације за метрику LP-PSNR према DMOS је у интервалу 0.771 за линеарни филтар дужине 2 до 0.799 за линеарни филтар дужине 13. Према слици, перформансе метрике MP-PSNR која користи морфолошку пирамидалну декомпозицију су знатно боље од перформанси метрике LP-PSNR засноване на линеарној пирамидалној декомпозицији. Разлика је већа за веће димензије филтара. На пример, применом SE димензија 11x11 за морфолошке операције код морфолошке пирамидалне декомпозиције, Пирсонов коефицијент корелације MP-PSNR према DMOS је 0.887, а применом линеарног филтра дужине 11 у линеарној

пирамидалној декомпозицији Пирсонов коефицијент корелације LP-PSNR према DMOS је 0.798.

Анализа перформанси PSNR метрике по сликама пирамиде

У неким случајевима (развучен контраст, померена средња вредност интензитета, додат бели Гаусов шум, импулсни шум, JPEG компресија, замућење слике, просторно скалирање, просторни померај, ротација) MSE се не поклапа најбоље са субјективном оценом квалитета слике [68] када се израчунава за слику у основном опсегу у пуној резолуцији. Када се PSNR израчунава за слику ниже резолуције која је добијена филтрирањем оригиналне слике нископропусним филтром са децимацијом након филтрирања, објективна мера квалитета се боље поклапа са субјективном оценом него када се PSNR израчунава за слику у основном опсегу пуне резолуције [78]. Ниво декомпозиције на коме се остварује најбоље поклапање оцене мере PSNR са субјективним оценама зависи од удаљености посматрања слике и резолуције слике.

У овој секцији су анализиране перформансе PSNR метрике за слике пирамиде по нивоима декомпозиције. Референтна и синтетизована слика су декомпоноване на низ слика опадајуће резолуције применом морфолошке пирамиде ерозија/дилатација. На сваком нивоу декомпозиције, PSNR је израчунат између слика две пирамиде истог нивоа декомпозиције. Као мера поклапања резултата PSNR метрике са субјективним оценама, израчунат је Пирсонов коефицијент корелације, PCC, објективне мере PSNR према субјективној оцени DMOS. На графицима на слици 5.2.6 приказане су вредности PCC по нивоима пирамиде. Структурни елементи различитих димензија, 2x2, 3x3, до 13x13, су коришћени за морфолошко филтрирање код пирамидалне декомпозиције слике. На основу графика може се закључити да је поклапање оцене PSNR метрике са субјективном оценом веома различито на различитим скалама пирамиде: значајно је боље на вишим скалама, за слике ниже резолуције. За МПД ЕД у којој је коришћен СЕ димензија 2x2 и 3x3, најбоље поклапање PSNR са субјективним оценама је на скали 5 за слику детаља d4. Када је коришћен СЕ димензија 5x5 најбоље поклапање PSNR са субјективним оценама је на скали 4 док је применом већих структурних елемената најбоље поклапање PSNR са субјективним оценама на скали 3.

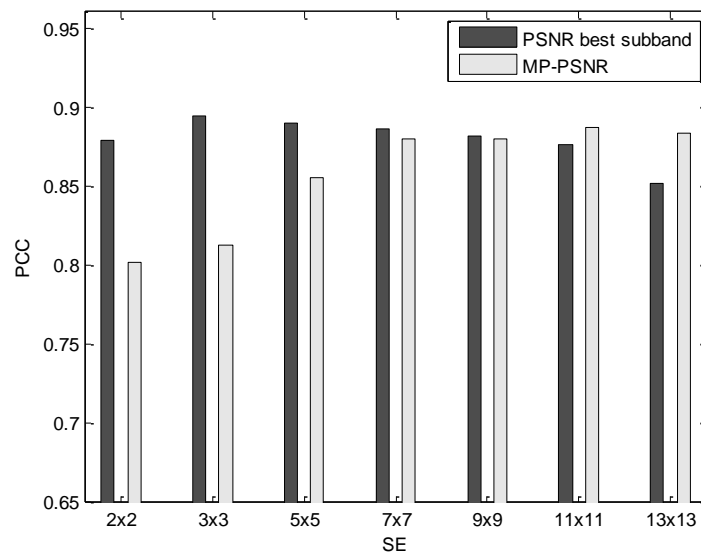


Слика 5.2.6. Пирсонов коефицијент корелације PSNR према DMOS за слике пирамиде по нивоима декомпозиције. Структурни елементи различитих димензија су коришћени за морфолошко филтрирање.

За различите димензије структурних елемената, у табели 5.2.4 су приказане перформансе PSNR метрике за ону слику пирамиде за коју су PSNR перформансе најбоље. Најбоље перформансе PSNR метрике, PCC=89.39% и SCC=86.71%, су постигнуте за слике детаља на нивоу декомпозиције 5 када је коришћен структурни елемент димензија 3x3 за морфолошко филтрирање у пирамидалној декомпозицији. Перформансе PSNR метрике су значајно боље за слике детаља виших нивоа пирамиде него за слику у основном опсегу у пуној резолуцији.

Табела 5.2.4. Перформансе PSNR метрике слике пирамиде креиране применом структурних елемената различитих димензија

СЕ	слика	RMSE	PCC	SCC
2x2	d4	0.3270	0.8792	0.8147
3x3	d4	0.3076	0.8939	0.8671
5x5	d3	0.3130	0.8899	0.8656
7x7	d2	0.3180	0.8862	0.8485
9x9	d2	0.3239	0.8816	0.8697
11x11	d2	0.3307	0.8763	0.8513
13x13	d3	0.3597	0.8517	0.7859
PSNR(оригинал/синтет.)		0.4525	0.7519	0.6766



Слика 5.2.7. Пирсонов коефицијент корелације PSNR слике детаља морфолошке пирамиде према DMOS и Пирсонов коефицијент корелације MP-PSNR према DMOS, за структурне елементе различитих димензија

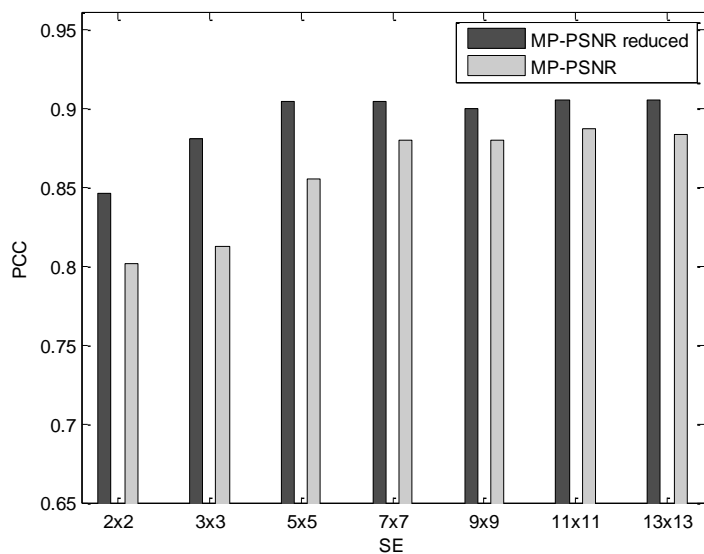
Пирсонов коефицијент корелације PSNR према DMOS за слику морфолошке пирамиде за коју су најбоље перформансе PSNR метрике за различите величине структурних елемената су приказане на слици 5.2.7. На истој слици су приказани и Пирсонови коефицијенти корелације метрике MP-PSNR према DMOS. Са графика видимо да је PSNR

слике пирамиде са виших нивоа декомпозиције боље корелисан са субјективним оценама него метрика MP-PSNR када се користи структурни елемент димензија мањих од 9x9 за морфолошке операције у пирамидалној декомпозицији.

Показано је да се за ДИБР-синтетизоване слике, PSNR веома добро поклапа са субјективним оценама када се рачуна за слике детаља морфолошке пирамиде виших нивоа декомпозиције.

Перформансе редуковане верзије MP-PSNR_r

На основу закључка да су перформансе метрике PSNR боље за слике виших нивоа морфолошке пирамиде, предложена је редукована верзија MP-PSNR_r која се заснива на MSE слика детаља само са виших нивоа пирамиде. На пример, када се користи SE димензија 2x2 за морфолошко филтрирање у пирамидалној декомпозицији, редукована верзија MP-PSNR_r користи слике детаља са нивоа пирамиде 4-6. Применом SE димензија 3x3, 5x5 и 7x7, MP-PSNR_r се израчунава на основу MSE слика детаља са нивоа пирамиде 3-5. Применом већих структурних елемената димензија 9x9, 11x11 и 13x13, MP-PSNR_r се израчунава на основу MSE слика детаља са нивоа пирамиде 2-4.



Слика 5.2.8. Пирсонов коефицијент корелације редуковане верзије MP-PSNR_r и основне верзије MP-PSNR према DMOS за структурне елементе различитих димензија 2x2 – 13x13

Перформансе редуковане верзије MP-PSNR_r су приказане у доњем делу табеле 5.2.3. Поуздана оцена синтетизованих слика се добија применом MP-PSNR_r која се заснива на морфолошкој пирамидалној декомпозицији са CE=5x5 (Пирсонов коефицијент корелације 90.39%, Спирманов коефицијент корелације 86.3%). Редукована верзија MP-PSNR_r има боље перформансе него основна верзија MP-PSNR, слика 5.2.8.

Анализа перформанси MW-PSNR

У овој секцији су анализирани перформансе метрике MW-PSNR (Morphological Wavelet Peak Signal-to-Noise Ratio) засноване на декомпозицији референтне и синтетизоване слике морфолошким таласићима. Испитан је утицај сепарабилне декомпозиције морфолошким таласићима minHaar и minLift и несепарабилне декомпозиције морфолошким таласићима minLiftQW са узорковањем на решетки шаховско поље (quincunx) на перформансе метрике MW-PSNR. Због поређења је анализиран утицај линеарних таласића, Харовог и cdf(2,2) у сепарабилној декомпозицији и cdfQ(2,2) у несепарабилној декомпозицији са узорковањем на “quincunx“ решетки. Анализиран је утицај различитог броја нивоа декомпозиције морфолошким таласићима на перформансе метрике MW-PSNR.

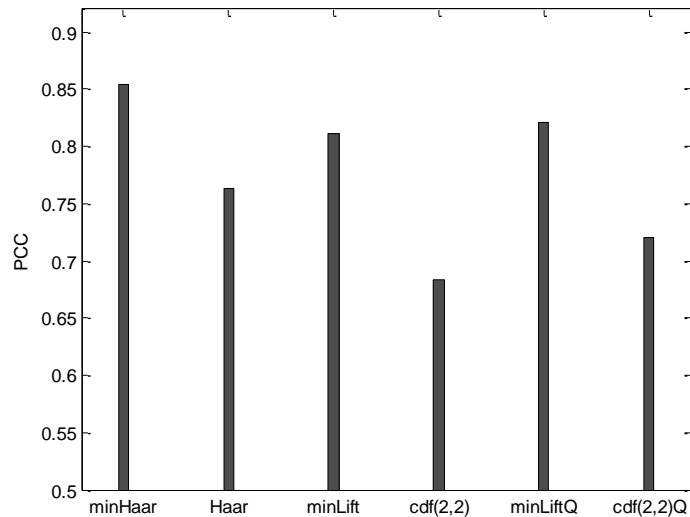
Анализиране су перформансе PSNR метрике израчунате за различите подопсеге слике добијене декомпозицијом слике морфолошким таласићима. На основу резултата перформанси метрике PSNR по подопсезима предложена је редукована верзија MW-PSNR_r која користи само подопсеге за које се PSNR боље поклапа са субјективним оценама.

Испитан је утицај броја нивоа декомпозиције од 1 до 8 на перформансе метрике MW-PSNR. Најбоље перформансе су остварене за 7 нивоа код сепарабилне декомпозиције чиме се добија 22 подопсега слике. Код несепарабилне декомпозиције се најбоље перформансе метрике MW-PSNR остварују такође за 7 нивоа декомпозиције чиме се добија 15 подопсега слике. Перформансе метрике MW-PSNR применом различитих таласића су приказане у табели 5.2.5. Перформансе метрике MW-PSNR применом морфолошких таласића за декомпозицију слика су боље него перформансе метрике засноване на декомпозицији линеарним таласићима. Најбоље перформансе метрике MW-PSNR су остварене применом морфолошких таласића minHaar. Перформансе метрике

MW-PSNR остварене применом несепарабилне декомпозиције са морфолошким таласићима minLiftQ су боље од перформанси метрике MW-PSNR остварене применом сепарабилне декомпозиције са аналогним морфолошким таласићима minLift.

Табела 5.2.5 Перформансе метрике MW-PSNR и редуковане верзије MW-PSNRr

	RMSE	PCC	SCC
таласићи			
MW-PSNR			
minHaar	0.3565	0.8545	0.7750
Haar	0.4435	0.7632	0.6491
minLift	0.4017	0.8108	0.6816
cdf(2,2)	0.5009	0.6836	0.5450
minLiftQ	0.3922	0.8206	0.7382
cdf(2,2)Q	0.4756	0.7210	0.5779
MW-PSNRr			
minHaar	0.3188	0.8855	0.8298
Haar	0.3935	0.8194	0.7695
minLift	0.3878	0.8251	0.6990
cdf(2,2)	0.4735	0.7239	0.5958
minLiftQ	0.3599	0.8514	0.7641
cdf(2,2)Q	0.4508	0.7541	0.6126

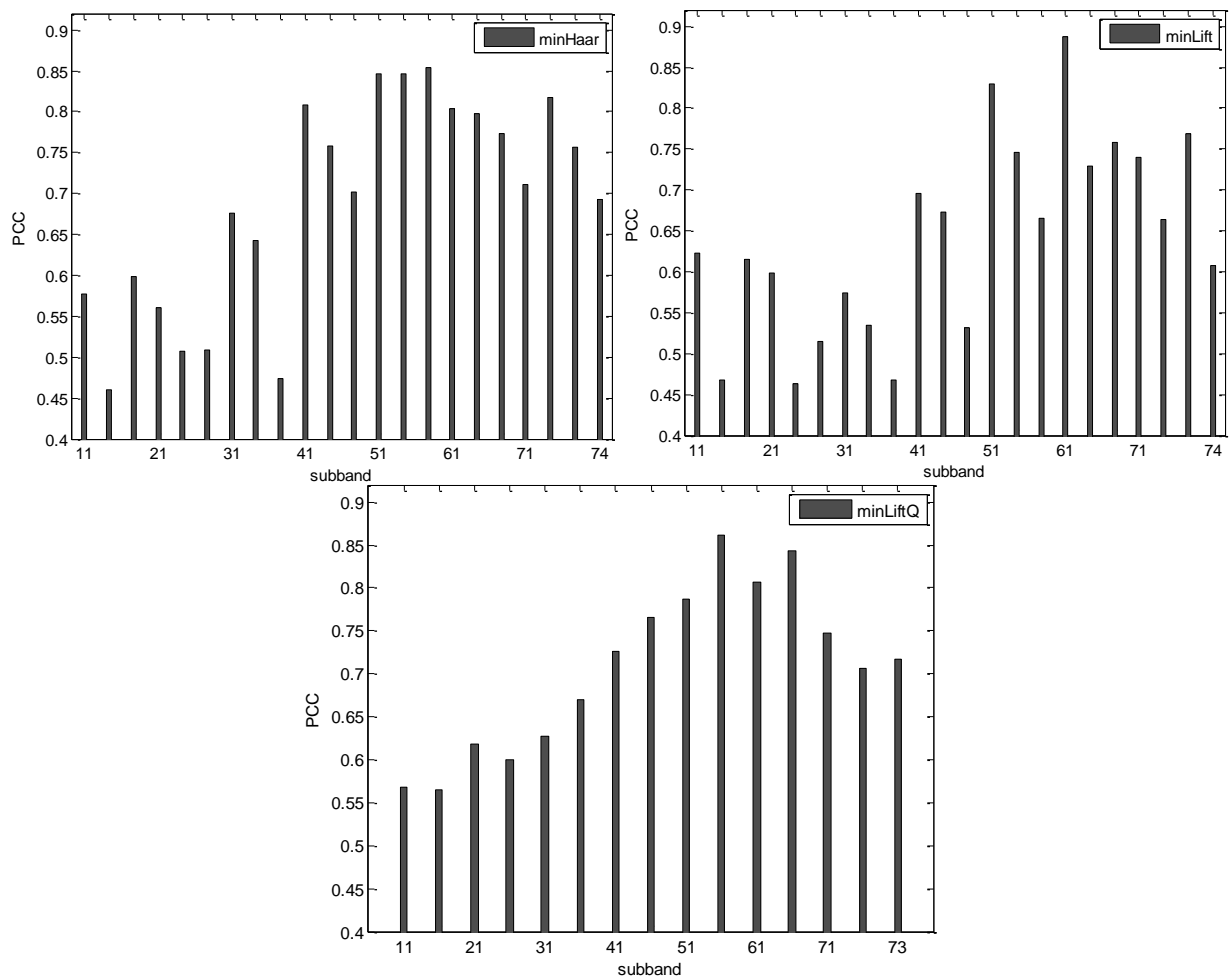


Слика 5.2.9. Пирсонов коефицијент корелације метрике MW-PSNR према DMOS за различите типове морфолошких и линеарних таласића

На слици 5.2.9 су приказане вредности Пирсоновог коефицијента корелације метрике MW-PSNR према DMOS за различите типове таласића.

Перформансе PSNR по подопсезима добијеним декомпозицијом слике таласићима

У овој секцији су анализиране перформансе PSNR метрике по подопсезима добијеним декомпозицијом слике морфолошким таласићима. За сваки подопсег синтетизоване слике израчунат је PSNR коришћењем одговарајућег подопсега референтне слике. Као мера поклапања резултата PSNR метрике са субјективним оценама, израчунат је Пирсонов коефицијент корелације, PCC, објективне мере PSNR према субјективној оцини DMOS.



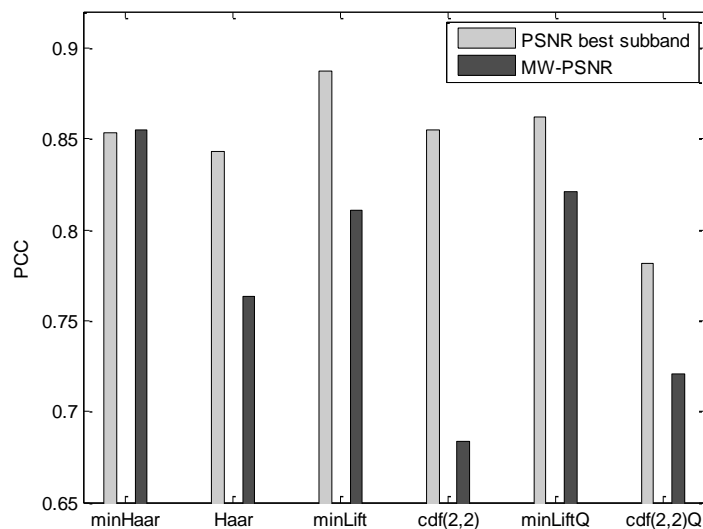
Слика 5.2.10. Пирсонови коефицијенти корелације (PCC) метрике PSNR према DMOS по подопсезима добијеним декомпозицијом синтетизованих слика морфолошким таласићима minHaar, minLift и minLiftQ

На графицима на слици 5.2.10 приказане су вредности PCC по подопсезима за три типа морфолошких таласића, minHaar, minLift и minLiftQ. На основу графика може се закључити да се PSNR знатно боље поклапа са субјективним оценама за подопсеге виших нивоа декомпозиције добијене декомпозицијом синтетизоване слике морфолошким таласићима. За сва три типа морфолошких таласића, Пирсонови коефицијенти метрике PSNR према DMOS су знатно виши за подопсеге виших нивоа декомпозиције, 4 до 7, него нижих нивоа декомпозиције, 1 до 3. Перформансе метрике PSNR за подопсег за који су PSNR перформансе најбоље, за све типове таласића, су приказане у табели 5.2.6. Применом сепарабилне декомпозиције слике морфолошким таласићем minLift, најбоље перформансе метрике PSNR се добијају за слику детаља d61, PCC=0.887, SCC=0.828. Перформансе метрике PSNR израчунате за подопсеге виших нивоа декомпозиције ниже резолуције су знатно боље од перформанси метрике PSNR израчунате за слику у основном опсегу у пуној резолуцији.

Табела 5.2.6. Перформансе PSNR за подопсег детаља добијен декомпозицијом слике таласићима

декомпозиција	таласић	подопсег	RMSE	PCC	SCC
сепарабилна	minHaar	53	0.3576	0.8535	0.7831
	Haar	61	0.3691	0.8431	0.7939
	minLift	61	0.3167	0.8872	0.8281
	cdf(2,2)	61	0.3558	0.8551	0.7671
несепарабилна	minLiftQ	52	0.3478	0.8621	0.7777
	cdf(2,2)Q	52	0.4279	0.7818	0.6493
без декомп.	референтна/синтетизована	основни опсег	0.4525	0.7519	0.6766

На графику на слици 5.2.11 су приказани Пирсонови коефицијенти корелације (PCC) метрике PSNR према DMOS за подопсег за који PSNR има најбоље перформансе. На истом графику су и вредности PCC метрике MW-PSNR према DMOS. За све типове таласића, PSNR израчунат за онај подопсег за који PSNR има најбоље перформансе, се боље поклапа с субјективним оценама од оцена метрике MW-PSNR.

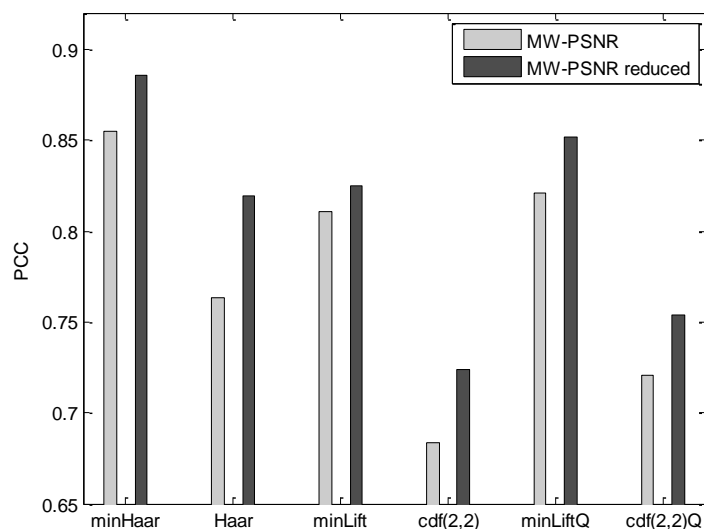


Слика 5.2.11. Пирсонов коефицијент корелације PSNR за најбољи подопсег према DMOS и MW-PSNR према DMOS за различите типове таласића

Анализа редуковане верзије MW-PSNR_r

При визуелном оцењивању квалитета слике, људи највећи значај дају ивицама. Човеков визуелни систем интегрише ивице слике од грубљих резолуција према финијим (од глобалног према локалном) [79]. У кори великог мозга се обавља интеграција кроз просторне фреквенције тако да се побољша репрезентација ивица. Како се ивице визуелно интегришу од грубљих скала према финијим скалама, визуелни квалитет слике се може очувати коришћењем слика грубљих резолуција.

На основу перформанси метрике PSNR по подопсезима, које су знатно боље за подопсеге виших нивоа декомпозиције од 4 до 7, предложена је редукована верзија MW-PSNR_r која користи само те подопсеге. Применом сепарабилне декомпозиције редукована верзија MW-PSNR_r се израчунава коришћењем само подопсега d41-d72. Применом несепарабилне декомпозиције са узорковањем на “quincunx“ решетки редукована верзија MW-PSNR_r се израчунава коришћењем само подопсега d42-d71. Перформансе редуковане верзије MW-PSNR_r су приказане у доњем делу табеле 5.2.5.



Слика 5.2.12.Пирсонов коефицијент корелације редуковане MW-PSNRr и основне верзије MW-PSNR према DMOS за различите типове морфолошких и линеарних таласића

Пирсонови коефицијенти корелације (PCC) редуковане верзије метрике MW-PSNRr према DMOS за различите типове таласића су приказане на графику на слици 5.2.12. Најбоље перформансе метрике MW-PSNRr су добијене применом сепарабилне декомпозиције са морфолошким таласићем minHaar, PCC=88.5%, SCC= 82.98%. За све типове таласића, перформансе редуковане верзије MW-PSNRr су боље од перформанси основне верзије MW-PSNR.

Преглед резултата

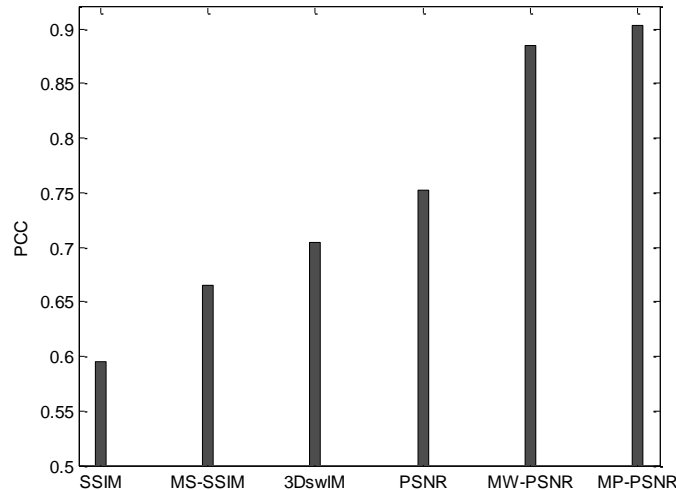
Перформансе стандардних метрика за оцену квалитета слика као што су PSNR, Universal Quality Index, UQI [80], Structural Similarity index SSIM [81], Multi-Scale Structural Similarity, MS-SSIM [42], Information Weighted PSNR, IW-PSNR [45], Information Weighted Structural Similarity IW-SSIM [45] и метрике намењене оцени квалитета синтетизованих слика помоћу ДИБР алгоритама синтезе, 3DswIM [28], су приказани у табели 5.2.7.

Табела 5.2.7. Перформансе стандардних метрика квалитета слика и предложених метрика

	IRCCyN/IVC DIBR		
	RMSE	PCC	SCC
стандардне мере			
PSNR	0.4525	0.7519	0.6766
IW-PSNR [99]	0.5267	0.6411	0.5320
UQI [97]	0.5199	0.6529	0.5708
SSIM [98]	0.5513	0.5956	0.4424
MS-SSIM [99]	0.5127	0.6649	0.5188
IW-SSIM [99]	0.5350	0.6265	0.4856
мера синтет. слика			
3DswIM [100]	0.4868	0.7049	0.6396
предложене мере			
PSNR (d4), SE=3x3	0.3076	0.8939	0.8671
MP-PSNRr SE=5x5	0.2936	0.9039	0.8634
MP-PSNR SE=7x7	0.3264	0.8796	0.8050
PSNR (d61), minLift	0.3167	0.8872	0.8281
MW-PSNRr, minHaar	0.3188	0.8855	0.8298
MW-PSNR, minHaar	0.3565	0.8545	0.7750

У табели су приказане перформансе предложених метрика заснованих на морфолошкој пирамидалној декомпозицији: PSNR слике детаља петог нивоа декомпозиције коришћењем SE димензија 3x3; редукована верзија MP-PSNRr коришћењем SE димензија 5x5; основна верзија MP-PSNR коришћењем SE димензија 7x7.

Од предложених метрика заснованих на декомпозицији морфолошким таласићима у табели су приказане перформансе следећих метрика: PSNR подопсега детаља који садржи вертикалне ивице шестог нивоа декомпозиције применом таласића minLift; редукована верзија MW-PSNRr и основна верзија MW-PSNR засноване на декомпозицији слика морфолошким таласићима minHaar. Перформансе предложених метрика су знатно боље од перформанси стандардних метрика за оцену квалитета слика и много боље од метрике намењене оцени квалитета синтетизованих слика 3DswIM. Пирсонови коефицијенти корелације ових метрика према DMOS су приказани на слици 5.2.13.



Слика 5.2.13.Пирсонов коефицијент корелацијестандардних метрика (SSIM, MS-SSIM, PSNR), метрике намењене сликама са изобличењем због ДИБР синтезе (3DswIM) и редукованих верзија предложене метрике MP-PSNRи MW-PSNRr

5.3. ТЕСТ ПРИМЕР 2: БАЗА СИНТЕТИЗОВАНИХ СТЕРЕО СЛИКА MCL-3D

У следећој секцији је описана база стерео слика синтетизованих применом ДИБР алгоритама, MCL-3D, а у секцији 5.3.2 су приказане перформансе предложене метрике при оцењивању ове базе.

5.3.1 Опис базе MCL-3D

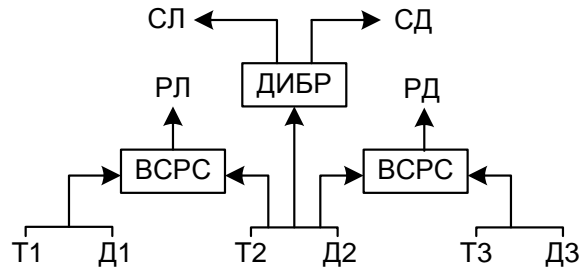
У тестовима је коришћен део базе MCL-3D [82], [83], који садржи 36 парова стерео слика са изобличењима због процеса ДИБР синтезе. Стерео слике су синтетизоване на основу фрејмова из 9 MVD видео секвенци приказаних на слици 5.3.1. Три видео секвенце, Kendo, Baloons и Lovebird1, садрже фрејмове резолуције 1024x728 док је резолуција фрејмова 6 видео секвенци Poznan street, Poznan Hall2, Shark, Microworld, Gt_fly, Undo_dancer 1920x1088 пиксела.

За креирање базе MCL-3D из сваког низа оригиналних MVD видео секвенци коришћене су 3 суседне секвенце и из сваке од њих су изабрани фрејмови текстуре (T1, T2, T3) и фрејмови мапа дубине (D1, D2, D3). На основу фрејма T2 и њему одговарајуће мапе дубине D2, слика 5.3.2, применом ДИБР алгорита синтезе синтетизоване су 2 слике, лева СЛ и десна СД, које чине стерео пар.

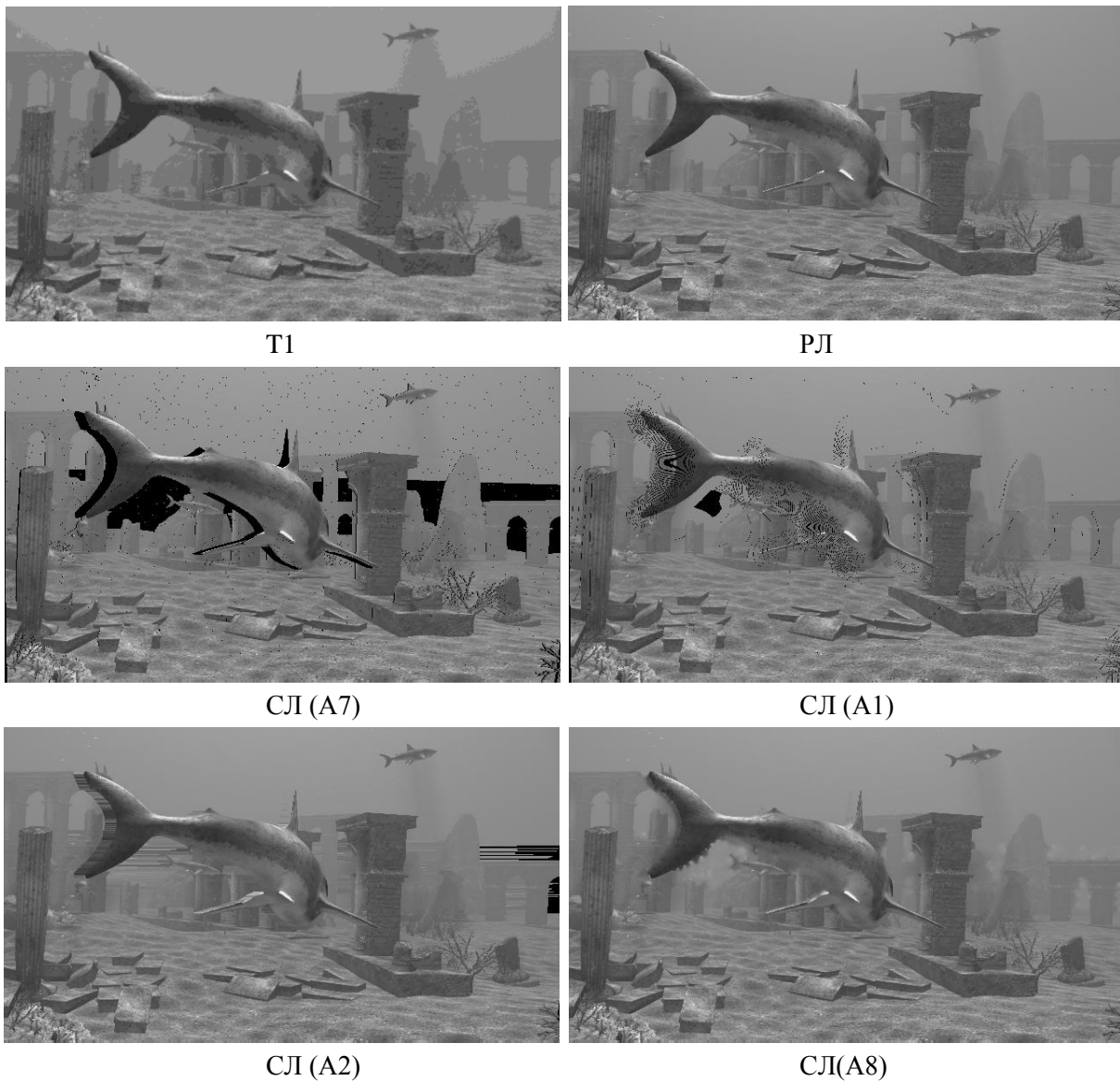


Слика 5.3.1. Оригиначне слике базе MCL-3D: Kendo, Baloons Lovebird1, Poznan street, Poznan Hall2, Shark, Microworld, Gt_fly, Undo_dancer

За генерисање базе MCL-3D коришћено је 4 ДИБР алгоритма синтезе: A1, A2, A7 и A8. Алгоритам A1 и A2 примењује метод [12]. Применом алгоритма A2 ивична област слике се попуњава применом алгоритма за попуњавање празнина [70]. A7 је алгоритам синтезе без попуњавања празних области које настају након пројекције у процесу ДИБР синтезе. A8 је мултирезолуциони алгоритам за попуњавање празних области [75]. Алгоритам A8 садржи пирамидалну декомпозицију синтетизоване слике са празним областима због “дисоклузија”. Празне области се естимирају на основу слика нижих резолуција и тако се добија слика са попуњеним празнинама без геометријских изобличења. На основу парова текстура, мапа дубине на две позиције, $(T1, D1)$ и $(T2, D2)$ синтетизована је лева референтна слика, РЛ, слика 5.3.2, и на основу $(T2, D2)$ и $(T3, D3)$ синтетизована је десна референтна слика, РД, референтног стерео пара, применом референтног софтвера за синтезу слика ВСРС [84]. Оригинална слика, T1, референтна лева слика стерео пара, РЛ, генерисана применом референтног софтвера ВСРС, леве слике стерео парова синтетизованих применом 4 ДИБР алгоритма синтезе СЛ(A1), СЛ(A2), СЛ(A7), СЛ(A8) су приказане на слици 5.3.3 за Shark и на слици 5.3.4 за Baloons.



Слика 5.3.2. Генерисање стерео слика базе MCL-3D. Стерео пар слика (СЛ, СД) је синтетизован на основу слике Т2 и њој одговарајуће мапе дубине Д2. Референтни стерео пар слика (РЛ, РД) је синтетизован на основу парова оригинално снимљених слика (Т1, Т2) и (Т2, Т3) и њима одговарајућих парова мапа дубине (Д1, Д2) и (Д2, Д3).



Слика 5.3.3. Shark: оригинална слика Т1, референтна лева слика РЛ, лева слике стерео парова синтетизоване применом ДИБР алгоритама А7, А1, А2, А8



T1



P1



СЛ (A7)



СЛ (A1)



СЛ (A2)



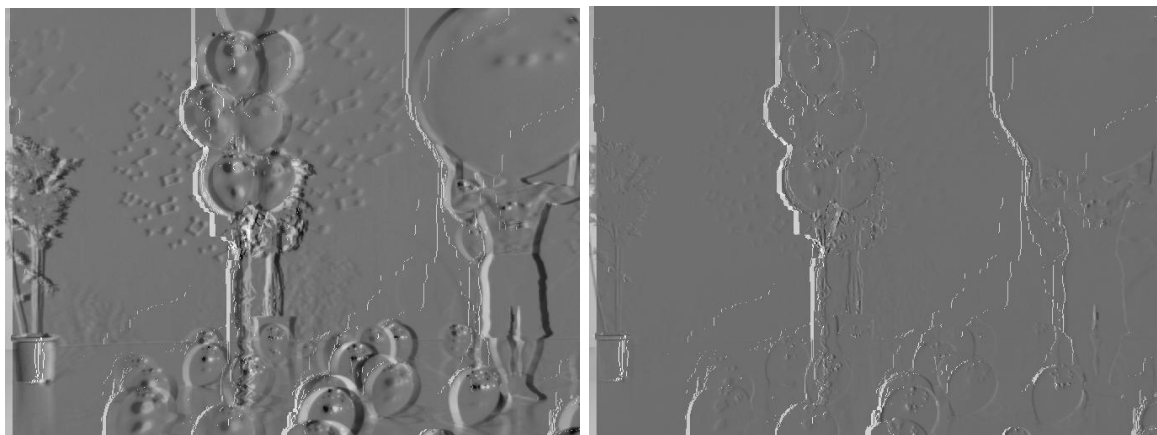
СЛ(A8)

Слика 5.3.4. Baloons: оригинална слика T1, референтна лева слика P1, леве слике стерео парова синтетизоване применом ДИБР алгоритама A7, A1, A2, A8

Поређењем синтетизованог фрејма СЛ и оригиналног фрејма T1 са слике 5.3.2 може се закључити да не одговарају истим позицијама, што се види на сликама у левој колони

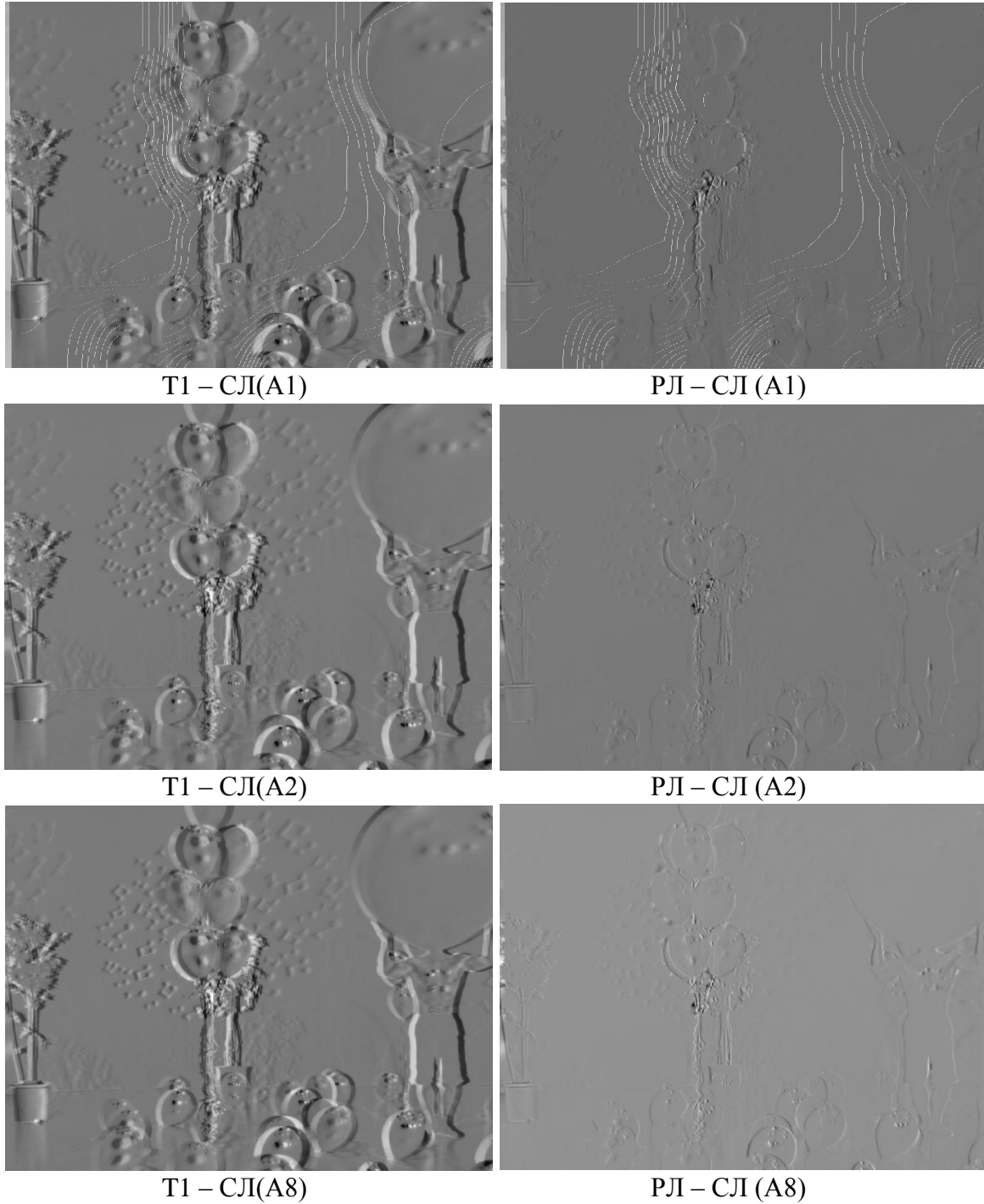
слике 5.3.5. На сликама у левој колони слике 5.3.5 које представљају разлику оригиналне слике и синтетизованих левих слика, T1 - СЛ(Ах), уочавамо бела и црна подебљања дуж вертикалних ивица која означавају померај, тј. разлику у позицијама. Слично важи и за десне слике СД, ТЗ, РД. Како на позицијама за које су синтетизоване слике СЛ и СД не постоје оригиналне слике, као референтне слике за израчунавање метрике се користе синтетизоване слике РЛ и РД, слика 5.3.2, генерисане применом референтног софтвера за синтезу слика ВСРС [84] на позицијама које одговарају позицијама слика СЛ и СР. Референтни стерео пар (СЛ, СД) је генерисан на основу оригиналних парова (Т1, Д1) и (Т2, Д2) за СЛ и (Т2, Д2) и (Т3, Д3) за СД. Разлика референтне леве слике, РЛ, и леве слике синтетизоване ДИБР алгоритмом, СЛ(Ах), приказана је у десној колони слике 5.3.5. На основу слика у десној колони слике видимо да нема помака позиција слике СЛ(Ах) и РЛ. Разлика ове две слике постоји због тога што је слика РЛ синтетизована на основу две слике на позицијама лево и десно у односу на синтетизовану слику па садржи мање изобличења због процеса синтезе него слика СЛ(Ах) која је генерисана на основу само једне оригиналне слике као и због различитих алгоритама ДИБР синтезе. Слично важи и за десне слике СД и РД.

Метрике су рачунате између синтетизованог стерео пара (СЛ, СД) и референтног стерео пара (РЛ, РД). Резултат стерео пара је добијен усредњавањем резултата метрике за леву и десну слику.



T1 – СЛ(А7)

РЛ – СЛ (А7)



Слика 5.3.5. Сlike синтетизоване ДИБР алгоритмима СЛ(A1), СЛ(A2), СЛ(A7), СЛ(A8) су на различитим позицијама у односу на оригиналне сlike Т1, лева колона; референтне сlike РЛ генерисане помоћу референтног софтвера ВСРС су на истој позицији као сlike синтетизоване ДИБР алгоритмима А1, А2, А7, А8 (десна колона);

База садржи и субјективне оцене стерео слика. Субјективно оцењивање базе стерео слика MCL-3D је урађено према ИТУ препоруци о субјективном оцењивању стерео ТВ слика [85]. За приказ стерео слика коришћен је екран 46.9” LG47LW5600. Посматрачу су приказиване по две стерео слике истовремено (pair-wise comparison method) и он је изабрао бољу и доделио јој оцену. Оцене из више приказа парова стерео слика су сабране и нормализоване и тако је добијена коначна оцена за тог посматрача. Оцене од више посматрача су усредњене и добијена је средња оцена (mean opinion score) за стерео пар слика. За сваку стерео слику је добијено 30 оцена од којих је 10% највећих и 10% најмањих одбачено тако да је МОС оцена израчуната на основу 24 појединачне оцене.

5.3.2 Перформансе предложене метрике при оцењивању слика базе MCL-3D

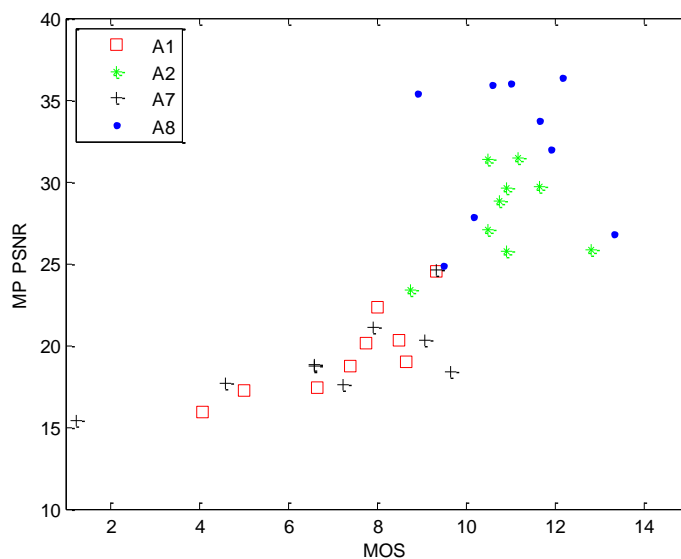
Предложена мултирезолуциона метрика са морфолошким филтрима је тестирана на бази MCL-3D која садржи стерео слике синтетизоване применом ДИБР алгоритама синтезе. Резултати метрика су рачунати посебно за леву и за десну слику. Резултат стерео пара је израчунат усредњавањем резултата метрике за леву и за десну слику.

Перформансе метрике MP-PSNR

Испитан је утицај величине структурног елемента и броја нивоа декомпозиције на перформансе предложене метрике. Перформансе метрике MP-PSNR су приказане у табели 5.3.1. За базу MCL-3D, перформансе метрике MP-PSNR не зависе од димензија структурног елемента морфолошких филтара пирамидалне декомпозиције. Како процес синтезе слике утиче на повећање високофреквентних компонената слике, референтна слика која је синтетизована применом референтног софтвера за синтезу такође садржи високофреквентне компоненте и разлика слика синтетизованих применом једног од 4 алгорита синтезе и референтне слике је мања него разлика синтетизоване и оригиналне (снимљене) слике. График расподеле MP-PSNR према MOS за стерео слике базе MCL-3D је приказан на слици 5.3.6. Свака тачка одговара једном пару стерео слика.

Табела 5.3.1. Перформансе основне верзије MP-PSNR и редуковане верзије MP-PSNRr при оцењивању базе MCL-3D

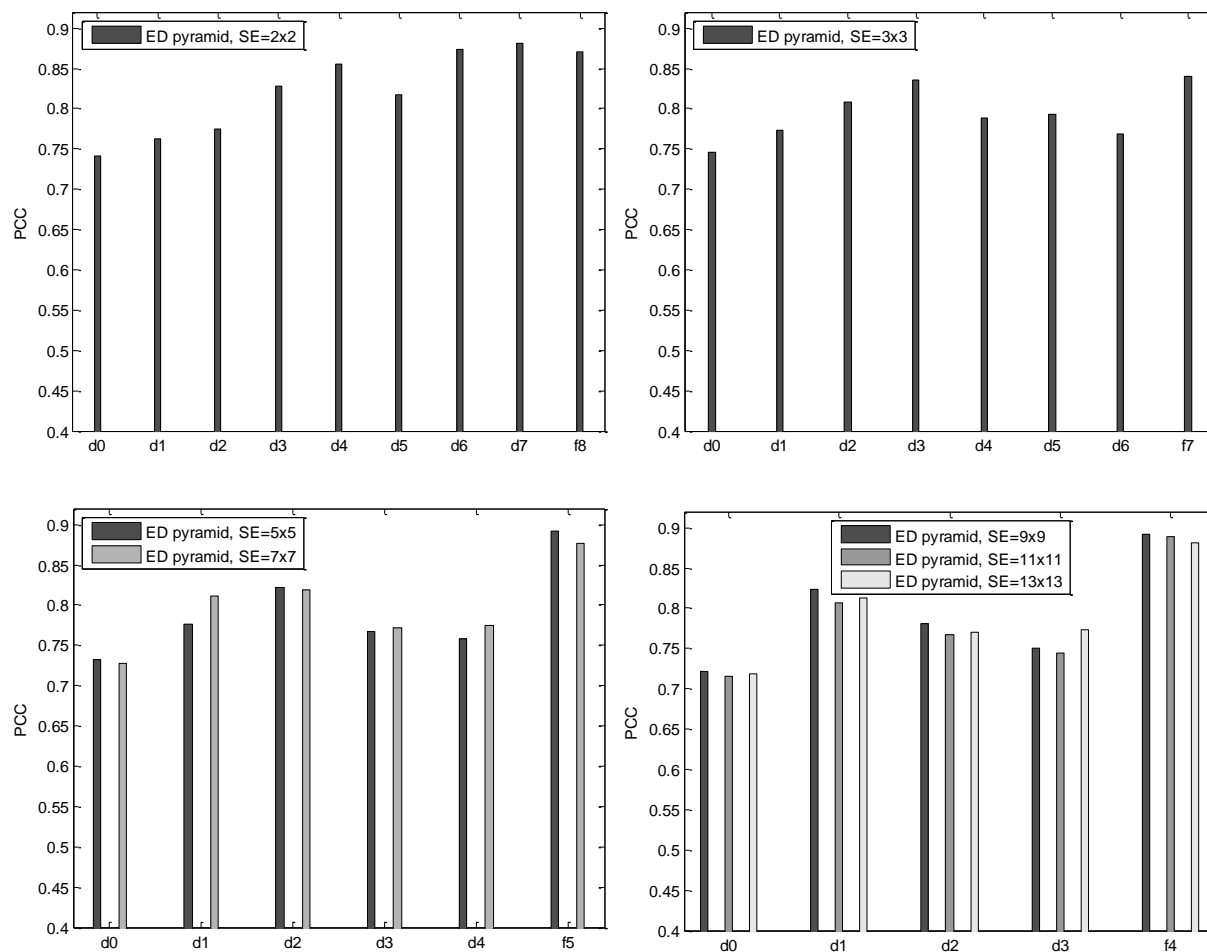
MCL-3D				
SE	нивоа	RMSE	PCC	SCC
MP-PSNR				
2x2	8	1.2506	0.8902	0.8131
3x3	7	1.2270	0.8946	0.8503
5x5	5	1.2305	0.8939	0.8360
7x7	5	1.2274	0.8945	0.8430
9x9	4	1.2304	0.8940	0.8395
11x11	4	1.2298	0.8941	0.8373
13x13	4	1.2121	0.8973	0.8389
MP-PSNRr				
2x2	3-9	1.2378	0.8926	0.8142
3x3	3-8	1.2014	0.8992	0.8481
5x5	3-6	1.2267	0.8946	0.8445
7x7	2-6	1.2097	0.8977	0.8502
9x9	2-5	1.2242	0.8951	0.8398
11x11	2-5	1.2261	0.8947	0.8389
13x13	2-5	1.2047	0.8986	0.8467



Слика 5.3.6. Расподела MP-PSNR према MOS за стерео слике базе MCL-3D.

Перформансе метрике PSNR по сликама морфолошке пирамиде

Анализиране су перформансе PSNR метрике за слике пирамиде по нивоима декомпозиције. На графицима на слици 5.3.7 су приказане вредности Пирсоновог коефицијента корелације метрике PSNR према MOS по сликама пирамиде за структурне елементе морфолошких филтара различитих димензија, 2x2 до 13x13, коришћених у пирамидалној декомпозицији слике. За MCL-3D базу, мале су разлике Пирсонових коефицијената корелације метрике PSNR према MOS по сликама пирамиде. Најмањи PCC је за слику детаља на првој скали, d_0 , а највећи за слику апроксимације на врху пирамиде.



Слика 5.3.7. Пирсонов коефицијент корелације мере PSNR по сликама пирамиде према MOS. Структурни елементи различитих димензија су коришћени за морфолошко филтрирање.

Перформансе PSNR метрике за слику пирамиде за коју су перформансе PSNR метрике најбоље, за различите величине структурних елемената, су приказане у табели 5.3.2.

Табела 5.3.2. Перформансе PSNR метрике слике пирамиде креиране применом структурних елемената различитих димензија за оцењивање базе MCL-3D

SE	MCL-3D			
	image	RMSE	PCC	SCC
2x2	d8	1.3095	0.8789	0.7988
3x3	s6	1.2327	0.8935	0.8427
5x5	s5	1.1959	0.9001	0.8583
7x7	s4	1.1696	0.9047	0.8733
9x9	s4	1.1777	0.9033	0.8724
11x11	s4	1.2012	0.8992	0.8691
13x13	s4	1.2431	0.8916	0.8556
-----	f0	2.4566	0.4465	0.4537

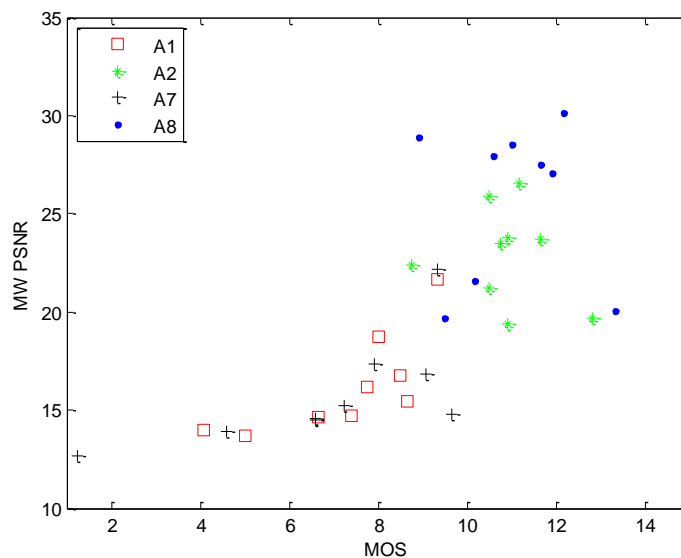
За базу MCL-3D, редукована верзија MP-PSNR_r је израчуната на основу слика пирамиде без слика детаља са прва два нивоа декомпозиције када је коришћен структурни елемент димензија мањих од 7x7. За структурне елементе димензија већих од 7x7, редукована верзија MP-PSNR_r је израчуната на основу слика пирамиде без слике детаља са првог нивоа декомпозиције. Перформансе редуковане верзије MP-PSNR_r су приказане у доњем делу табеле 5.3.1. Незнатно побољшање перформанси је постигнуто применом редуковане верзије MP-PSNR_r у односу на основну верзију MP-PSNR за базу MCL-3D.

Перформансе метрике MW-PSNR

Предложена метрика MW-PSNR заснована на декомпозицији слике морфолошким таласићима је тестирана на бази стерео слика MCL-3D синтетизованих применом ДИБР алгоритама. Перформансе метрике MW-PSNR за различите типове таласића су приказане у табели 5.3.2. Најбоље перформансе метрике MW-PSNR, PCC=0.88, SCC=0.80, су остварене применом несепарабилне декомпозиције са узорковањем на “quincunx“ решетки применом морфолошких таласића minLiftQ. Расподела вредности MW-PSNR применом морфолошких таласића minLiftQ за слике базе MCL-3D је приказана на слици 5.3.8.

Табела 5.3.2. Перформансе метрике MW-PSNR и редуковане верзије MW-PSNRr при оцењивању базе MCL-3D

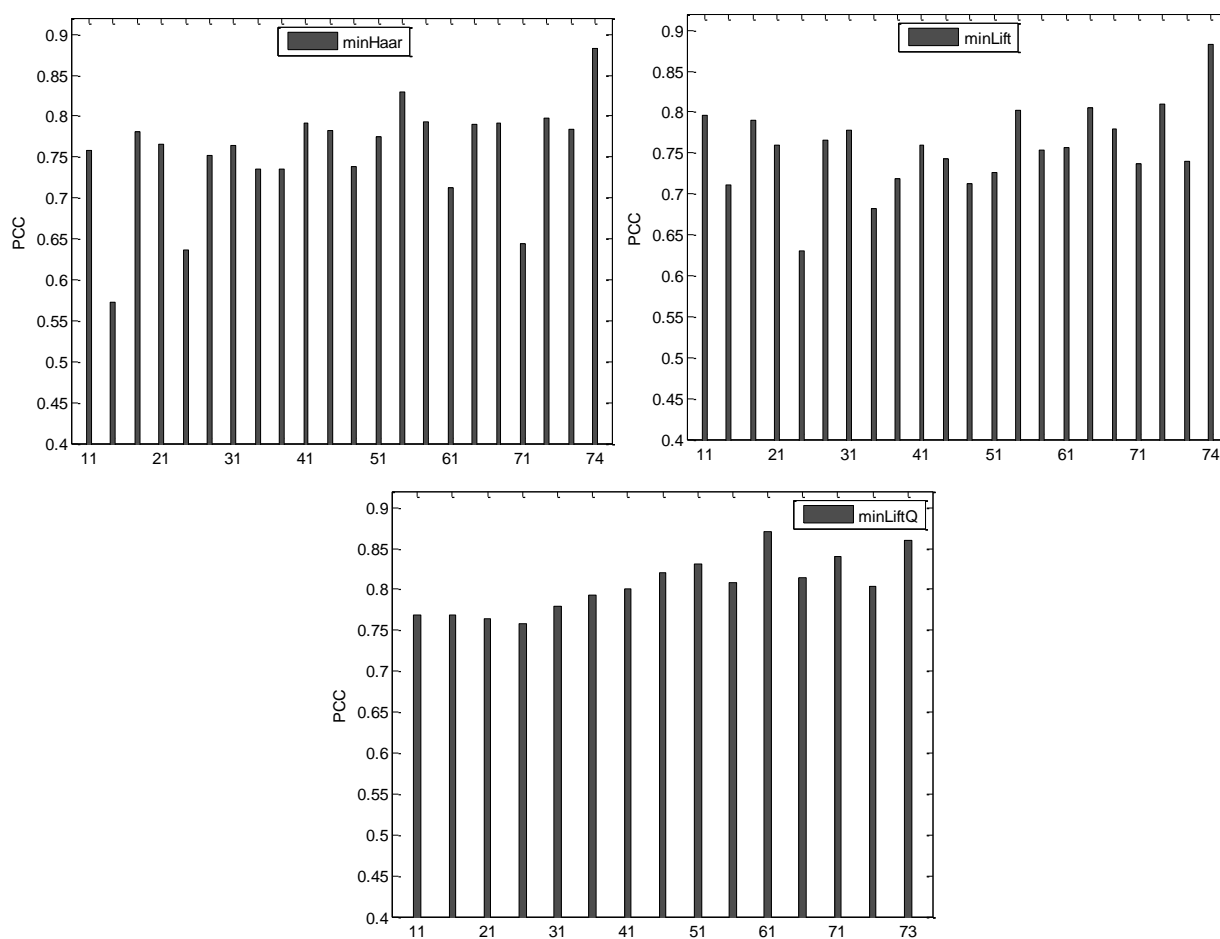
MCL-3D			
	RMSE	PCC	SCC
MW-PSNR			
minHaar	1.3063	0.8796	0.7970
Haar	2.2524	0.5718	0.6065
minLift	1.3823	0.8640	0.7909
cdf(2,2)	2.2916	0.5507	0.5849
minLiftQ	1.3035	0.8801	0.8064
cdf(2,2)Q	2.3691	0.5053	0.5101
MW-PSNRr			
minHaar	1.2962	0.8815	0.8070
Haar	1.8452	0.7405	0.7412
minLift	1.3786	0.8648	0.8130
cdf(2,2)	1.7043	0.7840	0.7749
minLiftQ	1.2700	0.8866	0.8142
cdf(2,2)Q	1.8794	0.7290	0.7485



Слика 5.3.8. Расподела MW-PSNR према MOS за стерео слике базе MCL-3D.

Перформансе метрике PSNR по подопсезима добијеним декомпозицијом слике морфолошким таласићима

Анализиране су перформансе PSNR метрике по подопсезима добијеним декомпозицијом слике морфолошким таласићима. Као мера поклапања резултата PSNR метрике са субјективним оценама, израчунат је Пирсонов коефицијент корелације, PCC, мере PSNR према DMOS. На графицима на слици 5.3.9 приказане су вредности PCC по подопсезима за три типа морфолошких таласића, minHaar, minLift и minLiftQ. Мале су разлике вредности PCC по подопсезима слике добијеним декомпозицијом слике таласићима.



Слика 5.3.9. Пирсонови коефицијенти корелације (PCC) метрике PSNR према DMOS по подопсезима добијеним декомпозицијом синтетизованих слика морфолошким таласићима minHaar, minLift и minLiftQ

Табела 5.3.3. Перформансе метрике PSNR за подопсег детаља добијен декомпозицијом слике различитим типовима таласића при оцењивању базе MCL-3D

декомпозиција	таласић	MCL-3D			
		подопсег	RMSE	PCC	SCC
сепарабилна	minHaar	s7	1.2858	0.8835	0.8465
	Haar	d13	1.7036	0.7842	0.7392
	minLift	s7	1.2866	0.8834	0.8408
	cdf(2,2)	d11	1.8400	0.7422	0.7440
	minLiftQ	s7	1.3051	0.8798	0.8354
несепарабилна	cdf(2,2)Q	d1	1.7024	0.7845	0.7674
-----	-----	основни	2.4566	0.4465	0.4537

Перформансе метрике PSNR за подопсег за који су PSNR перформансе најбоље, за све типове таласића, су приказане у табели 5.3.3. Применом морфолошких таласића у декомпозицији синтетизованих слика најбоље перформансе метрике PSNR су за слике апроксимација последњег нивоа декомпозиције. Применом сепарабилне декомпозиције слике морфолошким таласићем minHaar, остварене су добре перформансе метрике PSNR за слику апроксимације s7, PCC=0.883, SCC=0.846. Перформансе метрике PSNR израчунате за подопсеге слике добијене декомпозицијом слике морфолошким таласићима су знатно боље од перформанси метрике PSNR израчунате за слику у основном опсегу у пуној резолуцији.

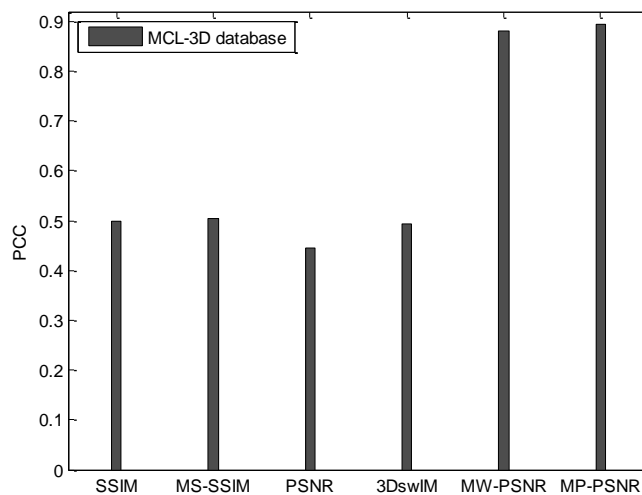
Перформансе редуковане верзије MW-PSNR_r су приказане у доњем делу табеле 5.3.2. Незнатно побољшање перформанси је постигнуто применом редуковане верзије MW-PSNR_r у односу на основну верзију MW-PSNR за базу MCL-3D.

Преглед резултата предложених метрика

Перформансе предложених метрика у оцењивању базе MCL-3D су знатно боље од перформанси стандардних метрика за оцену квалитета слика и много боље од метрике намењене оцени квалитета синтетизованих слика 3DswIM, табела 5.3.4. Пирсонови коефицијенти корелације ових метрика према MOS су приказани на слици 5.3.10.

Табела 5.3.4. Перформансе стандардних метрика квалитета слика и предложених метрика при оцењивању базе MCL-3D

метрике	MCL-3D		
	RMSE	PCC	SCC
стандардне метрике			
PSNR	2.4566	0.4465	0.4537
IW-PSNR [99]	2.4541	0.4483	0.2892
UQI [97]	2.3292	0.5294	0.3177
SSIM [98]	2.3797	0.4987	0.1739
MS-SSIM [99]	2.0834	0.6513	0.2091
IW-SSIM [99]	2.1908	0.6027	0.1578
за DIBR синтетиз. слике			
3DswIM [100]	2.3879	0.4935	0.3435
предложене мере			
PSNR(s6), SE=3x3	1.2327	0.8935	0.8427
MP-PSNRr SE=5x5	1.2267	0.8946	0.8445
MP-PSNR SE=7x7	1.2274	0.8945	0.8430
PSNR(s7), minLift	1.2866	0.8834	0.8408
MW-PSNRr, minHaar	1.2962	0.8815	0.8070
MW-PSNR, minHaar	1.3063	0.8796	0.7970



Слика 5.3.10. Пирсонови коефицијенти корелације стандардних метрика (SSIM, MS-SSIM, PSNR), метрике намењене сликама са изобличењем због ДИБР синтезе (3DswIM) и предложених метрика MP-PSNR и MW-PSNR за базу стерео слика MCL-3D

5.4. ТЕСТ ПРИМЕР 3: БАЗА СИНТЕТИЗОВАНИХ ВИДЕО СЕКВЕНЦИ FVSV

Прво је описана база видео секвенци FVSV а затим су приказани резултати предложене метрике при оцењивању видео секвенци базе FVSV.

5.4.1 Опис базе FVSV

База видео секвенци (Free-Viewpoint video database) FVSV, [30], [31], садржи видео секвенце са навигацијом кроз различите позиције синтетизоване на основу некомпримованих фрејмова из снимљених видео секвенци и одговарајућих декомпримованих мапа дубине. Видео секвенце базе садрже изобличења због процеса ДИБР синтезе и због компресије мапа дубине. Видео секвенце су креиране тако да садрже навигацију кроз различите позиције. Свака видео секвенца базе је креирана на основу два фрејма из две видео секвенце снимљене на различитим позицијама и њима одговарајућих декомпримованих мапа дубине. Применом референтног софтвера за синтезу слика (View Synthesis Reference Software 1D Fast, VSRS-1DFast [87]) у моду мешано (blended) синтетизовано је 49 фрејмова на 49 позиција између два снимљена фрејма. Софтвер за ДИБР-синтезу слика подржава два мода, мод са мешањем и мод без мешања. Мод са мешањем (blended) комбинује обе слике, леву и десну и додељује им тежинске факторе пропорционално удаћености док мод без мешања (non-blended) користи једну снимљену слику за пројекцију на нову позицију а другу за попуњавање празних делова у синтетизованој слици насталих због “дисоклузија”. На основу два фрејма из снимљених видео секвенци и 49 синтетизованих фрејмова креирана је видео секвенца која садржи 100 фрејмова тако да је 49 синтетизованих фрејмова поређано прво у једном смеру промене позиција а затим у супротном смеру, као што је приказано на слици 5.4.1. На тај начин се симулира навигација кроз позиције посматрања од једне до друге стране и обрнуто, враћање ка почетној позицији. Видео секвенца не садржи временску дисторзију јер сви фрејмови приказују исти временски тренутак.



Слика 5.4.1. Видео секвенца FVSV базе синтетизована у моду мешање (blend) садржи 100 фрејмова: r1, r2 снимљени фрејмови, s1-s49 ДИБР-синтетизовани фрејмови

Како су видео секвенце синтетизоване у моду са мешањем (blended) бољег квалитета од видео секвенци синтетизованих у моду без мешања (non-blended), у тестовима су коришћене само 103 видео секвенце синтетизоване у моду са мешањем. За генерисање видео секвенци FVSV базе коришћени су фрејмови из 5 MVD видео секвенци, од којих су 3 снимљене камерама а 2 рачунарски генерисане. Док мапе дубине реалних (снимљених) видео секвенци садрже грешке због несавршености уређаја за снимање или алгорита за естимацију, мапе дубине синтетички генерисаних секвенци су без тих грешака. У табели 5.4.1 су приказане MVD секвенце, број камера и њихова удаљеност при снимању сваке секвенце, резолуција снимљених слика, позиције видео секвенци из којих су узети фрејмови за синтезу и број фрејма у секвенци који је коришћен за синтезу.

Табела 5.4.1. MVD секвенце FVSV базе

секвенца	резолуција	број камера	удаљеност камера	позиције снимљених фрејмова r1, r2	број фрејма
Book Arrival	1024x768	16	6.5 цм	6, 10	33
Kendo		7	5 цм	1, 5	1
Balloons		7	5 цм	1, 5	1
Undo Dancer	1920x1080	рачунарски генерисане		1, 9	250
GT Fly				9, 1	157

Видео секвенце FVSV базе су генерисане на основу оригиналних некопримованих слика и њима одговарајућих декомпримованих мапа дубине. Мапе дубине су кодоване применом 7 кодека, сваки са по 3 квантизациона нивоа. Коришћени су следећи кодеци:

- C1: 3D-HEVC Test Model, 3D-HTM 0.4 [88], са предикцијом између погледа (inter-view prediction) и оптимизацијом синтезе (View Synthesis Optimization);
- C2: Multiview Video Coding (MVC), JM 18.4 [89];
- C3: HEVC Test Model, HM 6.1 [90];
- C4: JPEG2000, Kakadu имплементација [91];
- C5: кодек описан у [92], мапе дубине се кодују тако да су ивице кодоване без изобличења (lossless-edge depth map coding);
- C6: кодек описан у [93];
- C7: Z-LAR-RP [94].

Сви алгоритми кодовања су примењени у интра моду (intra coding). Изабрана су 3 параметра квантизације на основу визуелног квалитета синтетизованих слика [86].

База садржи и секвенце генерисане на основу оригиналних некомпримованих слика и њима одговарајућих некомпримованих мапа дубине које се користе као референтне секвенце код израчунавања метрика за оцену квалитета.

За све видео секвенце постоје субјективне оцене (Mean opinion score) МОС. За добијање субјективних оцена видео секвенци FVSV базе примењена је методологија ACR-HR[95]. Свака видео секвенца је приказана једном сваком посматрачу на екрану Panasonic BT-3DL2550 резолуције 1920x1080p према [96]. У оцењивању је учествовало 27 посматрача који су додељивали оцене од 1 (лош) до 5 (одличан) посматраним секвенцама. За сваког посматрача тест је трајао 30 минута. Оцене за 4 посматрача су одбачене као бескорисне (outliers). Референтне секвенце су такође оцењиване (сакривена референца).

У тестовима је коришћен део FVSV базе који садржи 103 видео секвенце синтетизоване коришћењем мода мешање (blended) на основу 5 MVD видео секвенци, 7 кодека за кодовање мапа дубине са по 3 нивоа квантизације (5 MVD секвенци * 7 кодека * 3 квантизациона нивоа – 2 неисправне сквенце).

5.4.2. Перформансе предложене метрике при оцењивању базе FVSV

Предложена мултирезолуциона метрика са морфолошким филтрима је тестирана на бази FVSV која садржи видео секвенце са изобличењима због процеса синтезе и компресије мапа дубине креиране тако да садрже навигацију кроз различите позиције. Како оригинално снимљене секвенце за 50 позиција не постоје, као референтне видео секвенце коришћене су видео секвенце базе синтетизоване на основу некомпримованих (оригиналних, снимљених) фрејмова и некомпримованих мапа дубине. На тај начин се мери утицај компресије мапа дубине на изобличења због процеса синтезе.

У табели 5.4.2 су приказане перформансе предложених метрика, MP-PSNR и MW-PSNR, и стандардних метрика за оцену квалитета слика (PSNR, SSIM, MS-SSIM) добијене при оцењивању дела базе FVSV који садржи видео секвенце синтетизоване применом мода мешање (blended mode). Метрика MP-PSNR је реализована применом морфолошке пирамидалне декомпозиције са 6 нивоа. Структурни елемент димензија 3x3 је коришћен за

морфолошко филтрирање пирамидалне декомпозиције. Метрика MW-PSNR је реализована применом морфолошких Харових таласића у декомпозицији са 7 нивоа. Морфолошке мултирезолуционе метрике показују боље перформансе од стандардних метрика. Мере квалитета слике засноване на локалној статистици као што је SSIM, нису погодне за мерење геометријских изобличења због грешака мапа дубине [15]. Када повећавамо интензитет изобличења у синтетизованој слици компримовањем мапе дубине већим кораком квантизације, вредност SSIM се мало мења. То је због тога што су дисторзије синтезе углавном геометријске дисторзије помераја које неће значајно променити просечну вредност осветљаја и варијансу у локалној области.

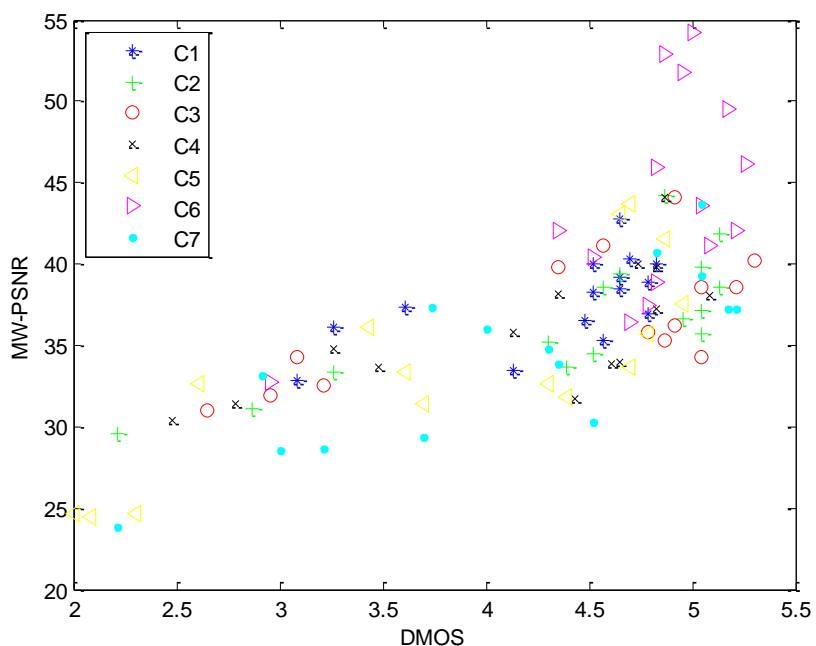
Табела 5.4.2. Перформансе стандардних метрика квалитета слика и предложених метрика у оцењивању дела базе FVSV синтетизоване применом мода мешање (blended mode)

метрика	RMSE	PCC	SCC
MW-PSNR	0.5259	0.8043	0.7343
MP-PSNR	0.5330	0.7983	0.7324
PSNR	0.5906	0.7448	0.6826
SSIM [98]	0.6661	0.6585	0.6090
MS-SSIM [99]	0.7152	0.5891	0.5912

Да би одредили поузданост метрике за оцењивање слика синтетизованих на основу мапа дубине кодovаних са 7 различитих кодера, део базе који се користи у тестовима је подељен на 7 делова. Сваки део садржи видео секвенце синтетизоване коришћењем мапа дубине кодovаних једним од кодера C1-C7. Перформансе метрике MW-PSNR у оцењивању 7 делова базе FVSV су приказани у горњем делу табеле 5.4.3. Перформансе метрике PSNR коришћене за оцењивање истих делова базе су приказане у доњем делу табеле 5.4.3. За све кодере осим C4, MW-PSNR показује боље перформансе него PSNR. Резултати метрике MW-PSNR се најбоље поклапају са субјективним оценама при оцењивању видео секвенци синтетизованих на основу мапа дубине кодovаних кодером C2. Расподела оцена метрике MW-PSNR према DMOS применом кодера C1-C7 за кодовање мапа дубине је приказана на слици 5.4.2. Свака тачка на графику одговара једној видео секвенци базе.

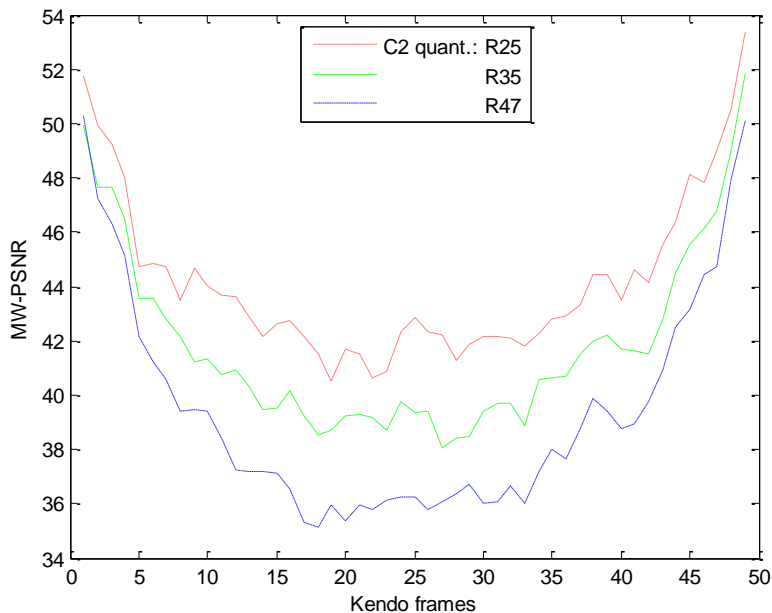
Табела 5.4.3. Перформансе метрика MW-PSNR и PSNR израчунатих за 7 делова FVSV базе који садрже видео секвенце синтетизоване применом мода мешање на основу мапа дубине кодovаних са 7 кодера C1- C7

кодер	RMSE	PCC	SCC
MW-PSNR			
C1	0.4787	0.6628	0.6679
C2	0.3461	0.9402	0.8032
C3	0.5430	0.8619	0.5705
C4	0.6112	0.7687	0.7437
C5	0.6140	0.8643	0.7757
C6	0.3067	0.8751	0.5541
C7	0.6122	0.8142	0.7542
PSNR			
C1	0.5804	0.4191	0.4165
C2	0.4919	0.8750	0.7299
C3	0.5845	0.8379	0.6013
C4	0.6091	0.7706	0.7459
C5	0.6845	0.8280	0.6988
C6	0.3197	0.8634	0.5326
C7	0.7032	0.7452	0.6184



Слика 5.4.2. Расподела резултата метрике MW-PSNR при оцењивању дела базе FVSV који садржи видео секвенце синтетизоване у моду мешање (blend) на основу мапа дубине кодovаних кодерима C1-C7

Утицај нивоа квантизације на квалитет синтетизованих слика је приказан на слици 5.4.3. Као што је и очекивано, на основу графика се може видети да је квалитет секвенци синтетизованих на основу мапа дубине кодованих мањим кораком квантизације бољи од квалитета секвенци синтетизованих на основу мапа дубине кодованих већим кораком квантизације. Поред тога, квалитет синтетизованих фрејмова опада са повећањем удаљености од референтних фрејмова.

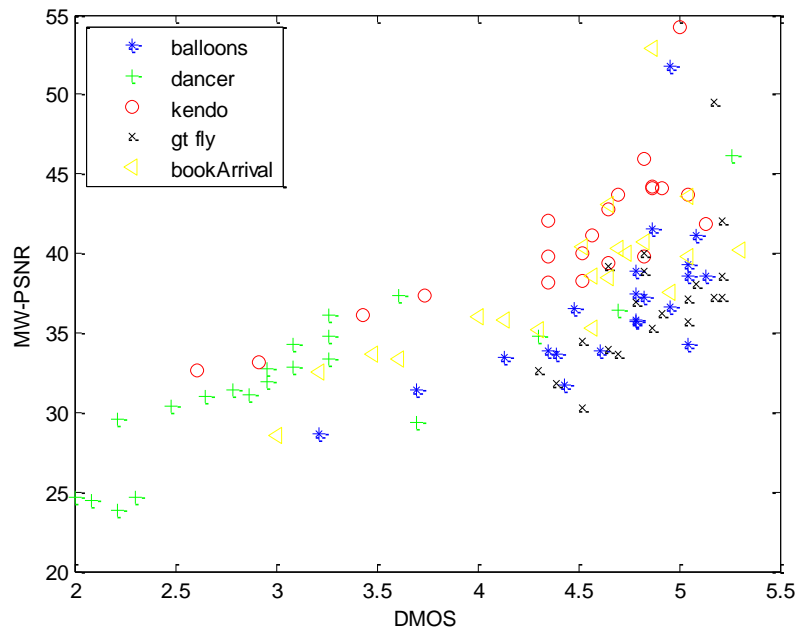


Слика 5.4.3. Квалитет видео секвенце Kendo синтетизоване на основу мапа дубине кодованих кодером C2 са 3 квантизациона нивоа: P25, P35, P47

Да би испитали зависност перформанси метрике по садржају, део FVSV базе синтетизован у моду мешање је подељен на 5 делова тако да су у сваком делу видео секвенце истог MVD садржаја. Перформансе метрике MW-PSNR за 5 делова базе различитог садржаја су приказане у табели 5.4.4. Метрика MW-PSNR показује најбоље поклапање са оценама људи за MVD садржај Kendo а најлошије за садржај Gt Fly. Расподела вредности метрике MW-PSNR према DMOS за различите садржаје је приказана на слици 5.4.4. Свака тачка на графику одговара једној видео секвенци базе.

Табела 5.4.4. Перформансе метрике MW-PSNR за 5 делова базе FVSV по MVD садржају

MW-PSNR	RMSE	PCC	SCC
Balloons	0.2220	0.9063	0.7994
Dancer	0.4566	0.8699	0.8526
Kendo	0.2330	0.9498	0.8184
BookArrival	0.3059	0.8994	0.7693
Gt_Fly	0.2092	0.7465	0.6767



Слика 5.4.4. Расподела вредности MW-PSNR по MVD садржају дела базе FVSV који садржи видео секвенце синтетизоване у моду мешање

Табела 5.4.5. Време израчунавања метрике за фрејм видео секвенце Kendo

метрика	време [s]
MW-PSNR [102]	0.15
MP-PSNR [101]	0.47
PSNR	0.01
SSIM [SSIMcode]	0.23
MS-SSIM [MS-SSIMcode]	0.41

Средње време израчунавања метрика по фрејму видео секвенце Kendo је приказано у табели 5.4.5. Време израчунавања је измерено на лаптопу HP ProBook 450 са процесором Intel Core i3, Dual Core и 4GB RAM. Коришћена софтверска платформа је Matlab R2010b. После PSNR, рачунски најефикаснија је метрика MW-PSNR.

6 ЗАКЉУЧАК

У докторској дисертацији је предложена мултирезолуциона метрика заснована на морфолошким филтрима за оцену квалитета слика и видео секвенци синтетизованих применом ДИБР (Depth Image Based Rendering) алгоритама синтезе. ДИБР технике синтезе слика узрокују геометријске дисторзије углавном у области ивица у синтетизованим сликама. Стандардне метрике за оцену квалитета слика не показују добре перформансе у оцењивању синтетизованих слика. Човеков визуелни систем обрађује визуелне информације на мултирезолуциони начин и примена мултирезолуционе декомпозиције слика при израчунавању квалитета слике доприноси бољим перформансама метрике. Човеков визуелни систем је веома осетљив на дисторзију ивица која значајно нарушава визуелни квалитет слике. Применом нелинеарних морфолошких филтара у мултирезолуционој декомпозицији слике важне геометријске информације као што су ивице остају очуване, без оштећења и без помака, у сликама кроз различите резолуционе скале. Израчунавањем средњих квадратних грешака између одговарајућих подопсега мултирезолуционе репрезентације који садрже ивице пиксел по пиксел, прецизно се мери разлика две мултирезолуционе репрезентације. Тако се највећи значај у оцени квалитета додељује области ивица које су склоне дисторзијама због процеса ДИБР синтезе.

Анализиране су две верзије мултирезолуционе метрике са морфолошким филтрима: MP-PSNR (Morphological Pyramid Peak Signal-to-Noise Ratio) која се заснива на морфолошкој пирамидалној декомпозицији и MW-PSNR (Morphological Wavelet Peak Signal-to-Noise Ratio) која користи декомпозицију слика морфолошким таласићима. Испитан је утицај сепарабилне декомпозиције за два типа морфолошких таласића применом лифтинг шеме на поузданост метрике. Утицај несепарабилне декомпозиције применом морфолошких таласића са узорковањем на “quincunx” решетки применом лифтинг шеме на поузданост метрике је испитан такође.

Развијене метрике су тестиране на три базе слика/видео: на бази слика синтетизованих различитим техникама ДИБР синтезе, IRCCyN/IVCDIBR, на бази стерео слика синтетизованих различитим техникама синтезе, MCL-3D, и на бази видео секвенци са

навигацијом кроз различите позиције, FVSV, које садрже изобличења због процеса синтезе и због компресије мапа дубине које се користе у процесу синтезе. Обе метрике, MP-PSNR и MW-PSNR, показују значајно боље перформансе од стандардних метрика за оцену квалитета слика као што су PSNR, SSIM, MS-SSIM и значајно боље перформансе од метрике намењене оцењивању квалитета синтетизованих слика 3DswIM.

Показано је да мера PSNR показује добро поклапање са субјективним оценама када се израчунава на подопсезима детаља виших нивоа морфолошке мултирезолуционе репрезентације синтетизоване слике.

Предложене су редуковане верзије морфолошких мултирезолуционих метрика, MP-PSNR_r и MW-PSNR_r, које користе само подопсеге виших нивоа декомпозиције. Редуковане верзије метрика показују боље перформансе од основних верзија.

Морфолошки оператори коришћени у мултирезолуционој декомпозицији слика су једноставни за имплементацију јер користе само операције минимум и максимум, раде са целим бројевима и као резултат дају целе бројеве. Израчунавање средње квадратне грешке која се користи као мера изобличења по подопсезима мултирезолуционе репрезентације слике је такође веома једноставно што за последицу има да је предложена метрика рачунски врло ефикасна.

У наставку истраживања предложена метрика ће бити тестирана на новој бази видео секвенци синтетизованих применом технике ДИБР синтезе које садрже и временску дисторзију. За боље прилагођење временским дисторзијама видео секвенце, предложена метрика ће бити модификована проширењем декомпозиције и на трећу димензију (време).

7 ЛІТЕРАТУРА

- [1] M. Tanimoto, M. P. Tehrani, T. Fujii, T. Yendo, “Free-Viewpoint TV”, IEEE Signal Processing Magazine, Vol.28, No.1, pp.67-76, January 2011.
- [2] K. Muller, P. Merkle and T. Wiegand, “3D video representation using depth maps,” Proc. IEEE, vol. 99, no. 4, pp. 643–656, April 2011.
- [3] Y Mori, N Fukushima, T Yendo, T Fujii, M Tanimoto, View generation with 3D warping using depth information for FTV. Signal Processing: Image Communication, 24(1-2), 65-72 (2009)
- [4] M Tanimoto, FTV Technologies and Standards, IEEE COMSOC MMTC E-letter, 10(2), 7-10, March (2015)
- [5] A. Vetro, T. Wiegand, and G. J. Sullivan, “Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard”, Proc. IEEE, Vol. 99, No. 4, pp. 626–642, Apr. 2011.
- [6] M. Hannuksela, D. Rusanovskyy, W. Su, L. Chen, R. Li, P. Aflaki, D. Lan, M. Joachimiak, H. Li, M. Gabbouj, “Multiview-video-plus-depth coding based on the advanced video coding standard,” IEEE Trans. Image Process., vol. 22, no. 9, pp. 3449–3458, Sep. 2013.
- [7] K. Muller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. H. Rhee, G. Tech, M. Winken, T. Wiegand, “3D high-efficiency video coding for multi-view video and depth data,” IEEE Trans. Image Process., vol. 22, no. 9, pp. 3366–3378, Sep. 2013.
- [8] M. Tanimoto, FTV (free-viewpoint television). APSIPA Transactions on Signal and Information Processing, 1, (2012) e4 doi:10.1017/ATSIP.2012.5
- [9] M. Levoy and P. Hanrahan, “Light field rendering,” in Proceedings of SIGGRAPH, ser. SIGGRAPH '96. New York, NY, USA: ACM, 1996, pp. 31–42. [Online]. Available: <http://doi.acm.org/10.1145/237170.237199>
- [10] H.-Y. Shum and L.-W. He, “Rendering with concentric mosaics,” in Proceedings SIGGRAPH, Los Angeles, California USA, 1999, pp. 299–306.

- [11] C. Buehler, M. Bosse, L. McMillan, and S. Gortler, "Unstructured Lumigraph Rendering," in Proc SIGGRAPH, Los Angeles, California USA, August 2001, pp. 425–432.
- [12] C Fehn, Depth image based rendering (DIBR), compression and transmission for a new approach on 3D-TV. Proc. SPIE, Stereoscopic Displays and Applications XV, 5291, 93-104, San Jose, CA, Jan 2004.
- [13] L. Do, S. Zinger, P. H. N. de With, "Objective quality analysis for free-viewpoint DIBR", International Conference on Image Processing (ICIP), Hong Kong, September 2010.
- [14] E Bosc, R Pepion, P Le Callet, M Koppel, P Ndjiki-Nya, M Pressigout, L Morin, Towards a New Quality Metric for 3-D Synthesized View Assessment. IEEE Journal on Selected Topics in Signal Processing. 5(7), 1332-1343 (2011).
- [15] X Liu, Y Zhang, S Hu, S Kwong, C C J Kuo, Q Peng, Subjective and objective video quality assessment of 3D synthesized views with texture/depth compression distortion, IEEE Trans. on Image Processing, vol. 24 no. 12, pp.4847-4861, Dec. 2015
- [16] A. Toet, "A morphological pyramidal image decomposition", Pattern Recognition Letters, vol.9, pp. 255-261, 1989.
- [17] P. Maragos, R. Schafer, "Morphological systems for multidimensional signal processing", Proceedings of the IEEE, April 1990.
- [18] D. Sandić-Stanković, D. Kukulj, P. Le Callet, "DIBR synthesized image quality assessment based on morphological pyramids", 3DTV-CON Immersive and interactive 3D media experience over networks, Lisbon, July 2015
- [19] J Goutsias, H Heijmans, Nonlinear Multiresolution Signal Decomposition Schemes—Part I: Morphological Pyramids. IEEE Trans. on Image Processing, 9(11), 1862–1876 (2000)
- [20] D. Sandić-Stanković, D. Kukulj, P. Le Callet, "DIBR synthesized image quality assessment based on morphological wavelets", International Workshop on Quality of Multimedia Experience QoMEX, Costa Navarino, Greece, May 2015
- [21] H. Heijmans, J. Goutsias, Multiresolution signal decomposition schemes-Part II: Morphological wavelets, IEEE Trans. on Image Processing, 9(11), 1897–1913 (2000)

- [22] D Sandić-Stanković, D Kukolj, P Le Callet, "Multi-scale synthesized view assessment based on morphological pyramid", Journal of Electrical Engineering, Vol. 67 (1), 2016, pp. 1–9, http://iris.elf.stuba.sk/JEEEC/data/pdf/1_116-01.pdf.
- [23] D Sandić-Stanković, D Kukolj, F. Battisti, P Le Callet, M. Carli, "Free Viewpoint Video Quality Assessment based on Morphological Multiscale Metrics", International Workshop on Quality of Multimedia Experience QoMEX, Lisbon, Jun 2016
- [24] IRCCyN/IVCDIBR images database: ftp://ftp.ivc.polytech.univ-nantes.fr/IRCCyN_IVC_DIBR_Images
- [25] Emilie Bosc, Romuald Pépion, Patrick Le Callet, Martin Köppel, Patrick Ndjiki-Nya, Luce Morin, Muriel Pressigout, "Perceived quality of DIBR-based synthesized views", SPIE Optics + Photonics, San Diego, 2011.
- [26] E Bosc, P Le Callet, L Morin, M Pressigout, An edge-based structural distortion indicator for the quality assessment of 3D synthesized views, Picture Coding Symposium, 249-252, May 2012
- [27] P-H Conze, P Robert, L Morin, Objective View Synthesis Quality Assessment. Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII, Febr 2012
- [28] F Battisti, E Bosc, M Carli, P Le Callet, S Perugia, Objective image quality assessment of 3D synthesized views. Elsevier Signal Processing: Image Communication. 30(1), 78-88 (2015)
- [29] N.El-Yamany, K.Ugur, M.Hannuksela, M.Gabbouj, "Evaluation of depth compression and view synthesis distortions in multiview-video-plus-depth coding systems", 2DTV-CON 2010
- [30] P. Hanhart, E. Bosc, P. Le Callet and T. Ebrahimi, "Free-Viewpoint video sequences: a new challenge for objective quality metrics", International Workshop on Multimedia Signal Processing (MMSP), Jakarta, Indonesia, 2014
- [31] Free-Viewpoint synthesized videos quality database: http://ivc.univ-nantes.fr/en/databases/Free-Viewpoint_synthesized_videos/
- [32] E.Bosc, F.Battisti, M.Carli and P.L.Callet, "A wavelet-based image quality metric for the assessment of 3D synthesized views", Proc. SPIE 8648, Stereoscopic Displays and Applications, (March 2013)

- [33] C T Tsai, H M Hang, Quality assessment of 3D synthesized views with depth map distortion, Visual communications and image processing (VCIP), 2013
- [34] H.Shao, X.Cao, G.Er, “Objective quality assessment of depth image based rendering in 3DTV system”, 3DTV-CON, 2009
- [35] A.B. Dehkordi, M. Pourzad and P. Nasiopoulos, “A study on the relationship between depth map quality and the overall 3D video quality of experience”, 3DTV-Conf, 2013
- [36] D. Silva, W.Fernando, S. Worrall and A. Kondo, “A novel depth map quality metric and its usage in depth map coding”, 3DTV/Conf., 2011
- [37] E.Ekmekcioglu, S.T.Worall, D.De Silva, W.A.C.Fernando and A.M.Kondo, “Depth based perceptual quality assessment for synthesized camera viewpoints”, Int. Conf. on User Centric Media, September 2010
- [38] M. Solh, G. AlRegib and J.M.Bauza, “3VQM: A vision-based quality measure for DIBR-based 3D videos, IEEE Int. Conference on Multimedia and Expo (ICME), July 2011, 1-6.
- [39] C. Hewage and M. Martini, “Edge-based reduced-reference quality metric for 3-D video compression and transmission,” IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 5, pp. 471–482, 2012.
- [40] E Adelson, C Anderson, J Bergen, P Burt, J Ogden, Pyramid methods in image processing. RCA Engineer, Nov/Dec 1984
- [41] S. Mallat, “Wavelets for a vision”, Proceedings of the IEEE, vol.84, no.4, April 1996
- [42] Z Wang, E Simoncelli, AC Bovik, Multi-scale structural similarity for image quality assessment. Asilomar Conference on Signals, Systems and Computers, Nov 2003
- [43] PJ Burt, EH Adelson, The Laplacian pyramid as a compact image code. IEEE Trans. on Communications, 31(4), 532-540 (1983).
- [44] D Sandić-Stanković, D Kukolj, P Le Callet, “Mera za ocenu kvaliteta slike zasnovana na piramidalnoj dekompoziciji i srednjoj kvadratnoj greški ”, TELFOR, Belgrade, November 2015
- [45] Z Wang, Q Li, Information Content Weighting for Perceptual Image Quality Assessment. IEEE Trans. on Image Processing, 20(5) 1185-1198 (2011).

- [46] J. Qian, D. Wu, L. Li, D Cheng, X Wang, Image quality assessment based on multi-scale representation of structure, *Digital Signal Processing* 33, 125-133, June 2014
- [47] Y.K.Lai, C.C.Jay, Kuo, Image quality measurement using the Haar wavelet , *Proc. SPIE* 3169, *Wavelet Applications in Signal and Image Processing V*, 127, October, 1997
- [48] S. Rezazadeh, S. Coulombe, A novel wavelet domain error-based image quality metric with enhanced perceptual performance, *Int. Journal of Computer and Electrical Engineering*, vol.4, no.3, June 2012.
- [49] E Adelson, E Simoncelli, W Freeman, Pyramids and multiscale representations. *Proc. European Conf. on Visual Perception*, Paris, Aug 1990
- [50] J. Ogden, E.Adelson, J.Bergan and P.Burt, "Pyramid-based computer graphics", *RCA Engineer*, Sept/Oct. 1985
- [51] W. Sweldens, "The lifting scheme: A construction of second generation wavelets", *SIAM J. Math. Anal.*, 27(2), 1997
- [52] I. Daubechies, W.Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier Analysis and Applications*, vol.4, num. 3, pp. 247-269, 1998.
- [53] W. Sweldens, P. Schroder, "Wavelets in Computer Graphics", *ACM SIGGRAPH Course Notes*, 1996
- [54] F.Meyer, P.Maragos, "Nonlinear scale-space representation with morphological levelings", *Journal of Visual Communication and Image Representation* 11, pp. 245-265, 2000
- [55] G. Matheron, "Random Sets and Integral Geometry," New York: Wiley, 1975
- [56] J. Serra, "Introduction to Mathematical Morphology," *Journal on Computer Vision, Graphics and Image Processing*, vol. 35, issue 3, pp. 283-305, Sept. 1986
- [57] S. Sternberg, «Gray scale morphology», *Comput. Graphics Image Process.* 35, 335-355, 1986
- [58] J. Roerdink, "Multiresolution maximum intensity volume rendering by morphological adjunction pyramids", *IEEE Trans. on Image Processing*, Jun 2003

- [59] H. Nobuhara, D. Trieu, T. Maruyama, B. Bede, "Max-plus algebra-based wavelet transforms and their FPGA implementation for image coding", *Journal of Information Sciences*, vol.180, issue 17, pp. 3232-3247, Sept. 2010
- [60] D. Bhowmik and G.C.K.Abhayaratne, "Morphological wavelet domain watermarking", *EUSIPCO*, September 2007
- [61] H. Heijmans, J. Goutsias, *Multiresolution signal decomposition schemes Part2: Morphological wavelets*. Tech. Rep. PNA-R9905, CWI, Amsterdam, The Netherlands, July 1999.
- [62] H. Heijmans, J. Goutsias, "Constructing Morphological Wavelets with the Lifting Scheme," *Int. Conf. on Pattern Recognition and Information Processing*, Belarus, pp. 65-72, 1999.
- [63] R. Andrews, D. Nguyen, "Separable and quincunx wavelet image coding", 1998
- [64] J. Kovačević, M. Vetterli, "Nonseparable Two- and Three-Dimensional Wavelets", *IEEE Trans. on Signal Processing*, vol. 43, No.5, May 1995.
- [65] H. Heijmans, J. Goutsias, "Morphological pyramids and wavelets based on the quincunx lattice", *Mathematical morphology and its applications to image and signal processing*, vol.18, pp. 273-281, 2000.
- [66] G. Uytterhoeven, A. Bultheel, "The Red-Black Wavelet Transform", *Proc. of IEEE Benelux Signal Processing Symposium*, 1997.
- [67] Z Wang, A Bovik, Mean squared error: love it or leave it. *IEEE Signal Processing Magazine*, 26(1), 98-117 (2009)
- [68] Z. Wang, A. Bovik and L. Lu, "Why is image quality assessment so difficult", *ASSP 2002*
- [69] VQEG HDTV Group, *Test Plan for Evaluation of Video Quality Models for Use with High Definition TV Content*. 2009.
- [70] A Telea, An image inpainting technique based on the fast matching method. *Journal of graphics, GPU and game tools*, 9(1), 23-34 (2004)
- [71] Y Mori, N Fukushima, T Yendo, T Fujii MTanimoto, View generation with 3D warping using depth information for FTV. *Signal Processing: Image Communication*, 24(1-2), 65-72 (2009)

- [72] K Muller, A Smolic, K Dix, P Merkle, P Kauff, T Wiegand, View synthesis for advanced 3D video systems. EURASIP Journal on image and video processing, 2008, 2008:438148
- [73] P Ndjiki-Nya, P Koppel, M Doshkov, H Lakshman, P Merkle, K Muller, T Wiegand, Depth image based rendering with advanced texture synthesis. IEEE Int. Conf. on Multimedia&Expo, 424-429, Suntec City, July 2010.
- [74] M Koppel, P Ndjiki-Nya, M Doshkov, H Lakshman, P Merkle, K Muller, T Wiegand, Temporally consistent handling of disocclusions with texture synthesis for depth-image-based rendering. IEEE Int. Conf. on Image Processing, 1809-1812, Hong Kong, Sept 2010.
- [75] M Solh and G AlRegib, Depth adaptive hierarchical hole filling for DIBR-based 3D videos, Proceedings of SPIE, 8290, 829004, Burlingame, CA, US, 2012
- [76] Methodology for the Subjective Assessment of the Quality of Television Pictures, document ITU-R BT.500, Nov. 1993.
- [77] M. Aubury, W. Luk, "Binomial filters", Journal of VLSI Signal Processing, 1-8, 1995
- [78] K Gu, M Liu, G Zhai, X Yang, W Zhang, Quality assessment considering viewing distance and image resolution, IEEE Trans. On Broadcasting, 61(3), 520-531 (2015)
- [79] D Chandler, S Hemami, VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images. IEEE Trans. on image processing, 16(9), 2284-2298 (2007)
- [80] Z Wang, AC Bovik, A universal image quality index. IEEE Signal Processing Letters, 9(3), 81-84 (2002)
- [81] Z Wang, AC Bovik, HR Sheikh E Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. on Image Processing, 13(4), 600-612 (2004)
- [82] R Song, H Ko, C C J Kuo, MCL-3D: A database for stereoscopic image quality assessment using 2D-image-plus-depth source, <http://arxiv.org/abs/1405.1403> (2014)
- [83] MCL-3D stereoscopic image quality database. <http://mcl.usc.edu/mcl-3d-database>
- [84] M. Tanimoto, T.Fujii, and K.Suzuki, "View synthesis algorithm in view synthesis reference software 3.5 (VSRS3.5)," (2009).

- [85] ITU, “Rec ITU-R BT.1438, Subjective assessment of stereoscopic television pictures,” tech. rep. (2000).
- [86] E. Bosc, P. Hanhart, P. Le Callet, T. Ebrahimi, “A quality assessment protocol for free-viewpoint video sequences synthesized from decompressed depth data”, QoMEX, 2013.
- [87] Zhang, G. Tech, K. Wegner, and S. Yea, “3D-HEVC test model 5,” ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 JCT3V-E1005, July 2013.
- [88] 3D-HTM [Online]. <http://hevc.hhi.fraunhofer.de/>
- [89] JM [Online]. <http://iphome.hhi.de/suehring/tml/>
- [90] HM [Online]. <http://hevc.hhi.fraunhofer.de/>
- [91] Kakadu [Online]. <http://www.kakadusoftware.com/>.
- [92] J. Gautier, O. Le Meur, and C. Guillemot, “Efficient depth map compression based on lossless edge coding and diffusion,” in Picture Coding Symposium (PCS), 2012, pp. 81–84.
- [93] F. Pasteau, C. Strauss, M. Babel, O. Deforges, and L. Bedat, “Adaptive colour decorrelation for predictive image codecs,” Proc. of EUSIPCO 2011, pp. 1–5, 2011.
- [94] E. Bosc, Compression of Multi-View-plus-Depth (MVD) data: from perceived quality analysis to MVD coding tools designing, Ph.D. thesis, INSA de Rennes, Oct. 2012.
- [95] ITU-T P.910, “Subjective video quality assessment methods for multimedia applications,” International Telecommunication Union, April 2008.
- [96] ITU-T BT.500, “Methodology for the subjective assessment of the quality of television pictures,” International Telecommunication Union, Jan. 2012.
- [97] UQI Matlab code: http://www.cns.nyu.edu/~zswang/files/research/quality_index/img_qi.m
- [98] SSIM matlab code: https://ece.uwaterloo.ca/~z70wang/research/ssim/ssim_index.m
- [99] MS-SSIM, IW-SSIM, IW-PSNR Matlab code:
<https://ece.uwaterloo.ca/~z70wang/research/iwssim/>
- [100] 3DSwIM matlab p-code. <http://www.comlab.uniroma3.it/3DSwIM.html>
- [101] MP-PSNR Matlab p-code:

<https://sites.google.com/site/draganasandicstankovic/code/mp-psnr>

[102] MW-PSNR Matlab p-code:

<https://sites.google.com/site/draganasandicstankovic/code/mw-psnr>

8 ПРИЛОГ

Прилог садржи списак слика, списак табела и списак скраћеница.

8.1 Списак слика

Слика 1.1. Категоризација телевизије	15
Слика 1.2. МВЦ систем	17
Слика 1.3. ЗДВ систем	17
Слика 1.4. ФТВ систем	18
Слика 1.5. Типична изобличења у сликама због процеса ДИБР синтезе	21
Слика. 3.1.1 Редундантна мултирезолуциона репрезентација: (а) шема декомпозиције са једним оператором за анализу и једним оператором за синтезу; (б) шема реконструкције са једним оператором синтезе	37
Слика 3.2.1 Нередундантна мултирезолуциона репрезентација: (а) шема декомпозиције са два оператора за анализу (б) шема реконструкције са једним оператором синтезе (в) шема реконструкције са два оператора синтезе	38
Слика 3.2.2 Трансформација сигнала таласићима применом лифтинг шеме садржи три корака: подела сигнала (split), процена (predict, P) и поправка сигнала (update, U). (а) декомпозиција (б) инверзна трансформација, синтеза	41
Слика 4.1. Израчунавање мултирезолуционе метрике кроз три фазе	43
Слика 4.1.1. Шема морфолошке пирамидалне декомпозиције: Е ерозија, D дилатација ..	46
Слика 4.1.2. Морфолошка пирамида МПД ЕД генерисана на основу синтетизованог фрејма видео секвенце Newspaper. За морфолошке операције коришћен је СЕ димензија 7x7	47
Слика 4.1.3. Сепарабилна декомпозиција слике таласићима применом две једнодимензионе декомпозиције таласићима у вертикалном и хоризонталном правцу	49

Слика 4.1.4. Декомпозиција сигнала s_j на сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} морфолошким и линеарним Харовим таласићима применом лифтинг шеме	51
Слика 4.1.5. Подопсеци првог нивоа декомпозиције фрејма из видео секвенце Newspaper морфолошким таласићем minLift садрже ивице по правцима: вертикалне, хоризонталне, угаоне	52
Слика 4.1.6. Декомпозиција сигнала s_j на сигнал детаља d_{j+1} и сигнал апроксимације ниже резолуције s_{j+1} применом морфолошког таласића minLift или линеарног таласићаcdf(2,2) коришћењем лифтинг шеме	54
Слика 4.1.7. Несепарабилна декомпозиција слике таласићима коришћењем “quincunx“ решетке применом непарног и парног корака	55
Слика 4.1.8. Непарни корак несепарабилне декомпозиције слике применом лифтинг шеме а) слика која се декомпонује садржи узорке на Cartesian решетки најпре се дели на два сигнала, оба са пикселима на quincunx решетки б) процена црног пиксела се израчунава на основу 4 бела пиксела који га окружују с) бели пиксел се ажурира на на основу 4 пиксела детаља који га окружују	55
Слика 4.1.9. Парни корак несепарабилне декомпозиције на quincunx решетки применом лифтинг шеме а) сигнал апроксимације са пикселима на quincunx решетки се дели на два сигнала, оба на Cartesian решетки б) процена сивог пиксела на основу 4 бела пиксела који га окружују по дијагоналним правцима с) бели пиксел се ажурира на на основу 4 пиксела детаља који га окружују по дијагоналним правцима	56
Слика 4.1.10. Подопсеци прва два нивоа несепарабилне декомпозиције фрејма из синтетизоване видео секвенце Newspaper са узорковањем према quincunx решетки, непарног и парног, применом морфолошког таласића minLiftQ. Слика детаља непарног корака декомпозиције је ротирана за 45°	58
Слика 4.2.1. Израчунавање MP-PSNR на основу MSE_j свих слика пирамиде: MPD- један ниво морфолошке пирамидалне декомпозиције	60

Слика 4.2.2. Израчунавање MP-PSNR _r на основу MSE_j слика детаља виших нивоа пирамиде: MPD-један ниво морфолошке пирамидалне декомпозиције	61
Слика 4.2.3. Израчунавање MW-PSNR на основу MSE_{ji} свих подопсега i свих нивоа декомпозиције j добијених сепарабилном декомпозицијом 2-D MWD синтетизоване слике таласићима	62
Слика 5.2.1. Фрејмови МВД видео секвенци базе IRCCyN/IVCDIBR: Book Arrival, Lovebird1 и Newspaper	69
Слика 5.2.2. Сlike синтетизоване на основу фрејма видео секвенце Newspaper применом алгоритама A2-A7 с лева на десно, одозго на доле	72
Слика 5.2.3. Структурни елементи различитих облика: квадратни, кружни, ромб, крстасти (с лева на десно)	74
Слика 5.2.4. Пирсонови коефицијенти корелације (PCC) метрике која користи пирамидалну декомпозицију слика са морфолошким филтрима MP-PSNR и са линеарним филтрима LP-PSNR према DMOS за различите величине структурног елемента за морфолошке филтре и различите величине прозора за линеарне филтре, 2x2 до 13x13 ...	75
Слика 5.2.5. База IRCCyN/IVCDIBR: DMOS _p (MP-PSNR) према DMOS	76
Слика 5.2.6. Пирсонов коефицијент корелације PSNR према DMOS за слике пирамиде по нивоима декомпозиције. Структурни елементи различитих димензија су коришћени за морфолошко филтрирање	78
Слика 5.2.7. Пирсонов коефицијент корелације PSNR слике детаља морфолошке пирамиде према DMOS и Пирсонов коефицијент корелације MP-PSNR према DMOS, за структурне елементе различитих димензија	79
Слика 5.2.8. Пирсонов коефицијент корелације редуковане верзије MP-PSNR _r и основне верзије MP-PSNR према DMOS за структурне елементе различитих димензија 2x2–13x13	80

Слика 5.2.9. Пирсонов коефицијент корелације метрике MW-PSNR према DMOS за различите типове морфолошких и линеарних таласића	82
Слика 5.2.10. Пирсонови коефицијенти корелације (PCC) метрике PSNR према DMOS по подопсезима добијеним декомпозицијом синтетизованих слика морфолошким таласићима minNaar, minLift и minLiftQ	83
Слика 5.2.11. Пирсонов коефицијент корелације PSNR за најбољи подопсег према DMOS и MW-PSNR према DMOS за различите типове таласића	85
Слика 5.2.12. Пирсонов коефицијент корелације редуковане MW-PSNR _r и основне верзије MW-PSNR према DMOS за различите типове морфолошких и линеарних таласића	86
Слика 5.2.13. Пирсонов коефицијент корелације стандардних метрика (SSIM, MS-SSIM, PSNR), метрике намењене сликама са изобличењем због ДИБР синтезе (3DswIM) и редукованих верзија предложене метрике MP-PSNR _r и MW-PSNR _r	88
Слика 5.3.1. Оригиналне слике базе MCL-3D: Kendo, Baloons и Lovebird1 Poznan street, Poznan Hall2, Shark, Microworld, Gt_fly, Undo_dancer	89
Слика 5.3.2. Генерисање стерео слика базе MCL-3D. Стерео пар слика (СЛ, СД) је синтетизован на основу слике Т2 и њој одговарајуће мапе дубине Д2. Референтни стерео пар слика (РЛ, РД) је синтетизован на основу парова оригинално снимљених слика (Т1, Т2) и (Т2, Т3) и њима одговарајућих парова мапа дубине (Д1, Д2) и (Д2, Д3)	90
Слика 5.3.3. Shark: оригинална слика Т1, референтна лева слика РЛ, леве слике стерео парова синтетизоване применом ДИБР алгоритама А7, А1, А2, А8	90
Слика 5.3.4. Baloons: оригинална слика Т1, референтна лева слика РЛ, леве слике стерео парова синтетизоване применом ДИБР алгоритама А7, А1, А2, А8	91
Слика 5.3.5. Сlike синтетизоване ДИБР алгоритмима СЛ(А1), СЛ(А2), СЛ(А7), СЛ(А8) су на различитим позицијама у односу на оригиналне слике Т1, лева колона; референтне слике РЛ генерисане помоћу референтног софтвера ВСРС су на истој позицији као слике синтетизоване ДИБР алгоритмима А1, А2, А7, А8 (десна колона)	93
Слика 5.3.6. Расподела MP-PSNR према MOS за стерео слике базе MCL-3D	95

Слика 5.3.7. Пирсонов коефицијент корелације мере PSNR по сликама пирамиде према MOS. Структурни елементи различитих димензија су коришћени за морфолошко филтрирање	96
Слика 5.3.8. Расподела MW-PSNR према MOS за стерео слике базе MCL-3D	98
Слика 5.3.9. Пирсонови коефицијенти корелације (PCC) метрике PSNR према DMOS по подопсезима добијеним декомпозицијом синтетизованих слика морфолошким таласићима minNaar, minLift и minLiftQ	99
Слика 5.3.10. Пирсонови коефицијенти корелације стандардних метрика (SSIM, MS-SSIM, PSNR), метрике намењене сликама са изобличењем због ДИБР синтезе (3DswIM) и предложених метрика MP-PSNR и MW-PSNR за базу стерео слика MCL-3D	101
Слика 5.4.1. Видео секвенца FVSV базе синтетизована у моду мешање (blend) садржи 100 фрејмова: r1, r2 снимљени фрејмови, s1-s49 ДИБР-синтетизовани фрејмови	102
Слика 6.4.2. Расподела резултата метрике MW-PSNR при оцењивању дела базе FVSV који садржи видео секвенце синтетизоване у моду мешање (blend) на основу мапа дубине кодованих кодерима C1-C7	106
Слика 6.4.3. Квалитет видео секвенце Kendo синтетизоване на основу мапа дубине кодованих кодером C2 са 3 квантизациона нивоа: P25, P35, P47	107
Слика 6.4.4. Расподела вредности MW-PSNR по MVD саржају дела базе FVSV који садржи видео секвенце синтетизоване у моду мешање	108

8.2 Списак табела

Табела 4.3.1. Број операција за израчунавање морфолошке пирамидеерозија/дилатацијаса L нивоаприменом структурног елемента величине $1 \times P$ пиксела по врстама и $P \times 1$ пиксела по колонама. Коришћене су операцијеодузимање (S) и поређење (C)	64
Табела 4.3.2. Број операција за израчунавање MSE слика пирамиде за основну верзију MP-PSNR и за редуковану верзију MP-PSNRr. Морфолошка пирамидална декомпозиција са L нивоа применом SE димензија $P \times P$. MSE се израчунава коришћењем операција одузимање (S), множење (M) и сабирање (A)	65
Табела 4.3.3. Број операција за израчунавање декомпозиције слике таласићима (Содузимање, C поређење, A сабирање, M множење)	66
Табела 5.2.1. База IRCCyN/IVCDIBR садржи фрејмове три MVD секвенце	69
Табела 5.2.2. База слика IRCCyN/IVCDIBR: 3 MVD садржаја, позиције оригиналне и синтетизоване слике	70
Табела 5.2.3. Перформансе основне верзије MP-PSNR и редуковане верзије MP-PSNRr	74
Табела 5.2.4. Перформансе PSNR метрике слике пирамиде креиране применом структурних елемената различитих димензија	79
Табела 5.2.5 Перформансе метрике MW-PSNR и редуковане верзије MW-PSNRr	82
Табела 5.2.6. Перформансе PSNR за подопсег детаља добијен декомпозицијом слике таласићима	84
Табела 5.2.7. Перформансе стандардних метрика квалитета слика и предложених метрика	87
Табела 5.3.1. Перформансе основне верзије MP-PSNR и редуковане верзије MP-PSNRr при оцењивању базе MCL-3D	95
Табела 5.3.2. Перформансе PSNR метрике слике пирамиде креиране применом структурних елемената различитих димензија за оцењивање базе MCL-3D	97

Табела 5.3.2. Перформансе метрике MW-PSNR и редуковане верзије MW-PSNR _r при оцењивању базе MCL-3D	98
Табела 5.3.3. Перформансе метрике PSNR за подопсег детаља добијен декомпозицијом слике различитим типовима таласића при оцењивању базе MCL-3D	100
Табела 5.3.4. Перформансе стандардних метрика квалитета слика и предложених метрика при оцењивању базе MCL-3D	101
Табела 5.4.1. MVD секвенце FVSV базе	103
Табела 5.4.2. Перформансе стандардних метрика квалитета слика и предложених метрика у оцењивању дела базе FVSV синтетизоване применом мода мешање (blended mode) ...	105
Табела 5.4.3. Перформансе метрика MW-PSNR и PSNR израчунатих за 7 делова FVSV базе који садрже видео секвенце синтетизоване применом мода мешање на основу мапа дубине кодovаних са 7 кодера C1- C7	106
Табела 5.4.4. Перформансе метрике MW-PSNR за 5 делова базе FVSV по MVD садржају	108
Табела 5.4.5. Време израчунавања метрике за фрејм видео секвенце Kendo	108

8.3 Списак скраћеница

ФТВ	Free-viewpoint TV, телевизија са избором позиције посматрања
ЗДТВ	телевизија која омогућује кориснику осећај дубине сцене
ЗДВ	ЗД видео
MVD	(Multiview Video plus Depth), скуп видео секвенци и њима одговарајућих мапа дубине
MVV	(Multiview Video), скуп видео секвенци
SVD	(Single View Video), једна видео секвенца
MPEG	Moving Picture Experts Group
ДИБР	(Depth-Image-Based-Rendering), техника генерисања слике на новој позицији на основу слике и мапе дубине на суседној позицији
MVC	(Multi-view Video Coding), стандард за кодовање скупа видео секвенци
MSE	(Mean Squared Error), средња квадратна грешка
PSNR	(Peak Signal-to-Noise Ratio), вршна вредност сигнал/шум
MP-MSE	Средња квадратна грешка пирамидалне репрезентације слике
MP-PSNR	Morphological Pyramid Peak Signal-to-Noise Ratio
MP-MSE _r	Средња квадратна грешка редуковане пирамиде
MP-PSNR _r	редукована верзија Morphological Pyramid Peak Signal-to-Noise Ratio
MW-MSE	Средња квадратна грешка репрезентације слике добијене декомпозицијом таласићима
MW-PSNR	Morphological Wavelet Peak Signal-to-Noise Ratio
MW-MSE _r	Средња квадратна грешка редуковане репрезентације слике генерисане декомпозицијом таласићима
MW-PSNR _r	редукована верзија Morphological Wavelet Peak Signal-to-Noise Ratio
СЕ	структурни елемент морфолошке операције
МПД ЕД	Морфолошка Пирамидална Декомпозиција Ерозија Дилатација

minHaar	морфолошки Харов минимум таласић
minLift	морфолошки таласић којим се чува локална минимална вредност сигнала кроз сигнале апроксимације на свим нивоима декомпозиције
minLiftQ	морфолошким таласић за несепарабилну декомпозицију на “quincunx” решетки применом лифтинг шеме
Haar	линеарни Харов таласић
cdf(2,2)	линеарни биортогонални таласић назван према истраживачима који су га представили, Cohen-Daubechies-Feauveau
cdf(2,2)Q	линеарни таласић за несепарабилну декомпозицију са “quincunx” узорковањем
SSIM	Structural Similarity index
MS-SSIM	Multi-Scale Structural Similarity
3DswIM	3D Synthesized view Image Quality Metric
IRCCyN/IVCDIBR	база слика синтетизованих применом ДИБР алгоритама
MCL-3D	база стерео слика синтетизованих применом ДИБР алгоритама
FVSV	база видео секвенци синтетизованих ДИБР техником са навигацијом кроз различите позиције
DMOS	Differential Mean Opinion Score
MOS	Mean Opinion Score
RMSE (Root Mean Squared Error)	корен средње квадратне грешке
PCC (Pearson’s Correlation Coefficient)	Пирсонов коефицијент корелације
SCC (Spearman’s correlation coefficient)	Спирманов коефицијент корелације ранга