

EKSPERIMENTALNI VIDEO KODEK DIRAK

Boban Kasalović, ETF Beograd

Sadržaj —Dirak je eksperimentalni video kodek baziran na wavelet tehnologiji, što je i osnovna razlika u odnosu na tradicionalne kompresione sisteme kao što su MPEG-x. Ovaj papir daje opis ovog kodeka. Takođe on daje originalne rezultate istraživanja koja se odnosi na subjektivnu procenu kvaliteta kodeka Dirak, koje je autor izvršio.

Ključne reči — Dirak, kompresija, kvalitet, subjektivna, video.

I. Uvod

Početak ovog veka, naglom ekspanzijom digitalne tehnologije, postalo je jasno da se uslovi emitovanja menjaju. Mnoge nove tehnologije, kao što je HDTV, internet *streaming* i IP TV su došle na scenu. Osnovni MPEG-2 video kompresioni sistem, koji se koristi za emitovanje DVD kvaliteta, je označen kao početak ovog doba. Novi sistemi sa boljim kompresionim performansama su postali neophodni za korišćenje u limitiranim opsezima.

Povećanjem slanja video sadržaja preko interneta, dovelo je do novih zahteva za javne servise emitovanja kao što je BBC. Za razliku od tradicionalnog emitovanja, cena distribucije preko interneta se povećava sa brojem korisnika. Jedna od cena je širina opsega neophodna za transmisiju video sadržaja, i ona se može minimizovati korišćenjem efikasnog kodeka. Još jedna cena koja se mora platiti je prateća tehnologija. Ta cena je rentabilna čak i za tzv. otvorene (*open*) standarde kao što je MPEG. Dok je inicijalna cena zanemarljiva, ona postaje značajna i veliki otežavajući faktor kada se skalira na milione korisnika. Šta više javni servisi emitovanja imaju obavezu da sadržaj učine slobodno dostupnim, sa nezavisnom platformom, bez pratećih zabrana. Upravo iz ovog razloga je započet razvoj projekta novog video kodeka.

Januara 2003, BBC R&D je izbacio prototip algoritma za video kodiranje, baziran na *wavelet* –talasić (u daljem tekstu jednostavno wavelet) tehnologiji, koji je drugačiji od dosadašnjih u smislu zaštićenosti glavnih osobina i standardizovanih video kompresionih sistema [1]. Ovaj algoritam, čini se, pruža dvo-slojnu redukciju bitskog protoka preko MPEG-2 za video visoke definicije (1920x1080 piksela), kao svoju originalnu ciljnu aplikaciju. On se i dalje unapređuje radi optimizacije za veličine internet protoka i postaje potpuno konkurentan sa ostalim video kodecima.

Kodek je dobio ime po čuvenom Britanskom fizičaru *Paul Dirac*-u, i razvijen je kao istraživački projekat, ne kao proizvod, nego baza za dalji razvoj. Izvasna eksperimentalna

verzija ovog koda, napisana u C++, je izdata pod *open source* licencnim dogovorom 11.marta 2004.

Osnovni princip Diraka je sačuvati jednostavnost (*keep it simple*), što je jedan od najambicioznijih slogana u svetu video kodovanja. Ipak BBC želi da sarađuje sa *open source* zajednicom, akademcima i drugima kako bi proizveli slobodan kodek.

Mnogi delovi kodeka se moraju konvertovati u povoljan algoritam, a eksperimentalne implementacije u praktično upotrebljiv kod. Ovo uključuje i optimizaciju kako bi dekodovanje bilo u realnom vremenu. Unapređenje algoritma je neophodno kako bi same kompresione performanse bile još bolje.

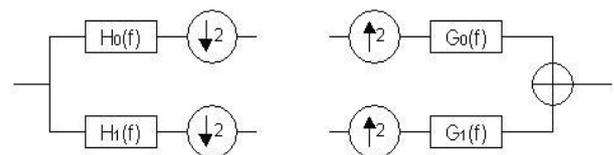
U drugom poglavlju rada data su neka od suštinskih unapređenja kod Diraka. Treće poglavlje je rezervisano za opis smernica potencijalnih unapređenja, čije je shvatanje od suštinske važnosti za izgled konačne verzija kodeka. Četvrto poglavlje je praktični doprinos ovog rada razvoju kodeka Dirak, jer su u njemu prikazani originalni rezultati procene trenutnog kodeka Dirak.

II. Unapređenja Diraka

A. Wavelet transformacija

Za razliku od DCT transformacije, koja je korišćena u većini dosadašnjih kodeka, ovde je upotrebljena wavelet transformacija i pokazala se kao pun pogodak. Kompresioni metodi koji se baziraju na matematičkoj tehnici poznatoj kao wavelet, koja je široko priznata, imaju konačni rezultat koji je znatno superijorniji od tradicionalnih blok-baziranih kompresionih šema kao što je JPEG i td [3].

Diskretna wavelet transformacija je sada veoma poznata i opisana u mnogim referencama, pa koristeći sve njene prednosti postaje sve više primenjivana u praktičnim realizacijama. U Diraku ona obavlja istu ulogu kao i DCT u MPEG-2. U jednoj dimenziji ona se sastoji od iterativne aplikacija komplementarnog para filtera vršeći pododabiranje sa faktorom 2.



Slika 1. Filterski parovi za analizu i sintezu sa perfektnom rekonstrukcijom [1]

Ovi filteri su nazvani analizatorski filteri. Odgovarajući filteri za sintezu poništavaju stepenasti efekat kritičnim

odabiranjem i tako dobijamo perfektnu rekonstrukciju. Filteri dele signal u LH i HF delove zatim wavelet transformacija vrši iterativnu dekompoziciju LF komponenata. Primenjujući je na dvodimenzionalnu sliku, wavelet filtri se normalno primenjuju na obe, i vertikalne i horizontalne, direkcije svake komponente. Tako dobijamo podopsege LL(Low-Low), LH(Low-High), HL(High-Low), HH(High-High) [4].

B. Estimacija pokreta

Estimacija pokreta je jedna od glavnih specifikacija kodera. To je uvek najkomplikovaniji deo celog sistema, i može da absorbuje sve sistemске resurse, tako da metodi koji se traže treba da pruže što jednostavnija rešenja. Dirak usvaja pristup u tri nivoa. U prvom nivou, vektori pokreta se traže za svaki blok i svaku referencu na nivou tačnosti od jednog piksela, koristeći hijerarhijsku estimaciju pokreta. U drugoj etapi vektori se predefinišu do tačnosti manjoj od piksela. U trećoj etapi se vrši izbor predikcije i grupisanje vektora pokreta u združene blokove sličnih pokreta. Estimacija pokreta je najtačnija kada su sve tri komponente uključene, ali je to znatno zahtevnije prilikom računanja.

C. Kompenzacija globalnog pokreta

Korišćenjem estimacije i kompenzacije globalnog pokreta u video kompresiji dobijamo parametarski model za aproksimaciju širokog spektra kretanja između dva frejma. Značajno smanjivanje veličine bitskog protoka se može postići prenosom parametara ovakvog modela umesto zasebnih blok vektora pokreta za svaki frejm.

Kompenzacija za pokrete između frejmova u video sekvenci veoma smanjuje veličinu podataka nove slike, koji su potrebni za osvežavanje svakog frejma. Ovaj proces često unosi tehniku povezivanja blokova. Analiza globalnog pokreta se sastoji od modelovanja pokreta koji se odnosi na ceo frejm. Takav globalni pokret se može ili raspoređivati za pojednostavljenje estimacije kretanja objekta ili koristiti direktno za smanjene veličine informacija korišćenih za opisivanje pokreta.

Pokreti malih regiona unutar frejma se često nazivaju pokreti objekta dok globalni pokret generalno opisuje pokrete kamere, mada može obuhvatiti pokrete velikog objekta. Pokreti same kamere su jedan od najsloženijih pokreta koji se mogu desiti. Mnogi načini pokreta moraju biti uključeni u pokušaju kvantifikovanja pokreta kamere: pan, paralelni švenk, zum i rotacija. Modelovanje pokreta kamere predstavlja opisivanje projekcije ovog pokreta na 2D ravni. Globalni pokret takođe može modelovati kretanje koje se ne odnosi na kameru ili objekat, već na pokret koji se odnosi na frejm npr. fejd ili odsecanje.

III. Smernice za unapređenja

Neka od potencijalnih unapređenja su već izbačena iz daljeg usavršavanja jer se ispostavilo da sa sobom donose ogromnu količinu podataka što, naravno, nije cilj razvoja Diraka. Zato i postoje smernice koje moraju održati osnovni moto-sačuvati jednostavnost.

A. Promenljivi broj nivoa wavelet dekompozicije

Trenutni kodek Dirak koristi tačno četiri nivoa podopsežne dekompozicije. Namera je da se Dirak može prilagoditi širokom opsegu aplikacija od povorka malih razolucija do HDTV. Optimalan broj nivoa wavelet

dekompozicije može varirati u raznim aplikacijama: više nivoa je moguće i poželjno za HDTV nego QCIF na primer. Predlaže se da može da se menja broj nivoa wavelet dekompozicije. Broj nivoa bi trebalo signalizirati u svakom frejmu. Korišćenje promenljivih nivoa wavelet dekompozicije bi trebalo poboljšati fleksibilnost, skalabilnost i kompresionu efikasnost za neke video formate. Postoji unutrašnja zavisnost između veličine blokova za kompenzaciju pokreta, rezolucije boja i broja korišćenih nivoa. Dodatni podaci uzrokovani promenljivim brojem nivoa wavelet dekompozicije je neznatan.

B. Promenljiva veličina blokova

Za sada je veličina pokretnih blokova i preklapanja postavljena jednom za celu sekvencu. Broj unapred definisanih blok/preklapanja okvira je definisan za različite video standarde: standardne definicije, progresivna 720-linijska HDTV, striming rezolucije, 1080-linijski HDTV. Više fleksibilnosti i znatno efikasnijeg kodovanja bi se moglo postići dopustajući promenljivost veličine bloka/preklapanja na jednom frejmu uz osnovicu frejma. Dodatno je neophodan samo jedan bit po frejmu ukoliko bi se koristio za promenu signala. Dodatni podaci bi se morali slati jedino kada se veličina bloka/preklapanja menja. Važno je da menjanje veličine bloka, unutar frejma je dozvoljen u određenim granicama, makroblokovskom strukturom [5].

C. Promenljiva tačnost vektora pokreta

Sadašnje performanse estimacije kompenzacije pokreta je preciznost od 1/8 piksela. Preciznost vektora pokreta može značajno uticati na efikasnost kompresije. Ukoliko je ona mala, predikcija ostatka (greške) je velika, što zahteva više bita za kodovanje. Ukoliko je preciznost vektora suviše velika onda bi se koristilo mnogo bita za njegovo kodovanje. Optimalna preciznost vektora zavisi od sastava slike, njenog formata (npr. protočni video standard ili HDTV), količine pokreta i tipa estimacije pokreta. Na primer, filmski sadržaj skeniran na rezoluciji od 1920x1080 linija zahteva tačnost od samo 1/2 piksela.

Za ovo unapređenje i konceptijska i implementacijska kompleksnost ja mala. Dodatni podaci su zanemarljivi. U isto vreme ovo može značajno unaprediti kompresionu efikasnost za neke aplikacije i formate slika. Na ovaj način broj aplikacija koji se mogu obaviti pomoću Dirak kodeka može biti uvećan.

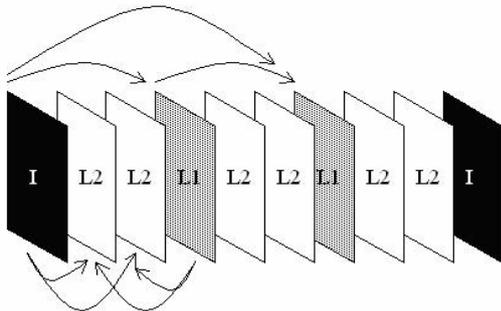
D. Opis globalnog pokreta

Postoje različiti pristupi za korišćenje kompenzacije globalnog pokreta u video kodovanju. Jedan, visokokompleksan pristup bio bi korišćenje parametara globalnog pokreta za prepravljanje referentnih frejmova. Racimo kad se kamera pomera pre uključivanja konvencionalne kompenzacije pokreta. Ovo je isuviše komplikovano, zato se odlučilo za korišćenje globalnog pokreta u prostijim slučajevima. Jedna alternativa za prepravljanje referentnih frejmova je da se pošalju parametri globalnog pokreta u dodatku vektora pokreta bloka. U ovom slučaju parametri se šalju opisujući globalne pokrete između referentnih i prediktovanih frejmova. Vektori pokreta bloka se šalju kao dodatna korekcija globalnom pokretu. Znači globalni pokret se može gledati kao globalna predikcija za vektore pokreta bloka. Dobitak ovakvog rada je da se pokreti jednostavnih scena (npr. kombinacija pan i zum) može

kodovati sa samo par bita. Parametri globalnog pokreta mogu se posmatrati kao jednostavan i efikasan način kodiranja pokretne scene korišćenjem parametarskog opisivanja. Sve što je potrebno je da se izračuna ofset globalnog pokreta između svakog bloka i pozicije bloka

Dirak koristi slike nivoa 0 (I frejmovi), nivoa 1 i nivoa 2. Grubo rečeno, to su ustveri I, P i B frejmovi korišćeni u MPEG kodovanju. Frejmovi nivoa 0 su nezavisno kodovane (intra) slike. Frejmovi nivoa 1 su prediktovani (iz 1 ili 2 referentna frejma) i oni sami se koriste kao referentni frejmovi. Frejmovi nivoa 2 su prediktovani, ali se ne koriste kao referentni frejmovi. Sada se frejmovi nivoa 2 uvek prediktuju iz dva referentna frejma ali bi bilo bolje kada bi se mogli prediktovati iz samo jednog frejma.

Tekuća implementacija kodeka Dirak uključuje frejm bafer koji prihvata frejmove i skladišti ih pre nego što se prikazuju ili koriste za referencu. Ovo je znatno fleksibilnije nego MPEG GOP struktura što potencijalno rezultira bolje kompresione performanse, na primer prilagodjevanje predikcione strukture. Ona takođe dopušta predikciju unazad ili dvosmernu predikciju frejmova nivoa 1 iz (budućih) I frejmova. Ona takođe dopušta korišćenje dugotrajnih referentnih slika koje mogu, na primer, sadržati pozadinu scene. Dugotrajne reference bi mogle biti korisne u progremima gde se učestalo prelazi između dva kadra, npr. intervjui. Namera je da se nastavi sa gene ralizovanjem ove predikcione strukture.



Slika 2. Raspored referentnih i prediktovanih frejmova [1]

Moglo bi se pokazati da dopuštanjem predikcije iz više od dva frejma bi bilo korisno, što je već iskorišćeno u H.264. Međutim, ovo znatno uvećava kompjuterski teret na koderu za estimaciju pokreta i način makroblokovske odluke. Odlučeno je da su dva frejma dovoljna.

IV. Subjektivna procena kvaliteta

Doprinos ovog rada razvoju konačne verzije kodeka, ogleda se u subjektivnoj proceni kvaliteta trenutne verzije kodeka. Izvršeno je subjektivno poređenje sa dva kodeka koja su u ranijim istraživanjima pokazala najbolje rezultate, a to su H.264 i DivX Pro [6].

Samo izvođenje testa nije bilo u optimalnim uslovima. Prvi razlog je bio taj što se sekvence kodovane Dirakom ne mogu pustati sa tradicionalnih plejera, već pomoću *Mplayera*. Ceo taj proces traje određeno vreme tako da ocenjivač nije mogao simultano da posmatra obe sekvence. Takođe sama činjenica da se samo ova sekvenca puštala iz drugog plejera na posredan način utiče na ocenjivača.

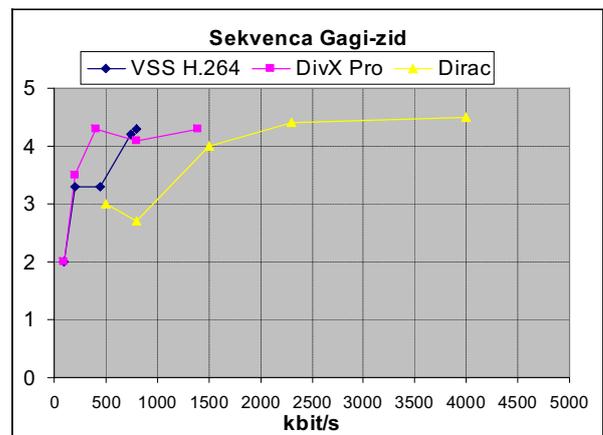
Tabela 1. Pregled sekvenci korišćenih za subjektivnu procenu kvaliteta

Naziv sekvence	Broj frejm.	Frejm. po sekundi	Veličina Fajla (KB)	Rezolucije i boje
Gagi-zid	300	25	91.141	360x288
Iz kola	500	25	151.895	360x288

Ipak pošto se radi o eksperimentalnoj verziji kodeka, koja je nesavršena u smislu korišćenja, nemoguće je stvoriti idealne uslove. Zato ove rezultate treba uzeti sa određenom rezervom. Ti rezultati mogu biti i te kako korisni za dalje usavršavanje kodeka. Izuzev ove činjenice, svi ostali uslovi u potpunosti ispunjavaju opšte priznate metodologije. Metodologija koja se koristi za subjektivnu procenu kvaliteta u ovom istraživanju zasnovana je na kombinaciji dve metode[7]: SAMVIQ metodologije i preporuke ITU-T P.910 uz određena prilagođenja eksperimentalnoj verziji[9].

A. Eksperimentalni rezultati

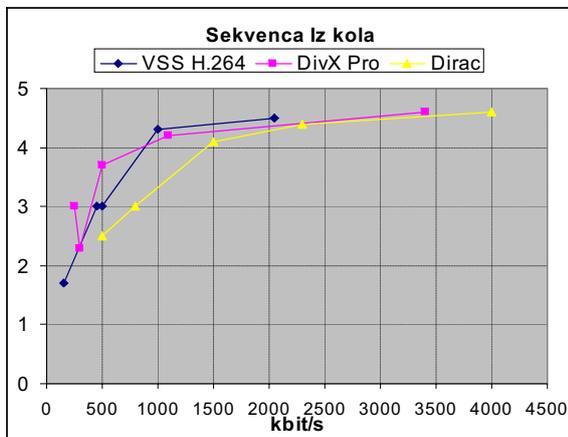
Trenutna verzija kodeka ne pruža mogućnost direktne promene veličine bitskog protoka, već se to radi izborom odgovarajućeg faktora kvaliteta. Zbog toga je vrlo teško rezultate prikazati tabelarno, jer dobijeni protoci nisu bili uvek isti. Dobijeni rezultati za izabrane sekvence su predstavljeni na sledećim grafikonima.



Slika 3. a) Sekvenca Gagi-zid i b) rezultati subjektivne procene kvaliteta za sekvencu iz kola

Rezultati dobijeni za H.264 i DivX Pro su veoma dobri, sa malim varijacijama koje su verovatno plod nesavršenosti

celog postupka. Rezultati nisu baš najbolji, moglo bi se reći, što se tiče Diraka. Na nižim protocima primetne su blokovske strukture zbog čega su ocene za Dirak nešto lošije. Na protoku od 1500 kbit/s ocene su gotovo identične sa H.264.



Slika 4.a) Sekvenca Iz kola i b) rezultati subjektivne procene kvaliteta za sekvencu Iz kola

U slučaju ove sekvence rezultati su ohrabrujući što se tiče Diraka. Na protoku iznad 1500 kbit/s rezultati su gotovo identični sa H.264 kodekom. Ovo potvrđuje njihovu uporedivost u performansama, koja je data u nekim tekstovima o Diraku. Pošto scena predstavlja zapravo globalni pokret, ovako relativno dobri rezultati eksperimentalne verzije su očekivajući, zbog brojnih unapređenja u kompenzaciji globalnog pokreta.

V. Zaključak

Dirak u sebi uključuje neka nova dostignuća i ideje, a neke od njih su još uvek u razmatranju. Eksperimentalna verzija Diraka sa sobom donosi i mnoge probleme pri korišćenju, ipak čini se da će krajnja verzija imati verovatno standardan interfejs, pa će i samo korišćenje biti veoma lakše. Ipak za one koji žele da u celu prišu uđu ranije onda ovaj rad može da pomogne u njihovim namerama. Rad obuhvata neke originalne rezultate dobijene subjektivnom procenom ovog kodeka. Može se reći da rezultati i nisu baš najpovoljniji, zbog čega mogu i te kako da posluže svima onima koji iskreno žele napredak ovog kodeka. Pošto je reč o eksperimentalnoj verziji Diraka, ovakvi rezultati mogu biti

posledica ne optimizacije, a tu pre svega se misli na još uvek ne definisanu sintaksu i drugo.

Praktičan doprinos rada leži u rezultatima subjektivne procene kvaliteta trenutnog Diraka. Rad pre svega predstavlja "ulaz" Diraka u Srbiju, da na posredan način može da utiče na sve zainteresovane, kako u pristupanju projektu tako i u svojim sopstvenim dostignućima u složenom svetu video kompresija.

Dalje istraživanje u oblasti kodeka Dirak leži i u njegovoj objektivnoj proceni, koja će biti objavljena po njihovom završetku na sledećim forumima.

Literatura

- [1] <http://dirac.sourceforge.net/>
- [2] Kasalović, B, *Eksperimentalni video kodek Dirak*, diplomski rad, ETF Beograd 2006.
- [3] Mojsilović, A, *Diskretna wavelet transformacija u analizi tekstone*, doktorski rad, ETF Beograd 1997.
- [4] R. Polikar, „Wavelet tutorial“, open tutorial <http://users.rowan.edu/wavelet>
- [5] www.bbs.co.uk
- [6] Ristić, D, Kasalović, B, *H.264/AVC među Srbima*, seminarski rad, ETF Beograd
- [7] Ristić, D, Josifović, I “Poređenje video kodnih postupaka: subjektivna procena kvaliteta”, *Zbornik Telfor 2005*, ETF Beograd
- [8] *ITU - R Recommendation BT. 500 - 9 (1998)*, “Methodology for the subjective assessment of the quality of television pictures”.
- [9] Josifović, I, Ristić, D, „Poređenje video kodnih postupaka: objektivna procena kvaliteta“, *Zbornik Telfor 2005*, ETF Beograd

Abstract: Dirac is experimental video codec based on wavelet technology; witch is different than traditional compression system such as MPEG-x. This paper presents an overview of this codec. Also it gives original results of research on subjective assessment of Dirac codec quality done by author.

**EXPERIMENTAL VIDEO CODEC DIRAC,
Boban Kasalović**