

OSVRT – BOJA I ZVUK U VIDEO KOMPRESIJI

UVOD

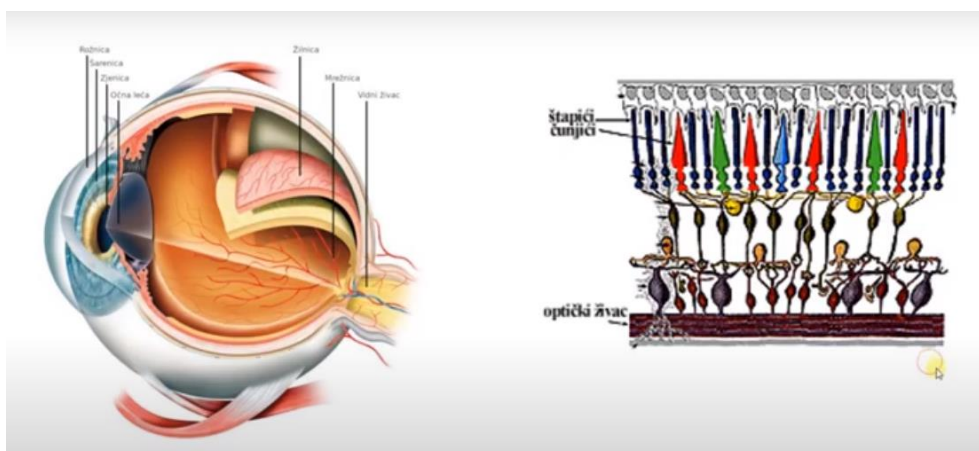
Bit rate (prijenos podataka) nekomprimiranog video signala je u rasponu od 270 Mbps (SDTV), 1.5 Gbit (HDTV) ili 3/> Gbit/s (Ultra-HD TV). Ova količina podataka je prevelika za učinkovitu pohranu ili transmisiju, stoga je moramo komprimirati (sažeti) pritom se koristeći **PERCEPTUALNIM KODIRANJEM**.

Prilikom sažimanja perceptualno kodiranje uzima u obzir karakteristike ljudskih organa za primanje signala kako bi se ustanovili limiti naše percepcije te shodno tome reducirali podaci koje mi ne možemo primijetiti. Ove se operacije izvode prilikom transmisije video signala ili prilikom samog snimanja videa. 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. Postoji i komprimiranje podataka u postprocesiranju kako bi se video prilagodio za različite medije.

Nakon kodiranja izvora količina podataka biti će svedena na 1 – 15 Mbita ovisno o algoritmu kodiranja (MPG-1, MPG-2, MPG-4, ...).

Proces kodiranja funkcionira na 2 načina: reduciranje suvišnih (redundantnih) i nevažnih (irelevantnih) podataka. Kod suvišnih podataka koristimo se **kompresijom bez gubitaka (lossless)**. U informacijskoj tehnologiji ovakvo se kodiranje naziva Huffmanovo kodiranje. S druge strane postoje nevažni (irelevantni) podaci, odnosno oni koje ljudsko oko ne može uopće percipirati. Kod video signala to su komponente koje ljudsko oko ne registrira zbog njegove anatomije. Dakle kod perceptivnog kodiranja uzimamo te podatke te ih brišemo kako bi se signal dodatno smanjio.

Kako to izgleda kod boje?



Naše oko ima puno manje receptora za informaciju o boji nego što ima o svjetlini. Receptori za crno-bijelo odnosno za svjetlinu se nazivaju **štapčići**, a receptori za boju su **čunjići**. Štapčići

imaju svoju primarnu ulogu u gledanju u uvjetima niskog svjetla (po noći), ima ih puno više od čunjića i mnogo su osjetljiviji. Zbog toga po noći slabije percipiramo boje (tonove), a dobro 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, potom je crvena i na kraju plava.

S obzirom da mi bolje vidimo promjenu svjetla-sjena, nego boje zbog toga količina boje u video signalu može biti reducirana, što posljedično znači i redukciju količine podataka o boji. Još jedna karakteristika oka je da *ne može vidjeti fine strukture u slici* (vrlo tanke linije, male promjene kontrasta..) te na ovom principu djeluju metode redukcije podataka kao što su **JPG** i **MPG** za video gdje se grublje strukture prenose sa puno većom preciznošću od finijih. Takvi se podatci nazivaju nevažni podaci i oni se **izostavljaju u procesu kodiranja**. Ovakav način kopresije se zove **lossy** tj **kompresija sa gubitcima**.

REDUKCIJA PODATAKA ZA BOJU

Video kamera nam daje izlazne signale u **RGB color sustavu**. Ti signali se matematički pretvaraju u **luminantne** (svjetlina; **Y**) i **krominantne** (ton; **Cb, Cr**) **signale**. Ovakav sustav boja označavamo kao **YCbCr (YUV)**.

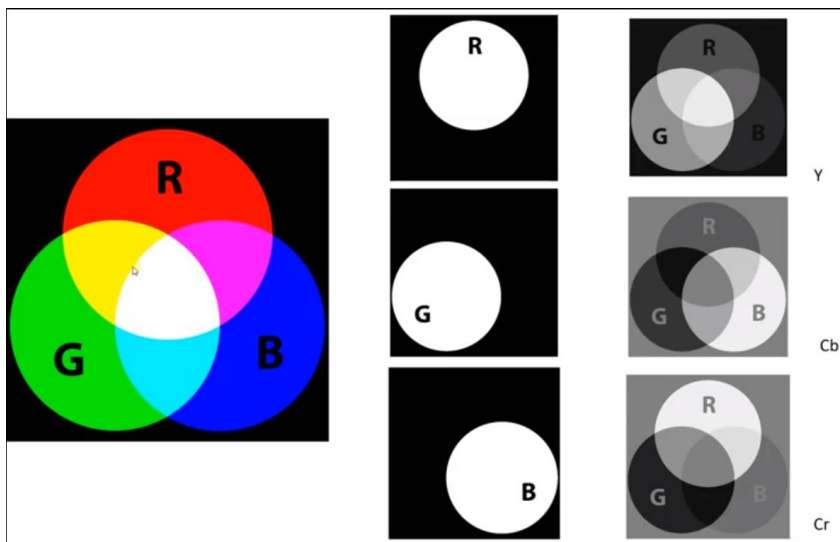
Konverzija se radi jednostavnim matematičkim operacijama:

$$Y = (0.3 * R) + (0.59 * G) + (0.11 * B)$$

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

$$R = 30\%, G = 59\%, B = 11\%$$

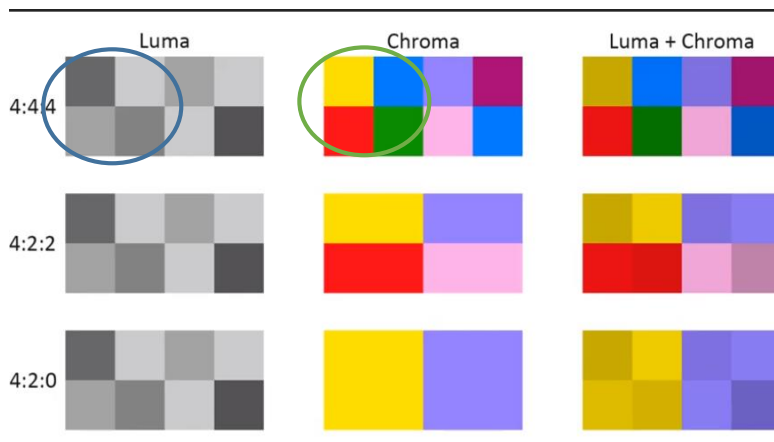
$$Cr = 0.71 * (R - Y)$$



Usporedba RGB kanala sa kanalima Y, Cb, Cr

➔ Bijeli krug predstavlja maksimalnu količinu neke boje dok crna predstavlja odsustvo te iste boje. Kod drugog reda slika Y daje podatke o svjetlini, a Cb i Cr o boji (tonovima)

U video rječniku često možemo čuti omjer tri broja (npr. 4:2:2), ti brojevi predstavljaju frekvenciju uzrokovanja za luminantnu i dvije krominantne komponente video signala. Taj proces zovemo i uzrokovanje boje odnosno **Chroma subsampling**.

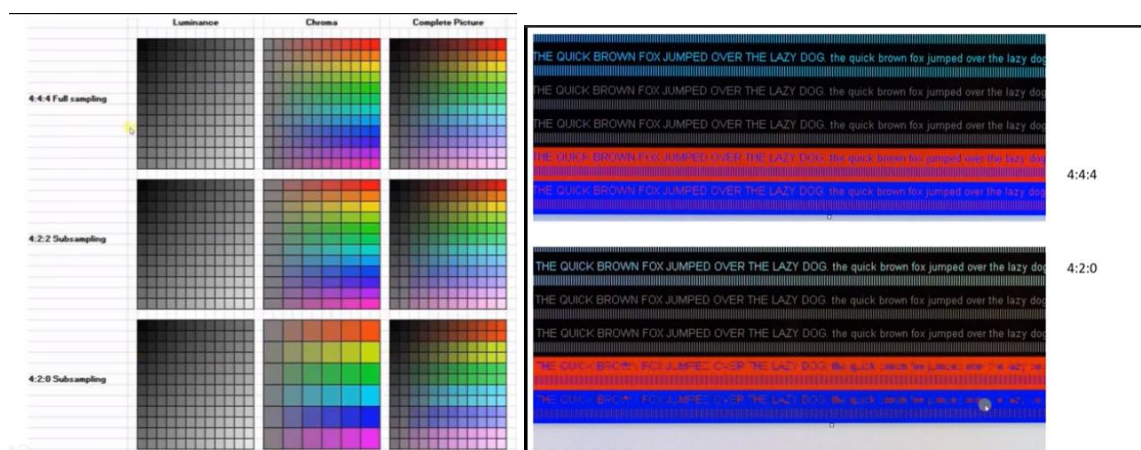


Za 4 uzorka Y komponente postoji i 4 uzorka Cb i Cr. 4 se u ovom slučaju odnosi na veličinu uzorka; obično su to 4 piksela. Druga 2 broja se odnose na chroma komponente; oni definiraju horizontalno i vertikalno uzorkovanje. Luma komponenta je crno-bijela i daje nam informaciju o svjetlini, a Chroma je ta koja nam daje informaciju o boji. Treća slika je zbroj prve dvije.

Pr. Ako pogledamo prvi piksel Luma kaže da je boja tamna, a Chroma da se radi o žutoj, dok u rezultatu vidimo da je to tamno žuta.

Kada govorimo o **reduciranju podataka** to možemo vidjeti u drugom redu slike. Kod redukcije boje, reduciraju se chroma podaci, a luminacija ostaje ista. Kod **omjera 4:2:2**, reduciramo svaki drugi horizontalni piksel. Kada ovakvu chromu spojimo sa nepromijenjenom luma komponentom dobit ćemo rezultat **4 različite boje sa 8 različitih svjetlina**. Ovakvo simpliranje još nazivamo **horizontalno simpliranje**.

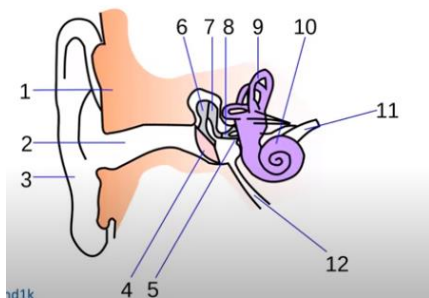
Kod omjera 4:2:0 izbacujemo i svaki 2. piksel sa vertikalne osi te na taj način reducirati sliku za 4 puta. Kao rezultat dobivamo **1 boju u uzorku 4 piksela sa 4 različite svjetline**.



→ *resempliranje statične slike*

DIGITALNI AUDIO SIGNAL

Slični principi koji su bili opisani za video/sliku funkcioniraju i kod zvuka.

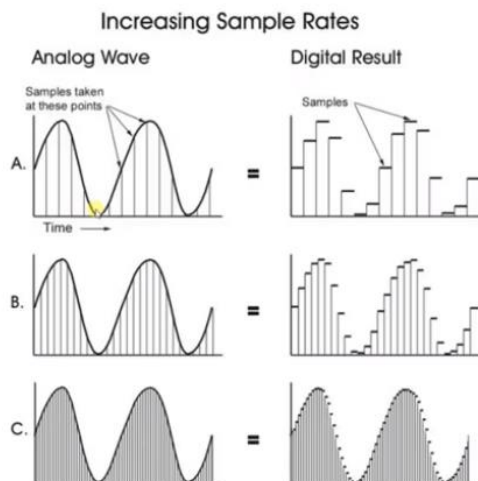


Ljudsko uho ima dinamički raspon od **14 db** sa rasponom frekvencija od **20 – 20 000 Hz**. Valovi zvuka koji nastaju pri frekvenciji većoj od 20 000 Hz se nazivaju **ultra zvukom**, a frekvencije manje od 20 Hz **infra zvukom**. Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha. Dio uha koji je zaslužen za prevođenje zvučne energije (zvuk) je **pužnica** (označena je ljubičastom bojom na slici). 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 te se tako **najveća osjetljivost nalazi u rasponu od 3-4 kHz**.

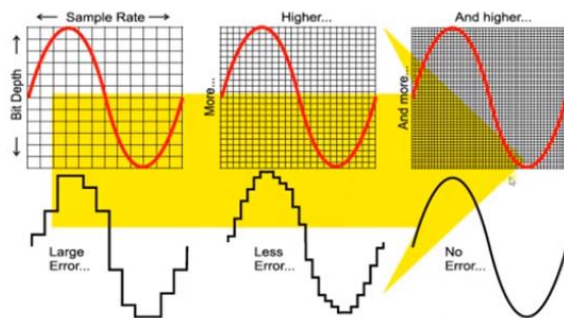
Analogni signal se snima različitim uređajima te se putem filtera limitira njegov raspon prije nego što se on digitalizira. Taj proces se naziva **sempliranje** ili **uzrokovanje signala**. Analogni zvuk koji je kontinuiran pretvaramo u mjerljive uzorke te na kraju dobivamo konačan broj uzoraka → **diskretni broj uzoraka**.

Sample rate (brzina uzokovanja) se mjeri kao broj snimljenih audio uzoraka unutar jedne sekunde ($f_s = 1/T$). Za audio signal najmanji broj uzoraka po sekundi iznosi **8 000 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.

➔ *Različiti sample rate u odnosu na analogni signal zvuka*



Druga mjera o kojoj govorimo kod zvučnog signala je **sample depth (preciznost uzorka)**. Ona se mjeri u bitovima po uzorku. Određuje koliko mogućih razina amplitude može postići zvučni signal. To je finoća podjele jačine zvuka na y osi.



Ako govorimo o 8-bitnom zvuku, to znači da možemo postići **256 različitih amplituda** (jačina zvuka), dok uzorak od 16 bita može postići **2¹⁶ različitih amplituda zvuka**. Za kvalitetan audio snimak se preporučuje **24 bitni zvuk**.

Treća veličina o kojoj pričamo je **bit rate**.

Bit rate = Sample rate * Sample depth

= uzorak / sek * bit / uzorku = bit / sek

Komprimirani audio signal uvijek ima bit rate koji je manji od umnoška sample rate i sample deptha tj frekvencije koje ljudsko oko ne može čuti se režu i komprimiraju.

Sve metode redukcije koriste nesavršenost ljudskog uha kako bi maknule nepotrebne signale iz prijenosa. Redukcija nevažnih i suvišnih podataka može smanjiti prijenos zvučnog signala čak i do 90% od originalnog, a da se značajnije ne naruši kvaliteta zvuka.

Najčešće korišteni sampling rate:

1. 8 kHz – telefonija
2. 44.1/48 kHz – TV/ CD
3. 96/192 kHz – blu-ray
4. > 300 kHz – specijalni zvukovi (ultra zvuk)

Kod CD kvalitete audio sampling rate je 44.1 kHz, a ljudsko uho čuje samo 20 kHz. Ta mjera je uzeta kao standard zbog teorema **Nyquist-Shannonov teorem**. Teorem kaže da se kontinuirani izvor signala može uzorakovati i savršeno rekonstruirati iz tih uzoraka, najmanje dvostrukom frekvencijom od najveće moguće frekvencije u signalu. Drugim riječima teorem tvrdi da bi se spriječio gubitak informacije, kada se signal digitalno uzorkuje (semplira), sample rate mora biti najmanje dvostruke veličine od najveće očekivane frekvencije signala. Naše uho ima najveći domet frekvencije od 20 kHz pa je 44 kHz uzeto kao prikladna frekvencija sempliranja.