

# Osvrt na predavanje: Boja i zvuk u video kompresiji

Lovro Vlašić, 22.5.2021.

Kompresija podataka je važan čimbenik u prijenosu video i zvučnog signala. Bit rate nekomprimiranog video materijala je u rasponu od 270 Mbps za SDTV, 1.5 Gbps za HDTV i 3 ili više Gbps za UHD TV.

Ta količina podataka je prevelika za efikasnu pohranu ili transmisiju, pa svi podaci moraju biti sažeti koristeći tzv. Perceptualno kodiranje da smanjimo količinu podataka koja se šalje.

Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa (oči i uši) za primanje signala, da bi se shodno tome reducirali podaci koje ne možemo primjetiti.

Kod transmisije također moramo voditi računa o količini podataka koje šaljemo kako bi primatelj mogao neometano pratiti signal bez prevelikih gubitaka.

Postoje dvije vrste kompresija: kompresija bez gubitaka (lossless) i kompresija s gubitcima (lossy).

Kompresija bez gubitaka – redundantni podaci se mogu lako izračunati matematičkim algoritmima, npr. Umjesto puštanja signala od deset nula, šalje se informacija 10 puta 0 ( $10^*$ ), kodom koji zauzima manje prostora od originala.

## Ljudsko oko i boja

Ljudsko oko ima puno manje receptora za boju nego što ima receptora za svjetlinu. Receptori za svjetlinu se nazivaju štapići, a receptori za boju su čunjići, oni se svi nalaze na mrežnici oka, iza leće. Štapići su mnogo osjetljiviji od čunjića, zato po noći slabije percipiramo boje i tonove, a lako percipiramo svjetlo.

Za percepciju tona boje su zaduženi čunjići, a njihova osjetljivost je na crveni, zeleni i plavi dio spektra (RGB). Osjetljivost na zelenu je puno veća nego na druge dvije boje. Druga je crvena po osjetljivosti, a zadnja je plava.

Lossy kompresija (s gubitcima) - nevažni, tj. Irelevantni podaci se izostavljaju tijekom procesa kodiranja, tj. Nemoguće ih je vratiti matematičnim operacijama.

## Redukcija podataka za boju

Video kamera nam daje izlazne signale u RGB color sustavu, a oni se pretvaraju u luminantne i krominantne signale, tj. Komponente koji se odnose na svjetlinu (luminancija,  $Y$ ) i ton (crominant,  $C_b$  i  $C_r$ ).

$$Y \text{ signal} = (0.3 * R) + (0.59 * G) + (0.11 * B)$$

$$C_b = 0.56 * (B - Y)$$

$$C_r = 0.71 * (R - Y)$$

Kada smo razlučili boju na luminantne i krominantne komponente, jedne možemo reducirati prema različitim odnosima ovisno o kvaliteti slike koju želimo postići. Video signal se razlučuje na luminantni dio, a drugi na krominantni dio.

U videu, omjeri 4:4:4, 4:2:2 te 4:2:0 predstavljaju odnos frekvencija uzorkovanja za luminantnu i dvije krominantne komponente video signala. Taj proces se zove uzorkovanje boje.

Omjer 4:4:4 je originalna rezolucija slike nakon preračunavanja iz RGB u YCbCr sustav. Za sva 4 uzorka  $Y$  komponente imamo 4 uzorka  $C_b$  i  $C_r$ . 4 znači veličina uzorka (tj. 4 piksela). Druga dva broja se odnose na kroma komponente, a oba su relativna prema prvom broju i oni definiraju horizontalno i vertikalno uzorkovanje.

Na primjer, kada sliku razlomimo na luminance i crominance, dobijemo jednu komponentu koja je crno bijela, a druga nam govori informacije o boji. Kada se one spoje dobijemo treću sliku, tj. Rezultat. Luma nam govori informacije o svjetlini, a chroma o boji. Prvi piksel luma je taman, a chroma govori da je to žuta boja. Kao rezultat toga, dobijemo tamno žutu boju.

Kod redukcije boje (tj. Chroma podataka) luminancija ostaje ista. Kod omjera 4:2:2, reduciramo svaki drugi horizontalni piksel. Kada usporedimo s originalnom slikom, oduzeli smo iz prve slike svaki drugi piksel. Kada takvu

chromu spojimo sa luma komponentom dobit ćemo 4 različite boje s 8 različitim svjetlina. Ovo se još naziva i horizontalno sempliranje.

Kod omjera 4:2:0, izbacit ćemo svaki drugi piksel sa vertikalne osi i reducirat ćemo sliku za 4 puta.

Dakle, video signal može podnijeti puno veću kompresiju slike zbog tromosti našeg oka. Ali ako gledamo statičnu sliku, možemo primijetiti razlike.

## Ljudsko uho i zvuk

Ljudsko uho ima dinamički raspon od otprilike 140 dB, sa rasponom frekvencija od 20 do 20 000 Hz. Valovi koji nastaju titranjem izvora većom frekvencijom 20 000 Hz se nazivaju ultrazvukom, a manje od 20 Hz se nazivaju infrazvukom.

Dio uha koji je zaslužen za prevođenje zvučne energije u ono što mi percipiramo kao zvuk je pužnica. U njoj postoje osjetne stanice za visoke, srednje i niske frekvencije. Osjetljivost uha je najovisnija o frekvenciji zvuka.

## Audio signal

Analogni signal se snima raznim uređajima te se putem filtera limitira njegov raspon prije nego se on digitalizira, to se naziva sempliranjem ili uzorkovanjem signala. Analogni zvuk koji je konvertiran pretvaramo u uzorke koji su mjerljivi, te dobivamo konačan broj uzoraka (diskretni broj uzoraka).

Sampling rate (sampling per second) = broj snimljenih audio uzoraka unutar jedne sekunde. 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 kod zvučnog signala zove se sampling depth (bit depth) – preciznost uzorka. Ona se mjeri u bitovima po uzorku. Ona određuje koliko mogućih razina amplitude može postići zvučni signal. To je finoća podjele jačine zvuka na y osi.

Ako imamo 8-bitni zvuk, to znači da možemo postići 256 različitih amplituda, a uzorak od 16 bita može postići  $2^{16}$  različitih amplituda.

Treća veličina je bit rate. Bit rate kod audio signala je produkt preciznosti uzorka (sampling depth) pomnožen sa brzinom uzorkovanja (sample rate).

Sve metode redukcije koriste nesavršenost ljudskog uha kako bi maknule nepotrebne signale iz prijenosa. Takva redukcija može smanjiti prijenos zvučnog signala do 90 % od originala, a da se značajnije ne ruši kvaliteta zvuka.

Primjeri nekih sampling rate iznosa su:

8 kHz – telefonija

44.1 / 48 kHz – TV/CD

96/192 kHz – blu-ray

>300 kHz

Nyquist-Shannonov teorem – kontinuirani izvor signala se može uzorkovati i savršeno rekonstruirati iz tih uzoraka najmanje dvostrukom frekvencijom od najveće moguće frekvencije u tom signalu.