

УНИВЕРЗИТЕТ У БЕОГРАДУ  
ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ

Владимир Б. Ковачевић

ЕСТИМАЦИЈА ПОКРЕТА У ВИДЕО  
СЕКВЕНЦАМА ПРИЛАГОЂЕНА  
АЛГОРИТМИМА ЗА ПОВЕЋАЊЕ БРОЈА  
СЛИКА У СЕКУНДИ

докторска дисертација

Београд, 2015

UNIVERSITY OF BELGRADE  
FACULTY OF ELECTRICAL ENGINEERING

Vladimir B. Kovacevic

MOTION ESTIMATION IN VIDEO SEQUENCES  
ADAPTED TO ALGORITHMS FOR FRAME-  
RATE UPCONVERSION

Doctoral Dissertation

Belgrade, 2015

## **ПОДАЦИ О МЕНТОРУ И ЧЛАНОВИМА КОМИСИЈЕ**

### **Ментор:**

Проф. др Ирини Рељин, Универзитет у Београду - Електротехнички факултет, редовни професор

### **Чланови комисије:**

Проф. др Милан Бјелица, Универзитет у Београду - Електротехнички факултет

Проф. др Ирини Рељин, Универзитет у Београду - Електротехнички факултет

Проф. др Лазар Сарановац, Универзитет у Београду - Електротехнички факултет

Проф. др Милан Прокин, Универзитет у Београду - Електротехнички факултет

Проф. Др Андреј Самчовић, Универзитет у Београду - Саобраћајни факултет

Датум одбране:

## Захвалница

Захваљујем се свом ментору, професорки Ирени Рељин која ме је подржала у истраживању и дала веома битне смернице посебно при објављивању рада.

Захвалност бих изразио и компанији Интел која је дозволила да мој рад буде искоришћен и публикован у овој тези.

Захваљујем се колегама из компаније Интел Здравку Пантићу, Радомиру Јаковљевићу, Александру Берићу, Владимиру Илићу и Милошу Марковићу који су ми помогли у изради и имплементацији описаног алгорита.

Посебно се захваљујем својим пријатељима и драгим људима који су ме окруживали све ове године, а пре свега родитељима Бориславу и Ђурђици на несебичној и бескрајној подршци, љубави и разумевању.

Владимир Ковачевић

## ЕСТИМАЦИЈА ПОКРЕТА У ВИДЕО СЕКВЕНЦАМА ПРИЛАГОЂЕНА АЛГОРИТМИМА ЗА ПОВЕЋАЊЕ БРОЈА СЛИКА У СЕКУНДИ

**Резиме** - Ова теза описује алгоритам за естимацију покрета у видео секвенцама прилагођену алгоритмима за повећање броја фрејмова у секунди. Предложена метода комбинује рекурзивни блоковски приступ и прилагођену фазну планарну корелацију. У поређењу са алтернативним приступима предложени алгоритам даје боље субјективне импресије и објективна мерења (*Peak Signal to Noise Ratio* и *Universal Image Quality Index*). Рачунарска комплексност алгоритма је смањена и погодна за извршавање у реалном времену на платформама са лимитираним меморијским и процесорским ресурсима. Као доказ описаног концепта дата је прва имплементација у реалном времену алгоритма за повећање броја слика (фрејмова) у секунди на мобилној платформи (*Intel Merrifield*). Она омогућава процесирање 1080п видео секвенце са 30 фрејмова у секунди и процесором учестаности 200 мегагерца. Поред предложеног алгоритма описан је и значај естимације покрета у видео секвенцама, типови, као и преглед постојећих решења, кроз историју.

Кључне речи: естимација покрета, фазна планарна корелација, блок мечинг, извршавање у реалном времену, мобилна платформа, повећање броја слика у секунди у видео секвенцама

Научна област: техничке науке, електротехника

Ужа научна област: Телекомуникације (мултимедија), Дигитална обрада сигнала

УДК број:

## MOTION ESTIMATION IN VIDEO SEQUENCES ADAPTED TO ALGORITHMS FOR FRAME-RATE UPCONVERSION

**Abstract** - This thesis describes a novel motion estimation algorithm adapted for frame-rate upconversion. The method combines recursive block-matching and customized phase plane correlation. In comparison with alternative approaches the proposed solution shows better subjective impressions and objective measurements, such as Peak Signal to Noise Ratio and Universal Image Quality Index. Computational complexity of the algorithm is constrained and suitable for real-time implementations on platforms with limited memory and processing resources. The proof of concept is the first real-time implementation of frame-rate up-conversion on a mobile platform (Intel Merrifield), which is capable of processing 1080p video at 30 frames per second while clocked at 200 MHz. Besides proposed solution, this thesis describes the importance and usages of motion estimation, its types and currently available solutions, throughout history.

Key words: motion estimation, phase plane correlation, block-matching, real-time execution, mobile platform, frame-rate up-conversion

Scientific area: technical sciences, electrical engineering

Specific scientific area: Telecommunications (multimedia), Digital signal processing

UDK number:

## Садржај

Списак слика.....	7
Списак табела .....	11
Списак скраћеница по редоследу коришћења .....	12
Предговор.....	13
Циљ истраживања .....	14
Полазне хипотезе истраживања .....	14
Резултати и допринос истраживања .....	14
Преглед садржаја докторске дисертације .....	15
<b>1. Увод .....</b>	<b>16</b>
1.1 Шта је естимација покрета и њене примене .....	17
1.2 Врсте естимација.....	17
1.3 Шта је <i>frame rate up-conversion</i> ? .....	21
<b>2. Преглед постојећих алгоритама за естимацију покрета ...</b>	<b>24</b>
2.1 Блоковске методе .....	26
2.2 Корелационе методе .....	30
2.3 Блоковско-корелационе методе .....	33
2.4 Методе са оптичким током .....	34
2.5 Методе са додатним процесирањем .....	35
2.6 Употреба естимације покрета у видео компресији .....	37

<b>3. Блоковска корелациона метода естимације кретања.</b>	<b>41</b>
3.1 Прилагођена фазна планарна корелација .....	43
3.2 Блок мечинг .....	47
3.3 Процесирање ивичних блокова .....	50
3.4 Пример употребе блоковске корелационе естимације .....	54
<b>4. Резултати и тестирање.....</b>	<b>59</b>
<b>5. Закључак .....</b>	<b>79</b>
<b>Додатак А: Методе за поређење сличности између слика.</b>	<b>83</b>
5.1 Peak signal to noise ratio (PSNR).....	83
5.2 Universal Image Quality Index (UIQI).....	83
<b>Додатак Б: Патентне апликације на којима је базирана теза .</b>	<b>85</b>
<b>Литература.....</b>	<b>88</b>



## Списак слика

Слика 1-1: Пример поделе фрејма на објекте, при чему је могуће сваком објекту доделити вектор помераја. Могуће је такође и да објекти деле исти вектор помераја, што је веома често и случај. .... 19

Слика 1-2: Унидирекциона (слика горе) и бидирекциона (слика доле) естимација покрета. На горњој слици референца (полазна тачка) је на претходном фрејму (*Previous frame*) одакле се тражи парњак сваког блока на следећем фрејму (*Next frame*). Уочава се да блок који се преноси вектором помераја из претходног на следећи фрејм не мора пасти тачно на GRID интерполираног фрејма (*Interpolated frame*), због чега је могуће да неки пиксели интерполираног фрејма остану без вектора помераја. На слици доле референца естимације је на интерполираном фрејму одакле се тражи најсличнији пар блокова са претходног и следећег фрејма. Слика преузета из [3] ..... 20

Слика 1-3: Илустрација повећања броја фрејмова у секунди креирањем интерполираних фрејмова са компензованим кретањем. .... 21

Слика 1-4: Пример потребног позиционирања бидирекционог поља вектора за неколико различитих фактора конверзије видео секвенце. .... 22

Слика 2-1: Преглед категорија алгоритама за естимацију кретања у видео секвенцама. Методе са додатним процесирањем користе блоковске или методе са оптичким током за иницијалну естимацију кретања. Блок мечинг корелација (БМК), предложена у овој тези, истакнута је у групи блоковско-корелационих метода. .... 25

Слика 2-2: Претраживање у три корака – поступак избора блокова за срачунавање суме апсолутних разлика. Срачунати вектор помераја на овом примеру је  $(v_x, v_y) = (4, -4) + (2, -2) + (-1, -1) = (5, -7)$ . .... 27

Слика 2-3: Патерн коришћен у *Diamond search* алгоритму. Лево - велики дијамант, десно - мали дијамант патерн. .... 28

Слика 2-4: Поступак претраживања блокова предложен у *diamond search* алгоритму. Алгоритам стартује од блока (0, 0), пронаћи најбољи парњак доле десно (слика горе лево). Затим, у кораку два (слика горе-десно) најбољи парњак је десно. У трећем кораку (слика доле-лево) најбољи парњак је у центру дијаманта. У четвртном кораку примењује се мали дијамант патерн за финално проналажење најсличнијег блока (слика доле десно). .... 29

Слика 2-5: Корелациона равна која одговара пару региона између којих нема кретања (слика горе) и корелациона равна која одговара пару региона између којих је померај -8 по хоризонталном смеру (слика доле). Преузето из [15] . .... 32

Слика 2-6: Приказ суседних региона из којих се узимају кандидати за вектор помераја. Смер срачунавања фазних корелација по регионима је од горе ка доле и од лево

ка десно. Из поменутог се закључује да су кандидати 1-4 из тренутног пара фрејмова, а 5-9 из претходног пара фрејмова (коришћена је временска зависност). Преузето из [17] .....	33
Слика 2-7: MFHME алгоритам базиран на три суседна фрејма. ....	36
Слика 2-8: Етапе извршавања MFHME алгоритма.....	37
Слика 2-9: Процес кодовања једног фрејма (слајса) заједнички за већину видео компресија. ....	39
Слика 3-1: Блок дијаграм блоковско корелационе методе предложене у овој тези (БМК). Улази у алгоритам су вредности пиксела са претходног фрејма, следећег фрејма и поље вектора из претходне итерације БМК алгоритма – $MVF(n-1)$ . Излаз је вектор поље за тренутни пар фрејмова – $MVF(n)$ . ....	42
Слика 3-2: Редукција броја пиксела на улазним фрејмовима коришћена код БМК која се састоји од изостављања пиксела. Пример приказује слику димензија 16x16 са приказаним вредностима пиксела. После скалирања два пута само ће маркирани пиксели остати у излазној слици. ....	43
Слика 3-3: Пример поделе фрејма на регионе код глобалне фазне планарне корелације.....	44
Слика 3-4: Пример поделе фрејма на регионе код локалне фазне планарне корелације.....	45
Слика 3-5: Пример поделе фрејма на локалне и глобалне ФПК регионе за 1080п резолуцију, уз постизање исте величине стандардног ФПК региона и код глобалне и код локалне фазне планарне корелације.....	45
Слика 3-6: Дијаграм улазних и излазних података Блок мечинг модула.....	47
Слика 3-7: Пример мапирања ФПК кандидата на блокове фрејма. Блок $B$ ће примити глобалне кандидате из региона $GR1$ и локалне кандидате из региона $LR4$ . Резолуција фрејма у овом примеру је 512x256 пиксела. ....	48
Слика 3-8: Начин избора позиције просторних кандидата ( $S1$ и $S2$ ) и темпоралног кандидата $MT$ за блок $B$ . Темпорални кандидат за блок $B$ у пољу вектора које се тренутно срачунава $MVF(n)$ се формира користећи векторе помераја из претходно срачунатог поља $MVF(n-1)$ . ....	49
Слика 3-9: Интерполирани фрејм <i>Harry Potter</i> видео секвенце. Слика горе не садржи специјално процесирање ивичних блокова због чега су видљиви блоковски артефакти на левој ивици фрејма. Слика доле представља излаз алгоритма за естимацију са примењеним процесирањем ивичних блокова. ....	51
Слика 3-10: Интерполирани фрејм <i>College Football</i> видео секвенце. Слика лево не садржи специјално процесирање ивичних блокова због чега су видљиви блоковски артефакти на десној ивици фрејма. Слика десно представља излаз алгоритма за естимацију са примењеним процесирањем ивичних блокова и не садржи блоковске артефакте на десној ивици. ....	52
Слика 3-11: Репрезентација стварног кретања за ивичне блокове $B_{1i}$ и $B_{2i}$ . Вектори $MV_1$ и $MV_2$ су стварни вектори помераја између претходног и следећег фрејма.	

Применивши вектор $MV_2$ на блок $B_{2i}$ блокови $B_{2p}$ и $B_{2c}$ су у границама оба фрејма. Међутим, применивши вектор $MV_1$ на блок $B_{2i}$ блок $B_{1c}$ је ван следећег фрејма. ....	53
Слика 3-12: Смер процесирања (скенирања) блокова у фрејму. Фрејм је подељен на одређени број вертикалних целина – страјпова. Блокови се процесирају од горе ка доле, пратећи назначене стрелице. Такав редослед процесирања блокова омогућава доделу стварних вектора помераја ивичним блоковима. Вектор кандидат блока означеним зеленом бојом је просторни кандидат за назначени ивични блок. Вектор кандидати за блокове означене црвеном бојом још увек нису срачунати. Вектори са блокова означених љубичастом бојом такође нису подобни као кандидати за ивичне векторе помераја. ....	54
Слика 3-13: Преглед распореда извршавања делова БМК алгоритма на <i>Intel Merrifield</i> векторском процесорском систему. ....	55
Слика 3-14: Телефон базиран на <i>Intel Merrifield</i> платформи на коме се извршава алгоритам за повећање броја фрејмова у секунди у реалном времену. ....	58
Слика 4-1: Визуелно поређење интерполираних фрејмова видео секвенце <i>Swans</i> , креираних алгоритмом за повећање броја слика у секунди у видео секвенцама користећи различите методе за естимацију кретања: горе-лево - <i>FS</i> , горе-десно - <i>3DRS</i> , доле-лево - БМК, доле-десно – оригинални фрејм. Артефакти настали усред погрешне или недовољно прецизне естимације су заокружени црвеном бојом. ....	63
Слика 4-2: Визуелно поређење интерполираних фрејмова видео секвенце <i>Fruit yard</i> , креираних алгоритмом за повећање броја слика у секунди у видео секвенцама користећи различите методе за естимацију кретања: горе-лево - <i>FS</i> , горе-десно - <i>3DRS</i> , доле-лево - БМК, доле-десно - оригинални фрејм. Артефакти настали усред погрешне или недовољно прецизне естимације су заокружени црвеном бојом. ....	64
Слика 4-3: Оригинални фрејмови видео секвенце <i>Street traffic</i> (горе и доле) и интерполирани фрејм између њих. ....	65
Слика 4-4: Оригинални фрејмови видео секвенце <i>Matrix</i> (горе и доле) и интерполирани фрејм између њих. ....	66
Слика 4-5: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Foreman</i> , креиран помоћу описаног БМК алгоритма. ....	67
Слика 4-6: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Belgrade aerial</i> , креиран помоћу описаног БМК алгоритма. ....	68
Слика 4-7: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Conference call</i> , креиран помоћу описаног БМК алгоритма. ....	69
Слика 4-8: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Danube islands</i> , креиран помоћу описаног БМК алгоритма. ....	70
Слика 4-9: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Forest in fall</i> , креиран помоћу описаног БМК алгоритма. ....	71
Слика 4-10: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Children playing</i> , креиран помоћу описаног БМК алгоритма. ....	72

Слика 4-11: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Stockholm pan</i> , креиран помоћу описаног БМК алгоритма.....	73
Слика 4-12: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Down the street</i> , креиран помоћу описаног БМК алгоритма. ....	74
Слика 4-13: Два оригинална и интерполирани фрејм између њих видео секвенце <i>Tango argentino</i> , креиран помоћу описаног БМК алгоритма.....	75
Слика 4-14: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце <i>I robot</i> , креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама $n$ и $n+2$ . ....	76
Слика 4-15: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце <i>Matrix</i> , креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама $n$ и $n+2$ . ....	77
Слика 4-16: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце <i>Spiderman</i> , креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама $n$ и $n+2$ . ....	78
Слика 5-1: Дијаграм публикација и патентних апликациј аутора ове тезе везаних за област естимације кретања у видео секвенцама. Тамно зеленом бојом истакнуто оно што је објављено, светло зеленом су означене две патентне апликације које су објављене крајем 2015. г. и жутом публикације које су у плану у будућности. ....	82

## Списак табела

Табела 1: Измерене перформансе извршавања БМК алгоритма на Интел Мерифилд ( <i>Intel Merrifield</i> ) мобилној платформи. ....	56
Табела 2: Дисипирана снага при повећању броја фрејмова у секунди на 1080п видео секвенци. ....	57
Табела 3: Објективна мерења која пореде три методе за естимацију покрета: <i>Full search</i> , 3-димензионално рекурзивно претраживање ( <i>3DRS</i> ) и предложену Блок мечинг корелацију (БМК). Веће <i>PSNR</i> и <i>UIQI</i> вредности представљају већу сличност оригиналној секвенци (бољи квалитет естимације). ....	60
Табела 4: Поређење објективних метрика пет метода за естимацију кретања: <i>Full search</i> , <i>3-dimensional search</i> , <i>Multi-frame hierarchical motion estimation</i> , <i>Variational Image Fusion</i> и блок мечинг корелације (БМК). Веће <i>PSNR</i> вредности (маркиране сивом бојом у табели) представљају већу сличност оригиналној секвенци (бољи квалитет естимације). 61	
Табела 5: <i>PSNR</i> и <i>UIQI</i> представљају поређење између две инстанце БМК, прве са прилагођеном фазном планарном корелацијом (ПФПК) и друге са имплементацијом Фазне планарне корелације у бројчаном систему са покретном тачком (ФФПК). ....	62

## Списак скраћеница по редоследу коришћења

Скраћеница	Пун назив
FPS	<i>Frames per second</i> (Број фрејмова у видео секвенци у секунди)
БМК	Блок мечинг корелација (алгоритам предложен у овој тези)
SAD	<i>Sum of absolute differences</i> , сума апсолутних разлика
FS	<i>Full Search</i> , алгоритам за естимацију кретања заснован на евалуацији свих суседних блокова у фрејму при тражењу најсличнијег блока
3DRS	<i>Three dimensional recursive search</i> , први алгоритам за естимацију који користи временску и просторну рекурзију
ФПК	Фазна планарна корелација
DFT	<i>Discrete Fourier Transformation</i>
FFT	<i>Fast Fourier Transformation</i>
MFHME	<i>Multi-frame hierarchical motion estimation</i>
FPGA	<i>Field Programable Gate Array</i>
VIF	<i>Variational image fusion</i>
MPEG	<i>Moving Picture Experts Group</i>
AVC	<i>Advanced video coding</i>
<i>Y</i>	Луминентна компонента фрејма (мапа осветљаја) YCbCr простора боја
<i>Cb</i>	Хроминентна компонента фрејма (плава разлика) YCbCr простора боја
<i>Cr</i>	Хроминентна компонента фрејма (црвена разлика) YCbCr простора боја
ПФПК	Прилагођена фазна планарна корелација
БМ	<i>Block matching</i>
<i>MVF</i>	<i>Motion vector field</i> , поље вектора помераја
<i>ВП</i>	Векторски процесор
<i>СП</i>	Скаларни процесор
<i>MHz</i>	Mega hertz
<i>PSNR</i>	<i>Peak signal to noise ratio</i>
<i>UIQI</i>	<i>Universal Image Quality Index</i>
dB	Decibels
ФФПК	Фазне планарне корелације у бројчаном систему са покретном тачком

## Предговор

У дисертацији је представљен део истраживачког рада реализован у компанији Интел у периоду од 2010. до 2015. године. Компанија Интел обезбедила је све потребне хардверске ресурсе, симулаторе, као и мобилну платформу на којој се извршава описани алгоритам у реалном времену.

Естимација покрета представља једну од главних области везаних за обраду и процесирање видео секвенци. Користи се у алгоритмима за видео компресију, видео стабилизацију, поравнање слика (*alignment*) и повећање броја слика (фрејмова) у секунди (у даљем тексту *frame-rate upconversion* конверзију). *Frame-rate upconversion* подразумева повећање броја фрејмова у видео секвенци креирањем нових и користећи претходно естимирано кретање и пикселе оригиналних фрејмова. На тај начин временско трајање видео секвенце остаје непромењено, али се повећава број фрејмова који се приказује на дисплеј уређају (монитор, телевизор) у јединици времена. Предмет предложеног истраживања управо је везан је за естимацију покрета прилагођену употреби у алгоритмима за *frame-rate upconversion* са фокусом на постизање ниске комплексности.

## Циљ истраживања

Циљ истраживања је показати могућност конструисања алгоритма високог квалитета, који ће подржавати данашње стандарде везане за резолуцију и учестаност фрејмова у видео секвенцама. Други циљ је показати да се поменути алгоритам може извршавати у реалном времену на мобилној платформи.

## Полазне хипотезе истраживања

На почетку истраживања дате су две, међусобно супротстављене хипотезе:

- Могуће је развити алгоритам високе тачности и прецизности за естимацију покрета у видео секвенцама
- Могуће је прилагодити алгоритам за естимацију покрета у видео секвенцама извршавању у реалном времену на мобилној платформи лимитираних меморијских капацитета оптимизовањем броја потребних процесорских циклуса и меморијских трансфера.

Као полазну тачку у развоју алгоритма такође је дата и хипотеза да је могуће применити блоковски приступ допуњен естимацијом базираној на корелацији у фреквенцијском домену за естимацију покрета у видео секвенцама.

## Резултати и допринос истраживања

Значај извршавања естимације покрета у реалном времену је редуковање потребе за огромним додатним меморијским простором који био потребан за смештање свих новокреираних фрејмова уколико би се цела секвенца испроцесирала пре саме репродукције. На пример, за конверзију видео секвенце пуне HD резолуције (*High Definition*, 1920п, телевизија високе резолуције) трајања 10 секунди са 30 на 60 фрејмова у секунди потребно би било сместити целу видео секвенцу која би заузимала меморијски простор од 900 мегабајта.

Значај постојања већег броја фрејмова у секунди у видео секвенцама (на пример 60 фрејмова у секунди) огледа се у томе да она поспешује утисак гледаоца видео материјала уклањањем неугодног сецкања (*judder*) које настаје услед високе резолуције и контраста



самих фрејмова и недовољног *frame-rate*-а. Висок проценат постојећег видео материјала садржи нижи *frame-rate* (на пример 24 или 30 фрејмова у секунди). Естимација покрета описана у предложеној теми омогућава репродукцију тих секвенци у високим учестаностима фрејмова (на пример 60 фрејмова у секунди).

Теза је базирана на следећим публикацијама:

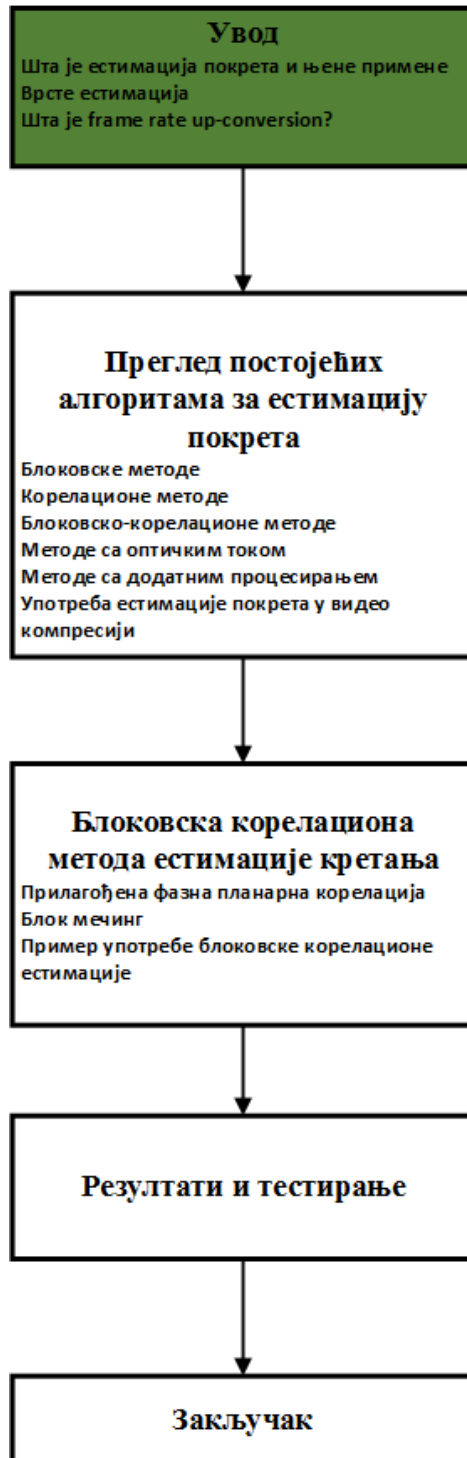
M23 – Vladimir Kovacevic, Zdravko Pantic, Aleksandar Beric, Radomir Jakovljevic, *Block-matching correlation motion estimation for frame-rate up-conversion*, The Journal of Signal Processing Systems, 2015.

Патент апликација: Beric, A., Pantic, Z., Kovacevic, V., *Motion estimation using hierarchical Phase Plane Correlation and Block Matching*, US Patent Application Number: 2014/0254678 A1, September 2014 (Додатак Б).

## **Преглед садржаја докторске дисертације**

У уводу докторске дисертације биће дата објашњења о томе шта представља естимација покрета у видео секвенцама и њен значај. Тренутно постојећа решења и алгоритми за естимацију кретања описани су у другом делу. У трећем делу описан је нови алгоритам, предложен овом тезом, који врши естимацију кретања засновану на блоковској корелацији. Четврти део описује резултате описаног алгоритма, као и његово поређење са постојећим решењима. Закључно поглавље описује значај предложеног решења и кораке за даљи рад. На крају је дата литература коришћена у изradi ове дисертације.

# 1. Увод



## 1.1 Шта је естимација покрета и њене примене

Естимација покрета (кретања) представља процес рачунања помераја сваког пиксела између два суседна фрејма видео секвенце. Сваки померај, по својој природи, има правац, смер и интензитет, па се због тога користи термин вектор помераја. Скуп свих вектора помераја за један фрејм (или пар фрејмова) зове се поље вектора помераја. Оно описује комплетно кретање свих пиксела тог фрејма (пара фрејмова).

Естимација покрета представља једну од главних области у оквиру обраде видео сигнала. Користи се у алгоритмима и апликацијама као што су повећање броја слика у секунди (*frame-rate up-conversion*), видео компресија, видео стабилизација, поравнавање слика (*image alignment*). Видео компресија, као најшире коришћена апликација са естимацијом покрета, ослања се на отклањање статистичкије редунадансе у видео садржају. Пошто су осветљај и боја у динамичким сликама највише корелисани у дирекцији кретања отклањање редунадансе је највише зависно од естимације покрета [1]. Са друге стране, компензовано повећање броја слика у секунди у видео секвенцама (*frame rate up-conversion*) користи пикселе постојећих фрејмова и срачунато поље вектора помераја за креирање нових (интерполираних) фрејмова између њих (на пример, конвертовање видеа од 30 фрејмова у секунди у видео од 60 фрејмова у секунди). Тачност и прецизност естимације покрета је критична због избегавања артефаката покрета, халоа око објеката, фликера, спречавања губитка делова објеката и настајања блура услед кретања на интерполираним фрејмовима. Сви поменути дефекти, а посебно изостанак глаткоће покрета, проузрокују негативан утисак код гледаоца видео секвенце. Квалитет слике, висока резолуција и број фрејмова у секунди у свим апликацијама и уређајима са естимацијом покрета се подразумева, што резултује огромним количинама података које је потребно процесирати. Због тога је изазов креирати алгоритам за естимацију покрета који помирује та два међусобно супротстављена захтева: срачунава поље вектора помераја које прецизно одговара реалним покретима и има ниску комплексност, идеално погодну за извршавање у реалном времену.

## 1.2 Врсте естимација

Естимација покрета у видео секвенцама је недетерминисан проблем због тога што се просторно кретање у природи трансформише у раван снимком камере и самим тим га је немогуће у потпуности описати. Међутим, и поред тога, могуће је дефинисати типове кретања. У зависности од броја параметара који описују кретање у видео секвенци могуће је дефинисати неколико типова естимација покрета:

- Транслација описује покрете само по хоризонталној и вертикалној равни. Вектор помераја садржи два параметра –  $(v_x, v_y)$ . Смер кретања одређен је знаком вектора помераја.
- Ротација омогућава, поред трансляторних помераја, дефинисање и угла за који се неки објекат (део фрејма) померио.
- Тилт омогућава опис помераја задње стране камере (горе-доле или лево-десно).
- Зум обезбеђује вектор помераја који садржи параметар за опис промене величине објекта од фрејма до фрејма.

У зависности од подржаних типова покрета формира се модел кретања. Он садржи матрицу са параметрима који садрже све потребне информације за везане опис покрета. У наставку је приказан пример матрице кретања:

$$\begin{bmatrix} \varepsilon \cdot \cos\theta & -\sin\theta & t_x \\ \varepsilon \cdot \sin\theta & \cos\theta & t_y \\ 0 & 0 & z \end{bmatrix}$$

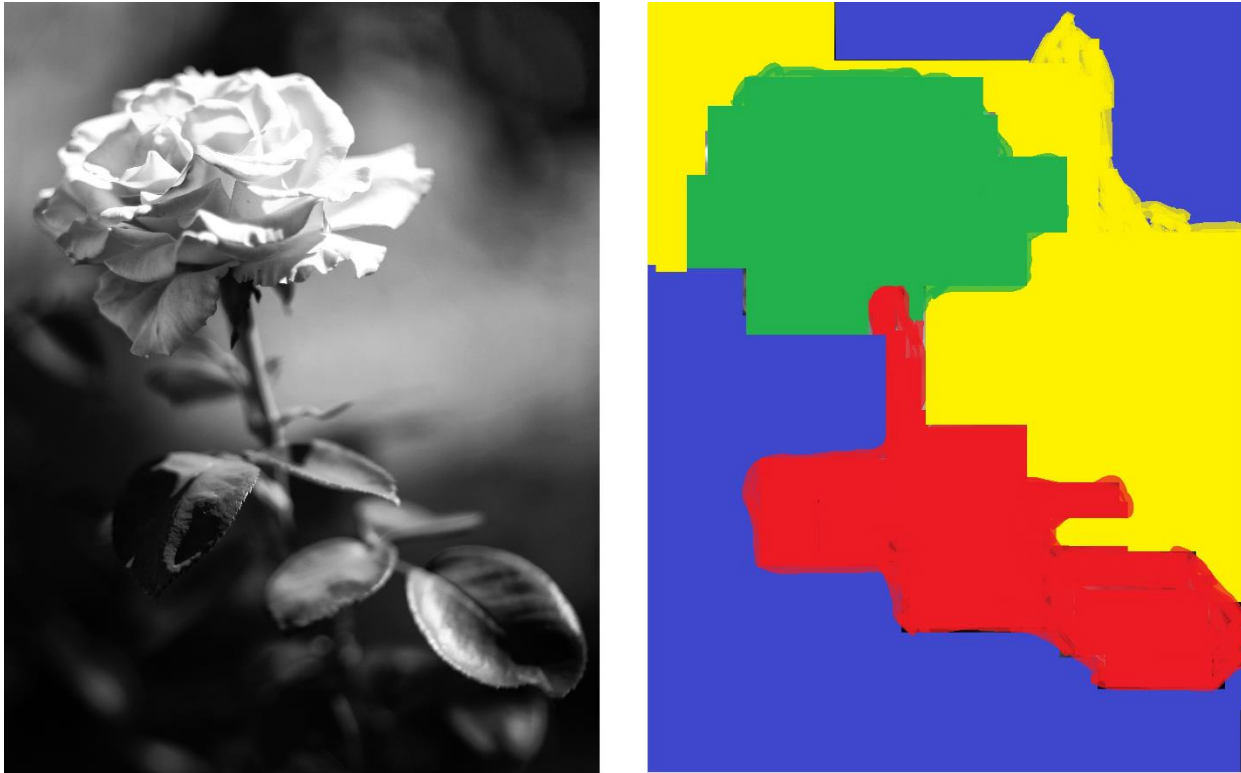
Параметар  $\varepsilon$  представља оријентацију кретања,  $\theta$  угао ротације,  $z$  зум фактор, а  $(t_x, t_y)$  вектор помераја translације.

Код естимације покрета коришћене у *frame rate up-conversion* алгоритмима, ради смањења комплексности, од интереса је најчешће translација јер је за креирање интерполираних фрејмова њоме могуће постићи висок квалитет интерполације са становишта корисника.

Друга подела естимације покрета везана је за тип, односно величину, региона фрејма за који је везан вектор помераја:

- Глобално кретање (сви пиксели фрејма крећу се на исти начин, односно цео фрејм има само један вектор помераја).
- Објектно кретање дели фрејм на произвољан број објеката произвољне величине и облика и сваки од објеката садржи вектор покрета који описује његово кретање. Пример поделе фрејма на објекте приказује Слика 1-1. Овај модел може обезбедити веома прецизну процену покрета и садржи минималан број објеката и вектора помераја, али му је мана прецизна сегментација фрејма односно одређивање тачних граница између објеката.
- Блоковско кретање дели фрејм на блокове исте величине (нпр. 16x16 пиксела) и сваки блок има вектор помераја који описује његово кретање. Овај тип је најчешће коришћен због једноставности и униформности при имплементацији.

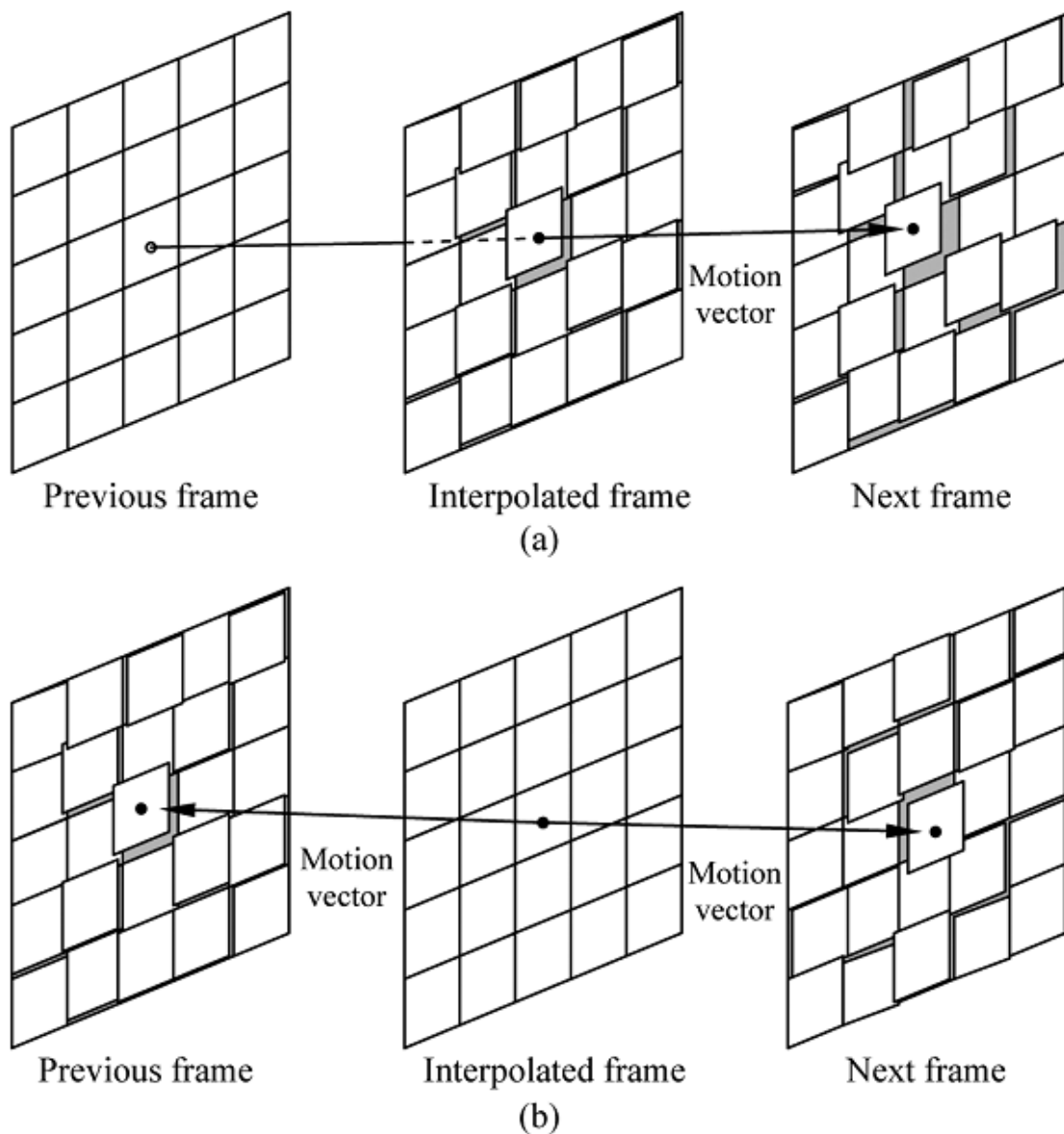
- Пиксел кретање подразумева да сваки пиксел фрејма садржи свој вектор покрета. Користи се понекад у комбинацији са неким од наведених типова и то када се изврши ерозија (рашчлањивање) објекта или блока до нивоа пиксела.



Слика 1-1: Пример поделе фрејма на објекте, при чему је могуће сваком објекту доделити вектор помераја. Могуће је такође и да објекти деле исти вектор помераја, што је веома често и случај.

Трећа подела естимације покрета, односно вектор поља помераја, је извршена на основу референце (почетне тачке) у односу на коју се процењују покрети. Уколико се посматрају два фрејма видео секвенце – претходни (*Previous frame*), временска позиција 0, и следећи (*Next frame*), временска позиција 1, ([2] , Слика 1-2), може се дефинисати:

- Унидирекционо. Вектори покрета почињу на претходном фрејму и завршавају се на следећем.
- Бидирекционо. Вектори покрета почињу на позицији између претходног и следећег фрејма на којој ће бити креиран интерполирани фрејм. Мана овог приступа је већа вероватноћа грешке, тј. добијања лажне предикције кретања.

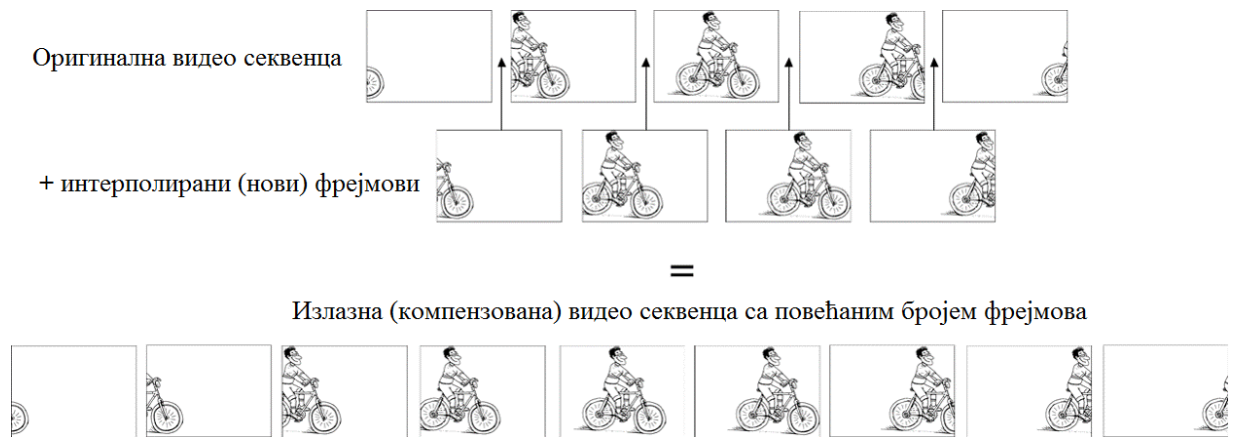


Слика 1-2: Унидирекциона (слика горе) и бидирекциона (слика доле) естимација покрета. На горњој слици референца (полазна тачка) је на претходном фрејму (*Previous frame*) одакле се тражи парњак сваког блока на следећем фрејму (*Next frame*). Уочава се да блок који се преноси вектором помераја из претходног на следећи фрејм не мора пасти тачно на гريد интерполираног фрејма (*Interpolated frame*), због чега је могуће да неки пиксели интерполираног фрејма остану без вектора помераја. На слици доле референца естимације је на интерполираном фрејму одакле се тражи најсличнији пар блокова са претходног и следећег фрејма. Слика преузета из [3] .

### 1.3 Шта је *frame rate up-conversion*?

Једна од главних особина видео материјала при оцени његовог квалитета, односно утиска гледаоца при гледању, је број фрејмова у секунди (*frames per second (FPS)*, *frame-rate*). Типично, видео материјал је најчешће доступан на 15, 24 или 30 фрејмова у секунди, док дисплеји данашњих уређаја, који врше њихово приказивање (телевизори, мобилни телефони, таблети), најчешће подржавају 60 фрејмова у секунди. Најједноставнија метода за повећање броја фрејмова у секунди је понављање фрејмова (*frame repeat*). Она обично изазива визуелно непријатно сецкање - цадер (*judder*), јер је стварна учестаност фрејмова остала непромењена. Већи број слика у секунди у видео секвенци побољшава утисак корисника (гледаоца) и минимизује непријатни изостанак глаткоће покрета која се најчешће јавља код објеката са мноштвом детаља који се крећу већом брзином или при померају камере.

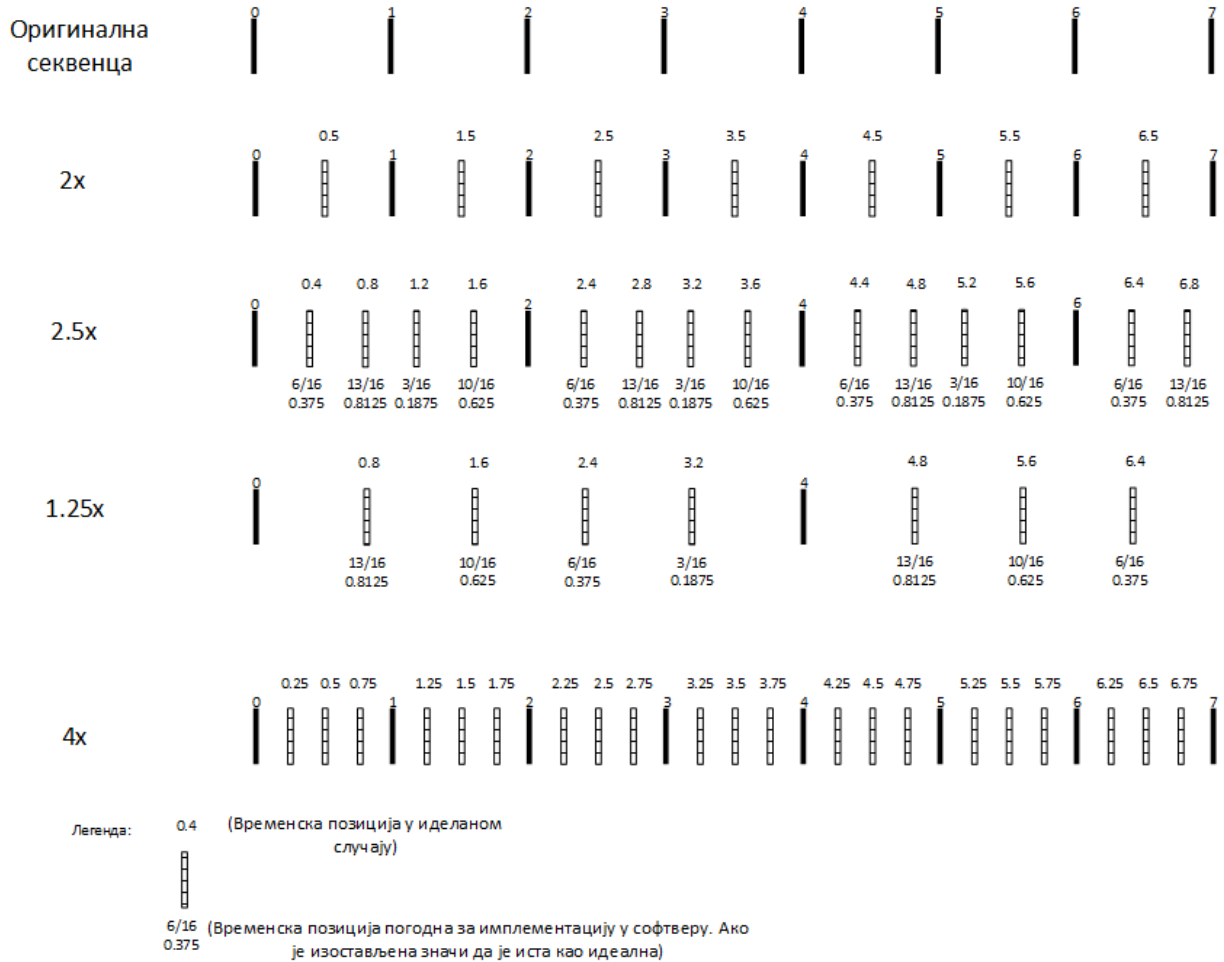
Компензовано повећање броја слика у секунди (*frame rate up-conversion*) управо подразумева естимацију покрета (креирање поља вектора) и компензацију (креирање интерполираног фрејма уз помоћ поља вектора помераја и пиксела постојећих фрејмова) – Слика 1-3.



Слика 1-3: Илустрација повећања броја фрејмова у секунди креирањем интерполираних фрејмова са компензованим кретањем.

Ако се посматра блоковска подела пиксела фрејма, да би се креирао интерполирани фрејм, потребно је доделити сваком блоку вектор помераја. Због тога је бидирекциона естимација погодна јер ће дати управо поменуто. На пример, за конверзију видео секвенце из 30 у 60 фрејмова у секунди временска позиција на којој се врши интерполација је тачно на средини између претходног и следећег фрејма (фактор конверзије 2x). Овај тип вектор поља је погодан код *frame rate up-conversion* алгоритама за креирање интерполираног фрејма. Примере срачунавања временских позиција за постизање неколико различитих фактора конверзије приказује Слика 1-4. Фактори конверзије 2 и 4 су једноставни за

компензацију јер подразумевају уметање једног, односно три интерполирана фрејма између оригиналних. Код фактора конверзије 2.5 (конверзија из 24 у 60 фрејмова у секунди) уочавамо да сваки други фрејм оригиналне секвенце не постоји у излазној секвенци. Због тога је посебно битно да квалитет интерполираних фрејмова буде веома висок.



Слика 1-4: Пример потребног позиционирања бидирекционог поља вектора за неколико различитих фактора конверзије видео секвенце.

Уколико постоји унидирекционо поље вектора покрета, креирање интерполираног фрејма могуће је извршити на два начина:

- На основу унидирекционог поља креирати бидирекционо поље. Тај поступак се зове ритајминг (*retiming*) [4]. При томе може се десити да поједини блокови или пиксели немају додељен вектор покрета или да неки пак имају неколико додељених вектора покрета. Прво је могуће решити тако што се пикселима (блоковима) без вектора покрета додели вектор са неког од суседних пиксела са интерполираног фрејма. Друго, пиксели (блокови) са



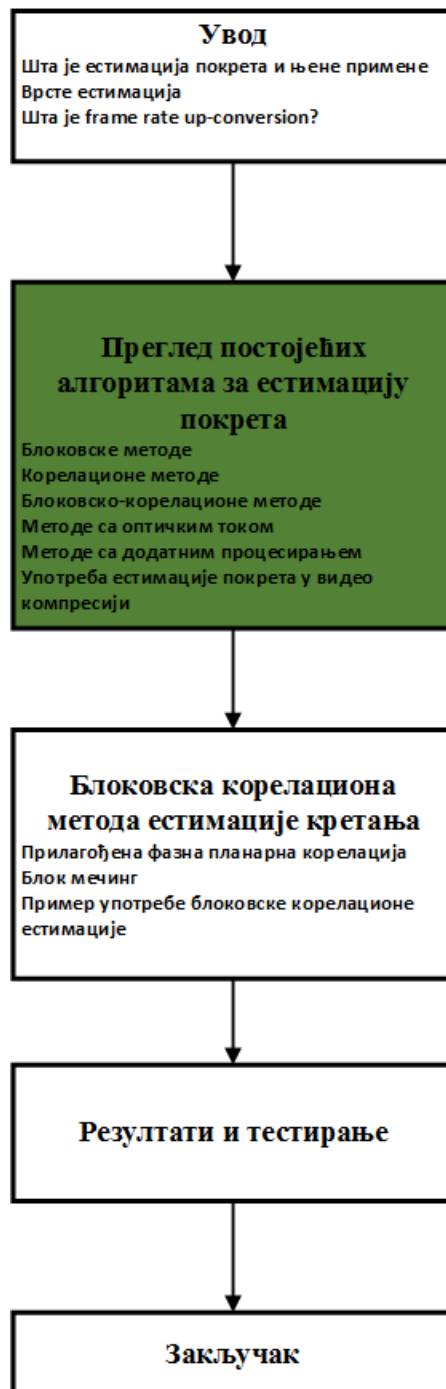
више вектора покрета могу изабрати један опет на основу анализе вектора са суседних пиксела (блокова).

- Ретко коришћен метод који подразумева директно пресликавање (преношење) пиксела претходног фрејма на интерполирани фрејм користећи векторе помераја. На тај начин, могућа су такође два сценарија, да неке позиције на интерполираном фрејму немају информацију о пикселу (празнине) и да неке позиције имају додељено неколико пиксела са претходног фрејма. И решење је слично као у претходном случају са векторима, празнине се могу попунити суседним пикселима, а позиције са више пиксела блендовањем и анализом суседних пиксела.

Лоша страна оба приступа је очигледна и огледа се у томе да је тешко извршити прецизан ритајминг који неће креирати нежељене артефакте на интерполираном фрејму.

Алгоритми за компензовано повећање броја фрејмова у секунди су рачунарски веома комплексни, најчешће реда величине хиљаду гига-операција у секунди. До сада, били су углавном лимитирани на телевизоре, где буџет дисипиране снаге није од пресудног значаја. Изазов је креирати алгоритам који ће омогућити визуелно добар утисак гледаоца, без непријатног цадера, и остати у буџету снаге мобилног уређаја.

## 2. Преглед постојећих алгоритама за естимацију покрета



До сада је било речи о типовима естимација и поделама, док ће у овом поглављу бити описани начини, односно алгоритми, којима је могуће срачунати саме векторе помераја. У зависности од начина срачунавања могуће их је груписати у више категорија (Слика 2-1):

- Блокковске методе;
- Корелационе методе;
- Блокковско-корелационе методе;
- Методе са оптичким током;
- Методе са додатним процесирањем.



Слика 2-1: Преглед категорија алгоритама за естимацију кретања у видео секвенцама. Методе са додатним процесирањем користе блокковске или методе са оптичким током за иницијалну естимацију кретања. Блок мечинг корелација (БМК), предложена у овој тези, истакнута је у групи блокковско-корелационих метода.

У наставку ће бити дат опис карактеристика сваке од категорија као и преглед алгоритама који им припадају. Поред тога, као што је већ поменуто, естимација покрета игра значајну улогу у алгоритмима за видео компресију. У последњем делу овог поглавља дат је опис генералне структуре алгорита за компресију у коме је истакнут значај естимације покрета.

## 2.1 Блокoвске методе

У оквиру алгоритама за естимацију покрета најзаступљеније су методе базиране на блокoвском приступу. У зависности од типа естимације по референци, блокoвски приступ код унидирекционе естимације подразумева поделу претходног фрејма на блокoве и, за сваки блок, проналажење његовог најсличнијег парњака на следећем фрејму. Код бидирекционе естимације врши се подела интерполираног фрејма (који ће бити креиран) на блокoве и проналажење пара најсличнијих блокoва на претходном и следећем фрејму које повезује исти вектор помераја, али са промењеним смером. То подразумева (Слика 1-2) да за селектовани блок на интерполираном фрејму вектором  $(v_x, v_y)$  долази се до селектованог блока на следећем фрејму, док се вектором  $(-v_x, -v_y)$  долази до селектованог блока на претходном фрејму. Вектор  $(v_x, v_y)$  који на описани начин повезује два најсличнија блока ће бити додељен блоку интерполираног фрејма. У наставку текста подразумеваће се да свака од описаних блокoвских метода естимације покрета може бити имплементирана и као унидирекциона и бидирекциона.

Што се тиче критеријума сличности између блокoва, најчешће коришћена и најподобнија за имплементацију је сума апсолутних разлика (*Sum of absolute differences*) вредности пиксела са претходног ( $P$ ) и следећег блока ( $C$ ):

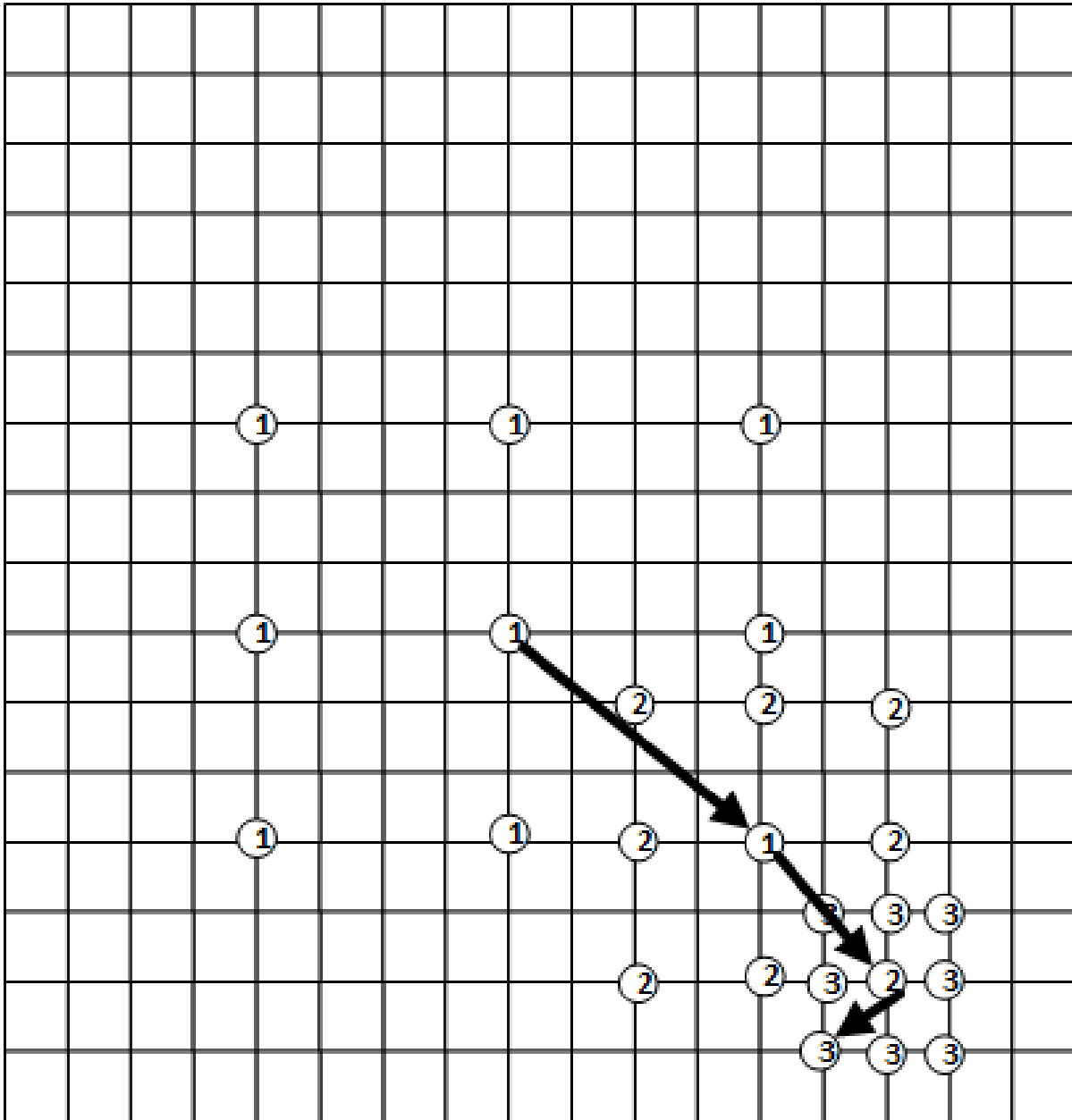
$$\sum_{i,j=1}^n |P(i,j) - C(i,j)| \quad 2.1$$

Мања вредност одговара већој сличности између блокoва. За идентичне блокoве она је очигледно једнака нули. Блокoвски приступ се широко користи због своје једноставности, отпорности на шум и регуларних патерна приступа подацима који су погодни за имплементацију у хардверу или софтверу.

Као најочигледније и алгоритамски најједноставније решење за естимацију кретања је потпуно претраживање (*Full search*) - *FS*. То подразумева, за сваки блок полазног фрејма, срачунавање суме апсолутних разлика између њега и свих блокoва на следећем фрејму. Од свих срачунатих сума апсолутних разлика бира се она најмања и њен вектор помераја се додељује тренутном блоку. Висока комплексност и вероватноћа процене лажног кретања најчешће дисквалификује употребу овог алгорита [5] .

Због генерално велике рачунарске комплексности алгоритама за естимацију покрета, у почетку су се фаворизовале једноставније методе. Пример поменутог је Претраживање у три корака [6] (*Three Step Search*). Слика 2-2 приказује ток претраживања. Тренутни блок или пиксел се пореди прво са 8 блокoва (пиксела) на растојању 4 и једним на растојању нула (први корак). Затим се позиционирање премешта на онај са најмањом сумом апсолутних разлика и понавља исти поступак само са растојањем једнаким два (други корак), па са растојањем један (трећи корак). На тај начин се уз 25 срачунавања суме апсолутних разлика

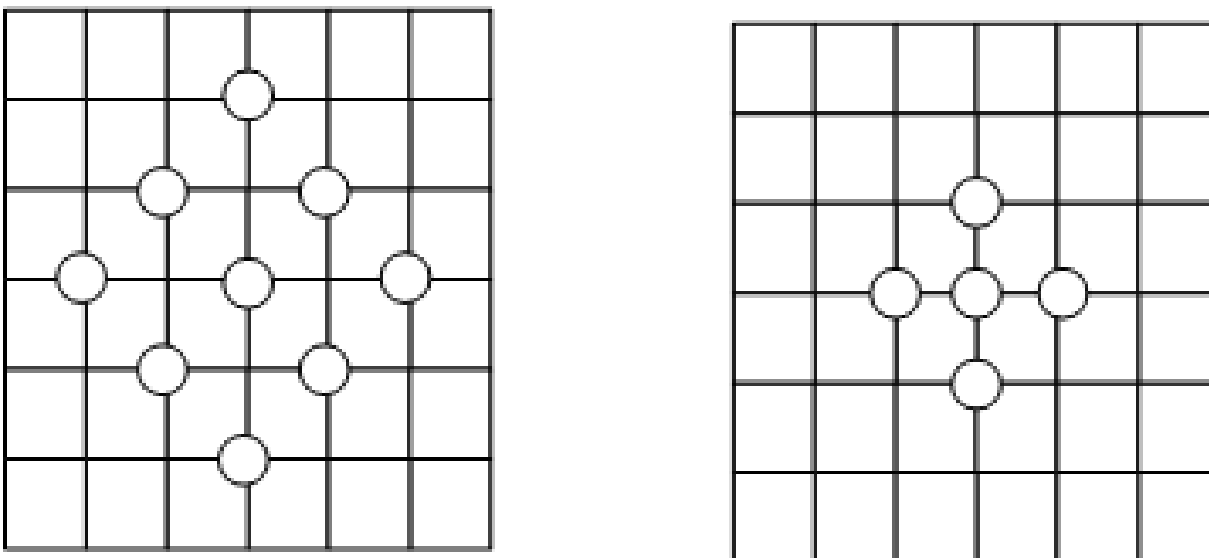
(9 + 8 + 8 = 25) оптимално добија најподобнији вектор помераја у околини од четири пиксела.



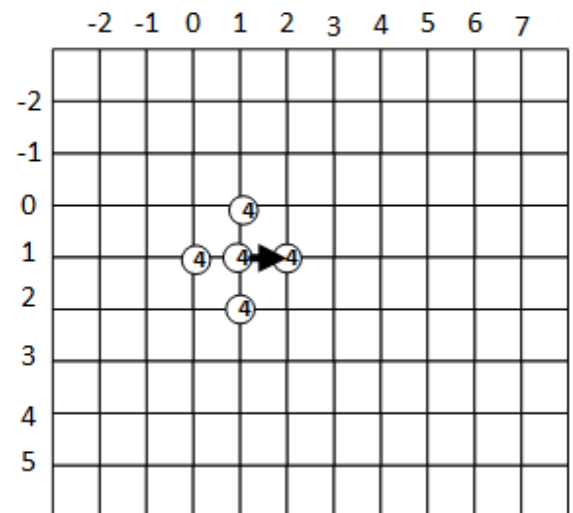
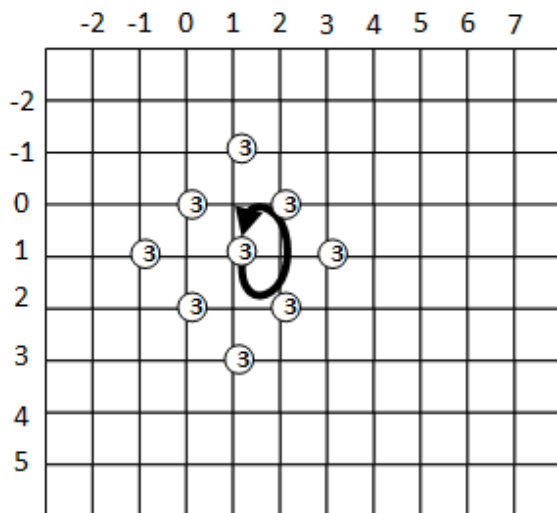
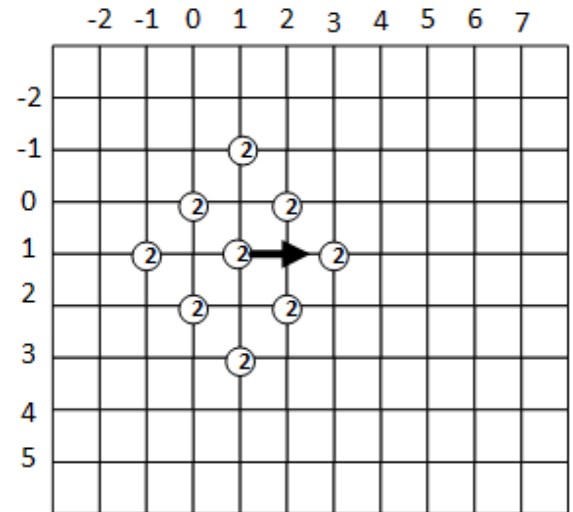
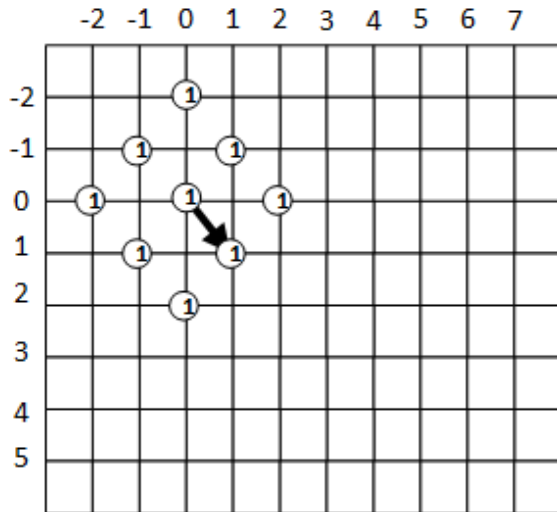
Слика 2-2: Претраживање у три корака – поступак избора блокова за срачунавање суме апсолутних разлика. Срачунати вектор помераја на овом примеру је  $(v_x, v_y) = (4, -4) + (2, -2) + (-1, -1) = (5, -7)$ .

Логаритамско претраживање (*2D-Logarithmic Search*) [7] ради по сличном принципу само са другим начином одабира суседних блокова (пиксела) са којима ће се поредити тренутни блок.

*Diamond Search* [8] користи за претраживање најсличнијег блока патерн у облику дијаманта. Логика тражења је изведена из функције сличности блока (траженог региона) која опада како му се приближавамо. Алгоритам стартује на блоку исте локације као што је блок на референтном фрејму за који се тражи најсличнији парњак и врши осам рачунања суме апсолутних вредности у околини центра дијаманта (Слика 2-3, лево). Овде се види разлика у односу на *Full Search*, који врши проверу за све блокове. Када се нађе локација са минималном сумом центар дијаманта се позиционира на њу и описани поступак се понавља све док блок са најмањом сумом апсолутних вредности буде у центру (Слика 2-4). Када се то деси примени се још једном мали дијамант патерн (Слика 2-3, десно) после чега се добија финални вектор помераја за тај блок.



Слика 2-3: Патерн коришћен у *Diamond search* алгоритму. Лево - велики дијамант, десно - мали дијамант патерн.



Слика 2-4: Поступак претраживања блокова предложен у *diamond search* алгоритму. Алгоритам стартује од блока (0, 0), пронаћи најбољи парњак доле десно (слика горе лево). Затим, у кораку два (слика горе-десно) најбољи парњак је десно. У трећем кораку (слика доле-лево) најбољи парњак је у центру дијаманта. У четвртном кораку примењује се мали дијамант патерн за финално проналажење најсличнијег блока (слика доле десно).

Рекурзија у алгоритме за естимацију покрета први пут је уведена уз 3-димензионално рекурзивно претраживање (*Three dimensional recursive search – 3DRS*) [9]. Позната је чињеница да суседним (блиским) блоковима на истом фрејму са великом вероватноћом одговара исти вектор помераја, односно, велика је вероватноћа да су део истог објекта који се креће или мирује (просторна зависност). Такође, блиским блоковима на суседним фрејмовима са великом вероватноћом одговара исти вектор помераја, односно, кретање у видео секвенцама је најчешће униформно (темпорална зависност). По подели фрејма на блокове дефинише се редослед блокова за које се срачунава вектор помераја (редослед скенирања). Управо због ове две чињенице уведена је рекурзија која подразумева за

тренутни блок проверу подобности вектора помераја са претходно срачунатих блокова истог фрејма, као и са суседних блокова претходног фрејма видео секвенце. Трећи додатак овог алгоритма је специјални случајни (рандом) вектор помераја који се добија додавањем малог псеудо-случајног шума просторном кандидату [10]. На тај начин постиже се постепена конвергенција вектора помераја ка крајњој вредности која одговара реалном померају у видео секвенци.

*Predictive Motion Vector Field Adaptive Search Technique* [11] и *Adaptive Rood Pattern Search* [12] такође користе рекурзивни приступ. Постоје и оптимизације алгоритма са потпуним претраживањем, као што је хијерархијско стохастичко брзо претраживање базирано на Калмановом филтрирању [13]. Хијерархијски приступ подразумева једну врсту пирамидалне декомпозиције фрејма скалирањем (*downscaling*) са два пута већим фактором од најнижег до највишег нивоа (нпр. 2, 4, 8, 16, 32 пута). Потом се естимација врши прво на највишем нивоу, а њени резултати се прослеђују до нижих нивоа. Друга могућност оптимизације алгоритма са потпуним претраживањем користи биогеографски приступ описан у [14].

## 2.2 Корелационе методе

Поред блоковских од значаја су и корелационе методе као што је Фазна Планарна Корелација (*Phase Plane Correlation*) - ФПК. Она се састоји од срачунавања корелације између два фрејма уместо тражења најсличнијих блокова. На основу корелације добија се вектор помераја присутан у видео секвенци. Мана овог приступа је постојање рачунарски захтевних, фреквенцијских трансформација, као што је Фуријеова трансформација. Још један недостатак је тешкоћа успостављања везе између срачаног вектора помераја и тачног региона фрејма (објекта) на који се он односи. Због поменутог неопходно је извршити значајна прилагођења и оптимизације ове методе да би се омогућило њено коришћење у алгоритмима за *frame-rate up-conversion*.

Традиционална ФПК метода је представљена у наставку. За два улазна суседна фрејма  $g_p$  и  $g_c$  видео секвенце, фазна планарна корелација може бити реализована пратећи следећу процедуру:

1. Рачунање дводимензионалне дискретне Фуријеове трансформације обе слике ( $g_p$  и  $g_c$ ) између којих се врши корелација:

$$G_p = DFT(g_p), G_c = DFT(g_c) \quad (2.2.1)$$

2. Рачунање спектра снаге (*Cross Power spectrum*) множењем, елемент по елемент, комплексног спектра прве слике и коњуговано комплексног спектра друге слике, као и нормализацију добијеног производа:



$$R = (Gp \cdot Gc^*) / |Gp \cdot Gc^*| \quad (2.2.2)$$

3. Рачунање инверзне, 2-димензионалне дискретне Фуријеове трансформације над спектром снаге:

$$r = DFT^{-1}(R) \quad (2.2.3)$$

4. Замена места (*shift*) између одговарајућих елемената првог и трећег и другог и четвртог квадранта претходно срачунате корелационе равни  $r$ . Као резултат добија се корелациона раван  $c$ :

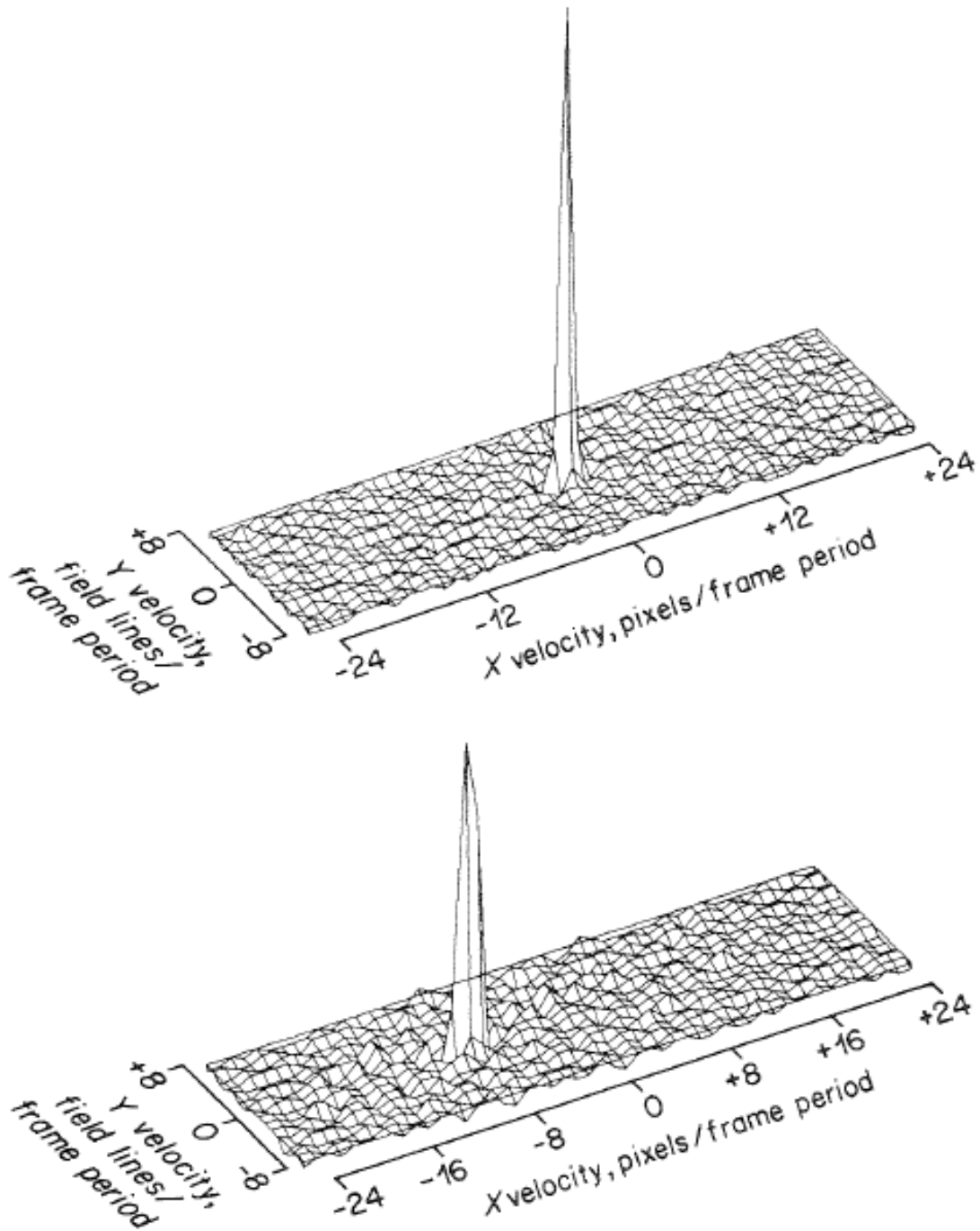
$$c = shift(r) \quad (2.2.4)$$

5. Проналажење максималних вредности, пикова корелационе равни (*peaks*). Индекси пикова у корелационој равни представљају померај између улазних слика, што је у основи вектор помераја  $(\Delta x, \Delta y)$  објекта унутар региона над којим је вршена корелација:

$$(\Delta x, \Delta y) = argmax(x, y) = \{c\} \quad (2.2.5)$$

Први пут употребљена корелациона метода у свом изворном облику и реализована је од стране BBC *Research department* [15]. У њиховом раду описана је и хардверска реализација са датим детаљима имплементације *Fast Fourier Transformation (FFT)* алгоритма коришћеног за срачунавање дискретне Фуријеове трансформације. Једина модификација у односу на традиционалну (теоријску) методу је нормализација спектра снаге замењена процесом фазне екстракције који подразумева проналажење фазе сваке од просторних учестаности и срачунавање њихове разлике. Димензије региона над којима се врши фазна корелација су 48 пиксела по хоризонтали и 16 пиксела по вертикали (Слика 2-5). Поменути

однос је установљен због тога што је заступљеност кретања у видео секвенцама већег интензитета по хоризонталним правцима него по вертикалним правцима фрејма.



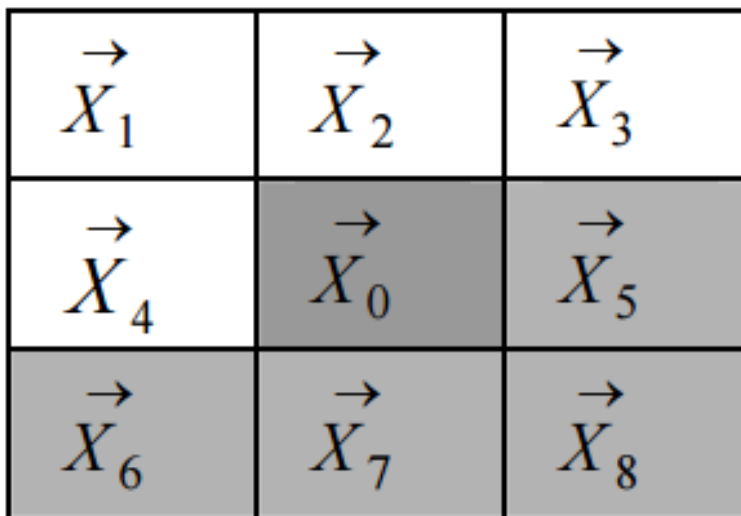
Слика 2-5: Корелациона равна која одговара пару региона између којих нема кретања (слика горе) и корелациона равна која одговара пару региона између којих је померај -8 по хоризонталном смеру (слика доле). Преузето из [15].

Још једно од решења које описује апроксимирање нормализације спектра снаге, као и Фуријеове трансформације, у процесу срачунавање фазне планарне корелације описано је у [16]. Овде предложена апроксимативна нормализација полази од идеје да се бројеви са покретном тачком сведу на најпогоднији експонент после чега је над њима могуће обављати операције као над бројевима са непокретном тачком.

### 2.3 Блокoвско-корелационе методе

Алгоритми ове групе садрже две јединице, блокoвску и корелациону, од којих корелациона врши независну естимацију кретања и своје резултате прослеђује блокoвској јединици која врши њихову додатну евалуацију срачунавањем поменуте суме апсолутних разлика.

У алгоритму предложеном у [17] врши се подела фрејма на регионе 24x24 пиксела над којима се срачунава фазна планарна корелација која као резултат даје  $N$  вектора помераја. Регион се дели на блокoве 2x2 или 4x4. Сваки од блокoва добија кандидате за вектор помераја од региона којем припада, као и од 8 суседних региона. Овде је такође искоришћена униформност кретања објеката у видео секвенцама тиме што се користе резултати срачунати над регионима из претходне итерације фрејмова (Слика 2-6). На описани начин сваки од блокoва прима укупно  $9N$  кандидата, али се не врши и толики број евалуација јер међу њима постоји већи број истих кандидата.



Слика 2-6: Приказ суседних региона из којих се узимају кандидати за вектор помераја. Смер срачунавања фазних корелација по регионима је од горе ка доле и од лево ка десно. Из поменутог се закључује да су кандидати 1-4 из тренутног пара фрејмова, а 5-9 из претходног пара фрејмова (коришћена је временска зависност). Преузето из [17].

Блок мечинг корелација (БМК), описана у овој тези, користи резултате фазне планарне корелације у блок мечинг модулу на јединствен начин, описан у следећем поглављу.

## 2.4 Методе са оптичким током

Могућност описивања свих поменутих типова кретања (транслација, ротација, тилт и зум) обезбеђују алгоритми засновани на оптичком току (*optical flow*). Заједничка полазна тачка свих алгоритама ове групе је претпоставка да се интензитети пиксела транслирају из претходног фрејма у следећи по просторно-временској функцији [18]. Ова претпоставка, на први поглед, изгледа неконзистентно, али се у пракси показала као веома основана. Да би се извео начин функционисања оптичког тока за дводимензионални сигнал дато је извођење за једнодимензионални сигнал.

Нека су дата два сигнала  $f_1(x)$  и  $f_2(x)$  у различитим временским тренуцима, и нека је:

$$f_2(x) = f_1(x - d) \quad (2.4.1)$$

где  $d$  представља транслаторни померај. Тејлоров развој функције  $f_1(x - d)$  је:

$$f_1(x - d) = f_1(x) - df_1'(x) + O(d^2 f_1''). \quad (2.4.2)$$

Са овим развојем могуће је представити разлику између два сигнала на локацији  $x$  као:

$$f_1(x) - f_2(x) = df_1'(x) + O(d^2 f_1''). \quad (2.4.3)$$

Занемаривањем последњег члана, као и чланова виших степена добијамо апроксимацију за  $d$ :

$$\hat{d} = \frac{f_1(x) - f_2(x)}{f_1'(x)} \quad (2.4.4)$$

Једнодимензионални случај интуитивно се генерализује на дводимензионални, уз исту претпоставку као малопре, да се дислоцирана слика може погодно представити Тејлоровим редом:

$$I(\vec{x} + \vec{u}, t + 1) \approx I(\vec{x}, t) + \vec{u} \cdot \nabla I(\vec{x}, t) + I_t(\vec{x}, t), \quad (2.4.5)$$

где су  $\nabla I \equiv (I_x, I_y)$  и  $I_t$  просторни и временски парцијални изводи слике  $I$ , а  $\vec{u}$  је дводимензионални померај (брзина). Занемеривањем виших степена Тејлоровог реда и одузимањем линеарне апроксимације и почетне претпоставке:

$$I(\vec{x}, t) = I(\vec{x} + \vec{u}, t + 1), \quad (2.4.6)$$

добија се:

$$\nabla I(\vec{x}, t) \cdot \vec{u} + I_t(\vec{x}, t) = 0. \quad (2.4.7)$$

Добијена једначина се односи на брзину просторно-временске примене на једној локацији на слици и зове се једначина константне разлике. Уколико су доступна само два фрејма или је немогуће естимирати  $I_t$  могуће је извести услов по коме се  $I_t(\vec{x}, t)$  у једначини изнад мења са:

$$\delta I(\vec{x}, t) \equiv I(\vec{x}, t + 1) - I(\vec{x}, t). \quad (2.4.8)$$

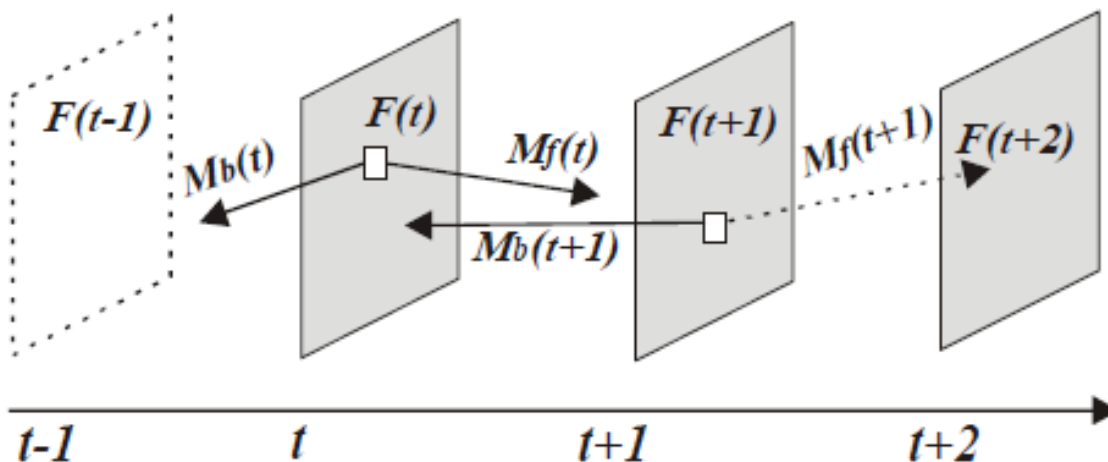
Један од често коришћених алгоритама овог типа је Лукас-Канаде метода [19]. Она проширује поменућу основу и на суседне пикселе и на тај начин даје могућност формирања скупа диференцијалних једначина чијим решавањем, методом најмањих квадрата, се добија модел кретања, а самим тим и вектор помераја.

## 2.5 Методе са додатним процесирањем

Новије методе најчешће уводе додатно процесирање на основне алгоритме за естимацију. Оне користе неке од постојећих метода (блоковске или методе са оптичким током) за срачунавање иницијалног поља вектора помераја над којим врше даље, додатно процесирање. У додатно процесирање најчешће спадају класификација, филтрирање и селекција заснована на естимационим грешкама претходно срачунатог поља вектора помераја [20]. На тај начин се додатно унапређује прецизност и тачност естимације. *Multi-frame hierarchical motion estimation (MFHME)* срачунава вероватноће и користи их за евалуацију и унапређење естимације [21]. Имплементација у реалном времену и додатни резултати овог решења доступни су у [22]. Да би се постигле што прецизније трајекторије кретања алгоритам користи два фрејма као референце и срачунава грешке у блок мечингу и за луминентну и за хроминентну компоненту. Такође, уместо најчешће коришћене суме апсолутних вредности за срачунавање разлике између блокова, *Multi-frame hierarchical motion estimation* користи суму квадратних разлика пиксела. Имплементација самог

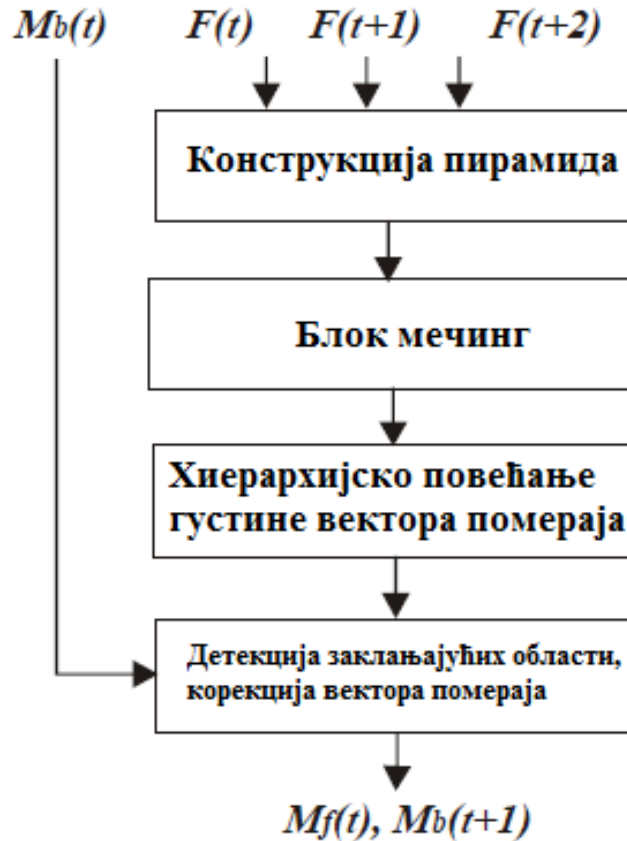
алгоритма, која користи хијерархијску структуру, извршава се у реалном времену на програмабилној платформи (*FPGA – Field Programmable Gate Array*).

*MFHME* користи три суседна фрејма видео секвенце за естимацију кретања  $F(t)$ ,  $F(t+1)$ ,  $F(t+2)$  и претходно естимирано поље вектора помераја  $M_b(t)$  за естимацију унидирекционих вектор поља  $M_f(t)$  и  $M_b(t + 1)$  (Слика 2-7).



Слика 2-7: *MFHME* алгоритам базиран на три суседна фрејма.

Дијаграм *MFHME* алгоритма приказује Слика 2-8. Пирамиде се конструишу за сваки од улазних фрејмова и врши се хијерархијска естимација кретања. Потом се над фрејмовима добијеним на највишим нивоима пирамида врши *Full search* блок мечинг алгоритам за естимацију кретања, чиме се добију три груба поља вектора -  $M_f(t)$ ,  $M_b(t + 1)$  и  $M_f(t + 1)$ .



Слика 2-8: Етапе извршавања MFHME алгоритма.

Користећи хијерархијски алгоритам повећава се густина вектора помераја по нивоима пирамиде, тако што се *Full search* блок мечинг извршава на сваком нивоу. Претходно, фрејм је подељен на блокове исте величине, ради лакшег процесирања. Последњи корак подразумева детекцију објеката који се преклапају и блокови којима они припадају се процесирају на посебан начин.

Још једно од новијих решења са додатним процесирањем је варирајућа фузија слика (*variational image fusion*) [23]. Она подразумева креирање више интерполираних фрејмова, чијим се блендовањем формира финални интерполирани фрејм. За сваки фрејм је формирано поље вектора користећи *optical flow* метод [24].

## 2.6 Употреба естимације покрета у видео компресији

Још од појаве дигиталних слика, и касније дигиталних видео секвенци, велики меморијски простор потребан за њихово смештање представља један од ограничавајућих фактора њиховог коришћења. Због тога се од најранијих дана прибегло техникама његове

минимизације увођењем концепта видео компресије. Видео компресија представља поступак редукције простора потребног за смештање видео секвенце у меморију уз потпуно или делимично очување њеног квалитета. Видео компресија треба да задовољи два међусобно супротстављена захтева:

1. Минимизована комплексност самих алгоритама за компресију и декомпресију како би репродукција видео секвенци била могућа у складу са њиховим бројем фрејмова у секунди.
2. Обезбеђивање што већег степена компресије.

Илустрације ради, за смештање 1080п видео секвенце са 30 фрејмова у секунди и три бајта по пикселу (која је данас стандард) трајања један сат, без компресије потребно би било:

$$1920 \times 1080 \times 3 \times 30 \times 3600 = 671.846.400.000 \text{ бајтова} \approx 625.7 \text{ гигабајта}$$

Чак и за данашње стандарде заузеће толиког простора није прихватљиво. Због тога се видео секвенце смештају у компресованом облику. Сам процес (алгоритам) који врши компресију сирових података видео секвенце зове се кодер. Процес распакивања ради припреме за приказивање кориснику (декомпресију) зове се још и декодер. Сам поступак компресије заснива се на:

- просторној зависности објеката унутар једног фрејма (интра-фрејм предикција, слична оној која се примењује код дигиталних фотографија)
- временској зависности (редунданси) између суседних фрејмова видео секвенце (интер-фрејм предикција), која значајније доприноси постизању већих фактора компресије

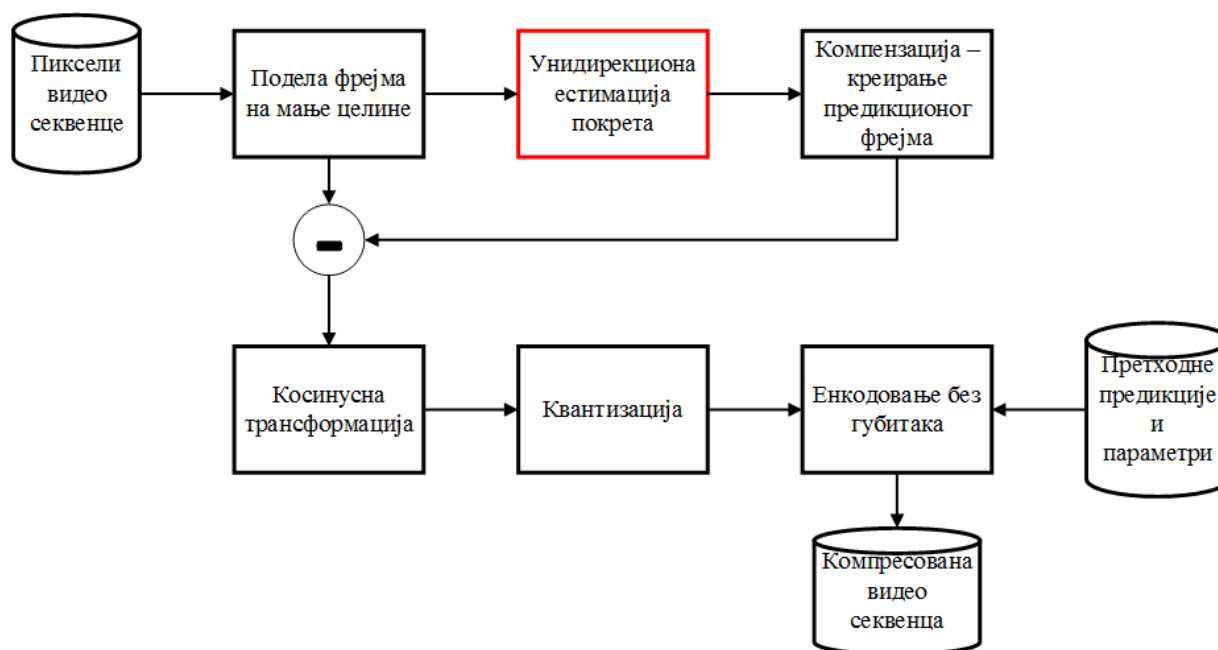
Ранији стандарди предложени од стране *Motion Picture Experts Group (MPEG)*, MPEG-1 [25], MPEG-2 [26], MPEG-4 *Advanced simple profile* [27] и MPEG-4 *Advanced video coding (AVC)* стандард [28] превасходно су користили временску редундансу за реализацију компресије. Иако сами стандарди предлажу и описују реализацију декодера у наставку ће бити дат општи опис поступка кодирања, заједнички за поменуте стандарде, како би се истакао значај саме естимације покрета.

Пиксели фрејма видео секвенце су груписани, као што је речено, у три канала, у формату луминансе ( $Y$ ) и два хроминентна канала ( $Cb$  и  $Cr$ ). При смештању саме видео секвенце искоришћена је особина људског визуелног система да је осетљивији на промену осветљаја него на промену боје тиме што је извршено семпловање хроминентних сигнала, односно редуковање броја пиксела коришћених за њено смештање на следећи начин: 4 суседна  $Y$  пиксела деле један пар  $Cb$  и  $Cr$  пиксела. То је такозвани 4:2:0 формат.



Процес саме компресије приказује Слика 2-9. На почетку компресије врши се подела фрејма на мање целине: макроблокове, слајсеве, блокове, субблокове, над којима се врши даље, униформно процесирање. Фрејмови, а у новијим стандардима и слајсеви, су груписани у неколико типова, према зависности од других фрејмова, односно слајсева:

- I - зависе само од пиксела са истог фрејма (слајса) – интра-фрејм предикција
- P - зависе од пиксела са претходних фрејмова (слајсева)
- B - зависе од пиксела са претходних и следећих фрејмова (слајсева)

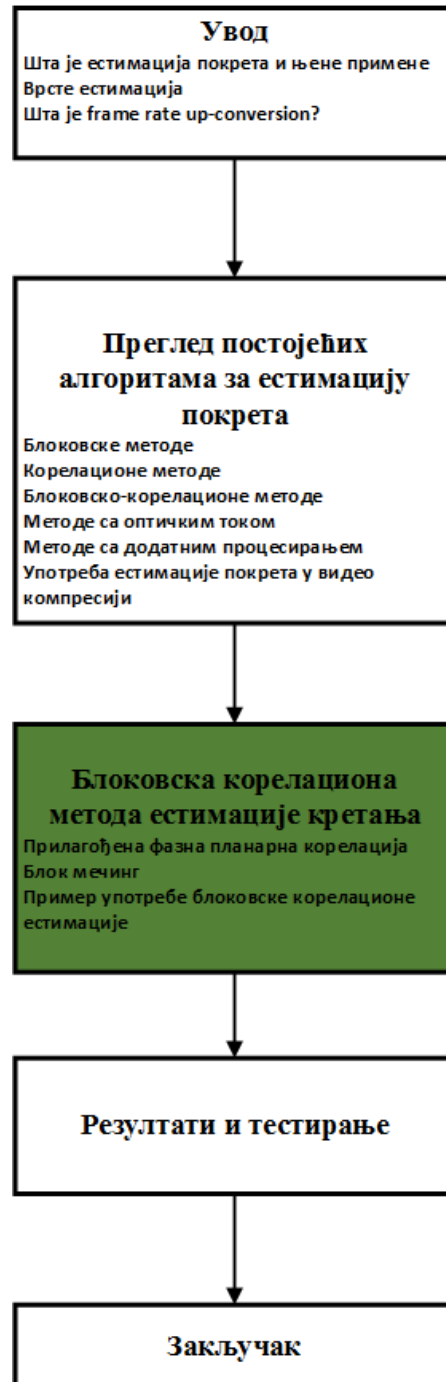


Слика 2-9: Процес кодовања једног фрејма (слајса) заједнички за већину видео компресија.

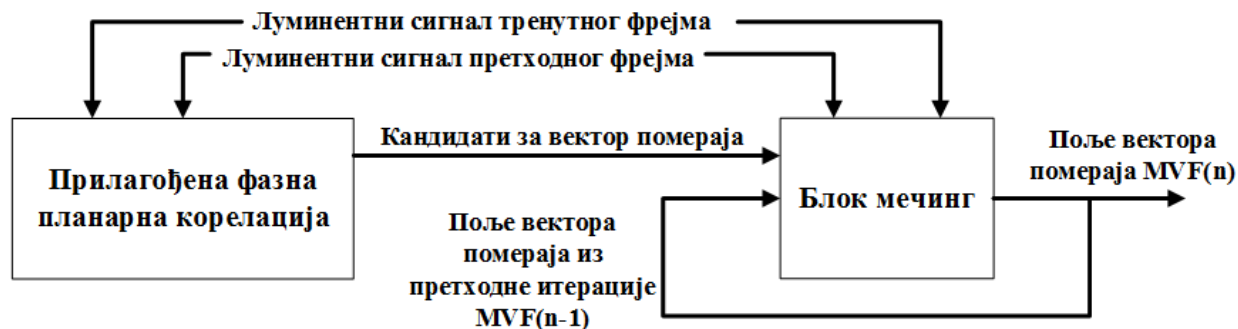
Потом се врши унидирекциона естимација кретања за фрејм (слајс) чија је референца (полазна тачка) на временској позицији тог фрејма. Овим се види директна употреба самих алгоритама за естимацију на месту којих се може наћи већина описаних алгоритама у претходном поглављу, а свакако и блоковска корелација предложена у овој тези. На основу естимације врши се физичко креирање самог предикционог фрејма (слајса). Предикциони фрејм (слајс) се затим одузима од оригиналног фрејма (слајса) креирајући диференцијалну разлику. У најидеалнијем случају диференцијална разлика је нула и тада није потребно додатно преносити информације о самом фрејму (слајсу). Уколико то није случај врши се дискретна косинусна трансформација над њом да би се омогућила што боља квантизација, односно компромис између преношења што мање информација потребних за реконструкцију диференцијалне разлике и одржања истодобности између улазне диференцијалне разлике и реконструисане диференцијалне разлике. Над излазом квантизације врши се статистичка

обрада и ентропијско кодовање без губитака, који креирају компримовану видео секвенцу на излазу.

### 3. Блоковска корелациона метода естимације кретања



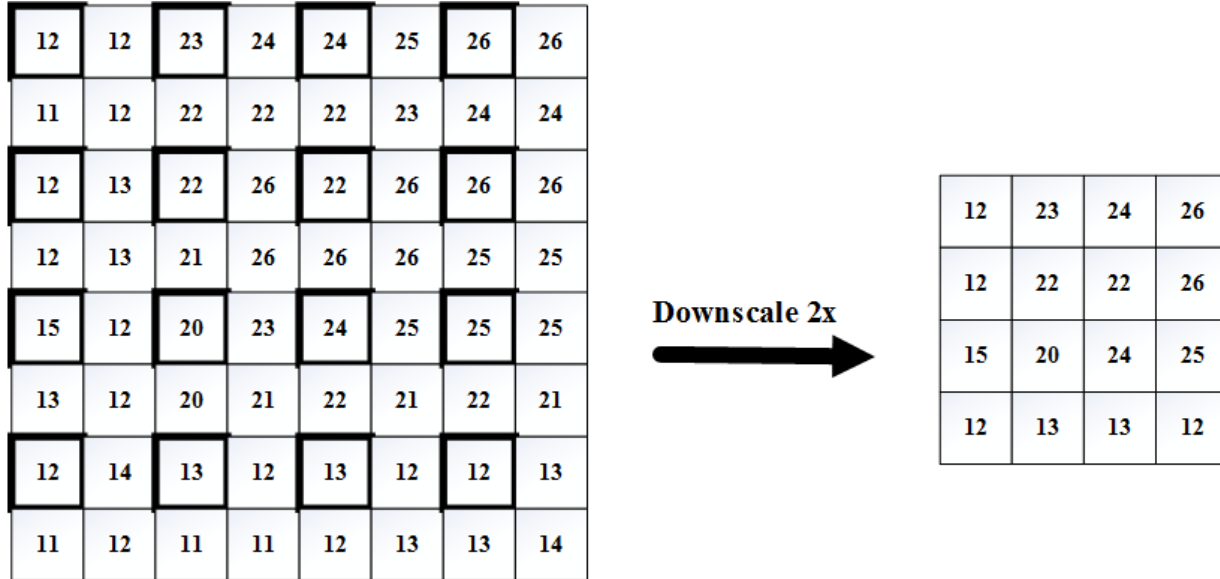
Метода за естимацију покрета предложена у овој тези заснива се на блок мечинг корелационом (*Block Matching Correlation*, БМК) приступу. Она комбинује снагу специјално прилагођене фазне планарне корелације (*Customized Phase plane correlation*), ПФПК и поменутог блок мечинг (*Block matching*, БМ) приступа на начин који приказује Слика 3-1 **Error! Reference source not found.**



Слика 3-1: Блок дијаграм блоковско корелационе методе предложене у овој тези (БМК). Улази у алгоритам су вредности пиксела са претходног фрејма, следећег фрејма и поље вектора из претходне итерације БМК алгоритма –  $MVF(n-1)$ . Излаз је вектор поље за тренутни пар фрејмова –  $MVF(n)$ .

ПФПК врши естимацију покрета и пропагира иницијалне процене у БМ модул као сет кандидата вектора покрета. Главна идеја је да ПФПК на тај начин врши добру иницијалну процену вектора помераја, а БМ даље фино подешава (*fine tuning*) и додељује векторе помераја одговарајућим блоковима. Да би се смањила рачунарска комплексност ФПК-а извршено је његово прилагођење на тај начин што је смањена величина улазних фрејмова (*downscaling*) и редукавана потреба за операцијама са покретним зарезом (*floating-point operations*). Утицај ове оптимизације на прецизност рачунања вектора помераја је занемарив, што су потврдила и извршена мерења и тестови дати у наредном поглављу. Могућност саме оптимизације постоји највише због саме робусности алгоритма за

корелацију сигнала. *Downscaling* је извршен без било каквог филтрирања, већ једноставним изостављањем пиксела (Слика 3-2).



Слика 3-2: Редукција броја пиксела на улазним фрејмовима коришћена код БМК која се састоји од изостављања пиксела. Пример приказује слику димензија 16x16 са приказаним вредностима пиксела. После скалирања два пута само ће маркирани пиксели остати у излазној слици.

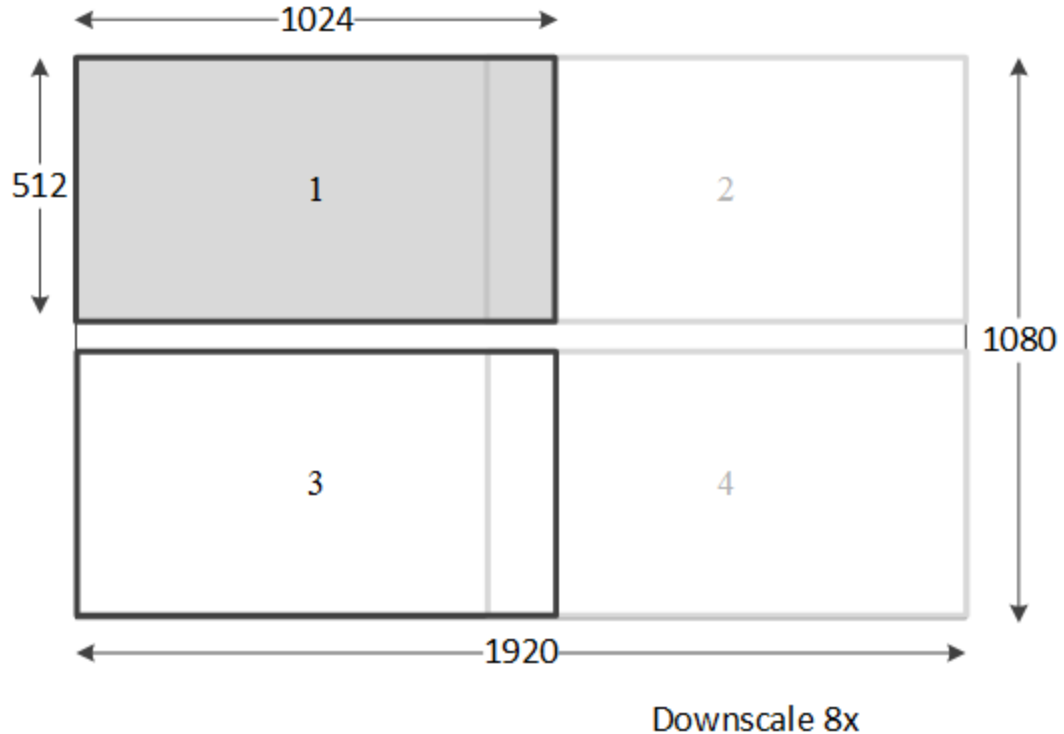
БМ је базиран на просторној и временској рекурзији између кандидата сваког од блокова, сличној као код 3-димензионалног рекурзивног претраживања [9] [10]. Излаз БМ-а је поље вектора помераја које садржи најподобнији вектор помераја за сваки блок у фрејму. Детаљан опис прилагођене фазне планарне корелације и блок мечинг модула, као саставних делова БМК-а, дати су у наставку.

### 3.1 Прилагођена фазна планарна корелација

Главна функционалност овог модула је естимација најзаступљенијих помераја између два улазна фрејма видео секвенце. ПФПК подразумева да су улазни фрејмови у *YCbCr* систему боја, најчешће коришћеном у алгоритмима за процесирање видео секвенци, и користи само лума компоненту (*Y*) која садржи више детаља од хрома компоненти (*CbCr*). Познато је да ће сигнали који садрже шири спектар учестаности (у случају слика то је већа количина детаља) дати поузданију и прецизнију корелацију. ПФПК има два нивоа хијерархије:

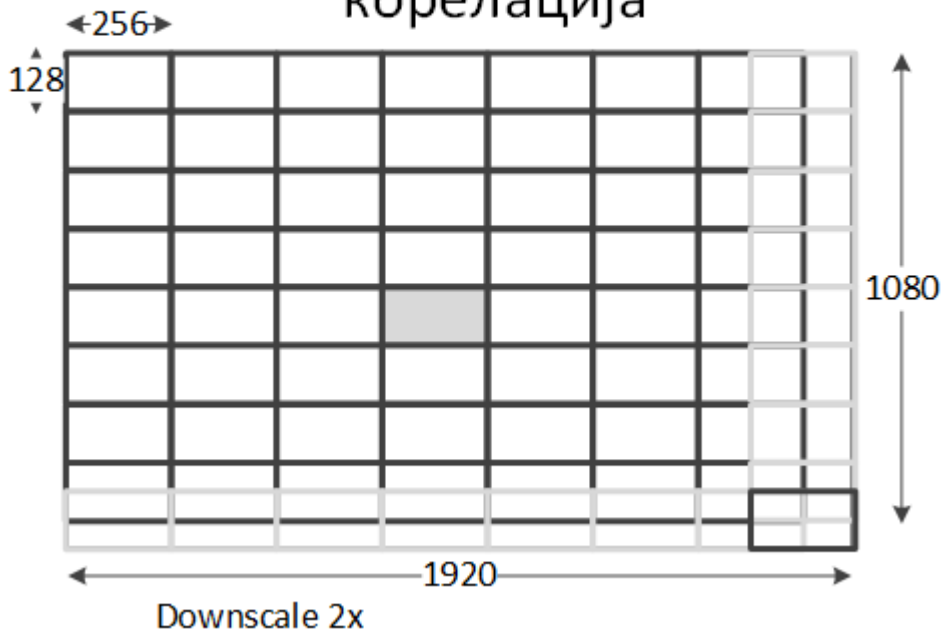
1. Глобална фазна планарна корелација (*Global Phase Plane Correlation*) дели фрејм на 4 региона једнаке величине који заједно покривају највећи могући део фрејма. Сваки од региона се смањује (*downscaling*) са фактором који је степен броја 2 (ради лакше имплементације) до величине стандардног ФПК региона (Слика 3-3, Слика 3-5). Фактор смањења зависи од резолуције фрејма. Потом се на смањеном региону врши ПФПК.
2. Локална фазна планарна корелација (*Local Phase Plane Correlation*) дели фрејм на једнаке регионе чија је величина једнака величини стандардног ФПК региона или 2 пута већег од њега по хоризонтали или вертикали, у зависности од резолуције фрејма (Слика 3-4, Слика 3-5). ПФПК се врши над сваким од добијених региона.

## Глобална фазна планарна корелација

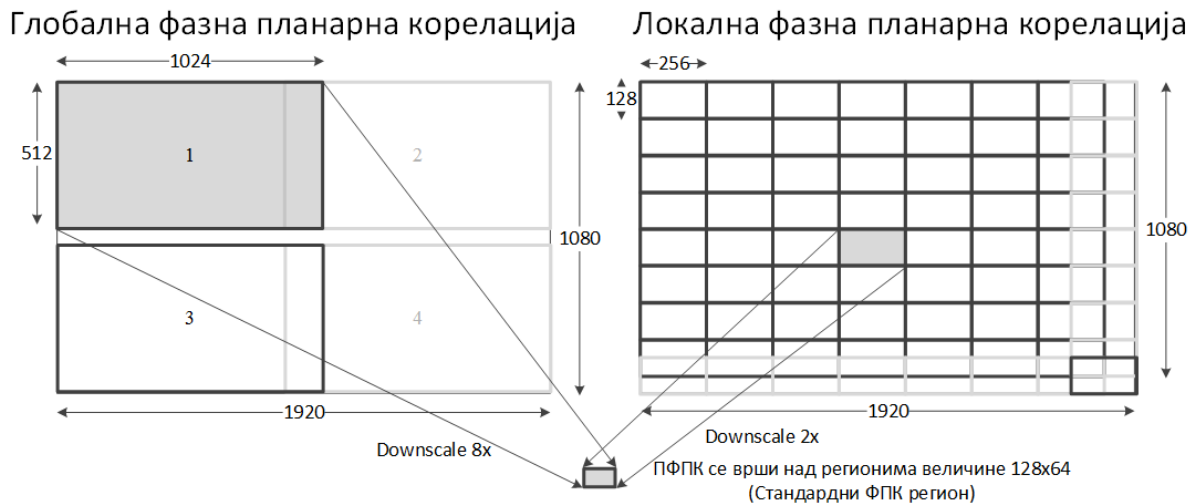


Слика 3-3: Пример поделе фрејма на регионе код глобалне фазне планарне корелације.

## Локална фазна планарна корелација



Слика 3-4: Пример поделе фрејма на регионе код локалне фазне планарне корелације.



Слика 3-5: Пример поделе фрејма на локалне и глобалне ФПК регионе за 1080п резолуцију, уз постизање исте величине стандардног ФПК региона и код глобалне и код локалне фазне планарне корелације.

Слика 3-5 показује пример поделе фрејма за глобалну и локалну фазну планарну корелацију за резолуцију 1920x1080 пиксела (1080п). Уочава се да оваквом поделом фрејма

на локалне и глобалне регионе и избор оваквих фактора скалирања омогућава извршавање фазне планарне корелације увек над регионом исте величине (у овом случају 128x64). Увођењем ове униформности знатно је убрзано извршавање фазне планарне корелације над целим паром фрејмова видео секвенце.

Робусност фазне планарне корелације омогућила је да се смањење фрејмова (*downscaling*) обави користећи једноставно изостављање (избацивање) пиксела уз занемаривање ефекта алајсинга. Прецизност овако добијених вектора помераја није тачна у пиксел, због претходно примењеног изостављања пиксела, већ је једнака фактору смањења региона коришћеног код глобалне и локалне фазне планарне корелације. Већа прецизност није потребна јер ће их БМ модул, коме ће естимирани помераји бити прослеђени, фино подесити и учинити тачним.

Димензије региона над којим се врши ПФПК је константа и дефинисана као стандардни ФПК регион. Ово омогућава још униформнију имплементацију самог алгорита и постојање неколико нивоа хијерархије који се разликују само у фактору скалирања примењеним на оригиналне улазне слике.

Приступ ФПК предложен у овој тези заснива се на корацима 1-5 наведеним у претходном поглављу и уводи неколико модификација. Дискретна Фуријеова трансформација имплементирана је користећи *radix-2 Fast Fourier Transformation* алгорита. Он је додатно оптимизован да ради са целим (*fixed point*) комплексним бројевима. Такође, постигнут је и висок ниво паралелизма инструкција и података који погодује имплементацији на паралелном векторском процесорском систему описаном у наставку. Поменути корак нормализације спектра снаге замењен је апроксимативном нормализацијом. Апроксимативна нормализација над сваким комплексним бројем подразумева:

- Срачунавање максимума између апсолутних вредности реалног и имагинарног дела;
- Проналажење броја водећих нула максимума;
- Шифтовање у лево максимума за срачунати број водећих нула.

На тај начин се добија апроксимативно нормализована вредност производа спектра.

Имплементација инверзне Фуријеове трансформације у оквиру ПФПК разликује се од директне једино у знаку улазног аргумента. Захваљујући томе могуће је користити исти код и за директну и за инверзну Фуријеову трансформацију, чиме је значајно редукован број потребних инструкција за рачунање ПФПК. Излаз модула прилагођене фазне планарне корелације је сет вектора помераја који садржи по два кандидата за сваки локални и глобални ФПК регион. С обзиром да постоји велики број локалних ФПК региона (нпр. 72 за 1080п резолуцију) вероватноћа постојања више од два различита кретања у оквиру једног региона је веома мала у реалним (несинтетизованим) видео секвенцама које су од интереса.



Са прва два ФПК кандидата, који одговарају двома највишим пиковима у корелационој равни, такође одговарају и најзаступљенија кретања у том региону. Због поменутог, чак и уз постојање више од два кретања по ФПК региону, објекти који њима одговарају су мали и њихово кретање се може занемарити (нпр. сцена бацања конфета или лишће на ветру, за примену у повећању боја фрејмова у секунди, њихово кретање је занемариво и неће додатно унапредити квалитет излазне секвенце).

Постигнути резултати су показали да целобројна имплементација предложена у оквиру ПФПК није смањила квалитет естимације, што је још један од значајних доприноса предложеног решења. Мерења која потврђују поменуто су дата у наставку.

### 3.2 Блок мечинг

Блок мечинг модул прима следеће улазе (Слика 3-6):

- Луминентне компоненте два суседна фрејма видео секвенце;
- Сет кандидата вектора помераја срачунатих у ПФПК модулу над истим паром улазних фрејмова;
- Срачунато поље вектора над претходним паром фрејмова.



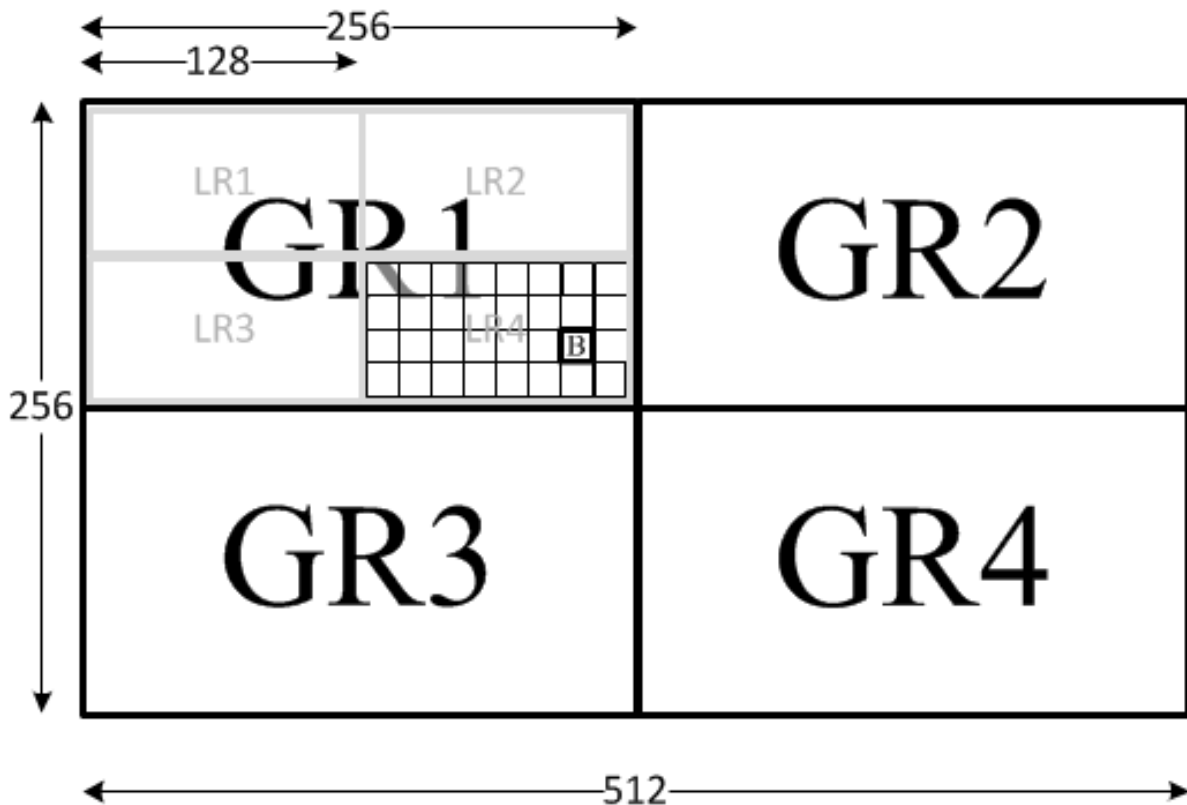
Слика 3-6: Дијаграм улазних и излазних података Блок мечинг модула.

Блок мечинг модул проналази најбољи вектор помераја за сваки блок у фрејму. Критеријум за избор најбољег кандидата је поменута сума апсолутних разлика (*Sum of absolute differences*). Укупно, за сваки блок, постоји осам кандидата за које се врши евалуација:

- Два локална ФПК кандидата ( $L1$  и  $L2$ );

- Два глобална ФПК кандидата ( $G1$  и  $G2$ );
- Кандидат из просторног домена са претходно евалуираног блока ( $S1$ );
- Кандидат из просторног домена са блока изнад оног који се тренутно евалуира ( $S2$ );
- Рандом кандидат формиран тако што се на  $S1$  дода мали, псеудослучајни шум ( $R$ );
- Медијан суседних кандидата из временског домена (претходне итерације блок мечинга) ( $MT$ ).

Блок мечинг користи излазе из ПФПК модула. Ови кандидати за вектор помераја нису довољно прецизни, али обезбеђују добру иницијалну процену за блок мечинг коју ће он, уграђеном рекурзијом кроз време и простор, тјуновати и учинити прецизном. Локални и глобални ФПК кандидати за блок се добијају из локалних и глобалних ФПК региона који тај блок обухватају. Пример поменутог мапирања приказује Слика 3-7.

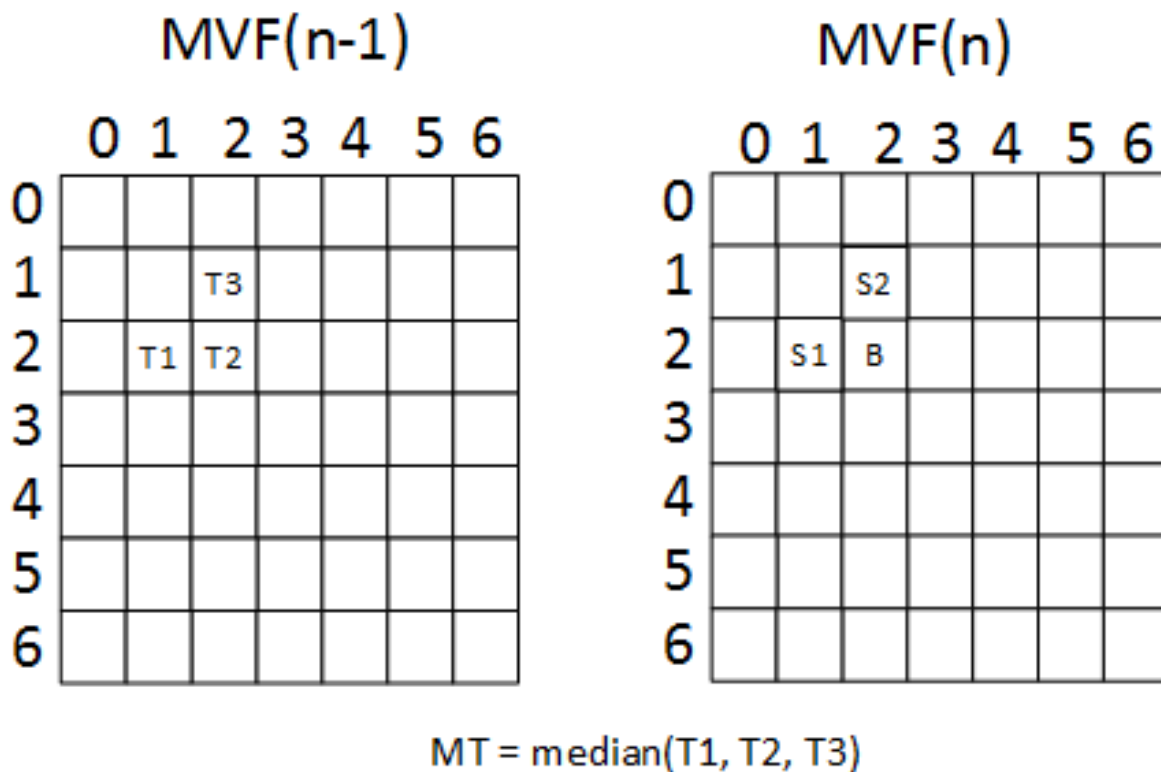


Слика 3-7: Пример мапирања ФПК кандидата на блокове фрејма. Блок  $B$  ће примити глобалне кандидате из региона  $GR1$  и локалне кандидате из региона  $LR4$ . Резолуција фрејма у овом примеру је 512x256 пиксела.

Вектор помераја који је био најбољи за претходни блок (имао најмању суму апсолутних разлика) представља просторни кандидат за тренутни блок ( $S1$  и  $S2$ ). Поменуто се оправдава чињеницом да у реалним видео секвенцама, као што је речено, постоји велика вероватноћа да суседни блокови имају исти померај. Блок мечинг такође евалуира

просторни кандидат на кога је додата мала, псеудослучајна, рационална вредност [10] . На тај начин могуће је добити векторе помераја који нису целобројни већ фракциони и њима се постиже додатна прецизност у естимацији кретања. Такође, овај концепт је широко заступљен у видео компресији [30] . Сума апсолутних разлика за фракционе помераје се рачуна користећи билинеарну интерполацију.

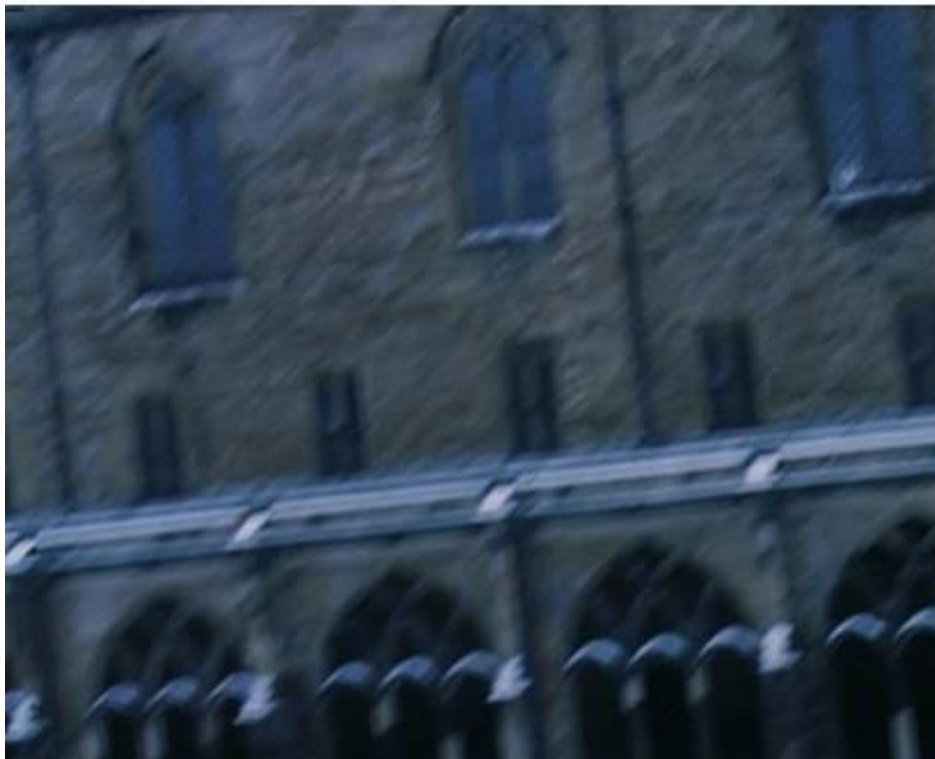
Кретање у видео секвенцама је најчешће униформно у времену. Евалуирање вектора помераја суседних блокова који су били победници у претходној итерацији блок мечинга такође има високу вероватноћу за проналажење најбољег кандидата за тренутни блок. Због тога блок мечинг узима победнике (типично три) из претходно срачунатог поља вектора и њиховим медијан филтрирањем добија временски кандидат ( $MT$ ). Пример срачунавања  $SI$  и  $MT$  кандидата приказује Слика 3-8.



Слика 3-8: Начин избора позиције просторних кандидата ( $SI$  и  $S2$ ) и темпоралног кандидата  $MT$  за блок  $B$ . Темпорални кандидат за блок  $B$  у пољу вектора које се тренутно срачунава  $MVF(n)$  се формира користећи векторе помераја из претходно срачунатог поља  $MVF(n-1)$ .

### 3.3 Процесирање ивичних блокова

Видео секвенце често садрже веће помераје настале усред кретања камере, саме позадине или објеката у њој. Код таквих секвенци, на ивицама фрејмова немогуће је вршити поређење блокова применом вектора који компензује кретање јер ће се изаћи из оквира фрејма. Због тога, већина естиматора помераја на ивицама фрејма дају нула вектор помераја, који ће у примени у компензацији у *frame-rate upconversion* алгоритмима дати веома уочљиве артефакте (Слика 3-9, горе, лева ивица фрејма и Слика 3-10, лево, десна ивица фрејма).

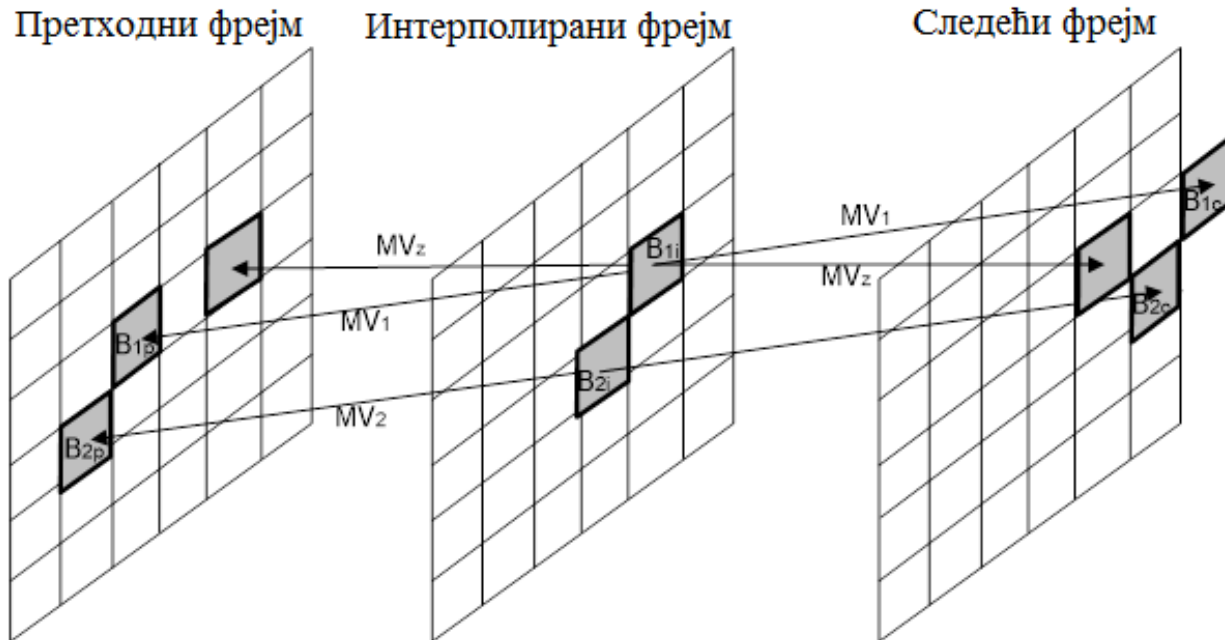


Слика 3-9: Интерполирани фрејм *Harry Potter* видео секвенце. Слика горе не садржи специјално процесирање ивичних блокова због чега су видљиви блоковски артефакти на левој ивици фрејма. Слика доле представља излаз алгоритма за естимацију са примењеним процесирањем ивичних блокова.



Слика 3-10: Интерполирани фрејм *College Football* видео секвенце. Слика лево не садржи специјално процесирање ивичних блокова због чега су видљиви блоковски артефакти на десној ивици фрејма. Слика десно представља излаз алгоритма за естимацију са примењеним процесирањем ивичних блокова и не садржи блоковске артефакте на десној ивици.

Ситуација када вектор који компензује кретање излази из оквира фрејма приказује Слика 3-11.



Слика 3-11: Репрезентација стварног кретања за ивичне блокове  $B_{1i}$  и  $B_{2i}$ . Вектори  $MV_1$  и  $MV_2$  су стварни вектори помераја између претходног и следећег фрејма. Применивши вектор  $MV_2$  на блок  $B_{2i}$  блокови  $B_{2p}$  и  $B_{2c}$  су у границама оба фрејма. Међутим, применивши вектор  $MV_1$  на блок  $B_{2i}$  блок  $B_{1c}$  је ван следећег фрејма.

Решење за естимирање стварног вектора помераја је препознати ситуацију коју описује Слика 3-11 за блок  $B_{1i}$ , доделити ивичном блоку вектор стварног помераја и при процесу компензације назначити да се за компензовање тог блока користе пиксели само са једног фрејма који је доступан. У предложеном блок мечинг решењу то је могуће урадити користећи просторно рекурзивни вектор помераја са претходно процесираним блоком ( $S1$  - Слика 3-8): Ако кандидат за вектор помераја добијен као најбољи за претходно процесирани блок ( $S1$ ) излази из граница фрејма прогласити га за најбољи блок без претходног срачунавања његове суме апсолутних вредности. Ово правило слободно се може применити за све блокове у фрејму, не само за одређени сет који је близак ивици фрејма.

Други фактор који омогућава доделу стварних вектора помераја ивичним блоковима је редослед процесирања блокова у фрејму. Блокови у фрејму су подељени на страјпове (Слика 3-12). У првом и последњем страјпу потребно је да смер процесирања блокова буде од унутра ка споља. На тај начин ће стварни вектори помераја, уграђеном рекурзијом, моћи стићи до ивичних блокова. Применом описаног процесирања ивичних блокова креиран је интерполирани фрејм без блоковских артефаката на ивицама (Слика 3-9).

Употреба временског рекурзивног кандидата у процесирању ивичног блока, иако изгледа као логично решење, даје артефакте и непријатно треперење ивичних блокова, које пропагира од фрејма до фрејма.



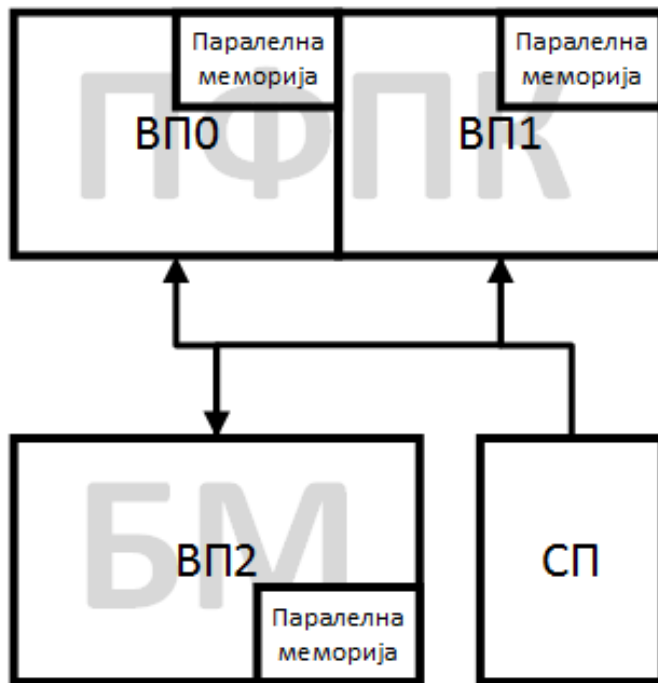
Слика 3-12: Смер процесирања (скенирања) блокова у фрејму. Фрејм је подељен на одређени број вертикалних целина – страјпова. Блокови се процесирају од горе ка доле, пратећи назначене стрелице. Такав редослед процесирања блокова омогућава доделу стварних вектора помераја ивичним блоковима. Вектор кандидат блока означеним зеленом бојом је просторни кандидат за назначени ивични блок. Вектор кандидати за блокове означене црвеном бојом још увек нису срачунати. Вектори са блокова означених љубичастом бојом такође нису подобни као кандидати за ивичне векторе помераја.

### 3.4 Пример употребе блоковске корелационе естимације

Описани БМК алгоритам за естимацију искоришћен је као главни део алгоритма за повећање броја фрејмова у секунди у реалном времену на Интел Мерифилд мобилној платформи. Он је имплементиран на Видео Сигнал Процесору који се састоји од више *Single-Instruction-Multiple-Data (SIMD)* и *Very Long Instruction Word (VLIW)* језгара као што приказује Слика 3-13. Извршавање БМК је подељено на три језгра која садрже паралелни меморијски подсистем погодан за процесирање блокова пиксела. Прилагођена фазна планарна корелација је подељена и извршава се на ВП0 и ВП1 векторским језгрима, док се БМ извршава на ВП2 језгру. Да би се постигло што ефикасније искоришћење брзе паралелне меморије унутар векторских процесора и смањио број приступа екстерној меморији у коју су смештени пиксели, поред блокова за које се тренутно срачунавају вектори помераја, у њу се смештају и сви суседни пиксели до којих се примењеним векторима може доћи (област претраживања). Скаларни процесор (СП) се користи за контролу извршавања, синхронизацију и интерфејс са Интел Мерифилд хост процесором



(*Intel Atom*). Постигнути паралелизам у процесирању података и извршавању инструкција, скалабилност, мало заузеће процесорске меморије и редукован број трансфера са екстерном меморијом омогућили су имплементацију комплетног БМК алгоритма у реалном времену на Интеловој мобилној платформи. Концепти и могућности поменуте архитектуре Видео Сигнал Процесора су описане у [31]. Изазови и захтеви за постизање извршавања алгоритама за повећање броја слика у секунди у реалном времену са минималном процесорском моћи су описане у [32].



Слика 3-13: Преглед распореда извршавања делова БМК алгоритма на *Intel Merrifield* векторском процесорском систему.

Изаз блок мечинга је, као што је речено, поље вектора помераја које садржи један вектор помераја за сваки блок у фрејму. Процес креирања новог (интерполираног) фрејма користећи срачунато поље вектора и вредности пиксела са претходног и следећег фрејма зове се компензација покрета (*motion compensation*). Нови фрејм је креиран на временској позицији видео секвенце између претходног и следећег фрејма, у зависности од траженог фактора конверзије.

Предложена блоковска корелациона естимација може бити реализована и као унидирекциона и бидирекциона естимација. Ради постизања имплементације у реалном времену изабрана је бидирекциона естимација која захтева мање меморије, а самим тим и мање меморијских трансфера. Такође она захтева и мањи број процесорских циклуса јер не

захтева ритајминг. Недостатак бидирекционе естимације, који се огледа у смањеној робусности односно већој вероватноћи за детекцију лажног кретања, није у потпуности превазиђен, али је значано минимизован самом структуром предложеног решења, а посебно просторном и временском рекурзијом коју садржи.

Табела 1 приказује перформансе описаног процесорског система при извршавању предложене естимације покрета. У првој врсти табеле приказан је број циклуса на сваком језгру потребних за процесирање 1080п секвенце на 30 фрејмова у секунди. Друга врста табеле приказује минималну учестаност такта процесора која омогућава процесирање у реалном времену. У досадашњим истраживањима не постоји објављен рад који описује имплементацију повећања броја фрејмова у секунди у видео секвенцама у реалном времену на мобилној платформи, што чини предложени алгоритам првим овог типа. Такође, због поменутог, није могуће извршити поређење перформанси са постојећим решењима.

Табела 1: Измерене перформансе извршавања БМК алгоритма на Интел Мерифилд (*Intel Merrifield*) мобилној платформи.

Језгро	ВП0	ВП1	ВП2
Број потребних циклуса за 1080п	6.552.446	6.552.446	3.765.124
Минимална учестаност такта језгра (MHz)	196.57	196.57	112.95

Табела 2 приказује дисипацију снаге измерену у току извршавања алгоритма за повећање броја слика у секунди у реалном времену. Поређење дисипиране снаге са другим алгоритмима за естимацију кретања или повећање броја слика у секунди не би било смислено пошто су разлике између мобилне платформе на којој се извршава предложено решење и хардверских, софтверских и *FPGA* платформи на којима се извршавају друга решења превелике.

Табела 2: Дисипирана снага при повећању броја фрејмова у секунди на 1080п видео секвенци.

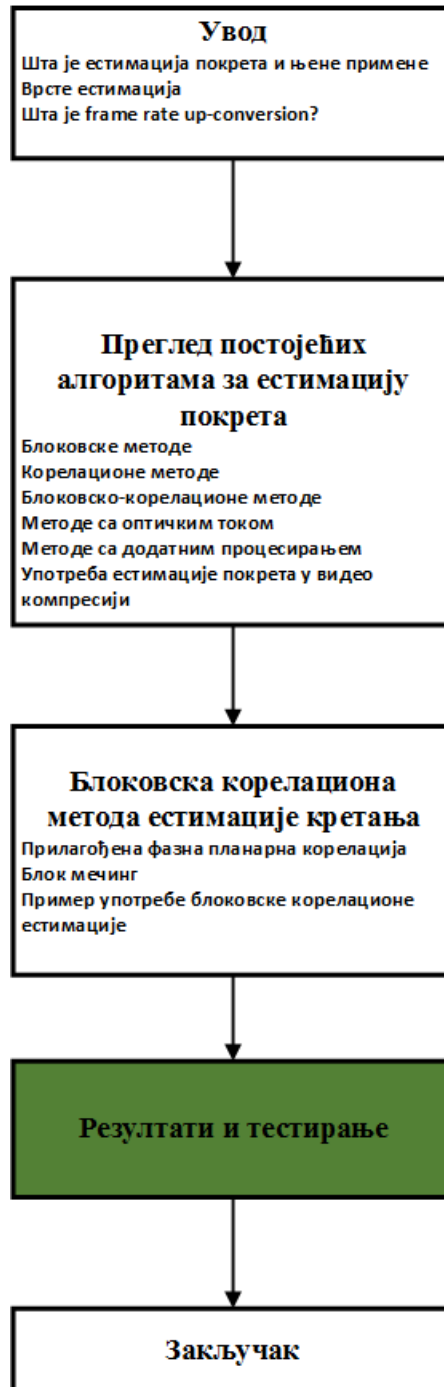
Дисипирана снага (1080п видео)	[mW]
Видео сигнал процесор	84
Трансфери са екстерном меморијом	48
Екстерна меморија	90
<b>Укупно:</b>	<b>222</b>

Слика 3-14 приказује прототип мобилног телефона базиран на *Intel Merrifield* платформи на коме се извршава алгоритам за повећање броја фрејмова у секунди у реалном времену. Све видео секвенце које су смештене у мобилном телефону, а кодоване на 24 или 30 фрејмова у секунди могуће је пуштати на 60 фрејмова у секунди, у реалном времену. Интерполирани фрејмови се, као што је раније напоменуто, не смештају трајно у меморију мобилног уређаја, већ се после приказивања кориснику уништавају. Иако су димензије дисплеја мобилног телефона релативно мале, разлике у глаткоћи приказа видео секвенце са и без алгоритма за повећање броја слика у секунди су веома уочљиве. Такође, пошто постоји могућност повезивања мобилног уређаја на већи дисплеј, на њему су разлике, односно глаткоћа кретања, још више уочљивија. Слика 3-14 приказује мобилни уређај на коме се извршава алгоритам за повећање броја слика у секунди у реалном времену, базиран на БМК естимацији кретања.



Слика 3-14: Телефон базиран на *Intel Merrifield* платформи на коме се извршава алгоритам за повећање броја фрејмова у секунди у реалном времену.

## 4. Резултати и тестирање



Предложена блоковска корелациона естимација се може користити у више различитих апликација. Њена евалуација је вршена на примеру алгоритма за повећање броја фрејмова у секунди у видео секвенцама. Вршено је поређење са поменутиим *Full search* и *3DRS*. *Full search* је изабран као скалабилно, конвенционално решење које се користи у многим хардверским системима, а *3DRS* такође због своје широке уопште и доброг односа између комплексности и квалитета естимације. Изабрана метода компензације покрета динамички медијан коришћена је код сваке методе за естимацију [33]. Као критеријум поређења коришћене су субјективне и објективне методе. Евалуација је вршена над сетом 1080п видео секвенци оригинално снимљеним са 60 фрејмова у секунди [34].

Објективна евалуација је извршена користећи следећу процедуру:

- Оригиналне тест секвенце су децимиране два пута, са 60 на 30 фрејмова у секунди (парни фрејмови су избачени);
- Над децимираним видео секвенцама је извршено повећање броја слика у секунди 2 пута користећи *FS*, *3DRS* и БМК као методе за естимацију кретања. (на тај начин су произведене излазне секвенце на 60 фрејмова у секунди);
- Креирани интерполирани фрејмови, на парним позицијама у излазној секвенци, се пореде са одговарајућим парним фрејмовима оригиналних видео секвенци користећи стандардне методе за мерење сличности фрејмова: *Peak Signal to Noise Ratio (PSNR)* и *Universal Image Quality Index (UIQI)* [35], који су детаљније описани у Додатку А.

Резултати објективне евалуације приказује Табела 3. Најбољи резултати су маркирани сивим оквиром.

Табела 3: Објективна мерења која пореде три методе за естимацију покрета: *Full search*, 3-димензионално рекурзивно претраживање (*3DRS*) и предложену Блок мечинг корелацију (БМК). Веће *PSNR* и *UIQI* вредности представљају већу сличност оригиналној секвенци (бољи квалитет естимације).

Видео секвенца (1080p, 60 fps)	PSNR [dB]			UIQI		
	FS	3DRS	БМК	FS	3DRS	БМК
Danube islands	42.23649	48.50988	<b>49.03328</b>	0.4898	0.547061	0.542702
Down the street	28.01018	32.81477	<b>34.27269</b>	0.717286	0.715841	<b>0.766175</b>
Forest in fall	22.0265	24.04839	<b>25.0733</b>	0.611143	0.603049	<b>0.659994</b>
Stockholm pan	30.45746	39.92172	<b>41.04493</b>	0.765433	0.807222	<b>0.809674</b>
Swans	35.10541	41.41942	<b>42.24626</b>	0.825245	0.831563	<b>0.848414</b>
Children playing	38.66661	<b>47.01015</b>	46.73785	0.963057	<b>0.973596</b>	0.973511
Conference call	33.87459	<b>48.44191</b>	48.30221	0.82082	0.849379	<b>0.848063</b>
Fruit yard	23.23485	25.19292	<b>26.12325</b>	0.653106	<b>0.667879</b>	0.723098
Tango argentino	34.93169	<b>48.39713</b>	48.2522	0.918017	0.952733	<b>0.952239</b>
<b>Укупно (сума):</b>	<b>288.5438</b>	<b>355.7563</b>	<b>361.086</b>	<b>6.763907</b>	<b>6.948323</b>	<b>7.12387</b>

Резултати које приказује Табела 3 недвосмислено приказују предност БМК у односу на FS и 3DRS. БМК има укупну предност у PSNR вредностима од 5.33 dB у односу на 3DRS и 72.54 dB у односу на FS на девет изабраних тест секвенци. Ако поредимо UIQI бројеве, БМК има укупну предност од 0,176 у односу на 3DRS и 0.360 у односу на FS на истом сету секвенци. Иако сумирање PSNR и UIQI вредности не представља прецизно и потпуно легитимно средство поређења ипак се може посматрати као груба, генерална процена односа квалитета естимације. Индивидуално, PSNR, БМК даје бољи резултат у односу на 3DRS на шест од укупно девет видео секвенци и бољи у односу на FS на свих девет секвенци. Што се тиче UIQI, БМК је бољи на шест секвенци, а скоро једнак на преостале три, док у односу на FS даје бољи резултат на свих девет секвенци.

Још једно поређење алгоритама за естимацију покрета приказује Табела 4. Сет секвенци и PSNR бројке за MFHME су преузети из [22] , а за VIF из [23] . Потпуна имплементација ова два решења није доступна због чега се поређење врши само на секвенцама чији резултати су приказани у наведеним радовима. Различитост платформе и хардверске архитектуре не дозвољавају квантитативно поређење БМК са MFHME и VIF. Са алгоритамског становишта, три фрејма коришћена за естимацију код MFHME и четири фрејма креирана код VIF у поређењу са само два фрејма која користи БМК, захтевају значајно више меморије за процесирање, а самим тим и меморијских трансфера, што продужава укупно време извршавања. Такође, додатно процесирање у MFHME за рачунање вероватноћа и пиксел базирани *optical flow* коришћен код VIF, додатно повећавају време процесирања и комплексност. Имајући то у виду, упоредиви PSNR резултати БМК у односу на ове две методе представљају значајно достигнуће предложеног алгорита.

Табела 4: Поређење објективних метрика пет метода за естимацију кретања: *Full search*, *3-dimensional search*, *Multi-frame hierarchical motion estimation*, *Variational Image Fusion* и блок мечинг корелације (БМК). Веће PSNR вредности (маркиране сивом бојом у табели) представљају већу сличност оригиналној секвенци (бољи квалитет естимације).

Видео секвенца	PSNR[dB]				
	FS	3DRS	MFHME	VIF	БМК
City	24.18	26.2	33.03	<b>34.41</b>	26.83
Mobile	20.34	26.83	25.75	<b>29.02</b>	27.13
Foreman	24.7	33.52	<b>35.9</b>	32.86	34.88
Crew	26.96	31.25	32.92	-	<b>33.17</b>
Harbor	20.87	23.27	24.82	<b>30.48</b>	23.39
Flower	17.61	26.41	27.97	<b>31.46</b>	28.08
Ice	17.9	27.19	31.59	<b>33.34</b>	25.79
Bus	20.06	23.37	29.37	27.22	<b>29.66</b>

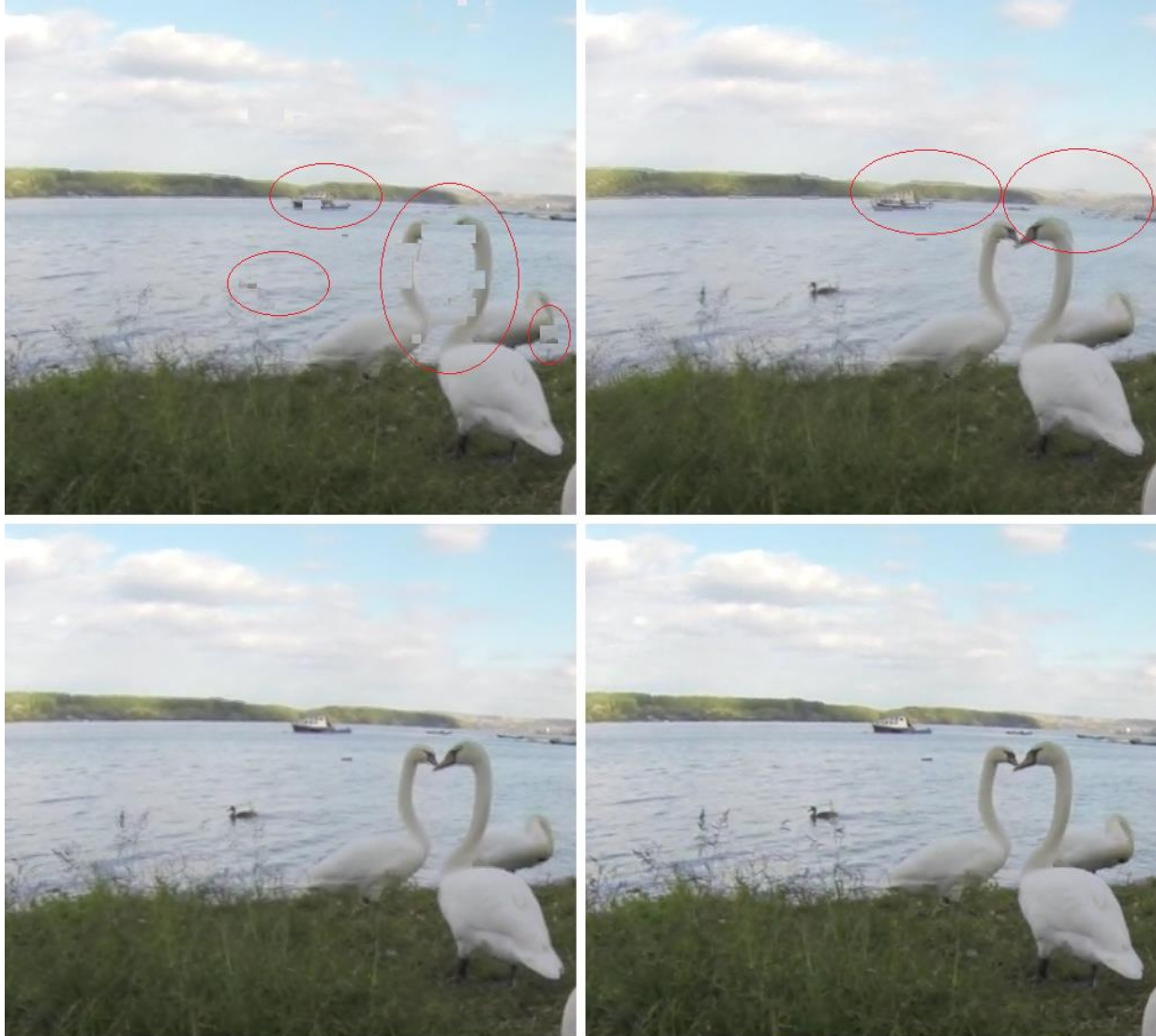
У једном од претходних поглавља описана је прилагођена фазна планарна корелација имплементирана у целобројном бројном систему и без дељења у кораку нормализације. Да би се измерила евентуално настала деградација квалитета естимације настала услед ове оптимизације дато је поређење PSNR и UIQI вредности две инстанце БМК (Табела 5): (1) целобројна имплементација Прилагођене Фазне Планарне Корелације (ПФПК) и (2) имплементација фазне планарне корелације у бројчаном систему са покретном тачком (*floating point*), ФФПК. Скоро идентичне измерене вредности обе метрике потврђују да квалитет самог БМК није нарушен услед описаних оптимизација.

Табела 5: PSNR и UIQI представљају поређење између две инстанце БМК, прве са прилагођеном фазном планарном корелацијом (ПФПК) и друге са имплементацијом Фазне планарне корелације у бројчаном систему са покретном тачком (ФФПК).

Видео секвенца (1080p, 60 fps)	PSNR [dB]		UIQI	
	ФФПК	ПФПК	ФФПК	ПФПК
Danube islands	49.040058	49.03328	0.541157	0.542702
Down the street	34.274691	34.27269	0.766267	0.766175
Forest in fall	25.0715	25.0733	0.659833	0.659994
Stockholm pan	41.04551	41.04493	0.809666	0.809674
Swans	42.254769	42.24626	0.84906	0.848414
Children playing	46.74247	46.73785	0.973534	0.973511
Conference call	48.304009	48.30221	0.848075	0.848063
Fruit yard	26.125974	26.12325	0.723206	0.723098
Tango argentine	48.26444	48.2522	0.952244	0.952239
<b>Average difference for all:</b>	<b>0.00456</b>		<b>0.00029</b>	

PSNR и UIQI мерења могу понекад одступати од субјективног утиска насталог после гледања видео секвенце. Због тога је извршена и субјективна евалуација визуелног утиска између фрејмова креираних коришћењем FS, 3DRS и БМК. Слика 4-1 и Слика 4-2 приказују издвојене интерполиране фрејмове продукване овим алгоритмима, као и одговарајући оригинални фрејм. Блокоски артефакти у горњим левим деловима обе слике настали су због поменуте неробусности бидирекционе естимације и самог FS алгоритма. Фрејмови креирани коришћењем 3DRS алгоритма не садрже тако видљиве артефакте, али назначена замућења (брод и врх брда на горњем десном делу, Слика 4-1, и танке гране, Слика 4-2) указују на непрецизност естимираног кретања. Интерполирани фрејмови настали коришћењем БМК, које приказују Слика 4-1 и Слика 4-2 доле-лево не садрже видљиве разлике у односу на оригинални фрејм (доле-десно). Процесирани секвенце су јавно доступне [34].





Слика 4-1: Визуелно поређење интерполираних фрејмова видео секвенце *Swans*, креираних алгоритмом за повећање броја слика у секунди у видео секвенцама користећи различите методе за естимацију кретања: горе-лево - *FS*, горе-десно - *3DRS*, доле-лево - *БМК*, доле-десно – оригинални фрејм. Артефакти настали усред погрешне или недовољно прецизне естимације су заокружени црвеном бојом.



Слика 4-2: Визуелно поређење интерполираних фрејмова видео секвенце *Fruit yard*, креираних алгоритмом за повећање броја слика у секунди у видео секвенцама користећи различите методе за естимацију кретања: горе-лево - *FS*, горе-десно - *3DRS*, доле-лево - *БМК*, доле-десно - оригинални фрејм. Артефакти настали усред погрешне или недовољно прецизне естимације су заокружени црвеном бојом.

Слика 4-3 и Слика 4-4 приказују пар оригиналних фрејмова и интерполирани фрејм између њих. Интерполирани фрејм је креиран предложеним алгоритмом за повећање броја слика у секунди. Увећани делови у горњем десном углу потврђују прецизност естимираног кретања. Уочава се да је квалитет интерполације довољно висок да није могуће визуелном инспекцијом утврдити разлику између оригиналних и интерполираних фрејмова, посебно не при гледању видео секвенце.





Слика 4-3: Оригинални фрејмови видео секвенце *Street traffic* (горе и доле) и интерполирани фрејм између њих.



Слика 4-4: Оригинални фрејмови видео секвенце *Matrix* (горе и доле) и интерполирани фрејм између њих.

Слика 4-5, Слика 4-6, Слика 4-7, Слика 4-8, Слика 4-9, Слика 4-10, Слика 4-11, Слика 4-12, Слика 4-13, Слика 4-14, Слика 4-15 и Слика 4-16 приказују оригиналне фрејмове видео секвенци и интерполирани фрејм између њих креиран коришћењем БМК алгоритма предложеног у овој тези. Визуелном инспекцијом се уочава да су интерполирани фрејмови, сем помераја који постоји, готово идентични оригиналним.



Слика 4-5: Два оригинална и интерполирани фрејм између њих видео секвенце *Foreman*, креиран помоћу описаног БМК алгоритма.





Слика 4-6: Два оригинална и интерполирани фрејм између њих видео секвенце *Belgrade aerial*, креиран помоћу описаног БМК алгоритма.



Слика 4-7: Два оригинална и интерполирани фрејм између њих видео секвенце *Conference call*, креиран помоћу описаног БМК алгоритма.



Слика 4-8: Два оригинална и интерполирани фрејм између њих видео секвенце *Danube islands*, креиран помоћу описаног БМК алгоритма.





Слика 4-9: Два оригинална и интерполирани фрејм између њих видео секвенце *Forest in fall*, креиран помоћу описаног БМК алгоритма.





Слика 4-10: Два оригинална и интерполирани фрејм између њих видео секвенце *Children playing*, креиран помоћу описаног БМК алгоритма.



Слика 4-11: Два оригинална и интерполирани фрејм између њих видео секвенце *Stockholm pan*, креиран помоћу описаног БМК алгоритма.





Слика 4-12: Два оригинална и интерполирани фрејм између њих видео секвенце *Down the street*, креиран помоћу описаног БМК алгоритма.



Слика 4-13: Два оригинална и интерполирани фрејм између њих видео секвенце *Tango argentino*, креиран помоћу описаног БМК алгоритма.



Већина филмова предвиђених преvasходно за биоскопске пројекције оригинално су снимљене са 24 фрејма у секунди. Да би се омогућило и њихово приказивање на 60 фрејмова у секунди потребно је извршити конверзију са фактором 2.5, као што приказује Слика 1-4. Конкретно, потребно је између сваког другог фрејма оригиналне видео секвенце убацити четири интерполирана фрејма, а оригинални фрејм избацити. Слика 4-14, Слика 4-15 и Слика 4-16 приказују пар оригиналних фрејмова видео секвенце и 4 интерполирана фрејма између њих креирани алгоритмом за повећање броја слика у секунди базираном на предложеној БМК естимацији кретања.



Слика 4-14: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце *I robot*, креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама  $n$  и  $n+2$ .



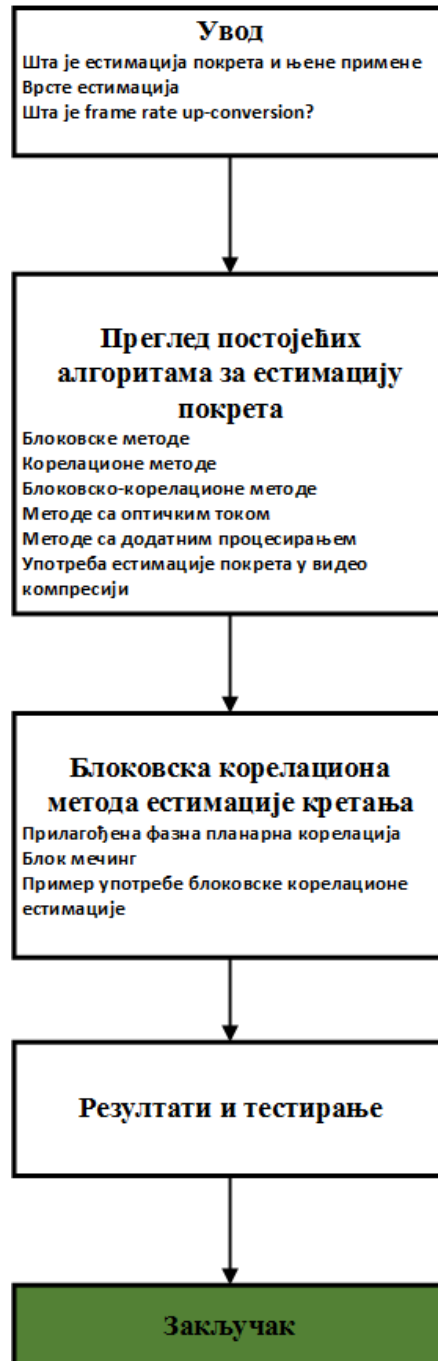
Слика 4-15: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце *Matrix*, креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама  $n$  и  $n+2$ .



Слика 4-16: Два оригинална (горе-лево и доле-десно) и четири интерполирана фрејма између њих видео секвенце *Spiderman*, креиран помоћу описаног БМК алгоритма. Оригинални фрејмови видео секвенце су на временским позицијама  $n$  и  $n+2$ .



## 5. Закључак



Блок мечинг корелација као описани алгоритам за естимацију покрета у видео секвенцама представља главни део алгоритма за повећање броја фрејмова у секунди у видео секвенцама ради уклањања непријатног џадера и повећања утиска гледаоца. Прецизност БМК, измерена преко креирања нових, интерполираних фрејмова је доказана субјективним поређењем и објективним (*PSNR* and *UIQI*) метрикама. Смањена рачунарска комплексност и потребан меморијски проток омогућили су имплементацију БМК алгоритма у реалном времену, која је демонстрирана на Интеловој мобилној платформи. Није потребно додатно нагласити очигледну предност извршавања повећања броја фрејмова у секунди у реалном времену у односу на ону која би се изводила у такозваном *off-line* режиму, која, поред тога што је спорија, захтева и огромну додатну количину меморије за смештање свих интерполираних фрејмова, пре него што би били приказани гледаоцу. Због поменутих особина може се закључити да БМК значајно повећава угођај гледаоца за време гледања видео секвенце на мобилним уређајима, као што су мобилни телефони и таблети.

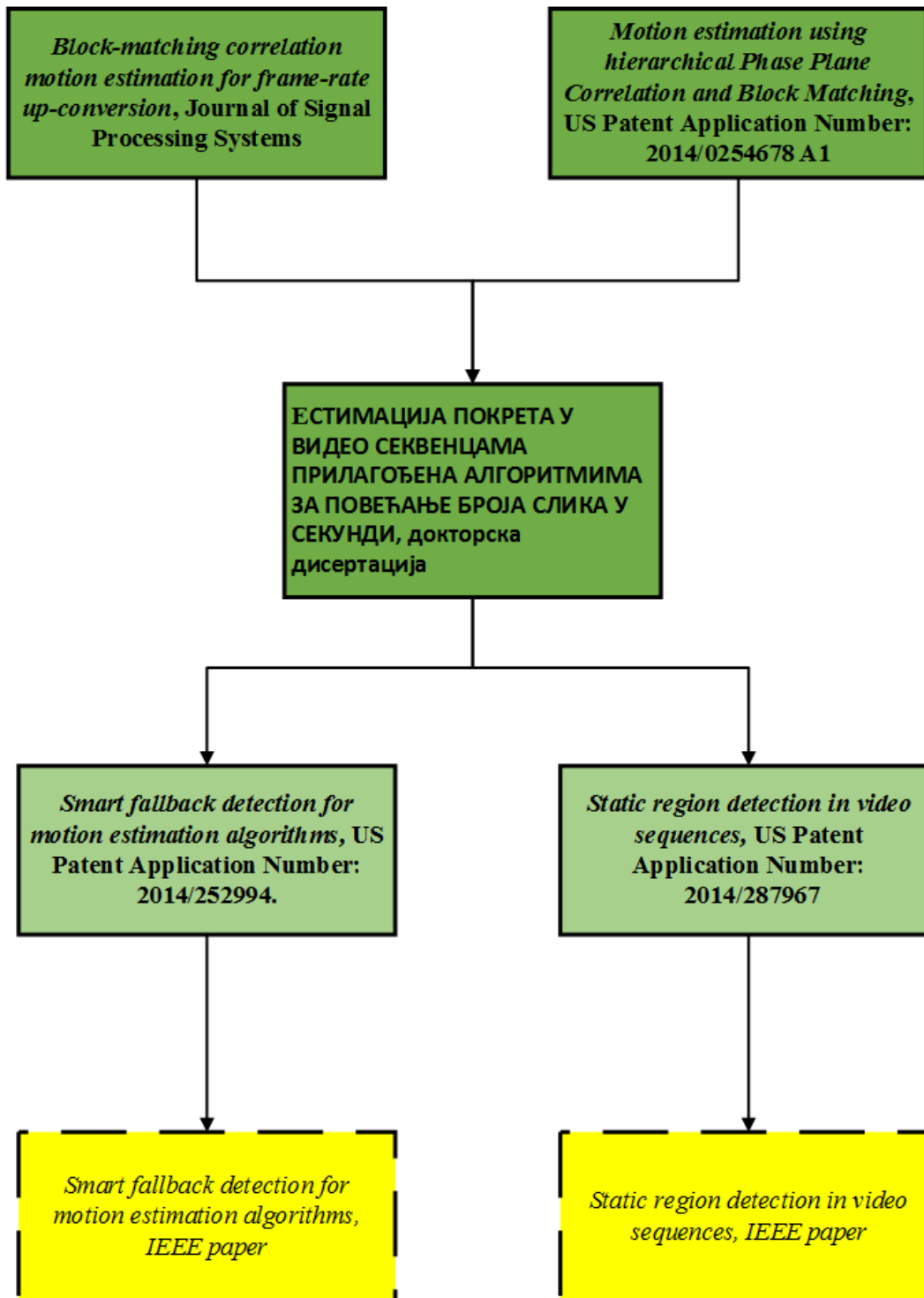
Блок мечинг и Фазна планарна корелација су добро познате методе често коришћене у ранијим решењима за естимацију покрета. Иновативност предложене блоковске корелације се огледа у јединственом начину њиховог повезивања и постизању сличног или бољег квалитета естимације од постојећих решења. Оно што највише издваја предложено решење од осталих алгоритама за естимацију кретања у видео секвенцама су модификације и оптимизације којима је значајно смањен број потребних операција и које нису смањиле квалитет саме естимације. Међу њима, најзначајније су:

- Коришћење само два суседна фрејма и бидирекције естимације која захтева мање меморијских ресурса од унидирекционе;
- Мали број кандидата вектора помераја (само осам по блоку);
- Могућност целобројне имплементације;
- Регуларни патерни за приступ меморији настали из коришћења блоковског приступа код блок мечинга и униформног стандардног региона код Прилагођене фазне планарне корелације;
- Редукован број потребних операција потврђен помоћу измерених процесорских циклуса.

Унапређења и додатни рад везан за описани алгоритам огледају се у препознавању и класификацији сцена у видео секвенцама и коришћење тих информација за контролу извршавања и перформанси БМК алгоритма [36].

Код алгоритама за естимацију покрета и повећање броја фрејмова у секунди дешава се да се статичним регионима грешком додели вектор помераја различит од нултог. Због тога, често се на посебан начин врши њихова детекција, чиме се коригују евентуалне грешке у естимацији. Детекција статичних региона описана у [37], представља још једно од могућих унапређења БМК алгоритма.

Идеје и доприноси описани у овој докторској тези су објављени као патент [38] .  
Дијаграм радова, публикација и патентних апликација које је аутор објавио везане за тему  
естимације кретања у видео секвенцама приказује Слика 5-1.



Слика 5-1: Дијаграм публикација и патентних апликациј аутора ове тезе везаних за област естимације кретања у видео секвенцама. Тамно зеленом бојом истакнуто оно што је објављено, светло зеленом су означене две патентне апликације које су објављене крајем 2015. г. и жутом публикације које су у плану у будућности.

## Додатак А: Методе за поређење сличности између слика

За евалуацију алгоритма описаног у овој тези коришћене су две методе описане у наставку.

### 5.1 Peak signal to noise ratio (PSNR)

*PSNR* представља однос између снаге максималне могуће вредности разлике сигнала и постојеће вредности сигнала. Закључује се да када је разлика између два сигнала једнака нули *PSNR* вредност ће тежити бесконачности. Математички, срачунава се на следећи начин:

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right),$$

где је *MAX* максимална вредност разлике сигнала који се пореде, а *MSE* средња квадратна грешка дефинисана као:

$$MSE = \frac{1}{m \cdot n} \cdot \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [P(i, j) - N(i, j)]^2,$$

где су *m* и *n* димензије сигнала (у нашем случају фрејмова) *P* и *N*.

### 5.2 Universal Image Quality Index (UIQI)

Формулација *UIQI* предложена у [35] заснива се на процени разлика између две слике на основу три фактора:

- Смањења корелације
- Дисторзије луминансе
- Дисторзије контраста

Сваки од поменутих фактора респективно је моделован кроз један од фактора у формули:

$$Q = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} \cdot \frac{2 \cdot \bar{x} \cdot \bar{y}}{\bar{x}^2 + \bar{y}^2} \cdot \frac{2 \cdot \sigma_x \cdot \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

што кад се среди добија се:

$$Q = \frac{4 \cdot \sigma_{xy} \cdot \bar{x} \cdot \bar{y}}{(\sigma_x^2 + \sigma_y^2) \cdot (\bar{x}^2 + \bar{y}^2)}$$

где је:

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=0}^N x_i, \bar{y} = \frac{1}{N} \cdot \sum_{i=0}^N y_i,$$

$$\sigma_x^2 = \frac{1}{N-1} \cdot \sum_{i=0}^N (x_i - \bar{x})^2, \sigma_y^2 = \frac{1}{N-1} \cdot \sum_{i=0}^N (y_i - \bar{y})^2,$$

$$\sigma_{xy} = \frac{1}{N-1} \cdot \sum_{i=0}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})^2.$$

Динамички опсег  $Q$  је  $[-1, 1]$ . Највишу вредност достиже када су улазни сигнали  $x$  и  $y$  једнаки за свако  $i = 1, \dots, N$ . Најнижу вредност добија када је:

$$y_i = 2 \cdot \bar{x} - x_i$$

Дефинисана формула примењена на цео фрејм свакако не би дала очекиване резултате. Због тога, за поређење фрејмова видео секвенце извршена је њихова подела на блокове димензија  $8 \times 8$  и за сваки блок је израчуната вредност  $Q$ . Финална вредност  $UIQI$  за цео фрејм се добија као аритметичка средина срачунатих вредности  $Q$  за сваки од блокова фрејма.

# Додатак Б: Патентне апликације на којима је базирана теза



US 20140254678A1

(19) **United States**

(12) **Patent Application Publication**  
Beric et al.

(10) **Pub. No.:** US 2014/0254678 A1  
(43) **Pub. Date:** Sep. 11, 2014

(54) **MOTION ESTIMATION USING HIERARCHICAL PHASE PLANE CORRELATION AND BLOCK MATCHING**

(22) Filed: Mar. 11, 2013

**Publication Classification**

(71) Applicants: Aleksandar Beric, Eindhoven (NL);  
Zdravko Pantic, Beograd (RS);  
Vladimir Kovacevic, Belgrade (RS);  
Radomir Jakovljevic, Belgrade (RS);  
Milos Markovic, Belgrade (RS)

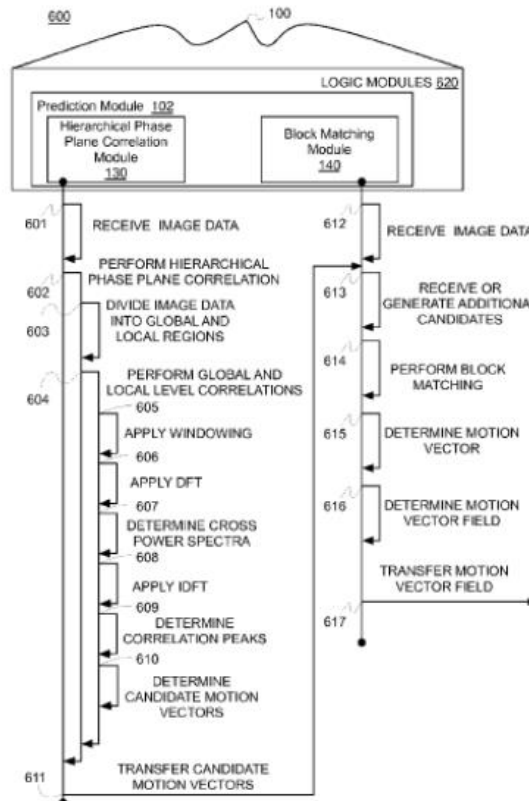
(51) **Int. Cl.** H04N 7/26 (2006.01)  
(52) **U.S. Cl.** CPC ..... H04N 19/00684 (2013.01)  
USPC ..... 375/240.14

(72) Inventors: Aleksandar Beric, Eindhoven (NL);  
Zdravko Pantic, Beograd (RS);  
Vladimir Kovacevic, Belgrade (RS);  
Radomir Jakovljevic, Belgrade (RS);  
Milos Markovic, Belgrade (RS)

(57) **ABSTRACT**

Systems, apparatus, articles, and methods are described related to motion estimation using hierarchical phase plane correlation and block matching.

(21) Appl. No.: 13/793,029



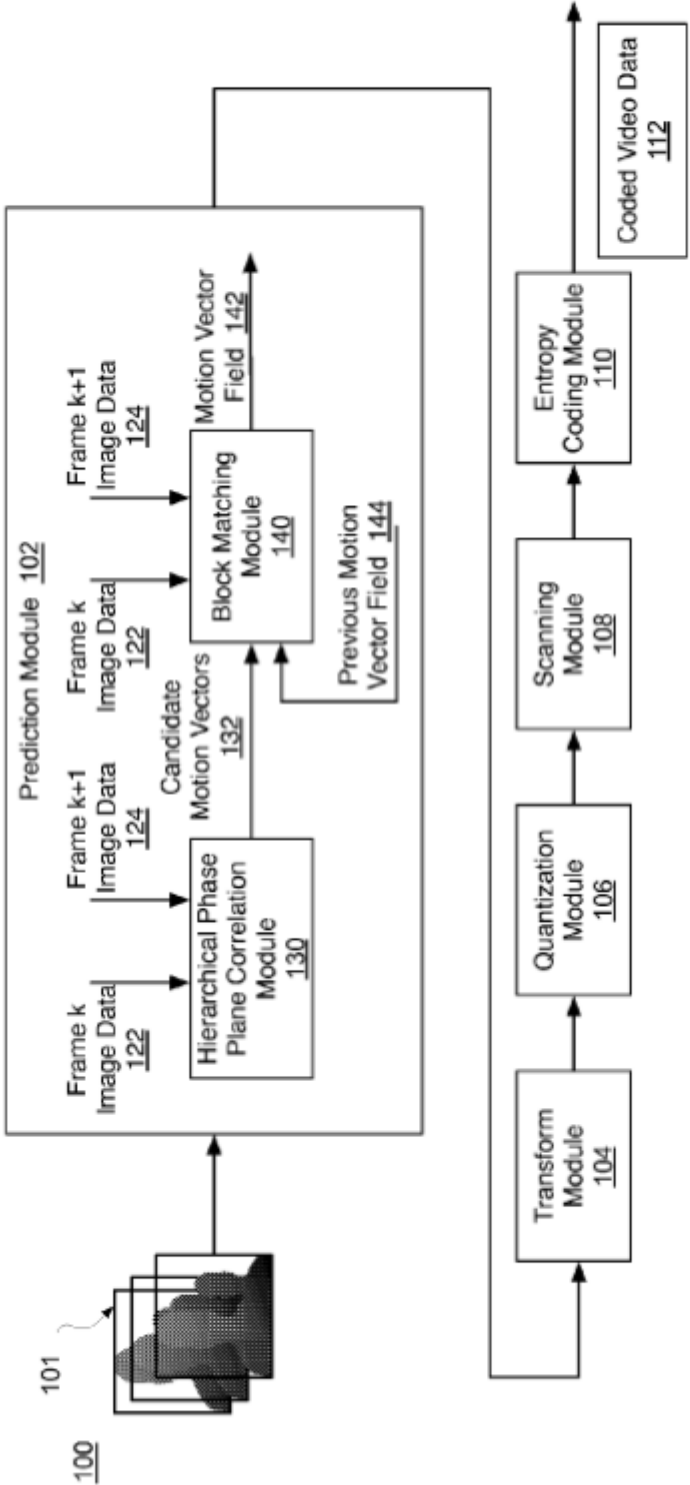


FIG. 1





US 20150294479A1

(19) **United States**

(12) **Patent Application Publication**  
**KOVACEVIC et al.**

(10) **Pub. No.: US 2015/0294479 A1**  
(43) **Pub. Date: Oct. 15, 2015**

(54) **FALLBACK DETECTION IN MOTION ESTIMATION**

*G06K 9/62* (2006.01)  
*G06K 9/46* (2006.01)

(71) Applicant: **Intel Corporation**, Santa Clara, CA (US)

(52) **U.S. CL.**  
CPC ... *G06T 7/20* (2013.01); *G06K 9/46* (2013.01);  
*G06K 9/52* (2013.01); *G06K 9/6215* (2013.01);  
*G06K 2009/4666* (2013.01)

(72) Inventors: **VLADIMIR KOVACEVIC**, Pancevo (RS); **ZDRAVKO PANTIC**, Belgrade (RS); **ALEKSANDAR BERIC**, Eindhoven (NL); **MILOS MARKOVIC**, Beograd (RS); **VLADIMIR ILIC**, Novi Beograd (RS)

(57) **ABSTRACT**

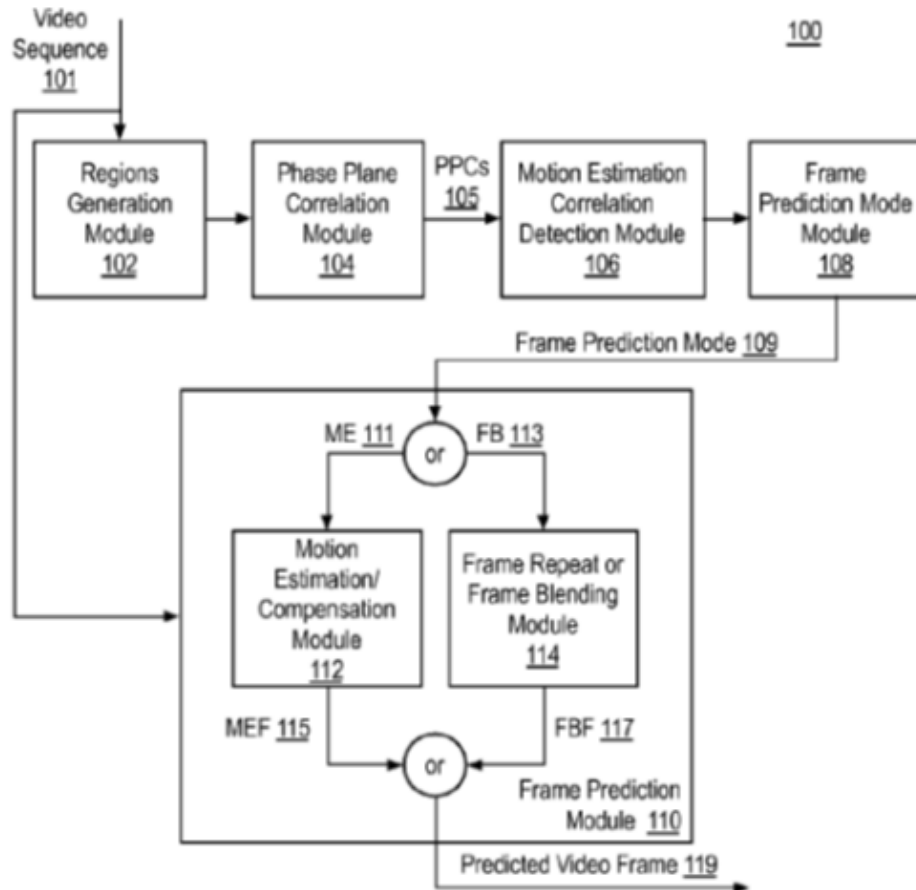
(21) Appl. No.: **14/252,944**

(22) Filed: **Apr. 15, 2014**

**Publication Classification**

(51) **Int. CL.**  
*G06T 7/20* (2006.01)  
*G06K 9/52* (2006.01)

Techniques related to managing the use of motion estimation in video processing are discussed. Such techniques may include determining dividing two video frames each into corresponding regions, generating phase plane correlations for the corresponding regions, determining whether the video frames are motion estimation correlated based on the phase plane correlations, and providing a video frame prediction mode indicator based on the determination.



## Литература

- [1] Cordes, C.N., de Haan, G. (2009). Key Requirements for High Quality Picture-Rate Conversion, SID Symposium Digest of Technical Papers, 41(1), 850-853.
- [2] Wang, D., Zhang, L., Vincent. A. (2010). A Motion-Compensated Frame Rate Up-Conversion - Part I: Fast Multi-Frame Motion Estimation, Broadcasting, IEEE Transactions on, 56(2), 133 – 141.
- [3] Wang, D., Vincent, A., Blanchfield P., Klepko, R. (2010). Motion-Compensated Frame Rate Up-Conversion - Part II: New Algorithms for Frame Interpolation, Broadcasting IEEE Transactions, 56(2), 142-149.
- [4] Quang, T. V., Kim, Y., Hong, S. (2009). Frame rate up-conversion using forward-backward jointing motion estimation and spatio-temporal motion vector smoothing, Computer Engineering & Systems, ICCES 2009, International Conference on, Cairo, 605-609.
- [5] Sorwar, G., Murshed, M., Dooley, L. (2004). Fast Block-Based True Motion Estimation Using Distance Dependent Thresholds, Journal of Research and Practice in Information Technology, 36(3) 157-169.
- [6] Li, R., Zeng, B., Liou, M.L. (1994). A new three-step search algorithm for block motion estimation, Circuits and Systems for Video Technology, IEEE Transactions 1994, 4(4), 438-442.
- [7] Jain, J. R., Jain, A. K. (1981). Displacement measurement and its application in interframe image coding, IEEE Trans. on Communications, Vol. COM-29, No. 12, pp. 1799-1808.
- [8] Zhu, S., Ma, K.K. (2000). A new diamond search algorithm for fast block-matching motion estimation, Image Processing, IEEE Transactions on, 9(2), 287-290.
- [9] De Haan, G., Biezen, P.W.A.C., Huijgen, H., Ojo, O.A. (1993). True Motion Estimation with 3-D Recursive Search Block-Matching, IEEE Transactions on Circuits and Systems for Video Technology, 3(5), 368-388.
- [10] De Haan, G., Biezen, P.W.A.C. (1994). Sub-pixel motion estimation with 3-D recursive search block-matching, Signal Processing: Image Communication, 6(3), 229-239.
- [11] Tourapis, A.M., Au, O.C., Liou, M.L. (2001). Predictive motion vector field adaptive search technique (pmvfast) - enhancing block based motion estimation, SPIE Proceedings Vol. 4310, Visual Communications and Image Processing, 883-892.
- [12] Nie, Y., Ma, K.K. (2002). Adaptive Rood Pattern Search for Fast Block-Matching Motion Estimation, Image Processing, IEEE Transactions, 11(12), 1442-1449.
- [13] Tedmori, S., N. Al-Najdawi (2012). Hierarchical stochastic fast search motion estimation algorithm, IET Computer Vision, 6(1), 21-28.
- [14] Zhang, P., Wei, P., Yu, H.-Y. (2012). Biogeography-based optimisation search algorithm for block matching motion estimation, IET Image Processing, 6(7), 1014-1023.
- [15] Research department report, the British Broadcasting Corporation (1990). Real-time motion measurement hardware: Phase correlation unit, 1-19.

- [16] Kumar, S., Biswas, M., Nguyen, T.Q. (2004) Efficient phase correlation motion estimation using approximate normalization, *Signals, Systems and Computers, Conference Record of the Thirty-Eighth Asilomar Conference on, Pacific Grove, CA, USA*, 1727–1730.
- [17] Biswas, M., Nguyen, T. Q. (2002). A Novel Motion Estimation Algorithm Using Phase Plane Correlation for Frame Rate Conversion, *Signals, Systems and Computers, 2002. Conference Record of the Thirty-Sixth Asilomar Conference on, Pacific Grove, CA, USA*, 492–496.
- [18] David J. Fleet, Yair Weiss (2006). Optical Flow Estimation. In Paragios et al. Handbook of Mathematical Models in Computer Vision. Springer. ISBN 0-387-26371-3.
- [19] Lucas, B. D., Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision, *Proceedings of Imaging Understanding Workshop*, pp. 121-130.
- [20] Huang, A. M., Nguyen, T. (2009). Correlation-Based Motion Vector Processing With Adaptive Interpolation Scheme for Motion-Compensated Frame Interpolation, *Image Processing, IEEE Transactions on*, 18(4) 740-752.
- [21] Wang, D. Zhang, L., Vincent. A. (2010). A Motion-Compensated Frame Rate Up-Conversion - Part I: Fast Multi-Frame Motion Estimation, *Broadcasting, IEEE Transactions on*, 56(2), 133 – 141.
- [22] Ho, H. Klepko, R., Ninh, N. Wang, D. (2011). A high performance hardware architecture for multi-frame hierarchical motion estimation, *Consumer Electronics, IEEE Transactions on*, 57(2) 794-801.
- [23] Won Hee Lee, Kyuha Choi, Jong Beom Ra, (2014). Frame Rate Up Conversion Based on Variational Image Fusion, *Image Processing, IEEE Transactions on*, 23(1), 399–412.
- [24] T. Brox, A. Bruhn, N. Papenbergh, J. Weickert, (2004). High accuracy optical flow estimation based on a theory for warping, in *Proc. 8th Eur. Conf. Comput. Vis.*, vol. 4. 25–36.
- [25] ISO/IEC JTC1/SC29. (1992). Coding of moving pictures and associated audio for digital storage media up to about 1.5Mbit/s. ISO/IEC 11172-2.
- [26] ISO/IEC JTC1/SC29. (1994). Generic coding of moving pictures and associated audio. ISO/IEC 13818-2.
- [27] [ISO/IEC JTC1/SC29. Coding of audio-visual objects. ISO/IEC 14496-2, January 2000.
- [28] Joint Video Team of ITU-T and ISO/IEC JTC1. (2003). Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification. ISO/IEC 14496-10 AVC.
- [29] Vladimir B. Kovacevic, Zdravko Pantic, Aleksandar Beric, Radomir Jakovljevic (2015). Block-matching correlation motion estimation for frame-rate up-conversion, *Journal of Signal Processing Systems*, 1-10.
- [30] Ndili, O., Ogunfunmi T. (2014). Fast Algorithm and Efficient Architecture for Integer and Fractional Motion Estimation, *Journal of Signal Processing Systems*, 75(1), 55-64.
- [31] Pinto, C. A., Beric, A. (2006). HiveFlex-Video VSP1 : Video Signal Processing Architecture for Video Coding and Post-Processing, *ISM'06. Eighth IEEE International Symposium on, San Diego, CA, USA*, 493-500.
- [32] Beric A., de Haan G., Sethuraman R., van Meerbergen J. (2005). An Efficient Picture-Rate Up-Converter, *Journal of VLSI signal processing systems for signal, image and video technology*, 41(1), 49-63.
- [33] Ojo, O.A., de Haan, G. (1997). Robust motion-compensated video upconversion', *Consumer Electronics, IEEE Transactions on*, 43(4), 1045-1056.

- [34] Test video sequences and frame rate up-conversion results, <http://1drv.ms/1ITCEp7>, accessed July 2015
- [35] Wang, Z. Bovik, A.C. (2002). A universal image quality index, *Signal Processing Letters, IEEE*, 9(3), 81-84.
- [36] Vladimir Kovacevic, Zdravko Pantic, Aleksandar Beric, Milos Markovic, Vladimir Ilic (2015). Smart fallback detection for motion estimation algorithms, 25. July 2013, US Patent Application Number: US2014/252994, <http://www.freshpatents.com/-dt20151015ptan20150294479.php>.
- [37] Vladimir Kovacevic, Zdravko Pantic, Aleksandar Beric et al. (2015). Static region detection in video sequences, 27. May 2014, US Patent Application Number: US2014/287967.
- [38] Aleksandar Beric, Zdravko Pantic, Vladimir Kovacevic (2014). Motion estimation using hierarchical Phase Plane Correlation and Block Matching, US Patent Application Number: 2014/0254678 A1, September 2014.

## Биографија аутора

Владимир Б. Ковачевић је рођен 1. 11. 1984. године у Вуковару. Школовање је започео у Ковину где се преселио 1991. године. По досељавању у Панчево и завршетку Основне школе “Васа Живковић” (где је био ђак генерације), уписао је Гимназију Урош Предић. У току школовања, све четири године је био учесник Републичког такмичења из математике, што је непосредно утицало и на избор његове будуће професије. Гимназију је завршио 2003. године као добитник Вукове дипломе. Исте године је уписао Електротехнички факултет у Београду, на одсеку за Рачунарску технику. Основне студије је завршио октобра 2007. године са просечном оценом у току студија 8.34, а одбравивши завршни рад са оценом 10 код професора Зорана Јовановића.

Мастер студије је завршио на Одсеку за електронику 2009. године просечном оценом 10. Мастер рад под називом "Дигитални филтри у систему за конверзију видео сигнала" одбранио је код професора Лазара Сарановца. У новембру 2010. године је уписао докторске студије на Одсеку за Телекомуникације.

Професионалну каријеру Владимир Ковачевић је започео у октобру 2007, у новооснованој компанији Омегатек. Од априла 2009. године до октобра 2010. године радио је на Институту Михајло Пупин на развоју везаном за бежичне сензорске мреже. Новембра 2010. године запослио се у компанији *Silicon Hive*, коју је четири месеца касније купила *Intel* корпорација. Од аквизиције (припајања) фирми, па до јула 2015. године Владимир је радио на развоју алгоритама за процесирање дигиталних слика и видео секвенци прилагођених извршавању на Интеловим мобилним платформама.

Као посебне целине на којима је Владимир радио могу се истаћи:

- Дизајн и имплементација комуникационог чвора за бежичне сензорске мреже
- *Frame-rate upconversion* за Интелову Merrifield платформу,
- Контрола HDR алгорита за видео (*Video HDR for Intel Skylake*)

Прилог 1.

### Изјава о ауторству

Потписани-а ВЛАДИМИР Б. КОВАЧЕВИЋ  
број уписа 5049 / 2010

#### Изјављујем

да је докторска дисертација под насловом

ЕСТИМАЦИЈА ПОКРЕТА У ВИДЕО СЕКВЕНЦАМА ПРИЛАГОЂЕНА  
АЛГОРИТМИМА ЗА ПОВЕЋАЊЕ БРОЈА СЛИКА У СЕКУНДИ

- 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/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 25. 06. 2015. г.

Владимир Ковачевић

Прилог 2.

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

Име и презиме аутора ВЛАДИМИР КОВАЧЕВИЋ

Број уписа 5049/2010

Студијски програм ТЕЛЕКОМУНИКАЦИЈЕ

Наслов рада ЕСТИМАЦИЈА ПОКРЕТА У ВИДЕО СЕКВЕНЦАМА ПРИЛАГОЂЕНА  
АЛГОРИТМИМА ЗА ПОВЕЋАЊЕ БРОЈА СЛИКА У СЕКУНДИ

Ментор ПРОФ. ДР. ИРИНА РЕЉИЋ

Потписани ВЛАДИМИР Б. КОВАЧЕВИЋ

изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу 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 одбране 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, 25. 06. 2015. г.

*Владимир Ковачевић*

Прилог 3.

### Изјава о коришћењу

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

ЕСТИМАЦИЈА ПОКРЕТА У ВИДЕО СЕКВЕНЦАМА  
ПРИЛОЖЕНА АЛГОРИТМИМА ЗА ПОВЕЋАЊЕ БРОЈА СЛИКА У СЕКUNДИ

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла 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/ла.

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, 25. 06. 2015. г.

*Владимир Ковачић*



1. Ауторство - Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. Ауторство – некомерцијално. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. Ауторство - некомерцијално – без прераде. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. Ауторство - некомерцијално – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. Ауторство – без прераде. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. Ауторство - делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.