Maskiranje u video signalu

Stupar, Irinej

Undergraduate thesis / Završni rad

2016

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:200:467292

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-28**

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
ELEKTROTEHNIČKI FAKULTET

Stručni studij

MASKIRANJE U VIDEO SIGNALU

Završni rad

Irinej Stupar

Osijek, 2016

SADRŽAJ

1. UVOD 1

1.1 Zadatak završnog rada 1

2. MASKIRANJE U VIDEO SIGNALU 2

2.1 Prostorno maskiranje 4

2.2 Vremensko maskiranje 7

3. ANALIZA UTJECAJA MASKIRANJA NA KVALITETU KOMPRIMIRANOG VIDEO SIGNALA 10

3.1 Baza testnih video signala 10

3.1.1 Mjerenje prostorne opažene informacije 12

3.1.2 Mjerenje vremenske opažene informacije 12

3.2 Subjektivni eksperimenti ocjene kvalitete video signala 14

3.3 Objektivno ocjenjivanje kvalitete video signala 15

3.4 Usporedba rezultata objektivne i subjektivne ocjene kvalitete videa 17

4. ZAKLJUČAK 21

LITERATURA 22

SAŽETAK 24

ABSTRACT 25

ŽIVOTOPIS 26

1. UVOD

Tema ovog završnog rada je maskiranje u video signalu. Osim što će se govoriti o maskiranju općenito, također će biti govora i o povezanosti maskiranja s ljudskim vizualnim sustavom te na koji način maskiranje utječe na percepciju izobličenja u videu. Analizirat će se utjecaj maskiranja na kvalitetu prikazanog video signala pomoću eksperimenata u kojima će se dobiti subjektivno mišljenje o kvaliteti nekolicine različitih video signala. Bit će predstavljeni i neki osnovni tipovi maskiranja koji se mogu pojaviti kod prikaza video signala. U prvom poglavlju opisan je zadatak završnog rada. U drugom poglavlju objašnjava se teorija maskiranja općenito kao i osnovni tipovi maskiranja. U trećem poglavlju analizira se kako učinci maskiranja djeluju na kvalitetu komprimiranog video signala te se prikazuju postignuti rezultati. Četvrto poglavlje donosi zaključak.

1.1 Zadatak završnog rada

U radu je potrebno općenito opisati pojam maskiranja te analizirati različite tipove maskiranja koji se pojavljuju kod prikaza video signala. Potrebno je na primjerima video signala različitih sadržaja te različite prostorne i vremenske aktivnosti prikazati učinke maskiranja i analizirati njihov utjecaj na doživljaj kvalitete prikazanog video signala od strane gledatelja.

2. MASKIRANJE U VIDEO SIGNALU

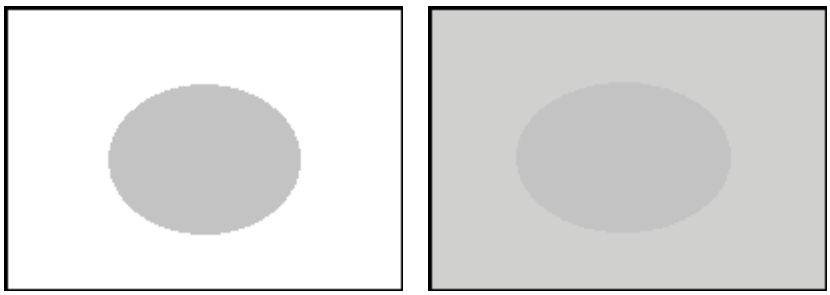
Maskiranje se manifestira kao smetnja kod promatranja ciljnog podražaja zbog umješanosti drugog maskirnog podražaja koji se u prostoru i vremenu nalazi blizu ciljnog podražaja. Omogućuje smanjenje kvalitete nekog videa ili slike koje je teže opaziti. Naravno, zbog načina na koji maskiranje funkcionira, vrlo je važno kojim se područjima u slici ili videu smije smanjiti kvaliteta, a da ta promjena gledateljima ostane neprimjetna. Tu se dolazi do povezanosti maskiranja i ljudskog vizualnog sustava. Na određena područja u slici ljudski vizualni sustav je osjetljiviji dok je na druga manje osjetljiv. O tome će nešto više govora biti u sljedećim podpoglavljima. Postoji nekoliko vrlo različitih pojava koje se vode pod pojmom „maskiranja“. Najčešće istraživanje maskiranja koristi vrlo kratak ciljni podražaj koji je u prostoru i vremenu popraćen tipično kratkim ili malo dužim maskirnim podražajem. Pritom se radi o vrlo kratkim vremenskim intervalima (0-200ms) koji odvajaju početak ciljnog i maskirnog podražaja (asinkronija početka podražaja – SOA) [1]. Asinkronija početka podražaja je mjera koja označava količinu vremena između početka jednog i početka drugog podražaja. Maskirni podražaj predstavljen je prije cilja u maskiranju unaprijed, nakon cilja u maskiranju unazad i tijekom cilja u istovremenom maskiranju. Prema svojstvima prostornog rasporeda maske, maskiranje se može podijeliti na maskiranje svjetlošću, maskiranje šumom, maskiranje uzorkom, maskiranje objektom ili metakontrast/parakontrast maskiranje. Dok se u prethodnim slučajevima maska obično preklapa s pozicijom cilja, metakontrast predstavlja raznovrsnost maskiranja unazad gdje se slike maske i cilja ne preklapaju u prostoru već su u prostoru jako blizu jedna drugoj. Parakontrast se odnosi na maskiranje unaprijed od strane prostorno ne-preklapajuće susjedne maske.

Osnovna podjela maskiranja bila bi podjela na prostorno i vremensko maskiranje. Prostorno maskiranje se odnosi na maskiranje unutar jedne individualne slike, gdje dio te slike utječe na percepciju drugog dijela te iste slike. Kod vremenskog maskiranja slike su predstavljene slijedno u vremenu te se stapaju jedna u drugu. Učinak maskiranja može biti smanjena svjetlina na slici ili narušena vidljivost pojedinih dijelova na slici. Maskiranje svjetline objašnjava se Weberovim zakonom koji govori kako je percepcija svjetline nelinearna funkcija svjetline. Ako u pozadini uniformne (jednolične) svjetline L neki dio ima drugačiju svjetlinu, $L+\Delta L$, prag vidljivosti odnosno najmanja vrijednost ΔL potrebna kako bi promatrač uočio razliku u intenzitetu raste sa svjetlinom u pozadini L [2]:

$$\Delta L = k * L \tag{2-1}$$

gdje je k približno 0.33.

Na slici 2.1 može se vidjeti jednobojni objekt s intenzitetom sive boje u obliku kruga koji se nalazi na pozadini s različitim intenzitetima sive boje (a i b). U ovom slučaju promatra se utjecaj jednog podražaja (pozadine odnosno maske) na vidljivost drugog podražaja (kruga). Moguća su dva slučaja. Kada je razlika između tih dvaju intenziteta sive velika, ljudski vizualni sustav nema poteškoća kod razlikovanja, odnosno krug se primjećuje. S druge strane kada su intenziteti sive boje kruga i pozadine približno isti ili isti, ljudski vizualni sustav ima poteškoća sa primjećivanjem kruga na pozadini ili ga uopće ne vidi. Prag vidljivosti koji označava trenutak u kojem se krug i pozadina počinju razlikovati objašnjen je upravo Weberovim zakonom. [3]



(a) (b)

Slika 2.1 Objekt u obliku kruga i pozadina (a) bijele boje i (b) sive boje

Maskiranjem u video signalu bavio se Turvey koji je 1973. god. u svojemu radu [4] izvršio niz eksperimenata. U jednom od eksperimenata ispitivao je dvije glavne varijable: vremenski interval između cilja i maske i relativnu „energiju“ (svjetlinu) dvaju podražaja. Ljudi su trebali identificirati slova s kratko prikazanih trigrama koji su se sastojali od suglasnika (točnije niz od 3 suglasnika, RBL – cilj). Svaki cilj bio je praćen maskom koja se sastojala od linija koje su bile slične dužine, orijentacije i širine onima u cilju. Maska je prikazana na način da se preklapa s lokacijom gdje su se pojavila ciljna slova. Cilj i maska bili su prikazani samo 10ms, a broj prijavljenih slova sa svakog trigrama mjerio se u odnosu na vremenski interval između početka cilja i početka maske. U jednom slučaju cilj je bio dvostruko svjetliji od maske (C:M omjer 2:1), dok je u drugom maska bila dvostruko svjetlija ond cilja (C:M omjer 1:2). Postojale su osjetne razlike u rezultatima između ta dva stanja. Kada je cilj imao manju energiju od maske (C:M = 1:2) i kada je bilo samo kratko razdoblje između podražaja, performanse su bile izrazito loše. Kako se SOA povećavala tako su se i performanse monotono poboljšavale. Taj slučaj se može nazvati monotona maskirajuća funkcija. Ta maskirajuća funkcija pokazuje da što se više povećava vrijeme između cilja i pojave maske povećava se i sposobnost identifikacije cilja. U drugom slučaju kada je cilj imao veću energiju od maske (C:M=2:1), gledatelji su točno prijavili cilj tamo gdje je SOA bila kratka. Međutim kako se interval (razmak) povećavao, prijave cilja su postale teže da bi se

kasnije kako je interval cilj-maska postajo duži opet poboljšale. Drugim riječima maskirajuća funkcija je u obliku slova „U“ te se ovaj rezultat smatra protu-intuitivnim (engl. *counter-intuitive*) odnosno neočekivanim. Breitmeyer i Ganz su 1976. god. u svome radu [5] raspravljali kako se između maski i ciljeva mogu pojaviti različite interakcije, kako unutar procesirajućeg kanala tako i preko kanala. To stvara različite efekte maskiranja. Na primjer, kada imamo relativno kratke intervale između maski i ciljeva, prolazni i neprekidni signali aktivirani od strane dva podražaja preklapati će se u vremenu. 1978. god. Taylor i Chabot su u svome radu [6] proveli zanimljivo istraživanje. Zatražili su od gledatelja da identificiraju ili jedno kratko prikazano slovo ili cijelu riječ. Slovo ili riječ bili su praćeni jednim od tri tipa maske: svjetlosni bljesak, maska sastavljena od nasumično orijentiranih fragmenata (dijelova) slova (maskiranje po uzorku) ili neka nepovezana riječ. Taylor i Chabot mijenjali su trajanje cilja i mjerili kritični ISI između cilja i maske koji je bio potreban gledateljima kako bi identificirali cilj. Došli su do zanimljivih rezultata. Kritični ISI (engl. *interstimulus interval* – vremenski interval između kraja jednog podražaja i početka drugog) potreban za identifikaciju cilja obično se povećavao kod maske po uzorku u odnosu na masku svjetlosnog bljeska te za masku rječju u odnosu na masku po uzorku. Prema tome maskiranje je bilo jače što je struktura maske više poprimala izgled uzorka – svjetlosni bljesak bio je najslabija maska, zatim je slijedio uzorak sa fragmentima slova i na kraju je bila riječ [7].

2.1 Prostorno maskiranje

Prostorno maskiranje ukazuje na činjenicu da se prag diskriminacije povećava s povećanjem detalja u slici. Drugim riječima manje je vjerojatno da će se greška na slici uočiti na područjima gdje ima puno detalja nego što je to slučaj kod područja s malo detalja. Razlog tome leži u karakteristici ljudskog vizualnog sustava, točnije oka, koje je puno osjetljivije na „glađa“ i ujednačena područja na slici (nebo na slici 2.2) dok je na „grublja“ područja s puno detalja (stijene na slici 2.2) znatno manje osjetljivo. Pomoću ove tri slike vrlo je jednostavno objasniti prostorno maskiranje. Prva slika je original. Na drugoj slici šum (smetnja) je dodan na gornji dio slike gdje ga je vrlo lako uočiti zbog jako malo detalja. Na posljednjoj slici jednaka količina šuma dodana je na donji dio slike gdje ima puno detalja pa ga je jako teško uočiti. Prostornim maskiranjem bavio se u Forster 1982. god. u svome radu [9]. Forster je od gledatelja zahtijevao da identificiraju ciljne fotografije koje su bile praćene maskom nasumičnog šuma filtriranom tako da s fotografije ukloni ili njene visoke prostorno-frekvencijske komponente (niska prostorno-frekvencijska maska) ili one niske prostorno-frekvencijske komponente (visoka prostorno-frekvencijska maska). Otkrio je da



(a) (b) (c)

Slika 1.2 Primjer prostornog maskiranja prikazan pomoću triju slika (a) originalna slika, (b) slika sa šumom dodanim u gornji dio slike, (c) slika sa šumom dodanim u donji dio slike. [8]

visoko-frekvencijska maska proizvodi monotoničnu maskirajuću funkciju dok nisko-frekvencijska maska proizvodi funkciju u obliku slova „U“ [8]. Prostorno maskiranje može se uočiti i kod slika s različitim brojem razina kvantizacije koje se koriste za prikaz slike. Povećanjem broja razina kvantizacije odnosno broja bita kojim se prikazuje svaki element slike na slici, povećava se i kvaliteta slike. Slika 2.3 koristi 256 razina kvantizacije odnosno 8 bita za prikaz svakog elementa slike, dok slika 2.4 koristi samo 16 razina kvantizacije odnosno 4 bita. Isto kao i kod uspoređivanja prethodne tri slike sa slike 2.2, lošija kvaliteta najviše se primjeti kod ujednačenih područja u slici, odnosno na zidu u pozadini. S druge strane kad se pogledaju područja s više detalja, kao npr. lica i odjeća kod ljudi, lošija kvaliteta je puno manje primjetna. Kako bi se dobila što bolja kvaliteta uz što manje bita za prikaz pojedinog elementa slike, slika se kvantizira na način da se više bita koristi za područja s manje detalja na koje je oko osjetljivije, dok se za manje osjetljiva područja s više detalja koristi manji broj bita. Učinci prostornog maskiranja obično se mjere tako da se izmjeri prag detekcije za ciljni podražaj u trenutku kada se on preklapa sa maskirnim podražajem koji ima promjenjiv kontrast. Maskiranje je najjače kada podražaji koji su u interakciji imaju



Slika 2.3 Originalna slika [3]



Slika 2.4 Slika smanjenje kvalitete na kojoj se koristila 4-bitna kvantizacija [3]

slične karakteristike odnosno slične frekvencije, orijentacije, boje i slično. Iako općenito slabije, maskiranje se također događa između podražaja s različitom orijentacijom, različitom prostornom frekvencijom i između kromatskih i nekromatskih podražaja [8].

2.2 Vremensko maskiranje

Vremensko maskiranje može se uočiti kod dva podražaja koji se prezentiraju slijedno odnosno jedan za drugim. Naime, svaki se video sastoji od niza okvira koji se izmjenjuju određenom brzinom. Ti okviri se razlikuju u nekim sitnim detaljima. Postoji mogućnost da neposredno prije i neposredno nakon izmjene okvira dođe do nekih nepravilnosti. Ljudsko oko nije u mogućnosti primjetiti te nepravilnosti upravo zbog toga što je brzina izmjene okvira prevelika. Ta pojava naziva se vremensko maskiranje. Nepravilnosti postoje ali se ne primjete. Postoje dvije glavne vrste vremenskog maskiranja. Kada dođe do nagle promjene scene, prođe dosta vremena prije nego se oko privikne na nove detalje. Tijekom tog vremena postoji mogućnost da se glavni artifakti (smetnje) na prikazanoj slici ne vide. Ta pojava naziva se maskiranje unaprijed [10]. Maskiranje unazad je pojava u kojoj prikazivanje jednog vizualnog podražaja (maske ili maskirnog podražaja) odmah nakon drugog kratkog ($\leq$ 50 ms) ciljnog vizualnog podražaja dovodi do neuspjeha da se svjesno opazi prvi podražaj. Nepoznato je kako kasniji podražaj može blokirati raniji. Međutim, jedna teorija za ovu pojavu, poznata kao teorija interakcije dvostrukog kanala predlaže da brzi signal stvoren drugim podražajem može dostići i prevladati sporiji signal poslan iz prvog impulsa. Iako svjesno neopažen, maskirni podražaj može ipak imati učinak na spoznajne procese [11]. Kao što je već i rečeno, svaki video sastoji se od određenog broja okvira (slika) koji se izmjenjuju određenom brzinom. Vremensko maskiranje može se objasniti promatranjem dvaju različitih video signala. Na slici 2.5 prikazan je jedan zaustavljeni okvir iz sporog video signala koji nema naglih odnosno brzih pokreta, dok je na slici 2.6 prikazan jedan zaustavljeni okvir iz video signala koji predstavlja brzi video s puno brzih pokreta. Kao što se može i vidjeti, zaustavljeni okvir iz sporog video signala ne sadrži nikakva izobličenja. S druge strane, kod zaustavljenog okvira iz brzog video signala može se primjetiti znatno pogoršanje u kvaliteti koje se manifestira zamućenjem. Vremensko maskiranje prisutno je u oba slučaja međutim kod brzog video signala ga je znatno lakše uočiti zbog toga što ljudsko oko ne stigne pratiti sve promjene koje se dogode u jednoj sekundi brzog videa (promjeni se između 20 i 30 okvira odnosno slika u sekundi). Tek kada bi se video zaustavio na jednom okviru (kao što se može vidjeti na slikama) primjetila bi se zamućenja koja su ljudskom oku neprimjetna dok je video u pokretu. Prema tome, može se reći da je video signal vremenski maskiran.



Slika 2.5 Uhvaćeni okvir iz video signala sa sporim sadržajem



Slika 2.6 Uhvaćeni okvir iz video signala s brzim sadržajem

Vremenskim maskiranjem bavili su se mnogi znanstvenici. U radu [12] iz 1965. god. Seyler i Budrikis su došli do zanimljivog otkrića. Koristili su prolazni niskopropusni filter kako bi smanjili propusnost i vidjeli subjektivni odgovor na smanjene prostorne detalje na slikama nakon

reza scene. Uz pretpostavku da se slika vratila na punu propusnost za 780 ms, gledatelji su ocijenili slike kao dovoljno dobre. To je značajno više u usporedbi sa rezultatom 50-100ms koji je dobio Turvey 1973. god. Međutim, važno je napomenuti da su Seyler i Budrikis koristili subjektivni kriterij ocjenjivanja umjesto objektivnog. Girod je 1989. god. u svome radu [13] istraživao uštede brzine prijenosa bita omogućene prostornim i vremenskim maskiranjem u video signalima. Otkrio je da je vremensko maskiranje unaprijed bilo značajno samo u prvih 100 ms nakon reza scene. Također je istraživao i vremensko maskiranje uslijed kretanja slike. Međutim otkrio je kako je to beznačajno zbog sposobnosti praćenja koje posjeduje vizualni sustav. Tam, Stelmach, Wang, Lauzon i Gray su 1995. godine u svojem radu [14] proučavali maskiranje unaprijed nakon reza scene. Uspostavili su psihofizičke pragove vidljivosti za oštećenja u prvim, drugim i trećim okvirima koji su slijedili rez scene. Otkrili su da se učinci maskiranja pojavljuju uglavnom u prvom okviru koji prati rez scene te da su ti isti učinci skoro neprimjetni u trećem okviru. Pored toga otkrili su da su oštećenja u prvom okviru (iako ispod praga) smanjila razinu praga u drugom okviru. 1996. godine Carney je skupa s ostalim znanstvenicima u svome radu [15] primjetio snažnu ovisnost kod polariteta podražaja. Otkrili su da je učinak maskiranja puno izraženiji kada su ciljni i maskirni podražaj istog polariteta. Također su otkrili da je maskiranje najjače za lokalne prostorne konfiguracije [16].

3. ANALIZA UTJECAJA MASKIRANJA NA KVALITETU KOMPRIMIRANOG VIDEO SIGNALA

Kvaliteta komprimiranog video signala prije svega ovisi o stupnju kompresije. Što je stupanj kompresije veći kvaliteta je lošija. Osim stupnja kompresije na kvalitetu utječe i maskiranje. Kako bi se proučio utjecaj maskiranja na kvalitetu komprimiranog video signala usporedit će se kvaliteta video signala različitih sadržaja, ali približno iste kvalitete odnosno približno istog vršnog omjera signala i šuma (engl. *Peak signal-to-noise-ratio* - PSNR).

3.1 Baza testnih video signala

Baza testnih video signala sastoji se od 8 neoštećenih, odnosno nekomprimiranih originalnih video signala različitog sadržaja. Na slici 3.1 prikazan je po jedan uhvaćeni okvir iz svakog video signala. Originalni signali imaju rezoluciju 352x288 elemenata slike te ne sadrže zvuk. Okviri se izmjenjuju brzinom od 25 okvira u sekundi, a s obzirom da svaki video signal traje 12 sekundi dolazi se do brojke od ukupno 300 okvira za svaki originalni video signal. Kako bi se što bolje ispitale mogućnosti metoda za mjerenje razlika kod obrade slike i videa, izabrano je 8 video signala vrlo različitog sadržaja. Osim po sadržaju, video signali se razlikuju i po prostornoj i vremenskoj aktivnosti (opaženoj informaciji) te po količini detalja i boja u slici. Izabrani video signali mogu se također podijeliti na one s pokretnom i nepokretnom kamerom [17]. Kratki opis svakog video signala nalazi se u tablici 3.1.

Tablica 3.1 Opis testnih video signala

Puni naziv video signala i skraćeni naziv u zagradi	**Opis video signala**
Container ship (cos)	Teretni brod sporo plovi, nepokretna kamera
Flower garden (flg)	Pokretna kamera prikazuje vrt s cvijećem, puno detalja i boja
Football (ftb)	Dio utakmice američkog nogometa uz brze pokrete i pokretnu kameru
Foreman (frm)	U prvom planu je glava radnika koji priča, pokretna kamera pomjera se na još neizgrađenu zgradu
Hall (hal)	Dvije osobe mimoilaze se u hodniku, nepokretna kamera
Mobile (mob)	Prikaz više različitih objekata od kojih se svaki kreće u drugom smjeru, pokretna kamera uz puno detalja
News (nws)	Dvije osobe čitaju vijesti dok se u pozadini kreću objekti, nepokretna kamera
Silent (sil)	Nepokretna kamera prikazuje osobu koja gestikulira rukama



(a) (b)



(c) (d)



(e) (f)



(g) (h)

Slika 3.1 Uhvaćeni okviri iz originalnog video signala iz baze. (a) cos, (b) flg, (c) frm, (d) ftb, (e) hal, (f) mob, (g) nws, (h) sil [17]

Količina aktivnosti u slici računa se pomoću prostorne i vremenske opažene informacije.

3.1.1 Mjerenje prostorne opažene informacije

Prostorna opažena informacija (engl. *spatial perceptual information* - SI) je mjera koja obično ukazuje na količinu prostornih detalja u nekoj slici. Obično je viša za prostorno složenije scene. Nije namijenjena da bude mjera entropije niti da bude povezana s informacijom definiranom u teoriji komunikacija [18].

Prostorna opažena informacija temelji se na Sobelovom filtru. Svaki okvir videa u vremenu *n* (F_n) prvo se filtrira Sobelovim filterom [$Sobel(F_n)$]. Zatim se preko elemenata slike u svakom okviru filtriranom Sobelovim filtrom računa standardna devijacija (std_{space}). Ova operacija ponavlja se za svaki okvir u video sekvenci (nizu) i rezultira u vremenskom nizu prostornih informacija scene. Bira se maksimalna vrijednost u vremenskom nizu (max_{time}) koja predstavlja sadržaj prostornih informacija scene. Ovaj proces može se zapisati u obliku jednadžbe:

$$SI = max_{time}\{std_{space}[Sobel(F_n)]\} \tag{3-1}$$

3.1.2 Mjerenje vremenske opažene informacije

Vremenska opažena informacija (engl. *temporal perceptual information* - TI) je mjera koja obično ukazuje na količinu vremenskih promjena video sekvence (niza). Obično je viša za sekvence s puno pokreta. Nije namijenjena da bude mjera entropije niti da bude povezana sa informacijom definiranom u teoriji komunikacija.

Vremenska opažena informacija temelji se na svojstvu razlike pokreta, $M_n(i,j)$, koja je zapravo razlika između vrijednosti elemenata slike na istoj lokaciji u prostoru, ali u uzastopnim okvirima. $M_n(i,j)$ kao funkcija vremena (n) definirana je kao:

$$M_n(i,j) = F_n(i,j) - F_{n-1}(i,j) \tag{3-2}$$

gdje je $F_n(i,j)$ element slike na *i*-tom retku i *j*-tom stupcu *n*-tog okvira u vremenu. Mjera vremenske informacije (TI) računa se kao:

$$TI = max_{time}\{std_{space}[M_n(i,j)]\} \tag{3-3}$$

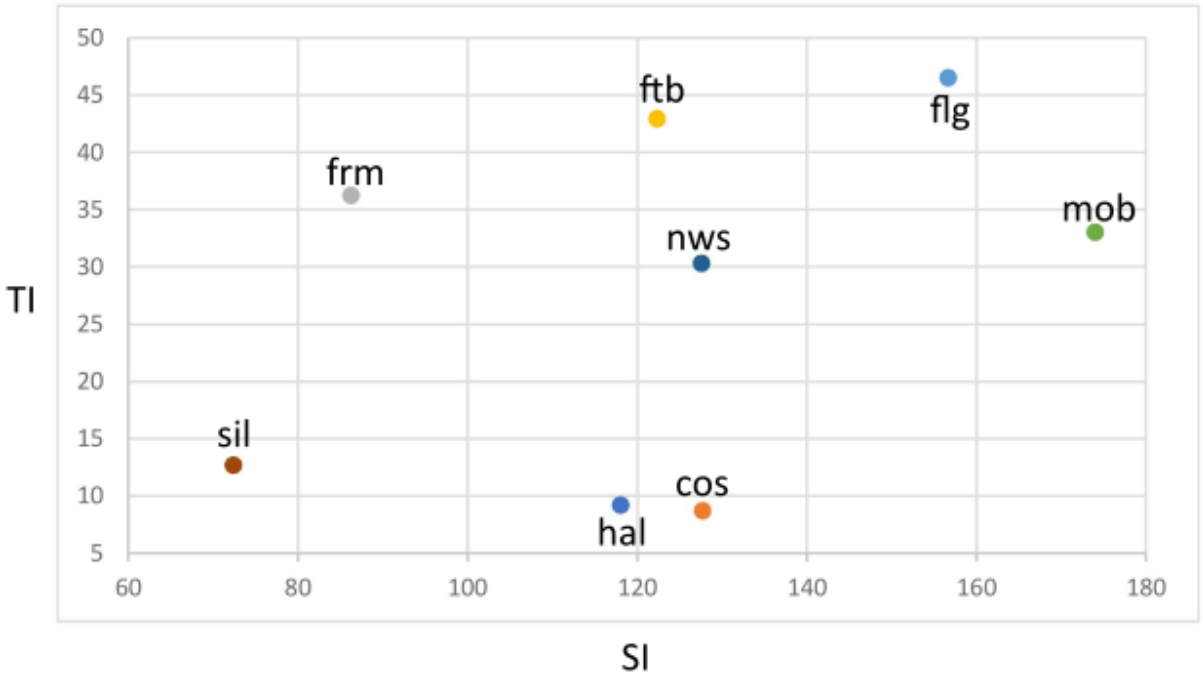
Više pokreta u susjednim okvirima rezultirat će većim vrijednostima vremenske opažene informacije.

Napomena – za scene koje sadrže rezove scena, mogu se dati dvije vrijednosti: jedna gdje je rez scene uključen u mjeru vremenske informacije i jedan gdje je iz te iste mjere isključen. Vrijednosti prostorne i vremenske opažene informacije originalnih video signala prikazane su u tablici 3.2.

Tablica 3.2 Vrijednosti prostorne i vremenske opažene informacije (aktivnosti) originalnih video signala

Izvorni video signal	**cos**	**flg**	**frm**	**ftb**	**hal**	**mob**	**nws**	**sil**
SI	127.72	156.67	86.26	122.33	118.04	174.03	127.60	72.40
TI	8.70	46.51	36.25	42.94	9.19	33.03	30.33	12.68

Na slici 3.2 dan je SI-TI graf za sve originalne video signale iz baze. Može se vidjeti da je odabranim originalnim video signalima dobro popunjen cijeli SI-TI prostor.



Slika 3.2 Graf s vrijednostima prostorne i vremenske opažene informacije (aktivnosti) originalnih video signala

3.2 Subjektivni eksperimenti ocjene kvalitete video signala

Subjektivno ocjenjivanje kvalitete video signala provedeno je na način da je 5 gledatelja bilo zaduženo za ocjenjivanje 24 od ukupno 32 video signala. Radi se o 8 različitih izvornih nekomprimiranih video signala od kojih svaki ima 3 komprimirane kopije koje se od izvornog video signala razlikuju samo po ciljanoj brzini prijenosa odnosno po kvaliteti. Svaka od 3 komprimirane kopije imala je drugačiju ciljanu brzinu prijenosa (64 kbit/s, 256 kbit/s i 768 kbit/s). Gledatelji su ocjenjivali video signale na način da su prvo pogledali izvorni video signal, a zatim i njegove 3 komprimirane kopije. Međutim, kako bi se što bolje proučio utjecaj maskiranja gledatelji prilikom ocjenjivanja nisu znali ciljane brzine prijenosa, odnosno nisu znali u kojoj mjeri je pojedina komprimirana kopija oštećena. O ocjenama gledatelja detaljnije će se govoriti u podpoglavlju 3.4. Rezultat srednje subjektivne ocjene (engl. *Mean Opinion Score* - MOS) je najčešće korištena subjektivna mjera kvalitete, gdje gledatelji ocjenjuju testni video signal pomoću fiksne skale. Skala je najčešće u rasponu 0-100. Za svaku razinu jedna prateća riječ opisuje kvalitetu videa. Sljedeće ocjene opisuju kvalitetu video signala.

80 - 100 = izvrsno

60 - 80 = dobro

40 - 60 = prosječno

20 - 40 = loše

0 - 20 = neprihvatljivo

Gledateljima se prezentiraju video signali jedan po jedan i traži se njihovo mišljenje o kvaliteti videa u skladu s prethodno objašnjenim uvjetima. Test se izvodi za svaki video s više gledatelja. Veličina grupe gledatelja kreće se u rasponu 15-20 pa do preko 40. Kao što je to slučaj sa svim testovima ove prirode, veća grupa gledatelja pruža pouzdanije rezultate. Sredina subjektivnih ocjena za svaki video računa se i zapisuje kao MOS za taj video. Kako bi se ocijenio sustav kodiranja videa, test uključuje veliki broj signala. Originalni nekodirani video signal trebao bi postići ocjenu 100. Većina sustava kodiranja videa postiže ocjenu između 40 i 80. Sustav kodiranja koji postiže ocjenu veću od 80 pruža jako visoku kvalitetu. Rezultati MOS testa međutim mogu jako varirati. Mišljenja o kvaliteti nekih video signala jako variraju kod različitih gledatelja pa čak i za istog gledatelja, ali neki drugi dan [19].

3.3 Objektivno ocjenjivanje kvalitete video signala

Srednja kvadratna pogreška (engl. *Mean squared error* - MSE) i PSNR su najpopularnije metode za objektivno mjerenje kvalitete kod obrade slike i videa. MSE je srednja vrijednost kvadratnih razlika između vrijednosti razina sive elemenata slike na dvije slike ili sekvence:

$$MSE = \frac{1}{TXY}\sum_t \sum_x \sum_y [I(t,x,y) - I_i(t,x,y)]^2 \tag{3-4}$$

za slike rezolucije $X \times Y$ elemenata slike i T okvira u sekvenci.

PSNR se definira kao:

$$PSNR = 10 \log \frac{m^2}{MSE} \tag{3-5}$$

gdje je *m* maksimalna vrijednost koju element slike može imati (npr. 255 za 8-bitne slike). Važno je imati na umu da su MSE i PSNR dobro definirane samo za informacije o osvjetljenju: kada se pojavi boja nema nekog dogovora oko računanja tih mjera [8]. MSE mjeri razliku slike dok PSNR mjeri vjernost slike odnosno koliko točno slika liči na referentnu sliku odnosno neoštećeni original. Popularnost ovih dviju metrika temelji se na činjenici da je svođenje MSE na minimum ekvivalentno optimizaciji najmanjih kvadrata u minimalnom energetskom smislu, a za koje su dostupni dobro poznati matematički alati. Osim toga računanje MSE i PSNR je vrlo brzo i lagano. Međutim, pošto su zasnovane na usporedbi slika element po element slike, imaju samo ograničenu približnu vezu s distorzijom ili kvalitetom opaženom od strane ljudskog vizualnog sustava. U određenim situacijama subjektivna kvaliteta slike može se poboljšati dodavanjem šuma čime se smanjuje PSNR. Osim toga vidljivost deformacija u velikoj mjeri ovisi o pozadini slike. Deformacije su često više uznemirujuće u relativno mirnim područjima slike nego što je to slučaj s područjima s teksturama i s puno aktivnosti, što je opisano u ranijim poglavljima. Problem je što se to ne uzima u obzir kod metrika zasnovanih na elementu slike. Stoga opažena kvaliteta slika s istim PSNR-om može ustvari biti vrlo različita. Primjer problema kod korištenja PSNR-a kao pokazatelja kvalitete prikazan je na slici 3.3.

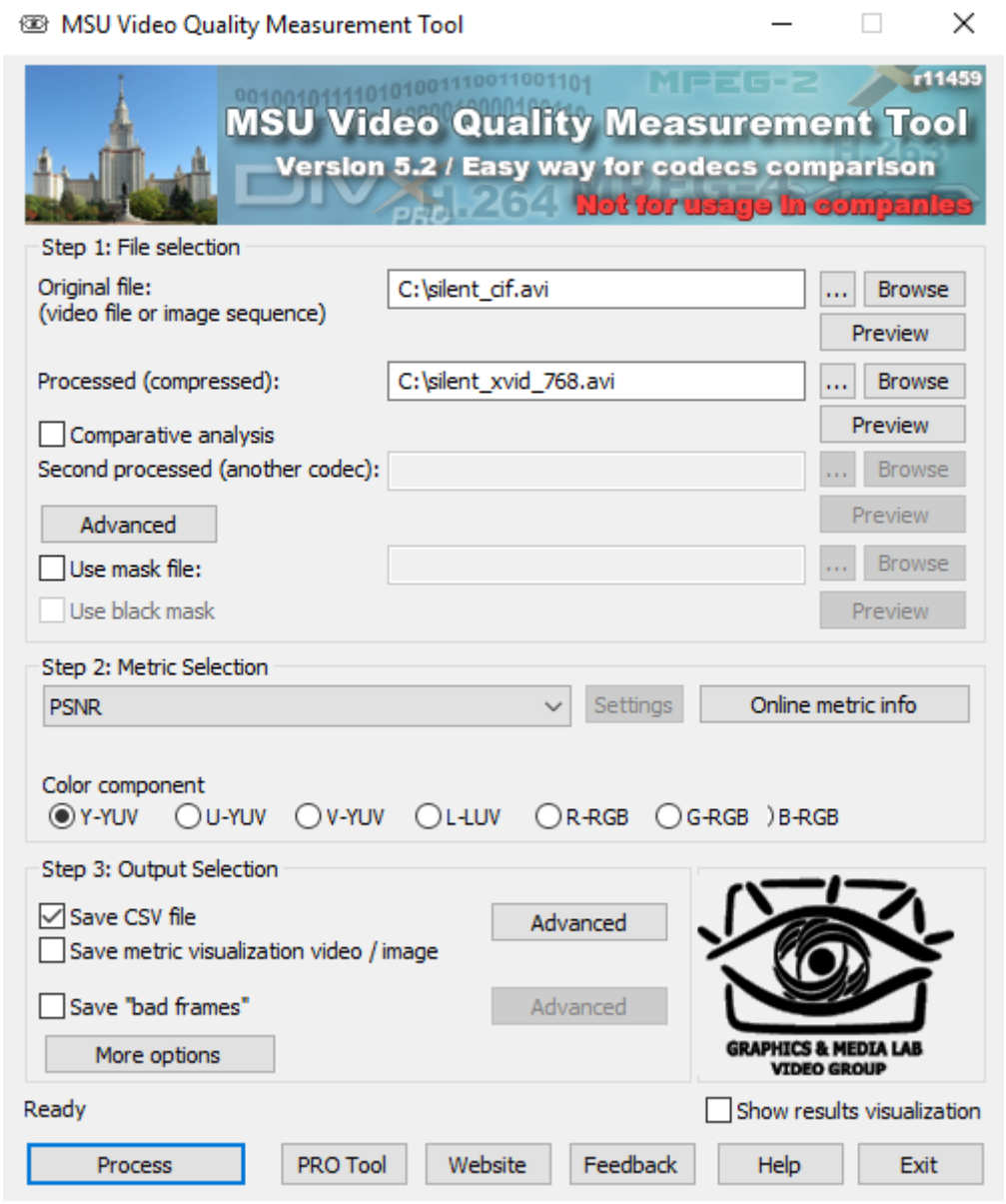


(a) (b) (c)

Slika 3.3 PSNR kod šuma na različitim područjima na slici. (a) originalna slika, (b) slika sa šumom dodanim u gornji dio uz PSNR = 32dB, (c) slika sa šumom dodanim u donji dio uz PSNR = 32dB [5]

Ista količina šuma umetnuta je u slike (b) i (c) tako da je njihov PSNR u odnosu na izvornu sliku (a) identičan. Pojasno-propusni filtrirani šum umetnut je u gornji dio slike (b), dok je visoko-frekvencijski šum umetnut u donji dio slike (c). Ljudska osjetljivost na strukturirani (nisko-frekvencijski) šum na slici (b) već je dosta visoka i šum se jasno vidi na mirnoj pozadini neba. Šum na slici (c) jedva je prepoznatljiv zbog niske osjetljivosti ljudskog vizualnog sustava na visoko-frekvencijske podražaje i zbog jakog maskiranja od strane visoko teksturiranog sadržaja u donjem području. PSNR nije u mogućnosti razlikovati nijedan od ovih učinaka.

Za objektivno ocjenjivanje kvalitete videa PSNR metrikom u ovom završnom radu korišten je alat *MSU Video Quality Measurement Tool* [20]. Pomoću tog alata provedena je usporedba originalnih nekomprimiranih video signala i onih komprimiranih te je tako dobivena PSNR ocjena za svaki od 24 izobličena video signala. Srednje objektivne ocjene nalaze se u podpoglavlju 3.4 u tablici 3.3. Na slici 3.4 prikazan je glavni prozor *MSU Video Quality Measurement Tool* alata.



Slika 3.4 Glavni prozor *MSU Video Quality Measurement Tool* alata [20]

3.4 Usporedba rezultata objektivne i subjektivne ocjene kvalitete videa

Rezultati subjektivnih eksperimenata, kao i PSNR-a dani su u tablici 3.3. Kod subjektivne ocjene kvalitete video signala najniža srednja zabilježena ocjena bila je 29 za video signal „ftb_64“ s ciljanom brzinom prijenosa od 64 kbit/s. Video signal je prikazivao jedan dio utakmice američkog nogometa uz puno brzih pokreta. Najviša ocjena 77 pripala je video signalu „frm_768“ kojemu je ciljana brzina prijenosa iznosila 768 kbit/s, a prikazivao je u svom većem dijelu glavu radnika koji nešto govori. Na tom video signalu nije bilo puno detalja niti brzih pokreta. Međutim, kod objektivne ocjene kvalitete video signala rezultati su nešto drugačiji, samim time što se za

dobivanje rezultata koristio programski alat odnosno sve se proračunavalo isključivo pomoću računala. Najniža ocjena je 24.11 dB koja pripada video signalu „mob_64“ ciljane brzine prijenosa od 64 kbit/s koji je prikazivao više različitih objekata koji se kreću svaki u drugom smjeru. Najvišu ocjenu 40.10 dB dobio je video signal „nws_768“ uz ciljanu brzinu prijenosa od 768 kbit/s koji je prikazivao dvoje ljudi koji čitaju vijesti dok se u pozadini vide kretnje. Ostale subjektivne i objektivne ocjene mogu se vidjeti u tablici 3.3.

Tablica 3.3 Srednje subjektivne (MOS) i objektivne (PSNR) ocjene kvalitete video signala

Video signal	MOS	PSNR [dB]
cos_64	52	31.31
cos_256	69	35.41
cos_768	71	37.51
flg_64	36	25.26
flg_256	57	26.95
flg_768	73	31.16
frm_64	30	29.06
frm_256	63	33.65
frm_768	77	37.76
ftb_64	29	27.64
ftb_256	46	28.52
ftb_768	71	32.34
hal_64	41	30.95
hal_256	63	32.80
hal_768	76	33.47
mob_64	47	24.11
mob_256	59	26.86
mob_768	72	30.72
nws_64	42	32.08
nws_256	75	37.28
nws_768	76	40.10
sil_64	34	32.84
sil_256	68	36.23
sil_768	72	40.07

Usporedbom više parova od po dva različita video signala koji imaju gotovo jednaku objektivnu ocjenu odnosno gotovo jednaki PSNR, dolazi se do zanimljivih rezultata. Naime, takvi video signali zbog različitog utjecaja maskiranja koje alat za mjerenje srednje ocjene PSNR-a nije u stanju razlikovati dobili su znatno drugačije subjektivne ocjene. Jedan takav primjer može se opisati kod video signala „cos_64“ i „flg_768“. Prvi video signal ima ciljanu brzinu prijenosa 64 kbit/s te prikazuje teretni brod koji plovi morem. Gledatelji su lako mogli uočiti lošiju kvalitetu uglavnom zbog male ciljane brzine prijenosa te zbog malog iznosa prostornog maskiranja zbog kojeg se na mirnoj pozadini koju je u ovom videu predstavljalo more vide jasna izobličenja. Srednja subjektivna ocjena za ovaj video signal bila je 52. Ciljana brzina prijenosa drugog video signala iznosila je 768 kbit/s, a prikazuje vrt sa cvijećem. Zbog velike ciljane brzine prijenosa, a i zbog prostornog maskiranja kod dosta strukturiranog videa sa puno detalja gledatelji ovoga puta nisu uspjeli uočiti neke veće smetnje u video signalu pa je srednja subjektivna ocjena ovoga puta bila 73. Treba podsjetiti da je računalo dalo iste ocjene za oba videosignala, iako se zbog količine maskiranja njihove kvalitete drastično razlikuju. Kod drugog primjera uspoređuju se video signali „ftb_768“ i „nws_64“. Kod prvog video signala ciljana brzina prijenosa iznosi 768 kbit/s, a video signal prikazuje dio utakmice američkog nogometa. Ovoga puta video signal je dobio dosta dobru ocjenu zbog vremenskog maskiranja odnosno zbog toga što sadrži puno brzih pokreta pa ljudsko oko ne stigne uočiti smetnje dok se video signal ne zaustavi. Naravno pomogla je i velika ciljana brzina prijenosa. Srednja subjektivna ocjena za ovaj video signal bila je 71. Kod drugog video signala ciljana brzina prijenosa iznosi 64 kbit/s, a u prvom planu video signala su dvije osobe koje čitaju vijesti dok se u pozadini kreću objekti. Ovaj video signal dobio je lošu ocjenu zbog male ciljane brzine prijenosa i zbog slabog prostornog maskiranja zbog kojeg su gledatelji lako mogli uočiti smetnje na mirnoj pozadini na kojoj su se kretali objekti. Srednja subjektivna ocjena bila je 42.

Drugi tip usporedbe može se napraviti ako se razmotre ocjene za videosignale različitog sadržaja s istim ciljanim brzinama prijenosa. Gledatelji su dali dobre ocjene svim sadržajima kada je ciljana brzina prijenosa bila dovoljno velika, tj. 768 kbit/s. Tada su svi signali dobili ocjene >70. Međutim, razlika se može vidjeti na niskim ciljanim brzinama prijenosa. Kod sekvence „sil_64“, čiji su SI i TI vrlo niski u odnosu na sekvencu „mob_64“ (pogledati sliku 3.2), ocjene gledatelja su znatno niže od onih za „mob_64“. U sekvenci „sil_64“ maskiranje izobličenja je manje nego kod „mob_64“, što su gledatelji znali vrednovati. Za razliku od gledatelja, PSNR je dao puno bolju ocjenu sekvenci „sil_64“, što potvrđuje često puta dokazanu činjenicu da PSNR ne daje rezultate

ocjene kvalitete slične subjektivnim rezultatima i zbog toga, unatoč svojoj jednostavnosti, nije dobar izbor za objektivnu ocjenu kvalitete slike/videa.

4. ZAKLJUČAK

Maskiranje se pokazalo kao vrlo zanimljiv i složen pojam koji je već dugo proučavan, a ovdje ga se nastojalo objasniti na što jednostavniji način. Zanimljiv je zbog različitih načina na koji ljudski vizualni sustav opaža maskiranu sliku ili video ovisno o njihovim karakteristikama i sadržaju. Upravo zbog maskiranja umetnuti šum (smetnja) u slici je na određenim dijelovima jasno vidljiv, dok je na drugim neprimjetan, odnosno kaže se da je maskiran prostorno. S druge strane kad se govori o videu i vremenskom maskiranju može se reći da se zbog prebrze izmjene okvira izobličenja (artefakti) uopće ne primjećuje, odnosno šum je maskiran vremenski. Izobličenja se uočavaju tek kod zaustavljanja videa. Također je zanimljivo da je kod ocjenjivanja kvalitete komprimiranih video signala puno pouzdanija subjektivna metoda odnosno ocjena gledatelja. Tu dolazi do izražaja utjecaj maskiranja koji je lako primjetan kod gledatelja dok je neprimjetan kod objektivne metode, odnosno računala. Međutim, provođenje subjektivne metode je dosta složeno pa se objektivna metoda koristi kad god je to moguće. Eksperimenti provedeni u ovom radu pokazali su da je maskiranje u video signalu izraženije kod video signala s većom prostornom i vremenskom aktivnošću. Također, pokazalo se da su učinci maskiranja izraženiji na nižim brzinama prijenosa kada je izobličenje video signala veće.

LITERATURA

[1] T. Bachmann, G. Francis, Visual Masking: Studying Perception, Attention, and Consciousness, Elsevir, Oxford, 2014.

[2] S. Rimac-Drlje, D. Žagar, G. Martinović, Spatial Masking and Perceived Video Quality in Multimedia Applications, 16th International Conference on Systems, Signals and Image Processing, DOI: 10.1109/IWSSIP.2009.5367719, 2009.

[3] Y. Q. Shi, H. Sun, Image and Video Compression for Multimedia Engineering, Fundamentals, Algorithms, and Standards, Seconde Edition, Taylor & Francis Group, Boca Raton, 2007.

[4] M. T. Turvey, On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli, Psychological Review, 80, 1-52, 1973.

[5] B.G. Breitmeyer, L. Ganz, Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing, Psychological Review, 83-1, 1-36, 1976.

[6] G.A. Taylor, R.J. Chabot, Differential backward masking of words and letters by masks of varying orthographic structure, Mem Cognit 6, 629-35, 1978.

[7] G. W. Humphreys, V. Bruce, Visual Cognition: Computational, Experimental and Neuropsychological perspectives, Lawrence Erlbaum Associates Ltd; London, 1989.

[8] S. Winkler, Digital Video Quality, Vision Models and Metrics, John Wiley & Sons, Ltd; Chichester 2005.

[9] K. I. Forster, An Investigation into the Representations Involved in Visual Masking, 1982.

[10] C. Fogg, D. J. LeGall, J. L. Mitchell, W. B. Pennebaker, MPEG Video Compression Standard, Kluwer Academic Publishers, New York, 2002.

[11] https://en.wikipedia.org/wiki/Backward_masking

[12] A. J. Seyler, Z. L. Budrikis, Detail perception after scene change in television image presentations, IEEE Transactions on Information Theory, January, IT-11-1, 31-43, 1965.

[13] B. Girod, The information theoretical significance of spatial and temporal masking in video signals. Proceedings of SPIE: Human Vision, Visual Processing, and Digital Display, 1077, 178-187, 1989.

[14] W. J. Tam, L. Stelmach, L. Wang, D. Lauzon, P. Gray, Visual Masking at video scene cuts, SPIE, 2411, 111-119, 1995.

[15] T. Carney, S. A. Klein, Q. Hu, Visual masking near spatiotemporal edges, In Proc. SPIE Human Vision and Electronic Imaging, vol. 2657, pp. 393-402, San Jose, CA, 1996.

[16] C. English, Visual Temporal Masking at Video Scene Cuts, Carol English, Ottawa, 1997.

[17] M. Vranješ, Objektivna metrika kvalitete slike zasnovana na prostorno – vremenskim značajkama videosignala i prostorno ovisnoj percepciji, Doktorska disertacija, Fakultet elektrotehnike, računarstva i informacijskih tehnologija, Osijek, 2012.

[18] ITU-T, P.910, 4/2008

[19] R. Goldberg, L. Riek, A Practical Handbook of Speech Coders, CRC Press LLC, Boca Raton, 2000.

[20] MSU Video Quality Measurement Tool, http://www.compression.ru/video/quality_measure/vqmt_download_en.html

SAŽETAK

Cilj ovog završnog rada je objasniti maskiranje u video signalu. Osim različitih tipova maskiranja, trebalo je objasniti i proučiti utjecaj maskiranja na kvalitetu prikazanog video signala. Maskiranje je objašnjeno kao smetnja kod promatranja ciljnog podražaja zbog umješanosti drugog maskirnog podražaja koji se u prostoru i vremenu nalazi blizu ciljnog podražaja. Glavni tipovi maskiranja su prostorno i vremensko maskiranje. Prostorno maskiranje objašnjava kako je zbog svojstva ljudskog vizualnog sustava šum (smetnja) u slici jasno vidljiv na „glatkim" područjima u slici dok je na „grubim" područjima gotovo neprimjetan. Zbog prevelike brzine izmjene okvira kod vremenskog maskiranja, ljudsko oko ne stigne pratiti sve promjene koje se dogode pa stoga ne primjećuje neke nepravilnosti kod izmjene okvira u videu. Utjecaj maskiranja na kvalitetu prikazanog video signala proučen je provođenjem subjektivnog i objektivnog ocjenjivanja kvalitete videa. Otkriveno je kako je subjektivna metoda pouzdanija. Naime, subjektivna metoda koja predstavlja ocjenu gledatelja puno je pouzdanija s obzirom da ljudsko oko ima sposobnost opažanja utjecaja maskiranja dok računalo to nije u stanju.

Ključne riječi: prostorno maskiranje, vremensko maskiranje, subjektivne metode, objektivne metode, video signal

MASKING IN VIDEO SIGNAL

ABSTRACT

Goal of this final work is to explain masking in video signal. Beside different types of masking, the influence of masking on quality of displayed video signal had to be explained and explored. Masking is explained as disturbance during observation of target stimulus because of interference of another masking stimulus which is near target stimulus in time and space. Main types of masking are spatial and temporal masking. Spatial masking explains that due to the characteristic of the human visual system, noise (disturbance) in an image is clearly visible on „smooth“ areas of the image while on „rough“ areas it is almost incospicuous. Due to very fast exchange of frames in temporal masking, human eye can't follow all changes that happen and therefore doesn't notice some irregularities at frame exchange in video. Masking effect on the quality of displayed video signal has been studied by conducting subjective and objective evaluation of video. It was revealed that subjective method is more reliable. Namely, subjective method that represents viewers evaluation is more reliable, considering that human eye has the ability to perceive the masking effect while the computer is not capable of such thing.

Key words: spatial masking, temporal masking, subjective method, objective method, video signal

ŽIVOTOPIS

Irinej Stupar je rođen 4. lipnja 1991. u Osijeku u Hrvatskoj gdje je proveo svojih prvih 7 godina života. Nakon toga se sa svojom obitelji seli u Umag gdje i danas živi. Tu je pohađao osnovnu školu pa je nakon završetka iste upisao srednju školu u Bujama, smjer opća gimnazija. 2010. godine završava srednju školu i dolazi privremeno živjeti kod rodbine u Beli Manastir te upisuje u Osijeku elektrotehnički fakultet, smjer informatika.