

# Boja i zvuk u video kompresiji

Prijenos podataka odnosno Bitrate nekomprimiranog videosignala je u rasponu od 270 megabita po sekundi za SDTV (Standard Definition TV), 1.5 gigabita HDTV ili 3 ili više gigabita po sekundi za High Definition TV. Takva količina podataka je prevelika za efikasnu pohranu ili transmisiju pa svi podatci moraju biti komprimirani odnosno sažeti koristeći perceptualno kodiranje da bismo smanjili količinu podataka koja se šalje.

Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa za primanje signala, a to su naše oči i uši da bi se ustanovili limiti naše percepcije i shodno tome reducirali podaci koje mi ne možemo primijetiti. Te operacije se izvode prilikom transmisije videosignala ili prilikom samog snimanja kamerama, fotoaparatom, mobilnim uređajima itd. Svi navedeni uređaji imaju mali kapacitet pohrane i moraju izvoditi neku vrstu kompresije da bi mogli spremati podatke. Također, kod transmisije moramo voditi računa o količini podataka koju šaljemo kako bi primatelj signala mogao neometano pratiti signal bez prevelikih gubitaka. Kada govorimo o sažimanju podataka direktno prilikom snimanja onda govorimo o kodiranju izvora (Source Coding) jer sažimamo analogni signal koji dolazi na senzore uređaja za snimanje za razliku od komprimiranja podataka u postprocesiranju kako i se on prilagodio za različite medije. Razlikujemo dvije vrste kodiranja odmah na izvoru i poslije u postprocesiranju. Nakon kodiranja izvora količina podataka će biti svedena na 1 do 15 megabita ovisno o algoritmu kompresije. Jačina kompresije ovisi o video Codecu i željenoj rezoluciji slike tj. zvuka.

Taj proces funkcionira na dva načina:

- Reduciranje suvišnih podataka (redundantnih); ponavljaju se više puta
- Reduciranje nevažnih podataka (irelevantnih); podatci koje ljudsko oko neće primijetiti da nedostaju

Redundantni podaci ponavljaju se u podatkovnom prijenosu te se mogu lako i bez gubitaka izračunati nekim matematičkim algoritmima prilikom dekodiranja. Dijelovi koda se zamjenjuju kraćim kodovima koji ih matematički opisuju Huffmanovo kodiranje. Takva kompresija se naziva kompresija bez gubitaka (lossless kompresija). Nevažne ili irelevantne podatke ljudsko oko ne može uopće percipirati. U slučaju videosignala to su komponente koje ljudsko oko ne registrira upravo zbog njegove anatomije, a u perceptivnom kodiranju uzimamo te podatke i brišemo ih kako bi se signal dodatno smanjio. Kompresija irelevantnih podataka naziva se kompresija s gubitcima (lossy kompresija).

Ljudsko oko se evolucijski razvilo da može bolje razlučivati pokrete u tamnom okruženju nego razlikovati sitne nijanse boja. Za percepciju tona boja su zaduženi čunjići, a njihova osjetljivost je na crveni, zeleni i plavi dio spektra odnosno RGB. S obzirom da ljudsko oko bolje vidi odnos svjetlo-sjena nego preciznost nijanse boje u videosignalu. Preciznost boje u videosignalu može biti reducirana što posljedično znači redukciju količine podataka o boji. Također, ljudsko oko ne može vidjeti fine strukture na slici npr. vrlo tanke linije, male promjene kontrasta itd.

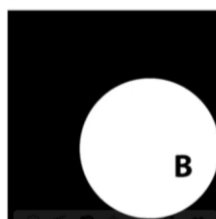
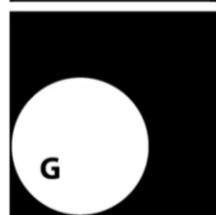
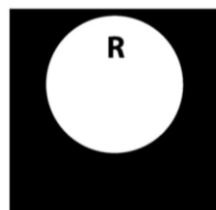
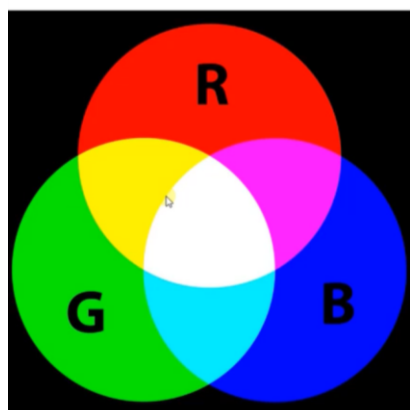
Videokamera nam daje izlazne signale u red, green i blue color sustavu odnosno RGB. Ti signali se matematički pretvaraju u luminantne i krominantne signale tj. u komponente koje se odnose na svjetlinu (luminance) i njih označavamo slovom Y. Podatci koji se odnose na ton (chrominance); krominantne komponente, a njih označavamo oznakama Cb i Cr; ponegdje se naziva YUV.

Konverzija se radi jednostavnim matematičkim operacijama:

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

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

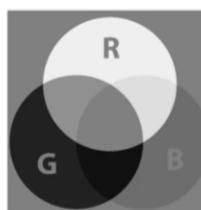
$$Cr = 0.71 * (R - Y)$$



Y



Cb



Cr



=



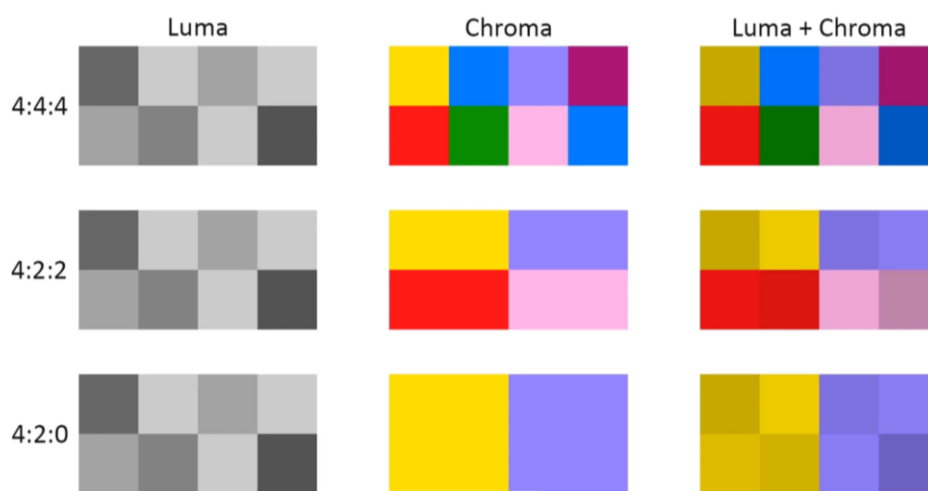
+



Video Signal

Luma

Chroma

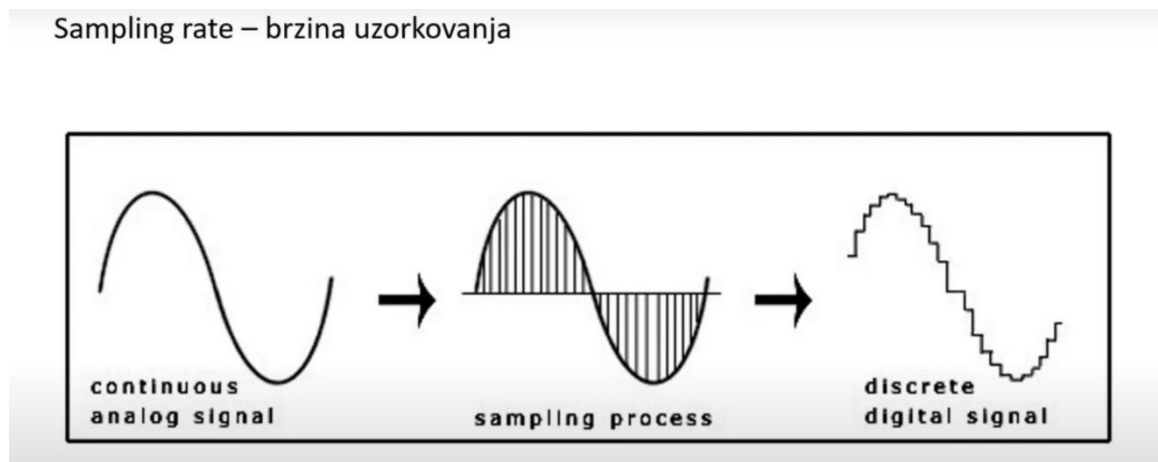


## Digitalni audio signal

Ljudsko uho ima dinamički raspon od otprilike 140 dB i raspon frekvencija od 20 Hz do 20 000 Hz. Valovi koji nastaju titranjem izvora frekvencijom većom od 20 000 Hz se nazivaju ultrazvukom, a frekvencije koje su manje od 20 Hz nazivamo infrazvukom. Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha. Dio uha koja je zaslužan za prevođenje zvučne energije u 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 3000 Hz do 4000 Hz.

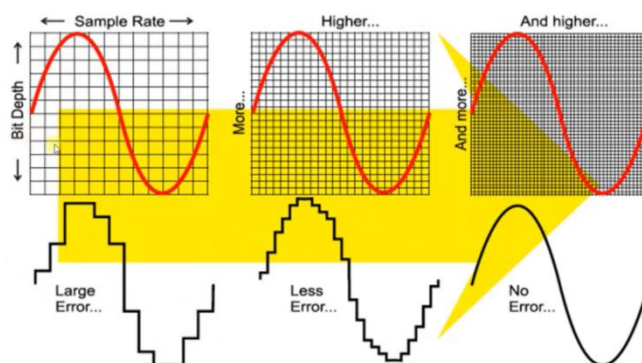
Analogni signal se snima različitim uređajima te se putem filtera limitira njegov raspon prije nego što se on digitalizira. Taj proces se zove sempliranje ili uzorkovanje signala. Analogni zvuk koji kontinuiran pretvaramo u uzorke koji su mjerljivi te na kraju dobivamo končan broj uzoraka koji još nazivamo i diskretnim brojem uzoraka. Sampling rate ili brzina uzorkovanje se mjeri kao sample per second odnosno broj snimljenih audio uzoraka unutar jedne sekunde. To je mjera slična brzini izmjena sličica u sekundi tj. frame per second. Za audio signal najmanji broj uzoraka je 8 000 Hz. Ako koristimo manju frekvenciju zvuk će se činiti isprekidanim. Sampling rate je mjera na x osi zvučnog vala koja nam kaže koliko je finoća uzoraka u jedinici vremena.

### Sampling rate – brzina uzorkovanja



Sampling depth ili preciznost uzorka se mjeri s bitovima po uzorku, a jednako je tako važna kada pričamo o kvaliteti zvuka kao i Sample rate. Sampling depth određuje koliko mogućih razina amplitude može postići zvučni signal. Sampling depth je mjera na y osi finoće podjele zvuka.

### Sampling depth – preciznost uzorka



Ako govorimo o 8 - bitnom zvuku to znači da možemo postići 256 različitih amplituda odnosno jačina zvuka, a uzorak od 16 bita može postići  $2^{16}$  različitih amplituda zvuka. U većini slučajeva se koristi 16-bitni zvuk dok se za kvalitetniji prijenos koristi 24-bitni ili 32-bitni zvuk.

Bit rate kod audiosignala je produkt preciznosti uzorka i brzinom uzorkovanja.

Bit rate = Sample rate \* Sample depth = uzorak/sek \* bit/uzorku = bit/sek

Najčešće korišteni sampling rate:

- 8 kHz – telefonija
- 44.1/48 kHz – TV/CD
- 96/192 kHz – blue-ray
- >300 kHz,

Nyquistov-Shannonov teorem tvrdi da bi se spriječio gubitak informacije kada se signal digitalno uzorkuje odnosno semplira Sample rate mora biti najmanje dvostruke veličine od najveće očekivanje frekvencije signala.

Natalija Talevski