Pregled algoritama za povećanje vremenske rezolucije videosignala

Žagar, Filip

Undergraduate thesis / Završni rad

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:200:743929

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Repository / Repozitorij:

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek

zir.nsk.hr

Image not found or type unknown

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I INFORMACIJSKIH TEHNOLOGIJA

Sveučilišni studij

PREGLED ALGORITAMA ZA POVEĆANJE VREMENSKE REZOLUCIJE VIDEOSIGNALA

Završni rad

Filip Žagar

Osijek, 2017.

FERIT
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I INFORMACIJSKIH TEHNOLOGIJA OSIJEK

Obrazac Z1P - Obrazac za ocjenu završnog rada na preddiplomskom sveučilišnom studiju

Osijek, 08.09.2017.

Odboru za završne i diplomske ispite

Prijedlog ocjene završnog rada

Ime i prezime studenta:	Filip Žagar
Studij, smjer:	Preddiplomski sveučilišni studij Elektrotehnika
Mat. br. studenta, godina upisa:	3987, 18.07.2014.
OIB studenta:	33452985705
Mentor:	Doc.dr.sc. Mario Vranješ
Sumentor:	Dr.sc. Denis Vranješ
Sumentor iz tvrtke:	
Naslov završnog rada:	Pregled algoritama za povećanje vremenske rezolucije videosignala
Znanstvena grana rada:	**Telekomunikacije i informatika (zn. polje elektrotehnika)**
Predložena ocjena završnog rada:	Vrlo dobar (4)
Kratko obrazloženje ocjene prema Kriterijima za ocjenjivanje završnih i diplomskih radova:	Primjena znanja stečenih na fakultetu: 2 bod/boda Postignuti rezultati u odnosu na složenost zadatka: 3 bod/boda Jasnoća pismenog izražavanja: 2 bod/boda Razina samostalnosti: 2 razina
Datum prijedloga ocjene mentora:	08.09.2017.
Datum potvrde ocjene Odbora:	11.09.2017.

Potpis mentora za predaju konačne verzije rada u Studentsku službu pri završetku studija:	Potpis:
	Datum:

FERIT
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA **OSIJEK**

IZJAVA O ORIGINALNOSTI RADA

Osijek, 20.09.2017.

Ime i prezime studenta:	Filip Žagar
Studij:	Preddiplomski sveučilišni studij Elektrotehnika
Mat. br. studenta, godina upisa:	3987, 18.07.2014.
Ephorus podudaranje [%]:	2

Ovom izjavom izjavljujem da je rad pod nazivom**: Pregled algoritama za povećanje vremenske rezolucije videosignala**

izrađen pod vodstvom mentora Doc.dr.sc. Mario Vranješ

i sumentora Dr.sc. Denis Vranješ

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija.
Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis studenta:

1. UVOD 1
1.1. Zadatak završnog rada 2
2. POSTUPCI ZA INTERPOLACIJU SLIKA KOD POVEĆANJA VREMENSKE REZOLUCIJE VIDEOSIGNALA 3
2.1. Metode koje koriste prostorne značajke videosignala 4
2.2. Metode koje koriste prostorne i vremenske značajke videosignala 5
2.3. Metode procjene pokreta 5
2.3.1. Vrste pretraživanja u BMA 7
2.3.2. Zaglađivanje vektora pomaka 9
2.4. Interpolacija slike utemeljena na jednosmjernoj procjeni pokreta 10
3. ANALIZA PROSTORNO-VREMENSKIH ZNAČAJKI VIDEOSIGNALA 11
3.1. Prostorna i vremenska aktivnost videosadržaja 11
3.2. Videosignali korišteni u testiranju algoritama 12
3.3. Metode objektivnog vrednovanja kvalitete videosignala 14
3.3.1. PSNR metoda 14
3.3.2. SSIM metoda 15
4. OCJENA KVALITETE INTERPOLIRANIH SLIKA 16
LITERATURA 21
SAŽETAK 22
ABSTRACT 23
KEYWORDS 23
ŽIVOTOPIS 24

1. UVOD

Razvojem elektroničkih uređaja koji se svakodnevno koriste (npr. pametni telefoni i tableti) dolazi i do razvoja digitalnih kamera te samim time i do porasta kvalitete videosadržaja koje snimaju. Razvojem društvenih mreža i mobilnih aplikacija omogućen je prijenos visokokvalitetnih videosadržaja komunikacijskom mrežom, kao i podrška drugim uslugama poput videopoziva ili videokonferencija. S ciljem očuvanja mrežnih resursa i omogućavanja slanja videosadržaja visoke kvalitete, u uvjetima malih brzina prijenosa, izvršava se kodiranje radi smanjenja prostorne i/ili vremenske rezolucije videosignala. Smanjenje vremenske rezolucije provodi se izbacivanjem određenog broja slika iz videosekvence. Nakon prijenosa komunikacijskom mrežom potrebno je vremensku i/ili prostornu rezoluciju videosignala povećati kako bi bio prikazan na uređaju primatelja u visokoj kvaliteti.

Ovaj rad orijentiran je na analizu postojećih rješenja problema povećanja vremenske rezolucije videosignala. U radu su analizirana dva algoritma koji koriste isključivo prostorne značajke videosignala te dva algoritma koji koriste prostorne i vremenske značajke videosignala. U okviru algoritama koji koriste prostorne i vremenske značajke videosignala opisana je metoda procjene pokreta, BMA (*Block Matching Algorithm*) te načini pretraživanja referentnog okvira.

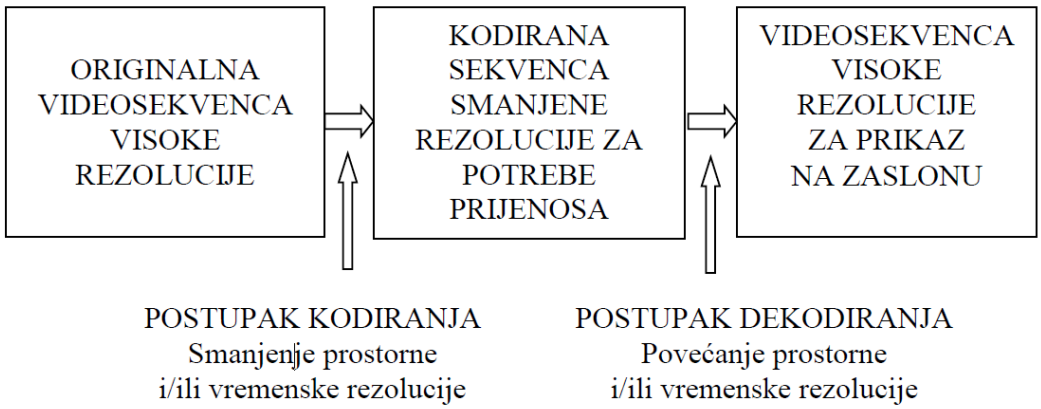
Rad se sastoji od pet poglavlja. Drugo poglavlje opisuje osnovne principe interpolacije slike kod povećanja vremenske rezolucije videosignala. Izvršena je podjela algoritama na one koji koriste samo prostorne i one koji koriste prostorne i vremenske značajke videosignala uz opis njihova načina rada. U trećem poglavlju opisani su videomaterijali koji se koriste u analizi te načini izračuna njihove prostorne i vremenske aktivnosti. Opisane su dvije objektivne metode vrednovanja kvalitete interpoliranih slika, PSNR (engl. *Peak Signal to Nosie Ratio*) i SSIM (engl. *Structural SIMilarity index*). U četvrtom poglavlju prikazani su rezultati ocjene kvalitete videosadržaja, čija je vremenska rezolucija povećana primjenom četiri analizirana algoritma. Za ocjenu kvalitete korištene su dvije objektivne metode, PSNR i SSIM. Posljednje poglavlje sadrži zaključke proizašle iz rezultata provedenih testova.

1.1. Zadatak završnog rada

U radu je potrebno dati pregled stanja postojećih algoritama za povećanje vremenske rezolucije videosignala i komentirati prednosti i mane pojedinih rješenja. Odabrati određena javno dostupna rješenja i testirati njihov rad na danom broju uzoraka videosignala.

2. POSTUPCI ZA INTERPOLACIJU SLIKA KOD POVEĆANJA VREMENSKE REZOLUCIJE VIDEOSIGNALA

Prilikom prijenosa videomaterijala velike vremenske i prostorne rezolucije mrežom poželjna je štednja mrežnih resursa te se u svrhu istoga izvršava smanjenje vremenske i/ili prostorne rezolucije videosignala. Nakon prijema videomaterijala reducirane vremenske i/ili prostorne rezolucije potrebno je istu ponovno povećati [1][2]. Videomaterijal s većom vremenskom rezolucijom koristi više okvira u sekundi te je jasno „uglađeniji“ u odnosu na onaj sa znatno manje okvira u sekundi, posebno ako sadrži puno pokreta [3]. Prema Slici 2.1. originalna videosekvenca visoke rezolucije postupkom kodiranja prilagođava se za prijenos smanjenjem prostorne i/ili vremenske rezolucije te se nakon prijenosa rezolucija povećava i prikazuje na zaslonu primatelja.



Slika 2.1. Pojednostavljeni prikaz sustava za prijenos videosignala uz smanjenje/povećanje prostorne i/ili vremenske rezolucije videosignala [4].

Postupci za povećanje vremenske rezolucije mogu se podijeliti u dvije osnovne skupine. Prvu skupinu predstavljaju postupci koji u obzir uzimaju samo prostorne značajke videosignala, dok druga skupina uz prostorne uključuje i vremenske značajke videosignala.

2.1. Metode koje koriste prostorne značajke videosignala

Prvu skupina metoda za povećanje vremenske rezolucije videosignala u obzir uzima samo prostorne značajke signala ne uzimajući u obzir pokrete objekata u sceni. Pri takvim postupcima elementi interpoliranih slika se predviđaju na temelju informacija o elementima na istim pozicijama unutar prethodne i sljedeće slike u odnosu na sliku koja se interpolira [1]. Najjednostavniji postupak, koji koristi samo prostorne značajke videosignala, predstavlja interpolacija slike ponavljanjem prethodne slike (engl. *Frame Repetition*-FR) prema izrazu (2-1).

$$F_t(x, y) = F_{t-1}(x, y), \tag{2-1}$$

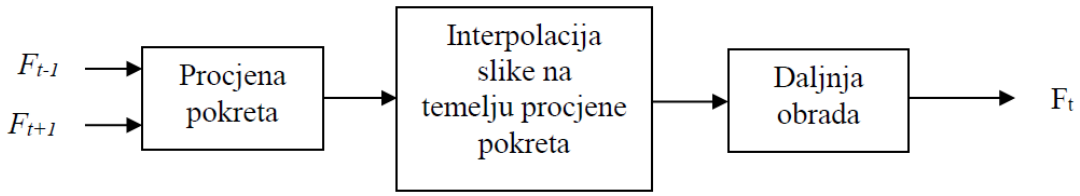
Gdje F_t predstavlja interpolirani element slike a F_{t-1} element slike koji mu prethodi. Pozicija elemenata slike unutar slike određena je koordinatama *x* i *y*. Vrlo često korištena metoda, koja pripada ovoj skupini, jest interpolacija slike korištenjem srednje vrijednosti elemenata na istim pozicijama u prethodnoj i sljedećoj slici (engl. *Frame Averaging*-FA).

$$F_t(x, y) = \alpha \cdot F_{t-1}(x, y) + (1+\alpha) \cdot F_{t+1}(x, y) \tag{2-2}$$

U izrazu (2-2), koji definira interpolaciju slike metodom FA, $F_t(x, y)$, $F_{t-1}(x, y)$, $F_{t+1}(x, y)$ redom predstavljaju interpolirane, prethodne i sljedeće elemente slike na pozicijama (*x*,*y*) unutar slike. Koeficijent α predstavlja vremenski omjer vremenske udaljenosti interpolirane slike od prethodne i vremenske udaljenosti sljedeće slike od prethodne. Ukoliko između dvije originalne slike interpoliramo samo jednu sliku α iznosi 0.5. U sekvencama s brzim i naglim pokretima nije moguće postići zadovoljavajuću kvalitetu interpolirane slike metodama koje koriste samo prostorne značajke videosignala pa se u takvim slučajevima koriste metode koje osim prostornih koriste i vremenske značajke videosignala.

2.2. Metode koje koriste prostorne i vremenske značajke videosignala

Druga skupina metoda uz prostorne koristi i vremenske značajke videosekvence uzimajući u obzir i pokrete objekata u sceni. Postupak interpolacije slike korištenjem ove skupine metoda može se podijeliti u tri cjeline: procjena pokreta, interpolacija slike na temelju procjene pokreta i daljnja obrada, kako je prikazano na Slici 2.2 [4].



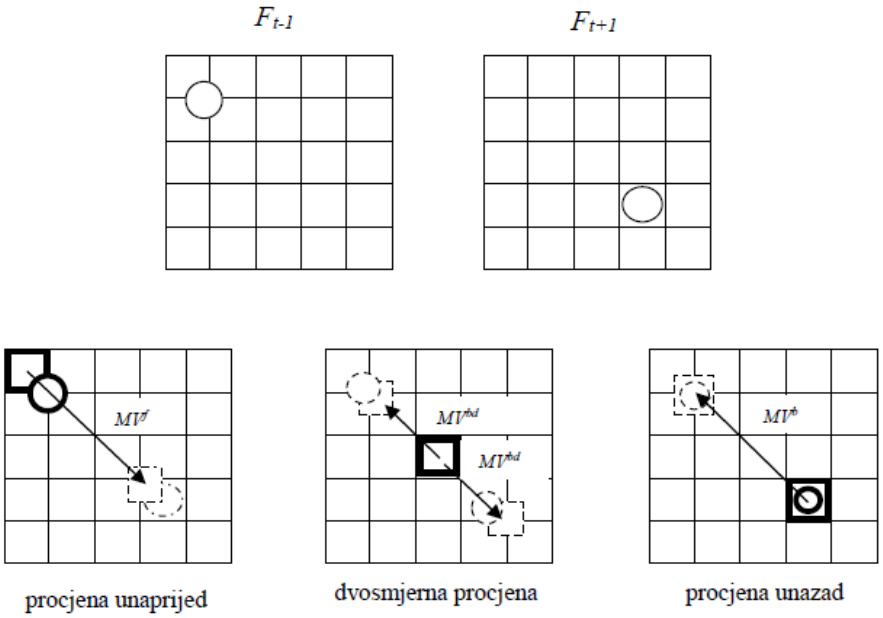
Slika 2.2. Prikaz procesa stvaranja interpolirane slike pri primjeni procjene pokreta.

2.3. Metode procjene pokreta

Procjena pokreta predstavlja prvi korak u interpolaciji slike primjenom metoda koje u obzir uzimaju prostorne i vremenske značajke videosignala. Praćenje položaja objekta unutar scene može se odrediti praćenjem pokreta objekata, optičkim tokom ili najčešće proračunom vektora pomaka za dijelove slike određene veličine. Precizno i točno izračunavanje vektora pomaka predstavlja vrlo bitan faktor pri krajnjem rezultatu kvalitete interpoliranih slika, stoga je potrebno izabrati što efikasniju metodu izračuna koja je u ovom slučaju *Block Matching* Alghorithm - BMA. U okviru BMA procjena pokreta može se izvršiti jednosmjerno ili dvosmjerno, a jednosmjerna procjena pokreta može se izvršti unaprijed ili unazad. Prilikom jednosmjerne procjene pokreta unaprijed prethodna slika F_{t-1} dijeli se na blokove te se traži najsličniji blok u sljedećoj slici, F_{t+1} dok se pri jednosmjernoj procjeni pokreta unazad sljedeća slika F_{t+1} dijeli na blokove te se traži najsličniji blok u prethodnoj slici F_{t-1} [4].

Izračunom razlike pozicija najsličnijih blokova u prethodnoj i sljedećoj slici dobiva se vektor pomaka predviđen unaprijed ili unazad. Moguće je vršiti i dvosmjernu procjenu pokreta pri kojoj se slika koja se interpolira dijeli na blokove te se istovremeno traže najsličniji blokovi u prethodnoj i sljedećoj slici. Pri dvosmjernoj

procjeni dobivaju se vektori pokreta istog iznosa amplitude, ali suprotnog smjera. Postupak procjene pokreta u BMA prikazan je slikom 2.3.



Slika 2.3. Postupci procjene vektora pokreta [4]

Interpolirana slika nalazi se između prethodne i sljedeće slike, stoga je potrebno dobivene vrijednosti vektora pomaka skalirati ovisno o vremenskoj udaljenosti prethodne i sljedeće slike. Određivanje vektora pomaka može se prikazati matematičkim izrazima (2-3) i (2-4) koji redom predstavljaju jednosmjernu procjenu unaprijed i jednosmjernu procjenu unazad[4].

$$MV^{f}(B_{i,j}) = \beta \cdot \left[x_{t+1}(B_{i,j}^{odg}) - x_{t-1}(B_{i,j}), y_{t+1}(B_{i,j}^{odg}) - y_{t-1}(B_{i,j}) \right] \tag{2-3}$$

$$MV^{b}(B_{i,j}) = (1-\beta) \cdot \left[x_{t-1}(B_{i,j}^{odg}) - x_{t+1}(B_{i,j}), y_{t-1}(B_{i,j}^{odg}) - y_{t+1}(B_{i,j}) \right] \tag{2-4}$$

Gdje je $MV^{f}(B_{i,j})$ vrijednost vektora pokreta dobivena jednosmjernom procjenom unaprijed za blok $B_{i,j}$ i njemu najsličniji blok u referentnoj slici $B_{i,j}^{odg}$ sa pozicijom na koordinatama *i* i *j*, a β koeficijent koji ovisi o broju interpoliranih slika između dvije originalne slike.

Za pronalazak najsličnijeg bloka u referentnoj slici upotrebljava se suma apsolutnih razlika(engl. *Sum of Absoulte Differences* – SAD). Suma apsolutnih razlika je vrijednost koja predstavlja sumu apsolutnih vrijednosti razlika elemenata slike na istim pozicijama unutar trenutnog i referentnog bloka, što implicira da je najsličniji blok u referentnoj slici onaj za koji je SAD minimalan. SAD je za jednosmjernu procjenu pokreta definiran izrazom:

$$SAD = \sum_{x,\,y} \left| F_{t-1}(x, y) - F_{t+1}(x + MV_x, y + MV_y \right|. \qquad (2\text{-}5)$$

MV_x predstavlja horizontalnu komponentu, a MV_y predstavlja vertikalnu komponentu vektora pomaka. Dvosmjerna procjena pokreta postiže nezadovoljavajuće rezultate ukoliko se primjenjuje na videomaterijale s puno pokreta, stoga će se u radu opisivati i primjenjivati jednosmjerna procjena unaprijed i unazad [6].

2.3.1. Vrste pretraživanja u BMA

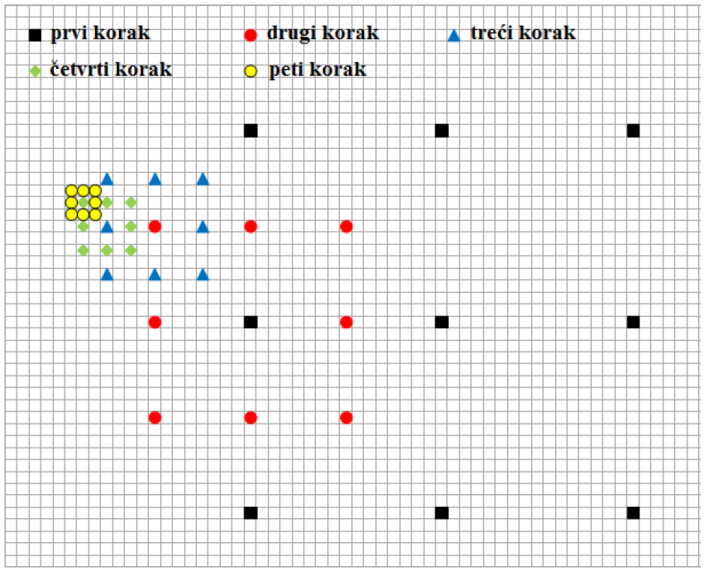
U BMA jedan okvir slike je podijeljen u skup nepreklapajućih pravokutnih blokova slike - B. Pojedini blok *b* koji je dio skupa B ($b \in B$), uspoređuje se s blokovima sadržanim unutar okolnih referentnih okvira slike. Najsličniji blok referentne slike bloku *b*, pronađen korištenjem nekih od prikladnih funkcija sličnosti, koristi se za određivanje vektora pokreta za taj blok. Najčešće korištene veličine blokova su 4x4, 8x8 i 16x16, jer su djelitelji standardno korištenih rezolucija slike [3]. Za svaki odabrani blok *b* određene veličine traži se njemu odgovarajući blok u referentnoj slici unutar područja pretrage (engl. *Search Range* - SR). Pronalazak odgovarajućih blokova je računalno zahtjevan zadatak, stoga je poželjno upotrijebiti BMA koji nastoji smanjiti broj potrebnih operacija uz zadovoljavajuću razinu preciznosti. Ukoliko se uzima veće područje pretraživanja povećava se preciznost određivanja vektora pomaka, ali i složenost postupka [4]. Cilj je postići optimalan omjer između preciznosti određivanja vektora pomaka i složenosti procesa pretrage.

Kao izbor nude se brojne mogućnosti pretraživanja a neke od njih su:

- Potpuno pretraživanje (engl. *Full Search* – FS)
- Logaritamsko pretraživanje (engl. *Log Search* – LS)
- Pretraživanje u tri koraka (engl. *Three Step Search* - TSS)
- Pretraživanje u četiri koraka (engl. *Four Step Search* - 4SS)
- Pretraživanje korištenjem strukture dijamanta (engl. *Diamond Search*- DS)
- Pretraživanje uz adaptivno područje pretrage (engl. *Adaptive Root Pattern Search* – ARPS)

Potpuno pretraživanje podrazumijeva izračun SAD-a za svaki referentni blok unutar područja pretrage u odnosu na trenutni, što rezultira najtočnijim rezultatima, ali je izuzetno računalno zahtjevno. Puno prihvatljiviju metodu predstavljaju metode pretraživanja u tri koraka i logaritamsko pretraživanje, koje za puno manji broj operacija daju približno jednaku točnost kao potpuno pretraživanje, stoga će se detaljnije opisati i primijeniti u nastavku rada [4]. Kao primjer će se uzeti slika rezolucije 352x288 elemenata slike i veličinom bloka 16x16 elemenata slike. Koristeći područje pretrage od 32 elementa slike odgovarajući blok tražio bi se koristeći potpuno pretraživanje na 3969 lokacije, dok bi se koristeći logaritamsko pretraživanje odgovarajući blok tražio na 41 lokaciji, što je drastično manje. BMA ovisi o svojstvu pokreta i veličini objekta u pokretu. U globalnim dijelovima pokreta poželjno je da je veličina blokova podjele slike što veća, a u lokalnim dijelovima pokreta poželjno je da veličina bloka odgovara veličini objekta u pokretu. Manja veličina okvira rezultira većom pogreškom procjene vektora pokreta ili netočnim vektorom pokreta. Ako se blok manje veličine nalazi u području slike bez pokreta njegov vektor pokreta neće iznositi nula zbog aditivnog šuma ili fliker šuma [5]. Prema [4] pokazano je da se podjelom slike na blokove veličine 16x16 elemenata slike i upotrebom logaritamskog pretraživanja sa područjem pretrage 32 elementa slike dobiva manja pogreška procjene vektora pomaka kao i manja pogreška prilikom interpolacije slike u usporedbi s blokovima veličina 4x4 i 8x8 elemenata slike. Uzimajući u obzir povećanje računalne složenosti smanjenjem veličine bloka dimenzije blokova veličine 16x16 elemenata slike predstavljaju optimalno rješenje. Veličina područja pretrage određuje broj lokacija pretrage. Odabere li se veličina područja pretrage od 32 elementa slike i logaritamsko pretraživanje u prvom koraku

dolazi do usporedbe trenutnog bloka s blokom na istoj poziciji u referentoj slici te sa 8 blokova udaljenih 16 elemenata slike od trenutnog bloka horizontalno i vertikalno. Uzimajući SAD kao mjerilo sličnosti područje pretrage svoje središte pozicionira na element slike s najmanje proračunatom vrijednosti SAD-a. Prije ponovne pretrage područje pretrage se prepolovi te tada iznosi 16 elemenata slike. Ponovno pretraživanje uspoređuje element u prvom pretraživanju označen kao onaj sa najmanjom vrijednosti SAD-a s elementom slike na istoj poziciji te na onima 8 elemenata slike udaljenima od središta pretrage horizontalno i vertikalno. Ovaj postupak se ponavlja dok se područje pretrage ne svede na jedan element slike što se vidi na Slici 2.4. Razlika između trenutne slike i dobivenog bloka sa najmanjom vrijednosti SAD u odnosu na trenutni blok čini vektor pomaka [4].



Slika 2.4. Logaritmsko pretraživanje područja veličine 32 elementa slike [4].

2.3.2. Zaglađivanje vektora pomaka

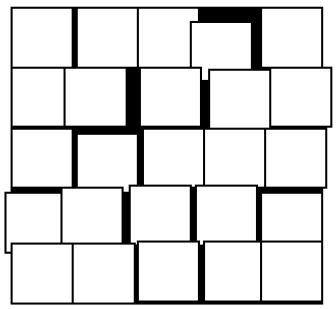
Za dobivanje vjernije putanje nekog gibanja unutar videomaterijala potrebno je stvoriti koherentno polje vektora pomaka u području koje zauzimaju objekti u pokretu. Koherentno polje vektora pomaka za objekte u pokretu postiže se prostornom i vremenskom korelacijom vektora pomaka susjednih blokova [4].

Postupci za zaglađivanje vektora pomaka u posljednjih desetak godina ugrađuju se u postupak procjene pokreta. Jednu vrstu postupaka predstavlja implicitno zaglađivanje nakon proračuna vektora pomaka putem neke vrste filtriranja. Među raznim

rješenjima navedenim u [4] ističe se medijan filtar kao optimalno rješenje zbog svoje jednostavnosti i visoke razine kvalitete slike. Zaglađivanje dovodi i do smanjenja artefakata pojave oštrih rubova na granicama blokova i objekata tj. povećanja kvalitete interpolirane slike.

2.4. Interpolacija slike utemeljena na jednosmjernoj procjeni pokreta

Prilikom jednosmjerne procjene pokreta kao prvotni rezultat dobivaju se dvije slike prediktori. Prediktori predstavljaju dvije slike dobivene jednosmjernim procjenama pokreta unaprijed i unazad. Slika koja se interpolira rezultat je težinske sume dva prediktora. S obzirom da se pri jednosmjernoj procjeni pokreta izračunavaju vrijednosti vektora pomaka za prethodnu i sljedeću sliku te se njihovi blokovi pomiču na temelju izračuna konačna slika će sadržavati šupljine i preklapanja blokova kao na Slici 2.5. [4].



Slika 2.5. Slika interpolirana jednosmjernom procjenom pokreta.

Šupljine i preklapanja potrebno je prikladno nadomjestiti kako bi se osigurala što bolja kvaliteta interpolirane slike. Prema [4] predloženo je nekoliko načina uklanjanja šupljina i preklapanja. Među predloženim načinima za otklanjanje ovog problema ističe se način procjene vektora pokreta za preklapajuća područja i proračun elemenata u području šupljina koji postiže dobre rezultate kvalitete interpolirane slike, a prikazan je u [7].

3. ANALIZA PROSTORNO-VREMENSKIH ZNAČAJKI VIDEOSIGNALA

Prostorno-vremenske značajke videosignala su kroz istraživanja pokazala bitan utjecaj na pogrešku prilikom interpolacije slike. Velika vremenska aktivnost videosadržaja uzrokuje veću pogrešku procjene interpolirane slike. S ciljem izrade najkvalitetnije moguće interpolirane slike potrebno je poznavati prostorne-vremenske značajke videosignala te njihov utjecaj na kvalitetu interpolirane slike.

3.1. Prostorna i vremenska aktivnost videosadržaja

Za opis vremenske i prostorne aktivnosti Međunarodna Telekomuikacijska Unija (engl. *International Telecommunication Union* – ITU) definira parametre *Temporal perceptual Information* – TI , za vremensku, te *Spatial perceptual Information* –SI za prostornu aktivnost.

$$SI = \max\left\{ std\left[Sobel(\mathrm{F}_t(\mathrm{x},\mathrm{y})) \right] \right\}, \tag{3-1}$$

$$TI = \max\left\{ std\left[F_t(x,y) - F_{t-1}(x,y) \right] \right\}, \tag{3-2}$$

gdje *std* predstavlja standardnu devijaciju, a *Sobel* sobelov operator. SI i TI za sekvence videosignala daju maksimalnu vrijednost tih veličina za svaku sliku pojedinačno, što rezultira neodgovarajućim vrijednostima za sekvence koje posjeduju nekoliko slika velike, a ostale sekvence male vremenske i/ili prostorne aktivnosti [4]. Iz tog razloga koristi se metoda izračuna prema izrazima (3-3) i (3-4), u kojoj se umjesto maksimalne vrijednosti uzima srednja vrijednost kroz sekvence, što doprinosi pouzdanijim rezultatima mjerenja prostorne i vremenske aktivnosti videosadržaja:

$$SI_{mean} = mean\left\{ std\left[Sobel(F_t(x,y)) \right] \right\}, \tag{3-3}$$

$$TI_{mean} = mean\left\{ std\, F_t(x,y) - F_{t-1}(x,y))] \right\}. \tag{3-4}$$

3.2. Videosignali korišteni u testiranju algoritama

Za analizu je korišteno pet videosadržaja različitih prostorno-vremenskih značajki: *container, football, news, mobile, foreman*. Navedeni materijali predstavljaju standardizirane sadržaje koji se koriste pri analizi i obradi videosignala. Videosadržaji su u *Common Intermediate Format*-u (CIF) s prostornom rezolucijom od 352x288 elemenata slike, te vremenskom rezolucijom od 25 slika u sekundi. Sadržaj *container*, *foreman* i *mobile* sadrže 300 slika, *news* 90 te *football* 260 slika. Videosekvence su nekomprimirane, u formatu YUV, s 4:2:0 shemom poduzorkovanja krominantnih komponenti U i V. U poustupku kodiranja sadržaja, tj. smanjivanja vremenske rezolucije odbačene su parne, a zadržane neparne slike, što je rezultiralo smanjenjem broja slika u videoskevencama na pola te je vremenska rezolucija smanjena na 12,5 slika u sekundi. Videosekvence *container* i *news* snimljene su mirnom kamerom dok je u sekvenci *football* prisutno kretanje kamere. Sekvencu *football* karakteriziraju vrlo brzi pokreti snimljenih objekata. *Container* i *News* sadrže značajno sporije pokrete dok se kod sekvence news objekti u pozadini kreću, a u prvom planu su voditelji koji čitaju vijesti. Videosadržaj *Mobile* sniman je pokretnom kamerom te s karakterstičnim pokretima objekata u pozadini. U sekvneci *Foreman* zabilježen je prikaz glave čovjeka koji govori te nakon toga uz pokret kamere objekt u izgradnji. Karakteristične slike iz navedenih videoskevenci nalaze se na Slici 3.1. Vrijednosti SI_{mean} i TI_{mean} izračunate prema formulama (3-3) i (3-4) nalaze se u tablici 3.1.

Tablica 3.1. Vrijednosti SI_{mean} i TI_{mean} videosekvenci [4].

Sekvenca	SI_{mean}	TI_{mean}
Container	125,98	5,39
News	122,72	7,95
Mobile	168,36	38,49
Foreman	78,90	20,90
Football	83,46	33,58



a)



b)



c)



d)



e)

Slika 3.1. Karakteristične slike iz sekvenci: a) *container* b) *news* c) *mobile* d) *foreman* e) *football.*

3.3. Metode objektivnog vrednovanja kvalitete videosignala

Objektivne metode vrednovanja kvalitete koje će biti obrađene u ovome radu koriste matematičke modele za ocjenu kvalitete videosignala bez potrebe korištenja ljudskih resursa. Objektivne metode dijele se u tri skupine: metode koja ne zahtijevaju nikakve informacije o referentnom signalu (engl. *No – Reference* – NR), metode koje zahtijevaju potpune informacije o referentnom signalu (engl. *Full Reference*-FR) i metode koja zahtijevaju samo neke značajke referentnog siganala (engl. *Reduced – Reference* – RR) [4]. Procijeniti kvalitetu interpoliranih okvira bez dostupnosti informacija ili s djelomičnim informacijama o referentnom signalu je vrlo teško dok nedostatke FR metoda predstavljaju potreba za potpunom dostupnosti referentnog signala te za vremenskim poravnavanjem referentnog i izobličenog signala jer se ocjenjivanje provodi sekvencijalno po slikama [3]. Najznačajnije metode iz skupine FR metoda su srednja kvadratna pogreška(engl. *Mean Squared error*-MSE) i omjer vršne snage i snage šuma (engl. *Peak Signal to Noise Ratio* – PSNR) [8][9]. U ovome radu provedeno je ocijenjivanje kvalitete videosignala korištenjem PSNR metode te mjerenjem indeksa strukturalne sličnosti (engl. *Structural SIMilarity indeks* – SSIM), koji pokazuje veću korelaciju sa subjektivnim ocjenama jer u obzir uzima svojstva ljudskog vizualnog sustava [10].

3.3.1. PSNR metoda

Objektivna metode se dijele u tri skupine prema načinu na koji analiziraju dekodirani signal: podatkovne metode, slikovne metode i metode zasnovane na paketima i toku bita. Podatkovna metode vjernost izobličenog signala u odnosu na referentni signal mjere bez uzimanja u obzir sadržaj videosignala. PSNR je iskazan u decibelima prema izrazu (3-5).

$$PSNR_{dB} = 10\log_{10}\frac{(2^m-1)^2}{MSE}, \tag{3-5}$$

gdje je MSE srednja kvadratna pogreška, 2^m-1 maksimalna vrijednost koju element slike može poprimiti za sliku od m bita. U ovome radu koristit će se vrijednost od m=8 bita po elementu slike što ukazuje da je maksimalna vrijednost slike 255. Unatoč tomu što često ne daje jednake rezultate kao subjektivne metode ocjene kvalitete, PSNR je u praksi standard za objektivnu metodu ocjene kvalitete videosignala. Razlog toga je brzina i jednostavanost za implementaciju u usporedbi sa drugim

kompleksnijim metodama [12]. Zbog prethodno navedenog nedostataka PSNR metode u radu će se obraditi i SSIM metoda.

3.3.2. SSIM metoda

SSIM, kao i PSNR pripada FR metodama objektivnog vrednovanja videosignala. SSIM za razliku od PSNR-a u obzir uzima razliku u strukturi i kontrastu. SSIM kao rezultat daje normaliziranu vrijednost u rasponu od -1 do 1, gdje vrijednost 1 označava identične slike. Proračun SSIM vrijednosti dobiva se prema (3-6) gdje su *x* i *y* okviri, μ_x i μ_y procijenjene srednje vrijednosti luminancije slika *x* i *y*, σ_x i σ_y su procijenjeni kontrasti od slika x i y, σ_{xy} je kovarijanca od σ_x i σ_y a C_1 i C_2 konstante koje se koriste kako bi se izbjegla nestabilnost funkcije.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \qquad (3\text{-}6)$$

Detaljniji postupak proračuna SSIM-a opisan je u [4]. Ocjena kvalitete interpolirane slike SSIM metodom daje znatno vjernije rezultate, tj. sličnije onima subjektivnih metoda u usporedbi s PSNR metodom, ali je kompleksnija i teža za implementaciju.

4. OCJENA KVALITETE INTERPOLIRANIH SLIKA

U ovom radu provedena je ocjena kvalitete slika interpoliranih s 4 različite metode na 5 prethodno spomenutih videosadržaja. Eksperiment je provedem primjenom dviju metoda koje koriste samo prostorne značajke videosignala FR i FA, te algoritam koji koristi prostorne i vremenske značajke videosignala AST (*Algorithm based on videosignal Spatio-Temporal features*), u dvije inačice koje se razlikuju u metodi pretraživanja najsličnijih blokova. Algoritam AST-FS koristi potpuno pretraživanje (FS) koji daje veliku točnost, ali je računalno vrlo zahtjevno u usporedbi s AST-LS koji koristi logaritamsko pretraživanje koje računalno znatno manje zahtjevno, a daje približno jednake rezultate. U AST algoritmima konačna interpolirana slika rezultat je srednje vrijednosti slika prediktora dobivenih odvojenim procjenama pokreta unaprijed i unazad. Kvaliteta interpoliranih slika određena je objektivnim metrikama za ocjenu kvalitete videosignala, PSNR i SSIM u programskom alatu *MSU Video Quality Measurment Tool ver.7.1*. Rezultati objektivne ocjene PSNR metrikom, uz vrijednosti prosječnih vremenskih aktivnosti, prikazani su u tablici 4.1. dok su rezultati ocjene kvalitete objektivnim SSIM metrikom prikazani u tablici 4.2.

Tablica 4.1. Prosječna vremenska aktivnost, TI_{mean}, videosadržaja i ocjena kvalitete interpoliranih slika za različite metode interpolacije korištenjem PSNR metode.

		PSNR			
SEKVENCA	TI_{mean}	**FR**	**FA**	**AST-LS**	**AST-FS**
container	5,39	39.79	43.2	41.75	41.75
news	7,95	34.29	36.83	36.82	36.82
foreman	20,90	27.56	29.78	32.51	33.74
football	33,58	21.77	23.74	25.14	25.49
mobile	38,49	24.6	31.05	31.86	31.82
prosjek	21.26	29.61	32.92	33.62	33.92

Tablica 4.2. Prosječna vremenska aktivnost, TI_{mean} ,videosadržaja i ocjena kvalitete interpoliranih slika za različite metode interpolacije korištenjem SSIM metode.

		SSIM			
SEKVENCA	TI_{mean}	**FR**	**FA**	**AST-LS**	**AST-FS**
container	5,39	0.989	0.994	0.993	0.993
news	7,95	0.986	0.992	0.992	0.992
foreman	20,90	0.875	0.901	0.949	0.959
football	33,58	0.745	0.768	0.812	0.836
mobile	38,49	0.873	0.929	0.976	0.976
prosjek	21.26	0.894	0.917	0.944	0.9512

Iz podataka prikazanih u tablici 4.1 vidljivo je da vremenska aktivnost videosignala utječe na kvalitetu interpoliranih slika bez obzira na korištenu metodu interpolacije slike. Uočava se razlika u kvaliteti između algoritama koji koriste samo prostorne značajke videosignala u odnosu na one koji koriste prostorno vremenske značajke, osobito u slučaju videosadržaja veće vremenske aktivnosti. FA uvijek daje bolje rezultate (u prosjeku oko 3 dB) od FR jer FA pri interpolaciji u obzir uzima prethodnu i sljedeću sliku tj. njihovu srednju vrijednost. Drastična razlika između ova dva algoritma se pojavljuje ukoliko se u sljedećoj slici pojavljuje neki objekt koji nije postojao u prethodnoj slici te se on zbog ponavljanja prethodne slike u FR algoritmu neće pojaviti u interpoliranoj slici dok će se pri korištenju FA algoritma pojaviti, iako zamagljen i s greškama. Usporedbom algoritma FA sa složenijim algoritmima AST-LS i AST-FS vidljivo je da za videosadržaje male vremenske aktivnosti daje približno jednake rezultate, a u slučaju najsporije sekvence *container* čak i bolje, dok za sekvence velike vremenske aktivnosti npr. Football složeniji algoritmi prednjače za oko 2 dB. Razlog ovakvog rezultata je pojava pogreške pri procjeni pokreta koja se izvodi u sklopu AST-LS i AST-FS logaritama, dok se kod FA izračunava samo srednja vrijednost prethodne i sljedeće slike što daje bolje rezulate ukoliko se radi o videosadržajima male vremenske aktivnosti. U analizi sekvence *mobile* dolazi do anomalije koja rezultira da za videosadržaj velike vremenske aktivnosti FA algoritam daje gotovo jednaku kvalitetu interpolirane slike kao i AST-LS i AST-FS algoritmi, a razlog je što velika vremenska aktivnost sekvence *mobile* proizlazi iz velike prostorne

aktivnosti koju zapravo posjeduje. Usporedbom AST-LS algoritma s AST-FS uviđamo kako ne postoji potreba koristiti algoritam AST-FS algoritam koji je računalno puno zahtjevniji jer u slučaju svih sekvenci daje neznatno bolje rezultate od AST-LS algoritma. Ocjena kvalitete korištenjem SSIM metode prikazana u tablici 4.2. potvrđuje rezultate ocjene kvalitete PSNR metodom.

5. ZAKLJUČAK

Videosadržaji snimljeni suvremenim elektroničkim uređajima imaju veliku vremensku i prostornu rezoluciju te predstavljaju problem u prijenosu mrežom s malom brzinom prijenosa. Problem se rješava postupkom kodiranja kojim se videosadržaju smanjuje vremenska i/ili prostorna rezolucija. Nakon prijenosa videosadržaja vremensku i/ili prostornu rezoluciju je potrebno ponovno povećati kako bi bio prikladan za reprodukciju na uređaju primatelja. U ovome radu ispitana su različita postojeća rješenja algoritama za povećanje vremenske rezolucije videosignala. Uz dva algoritma (FA i FR) koji koriste isključivo prostorne značajke videosignala ispitana su i dva algoritma (AST-LS i AST-FS) koja koriste prostorne i vremenske značajke videosignala. Zahtjevnost algoritama i kvaliteta interpoliranih slika ispitani su na pet videosadržaja različite prostorne i vremenske aktivnosti. Implementacijom algoritama uočena je znatno veća računalna zahtjevnost algoritama koji koriste prostorne i vremenske značajke videosignala. Razlog velike računalne složenosti AST-LS i AST-FS proizlazi iz njihove kompleksnosti tj. postupka prema kojem se izvršava procjena pokreta za blokove slike (BMA) i stvaranja interpolirane slike. Prilikom procjene pokreta izvršava se procjena pokreta unaprijed i procjena pokreta unazad na blokovima veličine 16x16 elemenata slike. Kako bi se dodatno zagladilo polje procijenjenih vektora pokreta primijenjen je medijan filtar. Procjenom vektora unaprijed i unazad dobivaju se dvije slike , a konačno dobivena interpolirana slika predstavlja srednju vrijednost između slika dobivenih procjenom pokreta unaprijed i unazad. Kvaliteta interpoliranih slika ispitana je objektivnim metodama za ocjenu kvalitete videosignala – PSNR i SSIM. Provedenim ispitivanjem pokazano je da pri povećanju vremenske rezolucije videosekvenci koje posjeduju malu srednju vremensku aktivnost (TI_{mean}) algoritmi koji koriste isključivo prostorne značajke videosignala daju malo lošije rezultate u slučaju FR algoritma ili čak i bolje u slučaju FR algoritma od AST-FS i AST-LS algoritama. Budući da su algoritmi koji koriste prostorne i vremenske značajke videosignala (AST-LS i AST-FS) računalno složeniji oni su neekonomični za korištenje u slučaju videosadržaja male vremenske aktivnosti. Ocjene kvalitete videosekvenci veće vremenske aktivnosti pokazuju dominantnost AST-LS i AST-FS algoritama nad FA i FR algoritmima. Rezultati ispitivanja kvalitete pokazuju neznatnu razliku u kvaliteti slika interpoliranih algoritmima koji

koriste vremenske i prostorne značajke videosignala, stoga je puno bolje koristiti AST-LS algoritam zbog manje računalne zahtjevnosti. Kao optimalno rješenje nameće se algoritam koji bi prvotno ispitao srednju vremensku aktivnost videosadržaja te nakon toga odlučivao o metodi povećanja vremenske rezolucije videosignala.

LITERATURA

[1] Y. Dar, A. M. Bruckstein, **Motion-compensated coding and frame rate up-conversion: models and analysis,** *IEEE Transactions on Image Processing*, vol. 24, no. 7, 2015, str. 2051-2066

[2] A. M. Bruckstein, M. Elad, R. Kimmel, **Down-scaling for better transform compression,** *IEEE Transactions on Image Processing*, vol. 12, no. 9, 2003, str. 1132-1144

[3] A. Isberg, J. Jostell, **Frame rate-up conversion of real-time high-definition remote surveillance video**, Göteborg, 2012

[4] D. Vranješ, **Algoritam za adaptivnu vremensku interpolaciju slika zasnovan na prostorno-vremenskim značajkama videosignala**, Osijek 2016.

[5] Fujiwara, S.,Taguchi A., **Motion-Compensated Frame Rate Up-Conversion Based on Block Matching Algorithm**, Intelligent Signal Processing and Communication Systems, 2005.

[6] B.T. Choi, S.-H. Lee, S.-J. Ko, **New frame rate up-conversion using bidirectional motion estimation,** *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3

[7] S. Dikbas, Y. Altunbasak, **Novel true-motion estimation algorithm and its application to motion-compensated temporal frame interpolation,** *IEEE Transactions on Image Processing*, vol. 22, no. 8

[8] A.M. Eskicioglu, P.S. Fisher, **Image quality measures and their perofrmance,** *IEEE Transactions on Communications*, vol. 43, 1995.

[9] M. Vranješ, S. Rimac-Drlje, D. Žagar, **Subjective and Objective Quality Evaluation of the H.264/AVC coded Video,** U:*Proc. Of International Conference on Systems, Signals and Image Processing*, Slovačka, 2008.

[10] Z. Wang, L. Lu, A.C. Bovik, **Video quality assessment based on structural distiortion measurement,** *Signal Processing: Image Communication*, vol. 19, no. 2, 2004.

[11] Wang, Z. & Bovik , A. **Mean squared error: love it or leave it? A new look at signal fidelity measures,** *Signal Pro-cessing Magazine, IEEE 26*, 2009.

[12] Oelbaum, T., Diepold, K. & Zia, **A generic method to increase the prediction accuracy of visual quality metrics, in 'Picture Coding Symposium (PCS)'**, 2007.

SAŽETAK

Radi očuvanja mrežnih resursa i omogućavanja dovoljno brzog prijenosa videosignala u uvjetima malih brzina prijenosa potrebno je smanjiti količinu podataka za prijenos. Smanjenje količine podataka za prijenos izvršava se kodiranjem videosignala s ciljem smanjenja vremenske rezolucije. Nakon prijenosa mrežom istim video sadržajima potrebno je povećati vremensku rezoluciju kako bi bili prikladni za prikaz na uređajima koji podržavaju videosadržaj visoke prostorne i vremenske rezolucije. U radu je ispitana kvaliteta interpoliranih slika dobivenih korištenjem nekih od postojećih algoritama za povećanje vremenske rezolucije videosignala. Za analizu su odabrana dvije algoritma koja koriste isključivo prostorne značajke videosignala (FA i FR) te dva algoritma koji koriste prostorne i vremenske značajke videosignala (AST-LS i AST-FS). U okviru usporedbe algoritama koji koriste samo prostorne značajke videosignala uočena je prednost u kvaliteti interpoliranih slika sa FA algoritmom uz približno jednaku računalnu zahtjevnost i vrijeme izvedbe. FA algoritam je pokazao najveću razinu kvalitete interpolirane slike za videosadržaje male vremenske aktivnosti dok za sadržaje velike vremenske aktivnosti prednjače algoritmi koji koriste vremenske i prostorne značajke videosignala. Usporedbom AST-LS i AST-FS algoritama kao bolje rješenje pokazao se AST-LS zbog manje računalne kompleksnosti uz gotovo jednake rezultate kvalitete interpoliranih slika. Za objektivnu ocjenu interpoliranih slika korištene su PSNR i SSIM metrike.

KLJUČNE RIJEČI

procjena pokreta, interpolacija slike, prostorno-vremenska aktivnost videosignala, kvaliteta slike, složenost algoritma

ABSTRACT

For the purpose of efficient usage of network resources and enabling fast transmission it is needed to reduce amount of video signal data. Amount of the data for transmission is reduced by reducing temporal resolution in the coding process. After transmission over the network, it is needed to perform algorithms for frame rate-up conversion to make the video signal suitable for reproduction on devices supporting videos with high spatial and temporal activity. In this thesis, quality of interpolated images was measured. Interpolated images were made with some of already proposed algorithms for frame rate up-conversion. Two of the analyzed algorithms (FR and FA) use only spatial features of video signal and other two algorithms use spatial and temporal features of video signal (AST-FS and AST-LS). The FA algorithm shows much better quality of interpolated images compared to the FR with similar computational complexity and execution time. In case of very slow video sequences, the FA algorithm shows the best results of interpolated image quality but in case of very fast video sequences, algorithms that use spatial and temporal features of video signal are leading. The AST-LS, in comparison with the AST-FS, was revealed as a better decision because of similar quality of interpolated images but less computationally complex. PSNR and SSIM metrics were used for objective interpolated image evaluation.

KEYWORDS

motion estimation, image interpolation, spatio-temporal activity of video signal, image quality, algorithm complexity

ŽIVOTOPIS

Filip Žagar rođen je 2. ožujka 1996. godine u Vinkovcima. Nakon završene Opće gimnazije u Vinkovcima, 2014. godine upisuje preddiplomski studij elektrotehnike na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku. Tijekom preddiplomskog studija aktivno sudjeluje u nastavi u ulozi demonstratora na kolegijima Programiranje II i Elektronika II te dobiva priznanje za nagrađeni studentski rad prigodom obilježavanja 39. godišnjice fakulteta.