

UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET

Goran J. Zajić

**DETEKCIJA PROMENA KADROVA U
VIDEO SEKVENCI PRIMENOM
MULTIFRAKTALNE ANALIZE**

doktorska disertacija

Beograd, 2014.

UNIVERSITY OF BELGRADE
SCHOOL OF ELECTRICAL ENGINEERING

Goran J. Zajić

**SHOT BOUNDARY DETECTION
IN VIDEO SEQUENCE USING
MULTIFRACTAL ANALYSIS**

Doctoral Dissertation

Belgrade, 2014.

Mentor:

Dr Irini Reljin, red. prof, Univerzitet u Beogradu, Elektrotehnički fakultet

Članovi komisije:

Dr Miodrag Popović, red. prof, Univerzitet u Beogradu, Elektrotehnički fakultet

Dr Andreja Samčović, vanr. prof, Univerzitet u Beogradu, Saobraćajni fakultet

Datum odbrane:

Прилог 1.

Изјава о ауторству

Потписани GORAN ZAJIC¹

број индекса 5054/07

Изјављујем

да је докторска дисертација под насловом

DETEKCIJA PROMENE KADROVA U VIDEO SEKVENCI
PRIMEKOM MULTIFRAKTALNE ANALIZE

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 12.06.2014.

Goran Zajic

Прилог 2.

**Изјава о истоветности штампане и електронске
верзије докторског рада**

Име и презиме аутора GORAN ZAJIC

Број индекса 5054/07

Студијски програм ELEKTROTEHNIKA I RAČUNARSTVO, TELEKOMUNIKACIJE I
INFORMACIONE TEHNOLOGIJE

Наслов рада DETEKCIJA PROMENE KADROVA U VIDEO SERVENCI
PRIMEKOM MULTIFRAKTALNE ANALIZE

Ментор PROF. DR. IRINI REČIĆ, ELEKTROTEHNIČKI FAKULTET, UNIVERZITET
U BEOGRADU

Потписани GORAN ZAJIC

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао за објављивање на портал Дигиталног репозиторијума Универзитета у Београду.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбрање рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, 12.06.2014

Goran Zajic

Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

DETEKCIJA PROMENE KADROVA U VIDEO SEKVENCI
PRIMENOM MULTIFRAKTALNE ANALIZE

која је моје ауторско дело.

Дисертацију са свим прилозима предао сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, 12.06.2014.

Goran Zajic

DETEKCIJA PROMENA KADROVA U VIDEO SEKVENCI PRIMENOM MULTIFRAKTALNE ANALIZE

REZIME:

Moderan i efikasan menadžment video sadržaja podrazumeva mogućnost pretraživanja videa na osnovu automatski generisanih oznaka i informacija dobijenih analizom sadržaja frejmova. Vremenskom segmentacijom videa generiše se hijerarhijska struktura video sadržaja, koja predstavlja osnovu za automatsku anotaciju videa. Prvi korak vremenske segmentacije predstavlja određivanje granica kadrova kao najniže hijerarhijske logičke strukture videa. Kadar predstavlja jedno kontinuirano snimanje kamerom između uključivanja i isključivanja kamere. Grupe kadrova kreiraju više hijerarhijske logičke strukture videa kao što su: scene, sekvence i programi. Pri detekciji granica kadrova razlikuju se dva osnovna tipa tranzicije između kadrova - nagle i gradijentne promene kadrova. Naglim promenama kadrova nazivaju se tranzicije između kadrova koje traju samo dva frejma, dok kod gradijentnih tranzicija prelazak između dva kadra traje više frejmova.

Tema ove disertacije je realizacija algoritma za efikasnu detekciju naglih promena kadrova u video sekvencama različitog sadržaja, dinamike i nivoa postprodukcije. U tu svrhu, sadržaj frejmova video sekvenci opisan je obeležjima za boju i teksturu. Korišćeni su histogram boje i vejllet obeležje kao elementarna obeležja izdvojena iz frejmova, na lokalnom i globalnom nivou. Odabrana obeležja obezbeđuju sa jedne strane dovoljnu osetljivost na nagle promene vizuelnog sadržaja na frejmovima. Sa druge strane, odabrana obeležja pokazuju određeni nivo tolerancije na promene sadržaja niskog intenziteta, kao što su: kretanje objekata, kamere i promene osvetljenja. Na obeležja frejmova primenjene su dve različite diferencijalne metrike, korelacija i multifraktalna analiza. Korelacija, kao diferencijalna metrika, kombinovanim pristupom je omogućila formiranje specifičnih artefakata koje je jednostavno detektovati u strukturi signala. Artefakti zbog dinamike promena unutar kadrova mogu biti degradirani i pomereni u odnosu na poziciju stvarne

nagle promene kadra. Za lokalizaciju nagle promene sadržaja u vremenu primenjena je multifraktalna analiza, tako što je iskorišćena sposobnost ove tehnike da dobro opiše lokalnu strukturu signala. U domenu grubih Hölder-ovih eksponenata multifraktalne analize, nagla promena signala se prikazuje kao lokalni maksimum. Kombinacijom ove dve metrike (ponovi metrike) detektuju se pozicije kandidata naglih promena kadrova. Klasifikacija kandidata izvršena je primenom klasifikatora baziranih na fiksnim pragovima i primenom dobro definisanih pravila. Kao dodatna mera u cilju smanjenja osetljivosti sistema na promene osvetljenja, u algoritam je implementiran i *flash* detektor. Realizovani algoritam za detekciju naglih promena kadrova testiran je na velikom broju frejmova (>1.500.000 frejmova), video sekvenci različitih sadržaja, žanrova i tehnologija proizvodnje. Algoritam je pokazao visoke performanse, odnosno tačnost i preciznost, među kojima se posebno ističe tačnost. Pokazalo se da preciznost algoritma zavisi od klase video sekvenci nad kojom je obavljeno testiranje. U disertaciji je prikazano poređenje realizovanog algoritma sa algoritmima za detekciju naglih promena kadrova visokih performansi iz literature. Poređenja su vršena na osnovu tačnosti, preciznosti i F-mere. Rezultati poređenja su pokazali da je predloženi algoritam porediv po performansama sa ostalim algoritmima, a za neke video sekvence pokazao je bolje rezultate od ostalih. Prikazani rezultati ukazali su na velike mogućnosti predloženog algoritma u pogledu daljeg razvoja.

KLJUČNE REČI: artefakti, detekcija, *flash* detektor, nagle promene kadrova, multifraktalna analiza, video.

NAUČNA OBLAST: Elektrotehnika i računarstvo

UŽA NAUČNA OBLAST: Telekomunikacije (multimedija) i Obrada signala

UDK: 621.3

SHOT BOUNDARY DETECTION IN A VIDEO SEQUENCE USING MULTIFRACTAL ANALYSIS

ABSTRACT:

Modern and efficient management of video content implies the possibility of video searching on the basis of automatically generated labels and information gained by the analysis of frame content. Temporal video segmentation generates hierarchical structure of video. This structure represents the basis for the automated annotation of video content. The first step of temporal segmentation is the determination of shot boundaries, as the lowest hierarchical logical structure of video. A shot refers to a continual recording with video camera between switch-on and switch-off of camera. Groups of shots create higher hierarchical logical structures of video, such as: scenes, sequences and programs. In the shot boundary detection there are two basic types of transitions between shots – abrupt and gradual shot changes. Abrupt shot changes refer to those transitions between shots which last only two frames, whereas in gradual shot changes the transitions between two shots last several frames.

The topic of this thesis is the experimental realization of algorithm for efficient abrupt shot changes detection in video sequences of different contents, dynamics and postproduction levels. For this purpose, the frame content of video sequences is described with color and texture features. Color histogram and wavelet feature are used as basic features extracted from frames, on both local and global levels. On the one hand, the selected features provide sufficient sensibility regarding sudden changes of visual content on frames. On the other hand, the selected features show certain level of tolerance for the low intensity content changes, such as: moving of objects, camera and light changes. Correlation and multifractal analysis have been used on frame features as differential metrics. With combined approach, correlation as a differential metrics applied on frames features, has enabled the formation of specific artefacts which can be easily detected in

signal structure. Due to the change dynamics within shots, artefacts can be degraded and shifted regarding the position of real abrupt shot changes.

Multifractal analysis has been used for temporal localization of abrupt content changes due to its ability to provide a detailed description of signal local structure. In the domain of rough Hölder multifractal analysis exponents, abrupt signal change is shown as a local maximum. By combining these two metrics, correlation and multifractal analysis, the positions of abrupt shot change candidates are detected. Candidate classification has been carried out by means of classifiers based on fixed thresholds and predefined rules. As an additional precaution aimed at reducing system sensibility related to light changes, *flash* detector has been implemented into the algorithm. The implemented algorithm for the abrupt shot change detection has been tested on a large number of frames (>1.500.000 frames), video sequences of different contents, genres and production technologies.

The algorithm has shown high performances, i.e. recall and precision, especially recall. It has been shown that precision of algorithm depends on the class of video sequence being tested. In this thesis, the comparisons between the proposed algorithm and the high-performance algorithms for the detection of abrupt shot changes in literature have been presented. Comparisons have been carried out on the basis of recall, precision and F-measure. The findings have shown that proposed algorithm is comparable by performances with other algorithms, showing even better results for some video sequences. The achieved results indicate great possibilities of the proposed algorithm for further development.

KEY WORDS: abrupt shot changes, artefacts, detection, *flash* detector, multifractal analysis, video.

SCIENTIFIC FIELD: Electrical Engineering and Computer Science

SPECIFIC SCIENTIFIC FIELD: Telecommunications (multimedia) and Signal Processing

UDK: 621.3

SADRŽAJ

	Strana
Spisak skraćenica	
Spisak slika	
Spisak tabela	
1. Uvod	1
1.1. Motivacija	2
1.2. Ciljevi istraživanja i polazne hipoteze	3
1.3. Organizacija disertacije	4
2. Segmentacija videa u vremenskom domenu i detekcija naglih promena kadrova	6
2.1. Segmentacija videa u vremenskom domenu	6
2.2. Detekcija promena kadrova	9
2.3. Detekcija naglih promena kadrova analizom sadržaja frejmova	11
3. Uvod u <i>MPEG</i> video kompresiju, <i>MPEG-7</i> alate za opisivanje sadržaja videa i multifraktalnu analizu	13
3.1. Uvod u <i>MPEG</i> standarde za kompresiju video sekvenci	14

3.1.1. <i>MPEG-1</i> standard	14
3.1.1.1. Kolor prostor, prostorna i vremenska redundansa video sekvence	14
3.1.1.2. Hijerarhijska struktura podataka u <i>MPEG-1</i> standardu	16
3.1.1.3. Intra-frejmska kompresija	19
3.1.1.4. Inter-frejmska kompresija	21
3.1.1.5. <i>MPEG-1</i> video format	24
3.1.2. <i>MPEG-2</i> standard	25
3.1.2.1. Tipovi slika	25
3.1.2.2. Hijerarhijska struktura podataka	27
3.1.2.3. Kompenzacija pokreta	27
3.1.2.4. Izračunavanje <i>DCT</i> i kvantovanje koeficijenata	28
3.1.2.5. Skalabilnost	29
3.1.2.6. Profili i nivoi u <i>MPEG-2</i> standardu	30
3.1.3. <i>MPEG-4</i> standard	32
3.1.3.1. Formati slike i bitski protoci u <i>MPEG-4</i> standardu	34
3.1.3.2. Hijerarhijska struktura podataka u <i>MPEG-4</i> standardu	34
3.1.3.3. Kodovanje pokreta u <i>MPEG-4</i> standardu	35
3.1.3.4. Kodovanje oblika i teksture u <i>MPEG-4</i> standardu	37
3.1.4. <i>MPEG-4/AVC p.10 (H.264)</i> standard	38
3.1.4.1. Struktura podataka i bitskog niza <i>H.264</i> standarda	38
3.1.4.2. Intrafrejmsko kodovanje u <i>H.264</i> standardu	39
3.1.4.3. Intrafrejmsko kodovanje u <i>H.264</i> standardu	41
3.1.4.4. Transformacija i kvantizacija u <i>H.264</i> standardu	43
3.1.4.5. Entropijsko kodovanje u <i>H.264</i> standardu	44
3.1.4.6. Profili u <i>H.264</i> standardu	45
3.1.5. <i>H.265/HEVC</i> standard	46
3.1.5.1. Struktura podataka u <i>H.265/HEVC</i> standardu	47
3.1.5.2. Intrafrejmska predikcija	48
3.1.5.3. Interfrejmska predikcija	49

3.1.5.4. Transformacija i kvantizacija	50
3.1.5.5 Entropijsko kodovanje	51
3.1.5.6 <i>In-Loop</i> filtri	51
3.1.5.7. Specijalni modovi kodovanja	53
3.1.5.8. Paralelizacija	53
3.1.5.9. Profili, rangovi i nivoi	54
3.2. MPEG-7 obeležja za opisivanje video sadržaja	56
3.2.1. Osnovne strukture	57
3.2.1.1. Mrežasta raspodela	57
3.2.1.2. Vremenske serije	57
3.2.1.3. 2D-3D višestruki izgledi	58
3.2.1.4. Prostorne 2D koordinate	58
3.2.1.5. Vremenska interpolacija	58
3.2.2. Kolor obeležja	59
3.2.2.1. Kolor prostor obeležje	59
3.2.2.2. Obeležje kolor kvantizacije	59
3.2.2.3. Obeležje dominantnih boja	60
3.2.2.4. Skalabilno kolor obeležje	60
3.2.2.5. Obeležje raspodele boje	60
3.2.2.6. Strukturno kolor obeležje	61
3.2.2.7. Grupno kolor obeležje	61
3.2.3. Obeležja teksture	62
3.2.3.1. Obeležje homogene teksture	62
3.2.3.2. Obeležje histogram ivica	63
3.2.3.3. Obeležje za teksturno pretraživanje	64
3.2.4. Obeležja oblika	64
3.2.4.1. Obeležje oblika regiona	65
3.2.4.2. Obeležje oblika konture	65
3.2.4.3. Spektralno 3D obeležje oblika	66
3.2.4.4. 2D/3D obeležje oblika	66

3.2.5.	Obeležja pokreta	67
3.2.5.1.	Obeležje pokreta kamere	67
3.2.5.2.	Obeležje trajektorije pokreta	68
3.2.5.3.	Obeležje parametarskog kretanja	69
3.2.5.4.	Obeležje aktivnosti pokreta	70
3.2.6.	Obeležja lokalizacije	71
3.2.6.1.	Obeležje lokalizacije regiona	71
3.2.6.2.	Obeležje prostorno-vremenske lokalizacije	72
3.2.7.	Obeležja za identifikaciju lica	72
3.2.7.1.	Obeležje za prepoznavanje lica	72
3.2.7.2.	Obeležje za napredno prepoznavanje lica	72
3.3.	Multifraktalna analiza	73
3.3.1.	Multifraktalni formalizam	73
3.3.2.	Primena multifraktalne analize u detekciji lokalnih/globalnih karakteristika signala	79
4.	Pregled publikovanih algoritama za detekciju naglih promena kadrova	81
5.	Novi algoritam za detekciju naglih promena kadrova	89
5.1.	Predloženi algoritam	89
5.2.	MPEG dekodovanje	91
5.3.	Odabir i izdvajanje obeležja za potrebe detekcije	92
5.3.1.	Odabir obeležja	92
5.3.2.	Izdvajanje obeležja	96
5.3.2.1.	Izdvajanje obeležja boje: kolor histogram	98
5.3.2.2.	Izdvajanje obeležja teksture: vevjlet obeležje	99
5.3.3.	Izračunavanje diferencijalne metrike	100
5.4.	Detekcija kandidata	103
5.4.1.	Detekcija artefakata	103
5.4.2.	Verifikacija kandidata	106
5.4.2.1.	Izračunavanje komplementarnog niza	106

5.4.2.2.	Primena multifraktalne analize	107
5.4.2.3.	Identifikacija pozicija artefakata sa pozicijama lokalnih maksimuma	109
5.5.	Gruba klasifikacija kandidata	111
5.5.1.	Poređenje globalnih histograma boje	111
5.5.2.	Provera okoline pozicije kandidata u alfa domenu multifraktalne analize	112
5.5.3	Provera oblika artefakta	114
5.6.	Fina klasifikacija kandidata	115
5.6.1.	Fina klasifikacija kandidata primenom empirijskih pravila	116
5.6.2.	Fina klasifikacija kandidata primenom fleš detektora	117
5.6.2.1.	Izračunavanje korelacionih matrica segmenata frejmova	117
5.6.2.2.	Izračunavanje klasifikacionog parametra	119
5.6.2.3.	Klasifikacija kandidata	119
6.	Eksperimentalni rezultati testiranja predloženog algoritma	121
6.1.	Video test sekvence	121
6.2.	Merenje performansi predloženog algoritma	127
6.2.1.	Metoda 1	127
6.2.2.	Metoda 2	128
6.3.	Rezultati simulacija	128
6.3.1.	Merenje performansi na prvoj test grupi	129
6.3.2.	Merenje performansi na drugoj test grupi	130
6.3.3.	Merenje performansi na trećoj test grupi	132
6.3.3	Usrednjeni rezultati merenja performansi po test grupama	132
6.4.	Poređenje sa postojećim algoritmima u literaturi	133
6.5.	Diskusija rezultata testiranja i poređenja	137
7.	Zaključak	144
	Literatura	151
	Biografija autora	161

SPISAK KORIŠĆENIH SKRAĆENICA

Skraćenica	Puni naziv
2D	<i>Two-dimensional</i>
3D	<i>Three-dimensional</i>
AC	<i>Alternating Current</i>
AVC	<i>Advanced Video Coding</i>
CABAC	<i>Context Based Adaptive Arithmetic Coding</i>
CAVLC	<i>Context Adaptive Variable Length Coding</i>
CB	<i>Coding Block</i>
CBIR	<i>Content Based Image Retrieval</i>
CSS	<i>Curvature scale-space</i>
CTB	<i>Coding Tree Block</i>
CTU	<i>Coding Tree Unit</i>
CU	<i>Coding Unit</i>
DC	<i>Direct Current</i>
DCT	<i>Discrete Cosine Transform</i>
DPCM	<i>Differential Pulse Code Modulation</i>
DST	<i>Discrete Sine Transformation</i>

DWT	<i>Discrete Wavelet Transform</i>
ECR	<i>Edge Change Ratio</i>
FOC	<i>Focus of Contraction</i>
FOE	<i>Focus of Expansion</i>
FPS	<i>Frame per Second</i>
GOF	<i>Group of Frames</i>
GOP	<i>Group of Pictures</i>
GVOP	<i>Group Video-Object Plane</i>
HEVC	<i>High Efficient Video Coding</i>
JPEG	<i>Joint Photographic Experts Group</i>
LSU	<i>Logical Story Unit</i>
MA	<i>Multifractal Analysis</i>
MaxDpb	<i>Maximal Decoded Picture Buffer</i>
MPEG	<i>Motion Photographic Experts Group</i>
NAL	<i>Network Abstraction Layer</i>
NTSC	<i>National Television System Committee</i>
PAL	<i>Phase Alternating Line</i>
PB	<i>Prediction Block</i>
PPS	<i>Picture Parameter Set</i>
PSNR	<i>Peak Signal to Noise Ratio</i>
SAO	<i>Sample Adaptive Offset</i>
SI	<i>Switching I</i>
SIF	<i>Source Input Format</i>
SNR	<i>Signal to Noise Ratio</i>
SP	<i>Switching P</i>
SPMM	<i>Sub-pixel Multi-fractal Measure</i>
SPS	<i>Sequence Parameter Set</i>
SVC	<i>Scalable Video Coding</i>
SVD	<i>Singular Value Decomposition</i>
SVM	<i>Support Vector Machine</i>

TB	<i>Transformation Block</i>
TU	<i>Transformation Unit</i>
URQ	<i>Uniform Reconstruction Quantization</i>
VLC	<i>Variable Length-Coding</i>
VO	<i>Video-Object</i>
VOP	<i>Video Object Plane</i>
VS	<i>Video-Object Sequence</i>

SPISAK SLIKA

Broj	Naziv	Str.
Slika 2.1.	Hijerarhija video segmenata u vremenskom domenu.	2
Slika 2.2.	Hijerarhijska segmentacija emitovanog video strima u vremenskom domenu.	3
Slika 2.3.	Osnovni tipovi promena kadrova: a) nagla promena kadra <i>hard cut</i> , b) gradijentna promena kadra <i>dissolve</i> , c) gradijentna promena kadra <i>fade in</i> , d) gradijentna promena kadra <i>wipe</i> .	4
Slika 3.1.	Različiti domeni video sekvence: (a) Kolor prostor jednog frejma, (b) prostorni i vremenski domen video sekvence.	15
Slika 3.2.	Hijerarhijska struktura podataka izlaznog signala u <i>MPEG-1</i> standardu.	17
Slika 3.3.	Grupa slika (<i>GOP</i>) prikazana u okviru <i>MPEG</i> frejm sekvence.	18
Slika 3.4.	Primer isečaka za jedan frejm video sekvence.	18
Slika 3.5.	Principijska šema intrafrejm kompresije.	19
Slika 3.6.	Koeficijenti <i>DCT</i> transformisanog bloka slike.	20
Slika 3.7.	Kvantizacija u interfrejm kompresiji: (a) Kvantizaciona matrica za intrafrejm kompresiju, (b) cik-cak putanja skeniranja.	21

Slika 3.8.	Principska šema interfrejm kompresije za P slike.	22
Slika 3.9.	Principska šema interfrejm kompresije za B slike.	23
Slika 3.10.	Tipovi skeniranja slika u $MPEG-2$ standardu: a) poluslike, b) kompozitne slike.	26
Slika 3.11.	Estimacija pokreta kod skeniranja sa proredom i razdvojenim poluslikama.	28
Slika 3.12.	Putanje za skeniranje DCT koeficijenata u $MPEG-2$ standardu (a) cik-cak skeniranje, (b) alternativno skeniranje.	29
Slika 3.13.	Manipulacija na osnovu sadržaja u $MPEG-4$ standardu.	33
Slika 3.14.	Hijerarhijska struktura niza podataka u $MPEG-4$ standardu.	35
Slika 3.15.	Kodovanje pokreta u $MPEG-4$ standardu.	36
Slika 3.16.	a) Prikaz intra predikcije za blok lumentne kolor komponente dimenzije 4×4 piksela, b) direkcioni modovi za intra predikciju.	40
Slika 3.17.	Kodovanje pokreta u $MPEG-4$ standardu (blokovi).	42
Slika 3.18.	Kodovanje pokreta u $MPEG-4$ standardu: a) raspodela na blokove, b) realan primer kodovanja.	42
Slika 3.19.	Podela CTB blokova na kodovane blokove (CB) i transformacione blokove (TB). Pune linije označavaju granice CB blokova, a isprekidane linije granice TB blokova.	48
Slika 3.20.	Modovi za podelu CB blokova na PB blokove, postoje određena ograničenja veličine. Za <i>intrapicture</i> predikciju CB blokova, podržani su samo $M \times M$ i $M/2 \times M/2$ modovi.	48
Slika 3.21.	Multifraktalni spektar singulariteta.	76
Slika 4.1.	Algoritmi za detekciju naglih promena kadrova, pregled literature i pozicija predloženog algoritma u odnosu na postojeće pristupe.	87
Slika 5.1.	Blok šema algoritma za detekciju naglih promena kadrova.	90
Slika 5.2.	Dekodovanje video sekvence u niz frejmova (slika).	91
Slika 5.3.	Rezultati klasifikacije realnih i animiranih frejmova na osnovu integralnog vektora obeležja: a) multifraktalni spektri singulariteta	94

	za realne i animirane frejmove različitih video sekvenci,	
	b) multifraktalni spektri singulariteta za realne i animirane	
	frejmove iste video sekvence, c) vrednosti Theta parametra za	
	prikazane multifraktalne spektre.	
Slika 5.4.	Izdvajanje obeležja niskog nivoa iz sekvence frejmova.	96
Slika 5.5.	Algoritam izdvajanja obeležja niskog nivoa za jedan frejm:	96
	a) prikaz podele frejma na regione, b) princip izdvajanja obeležja	
	iz regiona frejma, c) konkatencija izračunatih obeležja regiona u	
	vektore obeležja.	
Slika 5.6.	Dijagram algoritma kolor transformacije.	99
Slika 5.7.	Izdvajanje obeležja niskog nivoa (boja i tekstura) iz dekodovanih	99
	frejmova: a) primer piramidalne dekompozicije slike primenom	
	vevjlet-a, b) <i>Haar</i> vevjlet, c) način izračunavanja vrednosti	
	obelezja teksture, d) konkatencija izračunatih obeležja u niz	
	obelezja.	
Slika 5.8.	Prikaz izračunavanja različitih diferencijalnih metrika <i>HC</i> i <i>HW</i> .	101
Slika 5.9.	Diferencijalne metrike: a) metrika <i>CH</i> u intervalu koordinata	102
	[4150,4350], b) metrika <i>CW</i> u intervalu koordinata [4150,4350].	
Slika 5.10.	Detekcija artefakata: a) prikaz artefakta, b) algoritam detekcije	104
	artefakta.	
Slika 5.11.	Detekcija artefakata (dodatni primer).	105
Slika 5.12.	Niz <i>ICW[t]</i> u intervalu koordinata [4150,4350].	106
Slika 5.13.	Niz <i>MICW[t]</i> u intervalu koordinata [4150,4350].	108
Slika 5.14.	Niz <i>MW[t]</i> u intervalu koordinata [4150,4350].	108
Slika 5.15.	Primer analize okoline kandidata i dodeljivanje lokalnog	109
	maksimuma artefaktu.	
Slika 5.16.	Provera okoline kandidata: a) prikaz analiziranih okolina	110
	kandidata, b) algoritam za prihvatanje lokalnih maksimuma i	
	dodeljivanje artefaktima.	
Slika 5.17.	Prikaz okolina lokalnog maksimuma na poziciji kandidata koje se	113

	proveravaju u alfa nizu vejevlet korelacionog niza.	
Slika 5.18.	Provera oblika artefakta: a) slika jednog artefakta, b) uslovi prihvatanja artefakta.	114
Slika 5.19.	Rezultat detekcije naglih promena kadrova na sedmom segmentu videa “ <i>A Good Year</i> ”, <i>OdbPK</i> -odbijene pozicije kandidata, <i>PotPK</i> - potencijalne pozicije naglih promena kadrova, <i>KORMAX</i> -klasifikacioni prag korelacije.	115
Slika 5.20.	Izračunavanje korelacione matrice <i>MFD23</i> za segment (2,3) frejma kandidata.	118
Slika 5.21.	Rezultat detekcije naglih promena kadrova na sedmom segmentu videa “ <i>A Good Year</i> ”, <i>OdbPK</i> -odbijene pozicije kandidata, <i>DetNPK</i> -detektovane nagle promene kadrova, <i>FDprag</i> -klasifikacioni prag <i>flash</i> detektora.	120
Slika 6.1.	Primeri frejmova iz četiri klase video sekvenci koje su korišćene pri testiranju predloženog algoritma: a) <i>Film1</i> , b) <i>Film2</i> , c) <i>Film3</i> , d) <i>Film4</i> .	123
Slika 6.2.	Primeri frejmova iz treće grupe video test sekvenci, a) sportski program, b) jutarnji program.	126
Slika 6.3.	Uporedni prikaz performansi predloženog algoritma (<i>Zajic</i>) i <i>TRECVID2004</i> algoritama.	135
Slika 6.4.	Uporedni prikaz performansi predloženog algoritma (<i>Zajic</i>) i <i>TRECVID2005</i> algoritama.	135
Slika 6.5.	Uporedni prikaz performansi predloženog algoritma (<i>Zajic</i>) i <i>TRECVID2007</i> algoritama.	136
Slika 6.6.	Ispušteni kadrovi u procesu detekcije: a) usled promene osvetljenja celog frejma i ubacivanja veštačkih frejmova, b) usled specijalnog efekta unutar kadra.	138
Slika 6.7.	Ispušteni kadrovi u procesu detekcije: a) usled brzog kretanja kamere i objekata, b) i c) usled brzog kretanje kamere.	139
Slika 6.8.	Ispušteni kadrovi u procesu detekcije: a) i b) visoka korelisanosti	140

sadržaja oba kadra, c) mala promena ugla kamere između dva kadra.

- Slika 6.9. Prikaz pogrešne detekcije nagle promene kadra usled brzog pomeranja kamere: a) statična scena, b) dinamična scena. 140
- Slika 6.10. Prikaz pogrešne detekcije nagle promene kadra usled nagle promene sadržaja dela frejma: a) reklamni pano na košarkaškoj utakmici, b) reklamni pano na fudbalskoj utakmici. 141
- Slika 6.11. Prikaz pogrešne detekcije nagle promene kadra: a) i b) nagli pomeraj objekta, c) nagla promena veličine objekta. 141
- Slika 6.12. Prikaz pogrešne detekcije nagle promene kadra prilikom prolaska objekta u prvom planu: a) filmski program, b) sportski program. 142

SPISAK TABELA

Broj	Naziv	Str.
Tabela 3.1.	<i>Core MPEG-1</i> specifikacija kodaera.	25
Tabela 3.2.	Ograničenja parametara po nivoima u <i>MPEG-2</i> standardu za glavni profil.	31
Tabela 3.3.	Definisani nivoi po profilima u <i>MPEG-2</i> standardu	31
Tabela 3.4.	Ograničenja za najniži nivo 1 i najviši nivo 6.2 definisane u <i>HEVC</i> standardu.	56
Tabela 4.1.	Pregled najuspešnijih <i>TRECVID</i> (2004-2007) algoritama za detekciju naglih promena kadrova (sa znakom <i>x</i> su označeni osnovni principi na kojima se zasnivaju).	88
Tabela 6.1.	Prva grupa video test sekvenci (igrani filmovi).	123
Tabela 6.2.	Podgrupa akcionih video test sekvenci.	124
Tabela 6.3.	Podgrupa dramskih video test sekvenci.	124
Tabela 6.4.	Podgrupa kombinovanih video test sekvenci.	125
Tabela 6.5.	Podgrupa animiranih video test sekvenci.	125
Tabela 6.6.	Grupa televizijskih video test sekvenci.	126
Tabela 6.7.	Rezultati merenja performansi predloženog algoritma na prvoj	129

	test grupi video sekvenci.	
Tabela 6.8.	Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa akcione sekvence.	130
Tabela 6.9.	Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa dramske sekvence.	130
Tabela 6.10.	Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa kombinovane sekvence.	131
Tabela 6.11.	Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa animirane sekvence.	131
Tabela 6.12.	Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, usrednjeni po klasama.	131
Tabela 6.13.	Rezultati merenja performansi predloženog algoritma na trećoj test grupi video sekvenci.	132
Tabela 6.14.	Rezultati merenja performansi predloženog algoritma usrednjeni po test grupama.	133
Tabela 6.15.	Pregled algoritama za detekciju naglih promena kadrova za poređenje performansi sa predloženim algoritmom.	134
Tabela 6.16.	Poređenje F-mere izmerene za predloženi algoritam i TRECVID algoritme.	137

1. Uvod

*All truths are easy to understand once they are discovered;
the point is to discover them.*

Galileo Galilei (1564 - 1642)

Razmena informacija u savremenom društvu podrazumeva razmenu audio-vizuelnih informacija, kao što su tekst, audio, slike i video. Danas video ima sve značajniju ulogu u razmeni informacija. Nagli razvoj tehnologija za kreiranje video fajlova i razmenu putem interneta prouzrokovao je pojavu velikog broja video baza smeštenih na serverima širom sveta. Napredni algoritmi za kompresiju i visoki protoci računarskih mreža obezbedili su kompaktnije arhiviranje i efikasniju razmenu video fajlova. Sa enormnim porastom kapaciteta video baza pojavila se potreba za efikasnijim upravljanjem video informacijama smeštenim u istim, u smislu efikasnijeg pregledanja, pretraživanja i izdvajanja informacija. Da bi se obezbedilo efikasnije upravljanje video informacijama neophodno je izvršiti anotaciju video sadržaja.

Automatska segmentacija videa u vremenskom domenu predstavlja tehniku automatskog opisivanja video sadržaja, pronalaženjem vremenske strukture istog. Kao

rezultat vremenske segmentacije izdvaja se niz segmenata određenih karakteristika koji su hijerarhijski međusobno povezani: frejmovi, kadrovi, scene kao logičke celine i programski sadržaji. Kadrovi se smatraju segmentima drugog nivoa u hijerarhiji segmentacije a definišu se kao niz uzastopnih frejmova snimljenih tokom jednog neprekidnog rada kamere. Analiza sadržaja višeg nivoa, indeksiranje i klasifikacija videa zasnovana je na detekciji kadrova. Prelazi između kadrova mogu biti nagli i gradijentni. Nagli prelazi nastaju usled značajne promene sadržaja između dva uzastopna frejma i traju dva frejma. Gradijentni prelazi predstavljaju promenu kadra u okviru kratke sekvence frejmova i mogu se podeliti u nekoliko podgrupa. Jedna podgrupa nastaje kao posledica editovanja sadržaja videa kao što su *fading*, *dissolving* i *wiping*, a druga kao posledica pomeranja kamere (*pan*, *tilt* i *zoom*).

1.1 Motivacija

Standardni algoritmi za detekciju naglih promena kadrova zasnivaju se na detekciji naglih promena sadržaja između susednih frejmova ili grupe susednih frejmova. Video sekvence su prethodno dekodovane, a zatim su sadržaji frejmova opisani obeležjima niskog nivoa. U procesu detekcije naglih promena kadrova detektuju se pozicije frejmova u čijoj okolini se dešavaju naglašene promene sadržaja frejmova. Proces detekcije je veoma osetljiv na nagle promene osvetljenja u video sekvenci i nagla kretanja objekata i kamere u okviru kadra. Žanr video sekvence znatno utiče na efikasnost detekcije i na broj pogrešno detektovanih naglih promena kadrova, jer je kretanje kamere i objekata vrlo specifično za svaki od žanrova. Pored navedenih problema, kao dodatni problem pojavljuje se i različita tehnologija produkcije video sekvence. Primera radi, detekcija naglih promena kadrova za animiranu video sekvencu može dati lošije rezultate u odnosu na realni video zbog drugačije strukture teksture i boja na animiranim frejmovima. Pored pomenutog, sloboda kreatora animacije u realizaciji radnje u okviru kadrova, u vidu naglih promena perspektive scene, pozicija i veličine objekata u velikoj meri otežavaju tačnu detekciju pozicija frejmova u čijoj okolini se dešavaju nagle promene kadrova. Svi ovi problemi ukazuju na potrebu razvoja algoritma za detekciju naglih promena kadrova koji će biti u stanju da

efikasno izdvoji promene kadrova u video sekvenci u realnom vremenu bez obzira na žanr i tehnologiju produkcije.

1.2 Ciljevi istraživanja i polazne hipoteze

Cilj istraživanja je razvoj novog algoritma za detekciju naglih promena kadrova u video sekvenci, koji je u stanju da izvrši detekciju promena za različite tipove videa bez obzira na žanr ili tehnologiju produkcije. Namera je da se u algoritam implementiraju aktuelne tehnike obrade signala koje bi trebalo da obezbede efikasnu i kvalitetnu detekciju. Svrha istraživanja je razvoj kombinacije različitih tehnika detekcije sa relativno malom kompleksnošću izračunavanja u cilju postizanja dobrih performansi algoritma za njegovu primenu u realnom vremenu. Realizovani algoritam bi trebalo da obezbedi efikasnu segmentaciju video sekvenci u vremenskom domenu na osnovu naglih promena kadrova što će predstavljati osnovu za dalju analizu i indeksiranje.

U ovoj disertaciji biće izvršena analiza standardnih algoritama za detekciju naglih promena kadrova u video sekvenci. Analiziraće se obeležja niskog nivoa frejmova na kojima se vrši detekcija, diferencijalne metrike i same tehnike detekcije. Nagle promene kadrova u video sekvenci predstavljaju singularitete razlike sadržaja frejmova koje je neophodno uočiti i detektovati.

Radi realizacije navedenih ciljeva neophodno je postaviti sledeće hipoteze:

U cilju obezbeđivanja što izraženijih singulariteta neophodno je odabrati odgovarajuća obeležja niskog nivoa, koja će obezbediti da razlika obeležja susednih frejmova bude dovoljno naglašena.

Neophodno je odabrati odgovarajuću diferencijalnu metriku koja bi trebalo da obezbedi robusnost algoritma na promene osvetljaja frejmova, pomeranja objekata i kamere.

Primena multifraktalne analize bi trebalo da obezbedi dodatno strukturno opisivanje sekvence razlika sadržaja frejmova u domenu singulariteta, tako da singulariteti budu još izraženiji i uočljiviji.

Radi smanjenja kompleksnosti izračunavanja potrebno je primeniti tehniku detekcije koja će biti dovoljno jednostavna za realizaciju i omogućiti njeno brzo izvršavanje.

U cilju smanjenja pogrešno detektovanih singulariteta koji ne predstavljaju nagle promene kadrova neophodno je razviti algoritam koji će omogućiti što kvalitetniju klasifikaciju detektovanih singulariteta.

Algoritam za detekciju kadrova bi trebalo da bude univerzalan i primenjiv na video sekvence različitih žanrova i tehnologija produkcije. Očekuje se visoka efikasnost detekcije promene kadrova kako za realni video, tako i za animirane video sekvence, sa što manjim procentom kako pogrešno detektovanih naglih promena kadrova tako i nedetektovanih stvarnih promena. Očekuje se da performanse algoritma u pogledu brzine izvršavanja budu uporedive sa poznatim algoritmima. Jednostavna realizacija samog algoritma treba da omogući paralelizaciju procesa detekcije na odgovarajućem hardveru.

1.3 Organizacija disertacije

Disertacija je organizovana u sedam poglavlja.

U prvom poglavlju izložena je inicijalna motivacija za izradu disertacije i pregled definisanih ciljeva istraživanja. U skladu sa predstavljenim ciljevima postavljene su hipoteze kao smernice u toku istraživanja.

U drugom poglavlju dat je uvod u segmentaciju videa u vremenskom domenu i razrađena je problematika detekcije kadrova kao inicijalnog koraka segmentacije.

U trećem poglavlju dat je kratak uvod u *MPEG (Motion Photographic Experts Group)* kompresije radi uvida u strukturu video signala nad kojima su vršeni eksperimenti

sa jedne strane, a sa druge strane radi lakšeg razumevanja nekih algoritama za detekciju naglih promena kadrova koji su predstavljeni u literaturi. U drugom delu ovog poglavlja dat je uvod u *MPEG-7* standard, sa akcentom na opisivanje video materijala na osnovu sadržaja. Treći deo ovog poglavlja predstavlja pregled osnovnih teorijskih načela multifraktalne analize sa izdvojenim primerima primene ove tehnike u različim oblastima istraživanja.

U četvrtom poglavlju dat je pregled postojećih algoritama za detekciju naglih promena kadrova, na osnovu sadržaja, koji su predstavljeni u literaturi.

U petom poglavlju predstavljen je novi algoritam za detekciju kadrova u video sekvencama, sa detaljnim opisom funkcionalnih celina algoritma.

U šestom poglavlju dat je pregled rezultata testiranja predloženog algoritma za detekciju naglih promena kadrova za različite klase video sekvenci. Pored rezultata testiranja predstavljeni su i rezultati poređenja performansi predloženog algoritma sa postojećim algoritmima visokih performansi iz literature. Na kraju poglavlja data je diskusija rezultata testiranja i poređenja.

U sedmom poglavlju dat je pregled izvedenih zaključaka i navedeni su naučni doprinosi disertacije. U okviru ovog poglavlja predstavljene su smernice za buduća istraživanja u cilju daljeg razvoja algoritma za analizu, indeksiranje i pretraživanje videa.

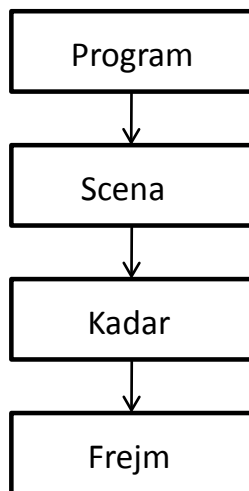
2. Segmentacija videa u vremenskom domenu i detekcija naglih promena kadrova

Ovo poglavlje predstavlja uvod u problematiku detekcije kadrova u video sekvenci i postavku problema koji je predmet disertacije. U prvom delu poglavlja dat je kratak uvod u segmentaciju videa u vremenskom domenu, sa posebnim osvrtom na detekciju kadrova. U drugom delu poglavlja razrađena je problematika detekcije naglih promena kadrova na osnovu izdvojenih obeležja sadržaja frejmova.

2.1. Segmentacija videa u vremenskom domenu

Video apstrakcija predstavlja kratku prezentaciju nekog originalnog videa koja se koristi za video katalogizaciju, indeksiranje i pretraživanje [SAG04]. Kreiranje video apstrakcije zasniva se na segmentaciji videa u vremenskom domenu, koja ima ulogu da detektuje hijerarhijsku strukturu video segmenata različitih nivoa i zastupljenosti detalja. Sa vizuelnog aspekta na najnižem nivou ove hijerarhijske strukture nalaze se pojedinačni frejmovi. Segmenti viših nivoa u redosledu ka najvišem su kadrovi, scene ili logičke celine,

sekvence i programi. Hijerarhijski redosled pomenutih video segmenata može se videti na Slici 2.1.



Slika 2.1. Hijerarhija video segmenata u vremenskom domenu

Više segmenata u okviru videa može biti objedinjeno u celinu na osnovu različitih karakteristika, zavisno od njihove pozicije u hijerarhijskoj strukturi. Takve karakteristike mogu biti vizuelne, strukturalne ili semantičke sličnosti. Jedan primer je kadar, video segment koji objedinjuje frejmove nastale tokom jednog snimanja kamerom. Dakle, u ovom slučaju frejmovi su objedinjeni u kadar po načinu proizvodnje.

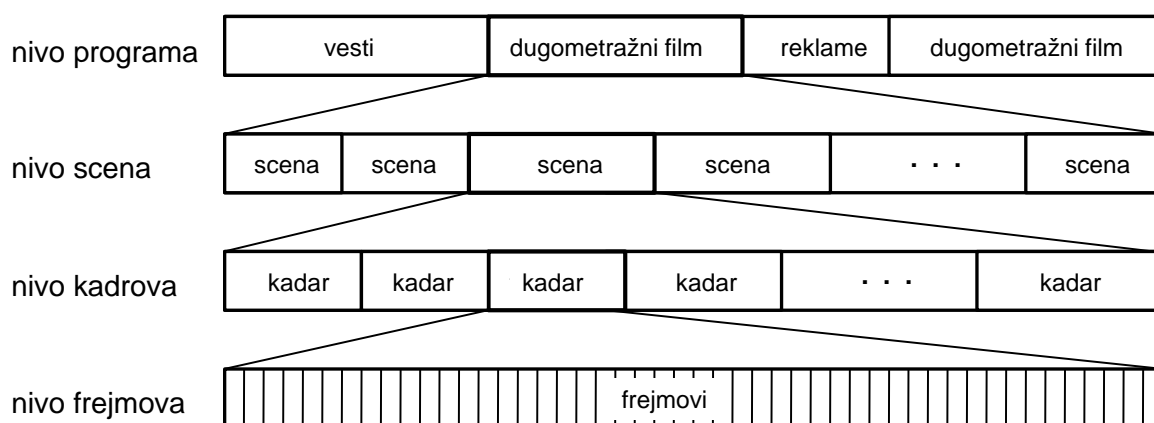
Video segment u vremenskom domenu se može definisati na sledeći način [PET10]: Video segment u vremenskom domenu predstavlja neprekidni deo video strima za koji je određen skup karakteristika konstantan. Dva susedna video segmenta u vremenskom domenu se razlikuju na osnovu definisanog skupa karakteristika. U skladu sa datom definicijom video segmenti se mogu definisati za različite nivoe zastupljenosti detalja i rezolucije na vremenskoj skali. S obzirom da video ima više modaliteta, segmente najnižeg nivoa za vizuelni modalitet predstavljaju frejmovi, za audio modalitet audio uzorci, a za tekstualni modalitet slova i reči. Naravno, audio i tekstualni modalitet se uzimaju u obzir samo ako su dostupni u okviru videa.

U hijerarhiji video segmenata u vremenskom domenu, kadrovi se smatraju video segmentima drugog nivoa [HAN02, LIE99] i definišu se kao niz uzastopnih frejmova

kreiranih jednim postupkom snimanja kamere. Kadrovi se smatraju primitivnim video segmentom za primenu u analizi videa visokog nivoa, indeksiranju i klasifikaciji [HAN02]. Scene kao video segmenti nastaju grupisanjem vremenski susednih kadrova. Pored ovog naziva, video segmenti ovog hijerarhijskog nivoa nazivaju se još i logičkim elementima događanja (*Logical Story Units - LSU*) [KEN98, HAN99, VEN02].

Scene se definišu kao događaji koji se odvijaju u jednoj postavci u toku kontinualnog vremenskog perioda [ASC99]. S obzirom da u filmskoj industriji postoje različita značenja vezana za izraz “scena” (kada je u pitanju film u fazi kreiranja i završen film), neophodno je naglasiti da se u ovoj disertaciji izrazom “scena” definiše video segment visokog nivoa u vremenskom domenu, u završenom filmu.

Sekvenca se kao video segment u hijerarhijskoj strukturi nalazi između scene i programa. Definiše se kao jedna ili više uzastopnih scena, a jedna ili više sekvenci čine programski segment. Na najvišem hijerarhijskom nivou video segmenata nalaze se programi. Programe mogu činiti igrani filmovi, vesti, dokumentarni filmovi, epizode televizijskih serija ili kućni video zapisi. U televizijskom emitovanju programi se nadovezuju jedan na drugi ili su međusobno isprekidani u različitim delovima, kao na primer slučaj kada je igrani film isprekidan reklamama. Na Slici 2.2 prikazan je primer hijerarhijske segmentacije emitovanog video strima u vremenskom domenu.

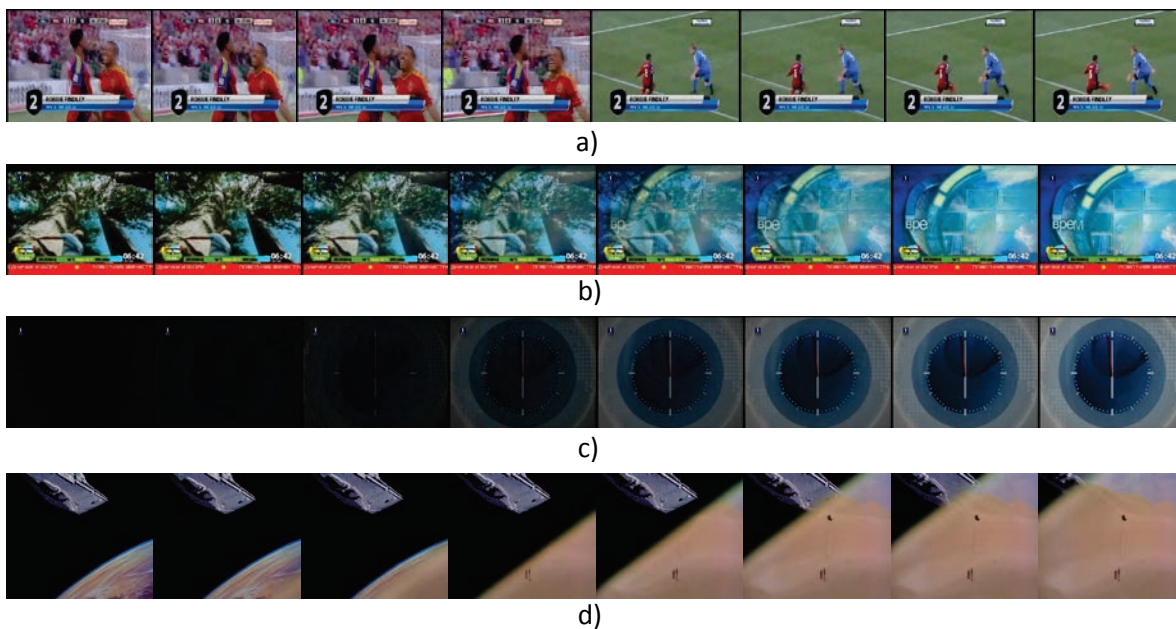


Slika 2.2. Hijerarhijska segmentacija emitovanog video strima u vremenskom domenu.

2.2 Detekcija promena kadrova

Segmentacija video sekvence na njene konstitutivne kadrove predstavlja proces koji je ekvivalentan pronalaženju granica kadrova u video sekvenci. Ovaj process pronalaženja granica kadrova uključuje detekciju promena unutar video sekvence. Kadar se može definisati kao sekvenca uzastopnih slika snimljenih kamerom bez prekidanja. Drugačije rečeno, kadar se sastoji od frejmova koji su kontinualno zabeleženi između pritiska na dugme “snimanje” i pritiska na dugme “pauza” kamere kojom se snima video.

Promene kadrova mogu biti nagle (*hard cut*) ili gradijentne (*gradual*). Pod gradijentnim promenama podrazumevamo tri tipa promena kadrova: *fade*, *dissolve* i *wipe*. Ako izuzmемо retke slučajeve gde se specijalni efekti koriste za tranziciju između kadrova, pomenute četiri promene kadrova možemo nazvati osnovnim. Na Slici 2.3 prikazana su sva četiri osnovna tipa promena kadrova.



Slika 2.3. Osnovni tipovi promena kadrova: a) nagla promena kadra *hard cut*, b) gradijentna promena kadra *dissolve*, c) gradijentna promena kadra *fade in*, d) gradijentna promena kadra *wipe*

Nagla promena kadrova (*hard cut*) nastaje kada posle poslednjeg frejma prethodnog kadra sledi prvi frejm narednog kadra. Ovo je kratka tranzicija između kadrova i traje samo dva frejma. Kod gradijentnih promena tranzicija između kadrova traje nekoliko frejmova. Kada je u pitanju gradijentna promena kadrova *fade* postoje tri tipa ove tranzicije. Tokom *fade-out* tranzicije slika gradijentno iščezava u monohromatsku, najčešće crnu sliku, u trajanju od nekoliko frejmova. Tokom *fade-in* tranzicije slika se postepeno pojavljuje iz monohromatske, najčešće crne slike. Treći tip ove gradijentne promene kadrova naziva se *fade-out/in*, kada je *fade-out* jednog kadra praćen *fade-in* promenom u narednom kadru. Tokom gradijentne promene kadrova *dissolve* slika jednog kadra gradijentno nestaje, dok se slika narednog kadra gradijentno pojavljuje. U toku gradijentne promene kadrova *wipe* kadrovi se smenjuju tako što sliku jednog kadra zamenjuje slika narednog kadra pomeranjem ivica frejma. Ivice generalno mogu imati proizvoljan geometrijski oblik. U najvećem broju slučajeva to su linije ili skupovi linija.

Nagle promene kadrova su najzastupljenije promene i koriste se za povezivanje različitih kadrova u video sekvencu. Gradijentne promene se koriste za uvođenje dodatnih efekata u video sekvencu. U ovoj disertaciji fokus je stavljen na detekciju naglih promena kadrova. Postoje dva različita pristupa u detekciji promena kadrova. Prvi pristup se zasniva na meta podacima koje poseduje svaki frejm. Naime, ako se radi sa „sirovim“ videom koji je snimljen digitalnom kamerom svaki frejm najčešće poseduje metapodatak o vremenu snimanja. Detekcija kadrova na osnovu vremenskih markera vrlo uspešno, bez greške, detektuje sve promene kadrova jer promene kadrova odgovaraju vremenskim prazninama između susednih frejmova. U sistemima koji se koriste za editovanje „sirovog“ videa vremenski markeri su najčešće dostupni, ali posle procesa editovanja oni se uglavnom uklanjaju i više nisu dostupni za proces detekcije promena kadrova. Zbog ovog razloga sistemi za detekciju kadrova se ne mogu osloniti samo na metapodatke, već moraju i analizirati sadržaj videa. Drugi pristup u detekciji kadrova se zasniva na analizi sadržaja videa. U prvom koraku se izdvajaju obeležja iz video sekvenci i taj proces se može izvršiti u prostornom domenu različitih dimenzija (piksel, region, cela slika). U sledećem koraku za detekciju promena kadrova i tranzicija mogu se primeniti različiti algoritmi koji koriste

vremenske prozore promenljivih dimenzija. Algoritmi za detekciju kadrova mogu se dodatno razlikovati po tome kako tretiraju pokret i koje tipove promena kadrova detektuju.

Sistemi za detekciju promena kadrova mogu samo detektovati poziciju promene kadra, u smislu da nalaze poziciju gde jedan kadar završava a naredni počinje. Dodatno, ovi sistemi mogu izvršiti klasifikaciju tipa tranzicije i dati informaciju o pozicijama početka i završetka tranzicije.

2.3. Detekcija naglih promena kadrova analizom sadržaja frejmova

Najčešći oblik tranzicije između dva kadra u video sekvenci predstavljaju nagle promene koje se pojavljuju između dva uzastopna frejma (k -tog i $k+1$ -og frejma), završnog frejma jednog kadra i prvog frejma narednog kadra. Nagla promena vizuelnog sadržaja između dva uzastopna frejma predstavlja diskontinuitet koji je vizuelno uočljiv, a proces detekcije ovakvih promena predstavlja proces prepoznavanja dovoljno izraženih diskontinuiteta u strimu vizuelnog sadržaja video sekvence.

Da bi diskontinuitet vizuelnog sadržaja bio merljiv neophodno je sadržaj frejmova opisati na odgovarajući način i zbog toga se u fazi pretprocesiranja vrši izdvajanje obeležja niskog nivoa. Diskontinuitet vizuelnog sadržaja reprezentuje se kao nagla promena vrednosti obeležja frejmova. Sledeći korak je uvođenje metrike koja treba da opiše razliku obeležja između dva susedna frejma. Vremenska sekvenca dobijena nakon ovog koraka u procesu detekcije predstavlja vremensku varijaciju amplitude razlike obeležja između dva susedna frejma, u oznaci $z(k, k+1)$ u kojoj se diskontinuiteti mogu jasno uočiti i detektovati. U procesu detekcije svaka vrednost iz sekvence amplitude razlika poredi se sa odgovarajućim pragom Tr . Ako je vrednost amplitude razlike obeležja veća od definisanog praga, nagla promena kadrova je detektovana.

Zavisno od korišćenih obeležja, na pozicijama diskontinuiteta vizuelnog sadržaja frejmova, amplitude razlike obeležja mogu se značajno razlikovati. Ove oscilacije otežavaju precizno definisanje praga Tr . Loše definisan prag može dovesti do lažne detekcije nagle promene kadra ili do situacije da stvarna nagla promena kadra ne bude

uopšte detektovana. Lažno detektovane nagle promene kadrova su posledica varijacije obeležja frejmova u okviru kadra. Ove varijacije nastaju usled naglih promena osvetljenja ili pomeranja objekata/kamere u video sekvenci, koje uzrokuju brze promene obeležja frejmova kreirajući tako dodatne naglašene diskontinuitete u okviru kadrova. Relativno nisko definisan prag uzrokuje detekciju određenog broja lažnih naglih promena kadrova. Sa druge strane, relativno visoko definisan prag Tr uzrokuje da diskontinuiteti manjih amplituda stvarnih naglih promena neće biti detektovani. Efikasnost detekcije naglih promena kadrova može da zavisi i od žanra video sekvenci koji se obrađuju s obzirom na stil snimanja i montiranja video sadržaja. Kod akcionih sekvenci pojavljuje se veliki broj brzih promena osvetljenja i pomeranja objekata/kamera koja prouzrokuju naglašene diskontinuitete u sekvenci razlika obeležja za razliku od stacionarnih dramskih sekvenci gde to nije slučaj.

Efikasnija i pouzdanija detekcija naglih promena kadrova može se postići odabirom odgovarajućih obeležja vizuelnog sadržaja frejmova i metrike razlike obeležja između frejmova na koje ne utiču u značajnoj meri promene osvetljenja i pomeranja objekata/kamere. U cilju unapređenja detekcije naglih promena kadrova mogu se primeniti tehnike detekcije pokreta i detekcije nagle promene osvetljenja čime se može značajno smanjiti broj pogrešno detektovanih naglih promena. Pored navedenog mogu se iskoristiti i neke dodatne informacije koje mogu ukazati na potencijalno prisustvo nagle promene. U nastavku teksta dat je kratak pregled tehnika za detekciju naglih promena kadrova.

3. Uvod u *MPEG* video kompresiju, *MPEG-7* video obeležja i multifraktalnu analizu

Nagle promene kadrova u video sekvenci mogu se detektovati korišćenjem potpuno dekodovanih frejmova proizvoljnog video formata i korišćenjem delimično dekodovanih frejmova u kompresionom domenu *MPEG* (*Motion Photographic Experts Group*) familije standarda za video kompresiju. Za potrebe detekcije, sadržaji potpuno dekodovanih frejmova moraju biti opisani obeležjima vizuelnih karakteristika frejmova. Kod delimično dekodovanih frejmova u kompresionom domenu, informacije o sadržaju se uzimaju iz kodovanih transformacionih koeficijenata svakog frejma u video strimu. U nastavku teksta ovog poglavlja dat je pregled *MPEG* video kompresija, radi lakšeg razumevanja principa na kojima se zasnivaju različiti algoritmi za detekciju naglih promena kadrova. Pre svega, radi razumevanja algoritama koji vrše detekciju kadrova u kompresionom domenu *MPEG-1* video kompresije. Pored ovog formata prezentovani su i ostali najbitniji video formati familije *MPEG* kompresija, kratkim pregledom karakteristika. U drugom delu ovog poglavlja dat je teorijski uvod u osnovne principe indeksiranja video sekvenci, sa pregledom vizuelnih obeležja definisanih u okviru *MPEG7* standarda. Na kraju poglavlja, u

trećem delu, predstavljena je multifraktalna analiza kratkim teorijskim uvodom. Predstavljeni su osnovni principi na kojima se zasniva ova tehnika nelinearne obrade signala, sa velikim brojem primera primene u različitim oblastima istraživanja.

3.1. Uvod u *MPEG* standarde za kompresiju video sekvenci

Pod *MPEG* familijom standarda za kompresiju video sekvenci podrazumeva se grupa od pet video formata: *MPEG-1* ([POP06],[LI03]), *MPEG-2* [ITU00], *MPEG-4 SVC p.2* [SCH07], *MPEG-4 AVC p.10(H.264)* [ITU10], *H.265 (HEVC)* [ITU13]. Ime standarda dolazi od naziva radne grupe formirane 1988. godine koja je pripremila pomenute standarde. Velika prednost ovih standarda je u tome što su to „generički“ [POP06] standardi, u standardima su definisani samo kompresioni bit strimovi koji implicitno definišu dekođer, dok su realizacije kompresionih algoritama i kodera potpuno prepuštene proizvođačima. U nastavku teksta biće reči o vizuelnim aspektima pomenutih standarda.

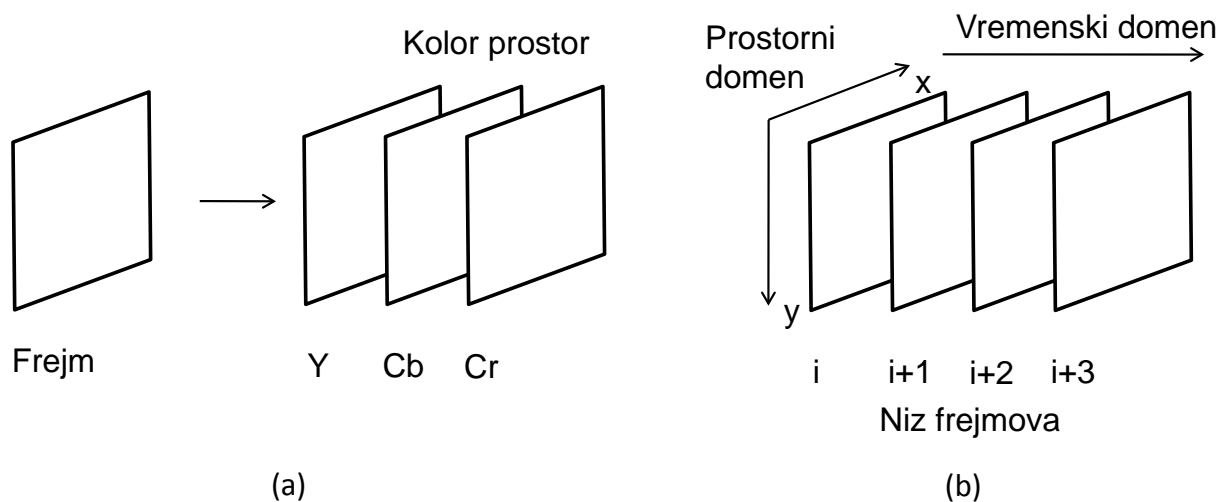
3.1.1. *MPEG-1* standard

MPEG-1 standard je prvi standard za video kompresiju iz *MPEG* familije, standard koji je uspostavljen i zvanično predstavljen javnosti 1991. godine. Uloga ovog standarda je bila da obezbedi veći kvalitet komprimovanog video sadržaja za veće stepene kompresije u odnosu na do tada već postojeće algoritme za kompresiju. Ovim standardom omogućena je kompresija video signala do *1,5 Mbit/s* uz zadovoljavajući subjektivno-perceptivni kvalitet komprimovanog sadržaja. U nastavku teksta prezentovani su osnovni principi *MPEG-1* standarda za video kompresiju.

3.1.1.1. Kolor prostor, prostorna i vremenska redundansa video sekvence

Vremenski niz slika koji predstavlja video sekvencu u nekomprimovanom obliku podrazumeva alokaciju velikog memorijskog prostora i zahteva širok propusni opseg prilikom prenosa. Kompresija videa predstavlja efikasnije kodovanje video sadržaja u

kompaktniji memorijski zapis. Kompresija video sekvence se može izvršiti u tri domena. U kolor domenu svake pojedinačne slike, zatim u prostornom domenu unutar svakog pojedinačnog frejma i u vremenskom domenu između susednih frejmova video sekvence. Na Slici 3.1 predstavljeni su kolor, prostorni i vremenski domen video sekvence.



Slika 3.1. Različiti domeni video sekvence: (a) kolor prostor jednog frejma, (b) prostorni i vremenski domen video sekvence.

Kolor domen jednog frejma video sekvence sastoji se od tri komponente $YCbCr$ kolor sistema [POP06]: lumentne komponente Y , hrominentne komponente Cb i hrominentne komponente Cr . Y -komponenta predstavlja reprezentaciju frejma u skali sivih tonova i sadrži informaciju o intenzitetu svakog piksela. Preostale dve komponente sadrže informaciju o boji. Svaka od komponenti u kolor prostoru pojedinačnog frejma nekomprimovane video sekvence ima istu dimeziju kao frejm. Unutar frejma čovek ima veću sposobnost razlikovanja sivih tonova nego boja [LI03]. Ova činjenica ukazuje da se može izvršiti redukovanje kolor prostora pododabiranjem hrominentnih komponenti. Primer formata odabiranja za $YCbCr$ kolor prostor je $4:2:0$ odnos, gde se za četiri bloka piksela iz Y komponente uzima se po jedan odgovarajući blok piksela iz svake od hrominentnih kolor komponenti.

Prostorni domen predstavlja dvodimenzionalni prostor svakog pojedinačnog frejma video sekvence. Kompresija signala u ovom domenu zasniva se na činjenici da se koristan

signal relativno sporo menja unutar frejma. Drugim rečima, vrednosti intenziteta ne variraju u velikom opsegu unutar malih površina (npr. u bloku slike dimenzija 8×8 piksela). S obzirom da se određena količina informacija ponavlja u okviru frejmova, te informacije predstavljaju prostornu redundansu sa aspekta kompresije. Psihofizičkim eksperimentima kod čoveka [LI03] pokazano je da gubitak komponente u visokofrekvencijskom domenu nije lako uočljiv, za razliku od gubitka komponente niskih frekvencija. Kompresija zasnovana na prostornoj redundansi može se realizovati redukcijom prostornog visokofrekvencijskog sadržaja.

S obzirom na mogućnost visoke korelacije sadržaja susednih frejmova može se konstatovati postojanje vremenske redundanse. Ona se može iskoristiti u cilju kompresije jer se ne mora svaki frejm nezavisno kodovati. Kompresija se u ovom domenu može izvršiti prediktivnim kodovanjem na osnovu susednih frejmova. Naime, koduje se razlika između tekućeg i referentnog frejma u sekvenci. Ovakve razlike su malih vrednosti što za posledicu ima malu entropiju i rezultuje efikasnijom kompresijom. Bolji pristup u odnosu na direktnu razliku između frejmova predstavlja određivanje pojedinih delova frejma koji će se oduzeti od susednog. Na ovaj način je iskorišćena visoka korelisanost pojedinih delova susednih frejmova u cilju smanjivanja količine informacija koje se koduju.

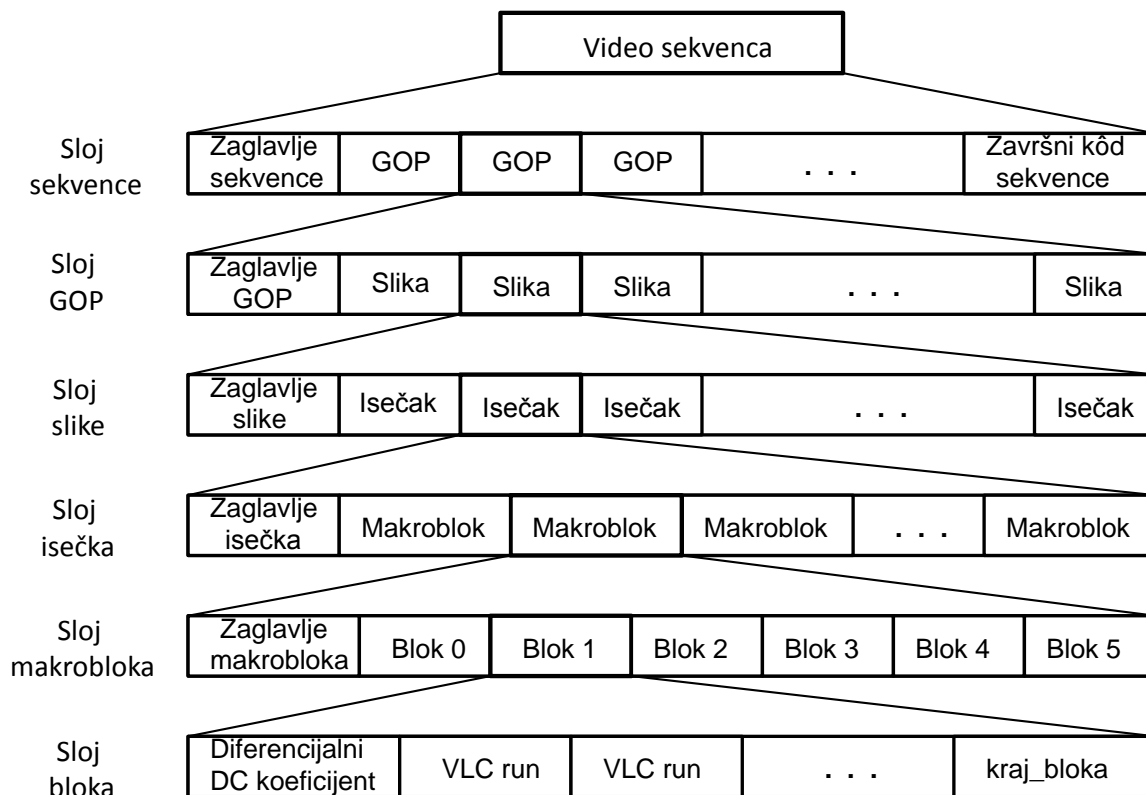
3.1.1.2. Hijerarhijska struktura podataka u *MPEG-1* standardu

Hijerarhijska struktura podataka izlaznog signala u *MPEG-1* standardu definisana je u šest slojeva, Slika 3.2 .

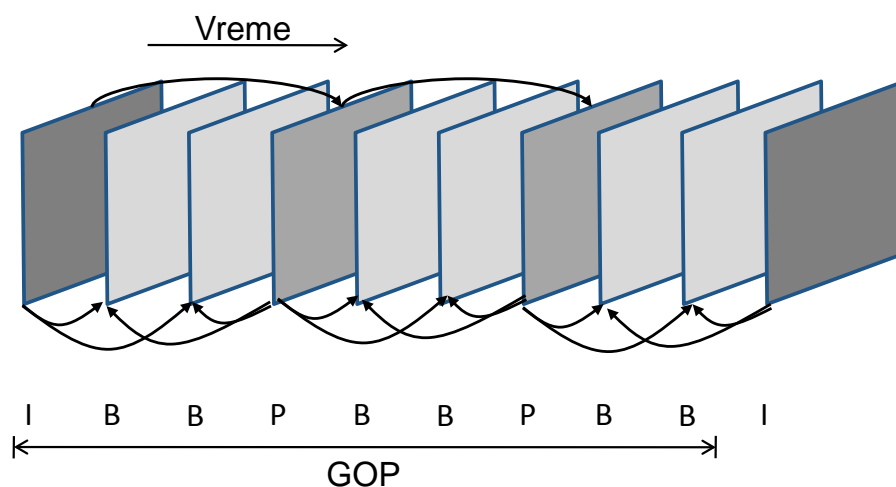
Prvi, najviši sloj ove hijerarhijske strukture čine podaci o sekvenci. U drugom sloju se nalaze podaci o grupama slika (*GOP*, *Group of Pictures*). Slučajan pristup podacima je moguć samo u prva dva sloja hijerarhijske strukture. Podaci o slikama se nalaze u trećem sloju. *MPEG-1* standardom definisana su četiri tipa slika: *I* slike, *P* slike, *B* slike i *D* slike. Pomenute slike se razlikuju po načinu kompresije. *I* slika predstavlja početnu sliku grupe slika (*GOP*), a kodovana je intrafrejm *DCT* (*Discrete Cosine Transform*) [RAO90] kompresijom. *P* slika je kodovana interfrejm *DCT* kompresijom sa predikcijom slike unapred korišćenjem već prenesenih *I* slika ili *P* slika. *B* slike se koduju interfrejm *DCT*

kompresijom sa predikcijama unapred, unazad ili bidirekcionom predikcijom korišćenjem već prenesenih I slika ili P slika. D tip slika je lošeg kvaliteta jer sadrži samo DC DCT koeficijente [POP06] transformisanih blokova. Ovaj tip slika se koristi za brzo pregledanje video sadržaja.

Grupa slika se sastoji od jedne I slike i više P i B slika, Slika 3.3. Broj slika u grupi zavisi od dozvoljenog bitskog protoka, zadatog vremena pristupa pojedinačnoj slici, dozvoljenog kašnjenja i drugih faktora. Standardom je definisano da jedna grupa slika može imati najviše do 132 slike da bi se sprečila nedozvoljeno velika akumulacija grešaka i veliko kašnjenje. U tipičnoj grupi slika I slike i P slike čine otprilike jednu trećinu ukupnog broja slika, dok se B slike interpoliraju na osnovu rekonstruisanih I i P slika.

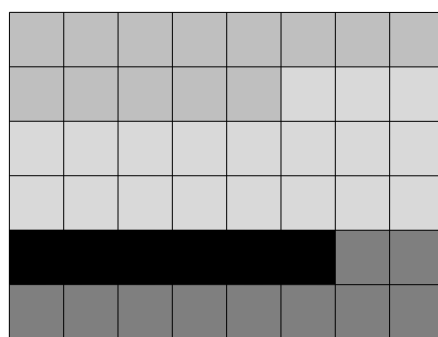


Slika 3.2. Hijerarhijska struktura podataka izlaznog signala u *MPEG-1* standardu



Slika 3.3. Grupa slika (*GOP*) prikazana u okviru *MPEG* frejm sekvence

Na Slici 3.3 prikazan je redosled prikazivanja slika u *MPEG* frejm sekvenci, međutim redosled kodovanja i prenosa je drugačiji zbog potrebe da se prvo prenese slika na osnovu koje se vrši predikcija. Redosled prenosa je *IPBBPBBIBB*, jer npr. ne može se preneti prvi *B* frejm dok nisu preneseni prvi *I* frejm i prvi *P* frejm.



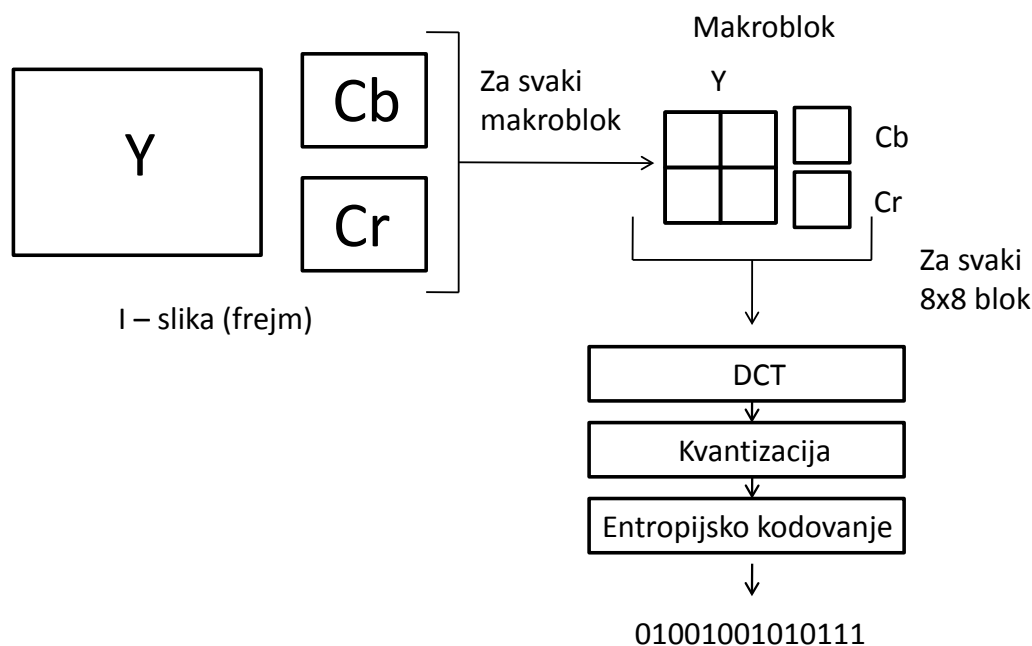
Slika 3.4. Primer isečaka za jedan frejm video sekvence

Četvrti sloj hijerarhijske strukture podataka *MPEG* izlaznog signala čine podaci o isečcima, slika 3.4. Isečak je segment frejma koji se može nezavisno dekodovati i sadrži određeni broj makroblokova. Svaki isečak može na proizvoljnom mestu početi i završiti, ako isecci prekrivaju celu sliku. Broj isečaka po frejmu je proizvoljan, ali ne sme biti prevelik jer utiče na smanjenje kompresije. Isečak se koristi u resinhronizaciji nakon pojave grešaka u prenosu. Peti sloj predstavljaju podaci o makroblokovima. Jedan makroblok

unutar frejma se sastoji od četiri bloka lumentne kolor komponente dimenzija 8×8 piksela i po jedan blok svake od hrominentnih komponenti istih dimenzija. Ovaj sloj je bitan jer se na ovom nivou vrši estimacija pomeraja, izbor načina transformacije i kvantovanja. Na najnižem, šestom sloju, nalaze se podaci o makroblokovima. U hijerarhijskoj strukturi podataka svaki od slojeva, izuzev najnižeg sloja blokova, sadrži zaglavlje u kojem su jedinstveno specificirani podaci koji slede u okviru sloja, Slika 3.2.

3.1.1.3. Intrafrejm kompresija

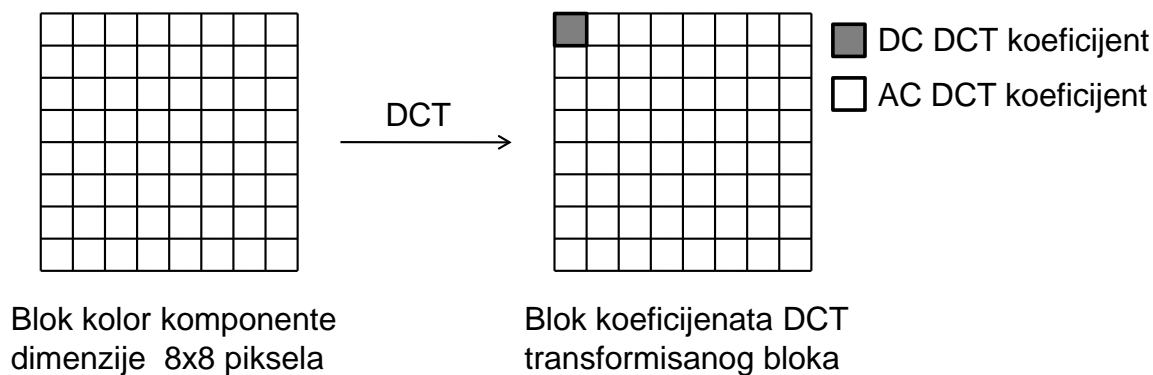
Intrafrejm kompresija se primenjuje na I slike video sekvence u cilju uklanjanja prostorne redundanse. Kompresija je realizovana primenom diskretne kosinusne transformacije na blokove slika dimenzija 8×8 piksela. Na Slici 3.5 prikazana je principska šema intrafrejm kompresije.



Slika 3.5. Principska šema intrafrejm kompresije

Primenom diskretne kosinusne transformacije na blok piksela kolor komponente slike dobija se blok DCT (*Discrete Cosine Transform*) [RAO90] koeficijenata, Slika 3.6. Vrednosti piksela bloka kolor komponente nalaze se u intervalu $[0,255]$. DC koeficijent

koji je označen na Slici 3.6 predstavlja nenegativnu vrednost u intervalu $[0,2040]$, a *AC* koeficijenti su vrednosti u intervalu $[-1024,1023]$. Srednja vrednost *AC* koeficijenata je nula.

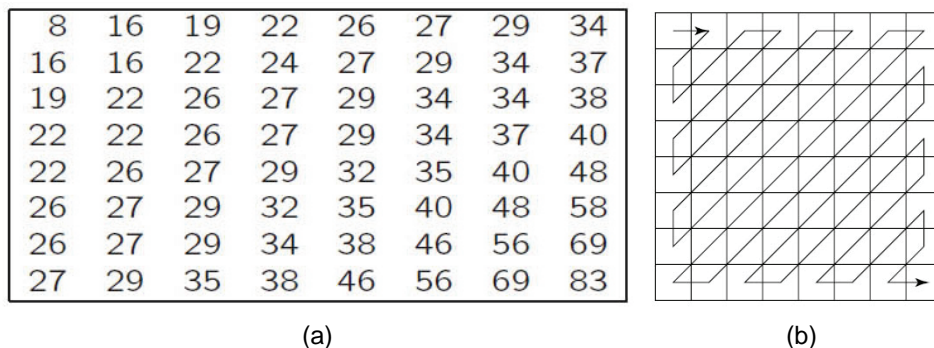


Slika 3.6. Koeficijenti *DCT* transformisanog bloka slike

Sledeći korak u intrafrejm kompresiji je kvantovanje *DCT* koeficijenata. Vrednosti koeficijenata se dele sa vrednostima iz kvantizacione matrice, a rezultat deljenja se zaokružuje na ceo broj. Ovaj postupak direktno utiče na kvalitet komprimovane slike jer dolazi do gubitka informacija. Vrednosti u kvantizacionoj matrici direktno utiču i na stepen kompresije. Na slici 3.7 (a) prikazana je kvantizaciona matrica za intrafrejm kodovanje. Na slici se može videti da vrednosti u kvantizacionoj matrici zavise od učestanosti koju predstavljaju *DCT* koeficijenti, i da su te vrednosti prilagođene vizuelno-perceptivnim karakteristikama ljudi. U *MPEG* standardu postoji mogućnost prostorno-adaptivnog kvantovanja [POP06] koje omogućava grublje kvantovanje *DCT* koeficijenata u makroblokovima gde se javlja velika dinamika pomeraja, i gde su greške kvantovanja manje uočljive. Koristi se ista kvantizaciona matrica koja se množi sa odgovarajućim parametrom, a parametar se prenosi u okviru zaglavlja makrobloka. Primenom ovakve kvantizacije stepen kompresije se dodatno uvećava.

Sledeću fazu u intrafrejm kompresiji predstavlja entropijsko kodovanje [POP06]. Primenom diferencijalne impulsne kodne modulacije (*Differential Pulse Code Modulation-DPCM*) redukuje se redundansa koja se pojavljuje kod *DC DCT* koeficijenata, a zatim se signal koduje Hafmanovim kodom koji je primenjen i u *JPEG (Joint Photographic Experts*

Group) [WAL91] kompresiji za mirne slike. *AC DCT* koeficijenti se nakon kvantizacije očitavaju po cik-cak putanji, Slika 3.7(b), a dobijeni niz vrednosti se konvertuje u parove brojeva (dužina niza istovetnih simbola, amplitudski nivo), kao u *JPEG* standardu. Sledi modifikovano Hafmanovo kodovanje signala koje se znatno razlikuje od primenjenog Hafmanovog kodovanja u *JPEG* standardu.

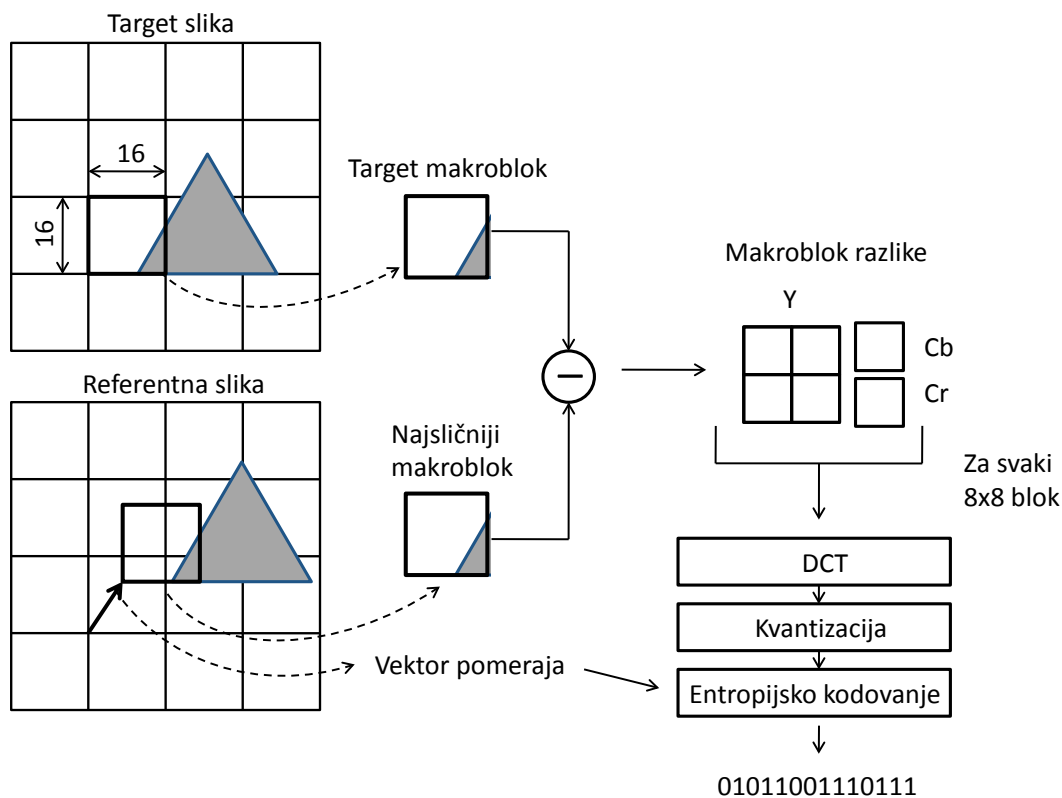


Slika 3.7. Kvantizacija u interfrejm kompresiji: (a) kvantizaciona matrica za intrafrejm kompresiju, (b) cik-cak putanja skeniranja

3.1.1.4. Interfrejm kompresija

Interfrejm kompresija se primenjuje na P i B slike u cilju uklanjanja vremenske redundanse. Koristi se kompenzacija pomeraja koja se izvodi na nivou makrobloka. Metode pretraživanja za određivanje vektora pokreta nisu specificirane u okviru MPEG-1 standarda i u tu svrhu se može primeniti bilo koja od postojećih metoda prezentovanih u literaturi [POP06], npr. algoritam nezavisnog pretraživanja po osama, algoritam paralelnog pretraživanja po osama, algoritam dvodimenzionalnog pretraživanja, algoritam trostepenog pretraživanja itd. Preciznost određivanja vektora pomeraja svodi se na tačnost od polovine piksela. Vektor pomeraja zavisi od dinamike promene sadržaja između frejmova, ali postoji ograničenje da vektor pomeraja ne sme da izađe van slike.

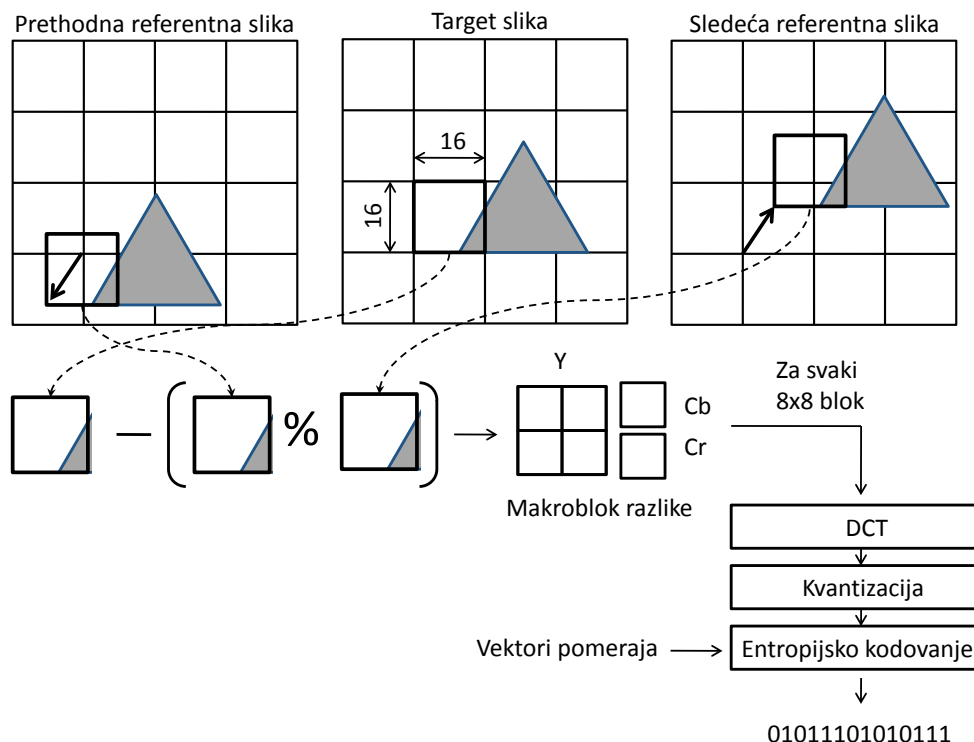
P slike se komprimuju interfrejm kompresijom pri čemu se za estimaciju pokreta koristi predikcija unapred. Pokret se procenjuje u odnosu na I slike ili prethodne P slike iz iste grupe slika. Na Slici 3.8 prikazana je principna šema realizacije interfrejm kompresije.



Slika 3.8. Principalska šema interfrejm kompresije za P slike

Target slika predstavlja tekuću sliku a referentna slika predstavlja prethodnu sliku u odnosu na koju se vrši određivanje vektora pomeraja. Standardom je predviđeno osam različitih načina realizacije interfrejm kompresije za P slike. U slučaju nepokretne pozadine, kada posmatrani makroblok nije promenio položaj u odnosu na prethodnu referentnu sliku, makroblok se ne prenosi. Ovakav slučaj se javlja kada se na slikama pojavljuje nepokretna pozadina. Sledeći slučaj je kada imamo makroblok koji se kreće, ali se pri pomeranju nije promenila njegova sadržina. Tada se prenosi samo vektor pomeraja. Ako se pri pomeranju makrobloka može utvrditi vektor pomeraja, a prilikom pomeranja došlo je do male promene njegove sadržine, na kompezovanu razliku odgovarajućih blokova iz dve slike primenjuje se DCT . Ako vektor pomeraja nije moguće odrediti, na razliku odgovarajućih blokova primenjuje se DCT ili se vrši intrafrejm kompresija kao kod I slike. Kao i kod intrafrejm kompresije kvantovanje DCT koeficijenata vrši se deljenjem sa elementima kvantizacione matrice i moguća je adaptivna kvantizacija.

B slike se takođe komprimuju primenom interfrejm kompresije. Na Slici 3.9 prikazana je principska šema realizacije interfrejm kompresije za B slike.



Slika 3.9. Principska šema interfrejm kompresije za B slike

U *MPEG* standardu B slike omogućavaju kombinaciju predikcije pomeraja i interpolacije. Naime, pored estimacije pokreta moguće je interpolirati B frejm između referentnih slika. Koduju se manjim brojem bita jer se ne koriste za dalju predikciju i ne proizvode akumulaciju grešaka. Predikcija se može izvršiti na tri načina, predikcijom unapred, predikcijom unazad i bidirekcionom predikcijom. Predikcija unapred se vrši na isti način kao za P slike. Predikcija unazad se vrši estimacijom pokreta makrobloka u odnosu na odgovarajući makroblok naredne P slike. Bidirekciona ili interpolaciona predikcija pomeraja makrobloka predstavlja aritmetičku sredinu procena unapred i unazad. Na Slici 3.8, operator $\%$ ima značenje: ili levi makroblok, ili desni makroblok, ili aritmetička sredina oba makrobloka. Primena bidirekcionom predikcije ima prednosti i mane. Sa jedne strane bidirekcionom predikcijom rešava se problem zaklonjenih ili otkrivenih

delova pozadine i dobija se bolji odnos signal/šum. Sa druge strane interpolacijom prevelikog broja B slika smanjuje se korelacija između I i P slika, pa ih je potrebno kodovati sa većim brojem bitova. Pored navedenog prevelikim brojem B slika unosi se i veće kašnjenje pri kodovanju i dekodovanju. Pamćenje po dve slike na prijemnoj i predajnoj strani za potrebe bidirekcionne predikcije predstavlja dodatni problem.

Za razliku od intrafrejm kompresije, gde su vrednosti ulaznog signala u DCT bile ograničene na interval $[0,255]$, kod interfrejm kompresije vrednosti ulaznog signala su ograničene na interval $[-256,255]$. Izlazni signali su ograničeni na interval $[-2048, 2047]$. Transformacijom greške predikcije bloka dobijaju se DCT koeficijenti koji sadrže visoke frekvencije, pa se može primeniti grublje kvantovanje korišćenjem širih koraka kvantizera.

Nakon kvantovanja u realizaciji interfrejm kompresije sledi faza kompaktnijeg pakovanja i kodovanja kvantovanih DCT koeficijenata. Kvantovani DC i AC DCT koeficijenti se primenom cik-cak putanje skeniraju i grupišu u niz, koji se konvertuje u parove brojeva (dužina niza istovetnih simbola, amplitudski nivo). Da bi se dobio što kompaktniji zapis signala, sa većim stepenom kompresije, primenjuje se entropijsko kodovanje. Kodovanje simbola se vrši na dva načina. Simboli sa većom verovatnoćom pojavljivanja se koduju sa Hafmanovim koderom sa kodom promenljive dužine, dok se simboli male verovatnoće pojavljivanja koduju pomoću simbola *escape* i kodom fiksne dužine u cilju izbegavanja kodnih reči velikih dužina. Vektori pokreta se koduju primenom $DPCM$ u odnosu na prethodne blokove.

3.1.1.5. *MPEG-1* video format

U *MPEG-1* standardu usvojen je *ITU-R. BT.601 (Y,Cr,Cb)* digitalni televizijski format poznat pod imenom *SIF (Source Input Format)*. Prilikom kodovanja u ovaj format, video sekvence se prilagođavaju *SIF* formatu odgovarajućim filtriranjem, decimacijom i ispuštanjem frejmova. *MPEG-1* podržava samo progresivno skeniranje [POP06], tako da je slika osnovni element za video kompresiju. Rezolucije slika u okviru *MPEG* standarda su: za *NTSC* [POP06] 352×240 piksela za lumentni signal odnosno 176×120 piksela za

hrominentne signale pri 30 frejm/s; za PAL [POP06] 352x288 piksela za lumentni signal odnosno 176x144 piksela za hrominentne signale pri 25 frejm/s.

Osnovna specifikacija Core MPEG-1 standarda, koja je prikazana u Tabeli 3.1, pruža mogućnost vrlo fleksibilnog definisanja parametara pri realizaciji svakog MPEG-1 koda ili dekodera.

Tabela 3.1. Core MPEG-1 specifikacija koda

Parametar	Vrednost
Horizontalna rezolucija	≤ 768
Vertikalna rezolucija	≤ 576
Broj makroblokova u slici	≤ 396
Broj makroblokova/s	≤ 9.900
Učestanost slika	≤ 30 f/s
Bitski protok	$\leq 1,856$ kb/s
Veličina bafera u dekoderu	≤ 376832 bita

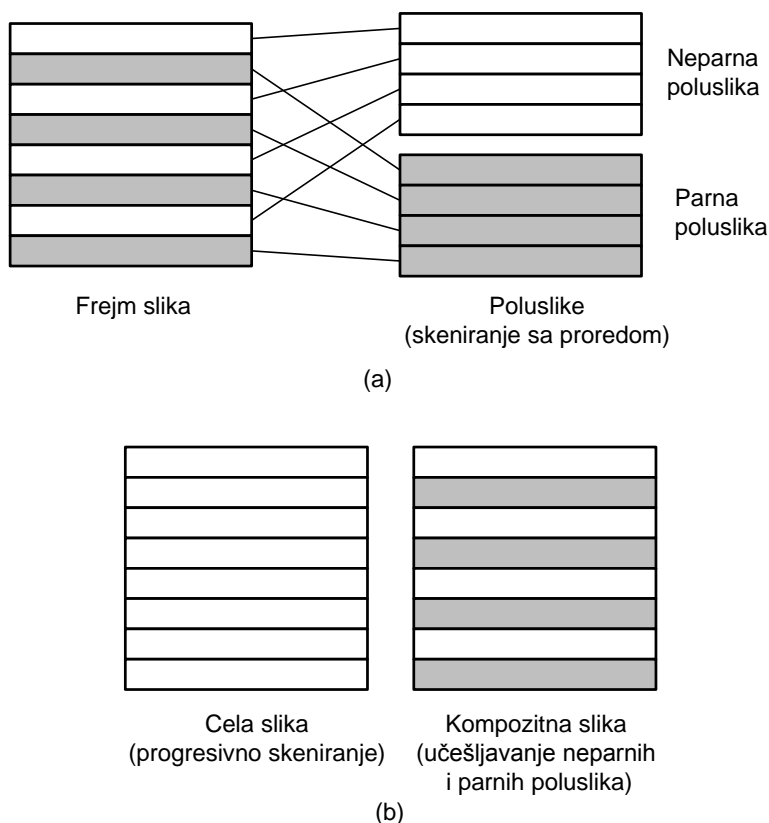
3.1.1. MPEG-2 standard

MPEG-2 standard predstavlja ekstenziju MPEG-1 standarda za video kompresiju i prezentovan je 1992. godine. Oba standarda su razvijana paralelno sa ciljem da se MPEG-2 standardom omogući širi spektar video rezolucija počevši od 352x288x30 do 1920x1152x60 i da se pokrije niz primena koje su prepoznate kao prezahtevne za MPEG-1 standard. Ovaj standard je pre svega namenjen primeni u distribuciji digitalnog televizijskog signala, ITU-R. BT.601 ili HDTV. Standard pokriva bitske protoke od 2 Mb/s do 20 Mb/s.

3.1.2.1. Tipovi slika

U MPEG-2 standardu koriste se dva tipa skeniranja slike: progresivno skeniranje i skeniranje sa proredom. Uvođenje skeniranja sa proredom je direktna posledica namere da se MPEG-2 standard prvenstveno odnosi na primenu u prenosu televizijskog signala. Kod progresivnog skeniranja očitava se linija po linija slike, dok kod skeniranja sa proredom

skeniranje ide u dva koraka. U prvom koraku skeniraju se neparne linije slike, a u drugom koraku parne linije slike, Slika 3.10(a). Struktura slike dobijene progresivnim skeniranjem u potpunosti je identična definisanoj slici sa progresivnim skeniranjem u *MPEG-1* standardu. Uvođenjem skeniranja sa proredom definisana su i dva nova tipa slika, kompozitna slika i poluslika.



Slika 3.10. Tipovi skeniranja slika u MPEG-2 standardu:
a) poluslike, b) kompozitne slike.

Kompozitna slika se formira učešljavanjem neparnih i parnih poluslika, pri čemu se novoformirana slika tretira kao celina. Ovaj tip slike pogodan je za kompresiju sekvenci sa malom dinamikom pokreta. Poluslike nastaju tako što se svaka od slika koja nastaje tokom skeniranja sa proredom, bilo neparnih ili parnih linija, tretira kao zasebna slika. Svaka od definisanih slika u *MPEG-2* standardu u procesu kompresije može se ravnopravno kodovati kao *I*, *P* ili *B* slika.

3.1.2.2. Hijerarhijska struktura podataka

Hijerarhijska struktura podataka u *MPEG-2* standardu generalno odgovara strukturi podataka u *MPEG-1* standardu prikazanoj na Slici 3.2. Pored velike strukturne sličnosti postoje velike razlike unutar pojedinačnih slojeva. Hijerarhijski na najvišem mestu se nalazi sekvenca, a zatim sledi prvi niži sloj a to je grupa slika. U okviru ovog sloja, s obzirom da su u okviru *MPEG-2* standarda definisane tri vrste slika, grupe slika mogu činiti mešavine sve tri vrste slika. Postoje odgovarajuća ograničenja po pitanju prisustva određene vrste slike unutar grupe slika. Naime, poluslike se uvek moraju pojavljivati u paru. Neparna i parna poluslika moraju biti kodovane kao par slika istog tipa, kao *P* ili *B* slike. Međutim, ako je neparna poluslika *I* tipa, u tom slučaju parna poluslika može biti *I* ili *P* tipa. Sledeći niži sloj predstavlja sloj slike koji čine informacije o slikama. Sledi sloj isečaka, a zatim sloj makroblokova i sloj blokova. Blok je definisan kao blok slike dimenzije 8×8 piksela.

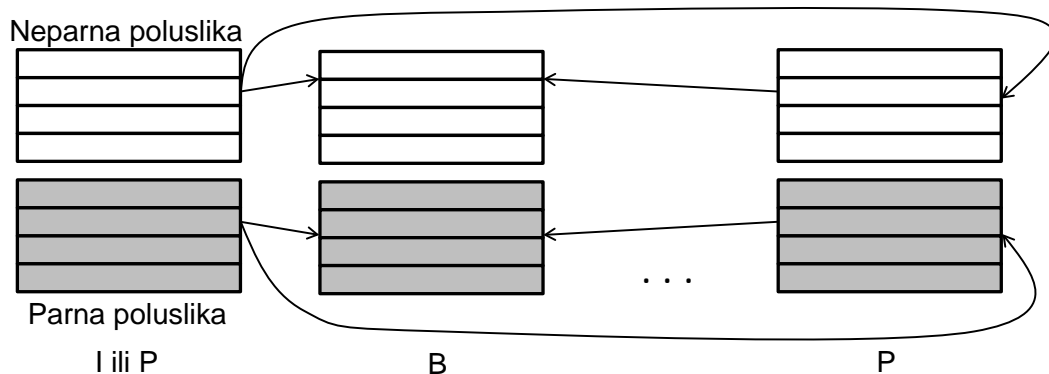
U *MPEG-2* standardu su definisana tri tipa makroblokova. Sva tri tipa sadrže isti broj blokova lumentne komponente *Y*, a razlikuju se po broju blokova hrominentnih komponentata, *Cb* i *Cr*. Prvi definisani format je $4:1:1$, gde pored četiri bloka lumentne komponente makroblok čine i po jedan blok iz svake od hrominentnih komponenti. Drugi format je $4:2:2$ koji se razlikuje od prvog formata po tome što makroblok čine po dva bloka iz svake od hrominentnih komponentata. Treći format je $4:4:4$, gde makroblok čine po četiri bloka iz svake od kolor komponenti signala.

3.1.2.3. Kompenzacija pokreta

Kompenzacija pokreta u *MPEG-2* standardu se vrši na nivou makroblokova. S obzirom da su u standardu definisane dve nove vrste slika, uvedena je dodatna mogućnost estimacije pokreta pomoću novih vrsta slika uz odgovarajuća ograničenja.

Estimacija pokreta kod slika dobijenih progresivnim skeniranjem vrši se na isti način kao u *MPEG-1* standardu. Primenjuje se estimacija unapred ili unazad, ili bidirekciona estimacija.

Kod skeniranja sa proredom, razdvajanjem poluslika dobijaju se dve nove mogućnosti za estimaciju pokreta. Moguće je izvršiti estimaciju pokreta samo sa neparnim poluslikama, odnosno samo sa parnim poluslikama. Na svaku od vrsta poluslika može se primeniti estimacija pokreta unapred ili unazad, ili bidirekciona estimacija pokreta, Slika 3.11.



Slika 3.11. Estimacija pokreta kod skeniranja sa proredom i razdvojenim poluslikama

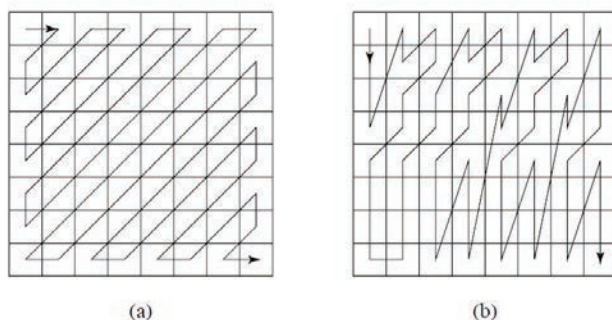
Estimacija pokreta kod kompozitnih slika je kompleksna. Kompozitne slike su dobijene skeniranjem sa proredom i kreirane su integriranjem dva različita signala. Iz ovog razloga dozvoljeno je da se estimacija pokreta za kompozitni signal može vršiti na osnovu kompozitnog signala ili na osnovu signala poluslika. S obzirom da se estimacija pokreta vrši na nivou makroblokova, na tom istom nivou se i donosi izbor o najboljem načinu estimacije. Izbor se donosi na osnovu procene dinamike pokreta u makroblokovima.

3.1.2.4. Izračunavanje *DCT* i kvantovanje koeficijenata

Izračunavanje *DCT* koeficijenata izvršava se na bloku slike dimenzije 8×8 piksela za slike dobijene progresivnim skeniranjem. *MPEG-2* standardom omogućeno je pored navedenog i izračunavanje *DCT* koeficijenta za signale poluslika koje su dobijene skeniranjem sa proredom. Izbor načina izračunavanja *DCT* koeficijenata vrši se na nivou makroblokova, jer u zavisnosti od formata makrobloka zavisi da li će se izračunati *DCT* koeficijenti poluslike ili kompozitne slike. Primer gde ne mogu da se izračunaju *DCT* koeficijenti za poluslike je $4:1:1$ format makrobloka. U ovom formatu makroblok sadrži

samo po jedan blok dimenzija 8×8 piksela iz svake od hrominentnih komponenti i *DCT* se može primeniti samo na kompozitnu sliku.

Nakon izračunavanja, *DCT* koeficijenti se očitavaju cik-cak putanjom. Pored cik-cak putanje u *MPEG-2* standardu je predviđena još jedna dodatna putanja za očitavanje koeficijenata. Njen naziv je “alternativno skeniranje”, Slika 3.12. Ovakav način očitavanja *DCT* koeficijenata je pogodan pri kompresiji signala dobijenog skeniranjem sa proredom.



Slika 3.12. Putanje za skeniranje *DCT* koeficijenata u *MPEG-2* standardu:
(a) cik-cak skeniranje, (b) alternativno skeniranje

U *MPEG-2* standardu je dozvoljena kvantizacija *DCT* koeficijenata sa većim brojem nivoa nego u *MPEG-1* standardu sa ciljem da se obezbedi bolji kvalitet slike. Kod makroblokova na koje se primenjuje intrafrejm kompresija povećan je broj nivoa za predstavljanje *DC* i *AC DCT* koeficijenata. *DC* koeficijenti se predstavljaju sa 11 bitova dok je dozvoljeni opseg vrednosti *AC* koeficijenata proširen na opseg $[-2048, 2047]$. Kod makroblokova na koje se primenjuje interfrejm kompresija dozvoljeni opseg *AC* koeficijenata je proširen na opseg $[-2048, 2047]$. Kontrola bitskog protoka vrši se na nivou makroblokova odgovarajućim parametrom, koji se prenosi u zaglavlju makroblokova.

3.1.2.5. Skalabilnost

Skalabilnost u *MPEG-2* standardu predstavlja mogućnost dekodovanja samo određenog dela iz niza podataka radi rekonstrukcije video signala određene rezolucije. Primenom skalabilnosti, iz jednog niza podataka sa izlaza koder, različiti dekoderi su u mogućnosti da rekonstruišu i prikažu video signal različitih prostornih rezolucija i učestanosti slika. Na ovaj način se kvalitet rekonstruisanog video signala prilagođava

potrebama korisnika, Signal minimalne prostorne rezolucije i minimalne učestanosti, u nizu podataka koji se može dekodovati, naziva se osnovnim slojem. Ostali slojevi se koriste pri rekonstrukciji signala veće prostorne i vremenske rezolucije signala iz osnovnog sloja. Skalabilni pristup u okviru *MPEG-2* standarda se koristi i u zaštiti od grešaka u prenosu, na način da se osnovni sloj prenosi sa boljom zaštitom od ostalih slojeva.

Oblici skalabilnosti koje podražava *MPEG-2* standard su sledeći:

1. *Prostorna skalabilnost* predstavlja sposobnost dekodera da dekoduje video signal različitih prostornih rezolucija bez prethodnog dekodovanja cele slike i decimacije dobijenog signala na zadatu rezoluciju. Osnovni sloj predstavlja kodovanu verziju video signala niske rezolucije, dobijenu decimacijom originalnog signala. U ostalim slojevima se nalaze informacije o razlici između signala odgovarajuće prostorne rezolucije i osnovnog sloja.
2. *Vremenska skalabilnost* predstavlja mogućnost dekodovanja izlaznog signala video sekvence sa različitim učestanostima slika, bez dekodovanja svih slika iz sekvence.
3. *SNR skalabilnost* predstavlja mogućnost dekodovanja izlaznog signala video sekvence korišćenjem različitih koraka kvantizera za kvantovanje *DCT* koeficijenata. Osnovni sloj se dobija kvantovanjem *DCT* koeficijenata sa najvećim korakom kvantizera.
4. *Hibridna skalabilnost* predstavlja kombinaciju prethodno opisanih tipova skalabilnosti.

3.1.2.6. Profili i nivoi u *MPEG-2* standardu

Profili u okviru *MPEG-2* standarda predstavljaju podskupove potpune sintakse standarda. Definisano je pet profila i namenjeni su različitim aplikacijama. Definisani su sledeći profili: jednostavan (*simple*) profil, glavni (*main*) profil, *SNR* skalabilni (*SNR scalable*) profil, prostorno skalabilni (*spatial scalable*) profil i visoki (*high*) profil. Naknadno su definisani *4:2:2* profil, i *multi-view* profil.

Podela unutar profila definisana je sa četiri nivoa i svaki nivo predstavlja izbor slobodnih parametara pri hardverskoj implementaciji. U Tabeli 3.2 prikazana su ograničenja parametara po nivoima i njihove primene, a u Tabeli 3.3 prikazani su definisani nivoi po profilima u *MPEG-2* standardu.

Tabela 3.2. Ograničenja parametara po nivoima u *MPEG-2* standardu za glavni profil [LI03]

Nivo	Maksimalna rezolucija (piksela/linija)	Maksimalna učestanost slika (frejm/s)	Maksimalan broj piksela/s	Maksimalni protok kodovanih podataka (Mbps)	Primena
Niski	352/288	30	$3,0 \times 10^6$	4	Korisničke trake
Glavni	720/576	30	$10,4 \times 10^6$	15	Studio TV
Visoki-1440	1440/1152	60	$47,0 \times 10^6$	60	Korisnička HDTV
Visoki	1920/1152	60	$62,7 \times 10^6$	80	Filmska produkcija

Tabela 3.3. Definisani nivoi po profilima u *MPEG-2* standardu [LI03]

Nivo	Jednostavan profil	Glavni profil	SNR skalabilni profil	Prostorno skalabilni profil	Visoki profil	4:2:2 profil	<i>Multi-view</i> profil
Niski		*	*				
Glavni	*	*	*		*	*	*
Visoki-1440		*		*	*		
Visoki		*			*		

Jednostavni profil podržava samo glavni nivo i ne dozvoljava upotrebu *B* slika. Bitski protok u ovom profilu je ograničen na maksimalnih *15 Mb/s*.

Glavni profil podržava sva četiri nivoa sa maksimalnim bitskim protocima od 4, 15, 60, i 80 Mb/s za niski, glavni, visoki-1440 i visoki nivo, respektivno.

SNR skalabilni profil podržava samo dva nivoa, niski i glavni. Maksimalni bitski protoci za ovaj profil po nivoima je 4 Mb/s i 15 Mb/s, respektivno. Bitski protok u osnovnom sloju za ova dva nivoa je 3 Mb/s i 15 Mb/s, respektivno.

U prostorno skalabilnom profilu podržan je samo visoki-1440 nivo. Maksimalni bitski protok je za ovaj nivo 60 Mb/s gde se 15 Mb/s odnosi na bitski protok u osnovnom sloju.

Visoki profil podržava tri nivoa, glavni, visoki-1440 i visoki nivo. Maksimalni bitski protoci za ove nivoe su 20 Mb/s, 80 Mb/s i 100 Mb/s, respektivno. Bitski protoci u osnovnom sloju za definisane nivo su 4 Mb/s, 20 Mb/s i 25 Mb/s, respektivno.

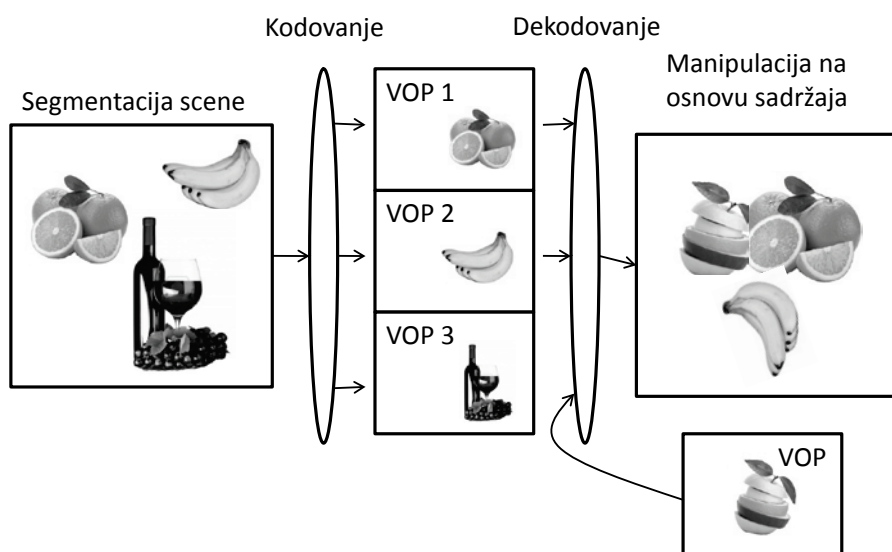
4:2:2 profil je uveden zbog potreba studijskih uređaja koji rade sa slikama u boji, u formatu 4.2:2. *Multi-view* profil je pripremi da bi se omogućilo efikasno kodovanje dve video sekvence koje potiču od dve kamere koje snimaju istu scenu, pomerene za mali ugao, dakle za potrebe stereovizije.

3.1.3. *MPEG-4* standard

MPEG-4 standard predstavlja generički standard za kodovanje audiovizuelnih informacija koji je uspostavljen 1994. godine. U ovom standardu, za razliku od njegovih prethodnika u familiji *MPEG* standarda, usvojen je novi princip kodovanja koji se zasniva na audiovizuelnim objektima. Scena se predstavlja kao skup audiovizuelnih objekata između kojih postoje hijerarhijske relacije u prostoru i vremenu. U hijerarhijskoj strukturi objekata na najnižem nivou nalaze se primitivni medijski objekti kao što su mirna slika (nepomična pozadina scene), vizuelni objekti (izdvojeni objekti bez pozadine), audio objekti (zvuk), itd. Novi princip kodovanja omogućio je veći stepen kompresije, veću

interaktivnost i mogućnost kombinovanja objekata različitih priroda u sceni, dakle integraciju mirne slike, videa, teksta, zvuka i grafike u jedan audiovizuelni sadržaj.

Kodovanje zasnovano na objektima omogućilo je, pored integracije različitih vrsta objekata u scenu, i manipulaciju sadržajem scene koja je principski prikazana na Slici 3.13. Na slici *VOP (Video Object Plane)* predstavlja ravan odgovarajućeg objekta koji je izdvojen iz scene. Takođe, ovakav pristup u kodovanju omogućava indeksiranje i pretraživanje video sekvence na osnovu sadržaja.



Slika 3.13. Manipulacija na osnovu sadržaja u *MPEG-4* standardu

U *MPEG-4* standardu definisano je osam novih funkcija koje nisu podržane od strane ostalih sličnih standarda i mogu se podeliti u tri klase:

a) Interaktivnost na bazi sadržaja:

1. Manipulacija sekvencom bita na bazi sadržaja i editovanje;
2. Alati za pristup multimedijalnim objektima na bazi sadržaja;
3. Slučajan pristup objektima;
4. Hibridno prirodno i sintetičko kodovanje podataka.

b) Kompresija:

5. Poboljšana efikasnost kodovanja signala.

c) Univerzalni pristup:

6. Robustnost na greške i gubitke informacija u šumnim sredinama;

7. Skalabilnost objekata (dodavanje i odstranjivanje video objekata);

8. Skalabilnost prostorne i vremenske rezolucije objekata.

3.1.3.1. Formati slike i bitski protoci u *MPEG-4* standardu

MPEG-4 standard podržava oba načina skeniranja slike: progresivno i sa proredom. Predstavljanje video signala u boji realizuje se u *YCbCr* kolor prostoru, sa prostornom rezolucijom *4:2:0*. Svaka komponenta se predstavlja sa 4 do 12 bitova po pikselu. Prostorna rezolucija lumentne komponente može se nalaziti u vrlo širokom opsegu, od 8×8 piksela do 2048×2048 piksela. U standardu su podržane različite vremenske rezolucije i kontinualno promenljiv broj slika u sekundi.

MPEG standard podržava kodovanje i dekodovanje pravougaonih slika i sekvenci slika, kao i kodovanje i dekodovanje slika i sekvenci slika proizvoljnog oblika.

U *MPEG-4* standardu dozvoljen je širok opseg bitskih protoka, od 5 kb/s do 38.4 Mb/s . Međutim, standard je optimizovan za primenu u tri opsega bitskih protoka: ispod 64 kb/s , od 64 kb/s do 384 kb/s i od 384 kb/s do 4 Mb/s . Standard podržava konstantni i promenljivi bitski protok.

3.1.3.2. Hijerarhijska struktura podataka u *MPEG-4* standardu

Hijerarhijska struktura *MPEG-4* niza podataka znatno se razlikuje od strukture niza podataka u standardima *MPEG-1* i *MPEG-2*. U *MPEG-4* standardu struktura podataka je orijentisana ka video objektu. Na Slici 3.14 predstavljena je hijerarhijska struktura niza podataka u *MPEG-4* standardu.

Video-objekat sekvenca (VS)
Video-objekat (VO)
Video-objekat sloj (VOL)
Grupa video-objekat ravni (GVOP)
Video-objekat ravan (VOP)

Slika 3.14. Hijerarhijska struktura niza podataka u *MPEG-4* standardu.

Hijerarhijska struktura niza podata u *MPEG-4* standardu ima petoslojnu strukturu. Na najvišem sloju nalazi se video-objekat sekvenca (*Video-Object Sequence, VS*) koja sadrži kompletnu *MPEG-4* vizuelnu scenu i koja može sadržati dvodimenzionalne ili trodimenzionalne objekte, prirodne ili sintetičke. Sledeći sloj predstavljaju podaci o video-objektu (*Video-Object, VO*), određenom objektu u sceni koji može biti proizvoljnog oblika u zavisnosti od ostalih objekata ili pozadine scene. Video-objekat sloj (*Video-Object Layer, VOL*) obezbeđuje podršku za skalabilno (višeslojno) kodovanje. Svaki video-objekat može se kodovati u jednom ili više slojeva. Video objekti su diskretizovani u vremenu i svaki vremenski odbirak predstavlja jednu ravan video objekta, na Slici 3.14 video-objekat ravan (*Video-Object Plane*). Grupisanjem ravni video objekata dobijaju se grupe, na Slici 3.14 grupa video-objekat ravni (*Group Video-Object Plane*).

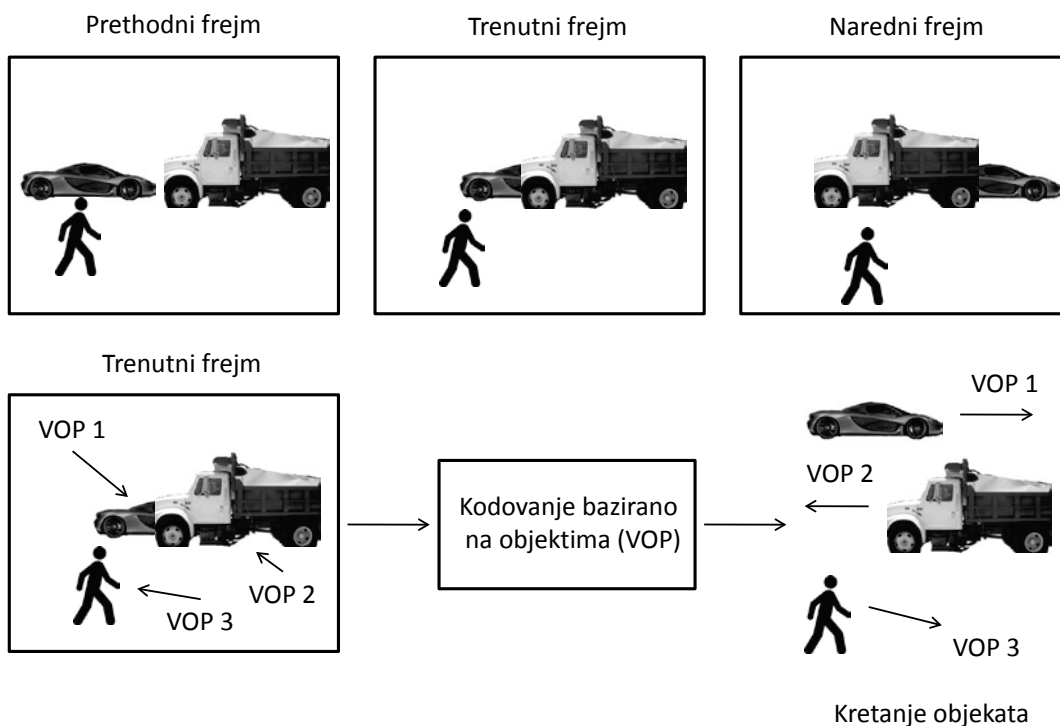
U *MPEG-4* standardu makroblok predstavlja osnovnu jedinicu za prenos podataka video signala. S obzirom da *MPEG-4* standard podržava samo jedan kolor video format, format *4:2:0*, makroblok čine 4 bloka lumentne komponente i po dva bloka hrominentnih komponenti dimenzije 8×8 piksela. Na nivou makroblokova kodovane su informacije o obliku, pokretu i teksturi (boji) piksela.

3.1.3.3. Kodovanje pokreta u *MPEG-4* standardu

Za predikciju pokreta i efikasnije kodovanje, u *MPEG-4* standardu su ponuđene dve vrste algoritama. Jedna vrsta vrši standardnu estimaciju i kompenzaciju pokreta. Druga

vrsta algoritama vrši globalnu kompenzaciju pokreta zasnovanu na prenosu statičke sprajt slike (*sprite*). Statička sprajt slika je velika mirna slika koja predstavlja panoramu pozadine.

Za standardnu estimaciju i kompenzaciju pokreta u *MPEG-4* standardu koriste se slične tehnike kao u standardima *MPEG-1* i *MPEG-2*, koje su modifikovane i primenjene na ravan video objekta (*VOP*). Definisana su tri načina kodovanja ulaznog *VOP*: Intra *VOP* (*I-VOP*), Procenjeni *VOP* (*P-VOP*) i Bidirekcionni *VOP* (*B-VOP*). Intrafrejm kodovanjem dobija se *I-VOP* pri čemu se svaki *VOP* nezavisno koduje. Interfrejm kodovanjem i procenom pokreta u odnosu na prethodno dekodovani *VOP* dobija se *P-VOP*. Procenom *VOP* u odnosu na prethodne i naredne *I-VOP* ili *P-VOP* dobija se *B-VOP*. Kompenzacija pokreta se koristi samo pri kodovanju *P-VOP* i *B-VOP*. Estimacija pokreta se obavlja na nivou makrobloka (16×16 piksela) ili na nivou bloka (8×8 piksela). Tačnost sa kojom se utvrđuje pomeraj je pola piksela.



Slika 3.15. Kodovanje pokreta u *MPEG-4* standardu.

S obzirom da objekat može imati proizvoljan oblik, pojavljuju se makroblokovi koji samo delimično pripadaju *VOP*. Za ove makroblobove vektor pomeraja se određuje modifikovanom metodom tako što se pri estimaciji uzimaju u obzir samo pikseli makrobloka koji pripadaju *VOP*. Kodovanje vektora pomeraja vrši se primenom *DPCM* i entropijskog kodovanja. Na Slici 3.15 je prikazano kodovanje pokreta na osnovu ravni video objekata koji su proizvoljnog oblika, i koji dobijaju jedinstven vektor pomeraja koji je saglasan sa stvarnim kretanjem objekta u sceni.

3.1.3.4. Kodovanje oblika i teksture u *MPEG-4* standardu

Kodovanje oblika u *MPEG-4* standardu može se izvršiti na dva načina. Binarni pristup u opisivanju oblika podrazumeva kreiranje pravougaone matrice koje obuhvata video objekat. Vrednosti matrice na mestima gde piksel pripada video objektu imaju vrednost 255, a gde ne pripada 0. Na binarnu matricu koja opisuje oblik primenjuje se kompresija bez gubitaka. Drugi pristup u kodovanju oblika naziva se generalisani binarni pristup. Objekat se u ovom slučaju opisuje pomoću nivoa sivog gde se koristi 8 *b/p* za predstavljanje vrednosti matrice koja pokriva objekat. Ovakav pristup u kodovanju oblika pruža mogućnost predstavljanja transparentnih objekata i omogućava smanjenje efekata preklapanja u spektralnom domenu. Informacije o obliku dobijene ovim pristupom koduju se primenom *DCT*.

Informacije o boji (teksturi) jednog *I-VOP* nalaze se u komponentama *YCbCr* kolor prostora, lumentnoj (*Y*) i hrominentnim (*Cb* i *Cr*). Kod *P-VOP* i *B-VOP* kodovanih ravni video objekta pojava teksture je posledica grešaka nastalih nakon kompenzacije pokreta. Informacije o teksturi koduju se standardnom 8×8 *DCT* ili primenom modifikovane *DCT* na blokove koji se nalaze na granici video objekta. Sledeći koraci u kodovanju informacija o teksturi su: kvantizacija *DCT* koeficijenata, predikcija vrednosti koeficijenata na osnovu susednih, uređivanje koeficijenata u jednodimenzionalni niz i entropijsko kodovanje razlika stvarnih i procenjenih vrednosti *DCT* koeficijenata.

3.1.4. MPEG-4/AVC p.10 (H.264) standard

H.264 standard za video kompresiju iz familije *MPEG* standarda predstavljen je 2003. godine. Video kompresija u ovom standardu zasniva se na sličnim principima kao u prethodnim *MPEG* standardima. Standard podrazumeva primenu *DCT*, kompenzacije pokreta i naprednog hibridnog kodovanja. Iako su principi slični, ovaj standard se bitno razlikuje od svojih prethodnika. Glavne prednosti u odnosu na prethodne standarde su unapređena kompenzacija pokreta, male dimenzije blokova, promenljiva veličina blokova za transformaciono kodovanje, unapređeni debloking filter i unapređeno entropijsko kodovanje. Standardom je definisana inteligentna kontrola svih algoritama u procesu kompresije.

Po performansama *H.264* standard nudi značajnu uštedu bitskog protoka u odnosu na ostale standarde za isti kvalitet reprodukovano video signala i bolji perceptivni kvalitet slike.

3.1.4.1. Struktura podataka i bitskog niza *H.264* standarda

U *H.264* standardu svaka video slika izdvojena je na makroblokove. Dimenzije makroblokova mogu biti različite. Blokova struktura unutar makroblokova definisana je sa tri kolor formata. U ovom standardu koristi se kolor prostor kroz tri različita kolor formata, *4:2:0*, *4:2:2* i *4:4:4*. Dimenzije blokova zavise od dimenzija makroblokova. Radi bolje zaštite od grešaka u prenosu video slike su dodatno izdvojene na celobrojan broj isečaka, tako da svaki isečak sadrži određen broj makroblokova. Isečci su veoma korisni za resinhronizaciju u slučaju gubitka podataka.

H.264 bitski niz je organizovan u diskretne pakete koji se nazivaju *NAL (Network Abstraction Layer)* jedinice. Svaki od ovih paketa može sadržati deo isečka slike, tako da se jedan isečak može naći u jednom *NAL* paketu ili može biti raspodeljen u više paketa. Sve *NAL* jedinice ne moraju da sadrže delove isečaka. Postoje *NAL* jedinice drugih namena kao što su jedinice za signalizaciju, zaglavlja i dodatne podatke. S obzirom da isečak sadrži deo video frejma, u uobičajenoj realizaciji bitskog niza ceo frejm je sadržan u jednom isečku, a

podaci isečka su smešteni u jednu *NAL* jedinicu. Raspodela frejmova na skoro proizvoljan broj *NAL* jedinica je veoma korisna u prenosu videa preko medijuma koji su skloni unošenju grešaka pri prenosu. Umesto ispuštanja celog frejma pri pojavi greške, dekodirer može izvršiti resinhronizaciju posle svake *NAL* jedinice.

H.264 standard podržava progresivno skeniranje slika i opciono skeniranje sa proredom. Pri skeniranju sa proredom slika se deli na dva polja. *H.264* podržava pet tipova slika: *I*, *P*, *B*, *SI* i *SP* slike. Definicija *I*, *P* i *B* slika je principski ista kao i u prethodnim *MPEG* standardima, međutim kodovanje ovih slika u *H.264* standardu značajno se razlikuje u odnosu na prethodne standarde. *SI* i *SP* slike (*Switching slike*) predstavljaju novi tip slika koji se koristi za tranziciju između dva različita *H.264* video bitska niza.

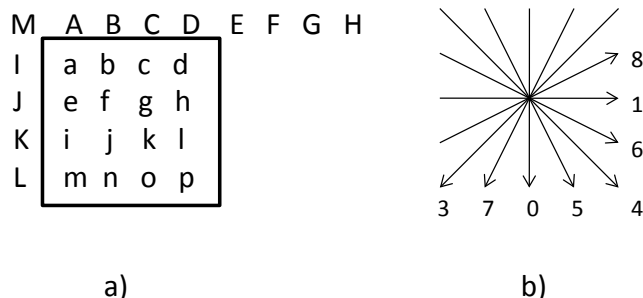
U bitskom nizu *H.264* standarda definisana su još dva skupa podataka, *Sequence Parameter Set (SPS)* i *Picture Parameter Set (PPS)*, koji sadrže osnovna zaglavlja bitskog niza. Svaki od ovih parametarskih skupova smešten je u posebnoj *NAL* jedinici i uobičajeno zauzimaju svega nekoliko bajtova. Oba parametarska skupa imaju svoje identifikacione vrednosti (*ID*) tako da se višestruki video bitski nizovi mogu prenositi u jednom osnovnom bitskom nizu.

3.1.4.2. Intrafrejm kodovanje u *H.264* standardu

U *H.264* standardu uvedena je novina u intrafrejm kodovanju koju predstavlja intra predikcija koja se izvodi pre primene *DCT*. U intra predikciji vrši se poređenje tekućeg makrobloka sa susednim makroblokovima, odnosno vrši se predikcija tekućeg makrobloka na osnovu makroblokova u njegovoj okolini koji su već kodovani. Uobičajeno je da se poređenje vrši sa susednim gornjim i levim makroblokom. Na razliku između tekućeg makrobloka i njegove predikcije primenjuje se *DCT*. Na ovaj način koduje se manja količina informacija.

H.264 standard predviđa devet modova za izvođenje intra predikcije. Za lumentni blok dimenzije 4×4 definisani su sledeći modovi: *DC* predikcioni mod (Mod 2) i 8

direkcionih modova označenih brojevima 0, 1, 3, 4, 5, 6, 7 i 8 koji su prikazani na Slici 3.16(b).



Slika 3.16. a) Prikaz intra predikcije za blok lumentne kolor komponente dimenzije 4×4 piksela, b) direkcionni modovi za intra predikciju.

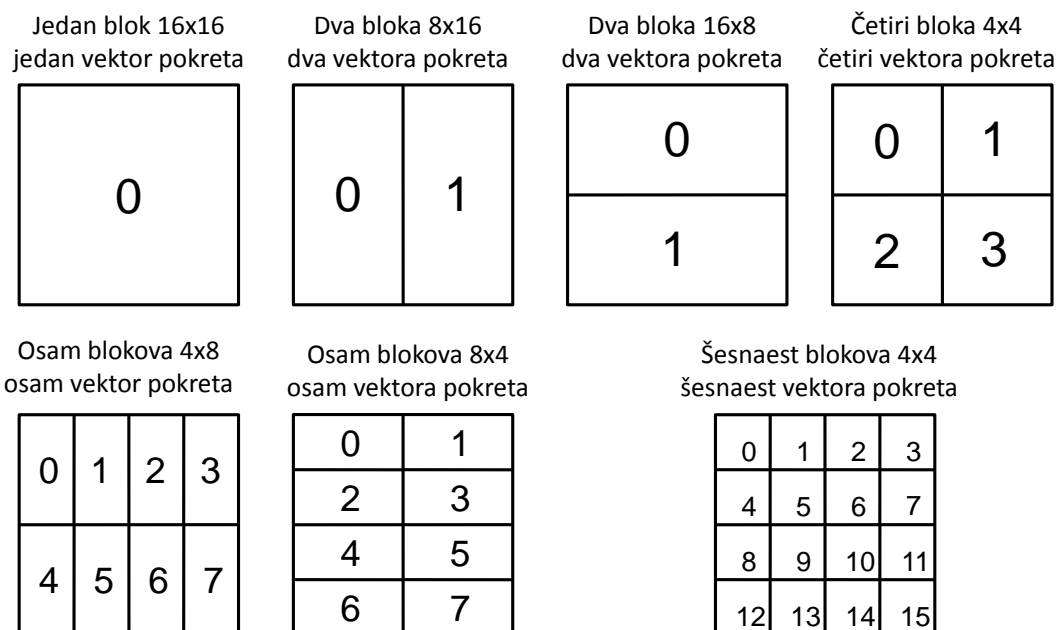
Na slici 3.16a prikazan je primer intra predikcije za lumentni blok dimenzije 4×4 piksela. Velikim slovima A, \dots, M označeni su pikseli susednih makrobloka koji su već kodovani. Unutar bloka koji se koduje vrednosti piksela su označene malim slovima a, \dots, p . Ako se pretpostavi da je odabran mod 0 (vertikalni mod) predikcija za sve vrednosti piksela u svakoj pojedinačnoj koloni unutar datog bloka je A, B, D, C po kolonama, respektivno.

Za kodovanje homogenih regiona sa manje prostornih detalja $H.264$ standard podržava 16×16 intra kodovanje, gde se bira jedan od četiri predikciona moda ($DC \pmod{2}$, vertikalni ($\pmod{0}$), horizontalni ($\pmod{1}$) i planarni ($\pmod{4}$)) koji se koriste za predikciju cele lumentne komponente makrobloka. U kodovanju hrominentnih blokova $H.264$ standard podržava 8×8 intra kodovanje pri čemu se može odabrati jedan od četiri predikciona moda kao u slučaju lumentnih blokova. Pri kodovanju slika sa regionima sa većim brojem detalja koristi se intra predikcija sa blokovima dimenzije 4×4 piksela sa mogućnošću odabira jednog od devet definisanih predikcionih modova. Predikcioni mod efikasno se koduje dodeljivanjem kraćih simbola modovima sa većom verovatnoćom pojavljivanja, gde se verovatnoća pojavljivanja određuje na osnovu modova iskorištenih za kodovanje susednih blokova.

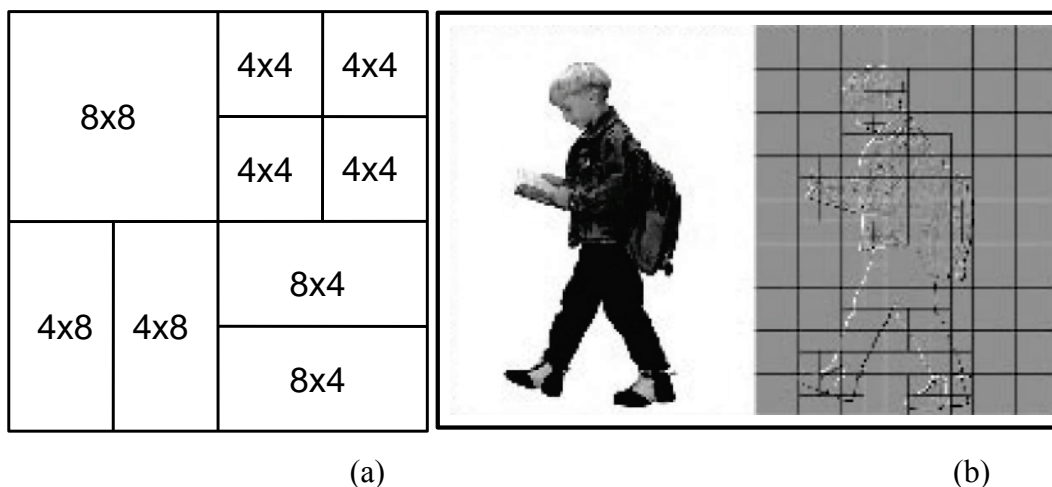
3.1.4.3. Interfrejm kodovanje u *H.264* standardu

Estimacija pokreta u *H.264* standardu podržava većinu ključnih elemenata estimacije pokreta u prethodnim *MPEG* standardima. Efikasnost estimacije u ovom standardu unapređena je dodavanjem fleksibilnosti i funkcionalnosti nasleđenim elementima iz prethodnih *MPEG* standarda. *H.264* standard podržava pored *P* i *B* slika (sa jednim ili više referentnih frejmova) s mogućnošću korišćenja proizvoljnih težinskih koeficijenata za bidirekcionu predikciju pokreta. U ovom standardu definisani su i novi tipovi *interstream* tranzicionih slika koje se nazivaju *SP* slikama. Uključivanjem *SP* slika u bitski niz omogućeno je efikasno prebacivanje između bitskih nizova sa sličnim sadržajem kodovanih sa različitim bitskim protokom, slučajan pristup i brzi modovi prikazivanja.

Kompenzacija pokreta za svaki makroblok dimenzije 16×16 piksela može se vršiti korišćenjem blokova različitih dimenzija i oblika koji su ilustrovani na Slici 3.17. Kao što se može videti na Slici 3.17 podržane su dimenzije blokova 16×16 , 16×8 , 8×16 i 8×8 piksela. Pored navedene podele makrobloka, definisana je i podela unutar bloka 8×8 piksela i podržane su dimenzije podblokova 8×4 , 4×8 i 4×4 piksela. Pojedinačni vektor pokreta može se definisati na bloku dimenzije 4×4 piksela, tako da se za jedan makroblok može preneti do 16 različitih vektora pomeraja. Deljenje makroblokova dozvoljava koderu da upravlja sa nekoliko tipova pokreta u skladu sa kompleksnošću pokreta i brzinom protoka. Generalno, velike površine particija su odgovarajuće za detektovanje pokreta u homogenim regionima frejma, dok su particije malih dimenzija korisne za detektovanje pokreta u delovima frejmova sa više detalja. Ovakav pristup rezultira boljim perceptivnim kvalitetom, manjim bloking efektom i manjim brojem artifakata. Na Slici 3.18a prikazana je jedna moguća kombinacija blokova za jedan makroblok. Na Slici 3.18b prikazan je primer korišćenja lumentne komponente originalnog frejma i upotreba delova makroblokova u oblastima sa mnogo detalja kako bi se unapredila kompenzacija pokreta.



Slika 3.17. Kodovanje pokreta u MPEG-4 standardu (blokovi).



Slika 3.18. Kodovanje pokreta u MPEG-4 standardu: a) raspodela na blokove, b) realan primer kodovanja.

U H.264 standardu tačnost kompenzacije pokreta je unapređena i podignuta na viši nivo u odnosu na prethodne MPEG standarde. Pored rezolucije tačnosti kompenzacije pokreta od polovine piksela za standardno kodovanje uvedena je i tačnost od jedne četvrtine piksela. Za potrebe porasta efikasnosti kodovanja pri velikim bitskim protocima i visokim rezolucijama primenjuje se tačnost kompenzacije pokreta od jedne osmine piksela.

Za predikciju kompleksnih pokreta, kao što su talasi, eksplozije *zoom in* i *zoom out*, korištenje samo dve referentne slike kao u prethodnim *MPEG* standardima nije dovoljno za dobru predikciju. *H.264* standard podržava predikciju pokreta pomoću višestrukih referentnih slika i daje mogućnost korištenja do pet referentnih slika pri predikciji pokreta. Ovakav pristup obezbeđuje veći subjektivni kvalitet videa i efikasnije kodovanje frejmova. Pored navedenog korištenje višestrukih referentnih slika obezbeđuje fleksibilnost bitskog niza po pitanju zaštite od grešaka.

Jedan od nedostataka kodovanja baziranog na blokovima jeste pojava diskontinuiteta na granicama blokova, naročito pri niskim bitskim protocima. Pojava ovog efekta je posledica obrade signala sa različitim kvantizerima, kodovanje različitih tipova kretanja i grešaka predikcije blokova. U cilju minimizacije ovog efekta *H.264* standardom predviđena je upotreba adaptivnog debloking filtra koji eliminiše bloking artefakte i generiše glatku sliku. Debloking filter obrađuje pojedinačno makroblokove dimenzije 16×16 piksela. Filter filtrira unutar makrobloka svaku granicu između blokova dimenzije 4×4 piksela. Debloking filter se sastoji od horizontalnog filtriranja duž svih vertikalnih ivica i vertikalnog filtriranja duž svih horizontalnih ivica između blokova dimenzija 4×4 piksela za lumentnu i hrominentne komponente makrobloka. Dva piksela sa različitih strana ivice bloka mogu se ažurirati primenom različitih filtera. Primena debloking filtra unapređuje *PSNR* (*Peak Signal to Noise Ratio*) i subjektivni kvalitet videa. Pravila upotrebe debloking filtra su komplikovana i kompleksna, a sama upotreba filtra je opciona na nivou isečka koji sadrži određeni broj makroblokova. Unapređenje subjektivnog kvaliteta je više nego dovoljno opravdanje za povećanje kompleksnosti algoritma uvođenjem ovog filtra.

3.1.4.4. Transformacija i kvantizacija u *H.264* standardu

Nakon primene intra ili inter predikcije makroblokova informacija sadržana u greški predikcije transformiše se u skup koeficijenata primenom određene transformacije. *H.264* je jedinstven po primeni prostorne transformacije bazirane na neoznačenim celim brojevima. Ova transformacija primarno se primenjuje na blokove dimenzija 4×4 piksela i predstavlja grubu aproksimaciju *DCT* transformacije. Primena transformacije na male blokove

doprinosi redukovanju bloking efekta i zvonećih artifakata. Sa druge strane precizna specifikacija koeficijenata kao neoznačenih celih brojeva eliminiše bilo kakav problem neslaganja između kodera i dekodera pri inverznoj transformaciji.

Na transformacione koeficijente, u *H.264* standardu, primenjuje se skalarna kvantizacija bez proširene mrtve zone. Definisano je pedeset i dve veličine kvantizacionih koraka koji se mogu odabrati na nivou makrobloka, što predstavlja veliku razliku u odnosu na prethodne standarde. Finoća hrominentnih komponenti je unapređena korišćenjem finijeg kvantizacionog koraka u odnosu na kvantizacioni korak koji se koristi za lumentnu komponentu, a posebno u slučajevima kada je lumentna komponenta grubo kvantovana. U standardu je dat veći opseg kvantizacionih skalirajućih faktora koji daju veću fleksibilnost koderskom mehanizmu za kontrolu bitkog protoka.

Kvantovani transformacioni koeficijenti odgovaraju različitim vrednostima, tako da se u levom gornjem uglu nalazi koeficijent koji predstavlja *DC* vrednost, a ostali koeficijenti odgovaraju različitim nenultim vrednostima frekvencije. Formiranje niza od kvantovanih transformacionih koeficijenata izvodi se na identičan način kao u prethodnim standardima, cik-cak putanjom za progresivno skeniranje slike i alternativnom putanjom za skeniranje sa proredom.

3.1.4.5. Entropijsko kodovanje u *H.264* standardu

H.264 standardom definisana su dva metoda za entropijsko kodovanje: *Context Adaptive Variable Length Coding (CAVLC)* i *Context Adaptive Based Arithmetic Coding (CABAC)*.

CAVLC metoda za entropijsko kodovanje koristi prednosti nekoliko karakteristika kvantovanih blokova dimenzije 4×4 piksela. Prvo, nenulti koeficijenti na kraju cik-cak skeniranja često imaju vrednosti ± 1 i *CAVLC* metoda koduje broj ovih koeficijenata (“prateće jedinice”) na vrlo kompaktan način. Drugo, *CAVLC* metoda efikasno koristi *run-level* kodovanje za reprezentaciju niza nula u kvantovanim blokovima dimenzije 4×4 piksela. S obzirom da su brojevi nenulatih koeficijenata u susednim blokovima često

korelisani, broj nenultih koeficijenata se koduje korišćenjem *look-up* tabele koja zavisi od brojeva nenultih koeficijenata u susednim blokovima. Sa druge strane, nivoi nenultih koeficijenata rastu u blizini *DC* koeficijenta i opadaju u okolini visokofrekvencijskih koeficijenata. *CAVLC* metoda koristi prednost ove činjenice tako što čini adaptivnim izbor *VLC (Variable Length-Coding) look-up* tabele na mestima gde izbor tabele zavisi od prethodno kodovanih nivoa.

CABAC metoda za entropijsko kodovanje koristi model verovatnoće i na strani koda i na strani dekoda za sve sintaksne elemente (transformacione koeficijente, vektore pomeraja, itd.). Povećanje efikasnosti aritmetičkog kodovanja postignuto je adaptacijom osnovnog modela verovatnoće na promenljivu statistiku unutar video frejma kroz proces kontekstnog modelovanja. Kontekstno modelovanje obezbeđuje estimaciju uslovne verovatnoće kodnih simbola. Postojeća intersimbolska redundansa može se eliminisati određivanjem odgovarajućih kontekstnih modela i prebacivanjem između različitih modela verovatnoće, u skladu sa već kodovanim simbolima u susedstvu tekućeg simbola koji se koduje.

3.1.4.6. Profili u H.264 standardu

U *MPEG-4 AVC* standardu definisana su tri profila: osnovni (*Baseline*), glavni (*Main*) i prošireni (*Extended*). Osnovni se odlikuje manjim mogućnostima sa elastičnim pristupom u pogledu zaštite od grešaka. Glavni nudi visok kvalitet kompresije. Kod proširenog dodate su još neke mogućnosti radi efikasnijeg prenosa bitskog niza.

Osnovni profil namenjen je prvenstveno za primene kao što su video konferencija i *wireless* video distribucija, dakle video znatno nižeg protoka. U ovom profilu podržane su samo *I* i *P* slike, a *B* slike nisu podržane. Implementiran je *In-loop* debloking filter i tačnost kompenzacije pokreta je četvrtina piksela. Segmentacija pokreta je realizovana po principu stabla sa najmanjim blokom za detekciju pokreta dimenzije 4×4 piksela. Entropijsko kodovanje je zasnovano na *VLC* entropijskom kodovanju (*CAVLC*). U ovom profilu zaštita od grešaka nudi vrlo napredne karakteristike kao što su fleksibilno ređanje makroblokova,

slučajno ređanje isečaka i redundantnost isečaka. U ovom profilu nije podržano skeniranje slika sa proredom.

Glavni profil je namenjen primeni u snimanju i distribuciji televizijskog signala, i u industriji zabave. Sadrži sve karakteristike osnovnog profila izuzev napredne zaštite od grešaka (elastičan pristup u zaštiti od grešaka). Podržava B slike, ali je implementirano i adaptivno težinsko označavanje za B i P slike. Podržano je skeniranje sa proredom gde je realizovano prebacivanje slika/polje na nivou slike i na nivou makrobloka. U entropijskom kodovanju koristi se *CABAC*. Treba istaći da Glavni profil nije nadskup osnovnog profila.

Prošireni profil ima namenu u strimingu, odnosno distribuciji video signala putem interneta. Podržava sve karakteristike osnovnog profila i podržava B slike. Uključuje i adaptivnu težinsku predikciju P i B slika, kao i skeniranje sa proredom sa mogućnošću prelaska iz režima u režim na nivou slike ili na nivou makrobloka. Ovaj profil nudi dodatnu fleksibilnost po pitanju zaštite od grešaka: deljenje podataka (*data partitioning*). U ovom profilu podržana su još dva tipa slika SP i SI . Prošireni profil je nadskup osnovnog profila ali ne i glavnog.

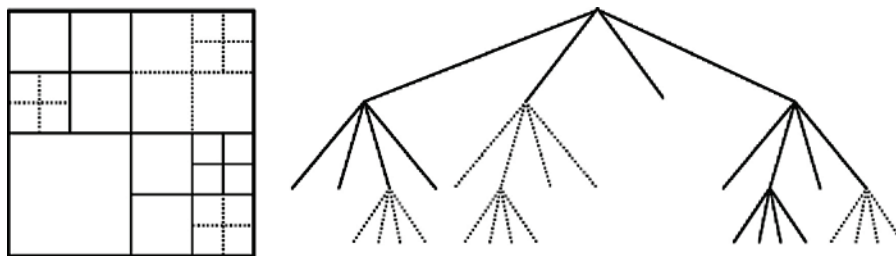
3.1.5. *H.265/HEVC* standard

H.265/HEVC (*High Efficient Video Coding*, HEVC) standard predstavlja novi standard za kompresiju video signala koji je uveden 2013. godine i predstavlja trenutno aktuelni novitet iz H-familije standarda. Ovaj standard je direktan naslednik *H.264* standarda i predstavlja hibridni algoritam kodovanja video signala zasnovan na blokovima. Osnovni kodni algoritam je hibrid intra predikcije, inter predikcije i transformacionog kodovanja. U suštini, sam standard ne donosi neku veliku novost u video kompresiji odnosu na prethodne standarde, već predstavlja skup manjih unapređenja u svim segmentima kodovanja video signala.

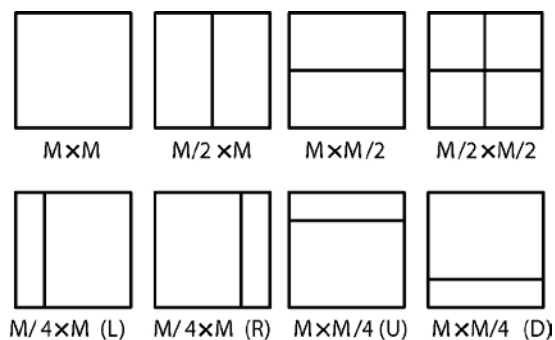
3.1.5.1. Struktura podataka u *H.265/HEVC* standardu

Za reprezentaciju kolor video signala *H.265/HEVC* standard koristi *YCbCr* kolor prostor u formatu $4:2:0$. Svaki uzorak pojedinačnih komponenti kolor prostora predstavljen je sa rezolucijom od osam do deset bitova po uzorku, pri kodovanju i dekodovanju. Video slike progresivno se skeniraju u pravouganom formatu dimenzija $W \times H$, gde W predstavlja širinu a H visinu slike za lumentnu komponentu. Hrominentne komponente za kolor format $4:2:0$ skeniraju se u pravougaonom formatu dimenzija $W/2 \times H/2$.

Kao što je već rečeno, *H.265/HEVC* standard je zadržao hibridnu arhitekturu kodovanja prethodnih *MPEG* standarda za kodovanje videa. Značajna razlika u pristupu leži u činjenici da za razliku od prethodnih *H* standarda koji kodovanje videa zasnivaju na makroblokovima, *H.265/HEVC* standard za kodovanje koristi adaptivnu *quadtree* strukturu zasnovanu na *coding tree* jedinice (*Coding Tree Unit, CTU*). U osnovi, *quadtree* struktura se sastoji od različitih jedinica i blokova. Blok se definiše kao matrica uzoraka različitih dimenzija, dok jedinica obuhvata jedan lumentni blok i odgovarajuće hrominentne blokove zajedno sa sintaksom neophodnom za njihovo kodovanje. U skladu sa tim *CTU* uključuje *coding tree* blokove (*Coding Tree Blocks, CTB*), sintaksu koja specificira podatke za kodovanje i dalju podelu strukture. Daljom podelom strukture dobijaju se kodirajuće jedinice (*Coding Units, CU*) i kodirajući blokovi (*Coding Blocks, CB*). Svaka *CU* sadrži više entiteta različitih namena, za potrebe predikcije predikcione jedinice (*Prediction Unit, PU*) a za potrebe transformacije transformacione jedinice (*Transformation Unit, TU*). Sličan sadržaj ima i svaki *CB* koji je podeljen na predikcione blokove (*Prediction Blocks, PB*) i transformacione blokove (*Transformation Blocks, TB*). Ovakav, adaptivni pristup sa korišćenjem struktura različitih dimenzija pogodan je za veće rezolucije tipa $4k \times 2k$, koje su i ciljne rezolucije za primenu nekih *HEVC* aplikacija. Primer *quadtree* strukture *CB* i *TB* prikazan je na Slikama 3.19(a) i 3.19(b). Na slikama je predstavljena hijerarhijska struktura stabla sa podelom *CTB* na *CB* i *TB*. Modovi za podelu *CB* na *PB* prikazani su na Slici 3.20.



Slika 3.19. Podela *CTB* na kodovane blokove (*CB*) i transformacione blokove (*TB*). Pune linije označavaju granice *CB* blokova, a isprekidane linije granice *TB* blokova. (a) podela *CTB* bloka. (b) odgovarajući *quadtree*.



Slika 3.20. Modovi za podelu *CB* blokova na *PB* blokove, postoje određena ograničenja veličine. Za *intrapicture* predikciju *CB* blokova, podržani su samo $M \times M$ i $M/2 \times M/2$ modovi.

Dekodovanje *quadtree* strukture ne predstavlja preveliko dodatno opterećenje jer se lako može prebaciti u hijerarhijsku strukturu primenom *z-scan* skeniranja. Predikcioni modovi za interfrejm kodovane *CU* koriste nekvadratne *PU* što zahteva neophodnu podršku pri dekodovanju u vidu dodatne logike u dekoderu koja vrši višestruke konverzije između *raster* skeniranja i *z-scan* skeniranja. U smislu očuvanja brzine bitskog protoka, sa koderske strane postoji jednostavan algoritam za analizu strukture stabla da bi se odredilo optimalno delenje blokova. Veličine *CTU* iznose 16×16 , 32×32 i 64×64 piksela.

3.1.5.2. Intrafrejm predikcija

Intra predikcija slika u *H.265/HEVC* standardu je vrlo slična kao u *H.264/AVC* standardu. Uzorci se prediktuju na osnovu rekonstruisanih susednih blokova. Kategorije modova su: *DC*, planarni, horizontalni, vertikalni i direkcioni. Nomenklatura modova

planarnog moda i direkcionih modova nešto je izmenjena u odnosu na *H.264/AVC* standard. Značajne promene unela je primena blokova većih dimenzija, jer blok može biti dimenzija 32×32 piksela i može se iskoristiti jedan od 35 modova za predikciju. Dimenzije najmanjeg bloka za intra predikciju je 4×4 piksela. Za *DC*, horizontalni i vertikalni mod *H.265/HEVC* standardom definisana je dodatna obrada, gde se kolone odnosno redovi filtriraju da bi se očuvao kontinuitet signala preko granica blokova. Kao u *H.264/AVC* standardu, referentni uzorci mogu biti filtrirani *smooth* filterom pre predikcije. Ovaj proces se primenjuje na isti način kao u prethodnom standardu ali vrlo selektivno, zavisno od predikcionog moda. Sa aspekta kodovanja povećani broj modova (35 u *H.265/HEVC* standardu u odnosu na 9 u *H.264/AVC* standardu) zahteva efikasnu logiku za odabir modova da bi se održao razuman nivo kompleksnosti algoritma za pretraživanje.

3.1.5.3. Interfrejm predikcija

Interfrejm predikcija ili kompenzacija pokreta je konceptualno vrlo slična u *H.265/HEVC* standardu ali je realizacija kompleksnija u odnosu na *H.264/AVC* standard. Implementacija kompenzacije pokreta u ovom standardu zahteva veću dužinu primenjenih filtara, veći kapacitet memorijskih bafera i veći broj operacija na hardverskom nivou. Postoje odgovarajuća ograničenja po pitanju broja bitova za predstavljanje koeficijenata filtera da bi se smanjila hardverska kompleksnost realizacije. Primena predikcije na blokove slike velikih dimenzija (npr. blok slike 64×64 piksela) ne mora uticati na povećanje hardverskih zahteva, jer se kompenzacija pokreta može izvoditi i na manjim blokovima da bi se postigao željeni kompromis između zahteva za memorijom i broja izvedenih operacija.

Kod *H.264/AVC* standarda uvedeno je ograničenje po pitanju podataka koji opisuju pokret da bi se ograničili zahtevi za memorijom. Na primer, ograničen je broj vektora pomeraja koja se mogu iskoristiti u dva susedna makrobloka. U *H.265/HEVC* standardu usvojeni su različiti pristupi i definisana su različita ograničenja, koja su mnogo jednostavnija za usklađivanje i primenu pri kodovanju. Jedan primer je definisanje najmanjeg bloka lumentne komponente za kompenzaciju pokreta koji je dimenzija 4×8 i

8×4 piksela. Dodatno ograničenje je da se za ovakve blokove može koristiti samo prva referentna slika iz liste referentnih slika, što znači da se za ove blokove ne može primeniti bi-predikcija.

U *H.265/HEVC* standardu uveden je *merge* mod, koji postavlja sve parametre pokreta nekog bloka kojeg treba interfrejm prediktovati na vrednosti parametara *merge* kandidata. *Merge* mod i proces predikcije vektora pomeraja za kodovanje vektora pomeraja opciono dozvoljavaju ponovnu upotrebu vektora pomeraja prethodnih slika što je u osnovni slično *H.264/AVC temporal direct* modu. U *H.264/AVC* standardu vektor pomeraja se određuje na nivou bloka dimenzija 8×8 piksela, dok u *H.265/HEVC* standardu vektor pomeraja se određuje na nivou bloka dimenzije 16×16 piksela. Utvrđivanje vektora pomeraja na većem bloku smanjeni su memorijski zahtevi pri interfrejm predikciji u *H.265/HEVC* standard.

H.265/HEVC standard nudi više načina za podelu slike na obrasce podeljenih po kompenzaciji pokreta. Za neki intra prediktovan *CB*, *PB* uzorci su dobijeni iz odgovarajućeg bloka regiona u referentnoj slici identifikovanog indeksom referentne slike, koji se nalazi na poziciji pomerenoj horizontalnim i vertikalnim komponentama vektora pomeraja. Izuzev u slučajevima kada vektor pomeraja ima celobrojnu vrednost, interpolacija frakcija uzoraka se koristi za generisanje predikcionih uzoraka za necelobrojne pozicije uzoraka.

3.1.5.4. Transformacija i kvantizacija

U *HEVC* standardu definisane su transformacije kao direktno množenje matrica (dimenzija 4×4 , 8×8 , 16×16 i 32×32) sa fiksnim brojem tačaka. Definisane transformacije iako predstavljaju celobrojne aproksimacije diskretne kosinusne transformacije, zadržale su osobinu simetričnosti *DCT* i omogućavaju parcijalnu *butterfly* implementaciju. Za fiksnu dimenziju transformacije 4×4 , pored celobrojne aproksimacije *DCT*, definisana je alternativna transformacija u vidu aproksimacije diskretne sinusne transformacije (*Discrete Sine Transformation, DST*). Kada je u pitanju kvantizacija u *H.265/HEVC* standardu primenjena je uniformna rekonstrukciona kvantizacija (*Uniform Reconstruction*

Quantization, URQ) sa mogućnošću primene kvantizacionih skalirajućih matrica koje se mogu primeniti na transformacione blokove različitih dimenzija.

3.1.5.5. Entropijsko kodiranje

H.265/HEVC standardom definisan je samo jedan metod za entropijsko kodovanje pod nazivom *CABAC* (*Context Adaptive Binary Arithmetic Coding*). Ovaj metod je sličan *CABAC* metodi koji se primenjuje u *H.264/AVC* standardu, ali sadrži i nekoliko unapređenja u cilju brzine izvođenja i smanjenja zahteva za hardverskim resursima. Prvo unapređenje *CABAC* metode u *H.265/HEVC* standardu, u odnosu na standard *H.264/AVC*, je znatno smanjenje broja konteksta čime se povećala efikasnost kodovanja i smanjili su se memorijski zahtevi. Sledeće unapređenje je adaptivno skeniranje koeficijenata koje se izvodi na podblokovima dimenzija 4×4 piksela za transformacione blokove svih dimenzija. Skenirajuće metode su: dijagonalno skeniranje u pravcu sporedne dijagonale od dole na gore, horizontalno i vertikalno skeniranje. Odabir metode skeniranja koeficijenata zavisi od pravca i smera intrafrejm predikcije. Pri kodovanju transformacionih koeficijenata, kao i u *H.264/AVC* standardu, u *H.265/HEVC* standardu prenosi se pozicija poslednjeg nenultog transformacionog koeficijenta, mapa značajnosti, bit znaka i nivoi transformacionih koeficijenata. Za svaki od ovih elemenata proces ekstrakcije je doživeo nekoliko promena u *H.265/HEVC* standardu u cilju veće efikasnosti kodovanja, a posebno da bi se obezbedilo bolje upravljanje celim procesom kada se koriste transformacioni blokovi većih dimenzija.

3.1.5.6. *In-Loop* filtri

Nakon dekodovanja videa, neophodno je dodatno filtriranje da bi se obezbedio bolji vizuelni kvalitet videa. Neophodno je redukovati blok artefakte koji su posledica kodovanja zasnovanog na blokovima. *H.265/HEVC* standardom definisana su dva koraka filtriranja unutar interfrejm predikcione petlje. Prvi korak predstavlja debloking filter koji je sličan debloking filteru u *H.264/AVC* standardu i primenjen je samo na uzorke slike na granicama blokova. U odnosu na prethodni standard dizajn ovog filtra je pojednostavljen u okviru procesa donošenja odluke i procesa filtriranja. Sa druge strane proces filtriranja je

prilagođen uslovima paralelnog procesiranja. Drugi korak u procesu poboljšanja vizuelnog kvaliteta slike je primena *Sample Adaptive Offset (SAO)* filtra. Ovaj filter se primenjuje adaptivno na sve uzorke slike u skladu sa odgovarajućim uslovima, npr. u skladu sa gradijentom. Cilj ove faze filtriranja je bolja rekonstrukcija amplituda originalnog signala korištenjem *look-up* tabele koja je opisana sa par dodatnih parametara koji se mogu odrediti analizom histograma sa koderske strane.

Kao što je već rečeno *deblocking* filter je primenjen na sve uzorke slike susedne sa granicama predikcionih ili transformacionih jedinica interfrejm predikcionih kodnih blokova. Izuzetak je u slučajevima kada se te granice poklapaju sa granicama slike ili kada je *deblocking* onemogućen na površinama jedinica slike koje se koriste pri paralelnom procesiranju (isečci). Dakle, *deblocking* filter može opciono biti onemogućen i tu informaciju koder signalizira dekomoderu sintaksnim elementima u *SPS* i hederu isečaka.

Za razliku od *H.264/AVC* standarda gde se *deblocking* filter primenjuje na osnovu dimenzije 4×4 piksela, *H.265/HEVC* standard primenjuje *deblocking* filter na ivice koje su poravnate sa mrežom osnove 8×8 piksela i za lumentnu i hrominentne komponente. Ovakvim pristupom izbegnuta je kompleksnost izračunavanja bez primetljive degradacije vizuelnog kvaliteta. Unapređena je paralelna primena ovog filtra, tako što su onemogućene kaskadne interakcije između susednih filtarskih operacija.

Redosled primene *deblocking* filtra je definisan tako da se prvo primenjuje horizontalno filtriranje na sve vertikalne ivice na slici, a zatim se primenjuje vertikalno filtriranje horizontalnih ivica. Ovakav redosled procesiranja *deblocking* filtrom omogućava primenu višestrukih horizontalnih ili vertikalnih filtriranja u paralelnim procesima. Nasuprot ovakvom paralelnom pristupu, filtriranje može biti obavljeno na *CTB* osnovi tako što se obrađuje jedan po jedan *CTB* blok čime se unosi određeno kašnjenje u proces filtriranja.

SAO filtriranje modifikuje dekodovane uzorke kontrolisanim dodavanjem ofset vrednosti svakom uzorku posle primene *deblocking* filtra na osnovu vrednosti u *look-up* tabeli koju je prosledio koder. Kontrolisano dodavanje je uslovljeno zadovoljavanjem

odgovarajućeg uslova. Ovaj tip filtriranja primenjuje se na regionalnoj osnovi, zavisno od odabranog tipa filtriranja za *CTB* blokove u odgovarajućem sintaksnom elementu.

3.1.5.7. Specijalni modovi kodovanja

U *H.265/HEVC* standardu definisana su tri specijalna tipa kodovanja koji se mogu primeniti kako na nivou kodnih jedinica tako i na nivou transformacionih jedinica.

Prvi mod se naziva *I_PCM* mod u kome je zaobiđena predikcija, transformacija, kvantizacija i entropijsko kodovanje, već su uzorci reprezentovani direktno sa predefinisanim brojem bitova. Glavna primena ovakvog moda je da se izbegne nepotrebna potrošnja bitova za video signale ekstremno neobičnih karakteristika, koji se ne mogu efikasno obraditi hibridnim kodovanjem (signali slični šumu). *Lossless* mod tip je tip kodovanja gde su zaobiđeni transformacija, kvantizacija i ostala procesiranja koja utiču na dekodovanu sliku (*deblocking* i *SAO* filtri), a ostatak signala koji preostaje posle interfrejm ili intrafrejm predikcije je direktno usmeren na entropijsko kodovanje. Pri entropijskom kodovanju koriste se isti konteksti susedstva koji se inače koriste za kvantovane transformacione koeficijente. Na ovaj način obezbeđena je rekonstrukcija matematički bez gubitaka, bez definisanja dodatnih alata za kodovanje. *Transform skipping* mod predstavlja model kodovanja u kojem je samo transformacija zaobiđena. Ovakav mod unapređuje kompresiju različitih tipova video sadržaja gde su kombinovani sadržaj sa kamere i računarski generisane slike ili grafika. Jedan primer je skrolovanje teksta. Ovaj mod se može primeniti samo na transformacione blokove dimenzija 4×4 piksela. *SAO* i *deblocking* filter nisu primenjeni na *lossless* mod regionima, dok se odgovarajućom kontekstnom informacijom kontroliše gde su primenjeni na *I_PCM* regione.

3.1.5.8. Paralelizacija

Visok stepen paralelizacije podrazumeva sposobnost simultanog procesiranja višestrukih regiona neke slike. Podrška za paralelizaciju mora biti obezbeđena i na koderskoj i na dekoderskoj strani u vidu višestrukih identičnih procesorskih jezgara koja se mogu koristiti paralelno. *H.265/HEVC* standard podržava tri koncepta koja omogućavaju

određen nivo visokog stepena paralizacije, i to su isečci (*slices*), pločice (*tiles*) i talasni frontovi (*wavefronts*).

Primena koncepta isečaka podrazumeva podelu slike na grupe uzastopnih *CT* jedinica u raster skenirajućem redosledu, gde se svaki isečak prenosi u posebnoj *NAL* jedinici koja se može nezavisno raščlaniti i dekodovati, izuzev u slučaju *interslice* filtriranja. Isečci prekidaju predikcione zavisnosti na svojim granicama, što izaziva smanjenje efikasnosti kodovanja i može stvoriti vidljive artefakte na prekidima između isečaka. Dizajn isečaka nije primarno određen potrebama paralelizacije tehnike kodovanja već zahtevima za što veću otpornost na greške u prenosu i maksimalnom veličinom jedinice za prenos.

Pločice kao koncept predstavljaju podelu slike na višestruke pravougaone regione, vertikalno i horizontalno. Pločice kao i isečci prekidaju predikcionu zavisnost na svojim granicama. Unutar slike uzastopne pločice su postavljene u raster skenirajućem redosledu. Redosled *CTB* jedinica unutar svake pločice je u raster skenirajućem redosledu, i ograničen je završecima svake pločice. Primenom horizontalne podele slike, pločice se mogu upotrebiti za smanjenje dimenzija linijskog bafera kodera, jer se radi sa regionom koji je uži nego cela slika. Primenom pločica omogućava se kreiranje slike iz višestrukih pravougaonih elemenata koji su nezavisno kodovani.

Talasni frontovi dele sliku na *CTU* vrste, gde se svaka *CTU* vrsta može procesirati u različitim programskim nitima. Predikciona zavisnost je očuvana između vrsta, izuzev u slučaju *CABAC* kontekstnog stanja, koje se iznova inicijalizuje na početku svake *CTU* vrste. Da bi se unapredila efikasnost kompresije, umesto standardne *CABAC* inicijalizacije, nasleđuje se kontekstno stanje iz druge *CTU* jedinice prethodne vrste, uz primenu jednostavne dvodimenzionalne adaptacije.

3.1.5.9. Profili, rangovi i nivoi

Profili, rangovi i nivoi određuju tačke usaglašavanja za primenu standarda na interoperabilan način na različite aplikacije koje imaju slične funkcionalne zahteve. Profil

definiše skup alata za kodovanje ili algoritama koji se mogu koristiti za generisanje usaglašenog bitskog niza, gde nivo postavlja ograničenja na određene ključne parametre bitskog niza u skladu sa mogućnostima procesiranja i memorijskim kapacitetom dekodera.

Parametri bitskog niza čija su ograničenja definisana nivoom su maksimalna brzina uzorkovanja, maksimalna veličina slike, maksimalna brzina bitskog protoka, minimalni nivo kompresije i kapaciteti bafera za dekodovanu sliku i kapacitet bafera za kodovanu sliku koji sadrži komprimovane podatke pre dekodovanja za potrebe upravljanja protokom podataka. Pri razradi dizajna *H.265/HEVC* standarda utvrđeno je da postoje neke aplikacije koje se razlikuju samo na osnovu maksimalne brzine bitskog protoka i kapaciteta bafera za dekodovanu sliku. Da bi se razrešio ovaj problem uvedena su dva ranga koja su specificirana samo za određene nivoe, Glavni rang (*main tier*) za većinu aplikacija, i Viši rang (*high tier*) za primenu u najzahtevnijim aplikacijama.

U *H.265/HEVC* standardu definisana su tri osnovna profila *Main*, *Main 10* i *Main Still Picture*, a početkom 2014. godine u standard je uvedena definicija još 16 profila. To su *Main 12*, *Main 4:2:2 10*, *Main 4:2:2 12*, *Main 4:4:4*, *Main 4:4:4 10*, *Main 4:4:4 12*, *Monochrome 12*, *Monochrome 16*, *Main 12 Intra*, *Main 4:2:2 10 Intra*, *Main 4:2:2 12 Intra*, *Main 4:4:4 Intra*, *Main 4:4:4 10 Intra*, *Main 4:4:4 12 Intra*, *Monochrome 12 Intra*, i *Monochrome 16 Intra* profili. Neke od karakteristika po kojima se profile razlikuju su broj bitova po uzorku, formati uzorkovanja hrominentnih kolor komponenti, monohromatski format, težinska predikcija visoke preciznosti itd. U okviru standarda definisano je 13 različitih nivoa: *1*, *2*, *2.1*, *3*, *3.1*, *4*, *4.1*, *5*, *5.1*, *5.2*, *6*, *6.1*, *6.2*. Rangovi su definisani od četvrtog nivoa pa na više. U Tabeli 3.4, kao primer, prikazana su ograničenja za najniži nivo *1* i najviši nivo *6.2* definisanih u *H.265/HEVC* standardu.

Maksimalna učestanost slika (*frame rate*) podržana u *H.265/HEVC* standardu je 300 fps (*frame per second*). The *MaxDpbSize* (*Maximal Decoded Picture Buffer Size*) je maksimalni broj slika u baferu za dekodovanu sliku predstavljen u Tabeli 3.4.

Tabela 3.4. Ograničenja za najniži nivo 1 i najviši nivo 6.2 definisane u HEVC standardu

Level	Max luma sample rate	Max luma picture size	Max bit rate for Main, Main 10, and Main 12 profiles (kbit/s)[A]		Example picture resolution @ highest frame rate[B]
	(samples/s)	(samples)			(MaxDpbSize[C])
			Main tier	High tier	
1	552.960	36.864	128	–	128x96@33,7 (6)
					176x144@15,0 (6)
6.2	4.278.190.080	35.651.584	240.000	800.000	3.840x2.160@300,0 (16)
					7.680x4.320@128,0 (6)
					8.192x4.320@120,0 (6)

3.2. MPEG-7 obeležja za opisivanje video sadržaja

MPEG-7 Visual standard [SIK01] predstavlja standard, definisan u okviru *MPEG-7* [CHA01] porodice standarda, čija je uloga specifikacija deskriptora zasnovanih na sadržaju. Ovi deskriptori omogućavaju korisnicima i sistemima za pretragu, merenje sličnosti između slika i video sekvenci na osnovu vizuelnih kriterijuma [SIK01]. Slike i video sekvence opisane ovakvim deskriptorima mogu se efikasnije identifikovati, filtrirati i pretraživati na osnovu sadržaja. Skup alata (*MPEG-7 Visual Description Tools*) za opisivanje sadržaja primenom vizuelnih kriterijuma sastoji se od elementarnih struktura i obeležja, koji pokrivaju elementarne vizuelne karakteristike kao što su boja, tekstura, oblik, pokret, lokalizacija i identifikacija lica. Za svaku od navedenih kategorija vizuelnih karakteristika, *MPEG-7* standardom definisana su osnovna i sofisticirana obeležja. U nastavku teksta dat je sažet opis struktura i obeležja definisanih *MPEG-7 Visual* standardom.

3.2.1. Osnovne strukture

MPEG-7 Visual standardom definisano je pet vizuelno povezanih osnovnih struktura: mrežasta raspodela (*grid layout*), vremenske serije (*time series*), višestruki izgled (*multiple view*), prostorne 2D koordinate (*spatial 2D coordinates*) i vremenska interpolacija (*temporal interpolation*). U nastavku dat je kratak opis svake od navedenih struktura.

3.2.1.1. Mrežasta raspodela

Mrežasta raspodela predstavlja deljenje slike na skup pravougaonih regiona jednakih dimenzija koji se mogu odvojeno opisati. Svaki od regiona raspodele može biti opisan u smislu neke vizuelne karakteristike kao što su boja ili tekstura, primenom odgovarajućih obeležja. Obeležja dozvoljavaju dodeljivanje podobeležja svim pravougaonim regionima ili proizvoljnom podskupu pravougaonih regiona raspodele.

3.2.1.2. Vremenske serije

Vremenske serije definišu vremenske serije obeležja u video segmentu i obezbeđuju funkcionalnost poređenja slike i video frejma, odnosno poređenje između video frejmova. Vremenske serije se dele na dve grupe, regularne vremenske serije (*regular time series*) i neregularne vremenske serije (*irregular time series*). Kod regularnih vremenskih serija obeležja su regularno lokalizovana, u konstantnim intervalima unutar nekog vremenskog prozora. Ovaj tip vremenskih serija omogućava jednostavnu reprezentaciju za aplikacije koje zahtevaju nizak nivo kompleksnosti. U slučaju neregularnih vremenskih serija obeležja su lokalizovana neregularno, u različitim intervalima unutar nekog vremenskog prozora. Ovakve strukture obezbeđuju efikasnu reprezentaciju za aplikacije koje zahtevaju uzak propusni opseg za prenos ili male mogućnosti skladištenja podataka. Takođe, vrlo su korisne pri kreiranju obeležja koja sadrže vremenske serije obeležja.

3.2.1.3. 2D-3D višestruki izgledi

Višestruki izgledi (*2D-3D*) predstavljaju strukturu koja kombinuje *2D* obeležja pri reprezentaciji vizuelnih obeležja *3D* objekta viđenih iz različitih uglova. Primenom ove strukture formira se kompletna *3D* reprezentacija objekta zasnovana na *2D* izgledu. Pri opisivanju *2D* izgleda može se koristiti bilo koji od *2D* vizuelnih obeležja kao što su obeležja oblika, teksture ili boje. Ova struktura podržava integraciju *2D* obeležja korišćenih za opisivanje karakteristika realnih *3D* objekata u ravni slike. Funkcionalnost poređenja omogućena je na nivou *3D* objekata tako što se porede izgledi objekata, i na nivou izgleda tako što se porede *2D* izgledi i *3D* objekti.

3.2.1.4. Prostorne 2D koordinate

Prostorne *2D* koordinate definišu *2D* prostorni koordinatni sistem i jedinicu koja će biti iskorištena kao referenca u drugim deskripcionim šemama, kada to bude relevantno. Koordinatni sistem je definisan mapiranjem između slike i koordinatnog sistema. Jedna od prednosti primene ovakve strukture je da promenom veličine slike ili isecanjem dela slike *MPEG-7* opis ne mora biti promenjen. U ovom slučaju neophodan je samo opis mapiranja između editovane i originalne slike. Ova struktura podržava dva tipa koordinatnih sistema: lokalni i integrisani. U lokalnom koordinatnom sistemu koordinate koje su korištene za izračunavanje opisa mapirane su u odnosu na tekući koordinatni sistem slike. U slučaju integrisanog koordinatnog sistema svaka slika (frejm), na primer nekog videa, može biti mapirana u različitim oblastima u odnosu na prvi frejm kadra ili videa.

3.2.1.5. Vremenska interpolacija

Vremenska interpolacija kao struktura podrazumeva vremensku interpolaciju korištenjem povezanih polinoma. Ova struktura se može iskoristiti za aproksimaciju vrednosti neke višedimenzionalne promenljive koja se menja u vremenu, kao što je recimo pozicija objekta u nekom videu. Veličina opisa vremenske interpolacije je obično znatno manja nego opis svih vrednosti promenljive.

3.2.2. Kolor obeležja

Boja predstavlja jednu od najšire korišćenih vizuelnih karakteristika u pretraživanju slika i videa. Kolor karakteristike su relativno otporne na promene u pozadinskim bojama i ne zavise od dimenzija slike niti od orijentacije. Ova obeležja mogu se koristiti za opisivanje sadržaja na mirnim slikama i videu. U okviru *MPEG-7* standarda definisan je niz kolor obeležja, od kojih je svako od obeležja pogodno za specifične funkcionalnosti poređenja sličnosti. Definisano je sedam kolor obeležja: kolor prostor (*color space*), kolor kvantizacija (*color quantization*), dominantne boje (*dominant colors*), skalabilno kolor obeležje (*scalable color*), kolor raspodela (*color layout*), strukturno kolor obeležje (*color-structure*) i kolor obeležje grupa frejmova/slika (*Group of Frames/Group of Pictures color descriptor, GoF/GoP color descriptor*). U nastavku teksta dat je kratak opis svakog od navedenih kolor obeležja.

3.2.2.1. Kolor prostor obeležje

Kolor prostor obeležje specificira odabir kolor prostora koji će biti korišćen u nekom drugom kolor obeležju, posebno pri korišćenju obeležja dominantnih boja. U *MPEG-7* standardu podržani su sledeći kolor prostori: *RGB*, *YCbCr*, *HSV*, *HMMD*, linearna transformacija matrica gde se kao referenca koristi *RGB* kolor prostor i monohromatski (*monochrome*) kolor prostor. Kolor prostor obeležje definiše kolor komponente kao entitet kontinualnih vrednosti. Za diskretnu reprezentaciju neophodna je kvantizacija.

3.2.2.2. Obeležje kolor kvantizacije

Obeležje kolor kvantizacije definiše uniformnu kvantizaciju kolor prostora, izuzev u slučaju *HMMD* kolor prostora koji se primenjuje kod strukturnog kolor obeležja. Ovo obeležje specificira broj kvantizacionih nivoa za svaku kolor komponentu kolor prostora. Broj binova kolor histograma, koji su rezultat procesa kvantizacije, moguće je podešavati tako da je obezbeđena velika fleksibilnost za širok opseg primena. Za potpunu primenu ovog obeležja u kontekstu *MPEG-7 Visual* standarda, neophodno je kombinovati ovo

obeležje sa obeležjem dominantnih boja. Kolor kvantizacijom definisane su vrednosti dominantnih boja.

3.2.2.3. Obeležje dominantnih boja

Obeležje dominantnih boja obezbeđuje kompaktan opis reprezentativnih boja na nekoj slici ili regionu slike. Pretraživanje baza slika na osnovu ovog obeležja svodi se na poređenje sličnosti skupova dominantnih boja. Pri izdvajanju ovog obeležja koristi se kolor kvantizacija, pri čemu se izdvaja mali broj reprezentativnih boja iz svake slike ili regiona. Pored definisane vrednosti izračunava se i procentno učešće odgovarajuće boje na slici. Nad celim obeležjem definisana je prostorna koherentnost, koja se koristi u pretraživanju na osnovu sličnosti.

3.2.2.4. Skalabilno kolor obeležje

Skalabilno kolor obeležje obezbeđuje opis raspodele boje na slici. Osnovu ovog obeležja predstavlja kolor histogram kodovan primenom *Haar* transformacije. Binarna reprezentacija ovog obeležja je skalabilna u smislu broja binova histograma i broja bitova za reprezentaciju. Ova reprezentacija je prilagodljiva širokom opsegu brzina prenosa podataka. Za opis boje koristi se *Hue Saturation Value (HSV)* kolor prostor uniformno kvantovan na 256 nivoa. Da bi se dobila kompaktnija reprezentacija, vrednosti histogramskih binova su neuniformno kvantizovane u opsegu od 16 bitova po histogramu za grubu reprezentaciju do 1000 bitova po histogramu za aplikacije gde je neophodan visok kvalitet reprezentacije raspodele boje. Poređenje skalabilnih kolor reprezentacija svodi se na poređenje *Haar* koeficijenata ili na poređenje vrednosti binova histograma primenom L_1 norme. Preciznost pretraživanja raste sa povećanjem broja bitova koji se koriste za reprezentaciju.

3.2.2.5. Obeležje raspodele boje

Obeležje raspodele boje predstavlja vrlo kompaktan opis prostorne raspodele boje na slici, koji ne zavisi od rezolucije slike. Slika se deli na regione primenom mrežaste (*Grid*

layout) strukture, a raspodela boje u svakom od regiona opisana je obeležjem dominantnih boja. Ovo obeležje može se koristiti za različita pretraživanja na osnovu sličnosti, filtriranje sadržaja i vizuelizaciju. Posebno korisnu primenu ovo obeležje može imati u pretraživanju baziranom na prostornoj strukturi. Na primer, u pretraživanju na osnovu crteža ili identifikaciji video segmenta. Funkcionalnosti obeležja raspodele boja svode se na poređenje slika i video zapisa, kao i poređenje crteža sa slikom odnosno videom.

3.2.2.6. Strukturno kolor obeležje

Strukturno kolor obeležje reprezentuje sliku sa dva različita aspekta, jedan je raspodela boje na slici a drugi je informacija o strukturi sadržaja. Drugim rečima ovim obeležjem izražavaju se lokalne kolor karakteristike slike. Glavna funkcionalnost ovog obeležja je primena u poređenju mirnih slika, gde se slika može sastojati od jednog pravouglog frejma ili proizvoljno oblikovanog regiona koji ne mora biti potpuno povezan. Izdvajanje ovog obeležja zasniva se na pomeranju strukturnog elementa, u formi pomerajućeg prozora, i brojanju ponavljanja boja unutar strukturnog elementa. Na ovaj način mogu se razlikovati slike koje imaju istu količinu piksela raspoređenih po bojama ali su ti pikseli različito prostorno grupisani. Vrednosti boje svakog piksela reprezentovane su u dvokonusnom *Hue Max Min Diff (HMMD)* kolor prostoru koji je neuniformno kvantovan na 32, 64, 128 ili 256 binova. Amplituda svakog bina reprezentovana je osmobicnim kodom. U odnosu na običan kolor histogram, strukturno kolor obeležje za prirodne slike obezbeđuje dodatnu funkcionalnost i unapređenje efikasnosti pretraživanja slika na osnovu sličnosti.

3.2.2.7. Grupno kolor obeležje

Grupno kolor obeležje izražava skalabilno kolor obeležje definisano za mirnu sliku, za kolor opis video segmenta ili kolekcije mirnih slika. Dodatni bitovi se koriste za definisanje načina na koji je izračunat ekvivalentni kolor histogram pre *Haar* transformacije, primenom srednje vrednosti, medijane ili presekom histograma. Srednji histogram dobijen je usrednjavanjem vrednosti svakog bina histograma za sve slike ili

frejmove. Ovaj pristup je ekvivalentan izračunavanju agregatnog kolor histograma svih frejmova ili slika, sa odgovarajućom normalizacijom. Medjian histogram predstavlja izračunavanje medjian srednje vrednosti nekog bina histograma za sve slike ili frejmove. Ovaj pristup je otporniji na greške zaokruživanja u odnosu na srednji histogram. Presečni histogram predstavlja izračunavanje minimalne vrednosti nekog stupca histograma za sve slike ili frejmove. Na ovaj način detektuje se najmanja zajednička vrednost neke boje za sve slike ili frejmove. Za poređenje GoF/GoP obeležja mogu se koristiti sve mere sličnosti/rastojanja koje se koriste za poređenje skalabilnog kolor obeležja.

3.2.3. Obeležja teksture

Tekstura predstavlja osobinu svake površine a manifestuje se kao vizuelni obrazac koji je posledica prisustva različitih boja i intenziteta na slici. Vizuelni obrazac koji reprezentuje teksturu može ali i ne mora imati osobinu homogenosti. Tekstura sadrži važne strukturne informacije o površinama i njihovom povezanošću sa sredinom koja ih okružuje. Opisom teksture na slikama primenom teksturnih obeležja obezbeđuje se vrlo važno razlikovanje vizuelnih struktura u procesu pretraživanja na osnovu sličnosti. U procesu pretraživanja obeležja teksture mogu se koristiti nezavisno ili u kombinaciji sa ostalim *MPEG-7* obeležjima. U okviru *MPEG-7* standarda definisana su tri obeležja teksture: obeležje homogene teksture (*homogeneous texture descriptor*), obeležje histogram ivica (*edge histogram descriptor*) i obeležje za teksturno pretraživanje (*texture browsing descriptor*). U nastavku dat je kratak opis svakog od navedenih obeležja.

3.2.3.1. Obeležje homogene teksture

S obzirom da sliku možemo posmatrati kao mozaik homogenih tekstura, vrlo značajnu ulogu u opisivanju teksture ima obeležje homogene teksture. Ovo obeležje opisuje usmerenost, finoću i regularnost teksture, čime se obezbeđuje kvantitativna karakterizacija homogene teksture. Opisivanje teksture u slučaju ovog obeležja zasniva se na primeni

banke *Gabor* filtra, koji predstavljaju modelovane *Gabor* funkcije podešene za odgovarajuću skalu i orijentaciju. Komponente teksturnog obeležja čine prvi i drugi moment energije odgovarajućih frekvencijskih podopsega, u frekvencijskom domenu. Ukupno se koristi 30 filtera, gde broj filtera predstavlja umnožak broja skala (5 skala) i broja orijentacija (6 orijentacija) koji se koriste u multirezolucionoju dekompoziciji primenom *Gabor* funkcija. Izračunavanje ovog obeležja bazira se na filtriranju slike selektivnim kernelima korišćenjem različitih skala i orijentacija. Obeležje homogene teksture koristi se za poređenje teksturnih slika.

3.2.3.2. Obeležje histogram ivica

Histogram ivica obeležje opisuje prostornu raspodelu ivica i predstavlja dobar teksturni pokazatelj, vrlo koristan u slučajevima poređenja tekstura koje nisu homogene. Izračunavanje komponenata ovog obeležja započinje deljenjem slike na regione (4×4), gde se za svaki od regiona izračunava lokalni histogram ivica. Ivice su grupisane u pet kategorija: vertikalne, horizontalne, dijagonalne ivice pod uglom 45 stepeni, dijagonalne ivice pod uglom 135 stepeni i izotropne (neorijentisane) ivice. Lokalni histogrami ivica regiona imaju pet stupaca u skladu sa definisanim kategorijama ivica. Za proizvoljnu sliku izdeljenu na 16 regiona histogram ivica obeležje sadrži 80 binova. Svaki od binova se nelinearno kvantuje korišćenjem tri bita. Detekcija orijentacije ivice zasnovana je na primeni jednostavnih filtera za detekciju orijentacije dimenzije 2×2 , a izvršava se na makroblokovima regiona sa dimenzijama koje su stepen broja dva. Makroblok se tretira kao slika dimenzije 2×2 piksela tako što svaka vrednost piksela ove slike predstavlja srednju vrednost intenziteta piksela u odgovarajućem regionu makrobloka koji ekvivalentira taj piksel. Na 2×2 sliku makrobloka primenjuju se filteri za detekciju orijentacije, svakom bloku se dodeljuje orijentacija i vrši se ažuriranje histograma. S obzirom da ivice imaju veliku ulogu u percipiranju slike, pretraživanje baza slika primenom ovog obeležja može kao rezultat izdvojiti slike sličnog semantičkog značenja. Histogram ivica obeležje je vrlo efikasno pri reprezentaciji prirodnih slika, a primarno je namenjen poređenju slika. Dodatno unapređenje performansi pri poređenju slika može se postići

kombinovanjem ovog obeležja sa ostalim obeležjima slika, npr. kolor obeležjima. Pored kolor obeležja, histogram ivica obeležja mogu se koristiti za detekciju promena scene i ključnih frejmova u videu. Nedostatak ovog obeležja je, što se za razliku od obeležja homogene teksture ne može primeniti u pretraživanju na osnovu objekata.

3.2.3.3. Obeležje za teksturno pretraživanje

Obeležje za teksturno pretraživanje predstavlja vrlo kompaktno obeležje teksture koje je semantički vrlo blisko ljudskoj percepciji teksture, jer karakterizuje regularnost, orijentaciju i finoću teksture u određenim skalama. Tekstura se može opisati ovim obeležjem u formatu bitskog niza dužine do 12 bitova, regularnost teksture (2 bita), usmerenost (3 bita \times 2) i finoća teksture (2 bita \times 2). S obzirom da tekstura može imati više od dva dominantna pravca ili skale, u okviru obeležja moguće je definisati maksimalno dve različite vrednosti orijentacije i skale finoće teksture. Izdvajanje ovog obeležja zasniva se na filtriranju slike primenom filtra propusnog opsega podešenih za odgovarajuće skale i orijentacije. Proces filtriranja je sličan filtriranju u slučaju obeležja homogene teksture. Rezultat filtriranja koristi se za izračunavanje komponenta obeležja za teksturno pretraživanje. Ovo obeležje može se koristiti u raznim primenama pretraživanja na osnovu teksture, a u kombinaciji sa obeležjem homogene teksture predstavlja skalabilno rešenje za predstavljanje homogene teksture u regionima na slici.

3.2.4. Obeležja oblika

Obeležja oblika obezbeđuju vrlo značajan alat za karakterizaciju identiteta i funkcionalnosti objekata, koji se može iskoristiti i za prepoznavanje objekata. Ljudski vizuelni sistem može prepoznati karakteristične objekte samo na osnovu njihovog oblika, što je dokaz da oblik često u sebi sadrži semantičku informaciju. Ova činjenica razlikuje oblik od ostalih elementarnih obeležja, kao što su boja, pokret ili tekstura. Ova elementarna obeležja su veoma važna ali uglavnom ne otkrivaju identitet objekta. Obeležje oblika ima

primenu u raznim aplikacijama pretraživanja na osnovu vizuelnih objekata. U okviru *MPEG-7* standarda definisana su četiri obeležja oblika: obeležje oblika regiona (*region shape descriptor*), obeležje oblika konture (*contour shape descriptor*), spektralno 3D obeležje oblika (*shape spectrum - 3D descriptor*) i *2D/3D* obeležje oblika (*2D/3D shape descriptor*). U nastavku teksta dat je kratak pregled *MPEG-7* obeležja oblika.

3.2.4.1. Obeležje oblika regiona

Obeležje oblika regiona opisuje distribuciju piksela unutar regiona *2D* objekta. Ovo obeležje oblika može opisati pored jednostavnih objekata sa ili bez unutrašnjih praznina, i kompleksne objekte koji se sastoje od višestrukih razdvojenih regiona. Razdvojenost regiona nekog objekta može biti posledica procesa segmentacije. Ovo obeležje obezbeđuje kompaktan i efikasan način za simultano opisivanje karakteristika razdvojenih regiona, tako što se pri izračunavanju komponenta obeležja koristi informacija o lokaciji svih regiona koji čine objekat. Obeležje je robustno na šumove koji su posledica procesa segmentacije, impulsni šum *salt&pepper* i minorne deformacije duž granica objekata. Obeležje oblika regiona spada u široku klasu tehnika za analizu oblika koje se zasnivaju na momentima [Sikora01]. Pri izračunavanju se koriste kompleksnu *2D* ugaono-radijalnu transformacija definisanu na jediničnom disku u polarnom koordinatnom sistemu.

3.2.4.2. Obeležje oblika konture

Obeležje oblika konture opisuje karakteristiku oblika objekta ili regiona na slici na osnovu njihove konture. U procesu izdvajanja ovog obeležja koristi se *Curvature Scale-Space (CSS)* reprezentacija konture koja reprezentuje značajne karakteristike oblika. *CSS* reprezentacija korišćena je i ranije vrlo uspešno za pretraživanje i prepoznavanje [SIK01], dok je u fazi razvoja *MPEG-7* standarda proširena i optimizovana. Karakteristike oblika izdvojene na osnovu konture vrlo su značajne za prepoznavanje semantički sličnih objekata. U slučajevima kada oblici imaju slične karakteristike na nivou regiona, obeležje oblika konture može ukazati na razliku između posmatranih oblika. Pretraživanje oblika na osnovu ovog obeležja semantički je vrlo slično ljudskom razlikovanju oblika, čak i kada

postoji značajna varijabilnost unutar odgovarajuće klase. Obeležje je robusno na značajne deformacije konture (rigidne i nerigidne) i na transformacije perspektive koje su uobičajene na slikama i videu. Obeležje se sastoji od ekscentričnih i cirkularnih vrednosti originalne i filtrirane konture (po 6 bitova svaka), indeksa koji ukazuje na broj “pikova” u CSS slici (6 bitova), visine najvećeg “pika” (7 bitova) i koordinatne pozicije preostalih pikova (9 bitova po “piku”). Obeležje oblika konture je vrlo kompaktno obeležje, a srednja dužina ovog obeležja je 112 bitova.

3.2.4.3. Spektralno 3D obeležje oblika

Spektralno 3D obeležje oblika zasniva se na spektru oblika. Ovaj pristup predstavlja proširenje koncepta koji se zasnivao na indeksu oblika, a ranije se koristio kao lokalna mera 3D oblika i 3D mreža. Indeks oblika neke prilagođene 3D površine u nekoj tački definiše se na osnovu primarnih zakrivljenosti u toj tački. Indeks oblika nije definisan za planarne regione. U principu, indeks oblika predstavlja informaciju o lokalnoj konveksnosti 3D površine. Spektar oblika definiše se kao histogram indeksa oblika izračunatog na celoj površini 3D objekta. Za 3D mreže indeks oblika se izračunava za svaki čvor mreže. Spektar je invarijantan na skaliranje i Euklidske transformacije. Broj binova u histogramu spektra je predefinisano sa vrednošću 100, a svaki bin se reprezentuje sa 12 bitova. U ovom obeležju koriste se dodatne dve promenljive. Prva promenljiva izražava odnos relativne površine planarnih površinskih regiona i ukupne površine 3D mreže. Druga promenljiva predstavlja odnos relativne površine svih poligonalnih elemenata gde se ne može pouzdano izračunati indeks oblika i ukupne površine 3D mreže.

3.2.4.4. 2D/3D obeležje oblika

2D/3D obeležje oblika predstavlja kombinaciju 2D obeležja za opisivanje vizuelnih karakteristika 3D objekta, viđenih iz različitih uglova posmatranja. Kombinovana obeležja formiraju kompletnu reprezentaciju objekta u 3D prostoru. Pri reprezentaciji mogu se koristiti proizvoljna 2D obeležja boje, teksture ili oblika. Za opisivanje karakteristika realnih 3D objekata, ovo obeležje podržava integraciju 2D obeležja korišćenih u ravni slike.

3.2.5. Obeležja pokreta

Sva MPEG-7 obeležja za boju, teksturu i oblik objekata, a koja su opisana u prethodnom delu teksta mogu se direktno primeniti u procesu indeksiranja slika u video sekvenci. Obeležja boje i teksture često se koriste za vremensku segmentaciju videa u smislu detekcije naglih promena kadrova, lokalizacije *dissolve* efekata kamere itd. Opis karakteristika pokreta u video sekvenci može obezbediti još egzaktniju reprezentaciju sadržaja. Generalno, opisivanje pokreta u video sekvencama upotrebom polja pokreta može biti vrlo zahtevno po pitanju neophodne količine bitova za reprezentaciju. U okviru *MPEG-7 Visual* standarda razvijena su obeležja koja izražavaju vrlo efikasno i precizno osnovne karakteristike iz polja pokreta. *MPEG-7 Visual* standard definiše četiri obeležja pokreta: obeležje pokreta kamere (*camera motion descriptor*), obeležje trajektorije pokreta (*motion trajectory descriptor*), obeležje parametarskog kretanja (*parametric motion descriptor*) i obeležje aktivnosti pokreta (*motion activity descriptor*). U nastavku teksta dat je kratak opis navedenih obeležja pokreta.

3.2.5.1. Obeležje pokreta kamere

Obeležje pokreta kamere opisuje pokret kamere ili virtuelne tačke posmatranja u sceni. Ovo obeležje pruža detaljne informacije o tome, koji su tipovi globalnih parametara pokreta prisutni i u kojem vremenskom trenutku unutar scene. Vrednosti ovih parametara, u nekim aplikacijama, mogu se dobiti direktno sa kamere. Takođe, ove parametre kamere moguće je odrediti iz intenziteta piksela na slici korišćenjem odgovarajućeg modela kamere i tehnika za estimaciju pokreta. Ovo obeležje može se koristiti u pretraživanju video sekvence na osnovu određenog globalnog parametra pokreta, npr. traženje scena sa velikom aktivnošću zuma (*zoom*) ili traženje scena sa pretežno translatorskim kretanjem. Pored navedenog moguće je pretraživati poklapanje sličnosti pokreta u određenim vremenskim intervalima, npr. traženje video sekvence gde se translatorsno pomeranje kamere pojavljuje u intervalu između tridesete i četrdesete sekunde. Ovo obeležje karakteriše parametre 3D pomeraja kamere. Zasnovano je na informacijama o parametrima 3D pomeraja kamere, koje se mogu automatski dobiti ili generisati sa uređaja za akviziciju.

Obeležje pokreta kamere podržava sledeće osnovne pokrete kamere: statična kamera (*fixed*), horizontalna rotacija (*panning*), horizontalno transversalno kretanje, *traveling* u filmskoj industriji (*tracking*), vertikalna rotacija (*tilting*), vertikalno transversalno kretanje (*booming*), promena žižne daljine (*zooming*), translacija duž optičke ose (*dollying*) i rotacija oko optičke ose (*rolling*).

Podkadrovi čiji su svi frejmovi okarakterisani nekim određenim tipom pokreta kamere, ili mešanim tipovima pokreta kamere, predstavljaju osnovne blokove obeležja pokreta kamere. Svaki od osnovnih blokova opisan je vremenom početka, dužinom trajanja, brzinom indukovanog pokreta slike, vremenskom frakcijom njegovog trajanja u poređenju sa zadatom veličinom vremenskog prozora, fokusom ekspanzije (*Focus of Expansion, FOE*) i fokusom kontrakcije (*Focus of Contraction, FOC*). Obeležje predstavlja uniju osnovnih blokova i može opciono opisivati mešavinu različitih pokreta kamere. Kombinovani mod (*mixture mode*) snima globalnu informaciju o parametrima pokreta kamere paralelno opisujući višestruke tipove pokreta, čak kada se ti tipovi pokreta pojavljuju simultano. Sa druge strane nekombinovani mod (*non-mixture mode*) snima pojavu pokreta pojedinačnog tipa i njihovu uniju unutar određenog vremenskog intervala. Simultana pojava višestrukih tipova pokreta opisuje se kao unija opisa pojedinačnih tipova pokreta. U ovom modu opisivanja pokreta kamere, vremenski prozor jednog elementarnog segmenta može se preklapati sa vremenskim prozorom drugog elementarnog segmenta.

3.2.5.2. Obeležje trajektorije pokreta

Kretanje svakog pojedinačnog objekta u video sekvenci može se opisati primenom obeležja trajektorije pokreta. Ovo obeležje opisuje detalje pomeranja objekata tokom vremena, opisivanjem trajektorije kretanja objekta. Omogućeno je poređenje trajektorija objekata za video pretraživanja, hiperlinkovanje videa i slične aplikacije. Primer primene ovog obeležja je primena u video nadzoru saobraćaja gde se kretanje vozila i pešaka može opisati i memorisati primenom *MPEG-7* obeležja trajektorije pokreta. Upiti se mogu realizovati na osnovu obeležja u cilju obezbeđivanja brzog pristupa scenama sa zahtevanim

kretanjem objekata. Mogući upiti mogu se definisati kao “prolazak objekta blizu određene oblasti” ili “objekti se kreću brzinom većom od dozvoljene”.

Obeležje trajektorije pokreta nekog objekta je jednostavno obeležje visokog nivoa, definisano kao lokalizacija neke reprezentativne tačke objekta u vremenu i prostoru. Ovo obeležje nalazi svoju punu primenu u pretraživanju na osnovu sadržaja u objektno orijentisanim bazama vizuelnih podataka. Takođe, može biti od pomoći u vrlo specifičnim primenama. U datom kontekstu, apriori znanje o trajektoriji omogućava mnoge funkcionalnosti. U video nadzoru, alarm se može aktivirati ako neki objekat ima trajektoriju koja se može okarakterisati opasnom (prolazak kroz zabranjenu zonu, kretanje neuobičajenom brzinom, itd.). U sportskim video zapisima mogu se prepoznati specifične akcije, npr. približavanje mreži učesnika u teniskom meču. Pored navedenog, ovakav pristup u opisivanju dozvoljava unapređenje interakcije sa podacima, odnosno manipulacije pri poluautomatskom multimedijском editovanju. Trajektorije mogu biti izmenjene u cilju prilagođenja kretanja objekta zadatom globalnom kontekstu video sekvence. Ovo obeležje u suštini predstavlja listu ključnih tačaka zajedno sa skupom opcionih interpolacionih funkcija koje opisuju putanju između ključnih tačaka u smislu ubrzanja. Brzina je implicitno zadata specifikacijom ključnih tačaka. Ključne tačke su specificirane vremenskim trenutkom i koordinatama u $2D$ i $3D$ prostoru, u zavisnosti od aplikacione namene. Interpolacione funkcije su definisane nezavisno za svaku od koordinata.

3.2.5.3. Obeležje parametarskog kretanja

Jedan od načina da se zabeleži kretanje na sceni je primena obeležja koje se zasniva na obeležju *Warping Parameter*. Globalno kretanje je izraženo u odnosu na globalne *sprite* ili *mosaic* video scene. Modeli parametarskog kretanja primenjivani su u velikom broju međusobno povezanih oblasti obrade i analize slika, uključujući estimaciju pokreta i segmentaciju baziranu na pokretu, estimaciju globalnog kretanja, kreiranje video mozaika i praćenje objekata. Modeli parametarskog kretanja već su korišteni u *MPEG-4* standardu, za estimaciju i kompenzaciju globalnog kretanja, i generisanje *sprite* slike. U okviru *MPEG-7 Visual* standarda pokret je vrlo bitno obeležje, povezano sa vremensko-

prostornom strukturom videa u smislu nekoliko specifičnih *MPEG-7* primena kao što su pamćenje i pretraga videa u bazama podataka, kao i hiperlinkovanje. Pored navedenog, pokret je ključno obeležje u vrlo specifičnoj primeni za jezično indeksiranje znakova. Osnovni princip ovog obeležja se sastoji od opisa kretanja objekata u video sekvenci pomoću *2D* parametarskih modela. Afini modeli uključuju translaciju, rotaciju, skaliranje i kombinaciju ovih transformacija, zatim planarni modeli perspektive omogućavaju uračunavanje globalnih deformacija povezanih sa projekcijama perspektive i kvadratni modeli omogućavaju opisivanje još kompleksnijih kretanja.

Parametarski model je povezan sa proizvoljnim objektima (prednja ravan ili pozadinski), definisanim kao regioni (grupa piksela) u slici unutar odgovarajućeg vremenskog intervalu. Na ovaj način pokret objekta je snimljen u kompaktnom opisu koji predstavlja skup svega nekoliko parametara. Na ovaj način omogućeno je opisivanje više tipova pokreta, jednostavnih kao što su translacija, rotacija i zumiranje, kao i vrlo kompleksnih pokreta koji predstavljaju kombinaciju već pomenutih elementarnih pokreta. Efikasno pretraživanje objekata na osnovu pokreta zahteva definisanje odgovarajuće mere sličnosti između modela kretanja. Takođe, neophodno je podržati pretragu niskog nivoa na osnovu primera scenarija kretanja i pretragu visokog nivoa na osnovu konteksta, kao na primer “traženje objekata koji prilaze kameri”, “traženje objekata koji rotiraju”, ili “traženje objekata koji se kreću transliranjem u levo” itd.

3.2.5.4. Obeležje aktivnosti pokreta

Obeležje aktivnosti pokreta snima ukupan nivo aktivnosti ili tempo pokreta, odnosno akcije u sceni unutar video segmenta (date grupe frejmova). Ovo obeležje opisuje gde je scena doživljena od strane posmatrača kao spora, sa brzim tempom ili sa brзом akcijom. Primeri scena sa visokim nivoom aktivnosti su, na primer, scene sportskih mečeva fudbala ili košarke, dok sa druge strane scene niskog nivoa aktivnosti su na primer kadrovi vesti kada je osoba koja čita vesti statična, ili kada su kadrovi statični. Obeležje intenziteta aktivnosti pokreta omogućava odabir video segmenata ili programa koji su kategorisani na osnovu ovog atributa. Princip izdvajanja ovog obeležja zasniva se na merenju intenziteta

pokreta na osnovu standardne devijacije amplitude vektora pokreta. Standardna devijacija je kvantovana na pet nivoa aktivnosti. Opciono, postoje pravac kretanja i prostorna raspodela aktivnosti pokreta. Takođe, vremenska raspodela aktivnosti pokreta može se izdvojiti kao opis i iskoristiti za poređenje video segmenata.

Video sadržaj generalno zahvata skalu od visoke do niske aktivnosti, i zato je neophodno obeležje koje omogućava precizno izražavanje aktivnosti date video sekvence (kadra), i koje sveobuhvatno pokriva prethodno pomenutu skalu. Obeležje aktivnosti pokreta je pogodno za aplikacije kao što su prenamena videa, video nadzor, brzo pregledanje, dinamička video sumarijacija, pretraživanje na osnovu sadržaja itd. Na primer, prezentaciona brzina prikazivanja frejmova može se smanjiti ako obeležje aktivnosti kretanja ukazuje na visoku aktivnost tako da takav sadržaj bude gledljiv za posmatrača. Drugi primer primene je pronalaženje svih kadrova sa visokim stepenom aktivnosti pokreta u video programu sa vestima koje se mogu videti pregledanjem i apstrakcijom.

3.2.6. Obeležja lokalizacije

Opisni alati za lokalizaciju koriste se za indikaciju regiona od interesa, proizvoljnog oblika, u prostornom i prostorno-vremenskom domenu. U okviru *MPEG-7* standarda definisana su dva obeležja za lokalizaciju: obeležje lokalizacije regiona (*region locator descriptor*) i obeležje prostorno-vremenske lokalizacije (*spatio-temporal locator descriptor*) obeležje. U nastavku teksta dat je kratak opis navedenih obeležja lokalizacije.

3.2.6.1. Obeležje lokalizacije regiona

Obeležje lokalizacije regiona omogućava lokalizaciju regiona unutar slika ili frejmova, obeležavanjem regiona kratkom i skalabilnom reprezentacijom pravougaonika ili poligona.

3.2.6.2. Obeležje prostorno-vremenske lokalizacije

Obeležje prostorno-vremenske lokalizacije opisuje prostorno-vremenske regione u video sekvenci, kao što su regioni pokretnih objekata, i obezbeđuje lokalizacionu funkcionalnost. Glavna primena ovog obeležja je u hipermediji koja prikazuje neophodne informacije u slučaju kada se označena tačka nalazi unutar objekta. Sledeća bitna primena je u pretraživanju objekata, na primer u video nadzoru, kada je moguće izvršiti proveru kada je objekat prošao određene tačke u prostoru. Obeležje prostorno-vremenske lokalizacije može opisati prostorno povezane i nepovezane regione.

3.2.7. Obeležja za identifikaciju lica

Lice predstavlja jednu od najvažnijih semantika ugrađenih u vizuelnom kontekstu. MPEG-7 standardom specificirana su dva alata za opisivanje identiteta lica: obeležje za prepoznavanje lica (*face recognition descriptor*) i obeležje za napredno prepoznavanje lica (*advanced face recognition descriptor*). U nastavku teksta date su osnovne informacije vezane za ova dva obeležja.

3.2.7.1. Obeležje za prepoznavanje lica

Ovo obeležje se koristi za pretraživanje baze slika lica na osnovu upitne slike lica. Izdvajanje obeležja se sastoji iz nekoliko koraka. Prvi korak je normalizacija slike lica na dimenziju 56x46 piksela, poravnate tako da su centri oba oka locirana na pozicijama piksela (24,16) i (24,31). Kao obeležje koristi se projekcija vektora lica na skup od 48 jediničnih vektora koji predstavljaju osnovu prostora mogućih vektora lica.

3.2.7.2. Obeležje za napredno prepoznavanje lica

Obeležje za napredno prepoznavanje lica je robusno na varijacije u pozama lica na slici i na različite uslove osvetljenosti na slici.

3.3. Multifraktalna analiza

Multifraktalna analiza (*MA*) primenjuje se u analizi signala kao pogodan alat za opisivanje neregularnih pojava u vremensko/prostornoj raspodeli signala. Visoko neuniformna raspodela verovatnoće signala koji se posmatra, koja proizilazi iz neuniformnosti sistema, često poseduje veliki broj skalabilnih osobina uključujući i samosličnost [MAN83]. Ova činjenica predstavlja osnovu multifraktalnog formalizma i omogućava da se primenom *MA* na odgovarajući način opišu lokalne i globalne neregularnosti posmatranih veličina ili funkcija, geometrijski ili statistički [EVE96]. Multifraktalna analiza se može primeniti u proučavanju dugotrajne zavisnosti (*long range dependency*) dinamike ponašanje nekih signala, u analizi skalabilnih osobina širokih neuniformnih statističkih raspodela verovatnoće i karakterizaciji fraktalnih osobina rezultata merenja različitih pojava. U nastavku teksta u sekciji 3.3.1 prikazan je opšti formalizam *MA*, a u sekciji 3.3.2 predstavljena je primena *MA* u detekciji određenih lokalnih i globalnih karakteristika signala.

3.3.1. Multifraktalni formalizam

Koncept fraktala uveden je od strane *Benoa Mandelbrot*-a sedamdesetih godina prošlog veka. Termin dolazi od latinske reči *fractus* što znači neregularno, isprekidano (*discontinuous*). Sama reč je izvedena iz glagola *frangere* što znači lomiti, razlagati na delove. Fraktali se grubo mogu definisati kao geometrijski oblici koji poseduju osobinu skaliranja. U suštini, kako i jedna od definicija fraktala kaže, fraktal predstavlja skup objekata koji su slični celini. Osnovna karakteristika fraktalnog objekta je njegova fraktalna dimenzija [MAN89, FAL03] koja opisuje kako skup objekata koji čine fraktal popunjava prostor.

Fraktalna dimenzija može se izračunati na različite načine. Jedan od uobičajenih načina je primena *box-counting* metode. Definisan je skup Q u prostoru B^n . Prostor B^n izdeljen je na n -dimenzionalne zapremine sa radijusom ε , a kao mera popunjenosti

prostora definisana je veličina $N(\varepsilon)$ koja predstavlja broj zapremina koje zauzimaju skup Q . Za definisani slučaj, vrednost *Hausdorff – Besicovitch* fraktalne dimenzije [FAL03], izračunate primenom *box-counting* metode, zadovoljava uslove definisane izrazom (3.1):

$$\lim_{\varepsilon \rightarrow 0} N(\varepsilon) \varepsilon^d = \begin{cases} 0, & d > D \\ \infty, & d < D \end{cases} \quad (3.1)$$

Ovakva definicija fraktalne dimenzije se može još pojednostaviti u oblik koji je pogodniji za praktičnu upotrebu. Za $\varepsilon \approx 0$ važi izraz (3.2) za fraktalnu dimenziju D ,

$$D \approx -\ln N(\varepsilon) / \ln \varepsilon \quad (3.2)$$

odnosno izraz (3.3),

$$D = \lim_{\varepsilon \rightarrow 0} \frac{\ln N(\varepsilon)}{\ln \left(\frac{1}{\varepsilon} \right)} \quad (3.3)$$

Fraktalna dimenzija, po pravilu, predstavlja necelobrojnu vrednost koja na određen način opisuje geometrijsku složenost objekta.

Realni fizički objekti i signali koji poseduju osobinu samosličnosti, vrlo retko se mogu opisati samo jednom vrednošću fraktalne dimenzije. Iz navedenog razloga za analizu nehomogenih fraktalnih objekata primenjuje se analiza zasnovana na multifraktalima. Multifraktal predstavlja skup sastavljen od mnoštva isprepletenih podskupova različitih fraktalnih dimenzija. Svaka tačka skupa je povezana sa merom ili funkcijom koja je tipično nenegativna realna vrednost, često i normalizovana. Podaci u multifraktalnom skupu često ne pokazuju osobinu samosličnosti jer se sastoje od velikog broja fraktalnih podskupova. Ideja primene multifraktalne analize jeste da se analizira dekompozicija procesa složene statistike na beskonačnom spektru homogenih fraktala (monofraktala) sa jasno izraženom fraktalnom dimenzijom.

Statističke osobine singularnih rezultata merenja ili funkcija mogu se opisati multifraktalnim formalizmom globalno u vidu njihovog spektra singularnosti (*Legendre*

spektra) ili lokalno u vidu njihove generalisane fraktalne dimenzije (*Rényi*) [FAL03]. Primenom *box-counting* metode mogu se izračunati multifraktalni parametri posmatranih veličina. Osnova posmatrane veličine se izdela na kvadrate dimenzija stranice ε centrirane u tački $(x, N_x(\varepsilon))$. S obzirom da gore prikazana standardna metoda za izračunavanje fraktalne dimenzije ne može da pokaže lokalnu neregularnost posmatrane veličine odnosno promenu broja zahvaćenih tačaka posmatrane veličine od tačke do tačke, u *box-counting* metod se uvodi mera μ koja treba da opiše varijaciju broja tačaka posmatrane veličine raspodeljenih po kvadratima. Prebrojavaju se samo kvadrati koji imaju određenu izmerenu vrednost $\mu(N_x(\varepsilon))$ i ova mera se može shvatiti kao verovatnoća da se vrednost merenja μ nalazi u kvadratu centriranom u tački x . Pokazuje se da varijacija izmerene vrednosti sa promenom dimenzije kvadrata odgovara takozvanom stepenom zakonu (*power law*) (3.4)

$$\mu(N_x(\varepsilon)) \approx \varepsilon^{\alpha(x)} \quad (3.4)$$

gde $\alpha(x)$ opisuje fraktalnost strukture, a u literaturi se može naći pod nazivom grubi Holderov eksponent. U graničnom procesu, kada $\varepsilon \rightarrow 0$, za datu meru μ , vrednosti eksponenta $\alpha(x)$ će varirati za različite veličine kvadrata, ali će na kraju konvergirati ka određenoj vrednosti za tačku x , izraz (5).

$$\alpha(x) = \lim_{\varepsilon \rightarrow 0} \frac{\mu(N_x(\varepsilon))}{\varepsilon} \quad (3.5)$$

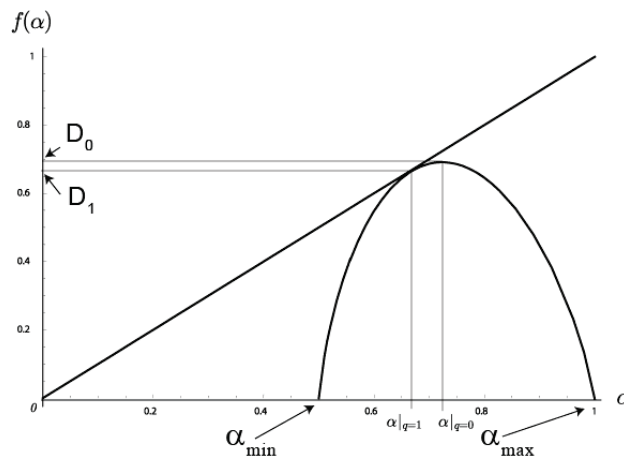
Kao što se može videti iz prethodne jednačine (3.5), Holderov eksponent zavisi od pozicije u strukturi posmatrane merene veličine i opisuje njenu lokalnu regularnost. Vrednost $\alpha(x)$ naći će se u većem broju kvadrata u strukturi (odnosno tačkama u graničnom procesu), tako da se može analizirati i raspodela vrednosti $\alpha(x)$ u strukturi. Ova raspodela dobila je ime multifraktalni spektar singulariteta $f(\alpha)$, koji daje opis globalne regularnosti strukture merene veličine. Spektar singulariteta povezuje u strukturi merene veličine sve tačke x čija lokalna fraktalna dimenzija ima vrednost α (3.6).

$$f(\alpha) = d_F(\{x \in A | \alpha(x) = \alpha\}) \quad (3.6)$$

Funkcija $f(\alpha)$ uobičajeno ima parabolično konkavni izgled [FAL03], Slika 3.21. Maksimalna vrednost krive je D i ona predstavlja fraktalnu dimenziju pozadine (*support*), a definisana je izrazom (3.7).

$$\max_{\alpha} f(\alpha) = D \quad (3.7)$$

Multifraktalni spektar singulariteta monofraktalne veličine predstavljen je jednom tačkom u spektru $f(\alpha) = \alpha = D$. Multifraktalni spektar singulariteta multifraktalne veličine je funkcija sa odgovarajućom širinom.



Slika 3.21. Multifraktalni spektar singulariteta

Spektar singulariteta statistički opisuje distribuciju α eksponenata u strukturi kojom je merena veličina ekvivalentno predstavljena pomoću α vrednosti. Primenom *box-counting* metoda možemo odrediti multifraktalni spektar ove strukture tako što njenu osnovu prekrijemo kvadratima dimenzije stranica ε . Definiše se mera $N_{\alpha}(\varepsilon)$ koja predstavlja verovatnoću da se merena vrednost intenziteta singulariteta nalazi u intervalu $(\alpha, \alpha + d\alpha)$. Prebrojavaju se kvadrati sa varijacijom ε^{α} za dato α i pokazuje se da za merenu veličinu važi stepeni zakon (3.8).

$$N_{\alpha}(\varepsilon) \sim \varepsilon^{-f(\alpha)} \quad (3.8)$$

Ako se malo detaljnije analizira funkcija $f(\alpha)$, ova funkcija u stvari predstavlja fraktalnu dimenziju strukture karakterisane sa vrednošću α . Dakle, ako bismo meru $N_{\alpha}(\varepsilon)$ predstavili kao histogram, funkcija $f(\alpha)$ opisuje način na koji se menja $N_{\alpha}(\varepsilon)$ u graničnom procesu kada $\varepsilon \rightarrow 0$.

Spektar singulariteta nije jednostavno odrediti u prezentovanoj formi. Iz praktičnih razloga izračunava se *Legendre*-ova transformacija $[\tau(q), q]$, koja služi za izračunavanje spektra singulariteta. *Legendre*-ova transformacija predstavljena je izrazima (3.9) i (3.10) [FAL03].

$$f(\alpha) = q\alpha - \tau(q) \quad (3.9)$$

$$\alpha = \frac{d\tau(q)}{dq} \quad (3.10)$$

U slučaju monofraktala važi $\alpha = f(\alpha) = D$, kada se izvrši smena u izrazu (3.9) *Legendre*-ove transformacije dobija se izraz $\tau(q) = D(q-1)$. Na osnovu ovog izraza definiše se generalisana fraktalna dimenzija (*Renyi dimension*) reda q , izraz (3.11).

$$D_q = \frac{\tau(q)}{q-1} \quad (3.11)$$

Za vrednost $q=0$, D_0 se naziva kapacitivnom dimenzijom koja odgovara Hausdorfovoj fraktalnoj dimenziji. Dimenzija $D_1 (q=1)$ odgovara informacionoj dimenziji, a dimenzija $D_2 (q=2)$ korelacionoj dimenziji. Generalno, dimenzija D_q odgovara skalirajućim eksponentima q -tih momenata neke mere μ merene veličine.

Izračunavanje generalizovane fraktalne dimenzije se opet može opisati *box-counting* metodom. Prekrijemo osnovu mere μ kvadratima dimenzije $(\varepsilon, B_i(\varepsilon))$. Particiona

funkcija $Z(q, \varepsilon)$ se može definisati izrazom (3.12), gde $N(\varepsilon)$ predstavlja ukupan broj kvadrata kojima je prekrivena mera, a promenljiva μ_i je definisana sa $\mu_i = \mu(B_i(\varepsilon))$.

$$Z(q, \varepsilon) = \sum_i^{N(\varepsilon)} \mu_i^q(\varepsilon) \quad (3.12)$$

Ako je mera μ multifraktalna, u graničnom procesu kada $\varepsilon \rightarrow 0$ particiona funkcija $Z(q, \varepsilon)$ se ponaša u skladu sa stepenim zakonom (3.13).

$$Z(q, \varepsilon) \sim \varepsilon^{\tau(q)} \quad (3.13)$$

Povezivanjem funkcije $\tau(q)$ i particione funkcije $Z(q, \varepsilon)$ dobija se novi izraz za generalisanu fraktalnu dimenziju (3.14).

$$D_q = \lim_{\varepsilon \rightarrow 0} \left(\frac{1}{q-1} \frac{\ln Z(q, \varepsilon)}{\ln \varepsilon} \right) \quad (3.14)$$

Generalisana fraktalna dimenzija obezbeđuje selektivnu karakterizaciju mere merene veličine. Vrednosti momenata q služe kao “mikroskop” [LOP09] za analizu različitih regiona neke merene veličine. Za vrednosti momenata $q > 1$, analiziraju se oblasti sa brzim promenama i D_q predstavlja oblast sa većom singularnošću. Za vrednosti momenata $q < 1$, analiziraju se oblasti sa sporijim promenama i D_q karakteriše manje singularnu oblast.

Postoji više metoda za izračunavanje multifraktalnog spektra singulariteta [LOP09]. Metode se mogu podeliti u dve grupe. Jednu grupu čine metode zasnovane na *box-counting* tehnici o kojoj je već bilo reči i analizi momenata mera merenih veličina, a drugu grupu čine metode zasnovane na *wavelet* transformaciji u kojima je prekrivanje osnove mere mrežom kvadrata zamenjeno sa skaliranjem *wavelet*-a. U ovoj disertaciji korišćena je metoda za izračunavanje spektra singulariteta zasnovana na histogramu [RELJ00a].

3.3.2. Primena multifraktalne analize u detekciji lokalnih/globalnih

karakteristika signala

Multifraktalna analiza primenjena je u velikom broju oblasti istraživanja kao alat za analizu lokalnih i globalnih karakteristika signala. Neki od primera su analiza klimatskih uticaja [RELJ00b, LIM09, SER10], segmentacija i klasifikacija teksture na slikama [ABA06, XIA06, EZE03], analiza širokog spektra signala u biologiji i medicini [POP05, KED02, HSU07, KOT03, IVA07, HE10], analiza medicinskih snimaka [RELJ02, YI07, SHI04, MAV06, MAN04], detekcija anomalija klasifikacijom fonokardiografskih signala [GAV13], karakterizacija oblika i modelovanje u dinamičkom evolucionom kontekstu [XIE10, BAC10], multifraktalna analiza finansijskih tržišta [CHE11, CAL08], mrežnog saobraćaja, analiza topologije i multimedijalnih signala [FUR11, WAN11, PAR00, RELJ09] i drugi.

Detekcija određenih fenomena u signalu primenom multifraktalne analize nije čest slučaj u literaturi, navešćemo nekoliko primera, detekcija promena u sekvencama slika [CAN96], analizom sekvence slika u domenu singulariteta α . Na osnovu razlike slika u različitim vremenskim trenucima u α domenu detektuje se razlika.

Naredni primer predstavlja detekciju mikrokalcifikacija primenom adaptirane multifraktalne analize [STO06]. U ovom radu predstavljena je primena adaptirane multifraktalne analize u detekciji mikrokalcifikacija na mamografskim snimcima koji predstavljaju lokalne singularitete. Ovi singulariteti su posebno naglašeni posmatranjem slike u α domenu primenom multifraktalne analize. Adaptacija metode se sastoji od primene dodatnih morfoloških operacija u procesu detekcije.

Detekcija ivica na slici primenom multifraktalne mere na nivou podpiksela (*Sub-Pixel Multi-fractal Measure, SPMM*) [WAN07]. Primenom *SPMM* tehnike dobija se tačnija vrednost eksponenta singulariteta za svaki piksel. Korišćenjem vrednosti singulariteta (α) i spektra singulariteta slike, slika se može segmentirati u niz skupova sa različitim vrednostima singulariteta omogućavajući tako automatsku detekciju ivica na slici.

Može se navesti i detekcija anomalija u mrežnom saobraćaju primenom multifraktalne analize [SHE11]. Pozicija graničnih parametara multifraktalnog spektra α_{\min} i α_{\max} iskorištena je za detekciju neregularne aktivnosti u mrežnom saobraćaju.

Multifraktalna analiza je primenjena i u akustici, za detekciju ranih refleksija [RIS13]. Autori su predstavili detekciju ranih refleksija u akustičkom impulsnom odzivu prostorija baziranu na multifraktalnoj analizi. Impulsni odziv je transformisan u α domen, gde je iskorišćena neregularnost lokalne strukture (singulariteti) za detekciju ranih refleksija.

4. Pregled publikovanih algoritama za detekciju naglih promena kadrova

U literaturi je predstavljen niz različitih algoritama za detekciju naglih promena kadrova. Sve prezentovane tehnike razlikuju se na osnovu obeležja koja opisuju vizuelni sadržaj frejmova, načina izračunavanja obeležja u prostornom i vremenskom domenu, principa detekcije i eventualne analize pokreta.

Različita obeležja se koriste za opisivanje vizuelnog sadržaja frejmova u detektorima naglih promena kadrova. Razlika piksela [LAW04] predstavlja jednostavnu metriku na osnovu koje se može izmeriti razlika između dva frejma, a izračunava se kao suma apsolutne razlike intenziteta (luminanse) ili boje odgovarajućih piksela. Ovo obeležje je osetljivo na promene osvetljenja i pokrete. Razlika histograma luminanse ili kolor histograma [GRE09, ABD08, PET04] predstavlja obeležje koje ima veoma nizak nivo kompleksnosti izračunavanja i najčešće sadrži manju količinu podataka od odgovarajućeg frejma. S obzirom da histogram ne odražava prostornu raspodelu piksela, osetljivost ovog

obeležja na pokrete je značajno manja u odnosu na ostala obeležja. Međutim, promena osvetljenja značajno utiče na histogram frejma.

Ivice, kao obeležje, korišćene su na više načina u detektorima naglih promena kadrova, prezentovanim u literaturi. Izdvajanje ivica je realizovano primenom *Canny Edge* detektora koji kao rezultat proizvodi mapu ivica. Postoje dva pristupa korišćenju mape ivica. Jedan pristup je izračunavanje *ECR*-a (*Edge Change Ratio*) [ZAB95] na osnovu ulaznih i izlaznih ivica dva uzastopna frejma ili izračunavanje razlika energija ivica na osnovu razlike koeficijenata *DCT* transformacije odgovarajućih blokova dva uzastopna frejma. Pokret značajno utiče na mapu ivica, a samim tim i celokupno obeležje ivica. Ovi algoritmi imaju veliku kompleksnost izračunavanja i njihove performanse ne prevazilaze performanse kolor histograma. Iz nadenih razloga ovi algoritmi se koriste samo u postobradi za uklanjanje pogrešno detektovanih promena kadrova kao *flash* detektori [CAO06, YUA07]. Drugi pristup koristi mapu ivica [LE08] svakog frejma za izračunavanje histograma orijentacije ivica za zadati broj orijentacija ili kao *MPEG-7 edge* histogram obeležje [MAN02, GRE09]. Na taj način je smanjena kompleksnost izračunavanja ovog obeležja. U [MOH12] iskorišćena je *scatter* matrica naglašenosti ivica, dobivena merenjem amplituda gradijenata u tačkama ivica između dva frejma, koja obezbeđuje reprezentaciju tranzicije intenziteta i kretanja.

Obeležje pokreta se može iskoristiti za detekciju naglih promena kadrova [LIU07] jer sama promena kadrova uzrokuje diskontinuitet kod ovog obeležja. Nedostaci u primeni ovog obeležja za detekciju zasnivaju se na tome što svako naglo pomeranje kamere ili objekta prouzrokuje diskontinuitet ovog obeležja. Ovi algoritmi su kompleksni za izračunavanja. Promene u statističkim momentima gore opisanih histograma mogu se iskoristiti za detekciju naglih promena kadrova. U literaturi, u svrhu detekcije naglih promena kadrova korišćeni su prvi i drugi statistički moment, srednja vrednost i varijansa [LIU07, AMI04].

MPEG statistika predstavlja statističku analizu informacija direktno izdvojenih iz MPEG strima koje se mogu iskoristiti za detekciju naglih promena kadrova. Kompleksnost

izračunavanja u ovakvom pristupu je značajno manja u odnosu na ostale pristupe jer video ne mora biti u potpunosti dekodovan. Primer ovakvog obeležja jeste upotreba *DC* koeficijenata [RAO90] intrakodovanih blokova, što odgovara upotrebi pododabiranih slika faktorom 8 po x i y osi. *DC* koeficijenti za ceo skup blokova slike su kodovani samo u I frejmu. Međutim u literaturi [SON01] je pokazano kako se *DC* koeficijenti mogu aproksimirati za interkodovane blokove u P i B frejmovima bez potpunog dekodovanja videa. Drugi primer upotrebe potpuno nedekodovanog videa u svrhu detekcije promena jeste slučaj kada se za potrebe detekcije nagle promene kadrova koristi analiza *MPEG GOP* (*Group Of Pictures*) u kojima se posmatraju tripleti frejmova [CAL02]. Dodatni primer predstavlja [REN09] gde se prvo iz MPEG makroblokova izdvoji nekoliko lokalnih indikatora, a nakon toga se primeni *Ada Boost* [KIM12] algoritam za selekciju i fuziju obeležja. Odabrana obeležja se koriste za klasifikaciju kandidata naglih promena kadrova. Globalni indikatori sličnosti frejmova, između graničnih frejmova i kandidata nagle promene kadra, analiziraju se primenom fazne korelacije *DC* slika.

Vejvlet obeležje je upotrebljeno u literaturi [LI09] za detekciju promena kadrova primenom diskretne vejvlet transformacije (*Discrete Wavelete Transformation, DWT*), dekompozicijom prvog nivoa na blokove frejmova. Frejm je podeljen na 4×4 regiona. Kolor obeležje je izvedeno na osnovu razlike koeficijenata *LL* dekompozicionih blokova odgovarajućih blokova frejmova. Obeležje ivica izvedeno je na osnovu razlike koeficijenata preostalih dekompozicionih blokova (*LH,HL,HH*) odgovarajućih blokova frejmova. Sledeći primer upotrebe obeležja kreiranih primenom diskretne vejvlet transformacije je [MIS13] gde se vrši ekstrakcija strukturnih obeležja iz svakog video frejma primenom *dual-tree complex* vejvlet transformacije.

U poslednje vreme pojavili su se algoritmi detekcije promena kadrova bazirani na promeni fraktalne dimenzije frejmova [ZEI10, GUN11]. Frejm je izdvojen na blokove i za svaki od blokova izračunata je fraktalna dimenzija, a za ceo frejm srednja fraktalna dimenzija. Detekcija naglih promena kadrova zasnovana je na pronalaženju diskontinuiteta sekvence razlika srednjih fraktalnih dimenzija uzastopnih frejmova. U [DON10] korišćeno je izdvajanje važnih obeležja iz video sekvence 3D merenjem. Obeležja se prate tokom

vremena u nameri da se odrede nagle promene kadrova. Jedan od novijih pristupa je [MOH12] u kojem se formuliše estimacija frejma korišćenjem prethodnog i narednog frejma. Parametri tranzicije između frejmova i greške, pri estimaciji frejmova na osnovu globalnih i lokalnih obeležja, koriste se za detekciju kadrova i njihovu klasifikaciju. Primena redukovano metričkog prostora za detekciju naglih promena kadrova prikazana je u [LU13]. Klasifikacijom i primenom adaptivnog praga odabira se segment kandidat koji sa određenom verovatnoćom sadrži naglu promenu kadrova. Za sve frejmove u ovom segmentu izdvoje se *HSV* kolor histogrami, koji kreiraju matricu obeležja frejmova. Primenjuje se *SVD* (*Singular Value Decomposition*) na ovu matricu i redukuje se dimenzija obeležja. Redukovana obeležja predstavljaju novi metrički prostor za detekciju nagle promene kadra.

Prethodno opisana obeležja korišćena u literaturi za detekciju naglih promena kadrova mogu se izračunati na tri različita načina: na nivou piksela, regiona ili celog frejma. Izdvajanje obeležja na nivou piksela [LIU07] obezbeđuje obeležja koja su osetljiva na promene sadržaja, sa vektorima obeležja velikih dimenzija. Ovako izračunata obeležja su veoma osetljiva na promenu osvetljenja i pokrete. Izdvajanje obeležja na nivou regiona [LE08, VOL04] podrazumeva deljenje frejma na pravougaone blokove jednakih dimenzija, regione koji ne smeju biti manjih dimenzija od 8×8 piksela. Ceo frejm ne mora biti pokriven blokovima. Povećanjem površine regiona smanjuje se osetljivost obeležja na pokrete objekata malih dimenzija, jer se vrši usrednjavanje vizuelnih karakteristika na velikoj površini. Ako se obeležje izdvaja na nivou frejma [ABD08, ROB04] umanjena je osetljivost na pokrete i osetljivost obeležja na promene sadržaja. Zahvaljujući ovome u literaturi postoje primeri korišćenja kombinacije izdvajanja obeležja [GRE09], gde je kolor histogram izdvojen na nivou regiona i celog frejma.

U vremenskom domenu algoritmi za detekciju naglih promena kadrova koriste analizu razlika obeležja na nivou dva uzastopna frejma, na nivou vremenskog prozora od N frejmova i na nivou tekućeg frejma. U slučaju izračunavanja razlike obeležja na nivou dva uzastopna frejma [GRE09, YUA07, MIS13] detekcija se svodi na pronalaženje velikih vrednosti koje predstavljaju promenu kadra. Ovakav način izračunavanja razlike obeležja

ima malu kompleksnost izračunavanja jer se analiziraju samo po dva frejma na svakoj poziciji. Primenom vremenskog prozora širine N frejmova, centriranog na određenoj poziciji, izračunavaju se razlike obeležja između svih parova frejmova zahvaćenih prozorom [LE08, COT06, LU13]. U [ADC04] izračunate razlike prosleđuju se klasifikatoru na dalju obradu u procesu detekcije. Za potrebe *flash* detekcije koristi se i poređenje dva frejma koja nisu uzastopna [QUE04]. Analiza tripleta frejmova prikazana je u [CAL02, MOH12]. Primena tekućeg frejma podrazumeva korišćenje prozora koji se računa od pozicije poslednje detektovane promene kadra do trenutnog frejma [PRI04] gde se srednja vrednost i varijansa razlike obeležja ažurira sa svakim novim frejmom, a resetuje posle promene kadra. Statistika obeležja se može značajno razlikovati unutar kadra, tako da prvih M frejmova ne može biti reprezentativno što pokazuje slabost ovakvog pristupa.

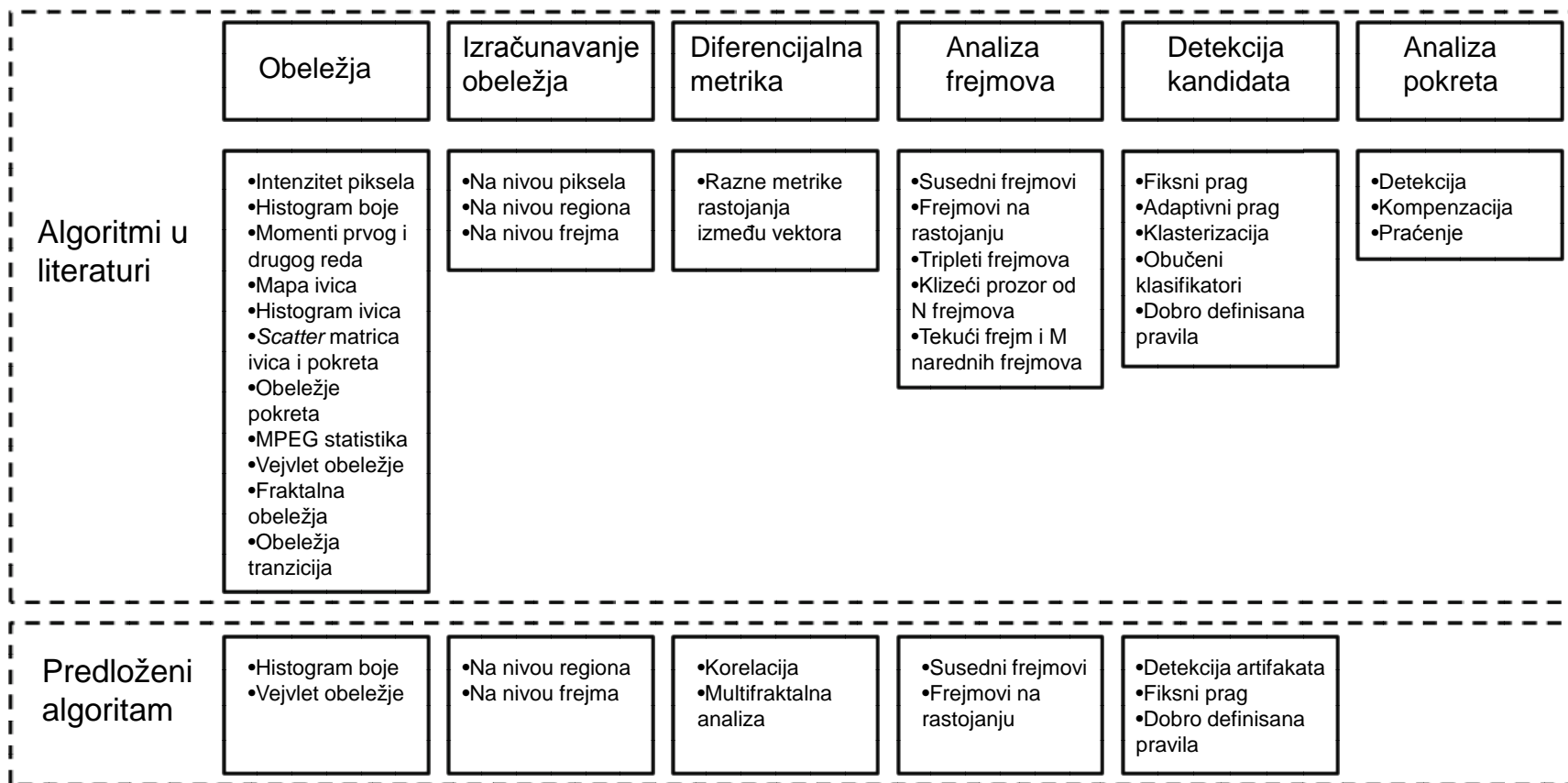
U algoritmima za pronalaženje naglih promena kadrova primenjene su različite metode detekcije, na osnovu fiksnog praga, na osnovu adaptivnog praga, klasterizacijom ili primenom obučenih klasifikatora. Primena fiksnog praga [LIU07, MIS13] u detekciji promena kadrova podrazumeva definisanje praga određenog tokom obučavanja sistema koji se ne menja tokom analize video sekvence. Vrednosti diskontinuiteta razlika obeležja se porede sa pragom u procesu detekcije naglih promena kadrova. Ovakav pristup ne uzima u obzir varijaciju diskontinuiteta sekvence koja se analizira i primena statičkog praga u procesu detekcije može proizvesti određen broj lažnih detekcija u raznim delovima video sekvence koje su posledica pomeranja objekta ili kamere. Adaptivni prag obezbeđuje lokalno prilagođenje statistici dela sekvence koji se analizira [AMI04, PRI04, LU13] tako što se trenutna statistika razlika vrednosti obeležja izračunava na osnovu tekućeg kadra ili vremenskog prozora od N frejmova. Kod primene klasterizacije postoje različiti pristupi, primena kao što su *k-means* algoritam [LEE00] za klasterizaciju razlike obeležja u dva klastera i spektralna klasterizacija [DAM07]. U poslednjoj tehnici kreira se matrica sličnosti frejmova u okviru klizećeg prozora od N frejmova i na osnovu normalizovanog *hard cut* algoritma se određuje optimalna promena kadrova. U slučaju treniranih klasifikatora u procesu detekcije [LIU07] koristi se neki od postojećih algoritama za klasifikaciju koji podrazumevaju prethodno obučavanje (na primer *Support Vectors*

Machines, SVM) [STE08] ili *multilayer perceptron* [MOH12]. Prednost ovakvih algoritama je što ne zahtevaju eksplicitno definisanje pragova. Nedostatak je kompleksno izračunavanje i samim tim dugo vreme izvršavanja algoritama. Klasifikacija kandidata nagle promene kadrova primenom preprocesiranja i odlučivanja na osnovu predefinisanih pravila prikazana je u [REN09]. Klasifikacija kandidata naglih promena, primenom poređenja paterna na osnovu definisanih mera sličnosti, prikazana je u [LU13].

Analiza pokreta u video sekvenci može biti implementirana u svrhu smanjivanja lažnih detekcija ukoliko je primenjeno obeležje osetljivo na pokret [LIU07]. Analiziran je pokret između parova frejmova a razlika frejmova je računata na nivou blokova. Drugi primer primene analize pokreta u detekciji naglih promena kadrova dat je u [MOH12], gde se za tu svrhu koristi matrica pokreta definisana *scatter* matricom ivica. Iako se koriste vrlo sofisticirani algoritmi za detekciju pokreta, pouzdanost detekcije nije velika pa se iz tog razloga analiza pokreta koristi u kombinaciji sa drugim obeležjima.

Stanje u literaturi, u oblasti detekcije naglih promena kadrova, ilustrovano je na Slici 4.1. Kao dodatak, na slici je opisana pripadnost predloženog algoritma po postojećim kategorijama iz literature. Algoritmi koji će se koristiti za poređenje, sa predloženim algoritmom u ovoj disertaciji, zasnivaju se na principima koji su prikazani u datom pregledu literature. Njihov pregled sa osnovnim karakteristikama dat je u Tabeli 4.1.

Algoritmi za detekciju naglih promena kadrova



Slika 4.1. Algoritmi za detekciju naglih promena kadrova, pregled literature i pozicija predloženog algoritma u odnosu na postojeće pristupe.

Tabela 4.1. Pregled najuspešnijih TRECVID (2004-2007) algoritama za detekciju naglih promena kadrova (sa znakom x su označeni osnovni principi na kojima se zasnivaju).

TRECVID	Algoritam	ML	KH	FL	LV	KOM	AP	KP	Ivice	VPK	Ostalo
TV2004	Tsinghua University [YUA04]		x	x	x						
TV2004	Fraunhofer/Heinrich Hertz Institute [PET04]			x	x		x	x	x		
TV2005	Tsinghua University [YUA05]	x	x	x			x				
TV2005	National ICT Australia [YU05]	x									
TV2005	KDDI R&D Labs Inc [MAT05]	x	x	x	x	x	x	x	x		
TV2005	Technical University Delft [NAC05]			x	x					x	
TV2005	Hong Kong Polytechnic University [CAI05]		x	x							x
TV2005	IBM Research [AMI05]	x	x	x	x		x		x		
TV2007	Beijing University of Posts and Telecommunications [ZHA07]	x	x	x	x			x	x		
TV2007	Science and Technical Research Laboratories, Tokio [KAW07]			x	x					x	
TV2007	Tsinghua University [YUA07, CAO06]	x	x	x			x				
TV2007	University of Karlsruhe [EKE07]		x	x	x				x		
TV2007	University of Marburg [MUH07]	x	x		x			x	x		

ML Upotreba obučanih klasifikatora (*machine learning*)
 KH Upotreba kolor histograma
 FL Realizacija *flash* detektora
 LV Pri poređenju frejmova koriste se vrednosti luminanse
 KOM Detekcija u kompresionom domenu (*MPEG-1*)

AP Detekcija primenom adaptivnih pragova
 KP Kompenzacija pokreta
 Ivice Detekcija i upotreba ivica
 VPK Upotreba ostalih vremensko-prostornih karakteristika
 Ostalo Primena ostalih tehnika

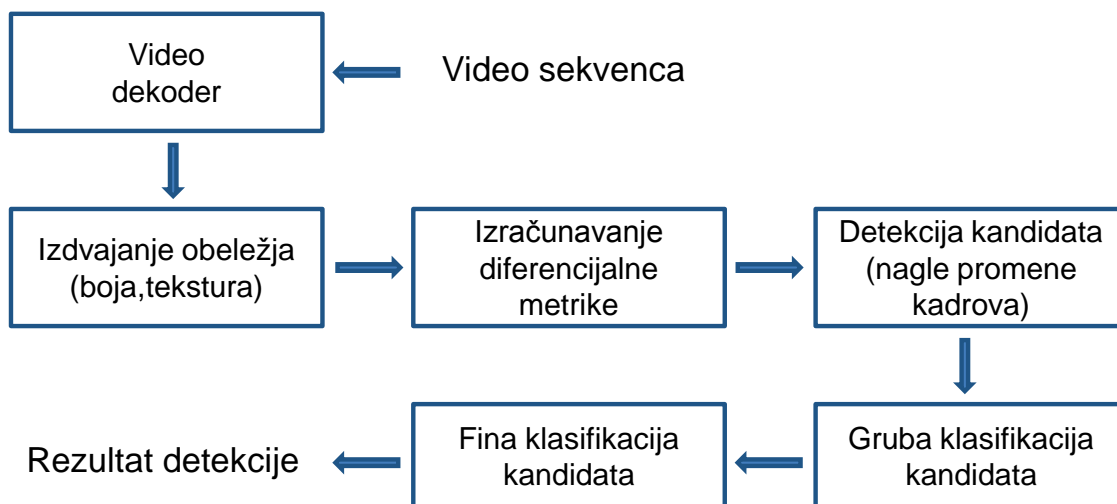
5. Novi algoritam za detekciju naglih promena kadrova

U ovom delu disertacije predstavljen je predlog novog algoritma za detekciju naglih promena kadrova u video sekvenci, zasnovan na primeni multifraktalne analize. Na pažljivo odabrane vektore obeležja primenjuje se odgovarajuća diferencijalna metrika koja obezbeđuje isticanje naglih promena kadrova. Primenom multifraktalne analize i drugih tehnika obrade signala vrši se detekcija potencijalnih kandidata naglih promena kadrova u vremenskom domenu. Naknadnom klasifikacijom iz grupe kandidata utvrđuje se konačni skup detektovanih naglih promena kadrova. Multifraktalna analiza je iskorišćena u ovom istraživanju kao pogodan alat za opisivanje singulariteta, odnosno naglih promena u signalu. U nastavku teksta dat je uopšteni uvid u celinu predloženog algoritma, a zatim su izneti detalji realizacije pojedinih faza u procesu detekcije.

5.1. Predloženi algoritam

Na Slici 5.1 predstavljena je blok šema predloženog algoritma za detekciju naglih promena kadrova u video sekvenci. Predložena rešenja pojedinačnih delova celokupnog

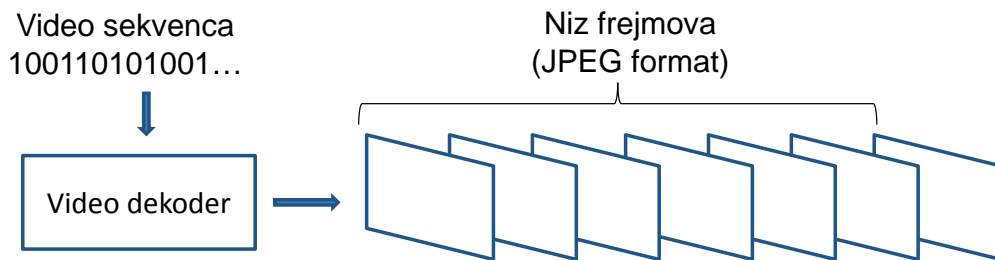
algoritma proizašla su kao rezultat intenzivnih istraživanja. Proces detekcije sastoji se od šest funkcionalno nezavisnih faza, u okviru kojih su primenjene različite tehnike obrade signala. Ulazna promenjiva predloženog algoritma je proizvoljna video sekvenca. U prvoj fazi procesa detekcije vrši se dekodovanje video sekvence i generisanje niza frejmova. Analiza kreiranih frejmova i izdvajanje obeležja niskog nivoa iz svakog pojedinačnog frejma izvodi se u drugoj fazi. Treća faza predstavlja poređenje izdvojenih obeležja izračunavanjem diferencijalne metrike koja treba da ukaže na razliku između susednih frejmova. Detekcija trenutaka u kojima je došlo do nagle promene sadržaja, tj. pozicije kandidata nagle promene kadrova obavlja se u četvrtoj fazi. Poslednje dve faze procesa detekcije predstavljaju grubu i finu klasifikaciju pozicija kandidata. U petoj fazi vrši se osnovna klasifikacija kandidata na osnovu zadatih klasifikacionih pragova. U šestoj fazi vrši se fina klasifikacija kandidata za nagle promene kadrova, na osnovu više dodatnih parametara. Rezultat procesa detekcije naglih promena kadrova su pozicije frejmova u video sekvenci, koji predstavljaju poslednje frejmove prethodnog kadra pri nagloj promeni. U nastavku teksta dat je detaljan opis realizacije svake od pomenutih faza procesa detekcije predloženog algoritma.



Slika 5.1. Blok šema algoritma za detekciju naglih promena kadrova.

5.2. MPEG dekodovanje

U prvoj fazi algoritma za detekciju naglih promena kadrova vrši se dekodovanje video sekvenci iz proizvoljnog video formata u niz frejmova, kao što je ilustrovano na slici 5.2. Dekodovanje se obavlja na osnovu originalnog video formata da bi se izbegli razni oblici degradacije videa, koji su posledica prekodovanja videa u neki drugi format [SHE08]. Promenom video formata može doći do degradacije videa u prostornom i vremenskom domenu. U prostornom domenu može doći do pogrešnog određivanja vektora pomeraja pri prekodovanju što proizvodi *blur* efekat na dekodovanom frejmu, koji može biti detektovan kao nagla promena sadržaja između dva susedna frejma. U vremenskom domenu može doći do ispuštanja frejmova zbog potrebe da se brzina protoka slika uskladi ili u slučaju da kodeci ne podržavaju promenljiv protok slika. Izostajanje frejmova proizvodi veliku razliku u sadržaju između susednih frejmova video sekvence u oba slučaja (u prostornom i vremenskom domenu). Takođe, pored standardnih promena kadrova i ovakve značajne promene sadržaja mogu biti detektovane kao nagle promene kadrova.



Slika 5.2. Dekodovanje video sekvence u niz frejmova (slika).

U ovom istraživanju za dekodovanje video sekvenci korišćen je *VirtualDub* softver [ref] sa neophodnim kodecima za dekodovanje različitih video formata u niz frejmova. U cilju smanjenja kompleksnosti izračunavanja obeležja koja se izdvajaju iz frejmova, nakon dekodovanja izvršeno je skaliranje frejmova na dimenzije 352x288 piksela. Frejmovi su izdvojeni iz video sekvence u slike *JPEG* formata (*JPEG* bez gubitaka [WAL91]).

5.3. Odabir i izdvajanje obeležja za potrebe detekcije

Odabir obeležja za opisivanje sadržaja frejmova video sekvence, predstavlja vrlo važan korak u procesu kreiranja algoritma za detekciju naglih promena kadrova. S obzirom da se razlika sadržaja između frejmova meri na osnovu odabranih obeležja niskog nivoa, njihova diskriminativnost direktno utiče na efikasnost detekcije. Sa druge strane, prevelika osetljivost obeležja na promene sadržaja može uticati na preciznost detekcije. Svaka značajnija promena sadržaja u smislu bržeg pokreta objekata ili brže promene osvetljenja može biti detektovana kao nagla promena sadržaja. Pri odabiru obeležja vođeno je računa o nekoliko zahteva. Prvi zahtev vezan je za broj obeležja koji se koristi za opisivanje sadržaja. Cilj je bio odabrati minimalan broj koji će zadovoljiti potrebe detekcije. Sledeći zahtev je bio da kompleksnost izračunavanja obeležja bude što je moguće manja. Zahtev za diskriminativnošću obeležja sa jedne strane i što manju osetljivost obeležja na pokret i brze promene osvetljenja, sa druge strane, predstavljali su poseban izazov pri izboru obeležja.

5.3.1. Odabir obeležja

U okviru ovog istraživanja uzet je u obzir skup od pet obeležja boje (kolor histogram obeležje, obeležje raspodele boje (*color layout*), skalabilno kolor obeležje (*scalable color*), obeležje kolor momenata (*color moments*), obeležja kolor korelograma (*color correlogram*) i četiri obeležja teksture (obeležja Gabor transformacije, obeležje histograma ivica, obeležja matrica uporednog pojavljivanja (*co-occurrence*), vejevlet obeležja). Razmatrani su i različiti načini izračunavanja nekih od pomenutih obeležja (globalni i lokalni histogram ivica, obeležja standardnih i radijalnih *co-occurrence* matrica). Izabrani skup obeležja proizašao je iz istraživanja vezanih za pretraživanje na osnovu sadržaja u okviru MPEG-7 standarda, CBIR (CBIR, *Content Based Image Retrieval*) sisteme [ZAJ06, ZAJ07, ZAJ08, ZAJ12]). U ovim istraživanjima pomenuti skup obeležja niskog nivoa pokazao se adekvatnim za potrebe detekcije.

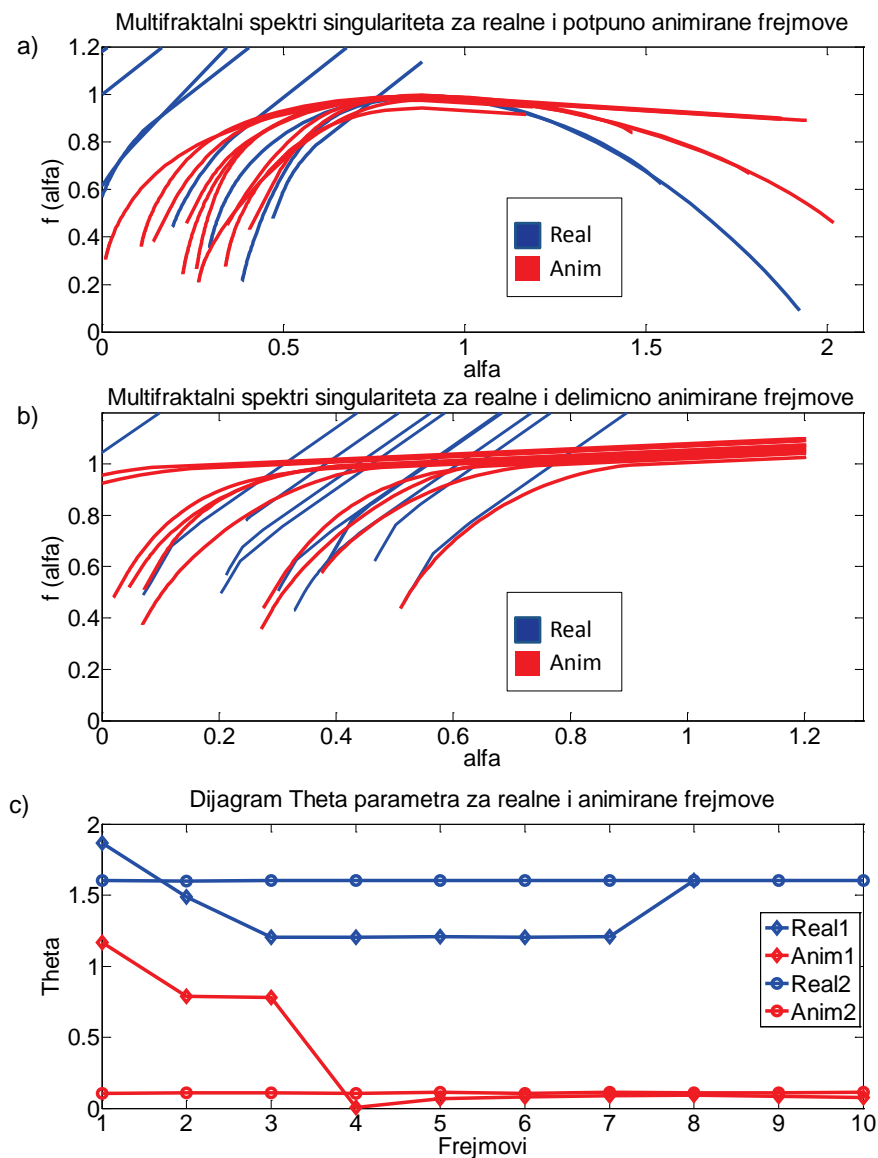
U okviru istraživanja [ZAJ07] korišćen je vektor obeležja dobijen konkatencijom vektora pojedinačnih obeležja. Pokazalo se da je moguće redukovati broj koordinata vektor

obeležja a da se ne degradiraju performanse pretraživanja. Odabir koordinata vrši se na osnovu detekcije naglih promena vrednosti koordinata unutar vektora obeležja. Vrednosti koje koordinata koje su se značajno razlikovale od okolnih koordinata odabirane su kao reprezentativne koordinate za posmatrani vektor obeležja. Pri odabiru vodilo se računa o ravnoteži broja koordinata koje reprezentuju boju, odnosno teksturu. Najveći broj koordinata, među odabranim, poticao je iz delova vektora obeležja koje su zauzimale koordinate kolor histograma i obeležja Gabor transformacije. Ovaj rezultat je pokazao da se može vrlo uspešno pretraživati klasteri sličnih slika korišćenjem informacija iz samo jednog obeležja boje i jednog obeležja teksture. Ako kadrove video sekvenci posmatramo kao klaster sličnih slika, analogno sa pomenutim rezultatom, za detekciju razlike između dva klastera slika sasvim su dovoljni po jedno obeležje boje i teksture.

Dodatna istraživanja [ZAJ11] pokazala su da se primenom multifraktalne analize može povećati osetljivost obeležja na promene sadržaja. Na vektore obeležja frejmova primenjena je multifraktalna analiza (histogramska metoda) i praćena je promena vrednosti koordinata obeležja na lokacijama naglih promena kadrova u video sekvenci. Rezultati su pokazali da se visoke vrednosti relativne promene ($>50\%$) koordinata, pri naglim promenama kadrova, pojavljuju na pozicijama koordinata kolor histograma i vejevlet obeležja. Ovaj rezultat je pokazao da kolor histogram i vejevlet obeležje pružaju sasvim zadovoljavajuću osetljivost na nagle promene sadržaja koja se može iskoristiti za potrebe detekcije naglih promena kadrova u video sekvenci.

Dodatni test osetljivosti obeležja na nagle promene sadržaja izveden je na jedinstvenom vektoru obeležja dobijenim konkatenacijom kolor histograma i vejevlet obeležja. Primenom multifraktalne analize (*Legendre* metoda [FRA12]) izvedena je klasifikacija realnih frejmova i animiranih frejmova na osnovu vektora obeležja. Na Slici 5.3 prikazani su rezultati klasifikacije, pomoću multifraktalnih spektara singulariteta za jedinstvene vektore obeležja frejmova video sekvenci. Na Slici 5.3(a) poređeni su spektri realnih i animiranih frejmova različitih video sekvenci (slučajan odabir frejmova iz sedmih segmenata videa “*Good Year*” i “*Little Mermaid*”). Jasno je uočljiva razlika između multifraktalnih spektara. Na Slici 5.3(b) poređeni su realni i delimično animirani iste video

sekvence (slučajno izabrani frejmovi iz sedmog segmenta videa “*Who Framed Roger Rabbit*”). Na Slici 5.3(c) prikazane su vrednosti parametra Θ izračunatog za svaki od prikazanih spektara na Slikama 5.3(a) i 5.3(b).



Slika 5.3. Rezultati klasifikacije realnih i animiranih frejmova na osnovu integralnog vektora obeležja: a) multifraktalni spektri singulariteta za realne i animirane frejmove različitih video sekvenci, b) multifraktalni spektri singulariteta za realne i animirane frejmove iste video sekvence, c) vrednosti Θ parametra za prikazane multifraktalne spektre.

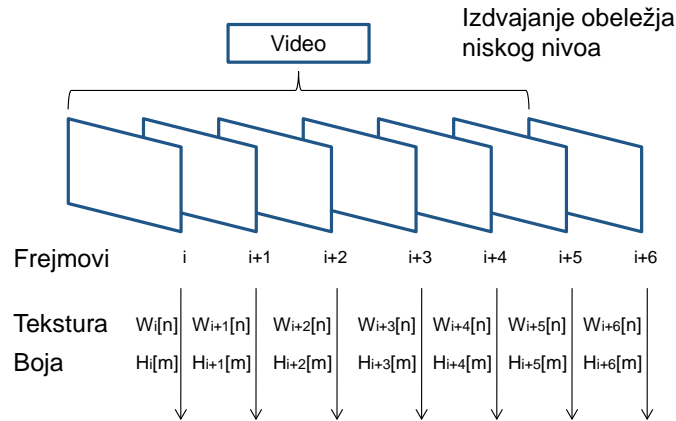
Prikazani parametar predstavlja ugao desnog završetka spektra koji ukazuje na fine varijacije unutar signala. Sa prikazanih slika može se videti da je na osnovu kolor histograma i vejtlet obeležja moguće razlikovati potpuno animirane frejmove od realnih, a posebno je značajno da se može razlikovati realni frejm od delimično animiranog frejma. Ovaj rezultat je posledica činjenice da animacija ima manji spektar boja od realnog frejma, dok je tekstura manje rustična u odnosu na realni frejm. Histogram boje i vejtlet obeležje omogućavaju uočavanje ovako finih razlika između frejmova. Princip klasifikacije razvijen u ovom eksperimentu iskorišćen je za unapređenje klasifikacije jednodimenzionalnog signala [GAV13], a rezultat eksperimenta je potvrdio zadovoljavajuću osetljivost odabranih obeležja na promenu sadržaja.

Kolor histogram predstavlja obeležje niskog nivoa kompleksnosti, a uključuje kolor transformaciju, kvantizaciju i sabiranje. Vejtlet obeležje predstavlja obeležje visoke diskriminativnosti, ali i veće složenosti od histograma, jer uključuje dekompoziciju signala primenom diskretne vejtlet transformacije (*DWT, Discrete Wavelet Transformation*) [EDW91].

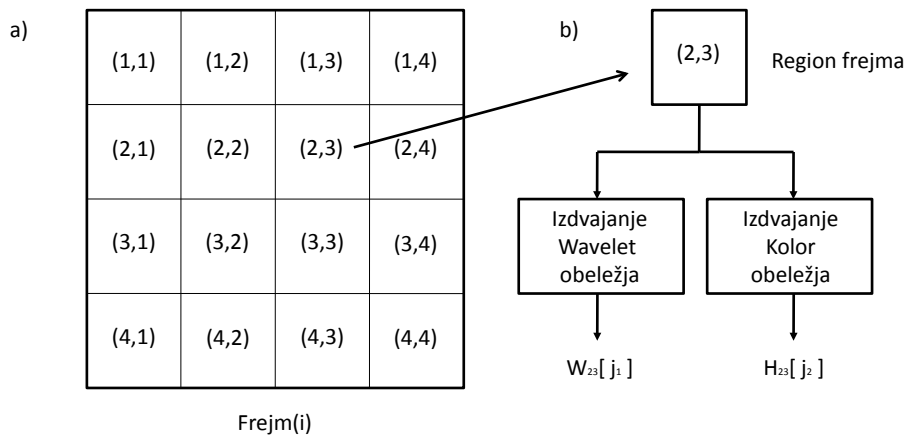
Histogram boje pokazuje određenu toleranciju na kretanje objekata unutar niza frejmova, jer raspodela piksela u histogramu ne zavisi od prostorne raspodele istih unutar frejma. Sa druge strane, na vejtlet obeležje utiče svako kretanje u nizu frejmova. Promena teksture izazvana kretanjem uzrokuje promene pojedinih delova ovog obeležja. Kombinovanjem histograma boje i vejtlet obeležja osetljivost vejtleta delimično se ublažava tolerancijom histograma. Međutim, kada su u pitanju nagle promene osvetljenja ili brza kretanja objekata, izbegavanje detekcije ovakvih događaja predstavlja pravi izazov. Na ovakav izazov nije dovoljno odgovoriti samo odabirom odgovarajućih obeležja [PET10], već je neophodno primeniti dodatne tehnike za ublažavanje uticaja pomenutih efekata na detekciju. Na primer, odgovarajuće diferencijalne tehnike ili *flash* detektore.

5.3.2. Izdvajanje obeležja

Kao što je već rečeno, razlika između sadržaja frejmova može se ustanoviti izračunavanjem razlike između obeležja sa kojima su frejmovi opisani. Iz navedenog razloga neophodno je izdvojiti odabrana obeležja, histogram boje i vejtlet obeležje, iz svakog pojedinačnog frejma. To je ilustrovano na Slici 5.4.



Slika 5.4. Izdvajanje obeležja niskog nivoa iz sekvence frejmova.



c)

$$W_i = [W_{11}[j_1] \ W_{21}[j_1] \ W_{31}[j_1] \ W_{41}[j_1] \ W_{12}[j_1] \ W_{22}[j_1] \ W_{32}[j_1] \ W_{42}[j_1] \ W_{13}[j_1] \ W_{23}[j_1] \ W_{33}[j_1] \ W_{43}[j_1] \ W_{14}[j_1] \ W_{24}[j_1] \ W_{34}[j_1] \ W_{44}[j_1]]$$

$$HFD_i = [H_{11}[j_2] \ H_{21}[j_2] \ H_{31}[j_2] \ H_{41}[j_2] \ H_{12}[j_2] \ H_{22}[j_2] \ H_{32}[j_2] \ H_{42}[j_2] \ H_{13}[j_2] \ H_{23}[j_2] \ H_{33}[j_2] \ H_{43}[j_2] \ H_{14}[j_2] \ H_{24}[j_2] \ H_{34}[j_2] \ H_{44}[j_2]]$$

$$H_i = \text{sum}(H_{11}[j_2], H_{21}[j_2], H_{31}[j_2], H_{41}[j_2], H_{12}[j_2], H_{22}[j_2], H_{32}[j_2], H_{42}[j_2], H_{13}[j_2], H_{23}[j_2], H_{33}[j_2], H_{43}[j_2], H_{14}[j_2], H_{24}[j_2], H_{34}[j_2], H_{44}[j_2])$$

Slika 5.5. Algoritam izdvajanja obeležja niskog nivoa za jedan frejm:
a) prikaz podele frejma na regione, b) princip izdvajanja obeležja iz regiona frejma,
c) konkatencija izračunatih obeležja regiona u vektore obeležja.

Da bi se obezbedila što veća osetljivost na promene sadržaja, obeležja se izračunavaju na nivou regiona frejmova. Svaki od frejmova je podeljen na pravougaone regione (4×4), kao na Slici 5.5.(a). Obeležja se izračunavaju za svaki od regiona, Slika 5.5.(b). Niz $W_{23}[j_1]$ predstavlja niz koordinata vejevlet obeležja dužine j_1 , a $H_{23}[j_2]$ predstavlja niz koordinata kolor histograma dužine j_2 . Od izračunatih obeležja regiona kreiraju se globalna obeležja za ceo frejm, a način formiranja globalnih obeležja boje i teksture prikazan je na Slici 5.5.(c). Kao što se može videti, na osnovu dva obeležja niskog nivoa svakog od regiona kreiraju se tri globalna obeležja za ceo frejm, W_i , HFD_i i H_i .

Globalno vejevlet obeležje W_i predstavlja niz dužine $k=192$ vrednosti. Ovaj niz je dobijen konkatencijom vejevlet obeležja regiona redosledom koji je prikazan na Slici 5.5(c). Regionalno izračunato teksturno obeležje obezbeđuje veliku osetljivost na nagle promene sadržaja, ali i na kretanje objekata i kamere unutar kadra. Globalno obeležje H_i predstavlja niz dužine $m=162$ vrednosti, a u suštini odgovara histogramu boje celog frejma koji je dobijen sabiranjem histograma regiona. Ovo globalno obeležje boje ima veću toleranciju na kretanje unutar kadra, ali je osetljivo na nagle promene osvetljenja. Globalno obeležje HFD_i predstavlja niz dužine $n=2592$ koordinata, a dobijen je konkatencijom histograma regiona redosledom koji je prikazan na Slici 5.5(c). Ovo obeležje boje koristi se za detekciju nagle promene osvetljenja (*flash* detektor) i sadrži informaciju o lokalnoj raspodeli boje.

Odabrana obeležja svih frejmova, za boju i teksturu, smeštena su u dve odvojene matrice MHi i MWi koje su prikazane izrazima (5.1) i (5.2). S obzirom da proces izdvajanje obeležja zahteva određeno vreme izvršavanja, ubrzanje celog postupka može se postići paralelizacijom. Moderni računarski sistemi nude više procesorske jedinice koje mogu izvršavati određene operacije paralelno. Da bi se iskoristile prednosti paralelnog procesiranja, video se može podeliti na segmente koji se mogu paralelno obrađivati. Pri određivanju veličine segmenata neophodno je voditi računa da segmenti sadrže dovoljan broj različitih kadrova po sadržaju i da u segmentima ne dominiraju određene dinamike promene sadržaja. Broj segmenata treba da bude prilagođen broju procesorskih jezgara na kojima se mogu paralelno obrađivati. Rezultati istraživanja su pokazali da veličina

segmenata od 20000 frejmova u potpunosti zadovoljava sve postavljene zahteve. Izdvojeni segmenti frejmova preklapaju se u dužini od 10 frejmova, da bi se izbeglo ispuštanje naglih promena kadrova koji se nalaze na prelazu između segmenata.

$$MH_j = [H_i[m]], \quad i=1:20000, \quad j=1:K \quad (5.1)$$

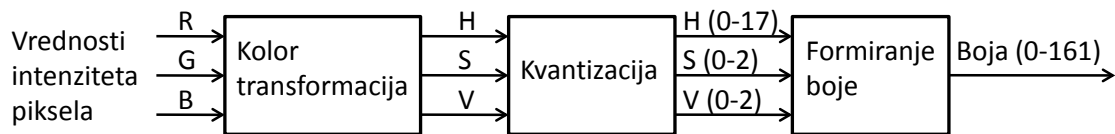
$$MW_j = [W_i[n]], \quad i=1:20000, \quad j=1:K \quad (5.2)$$

Poređenje obeležja je znatno efikasnije u normiranom metričkom prostoru [ZAJ07]. Razlika u nivoima koordinata može izazvati određenu dominaciju nekog dela određenog obeležja u odnosu na ostatak obeležja. Normalizacijom matrica obeležja frejmova postiže se da slični frejmovi po odgovarajućem obeležju budu što kompaktnije pozicionirani u metričkom prostoru. Normalizacija metričkog prostora realizovana je normiranjem matrica obeležja po kolonama, deljenjem kolona sa maksimalnom vrednošću kolone. Na ovaj način je obezbeđeno da se sve vrednosti unutar obeležja nalaze u intervalu između 0 i 1. U nastavku teksta predstavljeni su algoritmi za izračunavanje obeležja frejmova.

5.3.2.1. Izdvajanje obeležja boje: kolor histogram

Kao što je već rečeno, kolor histogram se izračunava za segment frejma. Sam algoritam za izračunavanje histograma preuzet je iz *MPEG-7* standarda [ISO01]. Algoritam se sastoji iz dva dela. Prvi deo predstavlja transformacija vrednosti intenziteta pojedinačnih piksela u indeks boje, pri čemu se formira slika indeksa boje. Drugi deo ovog algoritma predstavlja izračunavanje histograma slike indeksa boje. Principska šema algoritma za transformaciju vrednosti intenziteta piksela u indeks boje prikazana je na Slici 5.6. Vrednosti intenziteta piksela transformisane su iz *RGB* kolor prostora u *HSV* kolor prostor. Pri kvantizaciji vrednosti intenziteta u *HSV* kolor prostoru primenjen je odnos $H:S:V=18:3:3$. Na osnovu date kvantizacije formiraju se indeksi boje u intervalu [0-161] za svaki piksel. Za svaku sliku kreiran je histogram boje dužine $m=162$ binova.

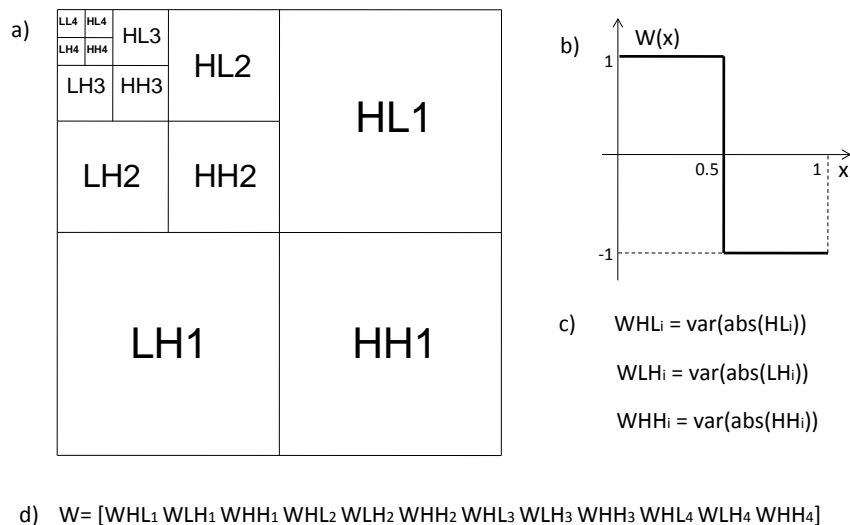
Izračunati kolor histogram normiran je ukupnim brojem piksela segmenta. Pri izračunavanju globalnog histograma sabiraju se kolor histogrami regiona, a dobijeni globalni histogram dodatno se normira sa brojem regiona, Slika 5.5(c).



Slika 5.6. Dijagram algoritma kolor transformacije.

5.3.2.2. Izdvajanje obeležja teksture: vevjlet obeležje

Algoritam za izračunavanje vevjlet obeležja preuzet je iz literature [EKE07]. Obeležje se izračunava na nivou regiona, odnosno na segmentu frejma. Na sliku segmenta primenjuje se *DWT* transformacija u vidu piramidalne dekompozicije slike četvrtog reda, Slika 5.7(a).



Slika 5.7. Izdvajanje obeležja niskog nivoa (boja i tekstura) iz dekodovanih frejmova.

- Primer piramidalne dekompozicije slike primenom vevjlet-a,
- Haar vevjlet, c) način izračunavanja vrednosti obeležja teksture,
- konkatenacija izračunatih obeležja.

Pri izračunavanju dekompozicije slike koristi se *Haar* vejevlet, Slika 5.7(b). Ovaj vejevlet je izabran zato što je najjednostavniji, odnosno najnižeg je reda. Pored navedenog ovaj vejevlet daje grubu aproksimaciju i čest je izbor za detekciju naglih promena u signalu [PAG07, WAL11, GER07]. Izračunavanje obeležja frekvencijskih podopsega prikazano je na Slici 5.7(c). Obeležje podopsega predstavlja varijaciju apsolutnih vrednosti koeficijenata unutar podopsega. Vejevlet obeležje segmenta dobijeno je konkatencijom obeležja podopsega redosledom koji je prikazan na Slici 5.7(d).

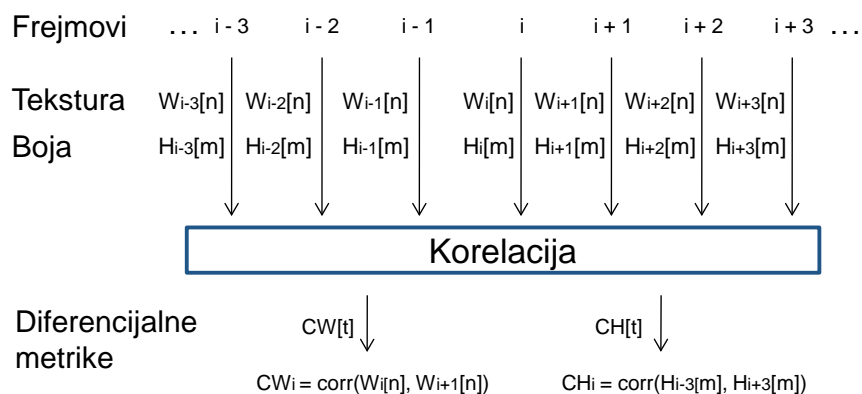
Vejevlet obeležje segmenta frejma sadrži 12 vrednosti. Globalno vejevlet obeležje, obeležje celog frejma, sadrži $n=192$ vrednosti. Kao što se može videti na osnovu dužina kolor i vejevlet obeležja, vrednosti m i n , dužine su uravnotežene. Dužina teksturnog obeležja veća od dužine kolor obeležja, ali ne toliko da je dominantnija u odnosu na njega.

5.3.3. Izračunavanje diferencijalne metrike

Diferencijalna metrika ima veoma važnu ulogu u detekciji naglih promena kadrova. Definisanjem načina merenja razlike između obeležja frejmova direktno se utiče na efikasnost detekcije naglih promena sadržaja na frejmovima. Ako definisana metrika nije dovoljno osetljiva na naglu promenu sadržaja, određen broj naglih promena kadrova može biti ispušten, odnosno nedektovan. Sa druge strane, preosetljiva metrika biće osetljiva generalno na promenu sadržaja, tako da će se pri detekciji naglih promena kadrova detektovati i određena kretanja objekata ili kamere unutar kadra.

Diferencijalna metrika pored dobre osetljivosti na nagle promene sadržaja mora imati malu zavisnost od nagle promene nivoa signala između frejmova, koji može biti uzrokovan naglom promenom osvetljenja (pojava *flash-a*). Dodatan problem pri definisanju metrike može biti nivo vrednosti koordinata obeležja, koji je posledica žanra video sekvence. U ovom slučaju promena nivoa signala može zahtevati prilagođavanje algoritma detekcije konkretnom video signalu, sa neizbežnom prethodnom detekcijom žanra video signala.

Kao odgovor na sve pomenute izazove odabrana je korelacija kao diferencijalna metrika za poređenje sadržaja susednih frejmova. Korelacija kao diferencijalna metrika nije često korišćena u sistemima za detekciju naglih promena kadrova. Napuštena je kao ideja, jer nije bila dovoljno precizna u lokalizaciji nagle promene sadržaja. U predloženom algoritmu za detekciju naglih promena kadrova korelacija se koristi isključivo kao diferencijalna metrika, a lokalizacija naglih promena sadržaja vrši se primenom drugih tehnika. Korelacija je kao metrika vrlo osetljiva na nagle promene vrednosti obeležja, samim tim i nagle promene sadržaja u frejmovima. Posедуje određenu toleranciju na promene vrednosti koordinata unutar obeležja, koje mogu biti posledice kretanja objekata unutar kadra. Nije mnogo osetljiva na nagle promene nivoa vrednosti obeležja, kao posledice nagle promene osvetljenja na frejmu. Daje normirani rezultat u intervalu [0-1] za proizvoljne nivoe vrednosti obeležja. Normiran rezultat dobijen merenjem razlike vrednosti obeležja, omogućava u okviru algoritma definisanje fiksnih parametara za detekciju naglih promena kadrova, bez obzira na žanzr, dinamiku i osvetljenje video sekvence.



Slika 5.8. Prikaz izračunavanja različitih diferencijalnih metrika HC i HW .

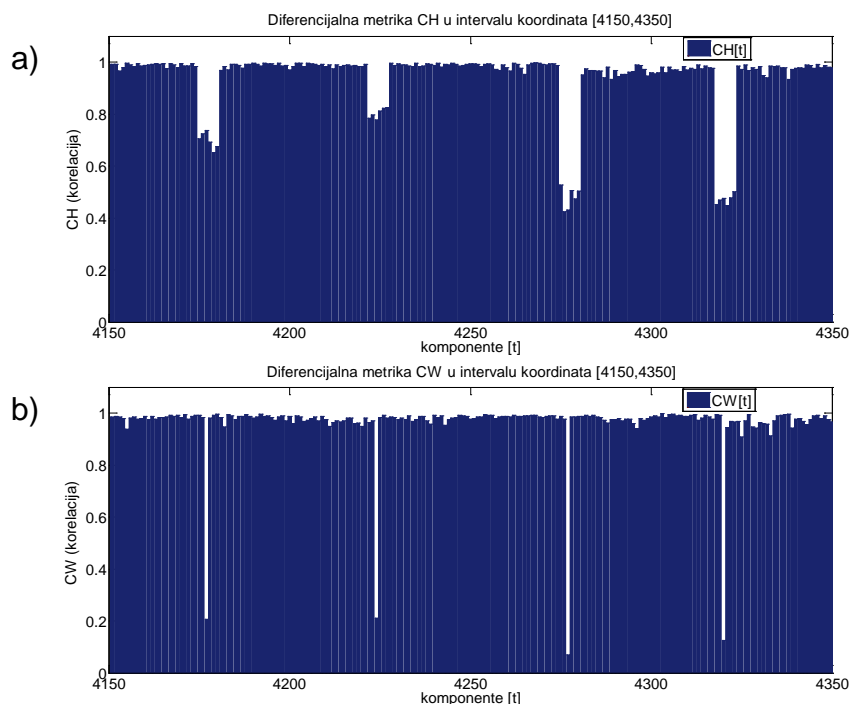
Izračunavanje razlike između obeležja frejmova primenom korelacije realizovano je na različite načine za kolor histogram i vejevlet obeležje. Funkcionalna šema ovog dela algoritma prikazana je na Slici 5.8. Za kolor histogram izračunata je korelacija obeležja frejmova koji se nalaze tri frejma ispred i tri frejma iza posmatranog frejma u vremenskom domenu (5.3). Cilj ovakvog izračunavanja je da se proizvedu odgovarajući artefakti na

mestima naglih promena kadrova koji će biti "uočljivi". Ovi artefakti ukazuju na naglu tranziciju između kadrova i specifičnog su oblika. Za vejevlet obeležje izračunata je korelacija između susednih frejmova (5.4). Cilj ovakvog načina izračunavanja diferencijalne metrike je precizna lokalizacija nagle promene kadrova.

$$CH_i = \text{corr}(H_{i-3}[m], H_{i+3}[m]), \quad i = 4:19997 \quad (5.3)$$

$$CW_i = \text{corr}(W_i[n], W_{i+1}[n]) \quad i = 4:19997 \quad (5.4)$$

Diferencijalna metrika označena sa CH primenjena je na matricu obeležja kolor histograma MH , i kao rezultat primene dobijen je niz $CH[t]$. Na matricu vejevlet obeležja MW primenjena je CW diferencijalna metrika i rezultat ovog postupka je niz $CW[t]$. Vrednost t je prirodan broj u intervalu [1-19996]. Preklapanjem segmenata obezbeđeno je da ne dođe do ispuštanja potencijalnih naglih promena kadrova usled smanjenja broja vrednosti koje će se koristiti u daljoj analizi.



Slika 5.9. Diferencijalne metrike: a) metrika CH u intervalu koordinata [4150,4350], b) metrika CW u intervalu koordinata [4150,4350].

Pozicija nagle promene kadrova definiše se pozicijom poslednjeg frejma prethodnog kadra, pre nagle promene. Ova pozicija će biti okarakterisana smanjenom korelacijom i biće uočljiva u oba niza koja su dobijena primenom diferencijalnih metrika, CH i CW . Na Slikama 5.9(a) i 5.9(b) predstavljeni su delovi nizova $CH[t]$ i $CW[t]$, respektivno, sedmog segmenta video sekvence "GoodYear" u dužini od dvesta vrednosti niza. Na Slici 5.9(a) mogu se uočiti artefakti koji ukazuju na moguće nagle promene kadrova, dok su na slici 5.9(b) uočljiva nagla opadanja vrednosti korelacije na određenim pozicijama, koja takođe ukazuju na nagle promene kadrova.

5.4. Detekcija kandidata

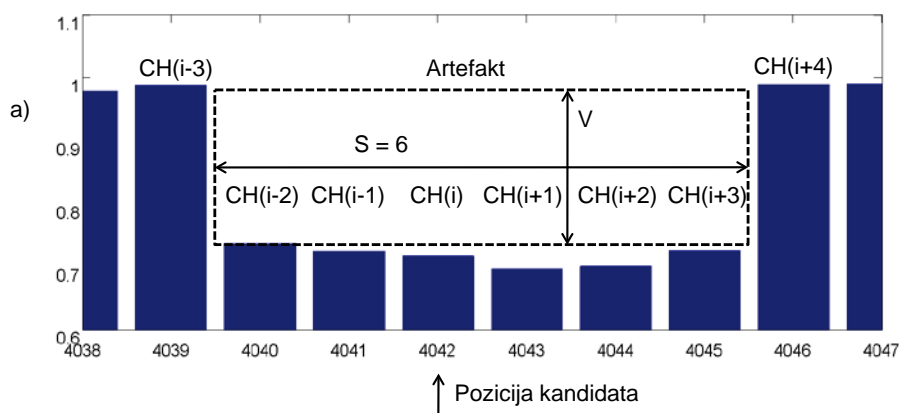
Algoritam za detekciju kandidata naglih promena kadrova sastoji se iz dva dela. Prvi deo predstavlja detekcija artefakata u nizu diferencijalne metrike $CH[t]$. Drugi deo ovog algoritma čini verifikacija detektovanih kandidata primenom transformisanog niza diferencijalne metrike $CW[t]$. Transformacija niza $CW[t]$ sastoji se od formiranja komplementarnog niza od originalnog, i primene multifraktalne analize na dobijeni komplementarni niz. Rezultat izvršavanja ove faze algoritma za detekciju naglih promena kadrova je skup lokacija u vremenskom domenu, koje predstavljaju potencijalne lokacije naglih promena kadrova.

5.4.1. Detekcija artefakata

U ovom delu algoritma detektuju se artefakti predefinisano oblika u $CH[t]$, nizu diferencijalne metrike za kolor histogram. Artefakt je definisan kao pravougaonik koji je moguće upisati u strukturu promene vrednosti niza $CH[t]$, Slika 5.10(a). Veličine koje karakterišu artefakt su njegova širina S i visina V , čija je definicija prikazana na slikama 5.10(a) i 5.10(b).

Vizuelnom inspekcijom velikog broja frejmova video sekvenci ($>1.500.000$ frejmova) identifikovane su različite širine artefakata koje odgovaraju naglim promena

sadržaja. Stvarnoj nagloj promeni kadrova odgovara širina $S=6$ komponenata. Međutim, kretanja i razni specijalni efekti unutar kadrova degradiraju ivice artefakata, pa broj vrednosti koji je zahvaćen artefaktima može varirati u intervalu 5-9 komponenata ($S \in [5,9]$). Visina artefakta mora biti veća od vrednosti $V = 0.001$ da bi se pozicija komponente niza proglasila kandidatom. Obavljena istraživanja pokazala su da je data vrednost minimalne visine artefakta dovoljna za efikasnu detekciju naglih promena. Mala vrednost minimalne visine detektovanog artefakta odgovara nagloj promeni između kadrova sa međusobno visoko korelisanim sadržajima.

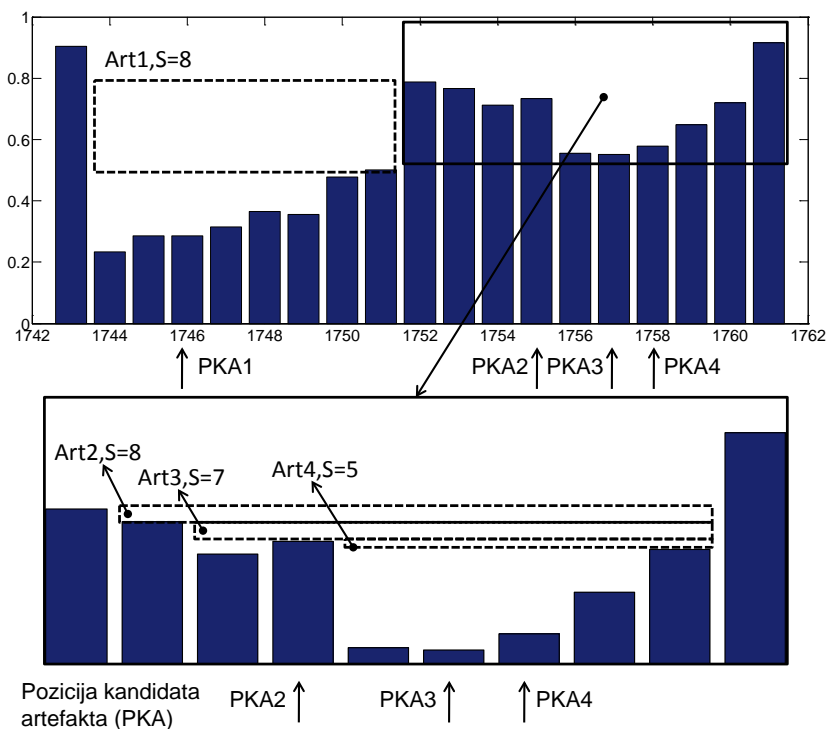


- b) $V = \min(\text{CH}(i-3), \text{CH}(i+4)) - \max(\text{CH}(i-2), \text{CH}(i-1), \text{CH}(i), \text{CH}(i+1), \text{CH}(i+2), \text{CH}(i+3))$
 $S = 6$ (komponenti)
 $(V > 0.001) \ \& \ (S = 6)$ uslov da pozicija $i = 4042$ bude detektovana kao kandidat

Slika 5.10. Detekcija artefakata: a) prikaz artefakta, b) algoritam detekcije artefakta.

Pozicija kandidata nagle promene kadrova definisana je kao centralna komponenta artefaktom zahvaćenog skupa komponenti niza diferencijalne metrike, Slika 5.10(a). Za širinu artefakata koja sadrži neparan broj komponenti niza analizira se simetrična okolina posmatrane komponente $HC(i)$. U slučajevima kada širina artefakta sadrži parni broj komponenti niza analizira se asimetrična okolina komponente $HC(i)$, Slika 5.10(a). U oba slučaja, ako su svi definisani uslovi zadovoljeni, lokacija komponente $HC(i)$ niza proglašava se kandidatom. Primer detekcije kandidata za artfakte koji zahvataju paran i neparan broj komponenata niza prikazan je na Slici 5.11. Nagla promena kadrova je

propraćena kretanjem objekata na završetku prethodnog kadra i na početku narednog kadra. S obzirom da je došlo do degradacije ivice artefakta širina artefakta je $S=8$. Kretanje objekata u prvom planu na početku narednog kadra uslovalo je velike oscilacije diferencijalne metrike i samim tim detekciju većeg broja kandidata. Grubom klasifikacijom u narednoj fazi algoritma za detekciju kandidati se klasifikuju i odbacuju oni koji ne zadovoljavaju određene uslove.



Slika 5.11. Detekcija artefakata (dodatni primer).

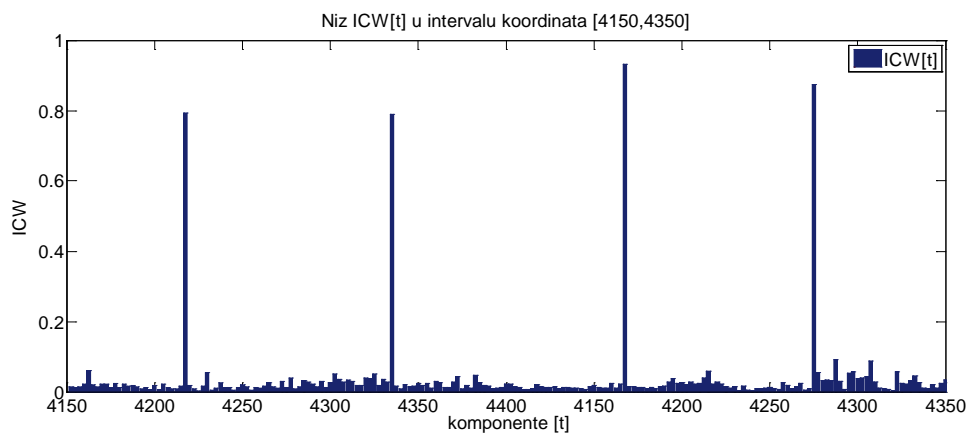
Detekcija artefakata u nizu diferencijalne metrike $CH[t]$ počinje od prve komponente koja se može testirati. Ivični efekti se zanemaruju, jer je mogućnost ispuštanja nekog od kandidata uklonjena preklapanjem segmenata frejmova. Rezultat ove faze algoritma je niz pozicija kandidata $PK[p]$ dužine p .

5.4.2. Verifikacija kandidata

Detektovana pozicija kandidata predstavlja uočenu anomaliju u signalu koja može biti posledica nagle promene kadrova. Detektovane pozicije kandidata pomoću artefakata ne moraju odgovarati pozicijima stvarnih naglih promena sadržaja između susednih frejmova. Kretanje objekata, specijalni efekti i nagla promena osvetljenja unutar kadrova mogu uticati na pomeranje artefakata u odnosu na poziciju stvarne nagle promene sadržaja. U ovoj fazi algoritma neophodno je povezati poziciju kandidata dobijenu na osnovu detekcije artefakta sa pozicijom stvarne nagle promene sadržaja. Kao što je već rečeno, vejevlet obeležje ima visoku osetljivost na promenu sadržaja između frejmova, zato se informacija o stvarnoj lokaciji nagle promene sadržaja može dobiti obradom niza diferencijalne metrike za vejevlet obeležje $CW[t]$. Prvo se izračunava komplementarni niz. Na ovaj niz se primenjuje multifraktalna analiza u cilju isticanja lokalnih maksimuma, koji predstavljaju pozicije naglih promena sadržaja. U sledećem koraku sledi identifikacija pozicija lokalnih maksimuma u multifraktalnom alfa nizu, na osnovu detektovanih pozicija kandidata.

5.4.2.1. Izračunavanje komplementarnog niza

Komplementarni niz $ICW[t]$ izračunava se u cilju pripreme niza $CW[t]$ za primenu multifraktalne analize.



Slika 5.12. Niz $ICW[t]$ u intervalu koordinata [4150,4350].

U okviru $CW[t]$ niza na pozicijama naglih promene sadržaja nalaze se mahom male vrednosti. Na tim pozicijama korelacija između susednih frejmova ima male vrednosti. Multifraktalna analiza koja treba da se primeni na ovaj niz, za tip signala sa skokovitim promenama intenziteta, najbolje performanse pokazuje kada su skokovite promene u stvari lokalni maksimumi u nizu. Transformacija lokalnih minimuma u lokalne maksimume obavlja se izračunavanjem komplementarnog niza $ICW[t]$. Princip izračunavanja ovog niza dat je izrazom (5.5), a deo izračunatog niza $ICW[t]$ predstavljen je na Slici 5.12.

$$ICW[t] = 1 - CW[t] \quad (5.5)$$

5.4.2.2. Primena multifraktalne analize

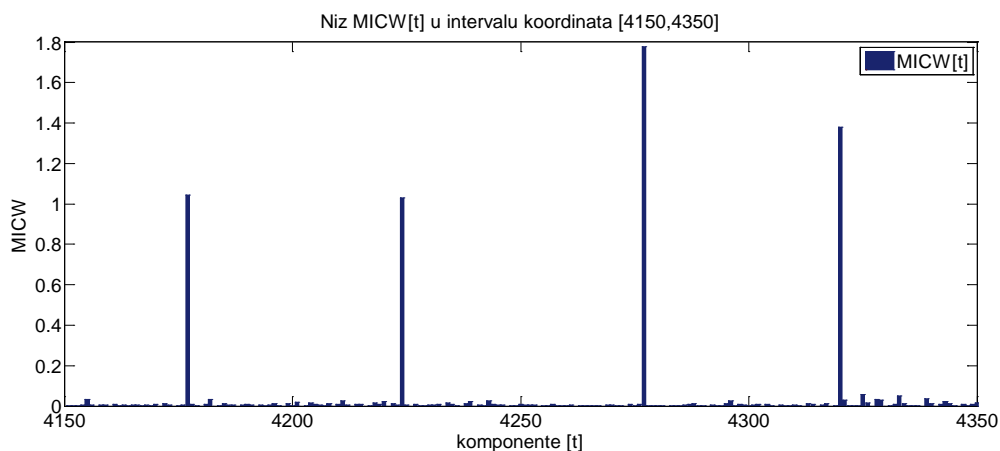
Pozicije naglih promena sadržaja okarakterisane su u nizu $ICW[t]$ kao lokalni maksimumi. Da bi se vrednosti ovih maksimuma istakle, neophodno je primeniti ogovarajuću tehniku obrade signala koja će ukazati na velike relativne promene vrednosti unutar niza. U ovom delu algoritma, za isticanje relativnih promena vrednosti unutar niza, primenjena je multifraktalna analiza zasnovana na histogramu. Ova analiza transformiše niz $ICW[t]$ u domen singulariteta, alfa domen. Izračunava se grubi Holderov eksponent, tj. alfa vrednost za svaku komponentu niza. Vrednosti ovih singulariteta karakterišu lokalno relativnu promenu signala, bez obzira na apsolutni nivo vrednosti komponenata signala. Lokalni maksimumi *alpha* vrednosti ukazuju na nagle promene vrednosti unutar analiziranog niza. Postupak analize niza $ICW[t]$ u *alpha* domenu naziva se inverznom multifraktalnom analizom (*IMFA*).

Primenom inverzne multifraktalne analize niz $ICW[t]$ se transformiše u niz $MICW[t]$, jednačina (5.6). Na Slici 5.13 prikazan je deo $MICW[t]$ niza. Nakon izračunavanja inverzne multifraktalne analize vrši se "pročišćavanje" niza, filtriranjem filtrom dužine 5. Cilj ovog filtriranja je da se u nizu izoluju pozicije dominantnih lokalnih maksimuma. Filtar je klizni prozor koji sadrži pet ćelija, čiji je korak pomeranja jedna ćelija. Od zahvaćenih vrednosti komponenata niza unutar filtra, zadržava se samo vrednost

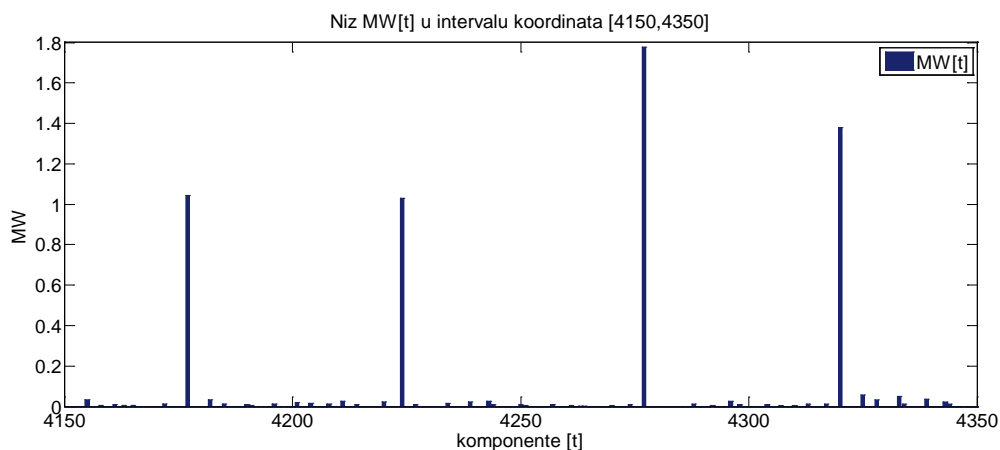
komponente sa maksimalnom vrednošću, koja se prenosi u novi niz sa zadržavanjem originalne pozicije. Vrednosti ostalih komponenata postavljaju se na nulu. Ovaj postupak u jednačini (5.7) označen je sa oznakom *filter*. Rezultat filtriranja niza $MICW[t]$ je niz $MW[t]$. Na Slici 5.14 predstavljen je deo niza $MW[t]$.

$$MICW[t] = IMFA(CW[t]) \quad (5.6)$$

$$MW[t] = filter(MICW[t]) \quad (5.7)$$



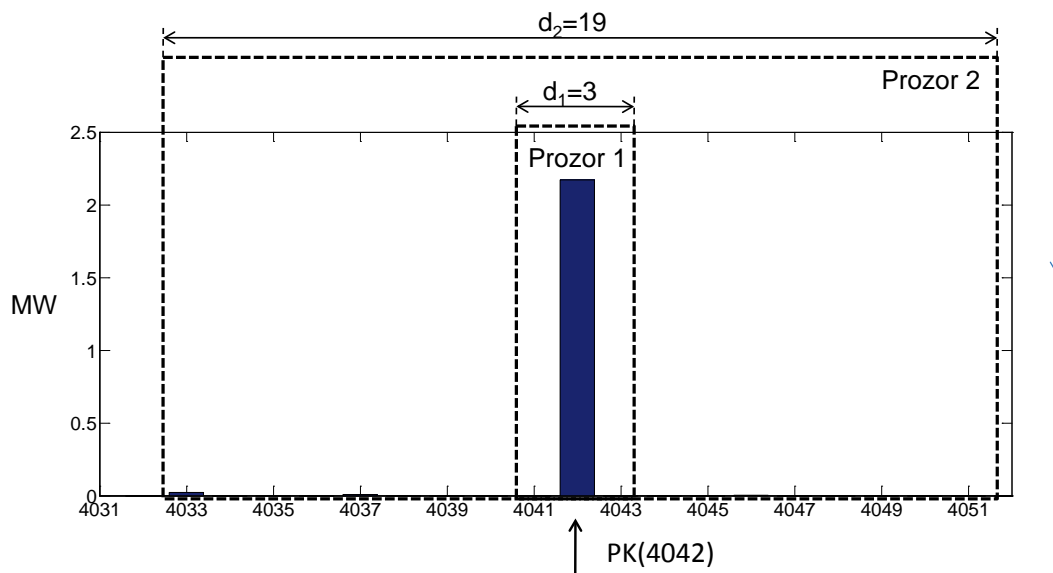
Slika 5.13. Niz $MICW[t]$ u intervalu koordinata [4150,4350].



Slika 5.14. Niz $MW[t]$ u intervalu koordinata [4150,4350].

5.4.2.3. Identifikacija pozicija artefakata sa pozicijama lokalnih maksimuma

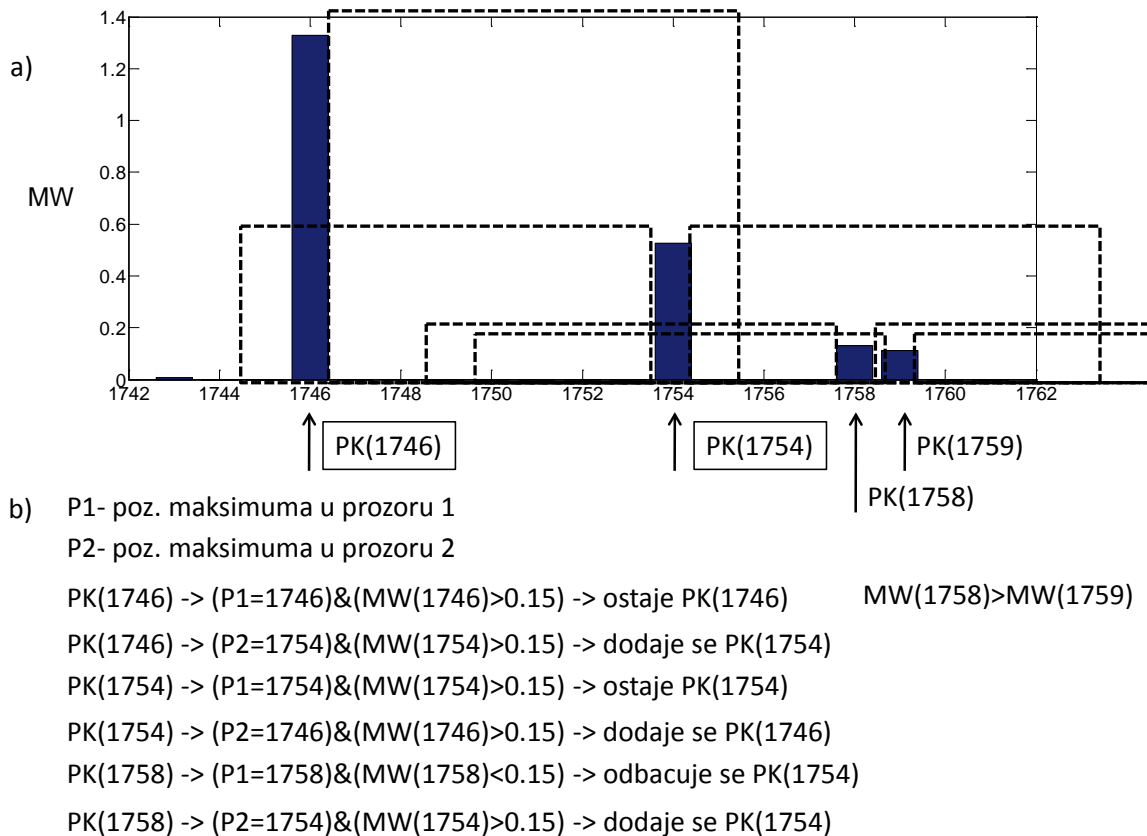
Detektovane pozicije kandidata u nizu $PK[p]$ ukazuju na naglu promenu sadržaja u njihovom okruženju. Verifikacija te pozicije ili pridruživanje nove stvarne pozicije nagle promene sadržaja, vrši se analizom okoline pozicije kandidata u nizu $MW[t]$. Okolina pozicije kandidata proverava se sa dva simetrična prozora različitih dimenzija, Slika 5.15.



Slika 5.15. Primer analize okoline kandidata i dodeljivanje lokalnog maksimuma artefaktu.

Prvi prozor (prozor1) ima širinu $d_1=3$ i proverava neposrednu okolinu pozicije kandidata, levu i desnu susednu komponentu niza. Proverava se vrednost maksimuma i ona mora biti veća od vrednosti $MW_{min_lok_max}=0.15$. Minimalna vrednost lokalnog maksimuma za verifikaciju utvrđena je empirijski, na osnovu eksperimenata. Data vrednost obezbeđuje da se verifikuju samo značajne nagle promene sadržaja. Kada je uslov vrednosti lokalnog maksimuma zadovoljen, njegova pozicija postaje nova pozicija kandidata. Ovaj prozor ima ulogu da zaštiti poziciju kandidata od mogućeg maskiranja lokalnog maksimuma dominantnim maksimumom u njenoj široj okolini. Drugi prozor (prozor2) širine $d_2=19$ ima ulogu da detektuje lokalni maksimum, čija je vrednost veća od lokalnog maksimuma detektovanog u prvom prozoru. Ako postoji takav lokalni maksimum i ako je njegova vrednost veća od $MW_{min_lok_max}$, pozicija lokalnog maksimuma uvrštava se

među verifikovane pozicije kandidata. Ovaj prozor obezbeđuje detekciju pozicije nagle promene sadržaja i u slučajevima kada je pozicija kandidata pomeren u odnosu na poziciju stvarne nagle promene sadržaja, van ivica prvog prozora. Kao što je već rečeno, ovaj se slučaj pojavljuje zbog degradacije artefakta kretanjem objekata ili pojavom specijalnih efekata unutar kadra. Na Slici 5.15 prikazan je primer verifikacije pozicije kandidata kada se unutar prvog prozora detektovani lokalni maksimum nalazi na poziciji kandidata, čija je vrednost veća od $MW_{min_lok_max}$. U široj okolini ne postoji dominantniji lokalni maksimum od detektovanog u prvom prozoru, pa se u ovoj proveru ne proširuje skup detektovanih pozicija kandidata. Pozicija kandidata $PK(4042)$ je verifikovana.



Slika 5.16. Provera okoline kandidata: a) prikaz analiziranih okolina kandidata, b) algoritam za prihvatanje lokalnih maksimuma i dodeljivanje artefaktima.

Na Slici 5.16(a) prikazan je postupak verifikacije pozicija kandidata za slučaj detekcije artifakata sa Slike 5.11. Analizirane su četiri pozicije kandidata: $PK(1746)$, $PK(1754)$, $PK(1758)$ i $PK(1759)$. Unijom rezultata, nakon provere okolina pozicija

kandidata i vrednosti lokalnih maksimuma, samo dve pozicije kandidata se verifikuju, $PK(1746)$ i $PK(1754)$. Na Slici 5.16(b) prikazan je postupak provere okoline pozicija kandidata. Verifikacija pozicija kandidata rezultira nizom $PKver[r]$ koji predstavlja uniju verifikovanih pozicija kandidata nakon izvršenih provera okolina inicijalnih pozicija kandidata za ceo video segment. Zbog neodređenosti deformisanih artefakata broj detektovanih pozicija naglih promena veći je od broja detektovanih artefakata. Detekcija naglih promena kadrova izvršiće se u narednim fazama algoritma nakon dva procesa klasifikacije. U ovoj fazi algoritma obezbeđen je polazni skup frejmova kandidata za proces klasifikacije koji sledi.

5.5. Gruba klasifikacija kandidata

Gruba klasifikacija kandidata sastoji se iz tri dela. Prvi deo predstavlja poređenje globalnih histograma boje, na verifikovanoj poziciji kandidata. Drugi deo klasifikacije predstavlja proveru okoline lokalnih maksimuma u multifraktalnom alfa vejevlet nizu na poziciji verifikovanih kandidata. Treći deo klasifikacije predstavlja proveru oblika artefakta na verifikovanoj poziciji kandidata. Rezultat grube klasifikacije može biti prihvatanje ili odbijanje pozicije kandidata. U nastavku teksta verifikovane pozicije kandidata biće nazivane pozicijama kandidata.

5.5.1. Poređenje globalnih histograma boje

Poređenje globalnih histograma boje na poziciji kandidata obavlja se izračunavanjem korelacije između globalnih histograma boje frejmova na poziciji kandidata i narednog frejma u odnosu na kandidata, izraz (5.8). Formira se niz $PKverKOR[r]$. Vrednost korelacije poredi se sa predefinisanim pragom $KORMAX = 0.98$, izraz (5.9). Za potrebe klasifikacije formiran je niz $PKverKlas$. Ukoliko je vrednost korelacije veća od praga $KORMAX$ pozicija kandidata se odbija ($PKverKlas=0$), u suprotnom prihvata ($PKverKlas=1$), izraz (5.9). Na osnovu vrednosti iz niza $PKverKlas$

redukuje se niz $PKver[r]$ i formira se novi niz $PKver1[r1]$ koji sadrži samo prihvaćene pozicije kandidata u ovom delu klasifikacije.

$$PKverKOR_i = corr(H_{PKver_i}, H_{PKver_{i+1}}), \quad i=1...r \quad (5.8)$$

$$PKverKlas_i = \begin{cases} 1 & PKverKOR_i < KORMAX \\ 0 & PKverKOR_i > KORMAX \end{cases}, \quad i=1...r \quad (5.9)$$

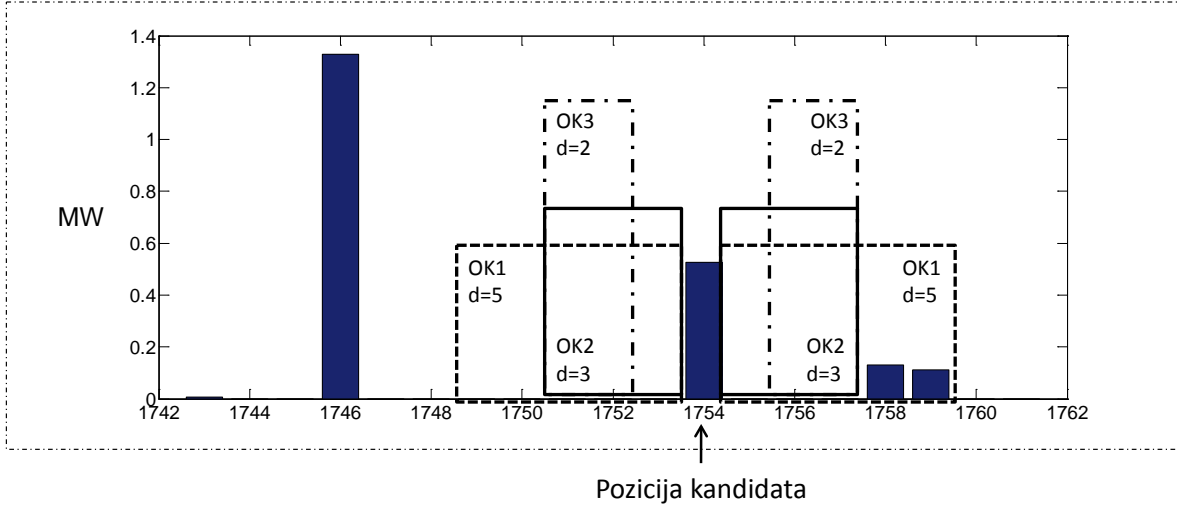
Vrednost predefinisano g praga $KORMAX$ određena je empirijski na osnovu analize velikog broja frejmova ($>1.500.000$). Na ovaj način izvršeno je ograničavanje dozvoljene korelacije graničnih frejmova na prelazu između dva kadra. Visoko korelisani kadrovi biće ispušteni. U cilju da se postigne veća preciznosti klasifikacije, prag klasifikatora je utvrđen na prikazanoj vrednosti.

5.5.2. Provera okoline pozicije kandidata u alfa domenu multifraktalne analize

U ovom delu grube klasifikacije analizira se alfa niz vejevlet korelacionog niza $MW[t]$. Na pozicijama kandidata, sadržanim u nizu $PKver1[r1]$, proverava se odnos lokalnog maksimuma i vrednosti u njegovoj okolini, Slika 5.17. Analizira se više okolina lokalnog maksimuma. Definisano je klasifikator kao količnik lokalnog maksimuma i maksimuma u njegovoj okolini. Predefinisano je prag klasifikacije $KP=3$. Za potrebe ove klasifikacije formiran je niz $Pkver1Klas[r1]$, izrazi (5.10)-(5.13). Sve pozicije kandidata čiji je količnik veći od praga KP prihvataju se ($Pkver1Klas_i=1$), u suprotnom odbijaju se ($Pkver1Klas_i=0$). Analizom vrednosti u nizu $Pkver1Klas[r1]$ redukuje se niz $PKver1[r1]$ i formira se niz $PKver2[r2]$ koji se sastoji od prihvaćenih pozicija kandidata u ovoj fazi klasifikacije.

Vrednost praga KP koja se koristi u ovom koraku klasifikacije određena je empirijski, analizom velikog broja frejmova ($>1.500.000$). Lokalni maksimumi u alfa nizu koji odgovaraju stvarnim promenama kadrova imaju znatno veću vrednost u odnosu na svoju okolinu. Međutim, kada unutar kadra postoje brza kretanja pojavljuju se dodatni

lokalni maksimumi u okolini lokalnog maksimuma na poziciji kandidata koji ga mogu maskirati. Iz ovog razloga proveravaju se različite okoline lokalnog maksimuma. Ovakav pristup je omogućio da se efikasnije detektuju stvarne nagle promene kadrova, a sa druge strane da se odstrane sve nagle promene koje imaju gradijentnu strukturu koja se znatno razlikuje od strukture koja zadovoljava definisane uslove.



Slika 5.17. Prikaz okolina lokalnog maksimuma na poziciji kandidata koje se proveravaju u alfa nizu vejevlet korelacionog niza

$$KOL1_i = MW(PKver1_i) / (\max([MW(PKver1_i - 5 : PKver1_i - 1)] \cup [MW(PKver1_i + 1 : PKver1_i + 5)])), i = 1 : r1 \quad (5.10)$$

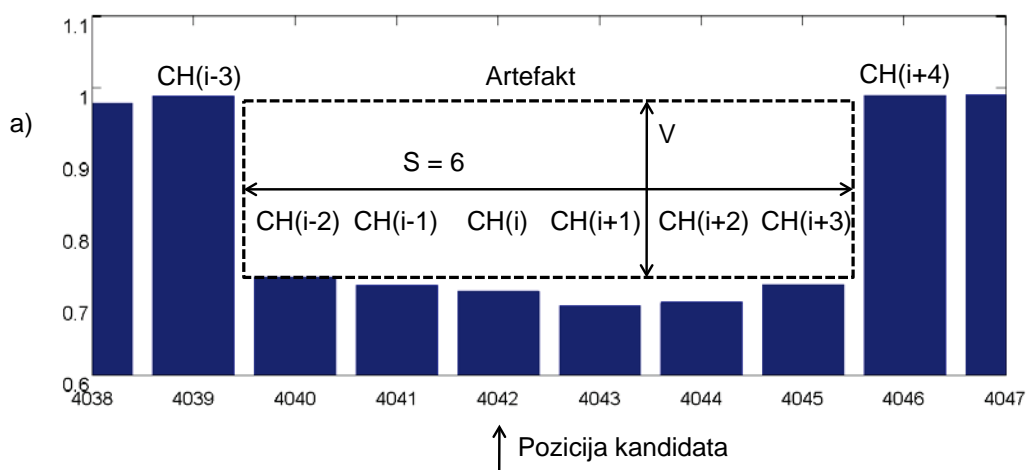
$$KOL2_i = MW(PKver1_i) / (\max([MW(PKver1_i - 3 : PKver1_i - 1)] \cup [MW(PKver1_i + 1 : PKver1_i + 3)])), i = 1 : r1 \quad (5.11)$$

$$KOL3_i = MW(PKver1_i) / (\max([MW(PKver1_i - 3 : PKver1_i - 2)] \cup [MW(PKver1_i + 2 : PKver1_i + 3)])), i = 1 : r1 \quad (5.12)$$

$$PKver1Klas_i = \begin{cases} 1 & (KOL1_i > KP) ili (KOL2_i > KP) ili (KOL3_i > KP) \\ 0 & ostalo \end{cases}, i = 1 : r1 \quad (5.13)$$

5.5.3 Provera oblika artefakta

U ovom koraku grube klasifikacije pozicije kandidata, sadržane u nizu $PKver2[r_2]$ podvrgavaju se proveru oblika artefakta. Pri osnovnoj detekciji zahvaćene su sve nagle promene čiji oblici artefakata su varirali po dimenziji. U ovoj klasifikaciji traži se vrlo specifičan oblik koji odgovara naglim promenama kadrova. Na Slici 5.18(a) prikazan je jedan reprezentativni artefakt, a na Slici 5.18(b) prikazani su uslovi prihvatanja. Uslovi prihvatanja su malo izmenjeni, zbog postojanja artefakata koji imaju degradirane ivice kao posledicu brzih kretanja u okolini samog prelaza iz kadra u kadar.

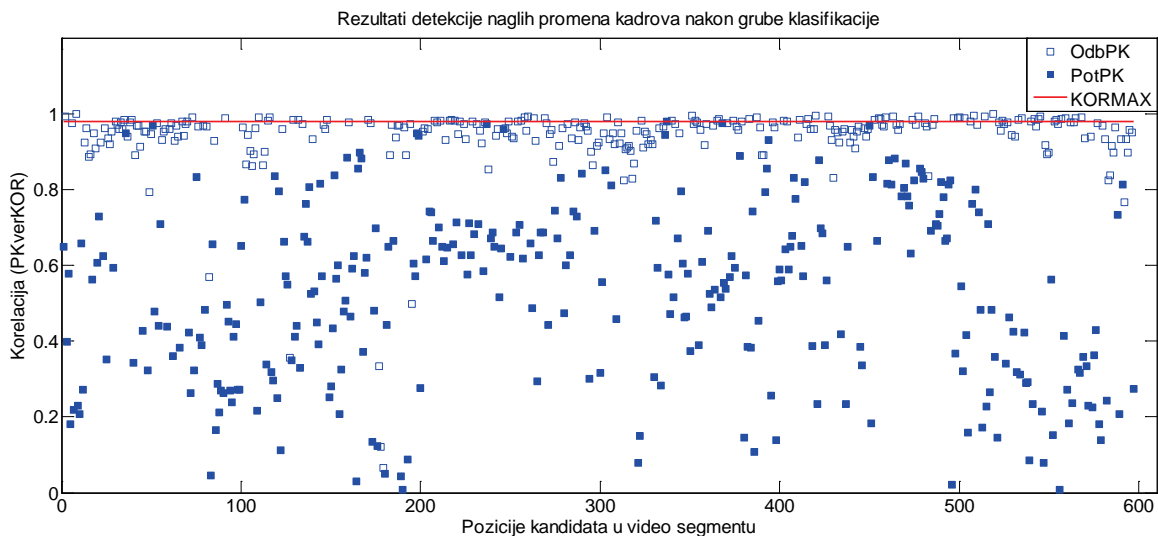


- b) $V = \min(\text{CH}(i-3), \text{CH}(i+4)) - \max(\text{CH}(i-2), \text{CH}(i-1), \text{CH}(i), \text{CH}(i+1), \text{CH}(i+2), \text{CH}(i+3))$
 $S = 6$ (komponenti)
 $(V > 0.001) \ \& \ (S = 6)$ uslov prihvatanja artefakta
- $V1 = \max(\text{CH}(i-3), \text{CH}(i+4)) - \max(\text{CH}(i-2), \text{CH}(i-1), \text{CH}(i), \text{CH}(i+1), \text{CH}(i+2), \text{CH}(i+3))$
 $S = 6$ (komponenti)
 $(V1 > 0.001) \ \& \ (S = 6)$ uslov prihvatanja artefakta

Slika 5.18. Provera oblika artefakta, a) slika jednog artefakta,
 b) uslovi prihvatanja artefakta

Kreiran je niz $PKver2Klas[r_2]$ za potrebe ove klasifikacije. Ako artefakt na poziciji kandidata iz niza $PKver2[r_2]$ zadovoljava definisane uslove, pozicija kandidata se prihvata ($PKver2Klas_i=1$), u suprotnom odbija se ($PKver2Klas_i=0$). Na osnovu informacija iz niza

$PKver2Klas[r_2]$ redukuje se niz $PKver2[r_2]$ samo na pozicije kandidata koje su prihvaćene u ovom koraku klasifikacije, od kojih se formira novi niz $PKver3[r_3]$ koji se prosleđuje dalje u proces fine klasifikacije. Na Slici 5.20 prikazani su rezultati detekcije nakon grube klasifikacije pozicija kandidata.



Slika 5.19. Rezultat detekcije naglih promena kadrova na sedmom segmentu videa "A Good Year", *OdbPK*-odbijene pozicije kandidata, *PotPK*-potencijalne pozicije naglih promena kadrova, *KORMAX*-klasifikacioni prag korelacije

5.6. Fina klasifikacija kandidata

U ovom delu klasifikacionog procesa pozicije kandidata iz niza $PKver3[r_3]$ podvrgavaju se dodatnom procesu klasifikacije koji se sastoji iz dva koraka. Prvi korak predstavlja poređenje globalnog histograma boje između frejma na poziciji kandidata i njegovih okolnih frejmova. Drugi korak predstavlja proveru pozicije kandidata primenom *flash* detektora.

5.6.1. Fina klasifikacija kandidata primenom empirijskih pravila

U ovom koraku klasifikacije izračunavaju se međusobne korelacije globalnih histograma boje između frejmova na poziciji kandidata i susednim pozicijama, izrazi (5.27)-(5.29). Korelacija se izračunava između frejma kandidata i prethodnog frejma (A), između frejma kandidata i narednog frejma (B), između prethodnog i narednog frejma u odnosu na frejm kandidat (C). Vrednosti izračunatih korelacija se koriste za odlučivanje u nizu empirijski utvrđenih pravila, izrazi (5.14)-(5.20). Za potrebe klasifikacije formira se niz $PKver3Klas[r_3]$.

$$A_i = corr(H_{PKver_i-1}, H_{PKver_i}), i = 1 \dots r_4 \quad (5.14)$$

$$B_i = corr(H_{PKver_i}, H_{PKver_{i+1}}), i = 1 \dots r_4 \quad (5.15)$$

$$C_i = corr(H_{PKver_{i-1}}, H_{PKver_{i+1}}), i = 1 \dots r_4 \quad (5.16)$$

$$srvred_i = (A_i + B_i) / 2, i = 1 \dots r_4 \quad (5.17)$$

$$param_i = |A-B| \cdot srvred_i, i = 1 \dots r_4 \quad (5.18)$$

$$R_i = (A_i < B_i) \text{ and } (param_i < 0.3), i = 1 \dots r_4 \quad (5.19)$$

$$PKver3Klas_i = \begin{cases} 0 & (A_i > B_i)(srvred_i > 0,97) \\ 0 & (R_i) \text{ and } (A_i < B_i) \text{ and } (srvred_i > 0,85) \\ 0 & (R_i) \text{ and } (A_i < B_i) \text{ and } (A_i < C_i) \text{ and } (srvred_i < 0,85), i = 1 : r_4 \\ 1 & (R_i) \text{ and } (A_i < B_i) \text{ and } (A_i > C_i) \text{ and } (srvred_i < 0,85) \\ 1 & \text{ostalo} \end{cases} \quad (5.20)$$

Empirijska vrednosti i pravila definisana izrazima (5.18)-(5.20) određeni su eksperimentalnim putem. Analizom rezultata uočena su određena pravila ponašanja vrednosti korelacija, što je predstavljalo osnovu za definisanje određenih pravila za klasifikaciju. Pravila prihvatanja pozicije kandidata definisana su izrazom (5.20), tako da vrednosti niza $PKver3Klas[r_3]$ pokazuju da li je pozicija kandidata prihvaćena ($PKver3Klas_i=1$) ili je odbijena ($PKver3Klas_i=0$). Na osnovu vrednosti niza

$PKver3Klas[r_3]$ vrši se redukovanje niza $PKver3[r_3]$ izuzimanjem odbačenih pozicija kandidata, a od prihvaćenih formira se novi niz $PKver4[r_4]$ koji se prosleđuje u sledeći korak fine klasifikacije. Sledeći korak u finoj klasifikaciji predstavlja detekcija fleša, odnosno detekcija nagle promena osvetljenja na frejmu.

5.6.2. Fina klasifikacija kandidata primenom *flash* detektora

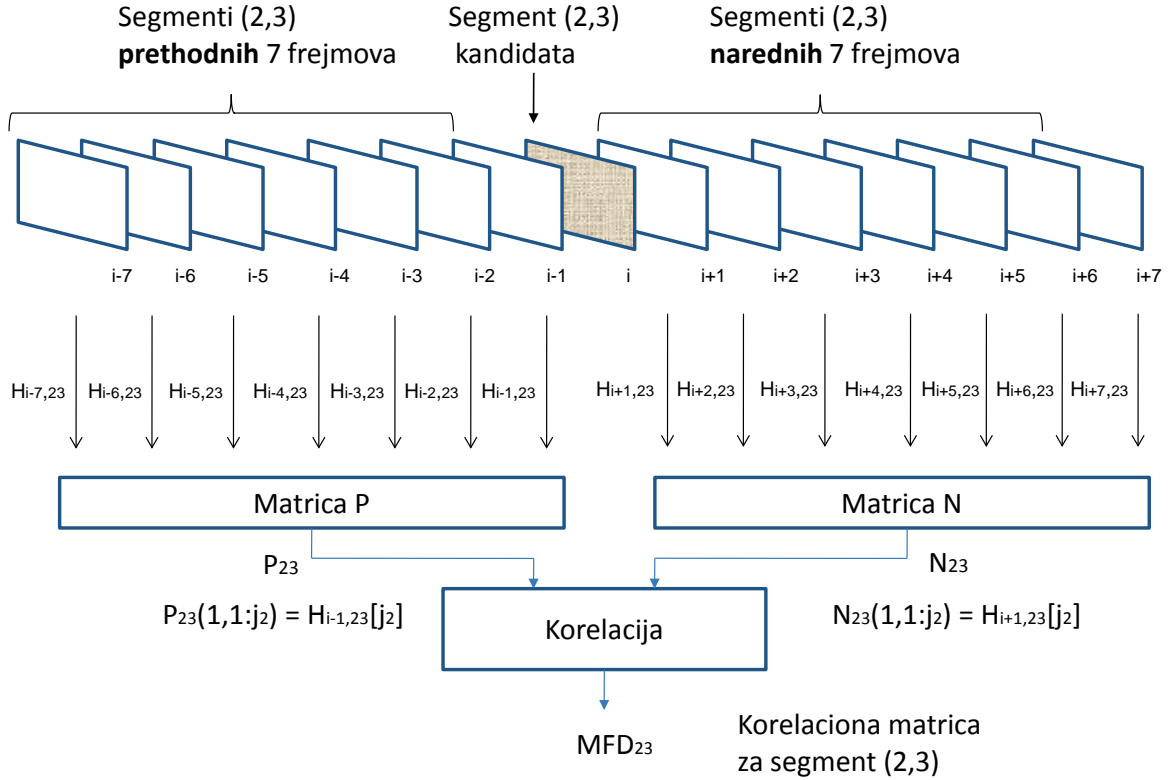
Flash detektor predstavlja algoritam za detekciju naglih promena osvetljenja koja ne degradiraju znatno vrednosti intenziteta piksela na čitavoj površini frejma. S obzirom da deo frejma nije znatno promenjen moguće je utvrditi lažnu detekciju nagle promene kadra. Sa druge strane fleš detektor omogućava i detekciju pomeraja objekta koja je nagla, ali ne utiče podjednako na vrednosti piksela na celoj površini frejma. Proces klasifikacije primenom fleš detektora odvija se u tri koraka. U prvom delu vrši se izračunavanje korelacione matrice segmenata frejmova kandidata. U drugom delu izračunava se klasifikacioni parametar za svakog kandidata i u trećem delu vrši se klasifikacija kandidata na osnovu klasifikacionog parametra.

5.6.2.1. Izračunavanje korelacionih matrica segmenata frejmova

U ovom koraku klasifikacije primenom fleš detektora vrši se izračunavanje korelacione matrice MFD za svaki segment frejmova, koji se nalaze na lokacijama detektovanih kandidata. Lokacije ovih frejmova sadržane su u nizu $PKver4[r_4]$. Algoritam za izračunavanje korelacione matrice predstavlja modifikaciju algoritma fleš detektora preuzetog iz literature [PET10].

Segmenti frejmova kandidata poklapaju se sa segmentima koji su korišćeni za izračunavanje obeležja boje i teksture. Testiraju se frejmovi u okolini posmatranog frejma. Okolina je definisana sa prethodnih $k_1=7$ frejmova i narednih $k_2=7$ frejmova. Za sve frejmove okoline posmatra se isti segment. Na Slici 5.20 predstavljen je primer formiranja niza segmenata od segmenata (2,3) frejmova okoline frejma na poziciji kandidata (pozicija i). Kolor histogrami segmenata prethodnih frejmova iz H_{FD} obeležja formiraju matricu

$P[k_1, j_2]$, jednačina (5.21). Analogno, kolor histogrami segmenata prethodnih frejmova iz H_{FD} obeležja formiraju matricu $N[k_2, j_2]$, jednačina (5.22).



Slika 5.20. Izračunavanje korelacione matrice MFD_{23} za segment(2,3) frejma kandidata.

$$P_{23}(k_1, 1:j_2) = H_{i-k_1}[j_2], k_1 = 1:7 \quad (5.21)$$

$$N_{23}(k_2, 1:j_2) = H_{i+k_2}[j_2], k_2 = 1:7 \quad (5.22)$$

Korelaciona matrica $MFD(k_1, k_2)$ za segment na poziciji (2,3) dobija se izračunavanjem međusobne korelacije svih vrsta matrica P i N , jednačina (5.23).

$$MFD_{23}(k_1, k_2) = \text{corr}(P_{23}(k_1, 1:j_2), N_{23}(k_2, 1:j_2)) \quad (5.23)$$

Izračunavanje korelacione matrice daje uvid o sličnosti segmenata na istoj poziciji u okviru frejmova u definisanoj okolini frejma kandidata za naglu promenu kadra. Korelacija se izračunava na osnovu vektora obeležja. Visoka korelacija ukazuje na sličnost

dva frejma u okolini frejma kandidata, što može biti posledica naglog pokreta u kadru, kretanju velikih objekata u kadru ili nagloj promeni osetljenja. Nagla promena kadrova manifestuje se kao mala korelacija testiranih segmenata u okolini pozicije frejma kandidata.

5.6.2.2. Izračunavanje klasifikacionog parametra

Izračunavanje korelacione matrice obavlja se za sve segmente frejma kandidata. Sledeći korak je izračunavanje srednje vrednosti za svaku korelacionu matricu segmenata, od kojih se formira matrica srednjih vrednosti MFD_{SV} , jednačina (5.24).

$$MFD_{SV}(s_1, s_2) = \text{mean}(MFD_{s_1s_2}), s_1 = 1:4, s_2 = 1:4 \quad (5.24)$$

Srednja vrednost korelacionih matrica izračunava se da bi se izbegla pojava neke visoke vrednosti korelacije, slučajno sličnih segmenata, koja bi maskirala ostale vrednosti korelacije u matrici. U matrici MFD_{SV} detektovana je usrednjena sličnost segmenata frejma kandidata sa segmentima frejmova u njegovoj predefinisanoj okolini. Klasifikacioni parametar MFD_{max} definisan je kao maksimalna vrednost u matrici srednjih vrednosti korelacije MFD_{SV} , jednačina (5.25). Vrednost ovog parametra ukazuje na prirodu frejma kandidata, odnosno na to da li taj frejm stvarno predstavlja početak novog kadra.

$$MFD_{max} = \max(MFD_{SV}) \quad (5.25)$$

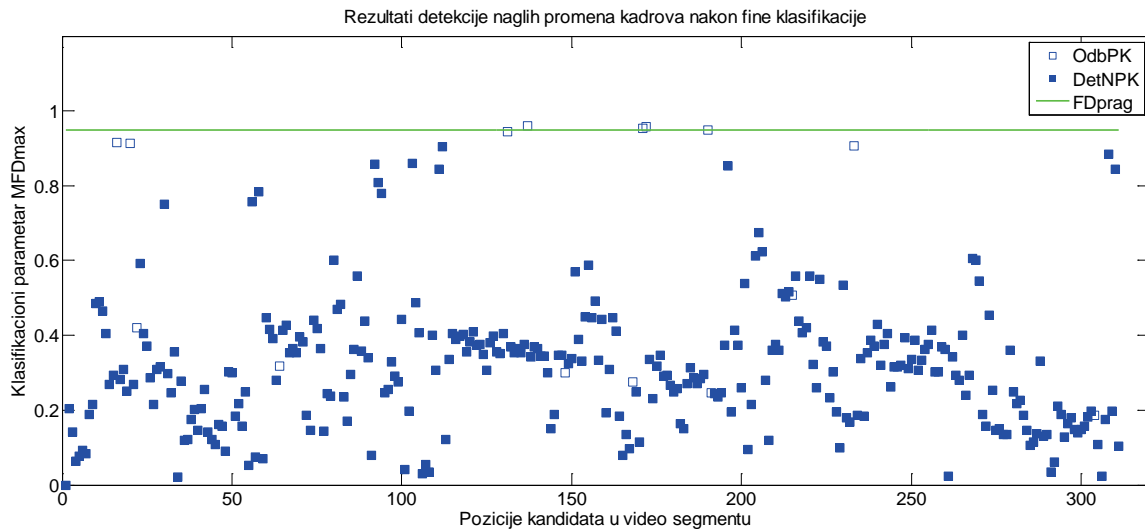
5.6.2.3. Klasifikacija kandidata

Klasifikacija kandidata se izvodi na osnovu vrednosti klasifikacionog parametra MFD_{max} primenom tehnike praga. Definisana je vrednost praga klasifikacije $FD_{prag}=0.95$. Vrednost praga je određena empirijski i ima prilično visoku vrednost zbog naglih promena kadrova visoke korelisanosti sadržaja. Ovako postavljen prag treba da obezbedi da i nagle promene kadrova između frejmova sličnog sadržaja zadovolje uslov klasifikacije i da budu klasifikovane kao stvarne nagle promene. Analizom velikog broja frejmova utvrđeno je da je mali broj naglih promena kadrova sa visoko korelisanim

sadržajima, međutim njihovo ispuštanje utiče na performanse sistema pa se ostalo pri utvrđenoj vrednosti praga klasifikacije.

Za svaku poziciju kandidata izračunava se vrednost $MFDmax_i$ koja se poredi sa vrednošću klasifikacionog praga, Slika 5.21. Za potrebe klasifikacije formiran je niz $PKver4Klas[r_4]$ čije vrednosti pokazuju da li je pozicija kandidata prihvaćena ($PKver4Klas_i=1$) ili odbijena ($PKver4Klas_i=0$). Način na koji se izračunava komponenta niza $PKver4Klas_i$ data je izrazom (5.26). Ispuštanjem neprihvaćenih pozicija kandidata redukuje se niz $PKver4[r_4]$ i od prihvaćenih pozicija kandidata formira se novi niz $Pkver5[r_5]$. Novi niz predstavlja rezultat fine klasifikacije pozicija kandidata i sam tim konačni rezultat detekcije naglih promena kadrova, slika 5.21.

$$PKver4Klas_i = \begin{cases} 1 & PKver4_i < MFDmax \\ 0 & \text{ostalo} \end{cases}, \quad i=1:r_4 \quad (5.26)$$



Slika 5.21. Rezultat detekcije naglih promena kadrova na sedmom segmentu videa "A Good Year", *OdbPK*-odbijene pozicije kandidata, *DetNPK*-detektovane nagle promene kadrova, *FDprag*-klasifikacioni prag *flash* detektora

6. Eksperimentalni rezultati testiranja predloženog algoritma

Testiranje predloženog algoritma za detekciju naglih promena kadrova izvedeno je na video test sekvencama različitog sadržaja i dinamike. Test sekvence razlikovale su se po dužini trajanja, žanru i tehnologiji kreiranja frejmova. Efikasnost detekcije je merena standardnim metodama objektivne i subjektivne procene kvaliteta. U nastavku teksta dat je detaljan opis video test sekvenci, opisane su metode merenja performansi algoritma i prezentovani su rezultati merenja istih. Rezultati testiranja izloženi su kroz tabelarni prikaz i izvršeno je poređenje performansi postojećih algoritama za detekciju naglih promena kadrova i predloženog algoritma. Detaljna diskusija predstavljenih rezultata testiranja i poređenja data je na kraju ovog poglavlja.

6.1. Video test sekvence

Video test sekvence na kojima su testirane performanse predloženog algoritma podeljene su u nekoliko grupa. Svaka od grupa test sekvenci omogućila je testiranje algoritma pod različitim uslovima, čime se obezbedila adekvatna provera performansi

sistema za detekciju naglih promena kadrova. Pri testiranju korišćene su filmske video sekvence kao video sadržaji visokog kvaliteta postprodukcije, i koji predstavljaju umetnički izraz reditelja. Drugi tip video sekvenci preuzet je sa distributivnih televizijskih mreža da bi se algoritam za detekciju testirao na svakodnevnom video sadržaju koji se nudi korisnicima, i koji je tehnički bogat velikim brojem različitih tranzicija između kadrova. Video sekvence svih test grupa izdvojene su u *MPEG-4* video formatu.

Prvu grupu video test sekvenci čine četiri dugometražna igrana filma različitih žanrova. Ova grupa test sekvenci kreirana je da bi se proverile performanse predloženog algoritma na video sekvencama dovoljno dugog trajanja, u okviru kojih se tokom trajanja sekvence često menja dinamika scene, kao i sadržaj između frejmova. Rezultat testiranja bi trebalo da pokaže usrednjen rezultat detekcije za veliki broj frejmova. Svaka od video sekvenci u ovoj test grupi predstavlja jednu klasu video sekvenci. Prvu klasu čine akcione video sekvence sa elementima drame i fikcije. Ovu klasu čine video sadržaji sa velikim brojem akcionih scena i specijalnih efekata. Akcione scene podrazumevaju brza pomeranja kamera i objekata unutar kadrova, a prisustvo specijalnih efekata podrazumeva nagle promene sadržaja unutar kadrova snimljenih kamerom (eksplozije, nagle promene osvetljenja) ili promene dodate postobradom frejmova.

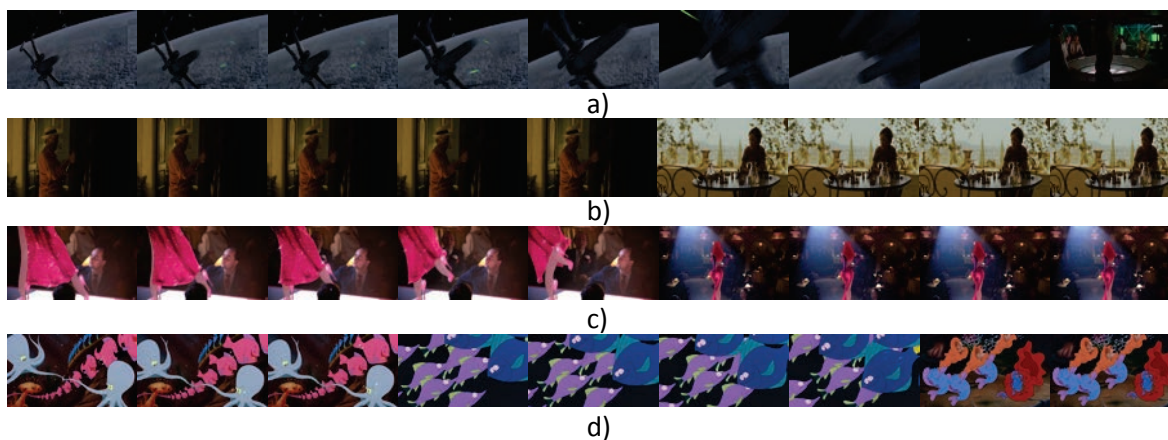
Drugu klasu video sekvenci predstavljaju dramske statične video sekvence. Ova klasa podrazumeva spore tranzicije radnje, spora pomeranja objekata i kamera. Treću klasu video sekvenci čine kombinovane video sekvence, koje su u stvari realne video sekvence sa dodatnom animacijom unutar frejmova. Ova klasa podrazumeva realni video sa ugrađenom računarskom grafikom ili sa dodatkom klasične animacije.

Poslednju, četvrtu klasu video sekvenci čine video sadržaji sa klasičnom animacijom, odnosno animirani filmovi. Ova klasa podrazumeva drugačiji stil tranzicije radnje, odnosno posebnu tehnologiju izrade frejma koja autoru daje dosta slobode pri kreiranju kadrova. Frejmovi iz ove klase značajno se razlikuju od prethodnih klasa na osnovu boje i teksture. Detaljan opis ovih video sekvenci, dužine trajanja, dimenzije frejmova i broj frejmova prikazani su u Tabeli 6.1. Na Slici 6.1 prikazani su primeri

frejmova video sekvenci iz svake od definisanih klasa. Svaki od primera, u obimu frejmova koji su prikazani na slici, zahvata jednu do dve nagle promene sadržaja koje predstavljaju nagle promene kadrova. Na slici se može primetiti razlika u primenjenoj tehnologiji za kreiranje sadržaja frejmova za različite klase video sekvenci.

Tabela 6.1. Prva grupa video test sekvenci (igrani filmovi)

Oznaka	Film	Reditelji	Tehnologija	Žanr	Godina	Format	Dužina (H:m:s)	Broj frejm.	Veličina frejma (p.)
Film1	Star Wars: Episode 1	George Lucas	Realni video & Ugrađene slike	Akcija, Fantazija	1977	H.264	2:17:14	187.118	352x288
Film2	A Good Year	Ridley Scott	Realni video	Romantična komedija, drama	2006	H.264	1:57:34	169.132	588x246
Film3	Who Framed Roger Rabbit	Robert Zemeckis	Realni video & Crtana animacija	Krimi komedija, fantazija	1988	H.264	1:39:40	140.021	544x304
Film4	Little Marmeid	Ron Clements i John Musker	Crtana animacija	Muzička komedija, fantazija	1989	H.264	1:22:39	118.901	640x352



Slika 6.1. Primeri frejmova iz četiri klase video sekvenci koje su korišćene pri testiranju predloženog algoritma: a) *Film1*, b) *Film2*, c) *Film3*, d) *Film4*.

Drugu grupu video test sekvenci čine sekvence kraćeg trajanja razvrstane po podgrupama. Svaka podgrupa odgovara jednoj od četiri definisane klase videa (žanru) i unutar svake grupe nalazi se po deset video sekvenci. Sekvence su birane vizuelno da odgovaraju žanrovima, pri čemu su izbegavane početne i završne sekvence. Izuzetak su kratki animirani filmovi koji su preuzeti u punom trajanju. Testiranje sistema za detekciju

na ovoj grupi test sekvenci trebalo bi da pokaže uticaj klasa kojima video sekvence pripadaju na performanse detekcije. U Tabelama 6.2-6.5 dat je detaljan opis video sekvenci, njihova raspodela po klasama, tačan vremenski okvir svake sekvence unutar filmova iz kojih su izdvojene, dimenzija frejmova i broj frejmova unutar sekvence.

Tabela 6.2. Podgrupa akcionih video test sekvenci

Oznaka	Film	Godina	Reditelji	Produkcija	Vremenski okvir [h:m:s]		Dimenzija frejma [p.]	Broj frejm.
					Početak	Kraj		
SEK1	Starwars: Ep. 1	1977	George Lucas	Lucas Film	01:20:00	01:33:19.960	320x240	20000
SEK2	Mission: Impossible 3	2006	J.J. Abrams	Paramount	00:55:33.333	01:09:26.625	320x240	20000
SEK3	The Matrix Revolutions	2003	Andy&Lana Wachowski	Warner Bros	01:09:30.838	01:23:24.963	320x240	20000
SEK4	Minority report	2002	Steven Spielberg	Twentieth Century Fox	01:09:30.838	01:23:24.963	320x240	20000
SEK5	Band of Brothers, Ep. 3 Carentan	2001	Mikael Salomon	HBO	00:13:54.168	00:27:48.293	320x240	20000
SEK6	Pearl Horbor	2001	Michael Bay	Touchstone Pictures	01:31:49.176	01:45:43.302	320x240	20000
SEK7	Battleship	2012	Peter Berg	Universal Pictures	01:28:53.333	01:39:59.967	320x240	20000
SEK8	The A-Team	2010	Joe Carnahan	Twentieth Century Fox	00:53:20	01:06:39.960	320x240	20000
SEK9	The 13th warrior	1999	John McTiernan	Touchstone Pictures	01:23:20	01:36:39.960	320x240	20000
SEK10	The Born Ultimatum	2007	Paul Greengrass	Universal Pictures	00:53:20	01:06:39.960	320x240	20000

Tabela 6.3. Podgrupa dramskih video test sekvenci

Oznaka	Film	Godina	Reditelji	Produkcija	Vremenski okvir [h:m:s]		Dimenzija frejma [p.]	Broj frejm.
					Početak	Kraj		
SEK11	A Good Year	2006	Ridley Scott	Fox 2000 Pictures	00:53:20	01:06:39.960	320x240	20000
SEK12	Trading Places	1983	John Landis	Cinema Group Ventures	01:06:40	01:19:59.960	320x240	20000
SEK13	Boomerang	1992	Reginald Hudlin	Paramount Pictures	00:40:00	00:53:19.960	320x240	20000
SEK14	Sweet November	2001	Pat O'Connor	Warner Bros.	00:40:00	00:53:19.960	320x240	20000
SEK15	Something's Gotta Give	2003	Nancy Meyers	Columbia Pictures Corporation	00:55:36.670	01:09:30.796	320x240	20000
SEK16	Notting Hill	1999	Roger Michell	Polygram Filmed Entertain.	01:28:29.434	01:42:23.570	320x240	20000
SEK17	Pretty Woman	1990	Garry Marshall	Touchstone Pictures	00:27:48.335	00:41:42.461	320x240	20000
SEK18	Working Girl	1988	Mike Nichols	Twentieth Century Fox	01:15:47.960	01:29:04.085	320x240	20000
SEK19	Gigi	1958	Vincente Minnelli	Metro-Goldwyn-Mayer	00:55:56.670	01:09:30.796	320x240	20000
SEK20	Love Actually	2003	Richard Curtis	Universal Pictures	00:26:40	00:39:59.960	320x240	20000

Tabela 6.4. Podgrupa kombinovanih video test sekvenci

Oznaka	Film	Godina	Reditelji	Produkcija	Vremenski okvir [h:m:s]		Dimenzija frejma [p.]	Broj frejm.
					Početak	Kraj		
SEK21	Who Framed Roger Rabbit	1988	Robert Zemeckis	Touchstone Pictures	00:26:40	00:39:59.960	320x240	20000
SEK22	Tron: Legacy	2010	Joseph Kosinski	Walt Disney Pictures	01:43:40	01:56:59.960	320x240	20000
SEK23	Casper	1995	Brad Silberling	Universal Pictures	00:26:40	00:39:59.960	320x240	20000
SEK24	The Mask	1994	Charles Russell	New Line Cinema	00:26:40	00:39:59.960	320x240	20000
SEK25	Dragonhearth	1996	Rob Cohen	Universal Pictures	00:53:20	01:06:39.960	320x240	20000
SEK26	Avatar	2009	James Cameron	Twentieth Century Fox	01:09:30.838	01:23:24.963	320x240	20000
SEK27	Wall-e	2008	Andrew Stanton	Pixar Animation Studios	01:06:40	01:19:59.960	320x240	20000
SEK28	Jumanji	1995	Joe Johnston	TriStar Pictures	00:53:20	01:06:39.960	320x240	20000
SEK29	Flight of the Navigator	1986	Randal Kleiser	Walt Disney Pictures	00:40:00	00:53:19.960	320x240	20000
SEK30	Ghostbusters 2	1989	Ivan Reitman	Columbia Pictures Corporation	00:27:48.335	00:41:42.544	320x240	20000

Tabela 6.5. Podgrupa animiranih video test sekvenci

Oznaka	Film	Godina	Reditelji	Produkcija	Vremenski okvir [h:m:s]		Dimenzija frejma [p.]	Broj frejm.
					Početak	Kraj		
SEK31	Little Mermaid	1989	Ron Clements John Musker	Walt Disney Pictures	00:27:48.335	00:41:42.461	320x240	20000
SEK32	Gullivers Travels	1939	Dave Fleischer	Fleischer Studios	00:26:40	00:40:00	320x240	20000
SEK33	Mickey Mouse: Boat Builders	1938	Ben Sharpsteen	Walt Disney Productions	00:00:00	00:07:20.320	320x240	11009
SEK34	Mulan	1998	Tony Bancroft, Barry Cook	Walt Disney Feature Animation	00:53:20	01:06:40	320x240	20000
SEK35	Pink Panther: The Pink Blueprint	1966	Hawley Pratt	DePatie-Freleng Enterprises (DFE)	00:00:00	00:06:14.541	320x240	9364
SEK36	The Princ Of Egypt	1998	Brenda Chapman, Steve Hickner, Simon Wells	DreamWorks Animation	01:06:40	01:20:00	320x240	20000
SEK37	Aladdin	1992	Ron Clements, John Musker	Walt Disney Pictures	00:53:22	01:06:39.960	320x240	20000
SEK38	The Jungle Book	1967	Wolfgang Reitherman	Walt Disney Productions	00:26:40	00:39:59.960	320x240	20000
SEK39	Tom And Jerry: Pent-House Mouse	1963	Chuck Jones, Maurice Noble	Metro-Goldwyn-Mayer (MGM)	00:00:00	00:07:21.00	320x240	11022
SEK40	Robin Hood	1973	Wolfgang Reitherman	Walt Disney Productions	00:27:48.335	00:41:42.544	320x240	20000

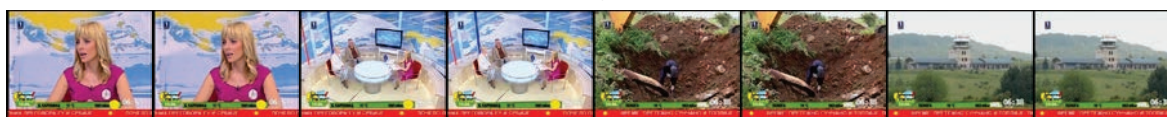
Treću grupu video test sekvenci čine sekvence sportskog programa lokalne sportske televizije i sekvence kombinovanog informativnog programa javnog televizijskog servisa. Sportski program se sastoji od dva različita sportska događaja, fudbalske i košarkaške utakmice. Pored navedenih sadržaja, ova test grupa sekvenci sadrži i reklame sa raznim ugrađenim grafičkim sadržajima. Video sekvence ove test grupe sadrže veliki broj različitih gradijentnih tranzicija kadrova i veliki broj tranzicija kadrova u uslovima visoke korelisanosti sadržaja između dva kadra. Testiranje predloženog algoritma na ovoj test grupi bi trebalo da pokaže uspešnost detekcije samo naglih promena kadrova, osetljivost detekcije na korelisanost sadržaja između dva susedna kadra i osetljivost detekcije na pojavu ugrađenih grafičkih sadržaja na frejmovima. Sekvence su preuzete u *MPEG'4* video formatu. U Tabeli 6.6 dat je detaljan opis video sekvenci, a na Slici 6.2(a)-(b) prikazani su primeri frejmova sekvenci iz ove test grupe.

Tabela 6.6. Grupa televizijskih video test sekvenci

Oznaka	TV program	Godina	Produkcija	Vremenski okvir [h:m:s]		Dimenzija frejma [px]	Broj frejm.
				Početak	Kraj		
SEK41	Sportski program	2013	Arena, kablovska televizija	00:00:00	02:33:55.200	320x240	200000
SEK42	Jutarnji program	2013	RTS, javni servis	00:00:00	01:02:01.440	320x240	93037



a)



b)

Slika 6.2. Primeri frejmova iz treće grupe video test sekvenci, a) sportski program, b) jutarnji program.

6.2. Merenje performansi predloženog algoritma

Performanse predloženog algoritma za detekciju naglih promena kadrova merena su primenom dve metode. Prva metoda, *Metoda1*, predstavlja merenje performansi izračunavanjem tri parametra: tačnosti, preciznosti i F-mere. Druga metoda, *Metoda2*, predstavlja procenu koliko objektivno pogrešne detekcije predstavljaju naglu promenu sadržaja unutar video sekvence.

6.2.1. Metoda1

Primenom ove metode, performanse predloženog algoritma mere se na osnovu tri parametra: tačnosti (*recall*), preciznosti (*precision*) i F-mere (*F-measure*). Prvo se izvrši vizuelna inspekcija video sekvenci i utvrđuju se pozicije stvarnih promena kadrova. Nakon odrađenih simulacija porede se rezultati detekcije i podaci o stvarnim pozicijama naglih promena kadrova. Način izračunavanja pomenutih parametara tačnosti (T), preciznosti (P) i F-mere (F), definisan je izrazima (6.1)-(6.3) [PET10], respektivno. U ovim izrazima t_p označava broj ispravno detektovanih naglih promena kadrova, f_n predstavlja broj nedetektovanih stvarnih naglih promena kadrova, a f_p predstavlja broj pogrešnih detekcija naglih promena kadrova. Definicija ukupnog broja detektovanih naglih promena kadrova i ukupnog broja stvarnih naglih promena kadrova data je izrazima (6.4) i (6.5), respektivno.

$$T = \frac{t_p}{t_p + f_n} \quad (6.1)$$

$$P = \frac{t_p}{t_p + f_p} \quad (6.2)$$

$$F = \frac{2TP}{T + P} \quad (6.3)$$

$$t_p + f_p = \text{ukupan broj detektovanih naglih promena kadrova} \quad (6.4)$$

$$t_p + t_n = \text{ukupan broj stvarnih naglih promena kadrova} \quad (6.5)$$

Parametar T pokazuje procentualnu uspešnost algoritma u detekciji stvarnih naglih promena kadrova. Parametar P pokazuje procentualnu uspešnost algoritma u smislu broja ispravno detektovanih naglih promena kadrova među svim detektovanim promenama. Parametar F pokazuje uravnoteženost parametara tačnosti i preciznosti. Visoke vrednosti ovog parametra pokazuju visoku uspešnost algoritma za detekciju po pitanju oba parametra. Ako postoji velika razlika između vrednosti parametara tačnosti i preciznosti vrednost parametra F značajno opada.

6.2.2. Metoda 2

Druga metoda merenja performansi pokazuje u kojoj meri pogrešno detektovane nagle promene kadrova predstavljaju naglu promenu sadržaja u odnosu na ispravno detektovane. U ovoj metodi koristi se korelacija kao objektivna mera sličnosti između frejmova. Na osnovu rezultata detekcije izračunavaju se korelacije frejmova na detektovanim pozicijama. Dve vrednosti korelacije se izračunavaju: korelacija između frejma na poziciji detektovane promene i prethodnog frejma, i korelacija između frejma na poziciji detektovane promene i narednog frejma. Uzima se minimalna vrednost od ove dve vrednosti i dodeljuje detektovanoj poziciji. Utvrđuje se minimalna vrednost korelacije za ispravno detektovane nagle promene kadrova, $KORMIN$. Porede se sve izračunate vrednosti korelacija frejmova koje su dodeljene detektovanim lokacijama naglih promena kadrova. Sve lokacije čije vrednosti korelacija frejmova prevazilaze vrednost $KORMIN$, proglašavaju se pogrešnom detekcijom. Ostale lokacije proglašavaju se ispravnim detekcijama. Na osnovu novih informacija o pogrešnim i ispravnim detekcijama izračunavaju se tačnost (D) i preciznost (P). Vrednost $KORMIN$ se izračunava za svaku video sekvencu posebno.

6.3. Rezultati simulacija

Predloženi algoritam za detekciju naglih promena kadrova testiran je na ukupno 1.619.604 frejma video sekvenci različitih klasa i dužina trajanja. Pre testiranja vizuelnom

inspekcijom izvršena je provera da li su sve sekvence ispravno detektovane. Integritet filmskih sekvenci nije narušavan, tako da su kompletni nizovi frejmova preuzeti iz originalnih sekvenci. Kod televizijskih programa, ignorisane su detekcije gradijentnih tranzicija između kadrova koje su kraće od 5 frejmova [PET10], neravnomerne gradijentne tranzicije i gradijentne tranzicije praćene ugrađenom grafikom koja se trenutno menja između frejmova tokom tranzicije. Rezultati merenja performansi biće prikazani tabelarno u nastavku teksta po test grupama video sekvenci.

6.3.1. Merenje performansi na prvoj test grupi

U Tabeli 6.7 predstavljeni su rezultati testiranja predloženog algoritma za četiri dugometražna igrana filma. Ukupno je testirano 615.172 frejmova.

Tabela 6.7. Rezultati merenja performansi predloženog algoritma na prvoj test grupi video sekvenci.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
FILM1	0,99347	0,980662	0,987024	0,993595	1
FILM2	0,998411	0,990544	0,994462	0,998426	1
FILM3	0,999323	0,972973	0,985972	0,999341	1
FILM4	0,997314	0,966168	0,981494	0,997404	1

Rezultati merenja performansi primenom *Metode1*, za prvu test grupu, pokazuju visoku tačnost (*T*) predloženog algoritma ($>0,99$). Preciznost (*P*) je varijabilna i zavisi od klase video sekvence. Kombinovane video sekvence i animirane video sekvence imaju niže vrednosti preciznosti zbog većeg broja pogrešno detektovanih naglih promena kadrova. Na ovu činjenicu ukazuje i vrednost parametra F-mere (*F*), iako su vrednosti ovog parametra vrlo visoke ($>0,98$). Rezultati merenja performansi primenom *Metode2* pokazuju da sve pogrešno detektovane lokacije naglih promena kadrova predstavljaju značajnu promenu sadržaja između frejmova čija je korelacija niža od korelacija frejmova na ispravno detektovanim naglim promenama kadrova.

U ovoj test grupi izmerene performanse su usrednjene zbog velikog broja promena kadrova unutar svake video sekvence. Pored navedenog, svaka od sekvenci ima promenljivu dinamiku kadrova po pitanju dužine kadra i kretanja unutar kadra.

6.3.2. Merenje performansi na drugoj test grupi

U Tabelama 6.8-6.12 predstavljeni su rezultati testiranja predloženog algoritma na četrdeset različitih video test sekvenci približno istih dužina. Rezultati su raspoređeni u četiri različite klase na osnovu klasne pripadnosti test sekvenci. Ukupno je testirano 711.395 frejmova.

Tabela 6.8. Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa akcione sekvence.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
SEK1	0,976945	0,965812	0,971347	0,977716	1
SEK2	0,994885	0,953431	0,973717	1	0,995122
SEK3	0,954819	0,838624	0,892958	1	0,961832
SEK4	0,992701	0,819277	0,89769	0,891566	0,993289
SEK5	0,991561	0,951417	0,971074	1	0,991935
SEK6	0,988889	0,929504	0,958277	1	0,989637
SEK7	0,984127	0,873239	0,925373	1	0,986014
SEK8	0,984733	0,962687	0,973585	1	0,985294
SEK9	0,97861	0,994565	0,986523	1	0,97861
SEK10	0,977876	0,938429	0,957746	1	0,97921

Tabela 6.9. Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa dramske sekvence.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
SEK11	0,997416	0,987212	0,992288	0,997449	1
SEK12	1	0,978261	0,989011	1	1
SEK13	1	0,988506	0,99422	1	1
SEK14	1	0,983607	0,991736	1	1
SEK15	1	0,993197	0,996587	1	1
SEK16	1	0,994819	0,997403	1	1
SEK17	1	0,990741	0,995349	1	1
SEK18	1	0,987654	0,993789	1	1
SEK19	1	0,966292	0,982857	1	1
SEK20	1	1	1	1	1

Tabela 6.10. Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa kombinovane sekvence.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
SEK21	0,994475	0,967742	0,980926	0,994652	1
SEK22	0,989247	0,913907	0,950086	1	0,990164
SEK23	1	0,939759	0,968944	1	1
SEK24	1	0,953488	0,97619	1	1
SEK25	0,995434	0,995434	0,995434	1	0,995455
SEK26	0,989744	0,960199	0,974747	1	0,990148
SEK27	0,979079	0,936	0,957055	1	0,980392
SEK28	0,996255	0,988848	0,992537	1	0,996296
SEK29	1	0,991111	0,995536	1	1
SEK30	0,984314	0,825658	0,898032	0,996711	0,986971

Tabela 6.11. Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, klasa animirane sekvence.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
SEK31	0,992481	0,97417	0,98324	0,992674	1
SEK32	1	0,842857	0,914729	1	1
SEK33	1	0,680412	0,809816	0,989691	1
SEK34	0,98008	0,904412	0,940727	1	0,981949
SEK35	0,982143	0,948276	0,964912	1	0,983051
SEK36	1	0,943925	0,971154	1	1
SEK37	0,996078	0,944238	0,969466	1	0,996296
SEK38	1	0,990566	0,995261	1	1
SEK39	1	0,858268	0,923729	0,984252	1
SEK40	1	0,985366	0,992629	1	1

Tabela 6.12. Rezultati merenja performansi predloženog algoritma na drugoj test grupi video sekvenci, usrednjeni po klasama.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
Akcionne	0,981865	0,926084	0,953159	0,983066	0,994494
Dramske	0,999383	0,987203	0,993256	0,999391	1
Kombinovane	0,992195	0,942023	0,966458	0,992641	0,999564
Animirane	0,994492	0,924346	0,958137	0,994898	0,998294

Iz predstavljenih rezultata merenja performansi predloženog algoritma po klasama video sekvenci, primenom *Metoda1* pristupa, može se zaključiti da je generalno visoka tačnost algoritma (>0.98) bez obzira na klasu, dok preciznost detekcije značajno zavisi od klase. Ovu činjenicu potvrđuju i usrednjeni rezultati merenja po klasama u Tabeli 6.12. Vrednosti F parametra ukazuju na neuravnoteženost tačnosti i preciznosti za klase akcionih

i animiranih sekvenci. Ako se pogledaju pojedinačni rezultati merenja po klasama može se uočiti da preciznost značajno varira od sekvence do sekvence.

Merenje performansi primenom *Metoda2* pristupa ukazuje da su pogrešno detektovane lokacije naglih promena kadrova po korelaciji frejmova i dalje u visokom procentu ispod nivoa korelacije frejmova ispravno detektovanih lokacija.

6.3.3. Merenje performansi na trećoj test grupi

U Tabeli 6.13 predstavljeni su rezultati testiranja predloženog algoritma na video sekvencama televizijskih programa različitih sadržaja i dužina trajanja. Ukupno je testirano 293.037 frejmova.

Tabela 6.13. Rezultati merenja performansi predloženog algoritma na trećoj test grupi video sekvenci.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
Sportski program	0,99894	0,983299	0,991057	0,998957	1
Jutarnji program	0,995434	0,98941	0,992413	0,995482	1

Prikazani rezultati merenja performansi predloženog algoritma, primenom pristupa *Metoda1*, pokazuju visoku efikasnost detekcije za ovu test grupu (>0.99). Preciznost detekcije je na visokom nivou (>0.83), a vrednosti *F*-mere ukazuju na uravnoteženost parametara tačnosti i preciznosti (>0.99).

Merenja performansi predloženog algoritma, primenom pristupa *Metoda2*, pokazuju da pogrešno detektovane lokacije naglih promena kadrova predstavljaju lokalno značajnu promenu sadržaja.

6.3.3 Usrednjeni rezultati merenja performansi po test grupama

U Tabeli 6.14 predstavljeni su rezultati merenja performansi predloženog algoritma, usrednjeni po test grupama.

Tabela 6.14. Rezultati merenja performansi predloženog algoritma usrednjeni po test grupama.

Oznaka	Metoda1			Metoda2	
	T	P	F	T	P
Prva test grupa	0,996916	0,979543	0,988153	0,996978	1
Druga test grupa	0,990259	0,941006	0,965005	0,990801	0,997545
Treća test grupa	0,9975	0,985794	0,991612	0,997535	1

Prikazani rezultati merenja performansi, primenom pristupa *Metoda1*, pokazuju da predloženi algoritam ima visoku tačnost (>0.99) u sve tri test grupe. Preciznost detekcije ima najveću vrednost za televizijske programe. Ovo je razumljivo s obzirom na tehnologiju proizvodnje video sekvenci. Najniži nivo preciznosti izmeren je za drugu test grupu (<0.95), gde su sekvence strogo profilisane po klasama, pa generalno smanjena preciznost kod akcionih i animiranih sekvenci utiče na opštu preciznost unutar test grupe. Ova činjenica je praćena padom vrednosti *F-mere*, koja ukazuje na neuravnoteženost vrednosti detekcije i preciznosti za drugu test grupu. Može se primetiti da je vrednost *F-mere* najviša za televizijske programe, čemu su doprinele visoke vrednosti tačnosti i preciznosti, odnosno uravnoteženost njihovih vrednosti.

Merenje performansi predloženog algoritma po test grupama, primenom pristupa *Metoda2*, ukazuje na generalni trend da pogrešno detektovane lokacije naglih promena kadrova zaista jesu lokacije značajnih promena sadržaja između susednih frejmova.

6.4. Poređenje sa postojećim algoritmima u literaturi

Rezultati testiranja predloženog algoritma za detekciju naglih promena kadrova poređeni su sa postojećim algoritmima koji su predstavljeni u okviru projekta *TRECVID*. Algoritmi razvijeni u okviru ovog projekta poseduju najbolje aktuelne performanse u mnogim oblastima multimedije. Kada je u pitanju detekcija kadrova, istraživači su se intenzivno bavili ovom oblašću u okviru *TRECVID* projekta u periodu od 2001. do 2007. godine. Svake godine je vršeno testiranje realizovanih algoritama. Video sekvence na kojima su testirani razvijeni algoritmi, sačinjene su od sekvenci vesti, komercijalnog

programa, informativnih priloga, snimaka sportskih događaja i reklama. Po sadržaju video sekvence na kojima su testirani pomenuti algoritmi odgovaraju trećoj test grupi video sekvenci na kojoj je testiran predloženi algoritam. Ukupna količina frejmova nad kojima je testiranje izvršeno kreće se reda nekoliko stotina hiljada frejmova. Na osnovu prikazanih performansi odabrane su tri godine *TRECVID* projekta (*TRECVID2004*, *TRECVID2005* i *TRECVID2007*). Iz svakog perioda odabrano je nekoliko najuspešnijih algoritama na osnovu performansi. U Tabeli 6.15 prikazani su odabrani algoritmi za poređenje raspoređeni po godinama učešća. Nazivi algoritama su preuzeti po imenu univerzitetima ili instituta na kojima su razvijeni.

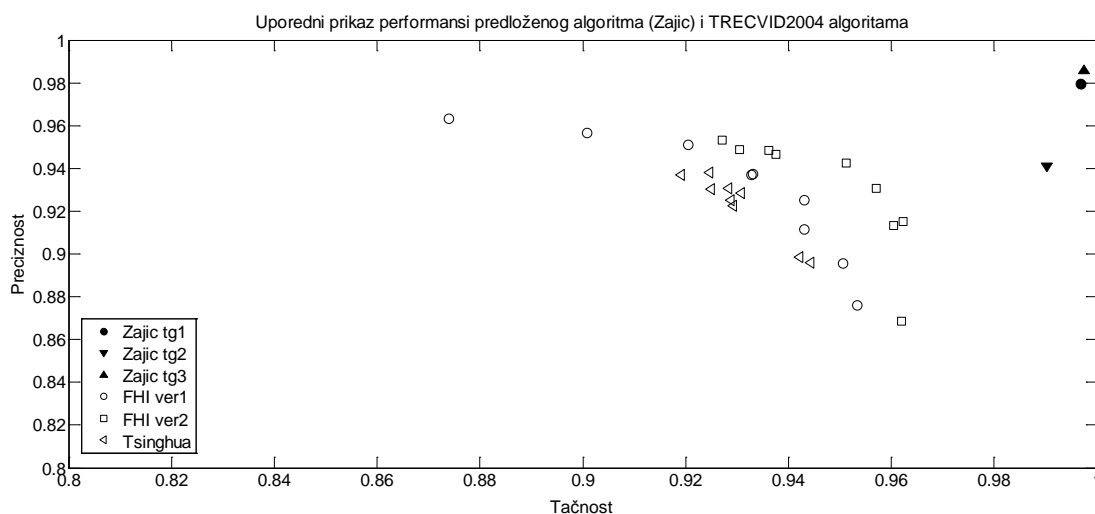
Tabela 6.15. Pregled algoritama za detekciju naglih promena kadrova za poređenje performansi sa predloženim algoritmom.

TRECVID period	TRECVID Algoritmi
TRECVID2004	Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut (ver1 i ver2) Tsinghua University, Beijing
TRECVID2005	Hong Kong Polytechnic University KDDI R&D Laboratories IBM Research National ICT Australia Tsinghua University, Beijing TU DELFT
TRECVID2007	Beijing University of Posts and Telecommunications Science and Technical Research Laboratories, Tokio Tsinghua University, Beijing University of Karlsruhe University of Marburg

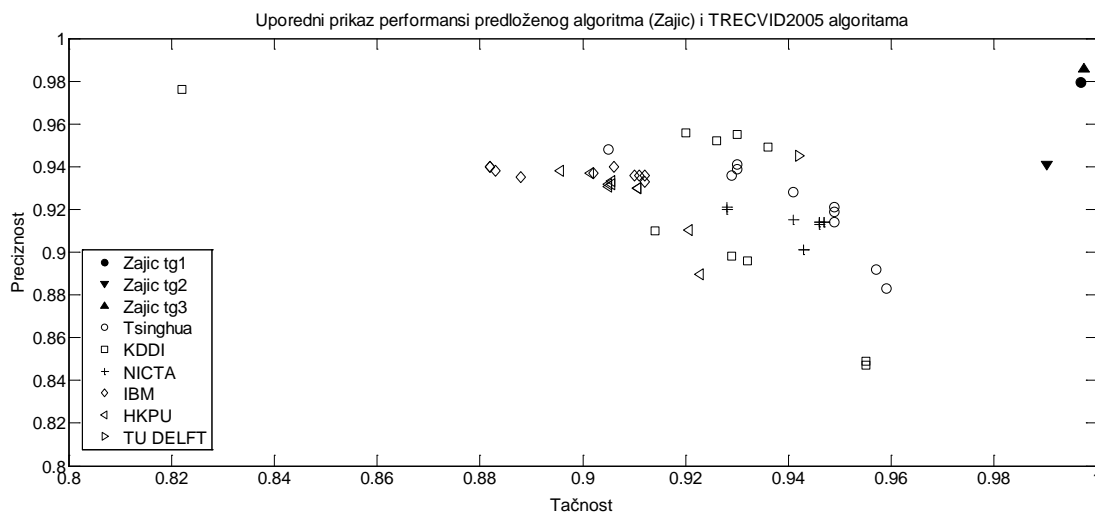
Poređenje algoritama je izvršeno na osnovu parametara tačnosti, preciznosti i *F-mere*. Poređenje algoritama na osnovu detektivnosti i preciznosti prikazano je na Slikama 6.3-6.5. Na prikazanim slikama predstavljeni su rezultati *TRECVID* algoritama dostupni iz literature. Svaki od učesnika predstavljen je sa više rezultata u skladu sa različitim segmentima video baze na kojima su algoritmi testirani. Predloženi algoritam je predstavljen sa tri merenja koja predstavljaju usrednjene rezultate za tri test grupe. Korišćene su vrednosti parametara tačnosti, preciznosti i *F-mere* iz Tabele 6.14, izmerene primenom pristupa *Metoda1*. S obzirom da do sada u literaturi za testiranje algoritama za detekciju naglih promena kadrova nisu korišćene video sekvence visoke postprodukcije, iskorišćena je prilika da se i ti rezultati uporede sa rezultatima poznatih algoritama iz

literature. Rezultati testiranja predloženog algoritma prikazani su za sve tri test grupe (*Zajic tg1*, *Zajic tg2*, *Zajic tg3*).

Poređenje performansi algoritama za detekciju naglih promena kadrova na Slici 6.3 pokazuje da predloženi algoritam ima znatno veću tačnost od *TRECVID2004* algoritama za sve tri test grupe. Kada je u pitanju preciznost, rezultati testiranja na prvoj i trećoj test grupi (*Zajic tg1* i *Zajic tg3*) bolji su od izmerene preciznosti *TRECVID2004* algoritama.



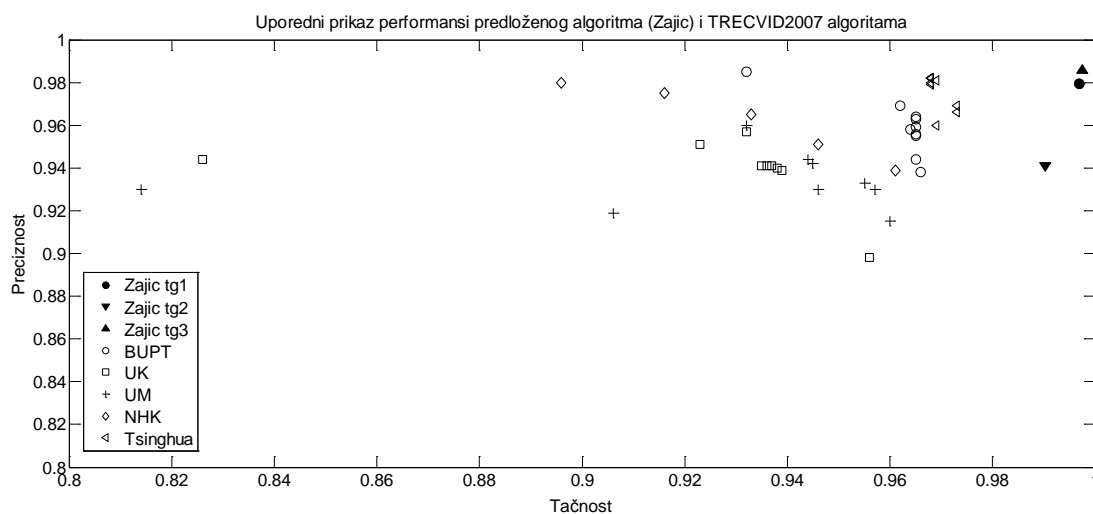
Slika 6.3. Uporedni prikaz performansi predloženog algoritma (Zajic) i *TRECVID2004* algoritama.



Slika 6.4. Uporedni prikaz performansi predloženog algoritma (Zajic) i *TRECVID2005* algoritama.

Preciznost izmerena pri testiranju predloženog algoritma na drugoj test grupi (*Zajic tg2*) nalazi se iznad srednjeg nivoa preciznosti prikazanih merenja *TRECVID2004* algoritama. S obzirom da vrednosti parametara za drugu test grupu predstavljaju usrednjene vrednosti preciznosti na nivou cele test grupe, preciznost pokazana pri testiranju na ovoj test grupi se može porediti sa *TRECVID* algoritmima.

Rezultati poređenja performansi *TRECVID2005* algoritama i predloženog algoritma prikazani su na Slici 6.4. Na slici se može videti dominantnost predloženog algoritma po pitanju tačnosti. Preciznost postignuta pri testiranju na prvoj i trećoj test grupi je veća od preciznosti svih merenja *TRECVID2005* algoritama. Srednja vrednost preciznosti svih merenja *TU DELFT* algoritma prevazišla je preciznost predloženog algoritma pri testiranju na drugoj test grupi. Performanse predloženog algoritma za ovu test grupu su uporedljive sa performansama *TRECVID2005* algoritama. Na Slici 6.5 prikazani su rezultati poređenja performansi predloženog algoritma i *TRECVID2007* algoritama.



Slika 6.5. Uporedni prikaz performansi predloženog algoritma (*Zajic*) i *TRECVID2007* algoritama.

Na gornjoj slici može se videti da predloženi algoritam ima visok nivo tačnosti u odnosu na *TRECVID* algoritme. Preciznost *TRECVID* algoritama je visoka, a pojedinačna merenja nekih algoritama dostižu vrlo visoke vrednosti (>0.98). Generalno, vrednosti preciznosti predloženog algoritma za prvu i treću test grupu više su od nivoa preciznosti

TRECVID algoritama. Preciznost prikazana na drugoj test grupi, nije poredljiva sa rezultatima *TRECVID* algoritama.

Poređenje na osnovu parametra F -mere prikazano je u Tabeli 6.16. U tabeli su prikazane tri najviše vrednosti F -mere svih *TRECVID* algoritama, bez obzira na *TRECVID* godinu. Vrednosti F -mere za tri test grupe na kojima je testiran predloženi algoritam prikazani su u drugom delu tabele. Vrednosti u tabeli su sortirane po opadajućem redosledu.

Tabela 6.16. Poređenje F -mere (F) izmerene za predloženi algoritam i *TRECVID* algoritme.

Algoritam	F (TRECVID)	Test grupa	F (ZAJIC)
Tsinghua(TV2007)	0,975	Treća	0,991612
NHK(TV2007)	0,968	Prva	0,988153
BUPT(TV2007)	0,965	Druga	0,965005

Prikazane vrednosti F -mere u Tabeli 6.16 pokazuju da je uravnoteženost performansi predloženog algoritma na visokom nivou, čak i iznad nivoa *TRECVID* algoritama za neke test grupe. Visoka tačnost predloženog algoritma najviše je uticala na postizanje visokih vrednosti F -mere bez obzira na test grupu, a zavisnost preciznosti od test grupa najviše je doprinela rasponu vrednosti F -mere po grupama.

6.5. Diskusija rezultata testiranja i poređenja

Prikazani rezultati testiranja predloženog algoritma na različitim video sekvencama pokazali su visok nivo tačnosti i relativno visok nivo preciznosti. Tačnost algoritma ne zavisi od klase video sekvence na kojoj se testira, što je direktna posledica zasnivanja algoritma na identifikaciji artefakata u korelacionom nizu. Artefakti su jasno uočljivi i lako se detektuju. Preciznost algoritma je generalno visoka, ali varira u zavisnosti od klase video sekvence na kojoj se vrši testiranje. Klasifikacija artefakata predstavlja složen proces u kojem nije jednostavno razdvojiti brzo pomeranje kamere ili objekata od stvarnih naglih promena kadrova. Poređenje algoritma sa najuspešnijim algoritmima za detekciju naglih promena kadrova, koji su prezentovani u literaturi (*TRECVID*), pokazalo je da su

performanse predloženog algoritma ne samo poredljive, već se mogu smatrati i konkurentnim. Na video sekvencama koje po sadržaju odgovaraju test sekvencama na kojima su testirani *TRECVID* algoritmi (informativni, komercijalni i sportski program), prezentovani algoritam je pokazao veću tačnost i preciznost od istih. Kada su u pitanju dugometražne video sekvence sa visokom postprodukcijom, preciznost algoritma je poredljiva sa najuspešnijim *TRECVID* algoritmima. Iako se radi o različitim tipovima video sekvenci, poredljivost je moguća jer u dugometražnim video sekvencama dominiraju dramske scene kod kojih je preciznost detekcije generalno veća. Manji udeo dramskih scena u test sekvencama, strogo profilisanim po klasama video sadržaja, utiče na smanjenje nivoa preciznosti.

Uspešnost detekcije naglih promena kadrova predloženog algoritma je posledica primene nove metodologije u procesu detekcije. Na performanse algoritma za detekciju utiču primene specijalnih efekata, brzih pomeranja kamere i objekata unutar kadrova. Uzastopne brze promene sadržaja između frejmova otežavaju detekciju. Sa druge strane, performanse sistema za detekciju umanjuju se i pogrešnom klasifikacijom detektovanih naglih promena sadržaja.



Slika 6.6. Ispušteni kadrovi u procesu detekcije: a) usled promene osvetljenja celog frejma i ubacivanja veštačkih frejmova, b) usled specijalnog efekta unutar kadra.

Primena različitih specijalnih efekata može izazvati maskiranje stvarnih promena kadrova, koje rezultira ispuštanjem stvarne promene. Neki od primera specijalnih efekata koji utiču na tačnost su: nagla promena osvetljenja u okolini stvarne promene kadra, zatim ubačeni frejm drugačijeg sadržaja u okolini stvarne promene kadra ili specijalan efekat koji traje niz frejmova i tokom trajanja značajno menja sadržaj frejmova. Kontinuirana značajna

promena sadržaja, koja traje više frejmova onemogućava formiranje artefakata u korelacionom nizu histograma frejmova, i samim tim izaziva ispuštanje stvarnih promena kadrova u procesu detekcije. Na Slici 6.6 prikazani su primeri specijalnih efekata koji otežavaju detekciju stvarne promene kadrova.

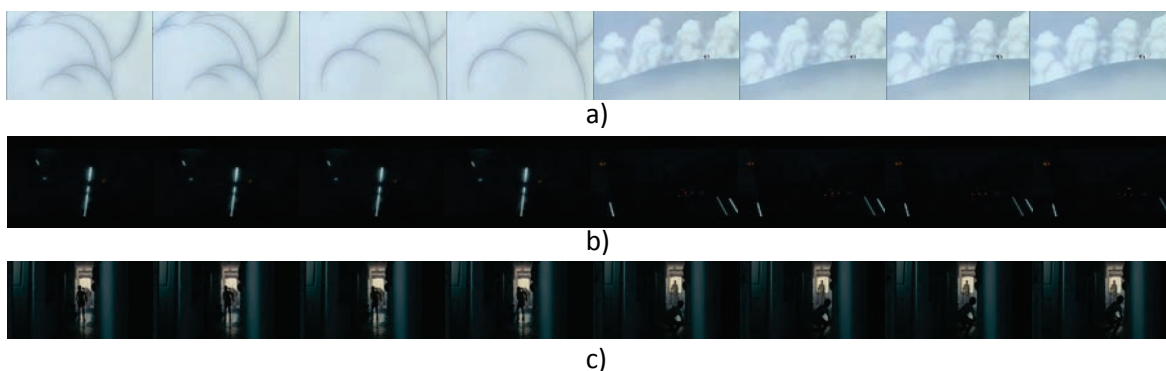
Brzo pomeranje kamere i objekata u blizini granica kadrova može uticati na tačnost algoritma. Pomeranje se može pojaviti sa obe strane granice između dva kadra. Ovakve kontinuirane promene sadržaja između frejmova degradiraju kontinuitet promene niskog intenziteta i onemogućavaju formiranje artefakata u korelacionom nizu histograma. Pored problema sa identifikacijom artefakta, kao dodatan problem pojavljuje se izostajanje izrazitog lokalnog maksimuma u alfa domenu, koji se koristi za pozicioniranje nagle promene sadržaja. U alfa domenu se pojavljuje veći broj lokalnih maksimuma koji odgovaraju značajnim promenama sadržaja između frejmova. Primer naglog pomeranja kamere i objekata koji maskiraju stvarnu promenu kadra prikazan je na Slici 6.7.



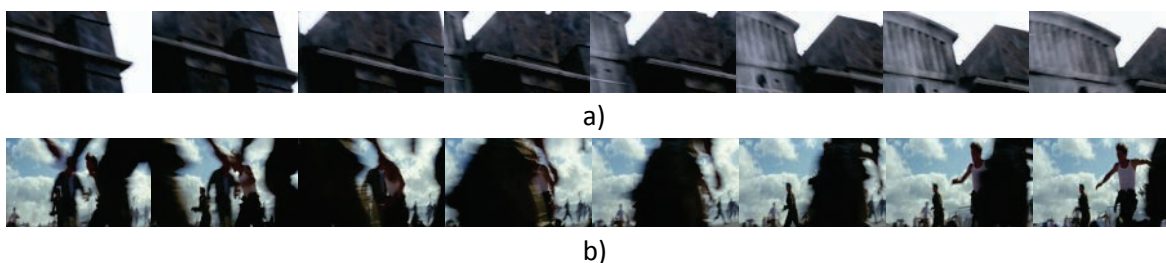
Slika 6.7. Ispušteni kadrovi u procesu detekcije: a) usled brzog kretanja kamere i objekata, b) i c) usled brzog kretanja kamere.

Detekcija naglih promena kadrova može biti otežana ako su grupe graničnih frejmova visoko korelisane. Ovakav slučaj je moguć kada se kamera naglo pomeri za mali ugao u istoj sceni. Drugi primer je visoka korelisanost sadržaja dva kadra. Visoka korelisanost sadržaja frejmova sa obe strane granice između kadrova, veličinu formiranog artefakta smanjuje na vrednost ispod praga detekcije. Na Slici 6.8 prikazani su primeri nagle promene kadrova koji nisu detektovani zbog visoke korelisanosti graničnih frejmova, odnosno male promene ugla kamere.

Preciznost detekcije naglih promena kadrova može biti degradirana iz nekoliko razloga. Prvi primer je brzo pomeranje kamere, koje ne predstavlja naglu promenu kadra. Ako se kamera dovoljno brzo pomera, razlika sadržaja između dva frejma postaje dovoljno značajna i algoritam za detekciju takvu promenu klasifikuje kao naglu promenu kadra. Detaljnija kvantifikacija ovakvih promena nije bila predmet analize u disertaciji. Primeri brzog pomeranja kamera prikazani su na Slici 6.9(a)-(b). Na slikama su uočljive promene sadržaja između frejmova. Na prikazanim slikama, između četvrte i pete pozicije u nizu frejmova, algoritam je pogrešno detektovao naglu promenu kadra.



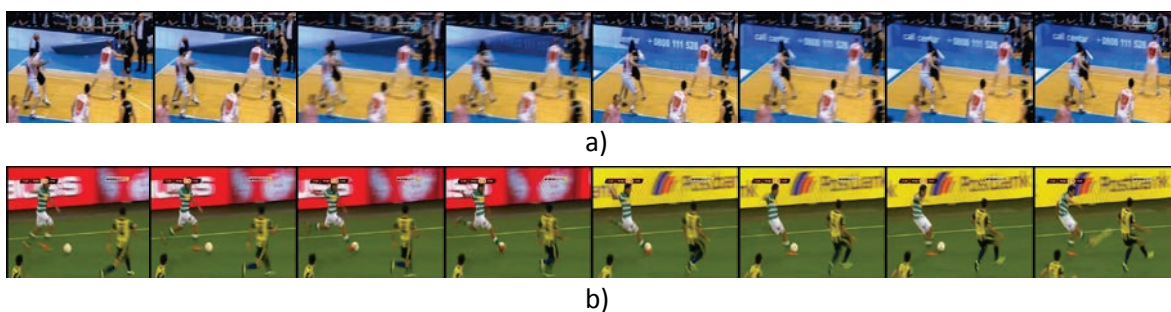
Slika 6.8. Ispušteni kadrovi u procesu detekcije: a) i b) visoka korelisanost sadržaja oba kadra, c) mala promena ugla kamere između dva kadra.



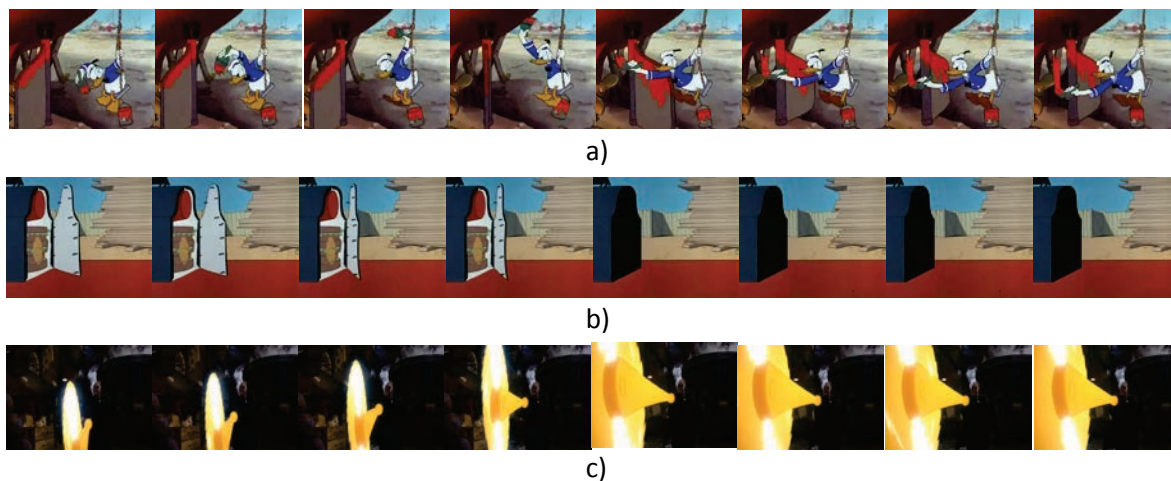
Slika 6.9. Prikaz pogrešne detekcije nagle promene kadra usled brzog pomeranja kamere: a) statična scena, b) dinamična scena.

Drugi primer degradacije preciznosti detekcije predstavlja naglo pojavljivanje ili naglo nestajanje objekata između dva susedna frejma. Ovakvi događaji najčešće prouzrokuju greške u procesu prekodovanja video sekvence kada koderi nisu u stanju da tačno odrede vektor pokreta zbog trenutnog pomeraja. Ove greške izazivaju na celoj površini frejma *blur* efekat i tako prouzrokuju značajnu promenu sadržaja koju algoritam

detektuje kao naglu promena kadra. Grafički elementi koji se naglo pojavljuju i nestaju u televizijskim programima takođe mogu izazvati lažnu detekciju nagle promene kadrova. Na Slici 6.10 prikazani su primeri nagle promene sadržaja dela frejma koja je izazvala *blur* efekat, zbog kojeg su u oba slučaja na pozicijama između četvrtog i petog frejma pogrešno detektovane promene kadrova.



Slika 6.10. Prikaz pogrešne detekcije nagle promene kadra usled nagle promene sadržaja dela frejma: a) reklamni pano na košarkaškoj utakmici, b) reklamni pano na fudbalskoj utakmici.

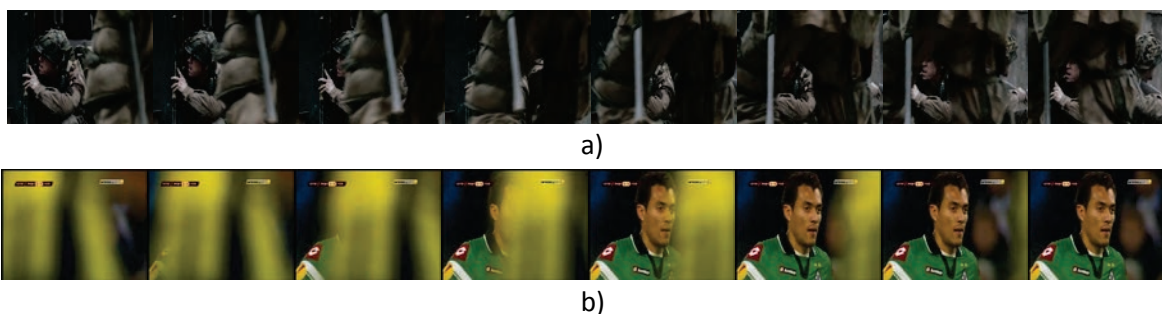


Slika 6.11. Prikaz pogrešne detekcije nagle promene kadra: a) i b) nagli pomeraj objekta, c) nagla promena veličine objekta.

Nagli pomeraj objekta koji izaziva aproksimativno dvostruku promenu sadržaja (na staroj i novoj poziciji), može izazvati pogrešnu detekciju nagle promene kadra. Ovo je čest slučaj kod animiranih video sekvenci u kojima autori koriste slobodu da na ovaj način dodatno ubrzavaju tok radnje u sceni. Pogrešna detekcija nagle promene sadržaja može se

dogoditi i u slučaju nagle promene veličine objekta. I ovo je čest slučaj u animiranim video sekvencama. Slika 6.11 prikazuje primere kadrova koji su pogrešno detektovani kao nagle promene kadrova usled naglog pomeranja objekata, odnosno nagle promene veličine objekta.

Brzo pomeranje objekata u prvom planu može dovesti do pogrešne detekcije nagle promene kadra. Objekat u prvom planu zauzima veliku površinu, te ulaskom ovakvog objekta u zahvat kamere dolazi do zaklanjanja velike površine frejma. Ako se objekat dovoljno brzo pomeri izaziva se značajna promena sadržaja između dva frejma koju algoritam za detekciju može pogrešno klasifikovati kao naglu promenu kadra. Primeri kadrova sa ulaskom objekta u prvi plan prikazani su na Slici 6.12. Oba prikazana prolaska objekta u prvom planu pogrešno su detektovana kao nagle promene kadrova na poziciji između četvrtog i petog frejma.



Slika 6.12. Prikaz pogrešne detekcije nagle promene kadra prilikom prolaska objekta u prvom planu: a) filmski program, b) sportski program.

Prezentovane probleme u detekciji naglih promena kadrova moguće je rešiti primenom dodatnih tehnika obrade signala. Brzo kretanje kamere i objekata unutar kadrova degradira stranice artefakata na pozicijama naglih promena kadrova. Oblik artefakta u tom slučaju nalikuje gradijentnim promenama kadrova, gde takođe postoji kontinualna značajna promena sadržaja. Detekcijom gradijentnih promena moguće je detektovati i degradirane nagle promene kadrova, koje bi se primenom odgovarajućeg algoritma klasifikacije izdvojile od gradijentnih promena. Visoka korelisanost frejmova između kadrova takođe je jedan od razloga neuspešne detekcije naglih promena kadrova. Ovaj problem mogao bi se rešiti primenom odgovarajuće postobrade niza frejmova da se ne bi ugrozile ukupne

performanse sistema. Postobradom bi se u ograničenom okruženju, na osnovu već detektovanih pozicija naglih promena kadrova, daleko uspešnije detektovale visoko korelisane nagle promene kadrova.

Klasifikacija kandidata za nagle promene kadrova može se unaprediti primenom odgovarajuće detekcije pomeraja unutar kadra, koja bi mogla da ukaže na pomeraj sadržaja u određenom pravcu. Na osnovu ove informacije moglo bi se odlučiti da li je u pitanju stvarna promena kadra ili pomeranje kamere, odnosno objekta. Bolja klasifikacija naglih promena sadržaja, koja odgovaraju naglim pomerajima objekata, mogla bi se postići postobradom grupe frejmova sa obe strane detektovane lokacije nagle promene.

Sva pomenuta moguća rešenja zahtevaju dodatnu obradu frejmova koja nije predmet analize u ovoj disertaciji, već su deo budućih unapređenja algoritma.

7. Zaključak

Predstavljeni algoritam za detekciju naglih promena kadrova u video sekvencama pokazao je određene napredne performanse pri intenzivnom eksperimentalnom testiranju, koje je detaljno opisano u poglavlju 6. U ovom poglavlju prikazan je skraćeni pregled izmerenih performansi algoritma i rezultati poređenja sa poznatim postojećim algoritmima u ovoj oblasti. Pored pregleda izmerenih performansi, dati su i izvedeni zaključci na osnovu rezultata testiranja. U nastavku teksta prezentovani su osnovni doprinosi ove disertacije sa kratkim analizama i buduće smernice u cilju daljeg razvoja i unapređenja performansi algoritma.

Rezultati intenzivnog testiranja algoritma za detekciju naglih promena kadrova pokazali su visoku tačnost algoritma bez obzira na klasu (akcione, dramske, animirane i kombinovane) video sekvenci nad kojima je vršeno testiranje. U akcionim video sekvencama prisutna je izražena dinamika pokreta kamere i objekata, uz prisustvo specijalnih efekata. Dramske sekvence se odlikuju statičnim položajem kamere i sporim kretanjem objekata unutar kadrova. Kombinovane sekvence, kao kombinacija realnog videa i ugrađene animacije (crtane ili računarske), sadrže tranzicije animiranih objekata u

realnom okruženju što posebno može uticati na detekciju kadrova. Animirane video sekvence predstavljaju izraz umetničke slobode autora u kreiranju brzih tranzicija scene i objekata. Visok nivo tačnosti je ostvaren pogodnim odabirom obeležja frejmova, diferencijalne metrike i primenom multifraktalne analize. Primenom histograma boje i *wavelet* obeležja, kao obeležja frejmova, obezbeđen je kvalitetan opis sadržaja frejmova i velika osetljivost na nagle promene u susednim frejmovima video sekvence. Primena korelacije kao diferencijalne metrike, implementirana na odgovarajući način, omogućila je kreiranje vrlo uočljivih artefakata na pozicijama naglih promena sadržaja. Na ovaj način detekcija naglih promena kadrova je pojednostavljena. Ujedno, korelacija je omogućila univerzalnost primene algoritma na različite video sekvence, bez obzira na klasu i tehnologiju proizvodnje video materijala. Uvođenjem multifraktalne analize omogućena je efikasna detekcija nagle promene sadržaja susednih frejmova i lokalizacija konkretnog frejma gde je nastala nagla promena sadržaja.

Testiranjem algoritma na velikom broju video sekvenci pokazalo se da preciznost detekcije naglih promena kadrova zavisi od video sekvenci na kojima se vrši testiranje. Konkretno, na video sekvencama sa sporijim promenama sadržaja, bez obzira na prisustvo različitih gradijentnih promena kadrova, dobijen je visok nivo preciznosti detekcije naglih promena kadrova. Testiranje predloženog algoritma na video sekvencama sa bržom tranzicijom radnje, specijalnim efektima, brzim pomeranjem kamera i objekata unutar kadra, pokazano je da preciznost detekcije naglih promena kadrova varira od sekvence do sekvence. Broj kadrova sa brzim pomeranjem kadrova, brzim pomeranjem objekata i specijalnim efektima, direktno utiče na preciznost algoritma. U ovom slučaju, uzrok niskog nivoa preciznosti klasifikacije naglih promena sadržaja nalazi se u činjenici da brze tranzicije kamere i objekata realno odgovaraju nagloj promeni sadržaja kada su u pitanju niske učestanosti frejmova ($<100f/s$) video sekvenci [SAL11]. Korišćenjem sekvenci visokih učestanosti frejmova ($\geq 100f/s$), brze tranzicije kamere i objekata imaju ravnomerniju tranziciju promene sadržaja kroz frejmove i ne predstavljaju značajnu naglu promenu sadržaja. U eksperimentalnim testiranjima korišćene su video sekvence sa niskim učestanostima frejmova ($25f/s$). Specijalni efekti predstavljaju vrlo zahtevne segmente

video sekvenci, u kojima postoji drastična promena sadržaja frejmova i još uvek nije usvojena jasna klasifikacija po pitanju promene kadra uz prisustvo specijalnih efekata.

Rezultati poređenja predloženog algoritma sa poznatim algoritmima za detekciju naglih promena kadrova visokih performansi, pokazali su da su performanse predloženog algoritma poredljive sa performansama poznatih algoritama. U nekim aspektima performansi, kao što je tačnost detekcije, predloženi algoritam se pokazao i boljim od algoritama sa kojima je vršeno poređenje. Poređenja su pokazala da F-mera nivoa performansi predloženog algoritma (tačnosti i preciznosti) varira u zavisnosti od video sekvenci na kojima je vršeno testiranje. Varijacija F-mere performansi je posledica promenljive preciznosti koja je već analizirana u pregledu rezultata.

Važni doprinosi ove disertacije mogu se izdvojiti u niz konstatacija koje oslikavaju razvojni put predloženog algoritma. U nastavku teksta dati su glavni doprinosi ove disertacije:

Doprinos 1 - Pregled i analiza obeležja frejmova video sekvenci implementiranih u standardnim tehnikama za detekciju naglih promena kadrova.

Obeležja frejmova predstavljaju polaznu osnovu za bilo kakvu klasifikaciju ili detekciju na osnovu sadržaja frejmova. U ovoj disertaciji je dat pregled i opširna analiza obeležja frejmova koja su implementirana u standardnim tehnikama za detekciju naglih promena kadrova i standardnim sistemima za pretraživanje na osnovu sadržaja frejmova [ZAJ06, ZAJ07, ZAJ08, ZAJ12]. Širok spektar različitih obeležja slike je analiziran, a posebno je razmatrana njihova osetljivost na promene sadržaja u slici [ZAJ11]. Razmatrani su razni aspekti kompaktnijeg zapisa [ZAJ08] obeležja slika uravnoteženih dužina u cilju efikasnije pretrage i izbegavanja dominacije nekog od obeležja. Kompleksnost izračunavanja obeležja je, takođe, bio jedan od važnih aspekata analize.

Doprinos 2 - Pregled i analiza standardnih tehnika za detekciju naglih promena kadrova.

Veliki broj algoritama za detekciju naglih promena kadrova je publikovan do danas. Najviše algoritama je razvijeno u okviru *TRECVID* [TRE07] projekta. Pojedini od algoritama u okviru ovog projekta pokazali su napredne performanse pri testiranjima na *TRECVID* video bazama. U ovoj disertaciji je dat pregled i analiza najuspešnijih algoritama *TRECVID* projekta, sa posebnim osvrtom na korišćena obeležja frejmova u okviru algoritama i performanse sistema. Pored navedenih, analizirani su još neki algoritmi za detekciju izvan *TRECVID* projekta. Predstavljeni algoritmi predstavljaju proširenje skupa obeležja i diferencijalnih metrika koje se koriste u procesu detekcije naglih promena kadrova.

Doprinos 3 – Odabir obeležja frejmova za detekciju naglih promena kadrova.

Na osnovu intenzivnih istraživanja u oblasti indeksiranja slika na osnovu sadržaja [ZAJ06] i razvoju sistema za pretraživanje slika na osnovu sadržaja [ZAJ07], odabir obeležja za detekciju naglih promena kadrova sveden je na samo dva obeležja. Odabrani su histogram boje i vejevlet obeležje. Histogram boje predstavlja standardno [PET10] obeležje u algoritmima za detekciju naglih promena kadrova. Nema mnogo primera u literaturi u kojima se vejevlet obeležje koristilo za detekciju naglih promena kadrova. Nedostatak vejevlet obeležja je velika osetljivost na promenu sadržaja [PET10]. U ovoj disertaciji vejevlet obeležje je upotrebljeno po prvi put za lokalizaciju naglih promena sadržaja. Visoka osetljivost obeležja na promene sadržaja iskorišćena je upotrebom diferencijalne metrike i multifraktalne analize za određivanje pozicije nagle promene. Kombinacijom ova dva obeležja postignuta je određena tolerancija na promene sadržaja malog intenziteta (histogram boje), a sa druge strane zadržana je visoka osetljivost na značajne promene (vejevlet obeležje).

Doprinos 4 – Višestruka upotreba jednog obeležja u okviru algoritma za detekciju naglih promena kadrova.

Stvarna iskorišćenost informacije dobijene pri izdvajanju određenog obeležja iz frejma, vrlo je diskutabilna. Mogućnost da se jedno obeležje iskoristi na više načina otvara prostor za smanjenje broja izračunavanja i na generalno smanjenje kompleksnosti sistema.

U ovoj disertaciji histogram boje je upotrebljen na dva načina, na lokalnom nivou i na globalnom nivou. Na lokalnom nivou izračunat je na nivou regiona, a na globalnom nivou dobijen je sabiranjem lokalnih histograma. Globalni histogram boje iskorišćen je za detekciju nagle promene sadržaja, a histogram na lokalnom nivou koristi se za potrebe *flash* detektora. Pored navedenog na ovo obeležje je primenjena korelacija kao diferencijalna metrika u dva moda, analizom susednih frejmova i analizom frejmova sa pomakom.

Doprinos 5 – Kreiranje uočljivih artefakata koji se jednostavno detektuju.

Upotrebom diferencijalne metrike u algoritmima za detekciju naglih promena kadrova, omogućava se lakše uočavanje i jednostavnija detekcija naglih promena sadržaja. Nagle promene različitog intenziteta uzrokuju promenljive vrednosti diferencijalnih metrike na lokacijama promene, što otežava detekciju. U ovoj disertaciji je prikazana primena korelacije obeležja kao diferencijalne metrike. Primenom korelacije, sa pomakom od nekoliko frejmova i bez pomaka, obezbeđeno je kreiranje artefakata i lokalnih minimuma na pozicijama naglih promene sadržaja. Upoređivanjem pozicija artefakata i lokalnih minimuma vrši se verifikacija pozicije nagle promene. Kombinovanjem pristupa u primeni diferencijalne metrike obezbedila se visoka tačnost detekcije promene kadrova.

Doprinos 6 – Primenom korelacije kao diferencijalne metrike obezbeđena je univerzalnost detekcije naglih promena kadrova, bez obzira na klasu video sekvenci.

U ovoj disertaciji je prikazana primena korelacije kao diferencijalne metrike koja kao rezultat merenja sličnosti između obeležja dva frejma uvek daje apsolutnu vrednost između nule i jedinice bez obzira na srednji nivo signala. Promene sadržaja mogu biti većeg i manjeg intenziteta, ali rezultat diferencijalne metrike je uvek u istom intervalu.

Doprinos 7 – Primena multifraktalne analize za lokalizaciju naglih promena sadržaja.

Primena multifraktalne analize u ovoj oblasti predstavlja novinu. U ovoj disertaciji je prikazana upotreba multifraktalne analize za lokalizaciju nagle promene sadržaja, gde je iskorišćena njena sposobnost lokalnog opisivanja signala. Multifraktalna analiza je primenjena na korelacioni niz vejevlet obeležja koji je vrlo osetljiv na promenu sadržaja između frejmova. Ceo niz je transformisan u alfa domen multifraktalne analize, gde se nagle promene sadržaja pojavljuju kao lokalni maksimumi. Pomenuti lokalni maksimumi iskorišćeni su za identifikaciju kreiranih artefakata u korelacionom nizu histograma boje sa stvarnim pozicijama naglih promena sadržaja.

Doprinos 8 – Primena jednostavne i brze klasifikacije upotrebom klasifikatora koji rade na principu pragova.

Klasifikacija detektovanih pozicija naglih promena sadržaja predstavlja završni deo procesa detekcije naglih promena kadrova. U ovoj disertaciji je prikazana klasifikacija rezultata detekcije artefakata primenom klasifikatora koji rade na principu pragova. Pragovi su empirijski utvrđeni na velikom uzorku frejmova ($>1.500.000$ frejmova). Ovakav pristup obezbeđuje brzu i jednostavnu klasifikaciju.

Doprinos 9 – Izvršena je uporedna analiza predloženog algoritma sa algoritmima visokih performansi publikovanim u dostupnoj literaturi

U ovoj disertaciji prikazano je poređenje performansi (tačnost i preciznost) predloženog algoritma za detekciju naglih promena kadrova sa postojećim algoritmima visokih performansi koji su predstavljeni u literaturi. Izvršena je analiza rezultata poređenja, i ukazano je na uočene prednosti i mane predloženog algoritma u odnosu na ostale algoritme.

Ova disertacija predstavlja rezime višegodišnjeg intenzivnog istraživanja u oblasti multimedije, konkretno indeksiranja i pretraživanja multimedijalnih baza. Predloženi algoritam uvodi određene novine u oblast koja se intenzivno istražuje poslednju deceniju. U literaturi je predstavljen veliki broj različitih realizovanih algoritama koji vrše detekciju naglih promena sadržaja. Bez obzira na veliku konkurenciju u ovoj oblasti, što predstavlja

poseban izazov, uvek se može pronaći nešto novo. Primarni cilj ovog istraživanja bio je kako realizovati algoritam visoke tačnosti za što veći broj klasa video sekvenci. Rezultati intenzivnih testiranja pokazali su da se u nameri uspelo, s tim da je ostalo još prostora za poboljšanje performansi. Naime, moguće je unaprediti klasifikaciju artefakata koja bi trebalo da rezultira povećanjem preciznosti algoritma. Sledeći korak u razvoju algoritma predstavlja proširenje mogućnosti detekcije algoritma na gradijentne promene kadrova. Ovaj razvojni korak bi trebalo da pomogne u unapređivanju primarne detekcije naglih promena kadrova u određenim slučajevima. Reč je o ispuštenim naglim promenama kadrova, maskiranim sa brzim kretanjima kamere i objekata unutar kadrova.

Literatura

- [CHE11] H. T. Chen, C. F. Wu, "Forecasting volatility in Shanghai and Shenzhen markets based on multifractal analysis", *Physica A-Statistical Mechanics and its Applications*, 390(16):2926-2935, 2011.
- [LAW04] S. Lawrence, D. Ziou, M-F. Auclair-Fortier, "Motion-insensitive detection of cuts and gradual transitions in digital videos", *Pattern Recognition and Image Analysis*, 14(1):109-119, 2004.
- [LE08] D.-D. Le, S. Satoh, T. D. Ngo, D. A. Duong, "A text segmentation based approach to video shot boundary detection", in *Proc IEEE international Workshop on Multimedia Signal Processing (MMSP)*, pages 702-706, 2008.
- [LEE00] H. C. Lee, C. W. Lee, S. D. Kim, "Abrupt shot change detection using an unsupervised clustering of multiple features", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 6. pages 2015-2018, 2000.
- [ABA06] M. Abadi, E. Grandchamp, "Texture features and segmentation based on multifractal approach", in *Proceedings of the 11th Iberoamerican Congress on Pattern Recognition*, 4225, pages 297-305, 2006.
- [ABD08] W. Abd-Almageed, "Online, simultaneous shot boundary detection and key frame extraction for sports videos using rank tracing", in *Proc. IEEE International Conference on Image Processing (ICIP)*, pp. 3200-3203, 2008.
- [ADC04] J. Adcock, A. Girgensohn, M. Cooper, T. Liu, L. Wilcox, E. Rieffel, "FXPAL experiments for TRECVID 2004", in *Proceedings of the*

TRECVID 2004 Workshop, Gaithersburg, MD, USA. NIST, 2004.

- [AMI04] A. Amir, J. O. Argillander, M. Berg, S-F. Chang, M. Franz, W. Hsu, G. Iyengar, J. R. Kender, L. Kennedy, C-Y. Lin, M. Naphade, A. P. Natsev, J. R. Smith, J. Tesic, G. Wu, R. Yan, D. Zhang, "IBM research TRECVID-2004 video retrieval system", *In Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [AMI05] A. Amir, J. Argillander, M. Campbell, A. Haubold, G. Iyengar, S. Ebadollahi, F. Kang, M.R. Naphade, A. P. Natsev, J.R. Smith, J. Tesic, T. Volkmer, "IBM Research TRECVID-2005 Video Retrieval System", *in Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [ASC99] S. Ascher, E. Pincus, "The filmmaker's handbook: a comprehensive guide for the digital age", Plume, New York. 1999.
- [BAC10] A. R. Backes, O. M. Bruno, "Shape classification using complex network and multi-scale fractal dimension", *Pattern Recognition Letters*, (31):44-51, 2010.
- [CAI05] C. Cai, K. M. Lam, Z. Tan, "TRECVID2005 Experiments in The Hong Kong Polytechnic University: Shot Boundary Detection Based on a Multi-Step Comparison Scheme", *in Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [CAL02] J. Calic, E. Izquierdo, "Temporal segmentation of MPEG video streams", *EURASIP Journal on Applied Signal Processing*, (6):561-565, 2002.
- [CAL08] L. Calvet, A. Fisher, "*Multifractal volatility: Theory, forecasting and pricing*", Academic Press, San Diego, CA, 2008.
- [CAN96] C. Canus, J.L. Vehel, "Change detection in sequences of images by multifractal analysis", *in the Proceedings of the IGARSS'96*, (3):1609-1611, 1996.
- [CAO06] J. Cao, Y. Lan, J. Li, Q. Li, X. Li, F. Lin, "Intelligent multimedia group of Tsinghua University at TRECVID 2006", *In Proceedings of the TRECVID 2006 Workshop*, Gaithersburg, MD, USA. NIST, 2006.
- [CHA01] S.-F. Chang, T. Sikora, and A. Puri, "Overview of the MPEG-7 standard", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp.688 - 695, 2001.
- [COT06] C. Cotsaces, N. Nikolaidis, I. Pitas, "Video shot detection and condensed representation. a review", *IEEE Signal Processing Magazine*, 23(2):28-37, 2006.
- [DAM07] U. Damnjanovic, E. Izquierdo, "Shot boundary detection using spectral clustering", *in Proceedings of the 15th European Signal Processing*

- Conference (EUSIPCO 2007), pp.1779-1783, 2007.
- [DEB05] S. Deb, ed. “*Video Data Management and Information Retrieval*”, IRM Press, 2005.
- [DEL09] M. I. P. de Lima, J. L. M. P. de Lima, “Investigating the multifractality of point precipitation in the Madeira archipelago”, *Nonlin. Processes Geophys.*, 16, pp. 299-311, 2009.
- [DON10] A. Donate, X. Liu, “Shot Boundary Detection in Videos Using Robust Three-Dimensional Tracking”, *First International Workshop on Three Dimensional Information Extraction for Video Analysis and Mining (in conjunction with CVPR)*. San Francisco, California, 2010.
- [EDW91] T. Edwards, “Discrete wavelet transforms: Theory and implementation”, *Stanford University*, September 1991.
- [EKE07] H.K. Ekenel, M. Fischer, H. Gao, K. Kilgour, J.S. Marcos, R. Stiefelhagen, “Universität Karlsruhe (TH) at TRECVID 2007”, in *Proceedings of the TRECVID 2007 Workshop*, Gaithersburg, MD, USA. NIST, 2007.
- [EVE96] C. J. G. Evertsz, H. O. Peitgen, R. F. Voss, “*Fractal Geometry and Analysis: The Mandelbrot Festschrift*”, Curacao 1996, World Scientific Pub Co Inc, London. 1996.
- [EZE03] S. Ezekiel, “Medical Image Segmentation using multifractal analysis”, *Proceedings of the Applied Informatics*, (378):220-224, 2003.
- [FAL03] K. J. Falconer, “*Fractal Geometry: Mathematical Foundations and Applications*”, Second Edition, John Wiley & Sons, New York, 2003
- [FRA13] FRACLAB: A fractal analysis toolbox for signal and image processing, <http://fraclab.saclay.inria.fr/>, 2013. (accessed in May 2013.)
- [FUR11] S. Furuya, K. Yakubo, ”Multifractality of complex networks”, *Phys. Rev. E* 84, 036118 (2011), arXiv:1110.0272v1 [physics.soc-ph], 2011.
- [GAV13] A. Gavrovska, G. Zajić, I. Reljin, and B. Reljin, “Classification of Prolapsed Mitral Valve versus Healthy Heart from Phonocardiograms by Multifractal Analysis”, *Computational and Mathematical Methods in Medicine*, Vol. 2013, Article ID 376152, 10 pages, 2013.
- [GER07] D. Gerónimo, A. López, , D. Ponsa, A. D. Sappa, “Haar wavelets and edge orientation histograms for on-board pedestrian detection”, in *Pattern Recognition and Image Analysis* (pp. 418-425), Springer Berlin Heidelberg, 2007.
- [GRE09] C. Grecos, M. Yang , “An improved rate control algorithm based on a novel shot detection scheme for the H.264/AVC standard”, *Real Time Image Processing*, 4(1). 2009.

- [GUN11] E. Gunal, S. Canbek, N. Adar, “Fractal Dimension Based Shot Transition Detection in Sport Videos”, *Journal of Software Engineering and Applications*, 4(4):235-243, 2011.
- [HAN02] A. Hanjalic, “Shot-Boundary Detection: Unraveled and resolved?”, *IEEE Trans. Circuits and Ssystems for Video technology*, 12(2):90-105, 2002.
- [HAN99] A. Hanjalic, R. L. Lagendijk, J. Biemond, “Automated high-level movie segmentation for advanced video-retrieval systems”, *IEEE Transactions on Circuits and Systems for Video Technology*, v.9 n.4, p.580-588, June 1999.
- [HE10] B. He, “The temporal structures and functional significance of scal-free brain activity”, *Neuron*, 66:353-369, 2010.
- [HSU07] W. Y. Hsu, C. C. Lin, M. S. Ju , Y. N. Sun, “Wavelet-based fractal features with active segment selection: application to single-trial EEG data”, *Journal of Neuroscience Methods*, (163):145-160, 2007.
- [ISO01] MPEG-7: Visual Experimentation Model (XM) Version 10.0. *ISO/IEC/JTC1/SC29/WG11*, Doc. N4062 (2001).
- [ITU00] International Telecommunications Union – Telecommunication Standardization Sector, “Information technology – Generic coding of moving pictures and associated audio information: Video”, *ITU-T H.262*, February 2000.
- [ITU10] International Telecommunications Union – Telecommunication Standardization Sector, “Advanced video coding for generic audiovisual services”, *ITU-T H.264*, March 2010.
- [ITU13] International Telecommunications Union – Telecommunication Standardization Sector, “Information technology – High efficiency video coding”, *ITU-T H.265*, April 2013.
- [IVA07] P. C. Ivanov, ”Scale-invariant aspects of cardiac dynamics”, *IEEE Eng. in Med. and Biol. Mag.*, 26(6):33-37, 2007.
- [KAW07] Y. Kawai, H. Sumiyoshi, N. Yagi, ”Shot Boundary Detection at TRECVID 2007”, in *Proceedings of the TRECVID 2007 Workshop*, Gaithersburg, MD, USA. NIST, 2007.
- [KED02] A. Kedzia, M. Rybaczuk, R. Andrzejak, “Fractal dimensions of human brain cortex vessels during the fetal period”, *Medical Science Monitor*, 8(3):46-51, 2002.
- [KEN98] J. R. Kender, B. L. Yeo, “Video scene segmentation via continuous video coherence”, in *Proc. IEEE computer vision and pattern recognition*, pp. 367:373, 1998.

- [KIM12] T. Kim, D.-C.Park, D.-M. Woo, T.Jeong, S.-Y. Min “,Multiclass Classifier-Based Adaboost Algorithm,"*Intelligent Science and Intelligent Data Engineering*, Vol.7202, pp.122-127, 2012.
- [KOT03] K. Kotani, K. Takamasu, L. Safonov, Y. Yamamoto,“Multifractal heart rate dynamics in human cardiovascular model”, *in Proceedings of the SPIE*, (5110):340-347, 2003.
- [LI03] Z. N. Li, M. Drew”, *Fundamentals of Multimedia*”, 1st ed, Prentice-Hall, 2003.
- [LI09] J. Li, Y. Ding, Y. Shi, Q. Zeng, ”DWT-Based Shot Boundary Detection Using Support Vector Machine”, *Proceedings of the Fifth International Conference on Information Assurance and Security*, pages 435-438, 2009.
- [LIE99] R. Lienhart, “Comparision of automatic shot boundary detection algorithms”, *in Proc. SPIE Storage and retrieval for Image and Video Databases*, pp. 290-301,1999.
- [LIU07] Z. Liu, E. Zavesky, D Gibbon, B. Shahraray, P. Haffner”, AT&T research at TRECVID 2007”, *in Proceedings of the TRECVID 2007 Workshop*, Gaithersburg, MD, USA. NIST, 2007.
- [LOP09] R. Lopes, N. Betrouni, “Fractal and multifractal analysis: a review. *Medical Image Analysis*”, 13(4):634-649, 2009.
- [LU13] Z.-M. Lu, Y. Shi, “Fast Video Shot Boundary Detection Based on SVD and Pattern Matching”, *IEEE Transactions on Image Processing*, 22(12): 5136-5145, 2013.
- [MAN02] B.S. Manjunath, P. Salembier, T. Sikora, “*Introduction to MPEG-7: Multimedia Content Description Interface*”, John Wiley & Sons, Inc, New York, 2002.
- [MAN04] Y. Mansury, T. Deisboeck, “Simulating "structure-function" patterns of malignant brain tumors”, *Physica A*, 331(1-2):219-232, 2004.
- [MAN83] B. B. Mandelbrot”, *The Fractal Geometry of Nature*”, WH Freeman, Oxford. 1983.
- [MAT05] K. Matsumoto, M. Sugano, M. Naito, K. Hoashi, H.Kato, M. Shshibori, K. Kita, F.Sugaya, Y. Nakajima, “Shot Boundary Detection and Low-Level Feature Extraction Experiments for TRECVID 2005,” *in Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [MAV06] M. Mavroforakis, H. Georgiou, N. Dimitropoulos, D. Cavouras , S. Theodoridis, ”Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers”, *Artificial Intelligence In Medicine*,

37, 145-162, 2006.

- [MIS13] R. Mishra, S. K. Singhai, M. Sharma, "Video Shot Boundary Detection Using Dual Tree Complex Wavelet Transform", *IACC, 3rd IEEE International Conference*, pp. 1201-1206, 22-23 Feb. 2013.
- [MOH12] P. P. Mohanta, S. K. Saha, B. Chanda, "A Model-Based Shot Boundary Detection Technique Using Frame Transition Parameters", *IEEE Transactions on Multimedia*, v.14 n.1, p.223-233, February 2012
- [MUH07] M. Mühlhling, R. Ewerth, T. Stadelmann, C. Zöfel, B. Shi, Bernd Freisleben, "University of Marburg at TRECVID 2007: Shot Boundary Detection and High Level Feature Extraction", in *Proceedings of the TRECVID 2007 Workshop*, Gaithersburg, MD, USA. NIST, 2007.
- [NAC05] U. Naci, A. Hanjalic, "TU DELFT at TRECVID 2005: Shot Boundary Detection", in *Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [PAG07] R. Le Page, J.M. Goujon, M. Billon, P. Guyader, "A simple and efficient algorithm with a coarse to fine approach for arterial beat pulse detection", in *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, Lyon, France, August 23–26, 2007.
- [PAR00] K. Park, W. Willinger, "Self-similar network traffic: An overview", in *K. Park and W. Willinger, editors, Self-Similar Network Traffic and Performance Evaluation*, pages 1-38. Wiley, 2000.
- [PET04] C. Petersohn, "Fraunhofer HHI at TRECVID 2004: Shot boundary detection system", in *Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [PET10] C. Petersohn, "Temporal Video Segmentation", Jörg Vogt Verlag, 2010.
- [POP05] D. Popivanov, S. Jivkova, V. Stomonyakov, G. Nicolova, "Effect of independent component analysis on multifractality of EEG during visual-motor task", *Signal Processing*, 85(11):2112-2123, 2005.
- [POP06] M. Popović, "Digitalna obrada slike", 1st Ed., Akademska misao. 2006.
- [PRI04] L. Primaux, J. Benois-Pineau, P. Kramer, J-P. Domenger, "Shot boundary detection in the framework of rough indexing paradigm", in *Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [QUE04] G. M. Quenot, D. Moraru, S. Ayache, M. Charhad, "CLIPS-LIS-LSR-LABRI experiments at TRECVID 2004", in *Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [RAO90] K.R. Rao, P. Yip, "Discrete cosine transform: Algorithms, advantages, applications", Academic Press, Boston, 1990.

- [RELJ00a] I Reljin, B Reljin, I Pavlovic, I Rakocevic, "Multifractal analysis of gray-scale images", in *Proceedings of the 10th Mediterranean Electrotechnical Conference, MELECON-2000*, May 29-31, Limasol, Cyprus, II:490-3, 2000.
- [RELJ00b] I. Reljin, B. Reljin, G. Jovanovic, "Multifractal analysis of El Nino climate influence", First Balkan Conference on Signal Processing, Communications, Circuits and Systems, Balkan '2000, in Conference CD, Istanbul (Turkey), June 2-3, 2000.
- [RELJ02] I. Reljin, B. Reljin, "Fractal geometry and multifractals in analyzing and processing medical data and images", *Archive of Oncology*, 10(4):283-293, 2002.
- [RELJ06] I. Reljin, A. Samcovic, B. Reljin, "H.264/AVC compressed video traces: Multifractal and fractal analysis", *EURASIP Journal on Applied Signal Processing – Special Issue on Advanced Video Technologies and Applications for H.264/AVC and Beyond*, 2006:1-13, Article ID 75217, 2006.
- [REN09] J. Ren, J. Jiang, J. Chen, "Shot Boundary Detection in MPEG Videos Using Local and Global Indicators", *IEEE Trans. Circuits Syst. Video Techn.* 19(8): 1234-1238 , 2009.
- [RIS13] D. Ristic, M. Pavlovic, D. Sumarac Pavlovic, I. Reljin, "Detection of early reflections using multifractals", *The Journal of the Acoustical Society of America*, 133(4) :EL235–EL241, 2013.
- [ROB04] O. D. Robles, P. Toharial, A. Rodriguez, L. Pastor, "Using adaptive thresholds for automatic video cut detection", in *Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [SAL11] R. A. Salmon, M. G. Armstrong, S. J. E. Jolly, "Higher Frame Rates for more Immersive video and television", BBC R&D Publications, October 2011.
Online available: <http://www.bbc.co.uk/rd/publications/whitepaper209>
- [SCH07] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, 17 (9), pp: 1103-1120, September 2007.
- [SER10] F. Serinaldi, "Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models", *Nonlin. Proc. Geophys.*, 17(2010), pp. 697-714. 2010.
- [SHE08] H. Shen , X. Sun , F. Wu, "Fast H.264/MPEG-4 AVC Transcoding Using Power-Spectrum Based Rate-Distortion Optimization", *IEEE Transactions on Circuits and Systems for Video Technology*, v.18 n.6,

p.746-755, June 2008.

- [SHE11] O. I. Sheluhin, A. A. Atayero, A. B. Garmashev, "Detection of Teletraffic Anomalies Using Multifractal Analysis", *International Journal of Advancements in Computing Technology, AICIT*, 3(4):174-182, 2011.
- [SHI04] Y. Shimizu, M. Barth, C. Windischberger, E. Moser, S. Thurner, "Wavelet based multifractal analysis of fMRI time series", *NeuroImage*, 22(3):1195-1202, 2004.
- [SIK01] T. Sikora, "The MPEG-7 visual standard for content description&mdash, an overview", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp.696 -702, 2001.
- [SON01] B. C. Song, J. B. Ra, "Automatic shot change detection algorithm using multi-stage clustering for MPEG-compressed videos", *Journal of Visual Communication and Image Representation*, 12(3):364-385, 2001.
- [STE08] I. Steinwart, A. Christmann, "*Support Vector Machines. Information Science and Statistics*", Springer, New York, 2008.
- [STO06] T. Stojic, I. Reljin, B. Reljin, "Adaptation of multifractal analysis to segmentation of microcalcifications in digital mammograms", *Physica A: Statistical Mechanics and its Applications*, (367):494-508, 2006.
- [TRE07] NIST, Homepage of Trecvid Evaluation, [online] Available: <http://www-nlpir.nist.gov/projects/trecvid/>
- [VEN02] J. Vendrig, M. Worring, "Systematic evaluation of logical story unit segmentation", *IEEE Trans. Multimedia*, 4(4):492-499, 2002.
- [VOL04] T. Volkmer, S. Tahaghoghi, H. E. Williams, "RMIT University at TRECVID 2004", *In Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD USA. NIST, 2004.
- [WAL11] A. C. Walbran, C. P. Unsworth, A. J. Gunn, L. Bennet, "Spike detection in the preterm fetal sheep EEG using Haar wavelet analysis", *In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pp. 7063-7066, 2011.
- [WAL91] G. K. Wallace, "The JPEG Still Picture Compression Standard", *Commun. ACM*, vol. 34, pp.30 -44, 1991.
- [WAN07] G. Wang, L. Xiao, A. He, "An improved computing method for the image edge detection", *Chinese Optics Letters*, 5(2):79-81, 2007.
- [WAN11] D. Wang, "Multifractal characterisation and analysis of complex networks", *Disertation*, Queensland University of Technology, 2011.
- [XIA06] Y. Xia, D. Feng, R. Zhao, "Morphology-based multifractal estimation for texture segmentation", *IEEE Transactions on Image Processing*,

15(3):614-624, 2006.

- [XIE10] S. Y. Xie, Q. M. Cheng, Q. C. Ling, B. Li, Z. Y. Bao, P. Fan, "Fractal and multifractal analysis of carbonate pore-scale digital images of petroleum reservoirs", *Marine and Petroleum Geology*, 27(2):476-485, 2010.
- [YI07] W. J. Yi, M. S. Heo, S. S. Lee, S. C. Choi, K. H. Huh, S. P. Lee, "Direct measurement of trabecular bone anisotropy using directional fractal dimension and principal axes of inertia", *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics*, 104(1):110-116, 2007.
- [YU05] Z. J. Yu, S. V. N Vishwanathan, A. Smola, "NICTA at TRECVID 2005 Shot Boundary Detection Task", in *Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [YUA04] J. Yuan, W. Zheng, L. Chen, D. Ding, D. Wang, Z. Tong, H. Wang, J. Wu, J. Li, F. Lin, B. Zhang, "Tsinghua University at TRECVID 2004: Shot Boundary Detection and High-level Feature Extraction", in *Proceedings of the TRECVID 2004 Workshop*, Gaithersburg, MD, USA. NIST, 2004.
- [YUA05] J. Yuan, H. Wang, L. Xiao, D. Wang, D. Ding, Y. n Zuo, Z. Tong, X. Liu, S. Xu, W. Z. X. Li, Z. Si, J. Li, F. Lin, B. Zhang, "Tsinghua University at TRECVID 2005", in *Proceedings of the TRECVID 2005 Workshop*, Gaithersburg, MD, USA. NIST, 2005.
- [YUA07] J. Yuan, Z. Guo, L. Lv, W. Wan, T. Zhang, D. Wang, X. Liu, C. Liu, S. Zhu, D. Wang, Y. Pang, N. Ding, Y. Liu, J. Wang, X. Zhang, X. Tie, Z. Wang, H. Wang, T. Xiao, Y. Liang, J. Li, F. Lin, B. Zhang, "THU and ICRC at TRECVID 2007", in *Proceedings of the TRECVID 2007 Workshop*, Gaithersburg, MD, USA. NIST, 2007.
- [ZAB95] R. Zabih, J. Miller, K. Mai, "A feature-based algorithm for detecting and classifying scene breaks", in *ACM Multimedia*, pp. 189-200, 1995.
- [ZAJ06] G. Zajić, N. Kojić, V. Radosavljević, S. Čabarkapa, B. Reljin, "Feature Vector Reduction in CBIR System with Relevance Feedback", in *Proc. 13th Inter. Conf. IWSSIP-2006*, pp. 479-482, Budapest, Hungary, 21-23, Sept. 2006.
- [ZAJ07] G. Zajić, N. Kojić, V. Radosavljević, M. Rudinac, S. Rudinac, N. Reljin, I. Reljin, B. Reljin, "Accelerating of Image Retrieval in CBIR System with Relevance Feedback", *EURASIP Journal on Advances in Signal Processing, Spec. Issue on Knowledge Assisted Media Analysis for Interactive Multimedia Applications*, 2007:1-13, Article ID 62678, 2007.
- [ZAJ08] G. Zajić, N. Kojić, N. Reljin, B. Reljin, "Experiment with Reduced Feature Vector in CBIR System with Relevance Feedback", *IET Conf. Publications*, vol. 2008, no. CP543, pp.176 -181, 2008.

- [ZAJ11] G. Zajić, I. Reljin, and B. Reljin, "Video Shot Boundary Detection based on Multifractal Analysis", *Telfor Journal*, Vol. 3, No. 2, pp.105-110, 2011.
- [ZAJ12] G. Zajić, N. Kojić, B. Reljin, "Improving CBIR Systems Using Automated Ranking," *Telfor Journal*, Vol. 4, No. 2, pp.116-121, 2012.
- [ZEI10] Z. Zeinalpour-Tabrizi, A. Aminian-Modarres, M. Fathy, M. Jahed-Motlagh, "Fractal Based Video Shot Cut/Fade Detection and Classification", *Active Media Technology, Lecture Notes in Computer Science*, (6335):128-137, 2010.
- [ZHA07] Z.C.Zhao, X.Zeng,T. Liu, A.-N. Cai, "BUPT at TRECVID 2007: Shot Boundary Detection", in *Proceedings of the TRECVID 2007Workshop*, Gaithersburg, MD, USA. NIST, 2007.

Biografija autora

Osnovni podaci

Goran J. Zajić rođen je 1972. godine, u Rijeci, Republika Hrvatska. Osnovnu školu “Veljko Vlahović” završio je 1987. godine u Splitu. Maturirao 1991. godine u Tehničkom školskom centru “Mirko Dumanić” u Splitu, smer Računarska tehnika. Diplomirao 2005. godine na Elektrotehničkom fakultetu, Univerziteta u Beogradu, na odseku Fizička elektronika, smer Optoelektronika i laserska tehnika sa prosečnom ocenom u toku studija 8,05, ocena na diplomskom 10. Upisao doktorske studije na Elektrotehničkom fakultetu 2007. godine, Univerzitet u Beogradu, smer Telekomunikacije i položio sve ispite sa prosečnom ocenom 10,00.

Dodatna usavršavanja i angažovanja:

Dodatna usavršavanja kroz veliki broj konferencija i radionica. Redovni posetilac *Gamax Ltd.* konferencija (*Matlab, Simulink*), itd. Usavršavao svoje sposobnosti kroz Cisco kurseve (CCNA). Kroz različite kurseve za potrebe projekata unapređivao svoje znanje (obuka za *Oracle 10g, Microsoft SQL Server 2008, Microsoft .NET Framework Applications, C#*, itd.). Posедуje certifikate ”*Database Design and Programming with SQL*” i “*JAVA Fundamentals*”, *ORACLE Academy*.

Usavršavao poznavanje jezika kroz tečajeve (tečajevi Instituta za strane jezike, Beograd). Poznaje: engleski i italijanski jezik.

Radom na međunarodnim projektima stekao značajno iskustvo u saradnji sa ljudima iz različitih kulturnih i etničkih sredina. Uvek zainteresovan za naučna istraživanja i inovacije, zaobilaznje standardnih metodologija i tehnika u cilju napretka i isplativosti.

Član IPTM grupe za digitalnu obradu slike, telemedicinu i multimediju od 2005 godine. Recenzent radova (TELFOR, HINDAWI *Publishing Corp.*) i predsedavajući sesije za multimediju (TELFOR). Angažovan kao član komisije na veštačenju autentičnosti video zapisa u krivičnim postupcima od strane Elektrotehničkog fakulteta u Beogradu.

Trenutno zaposlen u Visokoj školi strukovnih studija za informacione i komunikacione tehnologije u Beogradu, na poziciji saradnika u laboratoriji.

Projekti

Angažovan na realizaciji i programu istraživanja na domaćim i međunarodnim projektima:

- “COST Action IC0604, “*Anatomic Telepathology Network (EURO-TELEPATH)*”, 2007-2012, istraživač.
- "*Development of the System for Digitalization of Medical Images, Archiving in a Image Database and Database Searching*", Innovation project of the Ministry of Science of Serbia for the year 2008/2009, istraživač
- COST Action 292, “*Semantic Multimodal Analysis of Digital Media*”, 2004-2008, istraživač.

Objavljeni radovi

Autor je i koautor više od četrdeset naučnih radova izlaganih i objavljenih na domaćim i međunarodnim konferencijama, kao i u domaćim i međunarodnim časopisima (bez i sa impakt faktorom).