

## Osvrt – Boja i zvuk u video kompresiji

Ovo predavanje nastavlja se na prethodnu lekciju te informacije koje smo spominjali uz kompresiju zato što je kompresija vrlo važan čimbenik u prijenosu video i zvučnog signala.

Prijenos podataka ne komprimiranog video signala je u rasponu od 270 Mb po sekundi za SDTV, 1.5 Gb za HDTV ili 3 ili više 1.3 za UHDTV. Takva količina podataka je prevelika za efikasnu pohranu ili transmisiju i zato koristimo perceptualno kodiranje (perceptual coding) kako bismo smanjili količinu podataka koja se šalje. Budući da perceptualno kodiranje uzima u obzir ljudske organe onda se pri kompresiji reduciraju suvišni i nevažni podaci koje mi ne možemo primijetiti. Te operacije se izvode prilikom transmisije samoga signala ili prilikom snimanja. Kada govorimo o sažimanju podataka, direktno prilikom snimanja, govorimo o kodiranju izvora. Sažimamo analogni signal koji dolazi na senzore uređaja za snimanje. I u tome je razlika od komprimiranja podataka u post procesiranju. Nakon kodiranja izvora količina podataka će biti svedena na 1-15 Mb ovisno o algoritmu kompresije kojeg koristimo. Jačina kompresije ovisi o video CODEC-u i željenoj rezoluciji.

Neke podatke (redundantni podaci) možemo lako i bez gubitaka izračunati nekim matematičkim algoritmima prilikom dekodiranja. Npr. umjesto puštanja signala od deset nula šalje se informacija 10 puta 0 posebnim kodom koji je mnogo kraći. Ukratko, dijelovi koda se zamjenjuju kraćim kodovima koji ih matematički opisuju. Takva kompresija naziva se losless kompresija, tj. kompresija bez gubitaka. U informacijskoj tehnologiji ovakav tip kodiranja se zove variable length coding. Nadalje, postoje i nevažni podaci, oni koje ljudsko oko ne zamjećuje. U slučaju video signala, to su komponente koje ljudsko oko ne registrira zbog tromosti. Takvi podaci se brišu, npr. kada se govori o boji onda su nevažni podaci o tonu boje. S druge strane, dobro percipiramo svjetlo i sjenu. To se pokazalo na testovima percepcije gdje se varirala veličina i svjetlina jedne točke naspram pozadine. Pokazano je da u nekom trenutku naše oko ne može razlučiti malu točku koja se u nijansi razlikuje po boji od pozadine. Na ovom principu djeluju metode redukcije podataka poput JPEG, MPEG za video gdje se grublje strukture, veći kontrasti prenose sa puno većom preciznošću nego finije strukture. Takvi podaci se zovu nevažni podaci i oni se izostavljaju u procesu kodiranja. Odnosno nemoguće ih je vratiti matematičkim operacijama kao što je to moguće kod redundantnih podataka. Takav način kompresije naziva se lossy, tj. kompresija sa gubitcima.

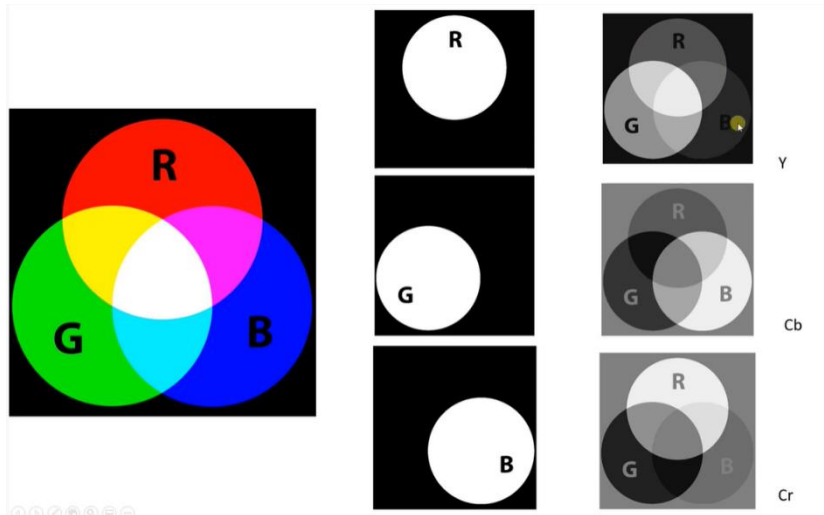
Video kamera nam daje izlazne signale u red, green i blue color sustavu. Ti signali se matematički pretvaraju u luminantne i krominantne signale, tj u komponente koje se odnose na svjetlinu (Y) i ton (Cb, Cr). Takav sustav boja označavamo kao YCbCr (YUV)sustav. Pri koverziji, Y signal se dobiva formulom:  **$Y=(0.3*R)+(0.59*G)+(0.11*B)$**

Svaka komponenta ima različit faktor množenja jer naše oko ima najveću osjetljivost na zelenu dio spektra, zatim crveni i na kraju plavi. Formule za druge komponente su :

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

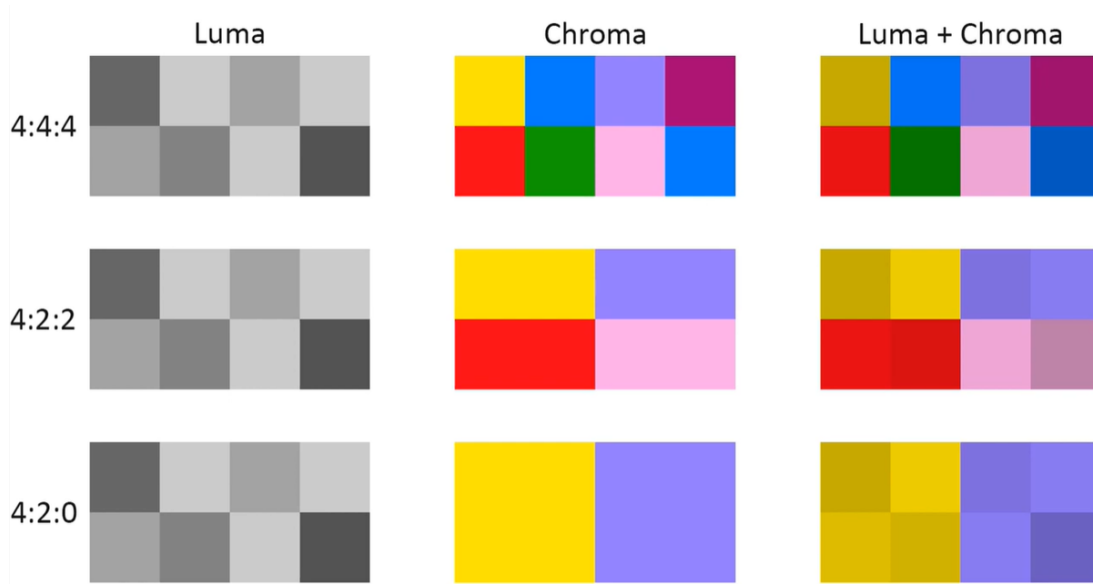
**$$Cr=0.71*(R-Y)$$**

Kada smo razlučili boju na luminantne i krominantne komponente jedne možemo reducirati prema različitim odnosima, ovisno o kvaliteti slike koju želimo postići. Zatim smo pogledali primjer s RGB kanalom. Vidjeli smo kako izgledaju odvojeno i kako izgledaju kada ih preračunamo u Y, Cb i Cr komponente. Tada su kanali izgledali drugačije.



Nadalje, taj način odvajanja pogledali smo na primjeru slike ženskog oka gdje je su bile dvije komponente- luminantna koja je bila crno-bijela i chroma u kojoj su bili podaci o boji i tonu.

Na sljedećoj slici prikazan je odnos frekvencija uzorkovanja za luminantnu i dvije krominantne komponente video signala.



Taj proces naziva se uzorkovanje boje, odnosno chroma subsampling. Omjer 4:4:4 je originalna rezolucija slike nakon preračunavanja iz RGB u YCbCr sustav. Za sva četiri uzorka Y komponente imamo četiri uzorka Cb i Cr. 4 se odnosi na veličinu uzorka ( 4 piksela). Druga dva broja se odnose na kroma komponente. Oba su relativna prema prvom broju i definiraju horizontalno i vertikalno uzorkovanje. U videu nećemo obraćati pažnju na detalje koji su na slici dok na statičnoj slici hoćemo. I to smo vidjeli na primjeru sa više piksela te na

konkretnom primjeru slike gdje se omjer 4:4:4 koristio kada se želimo koncentrirati na sitne detalje poput teksta. 4:2:0 je omjer koji je sasvim zadovoljavajuć za pokretnu sliku, video.

Nadalje, kada govorimo o digitalnom audio signalu bitno je definirati sljedeće pojmove. Prvi pojam je sampling rate (brzina uzorkovanja) je mjera koja nam kaže kolika je finoća uzoraka u jedinici vremena. Zatim smo usporedili rezultate analognog i digitalnog uzorkovanja. S većim sample rateom smo dobili rezultat sličniji analognom. Sample rate ne mora biti prevelik jer naše uho neće čuti tu kvalitetu. Druga mjera je sampling depth (preciznost uzorka) i određuje koliko mogućih razina amplitude može postići zvučni signal. Treći pojam je **bit rate=sample rate\*sample depth**. Zvuk se kao i slika komprimira jer je njegov kapacitet velik te se koristimo kompresijom nevažnih i suvišnih podataka. Neki zvukovi većih amplituda će zamaskirati zvukove manjih amplituda koje uho onda neće čuti te tako bez straha možemo ukloniti te zvukove manjih amplituda. Najčešće korišteni sampling rate su: 8 kHz-telefonija, 44.1/48 kHz-TV,CD, 96/192 kHz- blu-ray te >300kHz. Time se predavanje privelo kraju.