

Univerza v Ljubljani

Fakulteta za elektrotehniko

Rudi Hrvatin

**Analiza učinkovitosti kodiranja videa po
standardu HEVC**

Diplomsko delo univerzitetnega študija

Mentor: prof. dr. Andrej Košir

Ljubljana, 2015

Zahvala

Iskrena hvala mojemu mentorju, prof. dr. Andreju Koširju, za vso pomoč, podporo in nasvete pri izdelavi diplomskega dela ter viš. pred. dr. Urbanu Burniku za usmerjanje in koristne nasvete.

Zahvaljujem se tudi družini in prijateljem za oporo ter pomoč pri študiju in pisanju diplomskega dela.

Posebna zahvala tudi Oddelku za razvoj širokopasovnega dostopa v Iskratelu, ki mi je bil v pomoč in podporo pri praktični izvedbi testov.

Kazalo

1 Uvod	1
2 Video kodeki	3
2.1 Zgodovina.....	3
2.2 Standard H.264/MPEG4-AVC – delovanje in novosti	5
2.2.1 Notranja napoved.....	7
2.2.2 Vmesna napoved.....	8
2.2.3 Transformacijsko kodiranje	8
2.2.4 Entropijsko kodiranje.....	9
2.2.5 Filtriranje v zanki.....	9
2.3 Standard HEVC – delovanje in novosti	9
2.3.1 Razdeljevanje slike	10
2.3.2 Kodirne enote.....	10
2.3.3 Napovedne enote.....	10
2.3.4 Transformacijske enote.....	11
2.3.5 Rezine in polja	12
2.3.6 Vzporedno računanje v valovni fronti	13
2.3.7 Notranja napoved.....	14
2.3.8 Vmesna napoved.....	14
2.3.9 Transformacija in kvantizacija.....	15
2.3.10 Entropijsko kodiranje.....	15
2.3.11 Filtriranje v zanki.....	16
2.3.12 Vzorcju prilagojen odmik	17
2.4 Primerjava tehnologij H.264/MPEG4-AVC in HEVC	18
3 Mere za analizo video kvalitete	21

3.1 Kategorizacija.....	21
3.2 Subjektivne mere za analizo video kvalitete	21
3.2.1 Rezultat srednje ocene	22
3.3 Primerjava objektivnih in subjektivnih mer	22
3.4 Objektivne mere za analizo video kvalitete.....	23
3.4.1 Mere s polno referenco	23
3.4.2 Mere z zmanjšano referenco	24
3.4.3 Mere brez reference	24
3.5 Utemeljitev našega izbora PSNR-ja in SSIM-ja	25
4 Izvedba testa	27
4.1 FFmpeg.....	27
4.2 Video posnetki.....	27
4.3 Orodje za merjenje kvalitete videov	28
4.4 Avtomatizacija in iTest.....	29
4.5 Potek obdelave testnih podatkov v testih	29
4.6 Rezultati testiranja.....	30
5 Zaključek	37
6 Viri in literatura	39

Kazalo slik

Slika 1: Hibridno video kodiranje	4
Slika 2: Osnovni blok diagram za hibridni video kodirnik	6
Slika 3: Razdeljevanje makroblokov in podmakroblokov za kompenzacijo napovedi gibanja	8
Slika 4: Kompenzacija napovedi gibanja z referenco na več predhodnih slik	8
Slika 5: Razdeljevanje 64×64 LCU-ja na CU-je različne velikosti	10
Slika 6: Simetrični a) in nesimetrični b) PU-ji	11
Slika 7: Primer razporeditve TU-jev v enem LCU-ju	12
Slika 8: Primer razdeljevanja 32×32 CU-ja na PU-je in TU-je	12
Slika 9: Primer slike razdeljene na devet polj	13
Slika 10: Nesimetrično gibalno razdeljevanje CU-ja na velikost 32×32	14
Slika 11: Entropijsko kodiranje HEVC	16
Slika 12: Primer jakostne lestvice v BO načinu za 8 bitne podatke	18
Slika 13: Modeli uporabljeni pri EO-ju	18
Slika 14: Merilna shema testa	29
Slika 15: Primer slike z malo popačenja a) in veliko popačenja b), vir slike: video Testament 1, kodek H.264, glejte tabelo 4.	31
Slika 16: Primerjava med HEVC in H.264/MPEG4-AVC pri videu Full HD ločljivosti in pri veliki količini gibanja	32
Slika 17: Primerjava med HEVC in H.264/MPEG4-AVC pri videu Full HD ločljivosti in majhni količini gibanja	33
Slika 18: Primerjava med HEVC in H.264/MPEG4-AVC pri videu SD ločljivosti in veliki količini gibanja	34

Slika 19: Primerjava med HEVC in H.264/MPEG4-AVC pri videu SD ločljivosti in majhni količini gibanja.....	35
--	----

Kazalo tabel

Tabela 1: Primerjava H.264/MPEG4-AVC -ja in HEVC-ja.....	19
Tabela 2: Povprečen prihranek bitne hitrosti za enake vrednosti PSNR-ja za različne zabavne aplikacije (angl. Entertainment Applications) [2].....	20
Tabela 3: MOS lestvica.....	22
Tabela 4: Razdelitev videov po različnih sklopih: gibanje, ločljivost	28
Tabela 5: Parametri testa.....	29

Povzetek

Video kodiranje je že dlje časa uveljavljen postopek za prenašanje in shranjevanje video vsebin. Ključno vlogo pri video kodiranju igrajo kodirniki, ki na različne načine stiskajo video posnetke.

To diplomsko delo se osredotoča na primerjavo dveh kodirnih postopkov, H.264/MPEG4-AVC in HEVC. H.264/MPEG4-AVC je standard, ki se trenutno najpogosteje uporablja za kodiranje video posnetkov, pred kratkim pa je izšel nov standard HEVC, ki naj bi izboljšal kodirno učinkovitost.

Cilj je analizirati nov kodirni standard HEVC in ga primerjati s trenutnim standardom, H.264/MPEG4-AVC. Za primerjavo smo uporabili dve objektivni meri za oceno video kvalitete, PSNR in SSIM. PSNR se osredotoča na primerjavo posameznih slikovnih elementov, brez ozira na korelacijo med sosednjimi slikovnimi elementi. SSIM pa razširi območje primerjave in upošteva tudi območje slikovnih elementov ter tako vrne rezultate, ki so bolj primerljivi s človeško zaznavo video kvalitete.

Ugotovili samo, da nov standard dosega svoja pričakovanja pri Full HD vsebinah, kjer pri isti kvaliteti videa porabimo pol manjšo pasovno širino za prenos ali shranjevanje videa v primerjavi s H.264/MPEG4-AVC.

Ključne besede: video kodiranje, HEVC, H.264/MPEG4-AVC, video, SSIM, PSNR

Abstract

Video coding is a well-known procedure used for transporting and saving videos. The key role in video coding is performed by video codecs, which compress videos in a number of ways.

This diploma thesis is focused on the comparison between two different coding standards, H.264/MPEG4-AVC and HEVC. H.264/MPEG4-AVC is the most widespread video coding standard currently adopted by the industry. Recently, a new standard was released, HEVC, which is believed to improve the coding efficiency by 50 percent compared to H.264/MPEG4-AVC.

The goal is to analyse the new HEVC coding standard and compare it to the currently used H.264/MPEG4-AVC standard. Two objective video quality metrics were used in the process, PSNR and SSIM. PSNR focuses on the comparison between single pixels, with no regard to the neighbouring pixels, whereas SSIM enlarges the area of comparison by taking into account a bigger area of pixels and gives results that are well correlated to the human visual system and its perception of quality.

The results show that the new HEVC standard meets the expectations for Full HD videos, where compared to H.264/MPEG4-AVC half of the bandwidth is used for the same quality video.

Key words: video coding, HEVC, H.264/MPEG4-AVC, video, SSIM, PSNR

1 Uvod

Na internetu je v letu 2012 video predstavljal 66 % vsega prometa. Raziskave kažejo, da naj bi se ta odstotek še povišal, tudi do 79 % do leta 2018 [1]. Hkrati postaja ločljivost video vsebin vedno višja, HD je postal že množično razširjen standard, sedaj industrija stremi k vsebinam UHD.

Vse te zahteve porajajo vprašanje, kako bo internet prenesel ta trend naraščanja količine video vsebin, tako v ločljivosti kot tudi v sami količini podatkov.

Ena od rešitev je razvoj novih, bolj izpopolnjenih tehnik zgoščevanja video vsebin. S tem namenom so začeli razvijati nove tehnike kodiranja. Eden izmed večjih prelomov je bil razvoj kodeka H.264/MPEG4-AVC – to je naprednejši video kodek (angl. H.264/MPEG-4 Advanced Video Codec, H.264/MPEG4-AVC), ki je trenutno najbolj uporabljen kodek v industriji. H.264/MPEG4-AVC je začasno zadostil potrebam po zgoščevanju, saj je v primerjavi s svojim predhodnikom omogočal dvakrat boljše zgoščevanje, za video enake kvalitete je torej potreboval pol manjšo pasovno širino.

Svojo nalogo je opravljal odlično, dokler so videi temeljili na ločljivostih do Full HD. Vendar postajajo naprave vse močnejše, zasloni sedaj omogočajo ločljivosti že do 4K. Najnovejši telefoni lahko že snemajo in predvajajo vsebino 4K ločljivosti. H.264/MPEG4-AVC je standardiziran za vsebine vse do 4K ločljivosti, vendar je pasovna širina, ki jo ob tem porabi, previsoka, da bi lahko take vsebine bile dostopne širši množici uporabnikov. Zato so začeli razvijati nov kodek, ki bo omogočal shranjevanje in prenašanje vsebin vse do 8K ločljivosti, torej do 7680×4320 slikovnih elementov, kot tudi bolj učinkovito zgoščevanje videov nižjih ločljivosti. Zaradi večje procesorske moči današnjih naprav so tako standardizirali učinkovitejši, vendar tudi kompleksnejši kodek.

Novi kodek, imenovan HEVC ali H.265, naj bi omogočal v primerjavi s svojim predhodnikom, H.264/MPEG4-AVC, dvakrat boljše zgoščevanje, torej naj bi za video enake kvalitete potreboval pol manjšo pasovno širino. To velja predvsem za

vsebine visoke in ultra visoke ločljivosti. Cilj te diplomske naloge je analizirati, če HEVC izpolnjuje napovedi in pričakovanja [2].

2 Video kodeki

Signal, ki prenaša informacije, lahko zgostimo tako, da iz njega odstranimo redundantne podatke. Poznamo dva načina zgoščevanja, izgubno in brezizgubno. Pri brezizgubnem zgoščevanju odstranimo statistično redundantne podatke, tako da je lahko signal kasneje ponovno enolično rekonstruiran. Pri videih z brezizgubnim zgoščevanjem dosežemo le skromne rezultate. Zato vidoe v praksi zgoščamo izgubno, a s tem pri rekonstrukciji dobimo le približek originalnega videa.

Cilj video zgoščevalnih algoritmov je optimalno zgoščevanje videov in hkrati čim manjše popačenje med samim zgoščevalnim procesom. Video kodiranje pa je proces zgoščevanja in razširjanja digitalnega video signala [3].

2.1 Zgodovina

Prvi prelomni trenutek v teoriji kodiranja videov se je zgodil daljnega leta 1929, ko je Ray Davis Kell predlagal, da bi namesto prenašanja vseh podatkov video okvirjev lahko prenašali samo razliko med njimi. Tako je naredil prvi osnutek in kasneje standard, ki je opisoval video zgoščevanje [4].

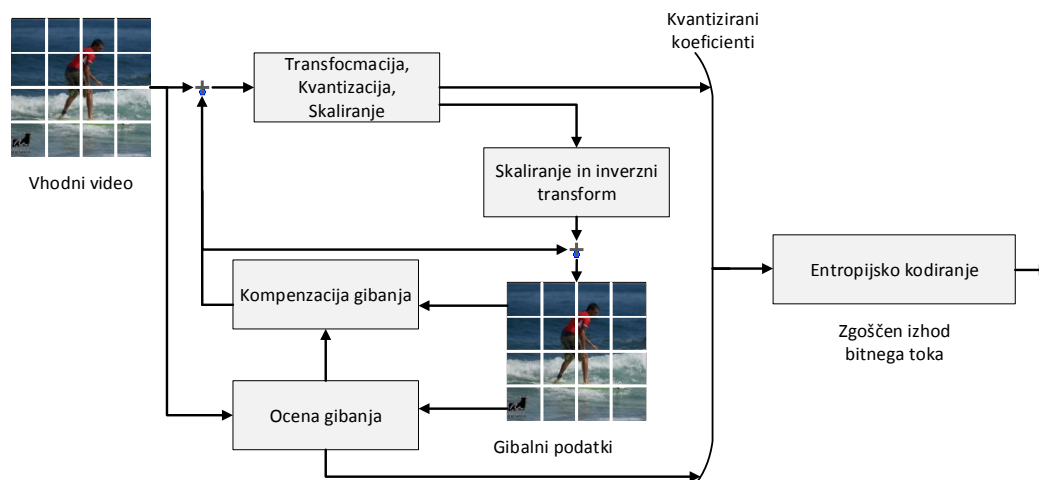
Kasneje je napredek v digitalnem procesiranju videa doprinesel veliko število kodekov, ki se še danes pogosto uporabljajo za zgoščevanje videa. Opisali jih bomo na podlagi njihovega nastanka, od starejšega do najnovejšega.

Med prvimi kodeki, ki so bili razviti in standardizirani je bil ITU-T-jev H.120 [5]. Nastal je leta 1984 in je temeljil na diferencialni impulzno-kodni modulaciji, skalarni kvantizaciji in kodiranju s spremenljivo dolžino. Z danim kodekom se je prenašalo vidoe NTSC in PAL formata preko namenske povezave točka-točka. Cilj kodeka so bile video konference. Njegovo obnovljeno različico lahko najdemo pod ITU-T-jevim standardom H.120 [6].

Prvi standard, ki je občutno izboljšal video kvaliteto v primerjavi s H.120, je bil razvit s strani ITU-T in se je imenoval H.261. Cilj standarda so bile video konference pri pasovnih širinah, ki so bile večkratnik 64 kbit/s. Osnova za to je bil ISDN priključek, ki je omogočal takšne pasovne širine. Že pri 384 kbit/s in ločljivosti 352×288 slikovnih elementov splošnega izmenjevalnega formata (angl. Common Interchange Format, CIF) ali ločljivosti 176×144 četrt CIF (QCIF) in pasovni širini okoli 1 Mbit/s je bila kvaliteta videa zadovoljiva. Ta standard je prvi uvedel hibridno video kodirno shemo in se še danes uporablja kot osnova za vse nadaljnje standarde.

Hibridno video kodiranje uporablja dve osnovni metodi [7]:

- napovedno kodiranje,
- transformacijsko kodiranje.



Slika 1: Hibridno video kodiranje

H.261 uporablja makrobloke velikosti 16×16 s kompenzacijo gibanja, diskretni kosinusni transform (angl. Discrete Cosine Transform, DCT) velikosti 8×8 slikovnih elementov, skalaro kvantizacijo, cikcak skeniranje in entropijsko kodiranje variabilne dolžine, ki temelji na Huffmanovem kodiranju [8]. Standard priporoča delovanje med $p \times 64$ kb/s ($p = 1$ do 30) [9].

V začetku leta 1990 je bila ustanovljena izvedenska skupina za gibljive slike (angl. Motion Picture Expert Group, MPEG), ki je kmalu razvila standard MPEG-1, katerega cilj je bil obdržati zadovoljivo video kvaliteto pri 1,5 Mbit/s in CIF ločljivosti. Njegova težava je ta, da podpira samo progresivno skeniranje. Večina takratnih standardov za analogno televizijo, kot so NTSC in PAL, pa temelji na skeniranju z vrstičnim preskokom. Ta manjkajoča karakteristika je botrovala razvoju in nastanku standarda MPEG-2, ki podpira skeniranje z vrstičnim preskokom.

MPEG-2 Part 2 / H.262 standard je bil razvit s skupnimi močmi ISO in ITU-T, ki je bil odobren leta 1994. Standard podpira video CIF ločljivosti, standardne ločljivosti (angl. Standard Definition, SD), torej video ločljivosti 720×576 , kot tudi video visoke ločljivosti (angl. Full High Definition, Full HD). Visoka ločljivost je definirana kot 1920×1080 . Podpora Full HD-ju v tem standardu je ustavila razvoj MPEG-3 standarda. MPEG-2 se še veliko uporablja, saj ga vsi standardi, kot so DVB-T, DVB-S in DVB-C podpirajo in uporabljajo. Uporaben je predvsem za video SD ločljivosti, saj porabi med 2-10 Mb/s pasovne širine. Za Full HD ločljivost se veliko manj uporablja, saj bi poraba pasovne širine bila občutno prevelika [10].

Dve leti kasneje, torej leta 1995, je ITU-T razvil H.263, ki se je dolga leta uporabljal kot glavni standard za video konference in video na mobilnih napravah.

Glavni cilj so bile nizke pasovne širine za mobilne brezžične komunikacije. H.263 temelji na predhodnih kodekih, predvsem H.261. Doprinesel je izboljšave v kodirni učinkovitosti ter odpornost na napake na račun večje kompleksnosti. H.263 ima več različic in veliko aneksov. Povezan je tudi z MPEG-4 standardom, saj sta se standarda začela razvijati ob istem času [5].

Razvoj MPEG-4 se je začel leta 1995. Sestavljen je iz več standardov, ki se osredotočajo na zgoščevanje zvočno vizualnih digitalnih podatkov. Njegova različica za kodiranje videa je definirana v MPEG-4 Part 2. Vgrajenih ima več funkcionalnosti predhodnih standardov in se je najpogosteje uporabljal za pretakanje videa preko interneta in na mobilnih omrežjih [11].

Naprave so se razvijale kot tudi kodeki za zgoščevanje. Tudi MPEG-4 Part 2 je postal star, nadomestil ga je MPEG-4 Part 10 / AVC ali H.264.

H.264/MPEG4-AVC je trenutno najbolj uporabljen video kodek. Uporablja ga YouTube, vsi proizvajalci in TV distributerji. Celotna video industrija temelji na njem. Zasnovan je bil leta 2003, zadnjih sprememb in dopolnitev je bil deležen 2014. H.264/MPEG4-AVC omogoča prenašanje video vsebin iste kvalitete pri polovični pasovni širini v primerjavi z MPEG-2 Part 2 / H.262. Drugače povedano, pri isti porabi pasovne širine bo vsebina zakodirana z H.264/MPEG4-AVC dvakrat boljše kvalitete kot pri MPEG-2 Part 2 / H.262 [2].

Industrija stremi k vedno višjim ločljivostim. Na trgu se pojavljajo vsebine UHD kvalitete, ki zahtevajo še učinkovitejše načine zgoščevanja. H.264/MPEG4-AVC ne omogoča dovolj učinkovitega zgoščevanja za video posnetke ultra visokih ločljivosti. Razvili so nov kodek, ki bo še za polovico zmanjšal potrebo po pasovni širini pri isti kvaliteti in tako omogočil prenašanje UHD vsebin. To je HEVC ali H.265.

2.2 Standard H.264/MPEG4-AVC – delovanje in novosti

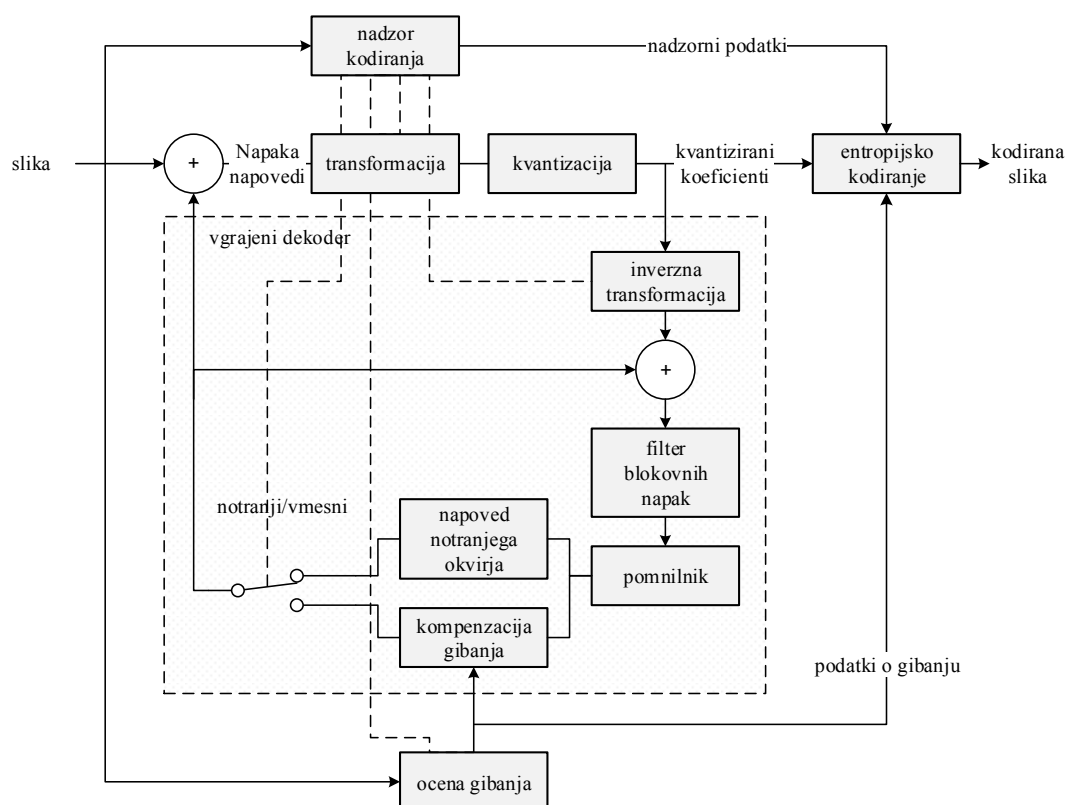
V industriji je trenutno najbolj razširjen standard za kodiranje video vsebin H.264/MPEG4-AVC. Uporablja se na veliko področjih, med drugim za digitalno satelitsko televizijo, digitalno prizemno televizijo, digitalno kabelsko televizijo, internetno televizijo itn. Njegovo delovanje in novosti, ki jih je doprinesel v primerjavi s prejšnjimi standardi, so zapisane spodaj.

Vhodna slika je najprej razdeljena na makrobloke. Vsak makroblok je sestavljen iz treh komponent; Y, Cr in Cb. Y komponenta predstavlja luminanco, Cr in Cb pa predstavljata krominantne komponente. Človeško oko je dosti bolj občutljivo na razlike v luminanci kot krominanci. Zato lahko oba krominantna

signala vzorčimo z 2-krat manjšim faktorjem v vertikalni in horizontalni smeri. Zaradi tega je en makroblok sestavljen iz 16×16 vzorcev luminance in 8×8 vzorcev krominance oz. barve.

Makrobloki so nato kodirani z uporabo vmesne ali notranje napovedi. Vmesni način pomeni, da makroblok napovemo s kompenzacijo napovedi gibanja. Informacijo zakodiramo v vektor premika, ki dekoderju pove, kje se je dani makroblok nahajal v predhodnih okvirjih. V notranjem načinu uporabljamo za napoved samo makrobloke znotraj istega okvirja. Tako izboljšamo odpornost na napake in hkrati omogočimo naključne točke vhoda na bitni tok. Razliko med originalnim makroblokom in napovedanim makroblokom imenujemo napovedna napaka. Napovedno napako transformiramo, kvantiziramo in entropijsko kodiramo. Na dekodirni strani kvantizirane koeficiente inverzno transformiramo in jih dodamo napovedanemu signalu. Tako dobimo rekonstruiran makroblok. Ta isti makroblok se bo pojavil tudi na dekodirni strani. Shranimo ga v spomin, saj z njegovo pomočjo napovedujemo naslednje makrobloke.

Osnovi diagram prikazuje slika 2.



Slika 2: Osnovni blok diagram za hibridni video kodirnik

Dekodirnik najprej entropijsko dekodira kvantizirane koeficiente in gibalne podatke, ki se uporabljajo za kompenzacijo napovedi gibanja. Napovedni signal dobimo iz notranje napovedi ali vmesne napovedi in ju dodamo inverzno

transformiranim koeficientom. Nato uporabimo še adaptivni filter blokovnih napak in dobimo makroblok, ki je dekodiran in shranjen v spominu za nadaljnjo napoved.

Makrobloke procesiramo v obliki rezin, kjer je rezina skupina več makroblokov in se skenira rastersko. H.264/MPEG4-AVC definira pet različnih vrst rezin. Te so I, P, B, SI in SP. I rezina vsebuje makrobloke, napovedane samo z notranjo napovedjo, v P rezini so vsi makrobloki napovedani z uporabo kompenzacije napovedi gibanja z enim referenčnim okvirjem in B rezine z dvema referenčnima okvirjema. SI in SP rezine so posebne vrste rezin, ki se uporabljajo za preklapljanje med dvema različnima bitnima tokovoma. [12]

Novosti, ki jih je H.264/MPEG4-AVC uvedel, obsegajo:

- uporaba adaptivnega filtra blokovnih napak;
- v spominu lahko hranimo več predhodnih video okvirjev, ne samo enega;
- napoved je uporabljena tudi pri notranjem načinu kodiranja;
- uporablja se celoštevilska transformacija, ki nadomešča DCT z natančnostjo v plavajoči vejici.

2.2.1 Notranja napoved

Vsi predhodniki H.264/MPEG4-AVC so se ukvarjali z napovedjo v časovnem prostoru in tako zanemarili pomembnost napovedi v krajevnem prostoru. H.264/MPEG4-AVC je prvi standard, ki je uvedel napoved znotraj video okvirja.

Pri notranji napovedi so vzorci makroblokov napovedani na podlagi informacij prisotnih znotraj že prenesenih makroblokov v isti sliki. H.264/MPEG4-AVC podpira dva načina notranje napovedi.

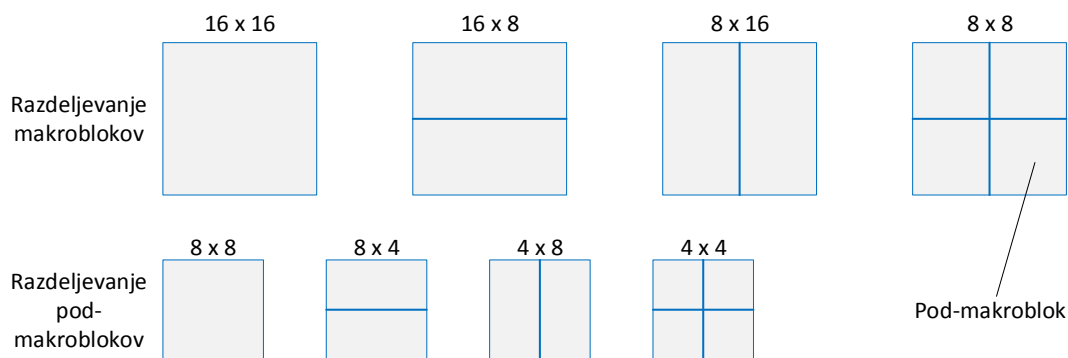
V prvem načinu, imenovanem INTRA_4 × 4, je makroblok velikosti 16 × 16 razdeljen na bloke velikosti 4 × 4, tako dobimo 16 podblokov. Za vsak podblok nato opravimo napoved luminance. Za napoved se lahko uporablja 9 različnih napovednih načinov.

V drugem načinu, imenovanem INTRA_16 × 16, imamo blok velikosti 16 × 16 in nad njim opravimo en napovedni način nad luminanco. Na razpolago imamo 4 različne napovedne načine.

Za krominanco uporabljamo način podoben INTRA_16 × 16 načinu, saj so krominantni signali v večini primerov zelo gladki. Opravljajo se na blokih velikosti 8 × 8.

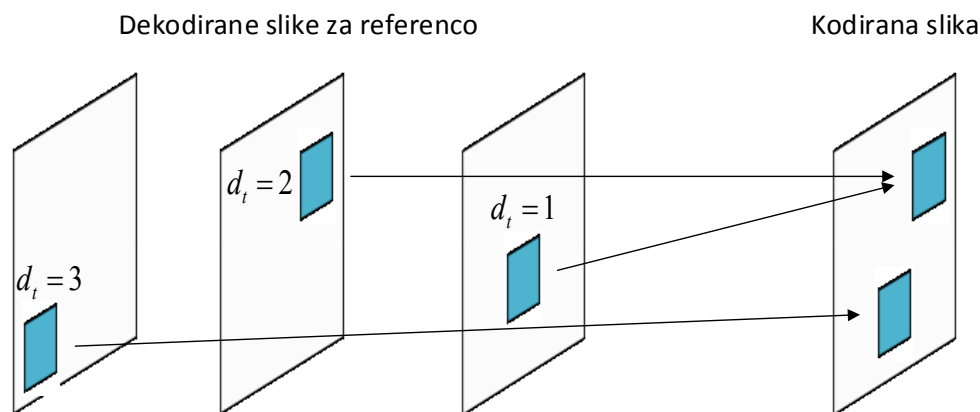
2.2.2 Vmesna napoved

V primeru, ko kompenziramo makrobloke z vmesno napovedjo, so makrobloki napovedani na podlagi že prenesenih referenčnih slik. Zaradi tega lahko vsak makroblok razdelimo na manjše dele, ki so lahko velikosti 16×16 , 16×8 , 8×16 in 8×8 . V primeru, da smo razdelili makroblok v P-rezini na 8×8 , lahko tega še naprej delimo na manjše dele, ki so lahko velikosti 8×4 , 4×8 in 4×4 . Razdeljevanje makrobloka na manjše delčke je vidno na spodnji sliki 3.



Slika 3: Razdeljevanje makroblokov in podmakroblokov za kompenzacijo napovedi gibanja

Poleg prej razdeljenih blokov se izračuna in pošlje še vektor premika samega bloka. Pri H.264/MPEG4-AVC lahko ta vektor premika vsebuje informacije o tem, kje se je dani blok nahajal na predhodnih slikah, tudi na večji množici predhodnih slik. Temu pravimo kompenzacija napovedi gibanja z referenco na več predhodnih slik. To je razvidno na sliki 4. Vektor premika se izračuna in pošilja z natančnostjo $\frac{1}{4}$ slikovnega elementa.



Slika 4: Kompenzacija napovedi gibanja z referenco na več predhodnih slik

2.2.3 Transformacijsko kodiranje

V H.264 so vpeljali nove transformacije, ki so podobne DCT. Imenujejo se celoštevilске transformacije. Uporablja se večinoma transformacija velikosti 4×4 , v

posebnih primerih se lahko uporabi tudi velikost 2×2 . Velikost blokov 4×4 , ki je v primerjavi s prejšnjimi kodeki za polovico manjša, omogoča večjo prilagodljivost.

2.2.4 Entropijsko kodiranje

H.264/MPEG4-AVC ima definirana dva načina entropijskega kodiranja. Kodiranje spremenljive dolžine s prilagajanjem okoliščinam (angl. Context-adaptive variable-length coding, CAVLC) in binarno aritmetično kodiranje s prilagajanjem vsebini (angl. Context-adaptive binary arithmetic coding, CABAC).

Obe metodi imata vgrajeno kontekstno modeliranje v svoji arhitekturni okvir, kar omogoča večjo prilagodljivost in s tem večjo kompleksnost.

CAVLC je izhodiščno entropijsko kodiranje v H.264/MPEG4-AVC, saj je podprto v vseh profilih, relativno enostavno za računanje in omogoča od 2 do 7 % prihranka pasovne širine [12]. Za bolj učinkovito kodiranje se uporablja CABAC. Temelji na binarizaciji, kontekstnem modeliranju in binarnem aritmetičnem kodiranju, vendar je računsko kompleksnejši, zato ni implementiran kot izhodiščni entropijski kodirnik.

2.2.5 Filtriranje v zanki

Še ena novost v H.264/MPEG4-AVC je filtriranje v zanki, ki je bila tukaj prvič postavljena v kodirni proces. Predhodniki so imeli filter vgrajen v dekodirnik, sedaj je ta proces vključen v kodirni proces, kar izboljša učinkovitost in hkrati omogoča ohranjanje kvalitete.

H.264/MPEG4-AVC temelji na blokovni shemi, zato pri transformaciji in blokovni kompenzaciji gibanja lahko prihaja do velikih blokovnih artefaktov. Zato so se odločili, da v kodirni zanki proces vključijo adaptiven filter blokovnih napak. S tem so vsi okvirji, ki so vključeni v kompenzacijo gibanja, podvrženi filtriranju rekonstruiranega okvirja.

2.3 Standard HEVC – delovanje in novosti

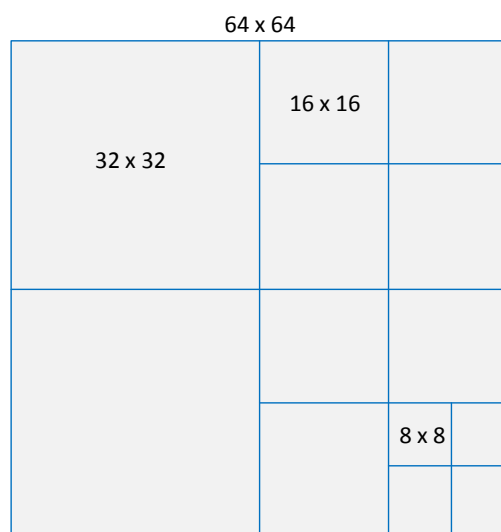
HEVC predstavlja novo prelomnico v razvoju video kodirnih standardov. Nastal je zaradi potrebe po učinkovitejšem zgoščevanju gibajočih slik. Narejen je tako, da je prilagodljiv množici različnih vrst omrežij in različnim aplikacijam. Njegove ciljne aplikacije so internetna televizija, digitalna televizija, mobilno pretakanje videov itd. Spodaj so opisane njegove novosti, poudarek je na novostih v primerjavi s H.264/MPEG4-AVC [13].

2.3.1 Razdeljevanje slike

Vsi moderni kodeki temeljijo na blokovni hibridni kodirni shemi, glej sliko 3, zato je bil tako zasnovan tudi HEVC. Osnovni blok je večji in bolj fleksibilen v primerjavi s predhodniki, kar je ključno za izboljšanje zmogljivosti in sposobnosti. Osnovni blok pri HEVC se imenuje največja kodirna enota (angl. Largest Coding Unit, LCU), ki je lahko razdeljena v manjše kodirne enote (angl. Coding Unit, CU), ki so lahko še dodatno razdeljene v napovedne enote (angl. Prediction Unit, PU) in transformacijske enote (angl. Transform Unit, TU).

2.3.2 Kodirne enote

HEVC razdeli vsako sliko v LCU-je velikosti do 64×64 . LCU-je lahko še naprej razdeli v manjše enote CU-je, ki so osnovna enota za notranje in vmesno kodiranje. Velikost CU-ja se giblje od maksimalne velikosti LCU-ja pa do 8×8 , odvisno od vsebnosti slike. V HEVC-u je zaradi rekurzivnega četrtinskega razdeljevanja definirana vsebini prilagojena struktura kodirnega drevesa, ki je sestavljena iz CU-jev. Slika 5 je primer razdeljevanja LCU-ja velikosti 64×64 na CU-je različnih velikosti.

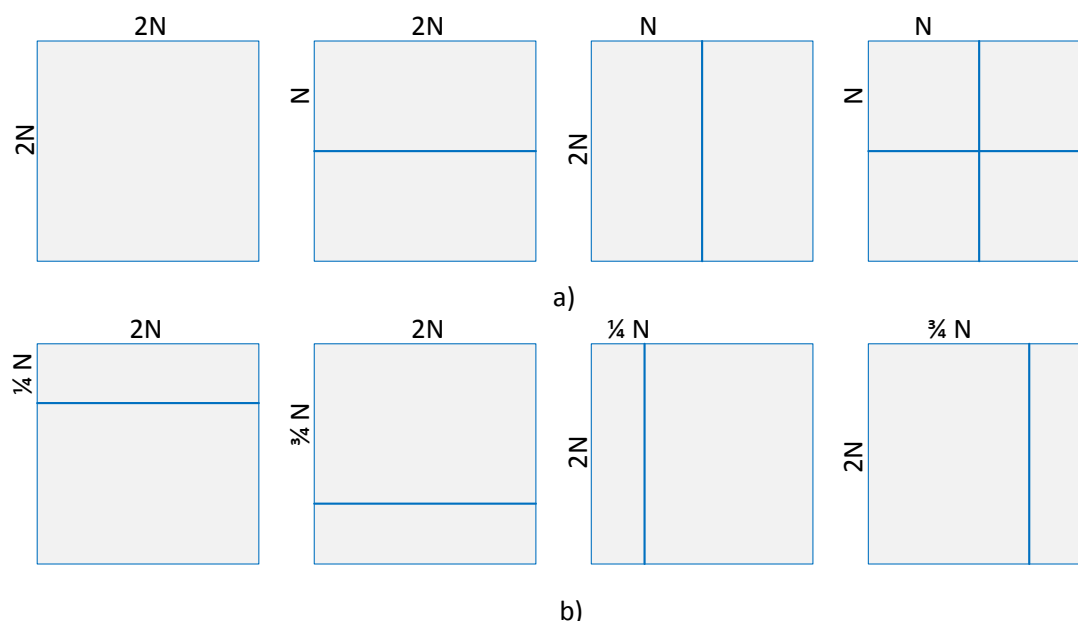


Slika 5: Razdeljevanje 64×64 LCU-ja na CU-je različne velikosti

2.3.3 Napovedne enote

Vsak CU je razdeljen na enega ali več PU-jev, katerih velikost se giblje od 4×4 pa do velikosti samega CU-ja. To so osnovne enote za opravljanje napovedi. Razdeljevanje CU-jev v PU-je se opravi samo enkrat, za razliko od razdeljevanja LCU-jev v CU-je, ki se lahko opravi večkrat.

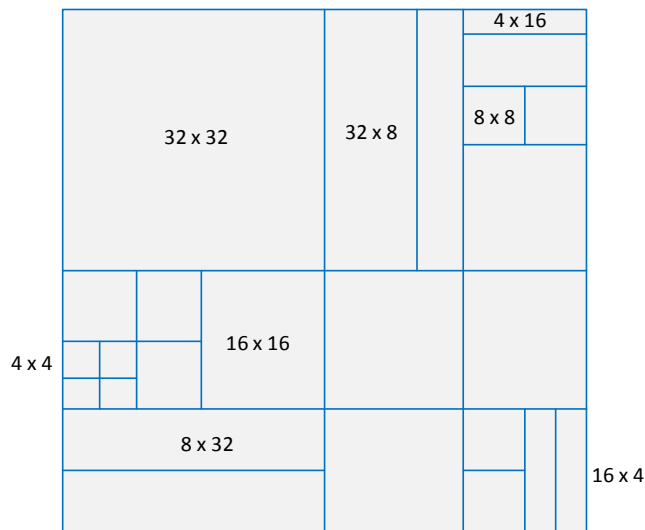
PU-ji so lahko simetrični ali nesimetrični. Simetrične se uporablja za notranjo in vmesno napoved, njegove oblike pa so lahko kvadratne ali pravokotne. Asimetrični PU-ji pa se uporabljajo samo pri vmesni napovedi. To omogoča razdeljevanje, ki točno pokriva robove objektov na slikah. Njihovo razdeljevanje je vidno na spodnji sliki 6.



Slika 6: Simetrični a) in nesimetrični b) PU-ji

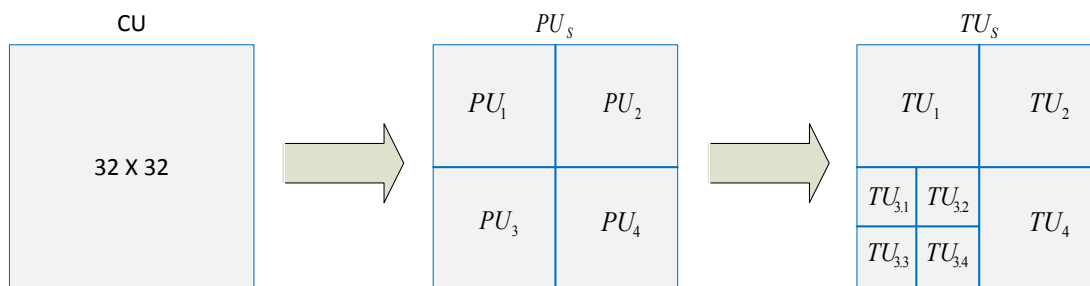
2.3.4 Transformacijske enote

HEVC opravlja celoštevilsko transformacijo nad odvečnimi podatki in jih tako dekorelira. Osnovna enota, ki se uporablja za transformacijo in kvantizacijo je transformacijska enota (angl. Transform Unit, TU). Velikost TU-jev je lahko od 4×4 pa do 32×32 . Lahko so tudi velikosti 32×8 , 8×32 , 16×4 in 4×16 luma vzorcev. Vsak CU lahko vsebuje enega ali več TU-jev. Na sliki 7 lahko vidimo primer, kako je več TU-jev razporejenih v enem LCU-ju.



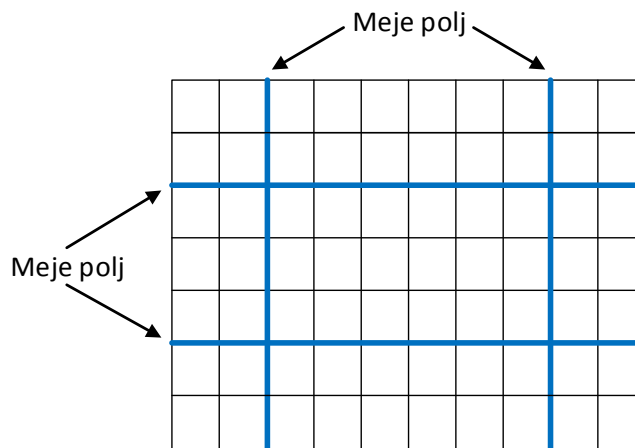
Slika 7: Primer razporeditve TU-jev v enem LCU-ju

Na sliki 8 pa vidimo primer, kako lahko 32×32 CU razdelimo na PU-je in TU-je.

Slika 8: Primer razdeljevanja 32×32 CU-ja na PU-je in TU-je

2.3.5 Rezine in polja

HEVC je za potrebe paralelnega procesiranja uvedel polja, ki omogočajo več fleksibilnosti v primerjavi z rezinami iz H.264/MPEG4-AVC in tudi očitno manj kompleksnosti. Polja so definirana s horizontalnimi in vertikalnimi mejami, s preseki, ki razdelijo sliko na pravokotne predele. Na sliki 9 vidimo primer razdeljevanja slike, ki vsebuje rezine.



Slika 9: Primer slike razdeljene na devet polj

Razdeljevanje robov polj na stolpce in vrstice ni nujno enakomerno, a s tem pristopom dodamo večjo fleksibilnost, ki je lahko uporabljena za večjo odpornost pri napakah. LCU-ji, ki se nahajajo znotraj enega polja, so procesirani rastersko, torej horizontalno od leve proti desni. Prav tako so rastersko procesirana polja znotraj slike.

HEVC podpira rezine na podoben način, kot jih podpira H.264/MPEG4-AVC, vendar brez fleksibilnega urejanja makroblokov (angl. Flexible Macroblock Ordering, FMO). Zato da lahko v HEVC-u vzporedno procesiramo, se rezine razdelijo na manjše rezine, imenovane entropijske rezine. Entropijske rezine so lahko dekodirane posamično, brez reference na ostale entropijske rezine. Tako lahko vsak procesor vzporedno procesira.

Prav tako lahko HEVC deluje s kombinacijo rezin in polj na isti sliki.

2.3.6 Vzporedno računanje v valovni fronti

Vzporedno računanje v valovni fronti (ang. Wavefront parallel processing - WPP) je mehanizem, ki poskrbi, da pri vzporednem procesiranju ne razbijemo napovednih odvisnosti ter da pri entropijskem kodiranju uporabimo čim večjo količino vsebine.

WPP deluje tako, da vsako vrstico LCU-jev začnemo procesirati v ločenem vzporednem procesu takoj, ko smo na trenutni vrstici že procesirali vsaj 2 LCU-ja. Dva zato, ker so lahko podatki za napoved gibanja in notranjo napoved odvisni od LCU-jev nad njima. Prav tako entropijsko kodiranje temelji na informacijah shranjenih na že kodiranih LCU-jih iz zgornje vrstice, kar omogoča uporabo največje količine vsebine pri novi kodirni niti.

2.3.7 Notranja napoved

HEVC ima zelo fleksibilno blokovno napovedno shemo znotraj slike, tako da lahko čim boljše izkorišča prostorsko korelacijo znotraj slike. HEVC ima 35 možnih luma vmesnihnapovednih načinov. V notranji napovedi so lahko tudi bloki različnih velikosti, od 4×4 pa vse do 64×64 . Število podprtih napovednih načinov je odvisno od velikosti PU-ja. HEVC podpira tudi ravninski notranji napovedni način, kar je zelo uporabno pri gladkih predelih slik.

Prav tako ima HEVC šest barvnih vmesnih napovednih načinov: direktni način (angl. Direct Mode, DM), linearni način (angl. Linear Mode, LM), vertikalni (način 0), horizontalni (način 1), DC način (način 2) in ravninski (način 3). Korelacija med kroma in luma komponentami je uporabljena pri DM in LM. Če sta kroma in luma komponenti podobni, je izbran DM. V tem primeru je isti način uporabljen za obe komponenti. V primeru, da intenziteta luma in kroma zelo korelirata, se uporablja LM. V tem primeru so kroma komponente napovedane iz luma komponent, ki so bile rekonstruirane iz zveze linearnega modela.

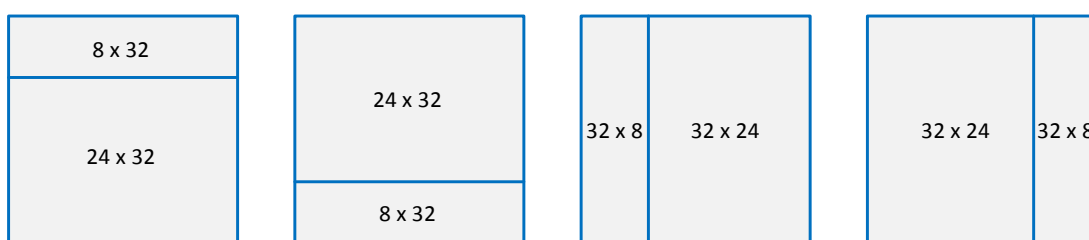
Oba načina izrabljata različne korelacije med luma in kroma komponentami. To je idealno, saj sta si na ta način komplementarna v kodirni zmogljivosti in sposobnosti. Zato sta najbolj v uporabi DM in LM.

2.3.8 Vmesna napoved

Variabilna PU velikost za kompenzacijo gibanja

Kot je bilo že zgoraj zapisano, se vsak LCU razdeli na manjše CU-je, ki se razdelijo na še manjše PU-je, ki so lahko kvadratni ali pravokotni. Vsak PU, ki ga kodiramo z vmesno napovedjo, ima skupek gibalnih parametrov. Sestavljajo jih vektor gibanja, referenčni indeks slike in referenčna zastavica.

Pri vmesno kodiranih CU-jih lahko uporabljamo simetrično ali nesimetrično gibalno delitev (angl. Asymmetric Motion Partitioning, AMP). S pomočjo AMP-ja lahko CU razdelimo na več manjših nesimetričnih PU-jev. Tako izboljšamo kodirno učinkovitost, saj lahko natančneje usklajujemo robove objektov v sliki brez dodatnih delitev. Primer na sliki 10.



Slika 10: Nesimetrično gibalno razdeljevanje CU-ja na velikost 32×32

Kodiranje gibalnih parametrov in izboljšan način izpuščanja

V HEVC-u so gibalni vektorji (angl. Motion Vector, MV) napovedani prostorsko ali časovno. Nova tehnika, ki jo je HEVC prinesel, je gibalno združevanje. Za vsak PU lahko kodirnik izbira med:

1. eksplisitnim kodiranjem gibalnih parametrov,
2. uporabo načina gibalnega združevanja,
3. uporabo izboljšanega načina izpuščanja (angl. SKIP).

Način gibalnega združevanja najprej ustvari spisek predhodno kodiranih sosedov (kandidatov) za PU, ki se kodira. Kandidati so prostorsko ali pa časovno blizu trenutnemu PU-ju. Kodirnik izbere kandidata iz spiska in kopira gibalne informacije v trenutno kodirani PU. Tako kodirniku ni potrebno shraniti in kodirati vektorja gibanja, ampak se lahko kodira samo indeks izbranega kandidata.

V SKIP načinu pa kodirnik tudi zakodira samo indeks izbranega kandidata, gibalni parametri za trenutni PU pa so skopirani iz izbranega kandidata. Na ta način lahko video okvirje, kjer je gibanje enakomerno, zakodiramo z zelo malo podatki.

2.3.9 Transformacija in kvantizacija

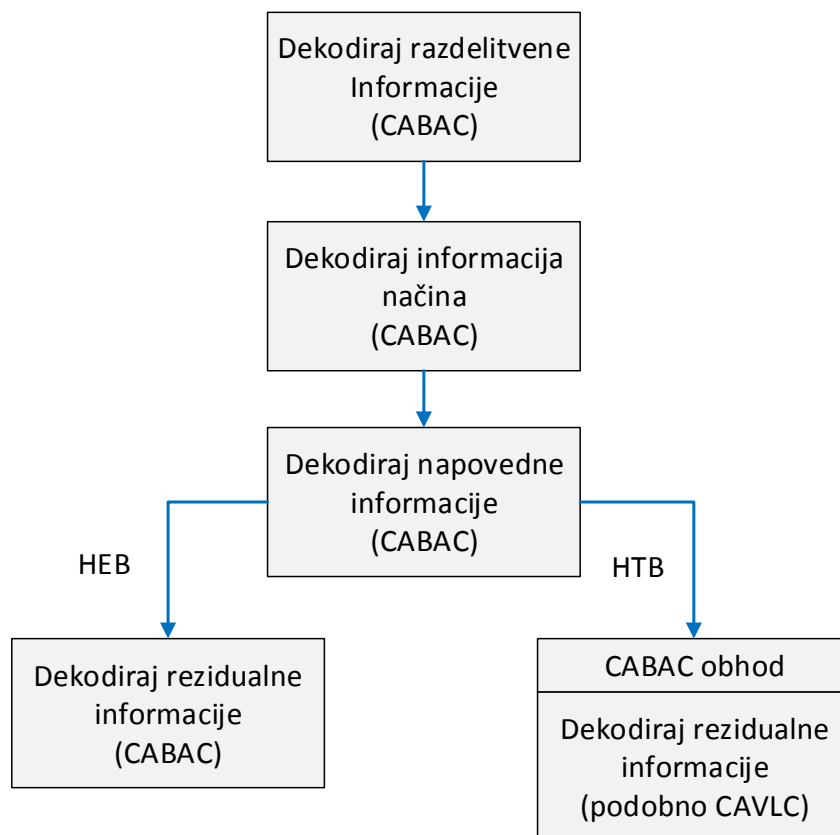
HEVC izvaja celoštevilsko transformacijo na ostankih napovedi. Transformacija je lahko aplicirana na blokih velikosti od 4×4 do 32×32 , HEVC pa podpira tudi transformacijo na nekvadratnih blokih, torej pravokotnih. Pravokotni bloki nimajo enakega števila transformacij po vrsticah in stolpcih. Osnovni vektorji imajo vsi enako energijo, zato ni potrebe po kompenzaciji različnih norm.

Še ena novost v HEVC-u je diskretna sinusna transformacija (angl. Discrete Sine Transform, DST), ki se uporablja pri notranji napovedi. Uporabna je zato, ker se osnovna funkcija DST-ja začne pri nizkih vrednostih in postopoma narašča do visokih. Ta lastnost je uporabna pri smerni notranji napovedi. V tem primeru je DCT popačil slikovne elemente, ki so bili zelo oddaljeni od napovedanih slikovnih elementov, DST pa se v danem primeru boljše obnese zaradi nizkih začetnih vrednosti.

2.3.10 Entropijsko kodiranje

HEVC ima dva načina za entropijsko kodiranje, in sicer visoko-pretočno binarizacijo (ang. High Throughput Binarization, HTB) in visoko učinkovito binarizacijo (angl. High-Efficiency Binarization, HEB). HEB način deluje samo s pomočjo CABAC. HTB pa poleg CABAC-a delno uporablja tudi CAVLC rezidualni način. HTB se uporablja za visoko pretočni način HEVC-a. Vsi sintaktični elementi

razen rezidualnih koeficientov so kodirani s pomočjo CABACA, rezidualne koeficiente pa zakodiramo s CAVLC-om. S to kombinacijo dobimo najboljše rezultate, saj uporabljamo najboljše lastnosti v CABAC-u in CAVLC-u. Več na sliki 11.



Slika 11: Entropijsko kodiranje HEVC

2.3.11 Filtriranje v zanki

Filtriranje v zanki se izvaja takoj za inverzno kvantizacijo in transformacijo, vendar ne pred rekonstrukcijo slike, ki se uporablja za kompenzacijo gibanja. Glej sliko 2. HEVC ima vgrajen filter blokovnih napak, podoben tistemu v H.264/MPEG4-AVC, vendar ima vgrajeni tudi dve novi funkciji: vzorčno adaptivno odmikanje (angl. Sample Adaptive Offset, SAO) in adaptivno zančno filtriranje (angl. Adaptive Loop Filtering, ALF). Ti dve funkciji naj bi zmanjšali popačenje, ki nastaja zaradi kodirnega procesa. Filter je vključen v napovedno zanko, saj bodo tako referenčne slike, ki se uporabljajo pri napovedi zaradi kompenzacije gibanja, boljše.

Filter blokovnih napak

Pri vseh blokovnih kompresijskih metodah prihaja do blokovnih artefaktov. Zato se pri le-teh uporablja nizkoprepustni filter, ki se aplicira na robovih blokov. Tako se izboljša kvaliteta, tako subjektivna kot objektivna.

HEVC uporablja filter blokovnih napak v zanki, podoben filtru v H.264/MPEG4-AVC. V HEVC-u imamo med bloki veliko število meja, od CU-jev pa do TU-jev. Za vsako od teh mej se odločamo, ali bomo uporabljali filter ter ali bomo uporabljali močno ali šibko filtriranje. Ta odločitev temelji na gradientih slikovnih elementov po mejah blokov in na QP samih blokov.

Adaptivno znančno filtriranje

ALF se aplicira na rekonstruiranem signalu, takoj za SAO-jem in filtrom blokovnih napak. Koeficienti filtra so signalizirani v bitnem toku, zato jih lahko prilagajamo vsebini slike ter imenujemo ta filter adaptivni filter. Cilj ALF-ja je, da je kvadratična napaka med izvorno in rekonstruirano sliko čim manjša.

V vsaki sliki lahko filter apliciramo na dva načina: na regijsko bazirani adaptaciji (angl. Region-Based Adaptation, RA) ali pa blokovno bazirani adaptaciji (angl. Block-Based Adaptation, BA). Pri RA sliko razdelimo na 16 regij enake velikosti. Te regije nato združimo med sabo. Vsaka regija, ki ostane po združevanju, bo imela svoj filter z unikatnimi koeficienti. BA pa bloke velikosti 4×4 razdeli v 16 kategorij glede na lastnosti gibanja na robovih in usmerjenosti. Te kategorije so nato lahko združene. Za vsako kategorijo posebej se nato izračuna filter, katerega koeficiente se lahko izračuna iz avtokorelacije in korelacije med originalnimi slikovnimi elementi in rekonstruiranimi slikovnimi elementi na tej regiji.

ACL se lahko izklopi ali vklopi, odvisno od razdelitve LUC-jev.

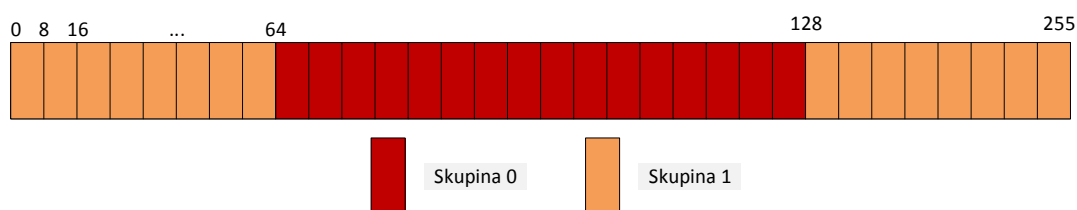
2.3.12 Vzorcju prilagojen odmik

Slikovne elemente kategoriziramo v različne kategorije in jim dodamo odmik, ki temelji na njihovi kategoriji. SAO klasificira rekonstruirane slikovne elemente v dve kategoriji, glede na intenziteto ali lastnosti robov. Nato doda odmik na podlagi kategorizacije, pasovni odmik (angl. Band Offset, BO) in robni odmik (angl. Edge Offset, EO). S tem zmanjša popačenje.

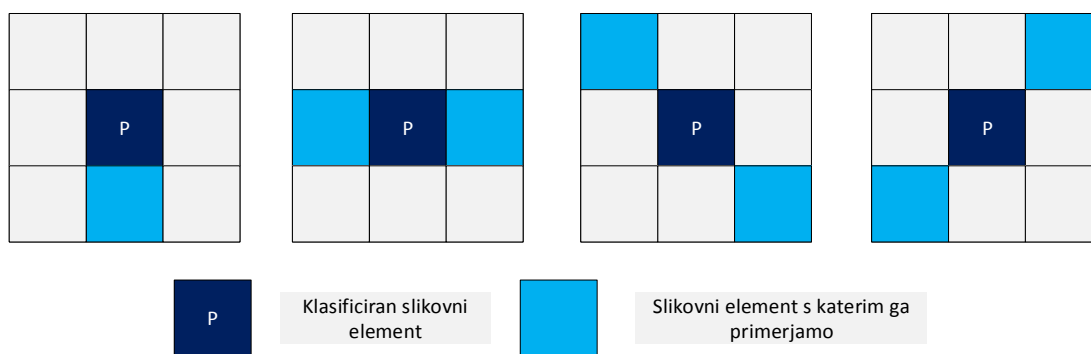
BO razdeli vse slikovne elemente iz nekega predela v več pasov, kjer vsak pas vsebuje slikovne elemente v istem intenzivnostnem intervalu. Ta interval je nato razdeljen na 32 enakih intervalov, ki gredo od nič pa do maksimalne intenzitete. Na sliki 12 je predstavljen primer za 8 bitne podatke. Ko vrednosti od 0 do 255

razdelimo, dobimo pasove širine 8 slikovnih elementov. Teh 32 pasov razdelimo na centralnih 16 in stranskih 16. Kodirnik se odloči, na katero skupino bo apliciral SAO, tako da bo kodirni odmik 16 v bitnem toku (glej sliko 12). EO razdeli slikovne elemente na osnovi smeri njihovih robov. Možni so štiri načini klasifikacije. Slikovni element je lahko vrh, dolina, rob ali nič od naštetega (glej sliko 13). Za te štiri kategorije bodo izračunani štirje odmiki.

Kodirnik se odloči, ali bo na določenih predelih slike uporabil BO ali EO. Lahko uporabi tudi oba.



Slika 12: Primer jakostne lestvice v BO načinu za 8 bitne podatke



Slika 13: Modeli uporabljeni pri EO-ju

2.4 Primerjava tehnologij H.264/MPEG4-AVC in HEVC

Kategorija	H.264/MPEG4-AVC	H.265 ali HEVC
Uporaba v industriji	Trenutno v prevladi	Se uvaja, glavni pobudniki so proizvajalci mobilnih telefonov in videa na zahtevo (angl. TV on-demand)
Ključne izboljšave	Za 40-50 % bolj učinkovit od svojega predhodnika MPEG-2. Omogočil priljubljenost HD vsebin.	Napovedujejo 40-50 % izboljšanje v primerjavi s predhodnikom H.264/MPEG4-AVC. Naj bi omogočil vsebine UHD velikosti.
Kompresijski model	Hibridno prostorsko-časovni napovedni model. Fleksibilno razdeljevanje makroblokov, podrazdeljevanje	Izpopolnjen hibridno prostorsko-časovni napovedni model. Fleksibilno razdeljevanje, uvedba elementov kodirnega

	za oceno gibanja. Notranja napoved. 9 načinov za notranjo napoved Makrobloki velikosti maksimalno 16×16 . Entropijsko kodiranje je CAVLC in CABAC.	drevesa (angl. Coding tree units) 35 načinov za notranjo napovedjo. Izboljšana arhitektura za vzporedno procesiranje, Podpora večjih blokov vse do 64×64 z večjo možnostjo razdeljevanja. Entropijsko kodiranje je CABAC.
Specifikacije	Podpira ločljivosti do 4K (4096x2304) in 60 fps	Podpira ločljivosti do 8K (8192x4320) in 300 fps
Slabosti	Podpira 4K, vendar je pasovna širina previsoka. Samo do 60 fps.	Procesorsko zahteven zaradi velikih blokov.

Tabela 1: Primerjava H.264/MPEG4-AVC in HEVC

Prednost HEVC-a pred H.264/MPEG4-AVC je predvsem v velikosti blokovnih struktur. HEVC podpira dosti večje blokovne strukture z veliko bolj fleksibilno razčlenitvijo, ki pa hkrati zahtevajo večjo procesorsko moč.

Napredek se vidi še v mehanizmu za vzporedno računanje v valovni fronti kodiranja in dekodiranja (WPP), pa tudi v več notranjih napovednih načinih (35 vsega skupaj). Več celoštevilskih transformacij z velikostmi od 4×4 pa vse do 35×35 , kot tudi pravokotne transformacije, ne samo kvadratne.

Izboljšali so tudi kodiranje informacij gibanja z novim načinom gibalnega združevanja, ki omogoča pošiljanje samo kazalca z informacijo o lokaciji predhodnega bloka.

Tudi procesiranje znotraj zanke so izboljšali z boljšim filtrom blokovnih napak, SAO-om in ALF [14], [15].

V spodnji tabeli 2 so prikazani rezultati, ki so bili pridobljeni v raziskavi [2] na videih, ki so bili zakodirani v kvaliteti primerni za zabavne aplikacije.

Kodirnik	Prihranek bitne hitrosti			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/H.262 MP
HEVC MP	35,4%	63,7%	65,1%	70,8%
H.264/MPEG-4 AVC HP	-	44,5%	46,6%	55,4%

MPEG-4 ASP	-	-	3,9%	19,7%
H.263 HLP	-	-	-	16,2%

Tabela 2: Povprečen prihranek bitne hitrosti za enake vrednosti PSNR-ja za različne zabavne aplikacije (angl. Entertainment Applications) [2]

3 Mere za analizo video kvalitete

Po spletu se pretaka vedno večja količina video vsebin. Industrija išče nove načine, kako prenašati video vsebino zadovoljive kvalitete s čim nižjo porabo pasovne širine.

Naloga mer za analizo video vsebin je med drugim ta, da določi še najmanjšo dopustno mejo zgoščevanja videa, da bodo gledalci še vedno zadovoljni z njegovo kvaliteto.

3.1 Kategorizacija

Mere za analizo video kvalitete v osnovi delimo na dve skupini: objektivne in subjektivne.

Objektivne mere se osredotočajo na razvoj analitičnih postopkov in algoritmov za določanje kvalitete videa, ki bi čim bolje odražali človeško zaznavo kvalitete videa in napak v njem.

Subjektivne mere temeljijo na človeku. Skupina ljudi ocenjuje video posnetke. Rezultat merjenja je povprečna ocena kvalitete različnih video posnetkov pri različnih stopnjah zgoščevanja [16].

3.2 Subjektivne mere za analizo video kvalitete

Subjektivne mere za analizo kvalitete posnetkov temeljijo na mnenju ljudi glede kvalitete slik in videa, torej na človeškem zaznavnem sistemu vida.

V subjektivnih merah je ocenjevalec človek. Skupina ljudi oceni kvaliteto video posnetkov. Najbolj poznana mera za subjektivno oceno je rezultat srednje ocene (angl. Mean Opinion Score, MOS).

Težava teh testov je, da je njihova izvedba zelo zahtevna, časovno kot tudi stroškovno. Ocenjevalce moramo tudi naučiti ocenjevati video posnetke.

Kljub vsemu so njihovi rezultati najbolj uporabni, saj človek kot ocenjevalec najboljše oceni kvaliteto, pod pogojem, da je dovolj ocenjevalcev in poteka ocenjevanje skladno s standardi.

3.2.1 Rezultat srednje ocene

MOS [17] je test, ki ga že dalj časa uporabljamo za ocenjevanje subjektivne kvalitete video vsebin. Množico ljudi (vsaj 16, da bodo rezultati statistično relevantni) posedemo v sobo, ki je opremljena za predvajanje in ocenjevanje videov po ITU-T P.910 [18] standardu ali ITU-T BT.500 [19] standardu in jim predvajamo posnetke. Ocenjevalci nato ocenijo vsak posnetek z lestvico od 1 do 5, kjer 1 pomeni zelo slabo kvaliteto, 5 pa odlično kvaliteto.

Na koncu izračunamo povprečno vrednost za vsak posnetek in dobimo lestvico kvalitete video posnetkov na osnovi kriterija, ki smo si ga zadali.

Tabela z lestvico interpretacije MOS ocen:

MOS	Kvaliteta	Poškodba
5	Odlična	Nezaznavna
4	Dobra	Zaznavno, a ne moteče
3	Zadovoljiva	Malo moteče
2	Slaba	Moteče
1	Neprimerno	Zelo moteče

Tabela 3: MOS lestvica

3.3 Primerjava objektivnih in subjektivnih mer

Človeška zaznava kvalitete videa je zelo zapletena, saj je sestavljena iz interakcije med človeških vizualnim sistemom, očmi in možgani. Dva glavna dejavnika, ki vplivata na subjektivno kvaliteto videa, sta točnost reprodukcije v časovnem in krajevnem prostoru. O krajevnem prostoru govorimo takrat, ko ugotavljamo, kako jasno so vidni deli scene, časovni prostor pa opredeljuje gladkost in naravnost gibanja [3]. Obstajajo še drugi dejavniki, kot so trenutno počutje ocenjevalca, okolica gledanja, pozornost ocenjevalca itd.

Zaradi zapletenosti človeške zaznave kvalitete videa so subjektivne ocene kvalitete optimalne, saj v primeru dovolj velikega števila ocenjevalcev in pravilnega

testiranja podajo najboljše rezultate. Vendar so ti testi zelo potratni, zato smo se odločili, da uporabimo objektivne mere za oceno kvalitete videa.

3.4 Objektivne mere za analizo video kvalitete

Objektivne mere za analizo kvalitete video posnetkov temeljijo na matematičnih modelih. Ti modeli poskušajo posnemati človeško zaznavo kvalitete videov in jo zapisati v algoritme.

Mere delimo na mere s polno referenco (angl. Full Reference Metrics, FR), mere z zmanjšano referenco (angl. Reduced Reference Metrics, RR) in mere brez reference (angl. No-Reference Metrics, NR).

3.4.1 Mere s polno referenco

Pri dani metodi imamo na voljo originalni (referenčni) video in poškodovani video. Mere s polno referenco se uporabljajo za vrednotenje in ocenjevanje video sistemov v nerealnem času [20].

Razmerje vršnih vrednosti signal-šum

Ena izmed preprostejših in najbolj uporabljenih objektivnih mer za analizo kvalitete video posnetkov je razmerje vršnih vrednosti signal-šum (angl. Peak Signal to Noise Ratio, PSNR). PSNR je logaritem razmerja med maksimalno možno vrednostjo moči signala in močjo šuma, ki kvari ta signal. Predstavljamo ga v logaritmični skali.

PSNR navadno definiramo s pomočjo srednje kvadratične napake (angl. Mean Squared Error, MSE), ki je tu mera za energijo. Če imamo referenčno enokanalno sliko I velikosti $m \times n$ in sliko K , s katero jo primerjamo, je MSE definiran kot:

$$MSE = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (3.1)$$

Tako torej dobimo PSNR:

$$PSNR = 10 \log_{10} \frac{MAX_I^2}{MSE} \quad (3.2)$$

PSNR je zelo preprost za računanje. Njegova glavna omejitev je, da temelji na numerični primerjavi in ne upošteva bioloških faktorjev človeškega vizualnega sistema. Zato smo uvedli še drugo mero, ki naj bi boljše odražala človeško zaznavo

napak v video posnetkih. Imenujemo jo indeks strukturne podobnosti (angl. Structural Similarity Index, SSIM).

Indeks strukturne podobnosti

Mera PSNR upošteva vrednosti med obema slikama po posameznih slikovnih elementih, zato lahko pri določenih oblikah šuma izračuna zelo slabo kvaliteto slike, v resnici pa na človeško oko ta šum zanemarljivo malo vpliva, saj ga sploh ne bi zmotil.

Zato so razvili novo mero SSIM, ki naj bi bolje modelirala način, kako ljudje zaznamo kvaliteto slike v primerjavi s prejšnjimi algoritmi.

SSIM primerja tri lastnosti, ki naj bi odražale človeško zaznavo kvalitete:

- Spremembe v luminanci. Primerjamo razlike v svetlosti obeh slik. Človeško oko ni pretirano občutljivo na svetlost, je pa zelo občutljivo na razlike v svetlosti.
- Spremembe v kontrastu. Primerjamo razpon med najsvetlejšimi in najtemnejšimi deli slike.
- Korelacijo med slikama. Primerjamo osnovno strukturo obeh slik in preverjamo, koliko se razlikujeta med sabo.

Mero SSIM izračunamo tako, da najprej obe sliki preuredimo do iste velikosti in ločljivosti. Tako lahko primerjamo slikovni element s slikovnim elementom. Nato izberemo okno določene velikosti. S tem oknom nato primerjamo obe sliki, za vse tri zgoraj omenjene lastnosti. Ta postopek opravimo na celotni sliki. Okno je ponavadi velikosti 8×8 . Na koncu dobimo tri vrednosti, katerim izračunamo povprečno vrednost.

Dobljeni rezultat je vrednost med -1 in 1, kjer vrednost 1 predstavlja primerjavo med dvema enakima slikama, torej najboljšo možno kvaliteto [21].

3.4.2 Mere z zmanjšano referenco

Pri tej metodi imamo na voljo samo del podatkov referenčnega posnetka, ki jih uporabimo za izračun.

Uporabljamo jih v podobnih primerih kot polne referenčne mere [20].

3.4.3 Mere brez reference

Poskušamo oceniti kvaliteto brez prisotnosti referenčnega video posnetka. Uporabljamo jih za ocenjevanje kvalitete videa v živo na sprejemni strani [20].

3.5 Utemeljitev našega izbora PSNR-ja in SSIM-ja

V našem testu smo želeli primerjati kvaliteto videov pri kodiranju z dvema različnima kodekoma. Najbolj optimalen rezultat bi dobili, če bi uporabili subjektivne mere, vendar za to nismo imeli sredstev in časa.

Odločili smo se za objektivne mere, vendar obstaja veliko število objektivnih mer, odločiti pa smo se morali za najprimernejše. Med seboj smo primerjali različne kodeke, zato mere, ki so prirejene določenemu kodeku, niso bile primerne. Nekatere mere so zahtevne za implementacijo ter so računsko zelo zahtevne, zato smo se med množico objektivnih mer odločili izbrati dve računsko enostavnejši, vendar najbolj razširjeni meri v primerjalnih testih različnih kodekov. Odločili smo se za PSNR in SSIM, saj je PSNR še vedno med najbolj uporabljenimi objektivnimi merami za primerjanje kvalitete videov med različnimi kodeki. SSIM pa se vedno bolj uveljavlja kot nova mera, ki naj bi bolje odražala človeško percepcijo kvalitete in bila primeren naslednik PSNR-ja [22].

4 Izvedba testa

V našem testu smo primerjali H.264/MPEG4-AVC in HEVC ter količino pasovne širine, ki jo porabita pri enaki kvaliteti videa. Zato smo uporabili FFmpeg za kodiranje in dekodiranje, orodje za merjenje kvalitete videov (angl. Video Quality Measurement Tool, VQMT) za računanje PSNR-ja in SSIM-ja ter profesionalno rešitev za avtomatizacijo testov iTest.

4.1 FFmpeg

FFmpeg je odprtokodni programski paket [23], ki omogoča shranjevanje, kodiranje, dekodiranje in prenašanje videa. Vse njegove funkcionalnosti so dostopne preko ukazne vrstice.

FFmpeg ima implementiranih veliko različnih video kodekov, med njimi tudi najnovejšo različico odprtokodnega kodeka x265, ki omogoča zgoščevanje videa skladno s standardom HEVC. Hkrati ima implementiran tudi kodek x264, ki je najbolj uporabljena odprtokodna rešitev za zgoščevanje videa po standardu H.264/MPEG4-AVC.

Izbrali smo ga, saj je brezplačen in omogoča vse funkcije, ki jih potrebujemo za analizo našega problema [23].

4.2 Video posnetki

Za testiranje smo pregledali množico videov in iz nje izbrali najprimernejše. Odločili smo se za podmnožico videov iz baze na Tehnični univerzi v Munchnu (nem. Technische Universität München, TUM) [24]. Videi so shranjeni v obliki slik s končnico ».tif«. Vsak video sestavimo tako, da 300 sličic zakodiramo v izbrani kodek. Videi so kategorizirani na podlagi ločljivosti, količine gibanja in metode za skeniranje slike. Uporabljali smo izključno progresivno skenirane videe, saj standard HEVC ne podpira več videov narejenih s skeniranjem z vrstičnim preskokom. V

tabeli 4 je prikazana podmnožica izbranih videov, razdeljenih na podlagi ločljivosti in količine gibanja [24].

Naslov video vsebine	Ločljivost	Gibanje
Testament 2	SD	Malo
Testament 3	SD	Malo
Carving 1	Full HD	Malo
Carving 4	Full HD	Malo
Testament 1	SD	Veliko
Testament 5	SD	Veliko
Ski 1	Full HD	Veliko
Ski 2	Full HD	Veliko

Tabela 4: Razdelitev videov po različnih sklopih: gibanje, ločljivost

Množica videov je bila velika, vendar smo se odločili za dano podmnožico zato, ker je reprezentativna. Vsebuje ločljivosti, ki se najpogosteje pojavljata na televiziji in na svetovnem spletu, SD in Full HD. Razdeljena je še po količini gibanja, kar je primerno za testiranje kodekov, saj se drugače obnašajo pod različnimi količinami gibanja. Za naša testiranja je taka majhna podmnožica zadovoljiva, saj smo vsak video zakodirali 13-krat v različni kvaliteti, z vsakim kodirnikom in tako dobili veliko količino videov različnih kvalitete.

4.3 Orodje za merjenje kvalitete videov

PSNR in SSIM smo računali z uporabo programa VQMT. Program omogoča izračun različnih objektivnih mer. Za nas je bilo pomembno, da program omogoča izračun PSNR-ja in SSIM-a.

Računanje PSNR-ja ali SSIM-a izvedemo s pomočjo ukazne vrstice. VQMT vrne rezultat v obliki tekstovne datoteke, ki vsebuje vrednosti PSNR-ja za vsak video okvir, prav tako pa povprečno vrednost.

VQMT temelji na implementaciji OpenCV knjižnice napisane v programskem jeziku C++. Omogoča zadovoljivo hitrost za računanje danih mer, rezultate vrne v obliki tekstovne datoteke z vrednostmi za vsak okvir posebej, na koncu pa nam poda tudi povprečno vrednost celotne primerjave.

4.4 Avtomatizacija in iTest

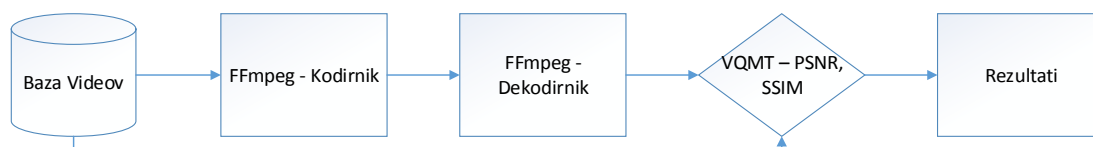
Ves postopek izvedbe testiranja smo avtomatizirali. Za avtomatizacijo so nam na podjetju Iskratel d.o.o. omogočili uporabo orodja Spirent iTest ali na kratko iTest. Glavna naloga iTest-a je, da omogoča hitro in preprosto načrtovanje testov pri razvoju novih produktov.

Z njegovo pomočjo lahko kličemo ukazno vrstico in v njo vnašamo ukaze. Prav tako lahko kličemo oddaljene naprave z uporabo protokolov Telnet ali varnostne lupine (angl. Secure Shell, SSH).

Z njegovo pomočjo smo avtomatizirali kodirni postopek kot tudi izračun mer. Tako smo lahko brez večjih težav avtomatsko kodirali posnetke z množico različnih parametrov.

4.5 Potek obdelave testnih podatkov v testih

Testiranje izvajamo na enem računalniku. Najprej zakodiramo vse videe s H.264/MPEG4-AVC in HEVC v različnih kvalitetah. Nato te posnetke dekodiramo in jih primerjamo z originali iz naše podmnožice videov s pomočjo programa VQMT. VQMT nam vrne rezultat za vsak video okvir posebej, kot tudi povprečno vrednost med vsemi video okvirji. Tako dobimo vrednosti za PSNR in SSIM za vse zakodirane posnetke. Merilna shema je narisana na sliki 14.



Slika 14: Merilna shema testa

Videe kodiramo z različnimi kodeki ter različnim faktorjem konstantne hitrosti (angl. Constant Rate Factor, CRF).

CRF bo spreminjal kvantizacijski parameter tako, da bo video ohranjal enakomerno kvaliteto skozi celoten potek [25].

Glej tabelo 5 za parametre testa.

Kodirnik	CRF
x.264	18 do 30
x.265	18 do 30

Tabela 5: Parametri testa

4.6 Rezultati testiranj

V tem poglavju predstavljamo rezultate testiranj. Primerjamo dva različna kodeka, pri enaki ločljivosti videov in primerljivi količini gibanja na videu. Rezultati so prikazani kot mera kvalitete v obliki PSNR in SSIM v primerjavi s porabljeno povprečno bitno hitrostjo videov pri različnih kvalitetah.

Pri kodirnem postopku prihaja do izgubnega zgoščevanja, zato je na videih vidno popačenje. Bolj stisnemo sliko, bolj popačena je. Na sliki 15 a) vidimo video okvir z manj popačenja, na sliki 15 b) pa z veliko popačenja. Vidna je razlika v barvah, najbolj očitno pa se popačenje vidi na človeški koži in v detajlih na video okvirjih. Robovi med objekti na zgornji sliki so dosti manj popačeni kot pa na spodnji.



a)



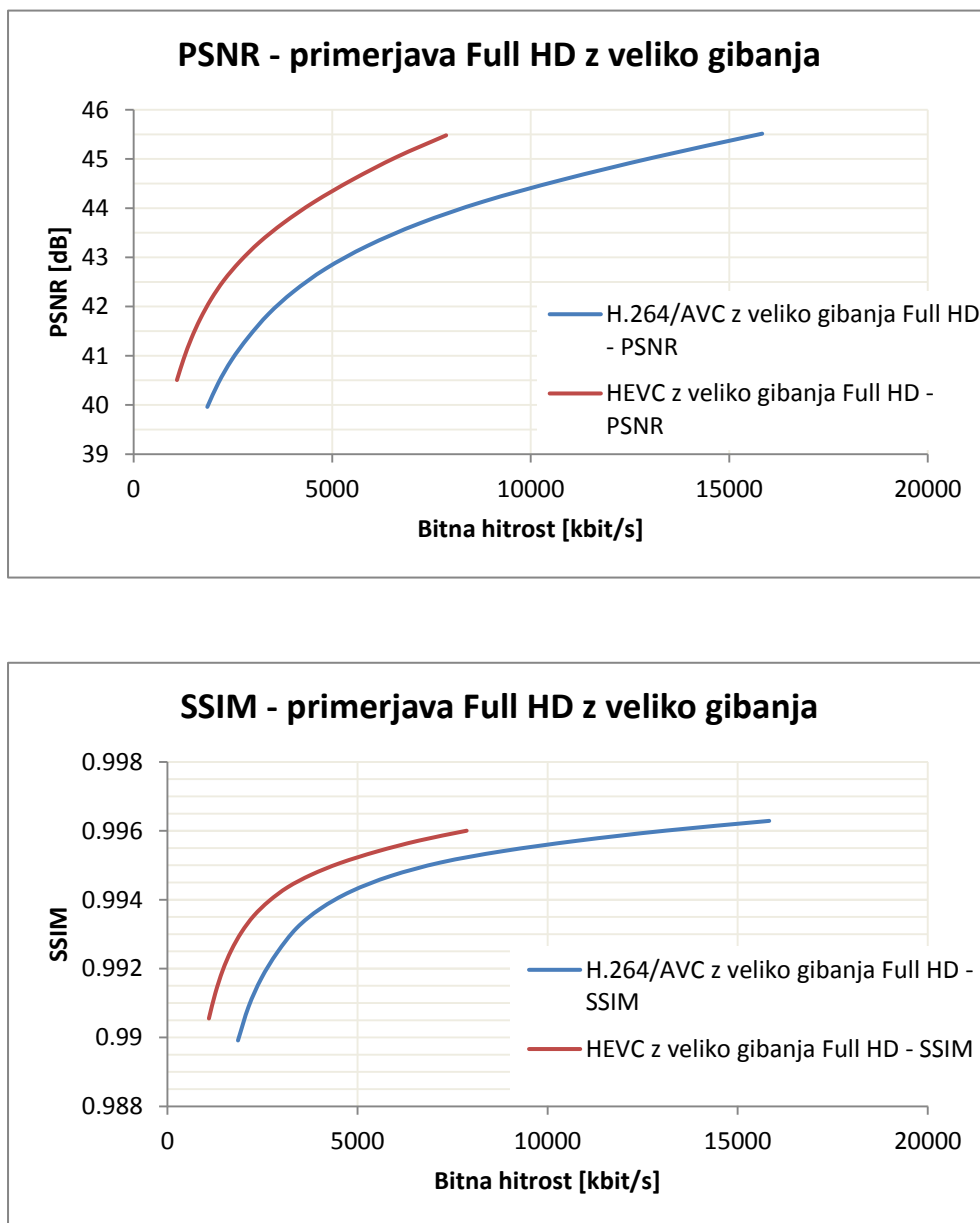
b)

Slika 15: Primer slike z malo popačenja a) in veliko popačenja b), vir slike: video Testament 1, kodirnik x.264, glej tabelo 4.

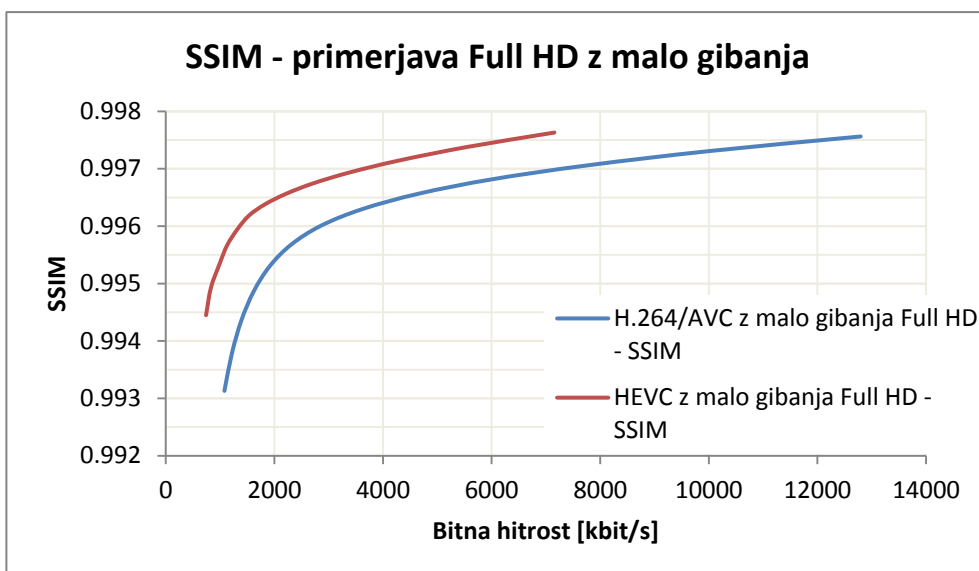
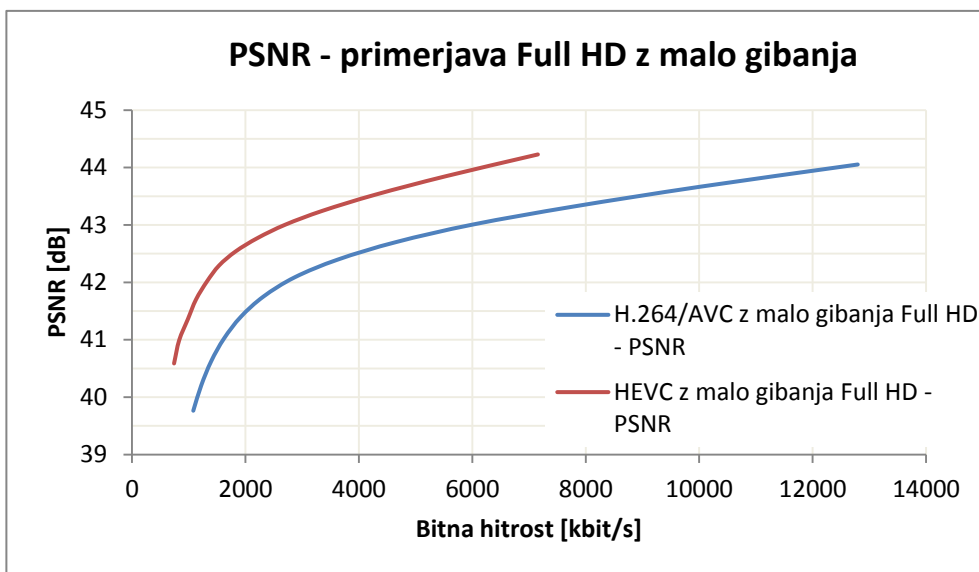
Na spodnjih grafih je razvidno, da kodirnik HEVC dosega pričakovane izboljšave pri videih Full HD ločljivosti v primerjavi s kodekom H.264/MPEG4-AVC, kar je razvidno na sliki 16 in sliki 17. Najbolj očitne izboljšave se pokažejo pri videih z veliko gibanja in Full HD ločljivostjo, kjer pridejo do izraza njegove izboljšave v obliki bolj fleksibilnega razdeljevanja slike na enote (oz. bloke, kot jih poimenuje H.264/MPEG4-AVC) ter v več možnih načinih za notranjo napoved.

Pri videih z manj gibanja in Full HD ločljivostjo so izboljšave še vedno zelo vidne. Najbolj se ponovno odražajo pri visoki kvaliteti, kar je tudi bil eden od ciljev razvoja HEVC-a.

Če meri PSNR in SSIM primerjamo med sabo, lahko vidimo tendenco krivulj pri PSNR-ju, ki dokaj enakomerno padajo, pri SSIM-ju pa se pokaže njegova korelacija s človeško percepcijo kvalitete v kolenu, ki se pojavi pri določeni vrednosti. Tam prihaja do hitrejšega padanja kvalitete, kar bi lahko nakazovalo na to, da bi tudi človeški opazovalci to padanje zaznali.

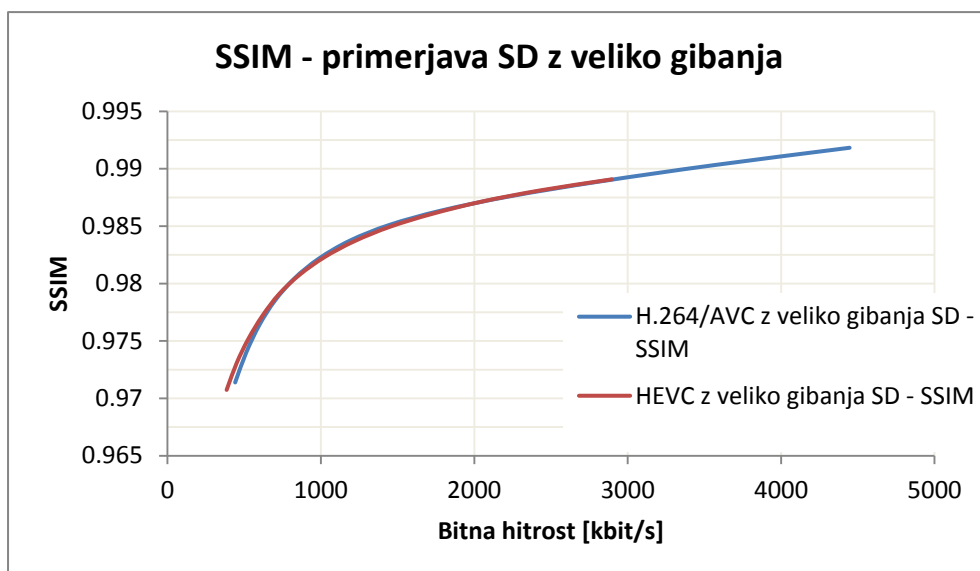
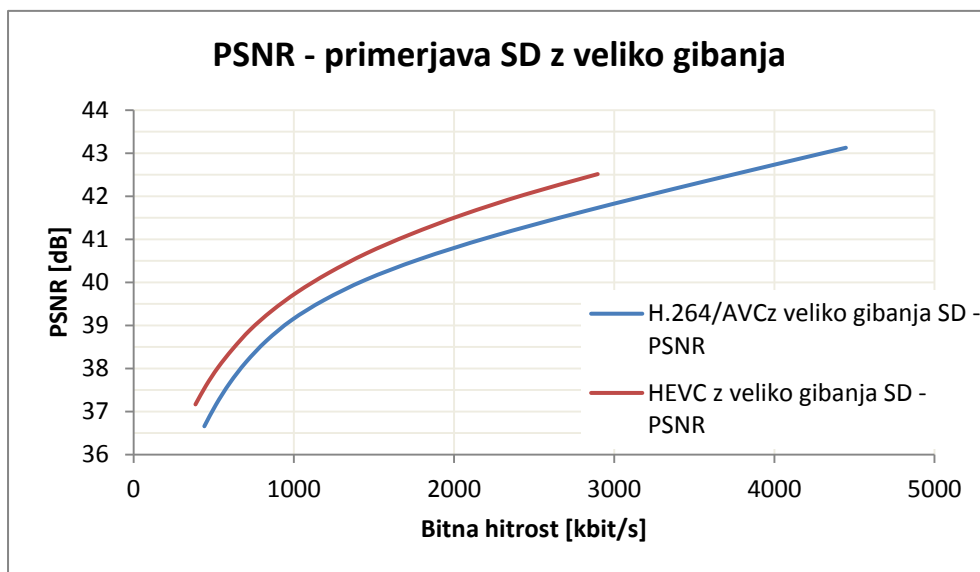


Slika 16: Primerjava med HEVC in H.264/MPEG4-AVC pri videu Full HD ločljivosti in pri veliki količini gibanja

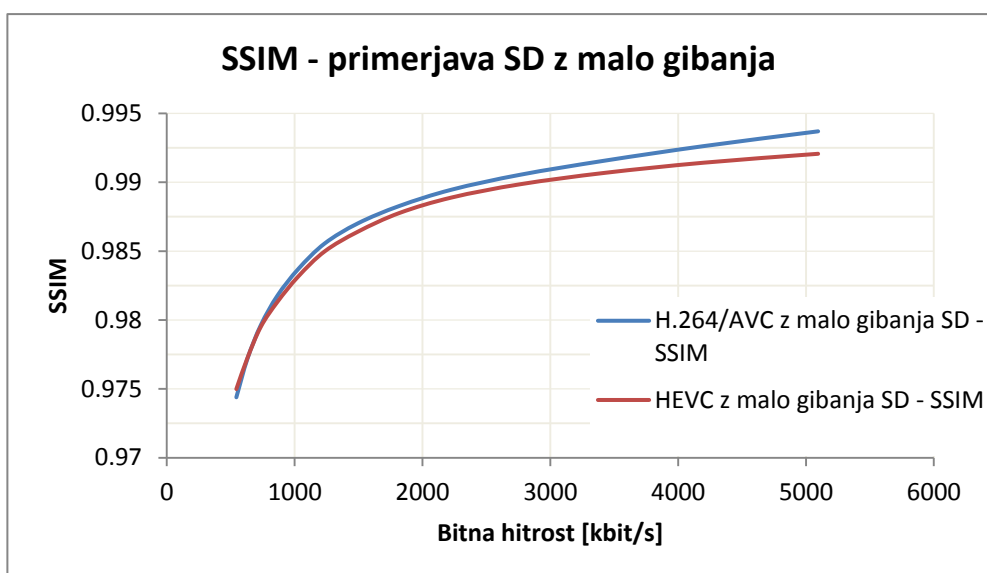
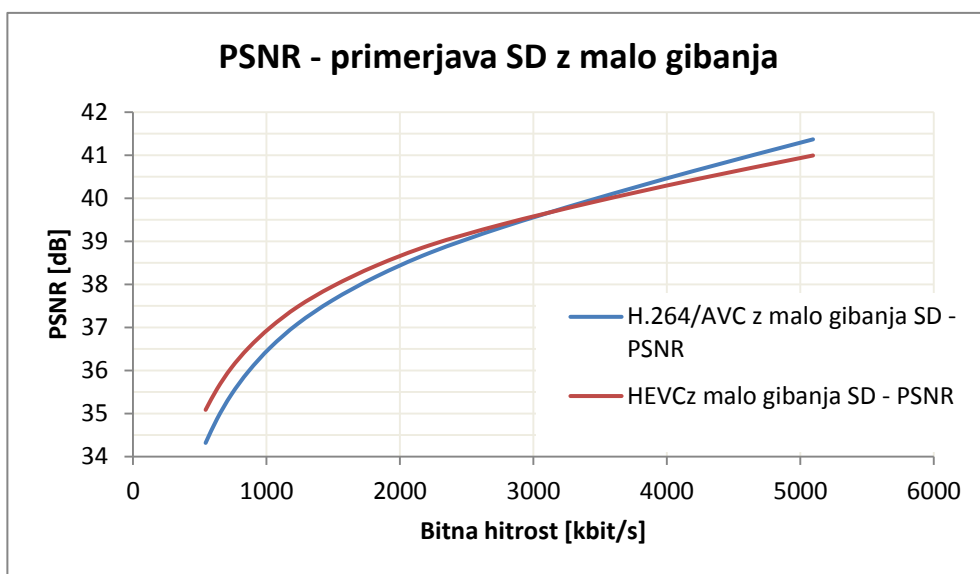


Slika 17: Primerjava med HEVC in H.264/MPEG4-AVC pri videu Full HD ločljivosti in majhni količini gibanja

Na videih SD kvalitete pa razlike niso tako očitne. Pri večji količini gibanja prihaja še vedno do manjših razlik, pri manjših količinah pa razlike niso več vidne. Razlog je v tem, da pri videih z nižjo ločljivostjo in z manjšo količino gibanja sam kodirni postopek ne more doprinesiti veliko izboljšanj. Za učinkovito kodiranje je namreč potrebna večja količina gibanja ali pa večja ločljivost, kjer je več redundantnih podatkov, ki jih lahko v kodirnem postopku odstranimo.



Slika 18: Primerjava med HEVC in H.264/MPEG4-AVC pri videu SD ločljivosti in veliki količini gibanja



Slika 19: Primerjava med HEVC in H.264/MPEG4-AVC pri videu SD ločljivosti in majhni količini gibanja

Tudi pri SSIM-ju vidno koleno pri padanju kvalitete, kar kaže na iste predpostavke, ki smo jih zapisali že pri Full HD videih.

Rezultati so pokazali, da HEVC dosega svoja pričakovanja pri videih visoke ločljivosti, pri videih standardne ločljivosti razlike niso opazne, za kar smo razloge navedli že zgoraj.

5 Zaključek

Zasnovali in razvili smo avtomatski sistem za testiranje videov z merama PSNR in SSIM. Sistem omogoča avtomatsko kodiranje, dekodiranje in računanje danih mer.

Za samo testiranje smo uporabili 8 videov, 4 SD ločljivosti in 4 Full HD ločljivosti. Razdelili smo jih v štiri podmnožice. Vsaka podmnožica je bila sestavljena iz videov iste ločljivosti in podobne količine gibanja.

Vsako podmnožico smo nato zakodirali s H.264/MPEG4-AVC in HEVC v različnih kvalitetah. Nato smo vsak video dekodirali v surovi video format in ga primerjali z originalnim. Dobili smo povprečno vrednost PSNR in SSIM za vsak video.

Dobljeni rezultati kažejo na veliko izboljšanje HEVC-a v primerjavi s H.264/MPEG4-AVC. Največje izboljšave so vidne pri videih Full HD vsebine, kjer pri isti kvaliteti videa prihranimo polovico pasovne širine. V nadaljevanju predlagamo, da se v test vnese primer izgubnega kanala, ki bo povzročal izgubo paketov. Tako bi preverjali obnašanje obeh kodekov v primeru izgube podatkov. Ugotovili smo, da ima odprtokodna rešitev FFmpeg ima hrošč v implementaciji vsebnika MPEG-TS. Ta v primeru izgub na prenosni poti ne zna pravilno dekodirati bitnega toka videa. Prihaja do izgub video okvirjev, kar pomeni, da so videi krajši za dolžino izgubljenih oz. zavrženih video okvirjev. Dekodirnik bi v primeru izgub moral kompenzirati to tako, da bi vnesel v video dvojnik nekega video okvirja ali na drugačen način zapolnil praznino v videu.

Druga možnost je, da bi v testiranje vključili vidoe UHD ločljivosti, ki bi morali pokazati še večjo razliko v učinkovitosti kodiranja. Videi takih ločljivosti zahtevajo zmogljivejši računalnik za kodiranje ter hkrati potrebujejo veliko časa za kodiranje, ki pa ga mi nismo imeli na razpolago.

6 Viri in literatura

- [1] “Cisco Visual Networking Index: Forecast and Methodology, 2013–2018,” *Cisco*. [Elektronski]. Razpoložljiv: http://cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html. [Poskus dostopa: 02-Mar-2015].
- [2] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, “Comparison of the Coding Efficiency of Video Coding Standards #x2014;Including High Efficiency Video Coding (HEVC),” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.
- [3] I. E. Richardson, *The H.264 Advanced Video Compression Standard*, 2nd ed. Wiley Publishing, 2010.
- [4] T. Wiegand and B. Girod, *Multi-Frame Motion-Compensated Prediction for Video Transmission*. Springer Science & Business Media, 2001.
- [5] L. L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. John Wiley & Sons, 2007.
- [6] ITU, “ITU-T Recommendation H.120 : Codecs for videoconferencing using primary digital group transmission.” Mar-1993.
- [7] T. Wedi and H. G. Musmann, “Motion- and aliasing-compensated prediction for hybrid video coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 577–586, Jul. 2003.
- [8] J. P. Marco Jacobs, “A Brief History of Video Coding.” ARC International, 2007.
- [9] S. Okubo, “Reference model methodology-A tool for the collaborative creation of video coding standards,” *Proceedings of the IEEE*, vol. 83, no. 2, pp. 139–150, Feb. 1995.
- [10] “MPEG-2 Video.” [Elektronski]. Razpoložljiv: <http://mpeg.chiariglione.org/standards/mpeg-2/video>. [Poskus dostopa: 23-Mar-2015].
- [11] “MPEG-4 Visual.” [Elektronski]. Razpoložljiv: <http://mpeg.chiariglione.org/standards/mpeg-4/video>. [Poskus dostopa: 23-Mar-2015].
- [12] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, “Video coding with H.264/AVC: tools, performance, and complexity,” *IEEE Circuits and Systems Magazine*, vol. 4, no. 1, pp. 7–28, First 2004.

- [13] "ITU-T Recommendation H.265: High efficiency video coding." Oct-2014.
- [14] M. T. Pourazad, C. Doutre, M. Azimi, and P. Nasiopoulos, "HEVC: The New Gold Standard for Video Compression: How Does HEVC Compare with H.264/AVC?," *IEEE Consumer Electronics Magazine*, vol. 1, no. 3, pp. 36–46, Jul. 2012.
- [15] "Difference between HEVC (H.265) and H.264 (MPEG-4 AVC), future of HEVC | Media & Entertainment Industry Trends, Technology and Research."
- [16] "Video quality," *Wikipedia, the free encyclopedia*, 19-Feb-2015. [Elektronski]. Razpoložljiv: http://en.wikipedia.org/w/index.php?title=Video_quality&oldid=647920663. [Poskus dostopa: 21-Feb-2015].
- [17] "Mean opinion score," *Wikipedia, the free encyclopedia*, 28-Jan-2015. [Elektronski]. Razpoložljiv: http://en.wikipedia.org/w/index.php?title=Mean_opinion_score&oldid=644494161. [Poskus dostopa: 16-Mar-2015].
- [18] ITU, "ITU-T Recommendation P.910 : Subjective video quality assessment methods for multimedia applications." Apr-2008.
- [19] ITU, "ITU-T Recommendation BT.500 : Methodology for the subjective assessment of the quality of television pictures." Jan-2012.
- [20] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 165–182, Jun. 2011.
- [21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [22] T. Zhao, K. Zeng, A. Rehman, and Z. Wang, "On the use of SSIM in HEVC," in *2013 Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 1107–1111.
- [23] "FFmpeg," *Wikipedia, the free encyclopedia*, 16-Feb-2015. [Elektronski]. Razpoložljiv: <http://en.wikipedia.org/w/index.php?title=FFmpeg&oldid=647367678>. [Poskus dostopa: 16-Mar-2015].
- [24] Technische Universität München, Institute for Data Processing, "TUM Multi Format Test Set," 2011. [Elektronski]. Razpoložljiv: <http://www.ldv.ei.tum.de/videolab>. [Poskus dostopa: 27-Feb-2015].
- [25] Werner Robitza, "CRF Guide." [Elektronski]. Razpoložljiv: <http://slhck.info/articles/crf>. [Poskus dostopa: 23-Mar-2015].
- [26] "Subjective video quality," *Wikipedia, the free encyclopedia*, 21-Feb-2015. [Elektronski]. Razpoložljiv: http://en.wikipedia.org/w/index.php?title=Subjective_video_quality&oldid=648155259. [Poskus dostopa: 21-Feb-2015].

Izjava

Izjavljam, da sem diplomsko delo izdelal samostojno pod vodstvom mentorja prof. dr. Andreja Koširja. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.

Rudi Hrvatin