

Boja i zvuk u video kompresiji

Bit rate nekomprimiranog video signala je u rasponu od 270 Mb u sekundi za SDTV, 1,5 Gb u sekundi za HDTV i 3 ili više Gb u sekundi za UHD TV. Pošto je ta količina podataka prevelika za transmisiju, svi moraju biti komprimirani/sažeti.

Za komprimiranje se koristi perceptualno kodiranje, koje u obzir uzima karakteristike ljudskih očiju i ušiju kako bi se postavili limiti ljudske percepcije, te time reducirali podaci koje ljudi nebi mogli primjetiti. Razlikujemo 2 vrste kodiranja na izvoru i na post procesiranju. Komprimiranjem podataka na post procesiranju se prilagođavaju za različite medije. Kodiranje na izvoru svodi veličinu podataka na 1 do 15 Mb ovisno o algoritmu kompresije i željenoj rezoluciji zvuka. Taj proces funkcioniše na 2 načina, reduciranje suvišnih i nevažnih podataka. Time se količina podataka može smanjiti i na preko 100x.

Kod suvišnih podataka dijelovi koda se zamjenjuju kačim kodovima koji ih matematički opisuju, a takva kompresija se naziva kompresija bez gubitaka ili lossless kompresija.

Kod nevažnih podataka se za smanjenje podataka dijelovi koda brišu. Anatomija našeg oka ima puno manje receptora za boju (čunjići), nego receptora za svjetlinu (štapići). Čunjići imaju osjetljivost na crveni, plavi i zeleni dio spektra. Naše oko je puno osjetljivije na zelenu boju u odnosu na ostale. Pošto mi bolje vidimo promjenu svijetlosti, preciznost boje u signalu može biti reducirana. Naše oko također ne primjećuje fine strukture u slici kao što su tanke linije, male promjene kontrasta, itd. Metode redukcije podataka za video rade na tom principu, gdje se grublje strukture prenose sa većom preciznošću, nego fine strukture. Ti podaci su nevažni i izostavljaju se u procesu kodiranja i nemoguće ih je matematički vratiti. Takva kompresija naziva se lossy kompresija.

Video kamera daje izlazne signale u RGB sustavu, koji se matematički pretvaraju u luminantne i krominantne signale, tj komponente koje se odnose na svjetlinu (luminance, označavamo s Y) i one koje se odnose na ton (krominantne, označavamo s Cb i Cr). Takav sustav naziva se YcbCr ili YUV.

Konverzija se dobiva jednostavnim matematičkim operacijama, gdje se Y dobiva zbrajanjem signala R (30%), G (59%) i B (11%), Cb se dobiva operacijom $56\% \times B - Y$, a Cr operacijom $71\% \times R - Y$. Nakon razlučenja boje na te komponente, jedne se mogu reducirati ovisno o kvaliteti slike koju želimo postići.

U video riječniku često se pojavljuju omjeri 3 broja, koji predstavljaju odnos frekvencija uzorkovanja za luminantrnu i dvije krominantne komponente video signala. Taj se proces naziva uzorkovanje boje. Omjer 4:4:4 je originalna rezolucija slike nakon preračunavanja iz RGB u YCbCr sustav. Za sva 4 uzorka Y imamo 4 uzorka Cb i 4 Cr. Broj 4 se odnosi na veličinu uzorka, tj. 4 piksela. Druga dva broja se odnose na kroma komponente, te oni definiraju horizontalno i vertikalno uzorkovanje. Kod reduciranja podataka luminancija ostaje ista, dok se krominacija mijenja ovisno o omjeru. Omjer 4:2:0 je 4 puta manji od omjera 4:4:4.

Slični principi funkcioniraju i na zvuku. Naše uho ima dinamički raspon od 140 decibela i raspon frekvencije od 20 Hz do 20000 Hz. Frekvencije manje od 20 Hz nazivaju se infazvukom, a frekvencije veće od 20000 Hz nazivaju se ultrazvukom.

Audio signal mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha. Pužnica je dio uha koji je zaslužan za prevođenje zvučne energije u ono što mi percipiramo kao zvuk. Ona ima osjetne stanice za nisku, srednju i visoku frekvenciju, koje pretvaraju zvučne signale u električne impulse.

Uzorkovanje signala je proces limitiranja raspona signala putem filtera prije nego što se digitalizira. Kontinuirani analogni zvuk pretvaramo u mjerljive uzorke, gdje će se vidjeti konačan broj uzoraka zvan indiskretnim brojem uzoraka. Brzina uzorkovanja mjeri se kao sample per second, tj broj snimljenih audio uzoraka unutar jedne sekunde. Najmanji broj uzoraka po sekundi je 8000 Hz. Kod manje frekvencije zvuk će se činiti isprekidan. Bit rate se dobije množenjem uzorka po sekundi x bit po uzorku.

Zvuk se također komprimira kako bi se maknuli nepotrebni signali iz prijenosa. Glasniji zvukovi maskirati će tiše zvukove koje uho neće čuti. Redukcija nevažnih i suvišnih podataka može smanjiti prijenos zvučnog signala i do 90% od originala bez da se značajnije naruši kvaliteta. Mjera 44.1/48 kHz uzeta je kao standard za TV i CD zbog Nyquist-Shannonovog teorema koji kaže da se kontinuirani izvor signala može se uzorkovati i savršeno rekonstruirati najmanje dvostrukom frekvencijom od najveće moguće frekvencije u signalu.