

(19) REPUBLIKA SRBIJA

(12) Patentni spis

(11) 53604 B1



ZAVOD ZA  
INTELEKTUALNU SVOJINU  
BEOGRAD

(51) Int. Cl.  
H 04 N 17/00 (2006.01)  
H 04 N 7/00 (2011.01)

(21) Broj prijave: P-2011/0251  
(22) Datum podnošenja prijave: 15.06.2011.  
(43) Datum objavljivanja prijave: 29.02.2012.  
(45) Datum objavljivanja patenta: 27.02.2015.

(73) Nosilac patenta:  
RT-RK D.O.O.,  
Fruškogorska 11, 21000 Novi Sad, RS

(72) Pronalazači:  
KUKOLJ, Dragan, dr.; ZLOKOLICA, Vladimir, dr.;  
PEKOVIĆ, Vukota; POKRIĆ, Maja, dr.

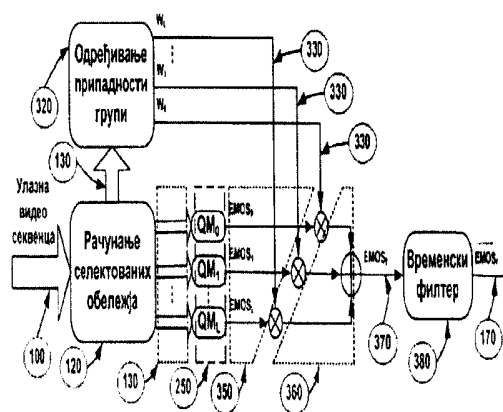
(54) Naziv: POSTUPAK ZA AUTOMATSKU  
OCENU VIDEO KVALITETA MULTI-  
MODELIMA

(51) Int. Cl.  
H 04 N 17/00 (2006.01)  
H 04 N 7/00 (2011.01)

(57) Apstrakt:

Postupak za automatsku ocenu video kvaliteta multi-modelima (250) ima za novost ocenjivanje kvaliteta video materijala zasnovano na percepcijskom pristupu primenom multi-modela (250). Svaki model iz multi-modela (250) se konstruiše i optimizuje nad odgovarajućim klasterom (241) odbiraka izračunatih objektivnih uzoraka (130) obeležja video frejmova u formiranom deskriptivnom prostoru (240) objektivnih uzoraka (130) obeležja. Pored nezavisnih parcijalnih modela ocene kvaliteta definisanih nad specifičnim klasterima (241) odbiraka u deskriptivnom prostoru (240) obeležja, pronalazak se odnosi i na uključivanje u multi-model (250) integralnog modela kvaliteta koji se konstruiše korišćenjem celokupnog skupa odbiraka u deskriptivnom prostoru (240) obeležja. Konačna, izlazna ocena (190) kvaliteta video materijala predstavlja ponderisana suma ocena pojedinačnih modela (360) koji čine multi-model (250). Uticaj pojedinih modela (360) određuje pripadnost vektora uzoraka (130) obeležja video frejmova odgovarajućim klasterima (241) odbiraka

deskriptivnog prostora (240) objektivnih uzoraka (130) obeležja.



RS 53604 B1

### **Области технике на коју се проналазак односи**

Проналазак припада области видео и мултимедијалних технологија. Проналазак се посебно односи на системе и поступке за естимацију квалитета видео сигнала без присуства оригиналног сигнала.

Ознака према међународној класификацији патената (МКП) је: **H04N**.

### **Технички проблем**

Мултимедијални садржаји се користе на широком спектру уређаја (на пример, TV, PC, мобилни телефони) са различитим екранима и резолуцијама. При томе се мултимедијални садржаји преносе путем разноврсних комуникационих мрежа и кодују са различитим стандардима (на пр. MPEG2, MPEG4, H264). Утицаји преносних путева и кодовања су често видљиви људском оку и утичу на људски доживљај квалитета видео садржаја.

Из наведених разлога је очигледно да постоји потреба за прецизном естимацијом и управљањем видео квалитетом у различитим мултимедијалним апликацијама, као што су интернет ТВ, видео на захтев и слично. Методе оцене квалитета видео материјала се могу класификовати у три основне групе: (а) субјективна оцена квалитета, (б) објективна оцена квалитета, и (в) перцепцијски заснована оцена видео квалитета. Субјективна оцена квалитета (позната под енг. називом Mean Opinion Score - MOS) тачно описује људску видео перцепцију, али је врло неефикасна због потребе за великим људским ресурсима приликом генерисања MOS оцена. Објективна оцена квалитета има за циљ мерење специфичних типова деградација на основу издвајања израчунатих обележја из видео материјала. Међутим, ове објективне мере често недовољно корелишу са субјективним оценама квалитета. Перцепцијски заснована оцена видео квалитета представља најприроднији поступак мерења видео квалитета који не захтева непосредно учешће човека приликом естимације. Побољшавање тачности и поузданости перцепцијског оцењивања видео квалитета је од суштинског значаја.

### **Стање технике**

Данас постоји много различитих решења перцепцијске оцене видео квалитета без познавања оригиналног видео материјала. Неколико познатих и заштићених решења је наведено у даљем тексту. Патент US 6,285,797 објављен 4. маја 2001. под називом, "Method and apparatus for estimating digital video quality without using a reference video", описује уређај за естимацију квалитета дигиталног видеа без познавања референтног сигнала, који се састоји од видео

предпроцесора, генератора сигнала виртуалног референтног сигнала, генератора естимације дисторизија и генератора метрике квалитета улазне видео секвенце, затим патентна пријава US 2007263897 објављена 15. новембра 2007. под називом, "Image and Video Quality Measurement", описује систем за оцену квалитета слике који у том циљу комбинује више мера као што су оштрина слике, разноврсност боја и изостанак блок-ефеката, затим патентна пријава WO/2010/103112 објављена 16. септембра 2010. под називом, "Method and Apparatus for Video Quality Measurement without Reference", која описује систем за оцену квалитета видеа без познавања оригиналног сигнала, која при естимацији квалитета комбинује више мера као што су информације издвојене из видео тока (квантизациони корак, битска брзина), нивои шума, блок-ефекти, замућеност слике итд., као и патент US 7,663,636 објављен 16. фебруара 2010. под називом, "Apparatus and method for automatically analyzing digital video quality", који описује уређај за аутоматску анализу дигиталног видеа, који се састоји од јединице за анализу видеа и јединице за приказивање, где јединице за анализу квалитета видеа анализира атрибуте текстуре и боје у случајевима када постоји и када не постоји референтни видео сигнал.

#### **Излагање суштине проналаска**

Приказано решење нуди нову перцепциски засновану методологију за аутоматску естимацију видео квалитета без референце које је робустно у односу на различит садржај видео сигнала. То се постиже декомпозицијом проблема естимације квалитета на више модела квалитета који се могу независно оптимизовати у односу на специфичне одлике слике-видеа, а потом заједно комбиновати на одређен начин ради добијање јединствене перцепцијски блиске оцене квалитета.

Представљени проналазак је базиран на мулти-модел приступу код кога се сваки модел независно конструише и оптимизује са одговарајућим групама узорака формираним у простору селектованих објективних обележја – објективној метрици видео сигнала. Групе (кластери) обележја се одређују у простору обележја које се састоји од редукованог и селектованог скупа обележја према њиховој релевантности у односу на субјективну перцепцију видео сигнала. Поред независних парцијалних модела квалитета дефинисаних над специфичним кластерима у селектованом простору обележја, представљени проналазак се односи и на укључивање у мулти-модел перцепцијске оцене квалитета тзв. интегралног модела квалитета који се конструише коришћењем целокупног скупа одбирака у простору обележја. Ово представља значајну новину у односу на претходна решења и значајно поправља тачност естимације квалитета видео сигнала. Представљени проналазак заснован на приступу са мулти-моделом је верификован коришћењем линеарних модела конструисаних помоћу методе линеарне регресије. Поред тога, мулти-модел

приступ је верификован и са нелинеарним моделима оцене квалитета конструисаним применом вештачких неуронских мрежа.

#### **Кратак опис слика проналаска**

Проналазак је детаљно описан помоћу нацрта датих у следећим сликама.

**Слика 1:** Представља општи приказ перцепцијског приступа у оцењивању видео квалитета

**Слика 2:** Представља детаљан приказ фазе пројектовања мулти-модалног приступа естимације квалитета видео сигнала

**Слика 3:** Приказује основну схему рада експлоатационе фазе мулти-модалног приступа за естимацију квалитета непознатог видео сигнала за који оригинални материјал није познат

#### **Детаљан опис проналаска**

Перцепцијски засновани приступ естимацији видео квалитета, којем припада и описани проналазак, састоји се увек из две основне фазе: (1) *фазе пројектовања* у којој се конструишу специфични модели видео квалитета на бази субјективне (већ познате) оцене квалитета, и (2) *експлоатационе фазе* намењена за аутоматску естимацију квалитета видео сигнала без коришћења субјективне оцене квалитета. Експлоатациона фаза поступка се може одвијати и у реалном времену.

Генерални опис пројектовања перцепцијски заснованих приступа оцене квалитета видео материјала је приказан на Слици 1. Слика показује како се одређује модел за предикцију квалитета видео сигнала на бази одабраних објективних обележја видео материјала и субјективне оцене квалитета. На приказаној Слици 1 улазна видео секвенца 100 служи за обучавање (тренирање) модела квалитета видео сигнала. За дате улазне видео секвенце 100 врши се њихова процена квалитета у блоку 140 за субјективно оцењивање квалитета. Са друге стране, унутар блока 120 за рачунање објективних обележја, прорачунавају се узорци 130 обележја за дате улазне видео секвенце 100. Као излаз из блока 120 за рачунање објективних обележја добијају се узорци 130 обележја (одбирци у форми вектора) прорачунатих објективних мерења 110. Коначно на основу узорака 130 обележја објективних мерења 110 и субјективних оцена за тренинг секвенце врши се моделирање видео квалитета унутар блока 160 за генерисање естимационог модела (енг. Video Quality Assessment model - VQA). Као излаз из блока 160 за генерисање естимационог модела добија се естимирана оцена 170 квалитета видео секвенце која се пореди са познатом субјективном оценом 150 на основу чега се унутар блока 180 за обуку модела квалитета видео сигнала помоћу кога се одређује корекција естимиране, излазне оцене 190 квалитета и прослеђује назад ради оптимизације модела квалитета помоћу блока 160 за

генерисање естимационог модела. Овај поступак се итеративно понавља док се не постигне минимална грешка естимације за дате субјективне оцене, рачуната објективна обележја и предпостављени модел квалитета који се генерише.

Детаљнији приказ фазе пројектовања је дат на Слици 2, где су приказане главне етапе генерисања компоненти овог проналаска са мулти-модел заснованом естимацијом квалитета видео сигнала: (1) Селекција објективних обележја слике у односу на субјективну оцену квалитета видео сигнала; (2) Груписање узорака (вектора) обележја у кластере у селектованом простору обележја; (3) Генерисање мулти-модела квалитета у специфичним кластерима обележја, коришћењем одговарајућих субјективних оцена квалитета. Као што се може видети са Сlike 2 иницијални скуп обележја, који може имати произвољно велик скуп објективних мерења 110 се прво корелише са субјективним оценама 150 квалитета унутар блока 220 за редукцију/селекцију узорака 130 обележја у циљу смањења броја обележја које ће се даље користити у поступку пројектовања естимационих модела. Резултат блока 220 за редукцију/селекцију је редуктовани подскуп обележја састављен од селектованих обележја која су најрелевантнија у односу на субјективну оцену квалитета. Тачније, сва разматрана иницијална обележја се израчунавају и пореде са субјективним оценама унутар блока 220 за редукцију/селекцију, на бази сваког видео фрејма, да би се одредило  $K$  изабраних обележја 230, коришћењем неке схема селекције обележја са елементима комбинаторне оптимизације, као на пример генетских алгоритама или селекције обележја унапред (енг. *wrapping-based feature forward selection scheme*). Сви одбирци расположиви за обуку (тренирање) састављени од изабраних обележја 230 се прво искористе у блоку 260 за пројектовање интегралног модела квалитета, а потом се  $K$ -димезионални одбирци групишу у  $L$  карактеристичних група-кластера у дескриптивном простору 240 обележја. У зависности од удаљености одбирка од центра кластера, сваки  $K$ -димензионални одбирка се додељује једном кластеру података са минималним Еуклидијановим растојањем и користи унутар блока 260 за пројектовање интегралног модела квалитета дефинисаног преко података из кластера 241 одбирака обележја. Сваком од  $L$  генерисаних кластера 241 одбирака у простору 240 обележја се додељује адекватан естимациони модел квалитета. Потом се врши паралелно тренирање  $L$  различитих модела квалитета и у блоку 260 за пројектовање интегралног модела квалитета сви модели квалитета, при чему сваки одговара специчном видео садржају и оштећењима у видео сигналу. Естимирана вредност 261 квалитета и стварна субјективна оцена 262 квалитета су везане за блок 260 за пројектовање интегралног модела квалитета. На основу та два сигнала се у блоку 260 за пројектовање интегралног модела квалитета врши итеративно моделовање на већ објашњен начин уз Сliku 1. Потпуно идентичан поступак се примењује и на сигналима блокова модела квалитета за пројектовање интегралног модела квалитета, који се

користе при генерисању појединих интегралних модела квалитета ( $QM_1, QM_2, \dots, QM_L$ ) који чине мулти-модел 250 квалитета видео материјала.

Слика 3 приказује основну схему рада *фазе експлоатације* представљеног проналаска за мулти-модел засноване естимације квалитета непознатог видео сигнала за који субјективна оцена квалитета није позната. Како је селекција обележја већ урађена у фази пројектовања, скуп обележја је већ познат и у почетку корака ове фазе једино је потребно израчунати одбирке за селектована објективна обележја помоћу блока 120 за рачунање објективних обележја, и то за сваки фрејм улазне видео секвенце 100 који се доводи на улаз система за оцену квалитета. Потом се на нивоу сваког фрејма одређује ниво припадности одговарајућег вектора обележја сваком од пројектованих и претходно дефинисаних кластера помоћу блока 320 за одређивање припадности групи. Вектор узорака 130 обележја се допрема сваком моделу квалитета ( $QM_1, QM_2, \dots, QM_L$ ). На бази овога се укупна естимација квалитета за сваки видео фрејм одређује преко тежински усредњене суме 370:

$$EMOS_f = \sum_{l=0}^L w_l(f) EMOS_l(f) \quad (1)$$

где  $EMOS_l(f)$  сигнал 350 представља естимовани MOS од сваког  $l$ -тог модела квалитета и  $L$  је укупан број укључених модела квалитета, а  $w_l$  представља сигнале 330 тежинских фактора. Овај тежински фактор је инверзно пропорционалан Еуклидијановом растојању узорака 130 обележја објективних мерења 110 текућег видео фрејма од центра од  $l$ -тог кластера, при чему је укупна сума вредности свих тежинских фактора једнака 1. Тачније, тежински фактор у описаном проналаску се одређује на следећи начин:

$$w_l = \frac{1}{d_l^3} \quad (2)$$

где  $d_l$  означава Еуклидијаново растојање  $K$ -димензионалног одбирка од  $l$ -тог центра кластера. Назначавамо да се у представљеном проналаску, интегрални модел означава као нулти модел и кластер центар се означава као 0-ти.

Коначно, финална субјективна оцена 150 квалитета текућег видео сигнала, односно естимирана оцена 170 се одређује као средња вредност естимираних оцена 170 квалитета последњих  $F$  фрејмова видео сигнала помоћу временског филтра 380:

$$EMOS_V = \frac{1}{F} \sum_{f=1}^F EMOS_f \quad (3)$$

Израз (3) може бити реализован на више начина, као филтар средњих вредности, рекурзивни временски филтар, итд.

Поступци и технике обраде видео сигнала из овог проналаска се могу имплементирати у различитим системима за естимацију квалитета видео сигнала када није расположив ни оригинални видео сигнал, ни субјективна оцена квалитета. Имплементација може бити реализована у хардверу (ASIC, FPGA), софтверу или комбиновано (DSP, микропроцесори) тако да омогућава извршавање описаних техника из овог проналаска.

Детаљи отривени у опису овог проналаска омогућавају било ком стручњаку у овој области да генеричке принципе овог проналаска може имплементирати у другим системима за перцепцијску естимацију квалитета видео материјала, чиме се не излази из оквира овог проналаска.

## Патентни захтеви

1. Поступак за аутоматску оцену видео квалитета мулти-моделима аутоматско који подражава људски субјективни доживљај видео квалитета, састоји се од прорачуна карактеристичних објективних узорака (130) обележја квалитета за сваки фрејм видео сигнала, и генератора вектора наведених узорака (130) обележја за дефинисање дескриптивног простора (240) обележја за карактеризацију садржаја и деградације квалитета видео сигнала и конфигурисаног мулти-модела (250) за естимацију оцена субјективног квалитета видео сигнала, **карактерисан тиме** да се састоји од следећих корака:

да сваки појединачни модел (360) наведеног мулти-модела (250) линеарно трансформише вредности наведеног вектора узорака (130) обележја у одговарајуће естимиране оцене субјективног доживљаја квалитета видео сигнала,

где је сваки појединачни модел наведеног мулти-модела (250) карактерисан меморисаним параметрима: прототип-вектором кластера у дескриптивном простору (240) обележја, и параметрима линеарног модела трансформације,

израчунава се значај појединачног модела у форми коефицијента пондерисања као функција растојања вектора обележја текућег видео фрејма од наведеног одговарајућег прототип-вектора сваког модела из мулти-модела (250),

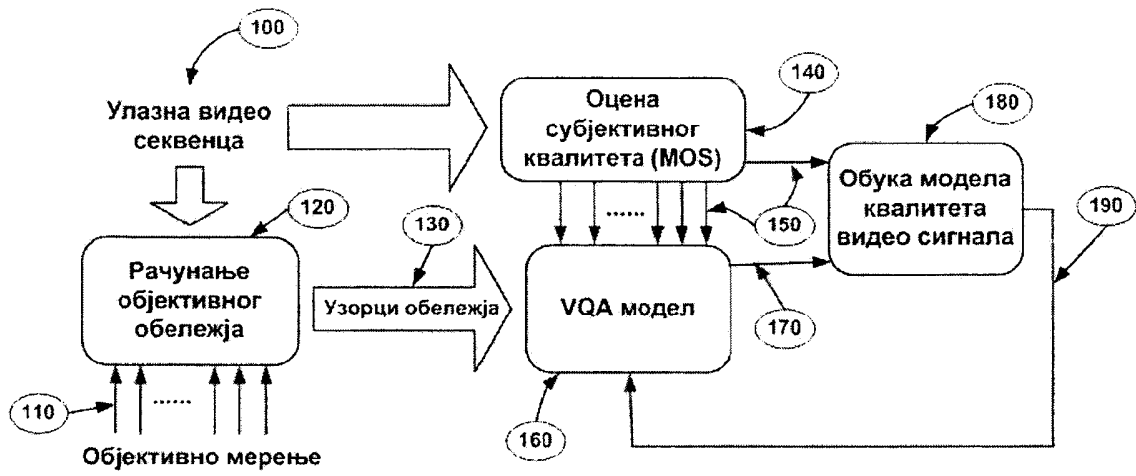
естимира се укупна субјективна оцена (150) квалитета видео фрејма као пондерисана сума оцена појединачних модела (360), и

генерише се финална естимирана оцена (170) квалитета видео сигнала применом на низ естимација за видео фрејмове временског филтра (380) са клизним временским прозором.

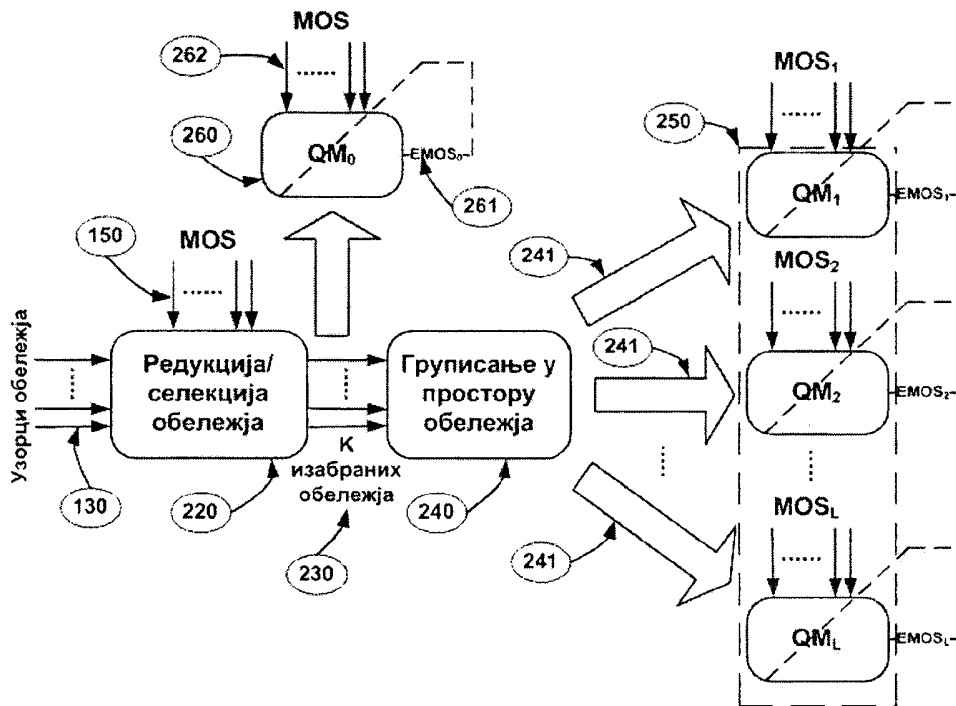
2. Поступак дефинисан према захтеву **1**, где су селекција карактеристичних узорака (130) обележја видео сигнала, конфигурација и параметри мулти-модела (250) претходно формирану помоћу улазних видео секвенци (100) расположивих за обуку.
3. Поступак дефинисан према захтеву **1** и **2**, где су карактеристични узорци (130) обележја видео сигнала одабрани у циљу мерења вредности нивоа различитих појава, попут контраста, блок-ефеката, замућености, зашумљености или покрета.
4. Поступак дефинисан према захтеву **1** и **2**, где је простор обележја састављен од кластера вектора обележја у дескриптивном простору (240) обележја.
5. Поступак дефинисан према захтеву **1**, **2** и **4**, где је сваком кластеру простора обележја додељен један одговарајући модел из конфигурације наведеног мулти-модела (250).



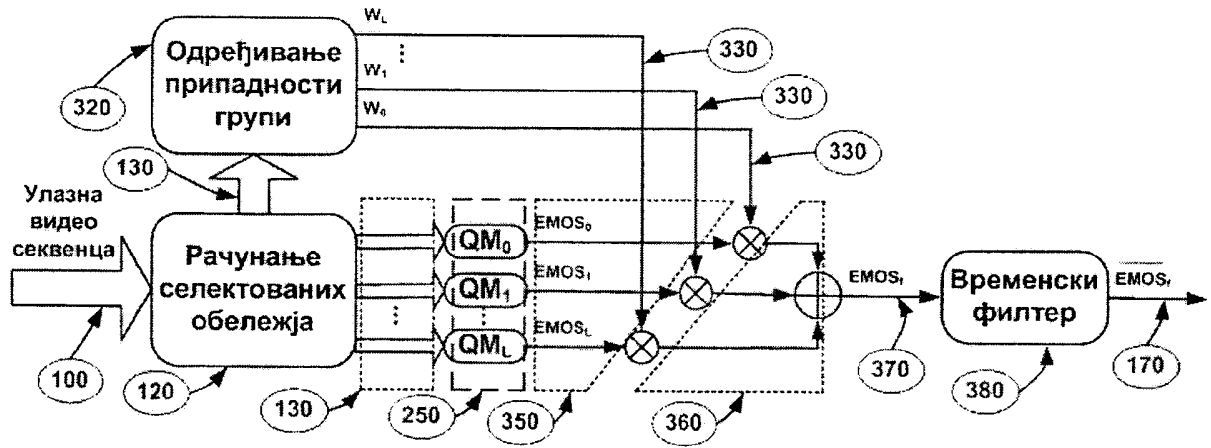
6. Поступак дефинисан према захтеву **1** и **4**, где прототип-вектор једног модела представља центар одговарајућег кластера у дескриптивном простору (240) обележја.
7. Поступак дефинисан према захтеву **1** и **5**, где линеарни математички модел трансформације има форму и нелинеарне зависности између узорака (130) обележја видео сигнала и субјективних оцена (150) квалитета.
8. Поступак дефинисан према захтеву **1**, реализован у форми рачунарског програма ускладиштеног у читљивом меморијском медијуму који се извршава на додељеном процесору.
9. Поступак за аутоматско оцењивање квалитета дигиталног видео сигнала без присуства референтног видео сигнала који подражава људски субјективни доживљај видео квалитета, дефинисан према свим наведеним захтевима од **1** до **8**, где је укључен и интегрални модел естимације квалитета конструисаним над целокупном дескриптивном простору (240) обележја, и где наведени интегрални модел естимације квалитета, представља члан у пондерисаној суми оцена појединачних модела (360) при естимацији укупне субјективне оцене квалитета видео фрејма.



Слика 1.



Слика 2.



Слика 3.