

Mjerenje kvalitete videa u sustavima mrežnog video strujanja

Štrk, Kristijan

Undergraduate thesis / Završni rad

2022

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://um.nsk.hr/um:nbn:hr:200:931841>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-10**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



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

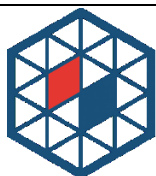
Sveučilišni studij

**Mjerenje kvalitete videa u sustavima mrežnog video
strujanja**

Završni rad

Kristijan Štrk

Osijek, 2022.

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

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

Osijek, 19.09.2022.

Odboru za završne i diplomske ispite

**Prijedlog ocjene završnog rada na
preddiplomskom sveučilišnom studiju**

Ime i prezime Pristupnika:	Kristijan Štrk
Studij, smjer:	Preddiplomski sveučilišni studij Računarstvo
Mat. br. Pristupnika, godina upisa:	R4195, 24.07.2018.
OIB Pristupnika:	15184463394
Mentor:	Prof.dr.sc. Snježana Rimac-Drlje
Sumentor:	,
Sumentor iz tvrtke:	
Naslov završnog rada:	Mjerenje kvalitete videa u sustavima mrežnog video strujanja
Znanstvena grana rada:	Telekomunikacije i informatika (zn. polje elektrotehnika)
Zadatak završnog rad:	Kvaliteta videa u sustavima mrežnog video strujanja ovisi o brojnim parametrima kao što su početno kašnjenje, trajanje i frekvencija zastoja, promjene razine kvalitete, prosječna razina kvalitete i drugi. U radu je potrebno dati pregled objektivnih metoda za mjerenje kvalitete videa u sustavima mrežnog video strujanja. Detaljnije opisati izabrane objektivne metode te prema njima napraviti program za proračun ocjene kvalitete.
Prijedlog ocjene završnog rada:	Dobar (3)
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: 2 bod/boda Jasnoća pismenog izražavanja: 1 bod/boda Razina samostalnosti: 2 razina
Datum prijedloga ocjene od strane mentora:	19.09.2022.
Datum potvrde ocjene od strane Odbora:	21.09.2022.
Potvrda mentora o predaji konačne verzije rada:	<i>Mentor elektronički potpisao predaju konačne verzije.</i>
	Datum:

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**IZJAVA O ORIGINALNOSTI RADA**

Osijek, 22.09.2022.

Ime i prezime studenta:

Kristijan Štrk

Studij:

Preddiplomski sveučilišni studij Računarstvo

Mat. br. studenta, godina upisa:

R4195, 24.07.2018.

Turnitin podudaranje [%]:

10

Ovom izjavom izjavljujem da je rad pod nazivom: **Mjerenje kvalitete videa u sustavima mrežnog video strujanja**

izrađen pod vodstvom mentora Prof.dr.sc. Snježana Rimac-Drlje

i sumentora ,

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:

SADRŽAJ

1. UVOD.....	1
1.1. Zadatak završnog rada	2
2. ADAPTIVNO VIDEO STRUJANJE ZASNOVANO NA HTTP PROTOKOLU	3
2.1. Adaptivni algoritmi za videostrujanje	5
2.2. Čimbenici koji utječu na kvalitetu videostrujanja	6
3. METRIKE ZA OCJENU KVALITETE VIDEO.....	8
3.1. Subjektivne metode mjerenja kvalitete videa	8
3.2. Opće objektivne metode mjerenje kvalitete videa.....	10
3.3. Metode za ocjenu kvalitete video signala nakon adaptivnog video strujanja.....	12
4. USPOREDBA OBJEKTIVNIH I SUBJEKTIVNIH OCJENA KVALITETE VIDEO SIGNALA ZA ADAPTIVNO VIDEO STRUJANJE.....	15
4.1. LIVE-NFLX-II QoE baza podataka	15
4.2. Analiza rezultata videostrujanja iz LIVE-NFLX-II baze	17
4.3. Korelacija ocjena objektivnih metrika sa subjektivnim ocjenama	18
5. ZAKLJUČAK	21

LITERATURA

SAŽETAK

ABSTRACT

ŽIVOTOPIS

PRILOZI

1. UVOD

Korisnici imaju visoke standarde kada je u pitanju video strujanje i današnji najpopularniji servisi poput Netflix-a, Youtube-a i Twitch-a moraju ispuniti očekivanja. Razlike između objektivnih i subjektivnih metrika ocjenjivanja, razni protokoli za što bolju kvalitetu i što veću ugodnost prilikom korištenja današnjih najpoznatijih aplikacija za video strujanje, razna mjerenja i njihovi rezultati su dio ovog završnog rada. Razni faktori utječu na kvalitetu usluge – QoS (eng. *Quality of Service*) poput stabilnosti i brzine interneta, gubitci i redosljed paketa. Ti faktori imaju veliki utisak na korisničko iskustvo – QoE (eng. *Quality of Experience*) pri gledanju različitih video sekvenci tijekom video strujanja. Postoje razne vrste protokola strujanja ali svi su zasnovani na HTTP-a npr. HTTP – DASH (eng. *Hypertext Transfer Protocol - Dynamic Adaptive Streaming*). U posljednje četiri godine predložen je niz algoritama za upravljanje prilagodbom DASH rješenja i predstavljeni su specifični klijenti i poslužiteljski program. Međutim, pažnja korisnika može biti pod negativnim utjecajem ako se razlučivost videozapisa često mijenja zbog drastičnih zagušenja mreže koja pogoršavaju QoE. Dakle, novi kriterij se treba dogovoriti kako bi se ocijenila QoE prilikom video strujanja. HTTP adaptivno video strujanje (eng. HAS – *HTTP-BASED adaptive video streaming*) je postao *de facto* standard za moderne uslužitelje video strujanja poput Youtube-a i sličnih.

Glavna ideja iza HAS-a je ta da se videozapis kodira u više različitih prezentacija sa različitim kodnim brzinama i različitim razinama kvalitete, kako bi se omogućilo klijentima odabir prezentacije koji će zadovoljiti vremenski promjenjivu propusnost mreže. Klijentske odluke se inače stvaraju na prethodnim vrijednostima protoka mreže, budućim procjenama protoka i ostalim informacijama koje se odnose na klijenta na primjer razina popunjenosti međuspremnika (eng. *buffer level*) [1]. Artefakti kompresije obično su vidljivi kao pogreška u obliku bloka koji dovodi do gubitka detalja i informacija o teksturi, dok su artefakti skaliranja obično vidljivi kao mutni rubovi i smanjene oštrine. Ako propusnost dosegne vrlo niske vrijednosti i međuspremnik se isprazni, klijent mora pauzirati video reprodukciju („*rebuffering*“), pričekati da se mreža oporavi i da se ispuni međuspremnik video podacima prije nastavka prikazivanja videozapisa. Postoje alternative za smanjenje brzine prijenosa videozapisa, kada propusnost pada, npr. korištenje niže brzine izmjene slika [2]. Veliki broj čimbenika i razni načini ocjenjivanja su bili povod dogovora oko zajedničkih metrika za ocjenjivanje kvalitete videa.

Ovaj završni rad je podijeljen u pet poglavlja. Nakon uvoda daje se pregled karakteristika HTTP protokola i adaptivnog video strujanja. Treće poglavlje opisuje metode mjerenja kvalitete videozapisa, kako se dijele i na koji način se provodi mjerenje kvalitete za adaptivno video strujanje. U četvrtom poglavlju je opisana baza podataka (LIVE-NFLX-II QoE) koja je dostupna svima, a u kojoj se nalaze parametri video strujanja za veliki broj različitih sekvenci, uz primjenu 4 algoritma za adaptivni izbor razine kvalitete segmenta, te različite uvjete propusnosti mreže. Na ovoj je bazi provedena usporedba objektivnih i subjektivnih ocjena kvalitete te su prikazani i analizirani rezultati usporedbe. U zadnjem poglavlju su dani zaključci.

1.1. Zadatak završnog rada

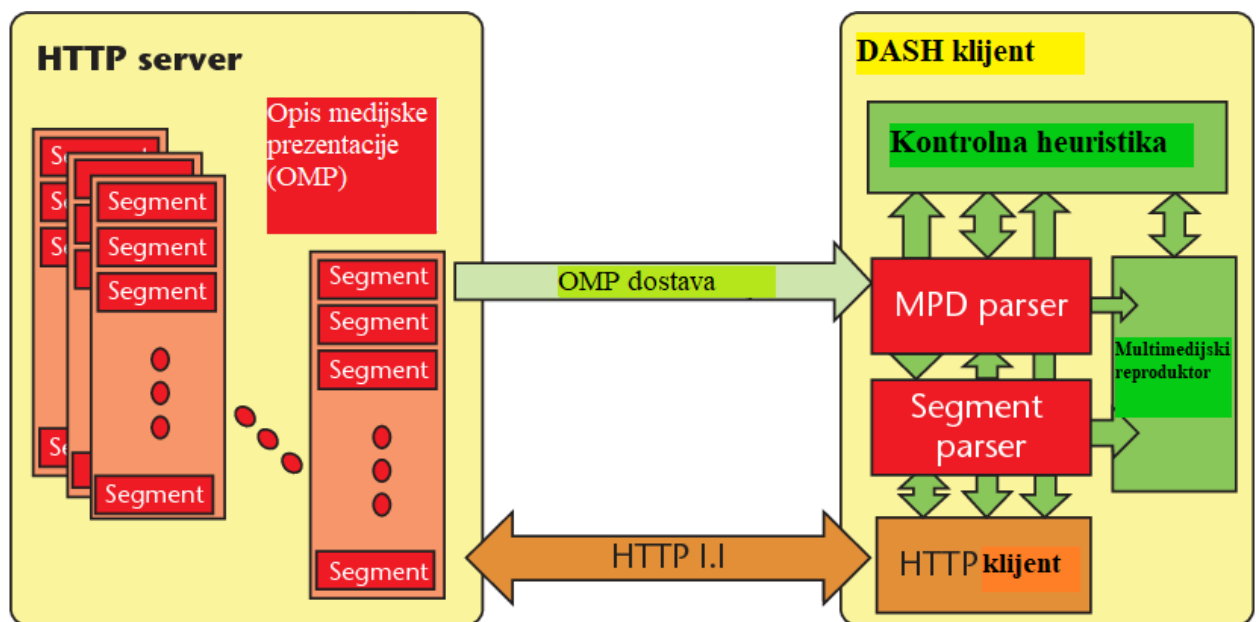
Kvaliteta videa u sustavima mrežnog video strujanja ovisi o brojnim parametrima kao što su početno kašnjenje, trajanje i frekvencija zastoja, promjene razine kvalitete, prosječna razina kvalitete i drugi. U radu je potrebno dati pregled objektivnih metoda za mjerenje kvalitete videa u sustavima mrežnog video strujanja. Detaljnije opisati izabrane objektivne metode te prema njima napraviti program za proračun ocjene kvalitete. Objektivne ocjene kvalitete dobivene proračunom treba usporediti sa subjektivnim ocjenama kvalitete.

2. ADAPTIVNO VIDEO STRUJANJE ZASNOVANO NA HTTP PROTOKOLU

Adaptivno video strujanje je tehnika koja je prisutna i koja se koristi u strujanju multimedija putem računalnih mreža. Protokol RTP (eng. *Real-time protocol*) zajedno sa protokolom RTPSP (eng. *Real Time Publish Subscribe Protocol*) korišten u prošlosti za većinu video ili audio strujanja, tehnološkim napretkom je zamijenjen prilagodljivim strujanjem koji se temelji na HTTP protokolu (eng. HTTP - *Hypertext Transfer Protocol*) i ujedno je dizajniran da radi na velikim distribuiranim HTTP mrežama npr. Internet. HTTP je protokol aplikacijske razine koji omogućava prijenos datoteka koje u sebi sadrže veze na druge dokumente. HTTP je kreiran uz HTML kako bi se napravio prvi interaktivni tekstualni internetski pretraživač odnosno originalna svjetska mreža (eng. WWW - *World Wide Web*). HTTP protokol i dalje ostaje primaran način za korištenje interneta. Protokol funkcionira na principu zahtjev-odgovor, daje korisniku način da pristupa web resursima kao što su HTML datoteke tako što prenosi hipertekstualne poruke između klijenata i poslužitelja. Klijenti koji koriste HTTP generalno koriste protokol kontrole prijenosa (eng. TCP - *Transmission Control Protocol*) konekcije za komunikaciju s poslužiteljima. Za izvođenje raznih naredbi i zadataka HTTP koristi specifične metode zahtjeva. Svi HTTP poslužitelji koriste GET (za zahtjeve određenih resursa u cijelosti) i HEAD (za zahtjeve određenih resursa bez sadržaja tijela) upite, kao i ostale upite poput POST, PUT, DELETE, TRACE, OPTIONS itd. koji mogu biti podržani, ali ne nužno na tim poslužiteljima.

Standard DASH temelji se na postojećim tehnologijama koje uključuju video kodek, formate enkapsulacije i transportne protokole, stoga je DASH lako prilagodljivo rješenje za strujanje video zapisa. Specifikaciju DASH razvili su 3GPP (eng. *3rd Generation Partnership Project*) i MPEG (eng. *Moving Picture Experts Group*). HTTP adaptivno strujanje (HAS) je napredna tehnologija koju koriste razni popularni servisi kojima je cilj riješiti problem promjena propusti mreže. HAS tehnologija funkcionira tako da video prilagodi trenutnoj propusnosti promjenom razine kvalitete sljedećeg segmenta koji izabire klijent. Ove promjene imaju veliki utjecaj na korisnika i pozitivno utječu na QoE. Napredak HAS-a naspram standardnog HTTP-a je u smanjenom broju prekidanja video strujanja i u boljoj iskorištenosti frekvencijskog pojasa. Video je prilagođen na razne načine npr. smanjenjem ili povećanjem broja slika u sekundi ili stupnjem kompresije.

HAS zahtjeva da video signali budu dostupni u više različitih brzina prijenosa tj. u različitim razinama kvalitete te da se video signal razlomi u male segmente koji sadrže par sekundi sadržaja. Mjerenjem trenutne brzine prijenosa i stanja međuspremnika, HAS zahtjeva sljedeći segment u prikladnoj brzini prijenosa kako bi se izbjeglo zaustavljanje video sadržaja. Način funkcioniranja DASH-a prikazan je na slici 2.1. HAS nudi više različitih brzina prijenosa video signala što omogućuje davateljima video usluga da prilagode isporučeni video signal prema zahtjevima korisnika. Kao primjer, video signal visoke brzine prijenosa, što obično žele korisnici u kući koji uživaju u brzom pristupu internetu i velikim zaslonima poput televizora, nije prikladan za korisnike mobilnih podataka s malim zaslonom mobitela i sporim pristupom internetu. Štoviše zbog adaptivnog strujanja korisnici proizvoljno mogu mijenjati rezoluciju i kvalitetu videa koju žele gledati. Da je video dostupan samo u jednoj kvaliteti bilo koja promjena na internetu negativno bi utjecala na QoE jer bi dolazilo do konstantnih zaustavljanja. Kako bi se ispoštovao zahtjev pružanja video sadržaja pri različitim brzinama prijenosa za HTTP adaptivno strujanje, može se koristiti jedna ili nekoliko dimenzija prilagodbe. Na slici 2.1. prikazano je djelovanje DASH-a i način na koji DASH klijent komunicira sa HTTP serverom.



Slika 2.1. Način funkcioniranja DASH-a [3]

2.1. Adaptivni algoritmi za videostrujanje

Razni adaptivni algoritmi se koriste za adaptivno video strujanje (eng. ABR – *Adaptive BitRate*), najpopularniji su: algoritam temeljen na međuspremniku (eng. BB – *Buffer-Based adaptor*) [4], algoritam temeljen na stopi (eng. RB – *Rate-Based adaptor*), algoritam temeljen na kvaliteti (eng. QB – *Quality-Based adaptor*), algoritam temeljen na oracle kvaliteti (eng. OQB – *Oracle Quality-Based adaptor*).

LIU algoritam [5] detektira promjene u širini pojasa koristeći izglađen HTTP protok izmjeren na temelju vremena dohvaćanja segmenata. Izglađeni HTTP je korišten da se provjeri dali je trenutna brzina prijenosa sadržaja ista kao i maksimalan kapacitet širine prijenosa sadržaja. Rezultati [5] pokazuju da se predloženi algoritam prilagodbe brzine brzo prilagođava tako da odgovara kapacitetu mreže od kraja do kraja, a također učinkovito kontrolira prelijevanje i prelijevanje međuspremnika. LIU algoritam ne zahtjeva da informacije o transportnom sloju kao što su stope gubitka paketa i vrijeme povratnog putovanja budu dostupni na aplikacijskom sloju.

Algoritam, zasnovan na propusnosti, za prilagodljivo video strujanje uz poboljšanu kvalitetu korisničkog iskustva (engl. *QoE-enhanced Adaptation Algorithm over DASH – QAAD*) [6]. QAAD čuva minimalnu duljinu međuspremnika kako bi se izbjegli prekidi i minimizirale promjene kvalitete video signala tijekom reprodukcije. QAAD pruža stabilan prijenos informacija sa stabiliziranom kvalitetom video sadržaja. Kada se stanje mreže naglo pogorša i smanji širina pojasa, QAAD polagano smanjuje razinu kvalitete kako bi se smanjio broj izmjena kvaliteta uzastopnih segmenata dok postojeći međuspremnik svjesnog algoritma prilagodbe naglo smanjuje razinu kvalitete. QAAD pruža stabilnu kvalitetu bez prekida videostrujanja čak i kada stanje mreže i širina pojasa povremeno oscilira. QAAD uvodi točniju procjenu propusnosti i glatke sheme odabira brzine prijenosa na temelju procijenjene širine pojasa. Ključna ideja sheme odabira brzine prijenosa u QAAD-u je minimiziranje naglog smanjenja odabranih brzina prijenosa za održivi QoE korisnika. QAAD poboljšava QoE korisnika sprječavajući prekid reprodukcije i minimiziranjem promjena kvalitete video sadržaja tijekom reprodukcije.

Algoritam koji u isto vrijeme sprječava pražnjenje međuspremnika ispod unaprijed određene granične vrijednosti i maksimizira ukupni QoE opisan je u [7]. Algoritam se koristi za prilagodljivo strujanje videosignala putem HTTP-a.

ABR algoritam za prilagodljivo strujanje videosignala opisan u [8] prvi je algoritam za video strujanje koji u obzir uzima složenost sadržaja videosignala (eng. *Content-Aware Rate*

Adaptation - CARA). Predložena [8] je prilagodba stope s obzirom na sadržaj shema koja poboljšava stabilnost preciznom propusnošću procjena. Predstavljen [8] je model međuspremnik reprodukcije koji uzima u obzir karakteristike sadržaja. Algoritam uzima u obzir složenost sadržaja videosignala kako bi smanjio utjecaj nepreciznog procjenjivanja pojasne širine. Preciznija procjena širine pojasa se postiže tako da se uzme u obzir izmjenjiva veličina segmenta. Za prevenciju oscilacija brzine prijenosa koriste se informacije o međuspremniku. Za dohvaćanje informacija o međuspremniku predložen [8] je model reprodukcije brzine prijenosa. Reproducirani međuspremnik je podijeljen na tri dijela za podešavanje agresivnosti predloženog algoritma. Predloženi algoritam postiže visoku kvalitetu videostrujanja, čak i pod nestabilnim mrežnim uvjetima. Shema poboljšava stabilnost koristeći pristup koji je svjestan sadržaja koji koristi stvarni segment informacija.

2.2. Čimbenici koji utječu na kvalitetu videostrujanja

Vrijeme koje je potrebno da video zapis započne sa reprodukcijom naziva se početnim kašnjenjem. Spremnik (eng. *buffer*) se za to vrijeme puni segmentima video zapisa. Početno kašnjenje je uvijek prisutno u multimedijским uslugama jer prije samog dekodiranja sadržaja i reprodukcije određene se količine podataka moraju prenijeti i obraditi. Razna istraživanja [9] su pokazala da korisnici žele duže čekati prije početka reprodukcije za traženu kvalitetu videozapisa radije nego da dolazi do promjene kvalitete videa ili zastoja u reprodukciji.

Ukoliko je propusnost mreže manja od kodne brzine, spremnik se počinje prazniti. Ako se spremnik skroz isprazni doći će do zaustavljanja (zamrzavanja) videozapisa (eng. *stalling*). ABR algoritmi tada smanjuju traženu kvalitetu sljedećeg segmenta, što znači manju kodnu brzinu, pa započinje ponovno punjenje spremnika (eng. *rebuffering*) sve dok poslužitelj ne odluči da je spremnik dovoljno pun za sadašnje stanje i uvjete u mreži. Prilagodljivo strujanje putem HTTP protokola omogućava izbjegavanje takvog zaustavljanja promjenom kvalitete videozapisa za razliku od postupnog preuzimanja putem HTTP protokola. Istraživanja pokazuju da korisnici preferiraju jednu dužu pauzu nego veći broj kraćih pauza [9]. Također zaustavljanja na početku videozapisa su manje značajna od onih u sredini i pri kraju.

Uvjeti u mreži nisu uvijek isti, razni faktori odlučuju o tome na primjer: broj korisnika povezanih na mreži, vremenski uvjeti, doba dana, udaljenost od modema, kvaliteta modema itd. Zbog tih raznih uvjeta u mreži mora postojati prilagodba kvalitete videa (eng. *adaptation*).

Smanjenje broja slika u sekundi negativno utječe na QoE ali slabije utječe nego zaustavljanje videozapisa. Postepeno smanjivanje broja slika u sekundi je znatno bolje prihvaćeno nego naglo smanjivanje broja slika u sekundi (eng. *FPS - frames per second*). Zaustavljanje videozapisa se može izbjeći smanjivanjem prostorne rezolucije ili smanjivanjem broja slika u sekundi. Korisnici bolje ocjenjuju videozapise koji na kraju prikaza imaju višu kvalitetu nego na početku. Bez adaptivnog strujanja, kvaliteta odnosno rezolucija i broj slika u sekundi bi morali biti stalni što bi uzrokovalo veliki broj zastoja i jako negativno utjecao na QoE.

Koraci obrade poput: stvarno snimanje videosignala, obrada, kodiranje, kompresija, prijenos, dekodiranje, imaju direktan i snažan utjecaj na ono što korisnik vidi.

Razni čimbenici imaju utjecaj na kvalitetu videosignala u procesu prijenosa, prikazivanja i obrade videosignala, na primjer: parametri dekodiranja i kodiranja, način snimanja videosignala, parametri i vrsta kompresije, prijenosni medij i mnogi drugi. Na korisnika utječu i drugi čimbenici poput: kontekst uporabe, emocije i tako dalje. [10]

Uvjeti gledanja videosignala, vrsta zaslona na kojem se prikazuje videosignal, kvaliteta i sinkronizacija pratećeg zvuka, pojedinačni interesi gledatelja u smislu sadržaja videosignala određuju razinu pozornosti itd. su samo neki od mnogobrojnih čimbenika koji utječu na QoE i na samu kvalitetu video strujanja. [10]

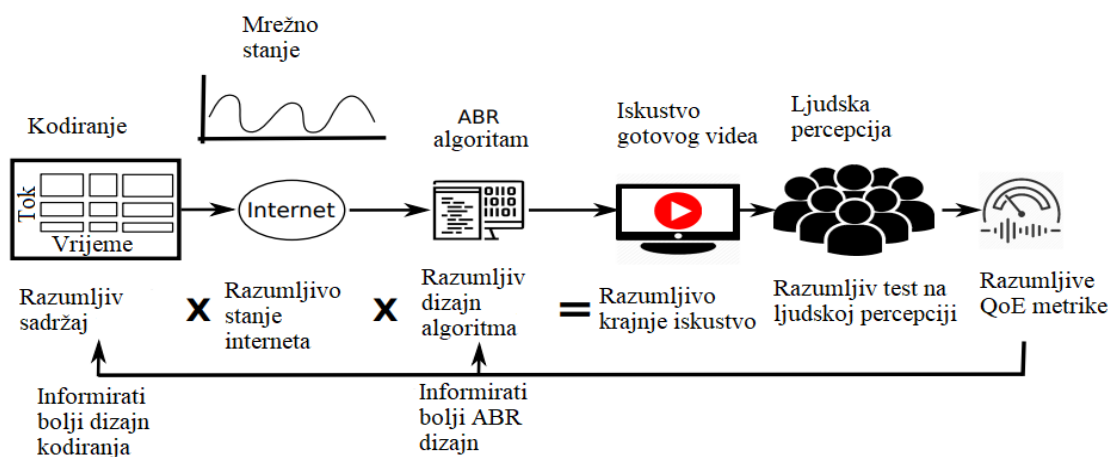
3. METRIKE ZA OCJENU KVALITETE VIDEA

Metrike za ocjenu kvalitete videa se dijele na subjektivne u kojima se ocjenjuje kvaliteta videa na rezultatima prosječne ocjene određenog broja gledatelja i objektivne gdje se ocjena dobije iz matematičkog modela.

Obje skupine metrika daju važan uvid u kvalitetu videosignala te se ostvareni rezultati koriste za unapređenje algoritama za videostrujanje, određivanje parametara kodiranja i segmentacije videosignala.

3.1. Subjektivne metode mjerenja kvalitete videa

Trenutno su najprihvaćenije metodologije subjektivnih testova za procjenu kvalitete videa opisane u ITU preporukama ITU-R BT-500 [11] i ITU-T P.910 [12]. Ovi testovi se provode u laboratorijskom okruženju koje ispunjava određene zahtjeve u pogledu osvjetljenja i akustike. Na slici 3.1 prikazana je međuovisnost kodiranja i ABR algoritama s ljudskim percipiranjem gotovog videosignala. Rezultati QoE metrike, dobivenih iz testova za razumijevanje ljudske percepcije, daju informacije za bolji ABR dizajn i bolji dizajn kodiranja za buduća testiranja.



Slika 3.1. Dijagram ocjenjivanja kvalitete videa [15]

Procesi subjektivnog vrednovanja koriste korisnikovo mišljenje o kvaliteti usluge. Mišljenje korisnika najčešće se opisuju prosječnom ocjenom mišljenja (eng. MOS – *Mean opinion score*). Kategorijske skale se mogu predstaviti li numerički, od 1 do 5 ili tekstualnim prikazom od „loše“ do „izvrsno“. Ocjene dobivene subjektivnim metodama smatraju se najtočnijom procjenom QoE. Jedini način na koji se može saznati kakva je iskustvena kvaliteta usluge je preko ispitivanja mišljenja korisnika na slučajno odabranom uzorku. Isto kao i u objektivnim metodama, izuzetno bitan korak je izbor faktora koji se ispituju. Nakon korištenja određenih usluga ispitivanja se provode anketiranjem korisnika.

Subjektivne metode mjerenja kvalitete videa temeljene su na ljudskom vizualnom sustavu i ocjenjivanju vizualnog sadržaja od strane anketiranih ljudi koji u unaprijed određenim, propisanim i zadanim uvjetima ocjenjuju kvalitetu promatranog sadržaja, u ovom slučaju, video sadržaja. Srednja ocjena svih anketiranih ljudi za pojedini sadržaj ekvivalent je kvalitete promatranog sadržaja. Rezultat kod ovakvih metoda i postupaka mjerenja nije egzaktan i potpuno pouzdan broj već produkt statističke distribucije niza pojedinačnih subjektivnih rezultata koji ovise o raznim faktorima poput: odabira ispitnog sadržaja, raspoloženja promatrača, osvjetljena i izgleda prostorije gdje se vrši promatranje itd.

Subjektivni postupci mjerenja kvalitete su podijeljeni na: jednopodražajni postupak s apsolutnom ocjenom kvalitete slike - SSCQS (eng. *Single stimulus continuous quality evaluation*), dvopodražajni postupak s ocjenom izobličenja slike – DDIS (eng. *Double stimulus impairment scale*), dvopodražajni postupak s ocjenom kvalitete slike – DSCQS (eng. *Double stimulus impairment scale*) i postupak apsolutnog kategorijskog ocjenjivanja – ACR (eng. *Absolute Category Rating*).[13]

Primjer subjektivnog mjerenja kvalitete videa je javno dostupan unutar LIVE-NFLX-II QoE baze podataka [14], kasnije u radu opširnije opisana baza podataka. Tijekom četiri tjedna na teksaškom fakultetu Austin provedeno je testiranje na 65 osoba (50 muških, 15 ženskih, od 18 do 30 godina). Sudionici testiranja su bili studenti bez iskustva u ocjenjivanju kvalitete videa. Za testiranje je korištena prostorija koja je zadržavala iste uvijete osvjetljenja, svjetlina monitora je uvijek bila postavljena na 50%, korišten je HP monitor. Traka za prikupljanje ocjena je prikazana cijelo vrijeme na dnu zaslona tijekom reprodukcije videozapisa kako bi subjekti mogli davati kontinuirani rezultat QoE u stvarnom vremenu. Sveukupni QoE rezultat prikazuje konačni QoE nakon gledanja svake video sekvence u potpunosti, kao srednju vrijednost kontinuiranih ocjena.

Svih 65 osoba je smješteno na udobnom sjedalu na udaljenosti od 0.76 metara od monitora. Sudionici su prije ocjenjivanja dobili upute o postupku. Za smanjivanje umora, eksperiment je podijeljen u 30 minutne sesije gledanja 50 videa (150 videa po ispitaniku). Između svake sesije je bila pauza od barem 24 sata da bi umor ispitanika bio minimalan, [15]. Korišten je Snellov test vida za testiranje ispitanika kako bi se utvrdilo da svi imaju normalan vid. Sveukupno je skupljeno $65 \cdot 150 = 9750$ rezultata QoE sveukupne kvalitete videa i 9750 valnih oblika kontinuiranih QoE ocjena.

3.2. Opće objektivne metode mjerenje kvalitete videa

Objektivne metode ocjenjivanja su zasnovane na matematičkim modelima koji za procjenu kvalitete videa koriste različite parametre videa. Objektivne metode ocjene kvalitete videa dijele se u tri skupine: metode ocjene kvalitete videa s potpuno dostupnim referentnim videom (eng. FR - *full reference methods*), metode ocjene kvalitete videa s djelomično dostupnim informacijama o referentnom videu (eng. RR - *reduced reference methods*) i metode ocjene kvalitete videa bez dostupnih informacija o referentnom videu (eng. NR - *no-reference*). U većini slučajeva referentni video nije dostupan stoga je ocjena kvalitete moguća samo putem NR metode. Pristup bez referentnog videa svodi se na pronalazak poznatih svojstava artefakata koji se mogu pojaviti u video sadržaju kao rezultat različitih distorzija tijekom produkcije i prijenosa videa, [16]. Osnovni je cilj razvoja metoda objektivne procjene QoE dizajniranje metode mjerenja kvantitativnih parametara i na takav način se pokušava s dobivenim rezultatima procijeniti osobno mišljenje korisnika. Razlozi za unaprjeđivanje objektivnih metoda su: jeftinije su za provedbu od subjektivnih metoda, lakša (brža) izvedba nego subjektivno ocjenjivanje. PSNR, MSAD[17], SSIM[18], MS-SIM[19], ST-RRED[20], VMAF[21] su neke od objektivnih metrika koje se koriste za ocjenu kvalitete videa.

PSNR (eng. *Peak signal to noise ratio*) metrika je jedan od sveprisutnih, najkorištenijih i najpopularnijih postupaka. Popularnost dobiva zbog svoje jednostavnosti i brzine izračuna, ali nije idealna jer rezultati dobiveni iz ovog postupka obično imaju manju korelaciju s rezultatima dobivenim iz subjektivnih ocjena kvalitete u odnosu na druge metrike. Formula za računanje kvalitete videa putem PSNR postupka dana je izrazom (3-1), [22].

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (3-1)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2$$

MAX_I = Maksimalna vrijednost elementa slike u originalnoj slici

m = broj redova u originalnoj slici

n = broj stupaca u originalnoj slici

Metrika MSAD temeljena je na vrijednosti koja predstavlja srednju apsolutnu razliku komponenti boja u referentnim točkama slike. Metrika je korištena prilikom testiranja kodeka i filtera. Prikaz MSAD metrike vidljiv je na slici 3.3. a formula za proračun dana je izrazom (3-2), [17].

$$d(X,Y) = \frac{\sum_{i=1}^m \sum_{j=1}^n |Y_{i,j} - X_{i,j}|}{mn} \quad (3-2)$$

m = broj redova u originalnoj slici, n = broj stupaca u originalnoj slici



Slika 3.2 MSAD metrika [17]

SSIM (eng. *Structural Similarity Index*) percepcijska je metrika koja kvantificira degradaciju kvalitete slike uzrokovanu obradom kao što je kompresija podataka ili gubici u prijenosu podataka. To je potpuna referentna metrika koja zahtijeva dvije slike istog sadržaja: referentnu sliku i obrađenu sliku. Obrađena slika obično je komprimirana. SSIM se koristi u video industriji, ali ima čestu primjenu i za mirne slike. SSIM zapravo mjeri perceptivnu razliku između dvije slične slike. Ne može procijeniti koja je od te dvije bolja: to se mora zaključiti na temelju saznanja koji je "original", a koja je podvrgnuta dodatnoj obradi kao što je kompresija podataka. Za razliku od PSNR-a, SSIM se temelji na vidljivim strukturama na slici.

MS-SSIM (eng. *Multi-Scale Structural Similarity*) koristi više skala (rezolucija) za proračun strukturne sličnosti za procjenu kvalitete slika, koja pruža veću fleksibilnost od pristupa s jednom skalom u uključivanju varijacija rezolucije slike i uvjete gledanja. Eksperimenti [23] pokazuju da s odgovarajućim postavkama parametara, metoda s više skala nadmašuje najbolje SSIM modele u jednoj skali. U razvoju „*top-down*“ modela kvalitete slike (kao što su algoritmi temeljeni na strukturnoj sličnosti), jedan od najizazovnijih problema je kalibracija parametara modela, koji su prilično apstraktni i ne mogu se izravno izvesti iz jednostavnih eksperimenata kao u modelima „*bottom-up*“.

ST-RRED (eng. *Spatio-Temporal Reduced Reference Entropic Differencing*) je model kvalitete videa koji je razvijen u LIVE-u [24] tijekom 2013. godine, a koristi smanjenu referentnu sliku te ima dobre performanse u smislu predviđanja kvalitete videa [25]. Osnovni element dizajna ST-RRED-a je fleksibilan okvir koji omogućuje jednostavno skaliranje u smislu dimenzije slike, brzine prijenosa i količine referentnih informacija potrebnih za provođenje predviđanja kvalitete videa. U načelu, ST-RRED se može implementirati u oblicima koji se kreću između korištenja pune reference i modela s jednim brojem, pri čemu bi bio potreban ili sav referentni video ili bilo koji iznos između, sve do samo jednog broja koji predstavlja pojedini referentni video okvir.

3.3. Metode za ocjenu kvalitete video signala nakon adaptivnog video strujanja

VMAF metrika je najutjecajnija metrika koja se koristi za ocjenjivanje kvalitete video sadržaja pri strujanju, a razvio ju je Netflix, [26]. Pokazalo se da rana verzija VMAF-a nadmašuje druge metrike kvalitete slike i videa kao što su SSIM, PSNR-HVS i VQM-VFD na tri od četiri skupa podataka u smislu točnosti predviđanja, u usporedbi sa subjektivnim ocjenama. Njegova izvedba također je analizirana u [27], gdje je pokazano da VMAF ima bolje rezultate od SSIM-a i MS-SSIM-a prilikom testiranja LIVE-Netflix Video QoE baze podataka. Godine 2017. inženjeri iz RealNetworksa izvijestili su o dobroj ponovljivosti Netflixovih nalaza performansi ove metrike, [28]. U MSU programu za mjerenje kvalitete videozapisa, gdje su testirane njegove različite verzije (uključujući VMAF NEG), VMAF je nadmašio sve ostale metrike na svim standardima kompresije.

Kao rezultat preliminarnih subjektivnih testova, [29], identificirana su 3 glavna čimbenika poremećaja u internet strujanju videa primjenom DASH standarda: slučajevi prebacivanja iz

jedne u drugu video kvalitetu (kombinaciju kodne brzine te vremenske i prostorne rezolucije), zastoji u reprodukciji zbog pražnjenja međuspremnika i početno kašnjenje reprodukcije zbog početnog pohranjivanja video sadržaja u međuspremnik. Na temelju niza eksperimenata razvijeno je više NR metrika koje koriste ove parametre za ocjenu kvalitete videa nakon video strujanja.

U [29] autori analiziraju korelaciju pojedinih čimbenika koji utječu na QoE u videostrujanju sa subjektivnom ocjenom kvalitete, kao što su početno kašnjenje, prosječna kodna brzina videa, postotak trajanja zastoja u odnosu na trajanje videosekvence, broj i dubina promjena kvalitete segmenata i sl. Na temelju korelacije sa subjektivnim ocjenama, odredili su težine u različitim aditivnim kombinacijama ovih parametra (dva i više) tako da se postigne najbolja korelacija. Najbolju korelaciju za model s tri parametra dobili su za kombinaciju postotka trajanja zastoja, prosječne kodne brzine i prosječne veličine promjene kodne brzine segmenata kako je dano izrazom (3-4), [29].

$$VQM = -63.1 P_r + 0.0079\bar{E} + 0.0010\bar{E}_s + 49.7 \quad (3-4)$$

P_r = postotak trajanja zastoja zbog punjenja međuspremnika

\bar{E} = prosječna kodna brzina segmenata

\bar{E}_s = prosječna veličina promjene kodne brzine segmenata tijekom prijenosa

U [25] autori predlažu metriku $VsQM_{DASH}$ (eng. *Video streaming Quality Metric for DASH*) koja uzima u obzir početno kašnjenje zbog punjenja međuspremnika, broj pauza i njihove duljine, te promjenu prostorne i vremenske rezolucije događaja koja se događa tijekom video strujanja. Također, uzeta je u obzir vremenska lokacija pauza i promjena događaja. Nadalje, $VsQM_{DASH}$ može raditi s pružateljima video sadržaja koji koriste nekoliko rezolucija videozapisa i može se koristiti za poboljšanje performansi DASH kontroliranih algoritama. Provedeni su subjektivni testovi [25] procjene kvalitete videa za modeliranje $VsQM_{DASH}$. Eksperimentalni rezultati u [25] su pokazali da pogoršanja u početnom temporalnom segmentu imaju veći negativan učinak na QoE. Također, videozapisi s različitim prostornim rezolucijama imaju veći negativan učinak na QoE korisnika u odnosu na događaje prebacivanja između videozapisa s različitim vremenskim rezolucijama. To se može objasniti time što tijekom događaja promjene prostorne rezolucije, kako bi se fiksirao novi prostorni položaj istaknute regije, potrebno je izvesti nekoliko pokreta očima. Očekivano, rezultati u [25] su pokazali da veliki broj malih izmjena VQL-a (eng. *Video Quality Level*) manje ometaju vizualnu pažnju nego jedna izmjena sa najvećeg na najmanji VQL. Također, dodatni testovi prikazuju kako $VsQM_{DASH}$ nema utjecaja

na performanse uređaja u smislu brzine obrade podataka i korištenja energije. $VsQM_{DASH}$ je jednostavan i može se lagano implementirati u bilo koji mobilni uređaj. $VsQM_{DASH}$ se računa prema (3-3), [30].

$$VsQM_{DASH} = C \exp \left[- \sum_{i=1}^k \frac{w_i}{T_i} (N_i L_i + \sum_{j=1}^n p_{ji} R_{ji} + \sum_{l=1}^m q_{li} S_{li}) \right] - I_{BD} \quad (3-3)$$

C = konstanta za skaliranje

i = trenutni vremenski segment

N_i = broj pauza

L_i = prosječno trajanje pauza

w_i = težinski faktor (u odnosu na stupanj degradacije segmenta i)

T_i = vremenski period segmenta i

k = ukupni broj temporalnih segmenata

p_{ji} = težinski faktor (za promjene vremenske razlučivosti tipa j)

R_{ji} = broj promjena tipa j

n = broj tipova promjene temporalnih rezolucija

q_{li} = težinski faktor (za promjene prostorne razlučivosti tipa l)

S_{li} = broj promjena tipa l

m = broj tipova promjene prostorne razlučivosti

I_{BD} = degradacija dodana početnim kašnjenjem međuspremnik

4. USPOREDBA OBJEKTIVNIH I SUBJEKTIVNIH OCJENA

KVALITETE VIDEO SIGNALA ZA ADAPTIVNO VIDEO STRUJANJE

Subjektivne ocjene kvalitete videa smatraju se referentnim vrijednostima (eng. *ground truth*) i služe za procjenu točnosti objektivnih metoda ocjenjivanja kvalitete. Glavna prednost objektivnog ocjenjivanja je jednostavnost izračuna, ali tako dobivene ocjene često ne predstavljaju u potpunosti kvalitetu videa kako je ljudi percipiraju. Kao mjera točnosti neke objektivne metrike najčešće se koristi Pearsonov koeficijent korelacije kojim se iskazuje stupanj korelacije subjektivnih ocjena kvalitete većeg broja videozapisa s onima dobivenima tom objektivnom metrikom za iste zapise. Pearsonov koeficijent korelacije računa se prema izrazu (4-1), [30].

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4-1)$$

n = veličina uzorka

x_i = subjektivne ocjene

y_i = objektivne ocjene

\bar{x} = prosječna subjektivna ocjena

\bar{y} = prosječna objektivna ocjena

Usporedba objektivnih i subjektivnih ocjena kvalitete videa nakon video strujanja za potrebe ovog završnog rada provest će se na podacima dostupnim u LIVE-NFLX-II QoE bazi podataka [24].

4.1. LIVE-NFLX-II QoE baza podataka

LIVE-NFLX-II QoE baza podataka razvijena je u Laboratory for Image & Video Engineering, University of Texas at Austin, [24]. Ova baza je napravljena upravo za ocjenu kvalitete video strujanja, a sadrži originalne sekvence, sekvence nakon video strujanja, subjektivne i objektivne ocjene kvalitete te još niz drugih podataka o sekvencama i video strujanju. Za bazu je korišteno 15 videozapisa različitog video sadržaja (Tablica 4.1), 7 zapisa mrežnih propusnosti snimljenih u stvarnim 3G mrežama te 4 adaptivna algoritama za

videostrujanje, na temelju čega je generirano 420 različitih video strujanja s 420 rezultirajućih videosekvenci na klijentskoj strani. Originalnih 15 video zapisa odabrano je na način da pokrivaju raznovrsne sadržaje, kao što su akcija, dokumentarci, sportovi, animacije i video igre, odnosno da imaju različitu prostornu i vremensku kompleksnost. Video sekvence također sadrže računalno generiran sadržaj, poput Blendera [4]. Te video sekvence su snimljene i obrađene pod različitim svjetlosnim uvjetima od najsvjetlije scene (SB) pa do onih tamnijih (CD). Rezolucije veće od 1920 x 1080 i broj slika u sekundi većih od 30 su smanjeni na spomenute vrijednosti.

Za razliku od drugih sličnih baza jedino ova baza podataka ima: klijentsku adaptaciju, kontinuirane ocjene QoE pojedinog videa, stvarne zapise mrežne propusnosti, model nadopunjavanja spremnika, javno je dostupna, te sadrži preko 400 videa generiranih iz 15 različitih video sadržaja. Popis video sadržaja koje sadrži ova baza dan je u tablici 4.1. Nadalje, u ostalim istraživanjima korištene su iste vrijednosti kodnih brzina za razine kvalitete različitih video sadržaja, bez uzimanja u obzir složenosti tog sadržaja. S obzirom da postoje razlike u karakteristikama svakog sadržaja, ne trebaju svi videozapisi biti kodirani s istom vrijednošću kodne brzine da bi bili kodirani na istoj razini kvalitete. Na primjer, video zapisi koji sadrže bogate teksture ili značajne pokrete zahtijevaju veći broj bitova za kodiranje (veću kodnu brzinu) za određenu kvalitetu, u usporedbi sa relativno jednostavnijom scenom gdje ima malo pokreta ili su teksture slabije izražene. Iz toga se može zaključiti da izbor razina kvalitete i kodnih brzina za stvaranje reprezentacija na poslužitelju uz uzimanje u obzir karakteristike sadržaja, može ostvariti uštedu u prostoru za pohranu videa pružateljima te pružiti bolju kvalitetu korisnicima usluge video strujanja.

Tablica 4.1. Kratice zapisa svih 15 videozapisa [24]

Video sadržaj	ID	Video sadržaj	ID	Video sadržaj	ID
AirShow	AS	ElFuenteDance	ED	SkateBoarding	SB
AsianFusion	AF	ElFuenteMask	EM	Soccer	SO
Chimera1102353	CD	GTA	GTA	Sparks	SP
Chimera1102347	CF	MeridianConversation	MC	TearsOfSteelRobot	TR
CosmosLaundromat	CL	MeridianDriving	MD	TearsOfSteelStatic	TS

Za modele video strujanja koriste se stvarne vrijednosti mrežne propusnosti i simulator popunjenosti međuspremnik klijenta, a ne samo jednostavna mreža i model zauzetosti međuspremnik. Za generiranje video strujanja za bazu LIVE-NFLX-II [24] korišteno je 4 ABR

algoritma za adaptaciju kodne brzine (algoritma za adaptivnu promjenu izabrane razine kvalitete segmenata): Huang algoritam temeljen na međuspremniku [1], Oracle VMAF Viterbi algoritam temeljen na kvaliteti [24], jednostavan algoritam temeljen na propusnosti [24] i VMAF Viterbi algoritam temeljen na kvaliteti [24].

S obzirom na mnoštvo mrežnih tragova i ABR algoritama, LIVE-NFLX-II baza podataka [24] sadrži više aspekata prilagodbe strujanja, kao što je razina kvalitete pojedinog segmenta, događaji zastoja prikazivanja videa različitog trajanja te prostorna i vremenska rezolucija pojedinog izabranog segmenta. Subjektivni podaci sastoje se od ukupnog rezultata i kontinuiranih rezultata mjerenja kroz stvarno vrijeme, što ih čini idealnim za razvoj različitih QoE modela.

4.2. Analiza rezultata videostrujanja iz LIVE-NFLX-II baze

Glavna zapažanja iz prikupljenih podataka prilikom testiranja mogu biti sažeti kako slijedi. Bolji model predviđanja propusnosti može poboljšati kvalitetu video strujanja ostvarenu u većini slučajeva, odnosno poboljšati parametre koji utječu na tu kvalitetu kao što je prosječna kodna brzina te broj i trajanje zastoja reprodukcije zbog pražnjenja međuspremnik. Početak je najzahtjevniji dio sesije za sve ABR algoritme, budući da tada ABR algoritmi još nisu popunili video međuspremnik, a time i mrežne varijacije mogu lako smanjiti QoE.

Iz analize rezultata sveukupne kvalitete video zapisa može se zaključiti da postojanje početnog kašnjenja, veliki broj zastoja i promjena kvalitete video zapisa negativno utječu na sveukupne ocjene QoE. Prilikom testiranja primijećeno je da duža zamrzavanja videa smanjuju QoE, ali ako je zamrzavanje duže od 4 sekunde, efekt zaboravljanja dužine trajanja dolazi u obzir i ublažava negativan utjecaj tih zamrzavanja, [7]. Iz testiranja se zaključuje da nema velike razlike u ukupnoj ocjeni QoE između 4 adaptivna algoritma jer se većina zamrzavanja događa rano u puštanju videozapisa, a prije nego što je videozapis gotov svi algoritmi uspiju dovoljno napuniti međuspremnik za bolje upravljanje s promjenama kvalitete/brzine prijenosa, čak i ako je mreža jako nestabilna. Ispitanici su u prosjeku davali višu ocjenu sveukupnog QoE jer su zaboravili to jest „oprostili“ početno kašnjenje i zamrzavanje jer se zamrzavanja nisu kasnije ponavljala.

Analiza valnih oblika kontinuiranih ocjena QoE pokazuje da unutar prvih par sekundi RB algoritam započinje sa boljim rezultatom QoE za razliku od BB, QB i OQB jer se oni fokusiraju

na bolju nadopunu međuspremnik. Unutar prvih 12 sekundi BB ima najnižu QoE ocjenu. Nakon 12 sekundi QB i OQB napreduju značajno, s tim da OQB dobija najbolje rezultate do kraja sesije. Nakon 25 sekundi QoE vrijednosti nemaju velike promjene jer su svi algoritmi učitali video u potpunosti i nema dodatnih zastoja. Nadalje, iz rezultata se može zaključiti da bolji internetski uvjeti postepeno poboljšavaju rezultat QoE jer adaptivni algoritmi uspijevaju brže prebaciti video zapis na višu rezoluciju.

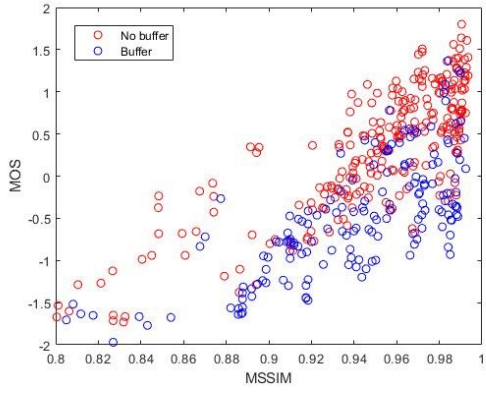
4.3. Korelacija ocjena objektivnih metrika sa subjektivnim ocjenama

Tablica 4.2. prikazuje Pearsonove koeficijente korelacije između objektivnih ocjena i subjektivnih ocjena. Za svaku objektivnu metriku izračunat je koeficijent korelacije za sekvence bez zastoja, sa zastojima i za sve sekvence.

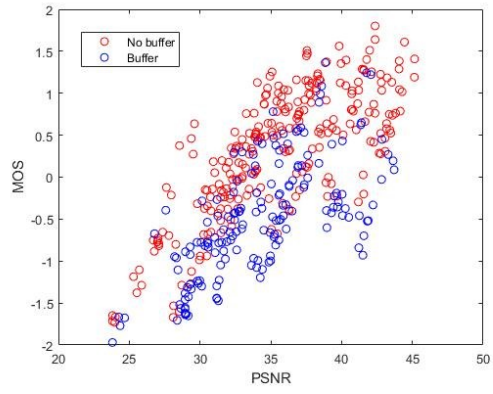
Tablica 4.2. PCC vrijednosti

Objektivna metrika	PCC za sekvence bez zastoja	PCC za sekvence sa zastojima	PCC za sve sekvence
PSNR	0.7527	0.6822	0.6982
SSIM	0.7842	0.7342	0.7334
MSSIM	0.7995	0.7321	0.7304
STRRED	- 0.7398	- 0.7414	- 0.7157
VMAF	0.8869	0.8436	0.8659
VsQM _{DASH}	0.2383	0.4471	0.2148
VQM _{Duanmu}	0.6038	0.6417	0.7032

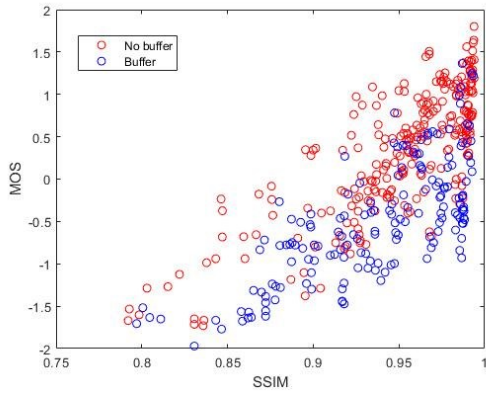
Za računanje PCC-a korišten je programski jezik MATLAB. Metrika VMAF ima najveći PCC za sve sekvence te također postiže najveći PCC za sekvence bez zastoja i sa zastojima. Najlošije rezultate u svim kategorijama postiže VsQM_{DASH}. Manja korelacija metrike VsQM_{DASH} je očekivana, jer samo trajanje zastoja i promjene kvalitete (rezolucije) nisu dovoljni za dobru procjenu QoE. STRRED jedini ima negativan koeficijent koji nam govori da postoji negativna korelacija na primjer smanjenjem objektivne ocjene raste subjektivna ocjena. PSNR metrika ima najmanje koeficijente korelacije iz NR skupine metrika. NR metrike (VsQM_{DASH} i VQM_{Duanmu}) imaju slabiji koeficijent korelacije nego FR metrike (PSNR, SSIM, MSSIM, STRRED i VMAF).



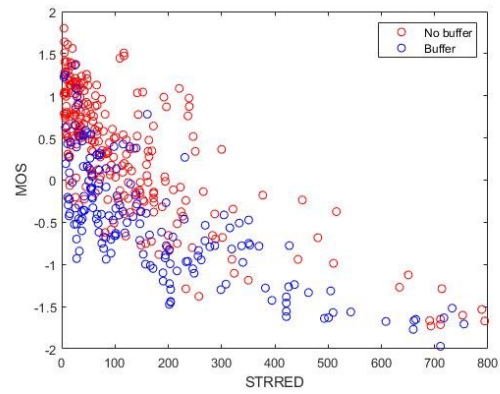
a)



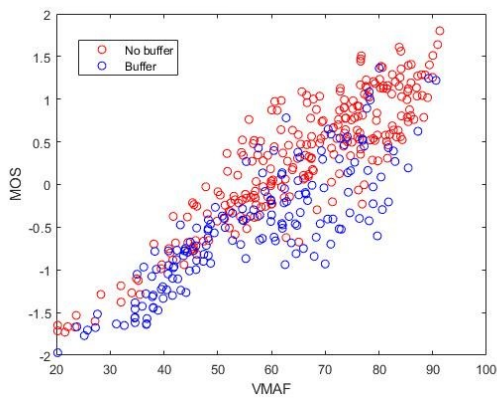
b)



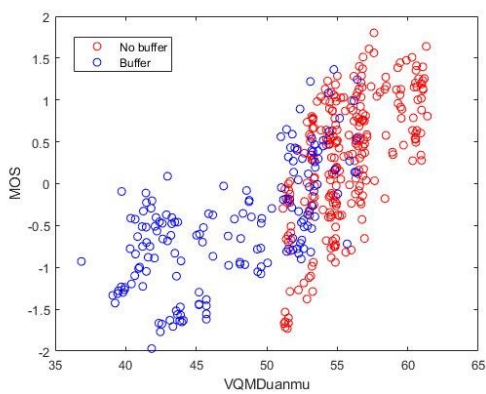
c)



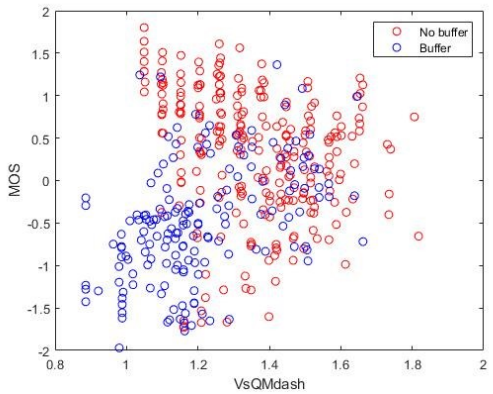
d)



e)



f)



g)

Slika 4.1. Raspršeni dijagrami modela predviđanja QoE vs. MOS podijeljeni na video signale sa zaustavljanjem i bez zaustavljanja: a) MMSIM metrika, b) PSNR metrika, c) SSIM metrika, d) STRRED metrika, e) VMAF metrika, f) VQMDunamu metrika, g) VsQMdash metrika

Raspršeni dijagrami objektivnih rezultata i MOS-a za algoritme korištene na cijeloj LIVE-NFLX-II QoE bazi podataka [24], su prikazani su na slici 4.1.

5. ZAKLJUČAK

U današnjem svijetu, ljudi imaju sve veće zahtjeve pri korištenju raznih aplikacija i društvenih mreža koje koriste video strujanje. Servisi poput Netflix-a rade testiranja kako bi ispunili očekivanja korisnika.

Adaptivno strujanje primjenom HTTP protokola (HAS) se pokazalo kao idealno rješenje jer se prilagođava stanjima u mreži. Za HAS video signali na poslužitelju su kodirani na više razina kvalitete, što podrazumijeva različite kodne brzine i različite rezolucije, obično prostorne, te podijeljeni u kraće segmente tako da za svaki segment klijent na korisničkoj strani odabire razinu kvalitete koja odgovara trenutnim mrežnim uvjetima i/ili razini popunjenosti međuspremnika. Ovisno o ABR-u, adaptivnom algoritmu za izbor razine kvalitete sljedećeg segmenta i ovisno o propusnosti mreže, kvaliteta videa na korisničkoj strani oscilira, a ponekad se pojavljuju i zastoji u prikazivanju. Kako bi osigurali što bolju iskustvenu kvalitetu gledateljima, operatori pokušavaju stvoriti kvalitetne objektivne metrike za ocjenu QoE, jer subjektivna testiranja traju dugo i zahtijevaju mnoga financijska ulaganja.

Velike baze podataka se koriste za testiranje i poboljšavanje objektivnih metrika. Uspoređivanjem MOS-a sa ocjenama dobivenih objektivnim metrikama saznaje se koliko su te metrike u korelaciji sa subjektivnim ljudskim doživljajem kvalitete strujanja.

Za ovaj rad korištena je javno dostupna baza podataka LIVE-NFLX-II. Uz osnovno znanje programskog jezika MATLAB testirano je 7 objektivnih metrika u odnosu na MOS. Kreirano je 7 programa čijim se pokretanjem dobije Pearsonov koeficijent korelacije za: video sekvence sa zaustavljanjem, bez zaustavljanja i za sve video sekvence. Također kao rezultat programi daju i raspršen dijagram objektivnih ocjena i MOS-a koji se koristi za dodatno proučavanje metrika za ocjenjivanje kvalitete video signala. VMAF metrika se pokazala najboljom jer ima najveći PCC u svim kategorijama, FR metrike su se pokazale točnije nego NR metrike.

LITERATURA

- [1] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, A buffer-based approach to rate adaptation: Evidence from a large video streaming service, *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, pp. 187–198, 2015.
- [2] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, Toward a practical perceptual video quality metric, Netflix, Los Gatos, CA, USA, Tech. Rep, 2016.
- [3] A. Vetro, The MPEG-DASH Standard for Multimedia Streaming Over the Internet, 2011. Dostupno na: <https://www.semanticscholar.org/paper/The-MPEG-DASH-Standard-for-Multimedia-Streaming-the-Vetro/99c01639bf9ee1098bdf676a5f5e34177be4a69d/figure/1> [16.8.2022.]
- [4] Te-Yuan Huang, Ramesh Johari, Nick McKeown, Matthew Trunnell, Mark Watson, A Buffer-Based Approach to Rate Adaptation: Evidence from a Large Video Streaming Service, Stanford University, Netflix, 2014.
- [5] C. Liu, I. Bouazazi, M. Gabbouj, Rate adaptation for adaptive HTTP streaming, Second annual ACM conference on Multimedia systems, pp. 169-174, San Jose, 2011.
- [6] D. Suh, I. Jang, S. Pack, QoE-enhanced Adaptation Algorithm over DASH for Multimedia Streaming, The International Conference on Information Networking, pp. 497–501, Phuket, 2014.
- [7] H. Zhang, X. Jiang, A QOE-driven approach to rate adaptation for dynamic adaptive streaming over http, IEEE International Conference on Multimedia & Expo Workshops, pp. 1-6, Seattle, 2016.
- [8] M. Kim, J. Park, K. Chung, Content-aware Rate Adaptation Scheme to Improve Stability in HTTP Adaptive Streaming, International Conference on Information Networking, pp. 401– 405, Da Nang, 2017.
- [9] B. Villa, P. Heegaard, Improving Fairness in QoS and QoE domains for Adaptive Video Streaming, The International Journal on Advances in Networks and Services, pp. 291-303, St. Maarten, 2012.
- [10] S. Winkler, P. Mohandas, The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics, *IEEE Transactions on Broadcasting*, 4, 54, pp. 660 –668, 2008.

- [11] ITU-R, Methodology for the subjective assessment of the quality of television pictures, Rec. BT.500, Geneva, Switzerland, 2012.
- [12] ITU-T, Subjective video quality assessment methods for multimedia applications, Rec. P.910, Geneva, Switzerland, 2008.
- [13] Wu, Hong Ren, and Kamisetty Ramamohan Rao, eds, Digital video image quality and perceptual coding, CRC press, 2017.
- [14] Bampis, Christos G., Zhi Li, Ioannis Katsavounidis, Te-Yuan Huang, Chaitanya Ekanadham, and Alan C. Bovik, Towards perceptually optimized end-to-end adaptive video streaming, arXiv preprint arXiv:1808.03898, 2018.
- [15] ITU Recommendation P.913, 2020. Dostupno na: <https://www.itu.int/rec/T-REC-P.913-201603-I/en> [16.8.2022.]
- [16] M. Shahid, A. Rossholm, B. Lövsström, and H.-J. Zepernick, Noreference image and video quality assessment: a classification and review of recent approaches, EURASIP J. Image Video Process., vol. 2014, no. 1, pp. 40, 2014.
- [17] MSU Quality Measurement Tool: Metrics information, 2022. Dostupno na: http://compression.ru/video/quality_measure/info.html [16.8.2022.]
- [18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004.
- [19] Z. Wang, E. P. Simoncelli, and A. C. Bovik, Multiscale structural similarity for image quality assessment, in Proc. IEEE Asilomar Conf. Signals Syst. Comput., vol. 2. Pacific Grove, CA, USA, pp. 1398–1402, 2003.
- [20] R. Soundararajan and A. C. Bovik, Video quality assessment by reduced reference spatio-temporal entropic differencing, IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 4, pp. 684–694, 2013.
- [21] Z. Li, A. Aaron, L. Katsavounidis, A. Moorthy, and M. Manohara, Toward a Practical Perceptual Video Quality Metric, 2016. Dostupno na: <http://techblog.netflix.com/2016/06/toward-practical-perceptual-video.html> [16.8.2022.]

- [22] Peak signal-to-noise ratio. Dostupno na: https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio [16.8.2022.]
- [23] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik, Multiscale structural similarity for image quality assessment, In The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, vol. 2, pp. 1398-1402, Ieee, 2003.
- [24] Bampis, Christos G., Zhi Li, Ioannis Katsavounidis, Te-Yuan Huang, Chaitanya Ekanadham, and Alan C. Bovik, Towards perceptually optimized adaptive video streaming-a realistic quality of experience database, IEEE Transactions on Image Processing, vol. 30, pp. 5182-5197, 2021.
- [25] R. Soundararajan and A.C. Bovik, Video quality assessment by reduced reference spatio-temporal entropic differencing, IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 4, pp. 684-694, 2013.
- [26] I. Katsavounidis, Dynamic optimizer — a perceptual video encoding optimization framework, Netflix Technology Blog, 2018.
- [27] Bampis, Christos G., and Alan C. Bovik, Learning to predict streaming video QoE: Distortions, rebuffering and memory, arXiv preprint arXiv:1703.00633, 2017.
- [28] Rassool, Reza, VMAF reproducibility: Validating a perceptual practical video quality metric, In 2017 IEEE international symposium on broadband multimedia systems and broadcasting (BMSB), pp. 1-2, IEEE, 2017.
- [29] Z. Duanmu, A. Rehman and Z. Wang, A Quality-of-Experience Database for Adaptive Video Streaming, in IEEE Transactions on Broadcasting, vol. 64, no. 2, pp. 474-487, 2018.
- [30] Pearson correlation coefficient. Dostupno na: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient [02.9.2022.]

SAŽETAK

U ovom završnom radu dan je pregled karakteristika adaptivnog videostrujanja primjenom HTTP protokola, čimbenika koji utječu na kvalitetu videa te objektivnih metoda za ocjene kvalitete videa u sustavima koji primjenjuju takvo videostrujanje. Provedena je analiza na bazi podataka LIVE-NFLX-II i proračunata korelacija ocjena dobivenih objektivnim i subjektivnim metrikama ocjenjivanja kvalitete video signala. Za analizu je napravljen program u MATLAB-u pomoću kojeg je izračunat Pearsonov koeficijent korelacije za sedam objektivnih metoda mjerenja QoE, na temelju podataka iz baze LIVE-NFLX-II.

Ključne riječi: adaptivno video strujanje, korisničko iskustvo, subjektivne i objektivne metrike

ABSTRACT

VIDEO QUALITY MESURMENT IN NETWORK VIDEO STREAMING SYSTEMS

This bachelor's thesis provides an overview of the characteristics of adaptive video streaming using the HTTP protocol, factors affecting video quality, and objective methods for video quality evaluations in systems that use such video streaming. An analysis was carried out on the LIVE-NFLX-II database and a calculated correlation of ratings obtained by objective and subjective metrics for evaluating video signal quality. For the analysis, a program was created in MATLAB, which was used to calculate the Pearson correlation coefficient for seven objective methods of measuring QoE, based on data from the LIVE-NFLX-II database.

Key words: adaptive video streaming, Quality of Experience, subjective and objective metrics

ŽIVOTOPIS

Kristijan Štrk rođen je 19. siječnja 2000. godine u Vinkovcima. Djetinjstvo provodi u Županji gdje pohađa Osnovnu školu Ivana Kozarca, a zatim upisuje Prirodoslovno – matematičku gimnaziju u Županji. Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek, preddiplomski studij računarstva upisuje 2018. godine.

PRILOZI

```
%MSSIM
Test1 = []; % average buffer MSSIM
Test2 = []; % with buffer MOS
Test3 = []; % average no buffer MSSIM
Test4 = []; % no buffer MOS
All = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = mean(T.MSSIM);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = mean(T.MSSIM);
        Test4(end+1) = MOS(k);
    end
    All(k) = mean(T.MSSIM);
end

PCC2 = corrcoef(Test3, Test4);
disp('pcc bez zastoja')
disp(PCC2(1,2));

PCC1 = corrcoef(Test1, Test2);
disp('pcc sa zastojem')
disp(PCC1(1,2));

PCC = corrcoef(All, MOS);
disp('pcc za sve')
disp(PCC(1,2));

plot(Test3, Test4, 'or', Test1, Test2, 'ob');
xlabel('MSSIM');
ylabel('MOS');
legend('No buffer', 'Buffer');
```

```

%PSNR
Test1 = []; % average buffer PSNR
Test2 = []; % with buffer MOS
Test3 = []; % average no buffer PSNR
Test4 = []; % no buffer MOS
All = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = mean(T.PSNR);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = mean(T.PSNR);
        Test4(end+1) = MOS(k);
    end
    All(k) = mean(T.PSNR);
end

PCC2 = corrcoef(Test3, Test4);
disp('pcc bez zastoja')
disp(PCC2(1,2));

PCC1 = corrcoef(Test1, Test2);
disp('pcc sa zastojem')
disp(PCC1(1,2));

PCC = corrcoef(All, MOS);
disp('pcc za sve')
disp(PCC(1,2));

plot(Test3, Test4, 'or', Test1, Test2, 'ob');
xlabel('PSNR');
ylabel('MOS');
legend('No buffer', 'Buffer');

```

```

%SSIM
Test1 = []; % average buffer SSIM
Test2 = []; % with buffer MOS
Test3 = []; % average no buffer SSIM
Test4 = []; % no buffer MOS
All = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = mean(T.SSIM);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = mean(T.SSIM);
        Test4(end+1) = MOS(k);
    end
    All(k) = mean(T.SSIM);
end

PCC2 = corrcoef(Test3, Test4);
disp('pcc bez zastoja')
disp(PCC2(1,2));

PCC1 = corrcoef(Test1, Test2);
disp('pcc sa zastojem')
disp(PCC1(1,2));

PCC = corrcoef(All, MOS);
disp('pcc za sve')
disp(PCC(1,2));

plot(Test3, Test4, 'or', Test1, Test2, 'ob');
xlabel('SSIM');
ylabel('MOS');
legend('No buffer', 'Buffer');

```

```

%STRRED
Test1 = []; % average buffer STRRED
Test2 = []; % with buffer MOS
Test3 = []; % average no buffer STRRED
Test4 = []; % no buffer MOS
All = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = mean(T.STRRED);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = mean(T.STRRED);
        Test4(end+1) = MOS(k);
    end
    All(k) = mean(T.STRRED);
end

PCC2 = corrcoef(Test3, Test4);
disp('pcc bez zastoja')
disp(PCC2(1,2));

PCC1 = corrcoef(Test1, Test2);
disp('pcc sa zastojem')
disp(PCC1(1,2));

PCC = corrcoef(All, MOS);
disp('pcc za sve')
disp(PCC(1,2));

plot(Test3, Test4, 'or', Test1, Test2, 'ob');
xlabel('STRRED');
ylabel('MOS');
legend('No buffer', 'Buffer');

```

```

%VMAF
Test1 = []; % average buffer VMAF
Test2 = []; % with buffer MOS
Test3 = []; % average no buffer VMAF
Test4 = []; % no buffer MOS
All = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = mean(T.VMAF);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = mean(T.VMAF);
        Test4(end+1) = MOS(k);
    end
    All(k) = mean(T.VMAF);
end

PCC2 = corrcoef(Test3, Test4);
disp('pcc bez zastoja')
disp(PCC2(1,2));

PCC1 = corrcoef(Test1, Test2);
disp('pcc sa zastojem')
disp(PCC1(1,2));

PCC = corrcoef(All, MOS);
disp('pcc za sve')
disp(PCC(1,2));

plot(Test3, Test4, 'or', Test1, Test2, 'ob');
xlabel('VMAF');
ylabel('MOS');
legend('No buffer', 'Buffer');

```

```

%VQMd
A = [];
B = [];
REGR = [];
Test1 = [];
Test2 = [];
Test3 = [];
Test4 = [];
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    P_r = T.rebuffer_duration_sec/T.video_duration_sec;
    broj_frames=size(T.payout_bitrate);
    BitRateSame=0;
    NumberOfChanges=1;
    StartCounter=1;
    bitrate_switch_mag = [];
    ScaledPayoutBitRate = mean(T.payout_bitrate);
    MOS(k) = T.retrospective_zscored_mos;
    for br_fram=1:(broj_frames(2)-1)
        btrc=double(T.payout_bitrate(br_fram));
        btrn=double(T.payout_bitrate(br_fram+1));
        if (btrc-btrn) == 0
            BitRateSame = BitRateSame+1;
        else
            BitRateSame = 0;
            StartCounter = br_fram+1;
            NumberOfChanges = NumberOfChanges+1;
        end
        bitrate_switch_mag(br_fram)=abs(btrn-btrc);
    end
    B_s = sum(bitrate_switch_mag)/NumberOfChanges;
    REGR(k)=-63.1*P_r+0.0079*ScaledPayoutBitRate+0.0010*B_s+49.7;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = REGR(k);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = REGR(k);
        Test4(end+1) = MOS(k);
    end
end

SCORR2 = corrcoef(Test3,Test4);
disp('pcc bez zastoja')
disp(SCORR2(1,2));

SCORR1 = corrcoef(Test1,Test2);
disp('pcc sa zastojem')
disp(SCORR1(1,2));

SCORR = corrcoef(REGR,MOS);
disp('pcc za sve')
disp(SCORR(1,2));

plot(Test3,Test4,'or',Test1,Test2,'ob');
xlabel('VQMDuanmu');
ylabel('MOS');
legend('No buffer', 'Buffer');

```

```

%VsQM
D = 'E:\Zavrsni database';
S = dir(fullfile(D, '*.mat'));
Test1 = [];
Test2 = [];
Test3 = [];
Test4 = [];
for k = 1:numel(S)
    T = load(fullfile(D, S(k).name));
    W=T.per_segment_encoding_width;
    H=T.per_segment_encoding_height;
    br_seg=size(T.per_segment_encoding_height,2)-1;
    sumSq=0;
    br_prom=0;
    for i = 1:br_seg
        Wc=double(W(i));
        Hc=double(H(i));
        Wn=double(W(i+1));
        Hn=double(H(i+1));
        if ((Wc-Wn)~=0) || ((Hc-Hn)~=0)
            br_prom=br_prom+1;
            Sr=((Wc-Wn)+(Hc-Hn))/(Wc+Wn+Hc+Hn);
            q=-0.006*Sr*Sr + 0.212*Sr - 1.239;
            sumSq=sumSq+q;
        end
    end
    L=T.rebuffer_duration_sec;
    Tdur=T.video_duration_sec;
    VsQMdash(k)= exp(-(1/Tdur)*(L+sumSq));
    MOS(k) = T.retrospective_zscored_mos;
    if (T.rebuffer_number ~= 0)
        Test1(end+1) = VsQMdash(k);
        Test2(end+1) = MOS(k);
    else
        Test3(end+1) = VsQMdash(k);
        Test4(end+1) = MOS(k);
    end
end

PCC2 = corrcoef(Test3,Test4);
disp('pcc bez zastoja')
disp(abs(PCC2(1,2)));

PCC1 = corrcoef(Test1,Test2);
disp('pcc sa zastojem')
disp(abs(PCC1(1,2)));
PCC = corrcoef(VsQMdash,MOS);
disp('pcc za sve')
disp(abs(PCC(1,2)));

plot(Test3,Test4,'or',Test1,Test2,'ob');
xlabel('VsQMdash');
ylabel('MOS');
legend('No buffer','Buffer');

```