

Osvrt na predavanje – boja i zvuk u video kompresiji

U ovome predavanju obrađujemo temu kompresije podataka koja je važan čimbenik u prijenosu video i zvučnog signala. Prijenos podataka, tj. bitrate ne komprimiranog video signala je u rasponu od 270 Mb po sekundi za SDTV(standard definition TV), 1.5Gb po sekundi za HDTV te 3Gb po sekundi i više za UltraHD TV. Takva količina podataka je zapravo pre velika za efikasnu pohranu ili transmisiju pa svi podatci moraju biti komprimirani odnosno sažeti koristeći nešto što zovemo perceptualno kodiranje da bismo smanjili količinu podataka koje šaljemo. Ono uzima u obzir karakteristike ljudskih organa za primanje signala da bi se ustanovili limiti naše percepcije i srodno tome uklonili podatci koje mi ne možemo primijetiti. Te operacije se izvode prilikom transmisije videosignala ili prilikom samog snimanja. Svi oni imaju mali kapacitet pohrane i moraju izvoditi neku vrstu kompresije da bi mogli pohraniti podatke. Također, kod transmisije ili kako sada kažemo streamanja video i zvučnog signala moramo voditi računa o količini podataka koje šaljemo kako bi primatelj signala mogao neometano pratiti signal bez prevelikih gubitaka. Razlikujemo dvije vrste kodiranja. Kada govorimo o sažimanju podataka direktno prilikom snimanja govorimo o kodiranju izvora (source coding) jer sažimamo analogni signal koji dolazi na senzore uređaja za snimanje te kodiranje kasnije u post produkciji. Nakon kodiranja veličina se smanjuje na 1 do 15Mb ovisno o algoritmu kompresije (MPEG-1, MPEG-2, ...). Proces funkcionira na 2 načina. Postoji reduciranje suvišnih i ne važnih podataka. Suvišni su oni koji se ponavljaju više puta u podatkovnom prijenosu, a ne važni odnosno irelevantni podatci su oni koje ljudsko oko neće primijetiti da ne dostaju (mogu smanjiti veličinu i do 100 puta). Prilikom kompresije suvišnih podataka vrši se zamjena kompliciranog koda sa nekim jednostavnijim i manje opterećujućim kodom i takva vrsta kompresije se naziva kompresije bez gubitaka (lossless), a s druge strane prilikom reduciranja ne važnih podataka sami brišemo elemente čiji gubitak ljudsko oko, radi svoje anatomije, neće primijetiti. Takav način kompresije se zove lossy odnosno kompresija s gubiticima.

REDUKCIJA PODATAKA U BOJI

Video kamera nam daje izlazne signale u RGB color sustavu. Ti signali se matematički pretvaraju u luminantne i krominante komponente koje se odnose na svjetlinu (Y) i ton (CB i CR). Takav sustav boja označavamo kao YCbCr ili YUV. Konverzija se radi matematičkim formulama gdje R sudjeluje sa 30%, G sa 59% i B sa 11%. Ovi postotci se razlikuju radi osjetljivosti ljudskog oka na pojedine boje. Kada smo pretvorili boje u luminantne i krominante komponente jedne možemo reducirati ovisno o kvaliteti slike koju želimo postići. U video rječniku često možemo čuti omjer 3 broja (npr. 4:2:2 ...) koji predstavljaju odnos frekvencija urozkovanja za luminantnu i 2 krominantne boje komponente video signala. Taj proces zovemo uzorkovanje boje ili chroma subsampling. Omjer 4:4:4 je originalna rezolucija slike nagon prebacivanja iz RBG u YCbCr sustav. Omjer 4:2:2 naziva se horizontalno sempliranje, a za veću redukciju boje možemo uzeti i omjer 4:2:0. Ovakve promjene u video teže ćemo primijetiti nego promjene na statičnoj slici upravo zbog tromosti ljudskog oka i veće fokusiranosti na pokreti nego na detalje slike.

KOMPRESIJA ZVUKA

Slični principi koje smo opisali za sliku koriste se i za redukciju zvuka. Ljudsko uho ima dinamički raspon od otprilike 140 dB i sa rasponom frekvencije od 20 do 20 000Hz. Valovi koji nastaju titranjem od izvora veći od 20 000Hz se nazivaju ultrazvukom, a frekvencije manje od 20Hz nazivamo infrazvukom.

Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike raspona čujnosti ljudskog uha. Dio uha koji je zaslužan za prevođenje zvučne energije ono što mi percipiramo kao zvuk je pužnica. U njoj postoje osjetne stanice za visoke, srednje i niske frekvencije gdje auditorni živci prevode taj signal u električni impuls. Osjetljivost uha je najovisnija o frekvenciji zvuka, pa se tako najveća osjetljivost uha nalazi u rasponu od 3 do 4 KHz, a van toga osjetljivost na više i niže frekvencije pada.

Analogni signal se snima različitim uređajima te se preko filtera limitira njegov raspon prije nego što se on digitalizira. Taj proces se zove sempliranje ili uzorkovanje signala. Analogni zvuk koji je kontinuiran pretvaramo u uzorke koji su mjerljivi te na kraju dobivamo konačan broj uzoraka koji nazivamo diskretan broj uzoraka. Sampling rate ili brzina uzorkovanja se mjeri kao sample per second, tj. Broj snimljenih audio uzoraka u jednoj sekundi. Za audio signal najmanji broj uzoraka po sekundi je 8000 Hz. Ako koristimo manju frekvenciju zvuk će se činiti isprekidanim. Sampling rate je mjera na x-osi zvučnog vala koja nam kaže kolika je finoća uzoraka u jedinici vremena.

Druga mjera koju koristimo kod zvučnog signala je sample depth odnosno preciznost uzorka. Ona se mjeri u bitovima po uzorku, a jednako je važna pri kvaliteti zvuka kao i sample rate. Ta mjera nam određuje koliko mogućih razina amplitude može postići zvučni signal. Finoća podjele jačine zvuka na y-osi, za razliku od sample rate-a koji je finoća na x-osi. Ako govorimo o 8-bitnom zvuku to znači da možemo postići 256 različitih amplituda tj. jačina zvuka, a uzorak od 16-bita može postići 2^{16} jačina zvuka odnosno 65 500 različitih amplituda zvuka. U većini slučajeva se koristi 16-bitni zvuk dok se za kvalitetniji prijenos koristi 24-bitni ili 32-bitni zvuk.

Treća veličina koju ćemo spomenuti je bit rate koji imamo i kod video signala. No bit rate kod audio signala je produkt preciznosti uzorka (sample depth) pomnožen sa brzinom uzorkovanja (sample rate). Kada pomnožimo uzorak po sekundi i bit po uzorku dobijemo mjernu jedinicu bit po sekundi. Komprimirani audio signal uvijek ima bit rate koji je manji od umnoška sample ratea i sample deptha tj. frekvencije amplitude koje ljudsko uho ne može čuti se režu i komprimiraju.

Zvuk se kao i video treba komprimirati jer je njegov ne komprimirani kapacitet vrlo velik. Kod zvuka se služimo tehnikama kompresije suvišnih i ne važnih podataka kao i kod kompresije video materijala. Sve metode redukcije koriste ne savršenosti ljudskog uha kako bi maknule nepotrebne signale iz prijenosa. Glasniji zvukovi tj. zvukovi većih amplituda će tako maskirati zvukove manjih amplituda koji se nalaze odmah pored njih te ih uho neće čuti niti će primijetiti razliku. Najčešće korišteni sampling rate iznosi su 8kHz u telefoniji, 44.1/48 na TV/CD uređajima, 96/192kHz za blu-ray, a može se ići i više 300kHz za neke specijalne svrhe.