

## OSVRT NA PREDAVANJE

### BOJA I ZVUK U VIDEO KOMPRESIJI

Kompresija podataka je važan čimbenik u prijenosu video i zvučnog signala. Prijenos podataka odnosno Bitrate nekog primiranog video signala je u rasponu od 270 mb/s za SD TV (standard definition TV) , 1.5 GB za harder tv ili 3 ili više GB po sekundi za ultra high definition TV. Takva količina podataka je prevelika za efikasnu pohranu ili transmisiju pa svi podatci moraju biti komprimirani odnosno sažeti koristeći perceptualno kodiranje (perceptual coding) da bismo smanjili količinu podataka koja se šalje

Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa za primanje signala a to su: oči i uši da bi s ustanovili limiti naše percepcije i shodno tome reducirali podatci koje mi ne možemo primjetiti. Te operacije se izvode prilikom transmisije video signala ili prilikom samog snimanja kamerama, fotoaparatima, mobilnim uređajima itd. Svi oni imaju mali kapacitet pohrane i moraju izvoditi neku vrstu kompresije da bi mogli spremiti podatke.

Također kod transmisije (streamanja) video ili zvučnog signala 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 govorimo o kodiranju izvora (source coding) jer sažimamo analogni signal koji dolazi na senzore uređaja za snimanje za razliku od komprimiranja podataka u post procesiranju kako bi se on prilagodio za različite medije

Tu razlikujemo dvije vrste kodiranja a to su odmah na izvoru i kasnije u post procesiranju

Nakon kodiranja izvora količina podataka će biti svedena na 1-15 MB zavisno o algoritmu kompresije koji koristimo (mpg1, mpg2, mpg4, h264, h264). jačina kompresije ovisi o video kodeku i željenoj rezoluciji slike odnosno zvuka. Taj proces funkcionira na dva načina. Postoji funkcioniranje na osnovu suvišnih (oni koji se ponavljaju više puta u podatkovnom prijenosu, i oni se mogu lako i bez gubitaka izračunati nekim matematičkim algoritmima preko kodiranja) i nevažni/relevantni (oni koje ljudsko oko neće primijetiti da nedostaju, i oni mogu smanjiti količinu podataka i preko sto puta).

Video kamera nam daje izlazne signale u red,green i blue color sustavu (RGB) m ti signali se matematički pretvaraju u luminantne krominantne signale tj u komponentne koje se odnose na svjetlinu luminancija i njih označavamo sa „Y“ i s druge strane imamo podatke koje se odnose na ton tj krominantne i označavamo ih sa oznakama „Cb“ i „Cr“.

Konverzija se radi jednostavnim matematičkim opracijama. Y signal se dobiva zbrajanjem signala RGB gdje crveni primar R sudjeluje 30%, primar G sudjeluje sa 59% i primar B sudjeluje sa 11%. Svaka komponenta ima različit faktor množenja zbog toga što ljudsko oko ima različitu osjetljivost, tj. ima najveću osjetljivost na zeleni dio spektra, drugi po osjetljivosti je crveni a plavi dio spektra je najmanje osjetljiv.

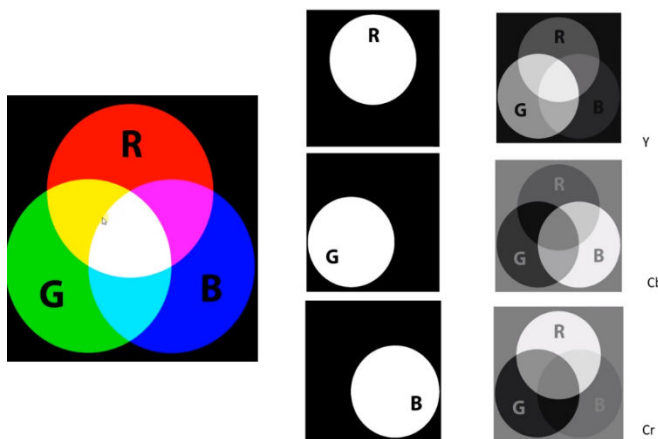
Formula za druge komponente koje označavaju ton :

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

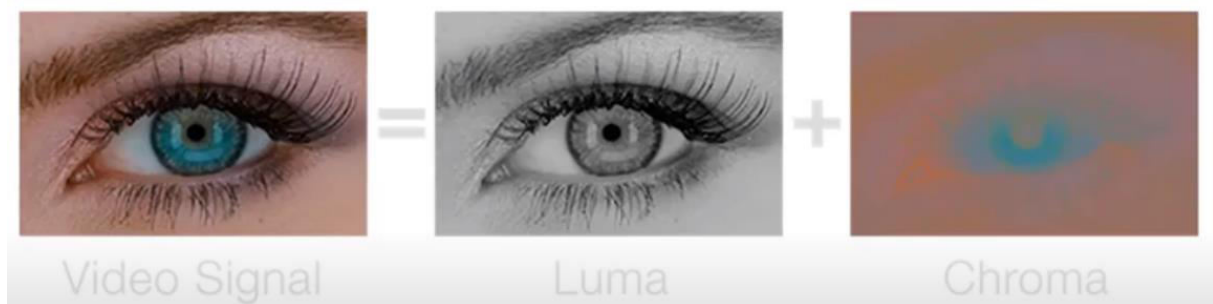
$$Cr = 0.71 * (R - Y)$$

Kada smo razlučili boju na luminantne i krominantne komponente, možemo ih reducirati prema

različitim odnosima ovisno o kvalitete slike koju želimo dobiti.

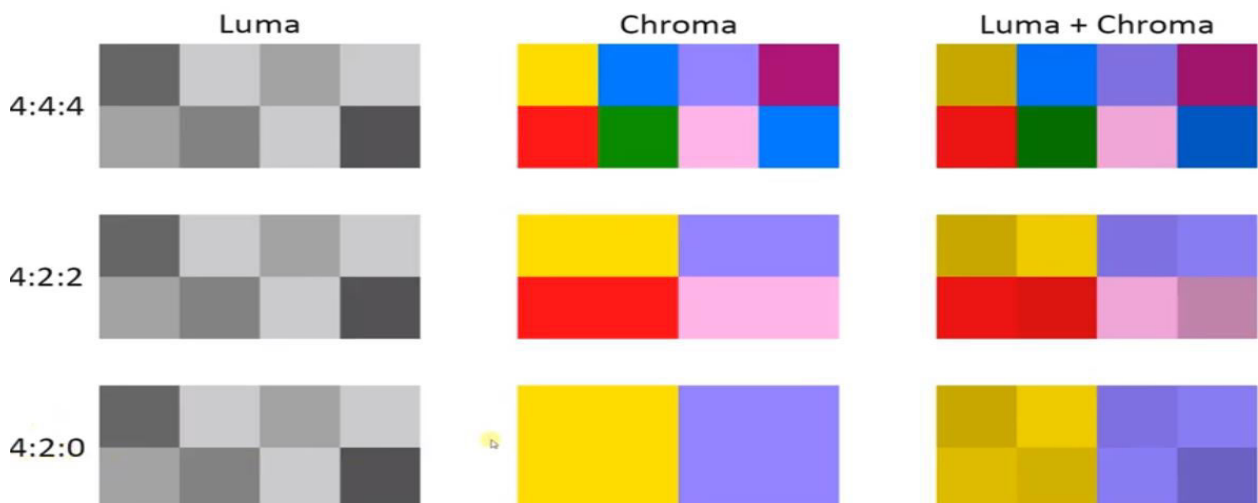


Na ovoj slici nam se prikazuje usporedba RGB kanala sa Cr. Kada rastavimo sliku u boji RGB na kanale RGB dobijemo za rezultat da bijela predstavlja maksimalnu količinu neke boje, dok crna predstavlja odsustvo te komponente. Tako da za RGB kanal u ovom slučaju imamo maksimalnu crvenu, zelenu i plavu boju u određenim krugovima kao što je prikazano na slici.

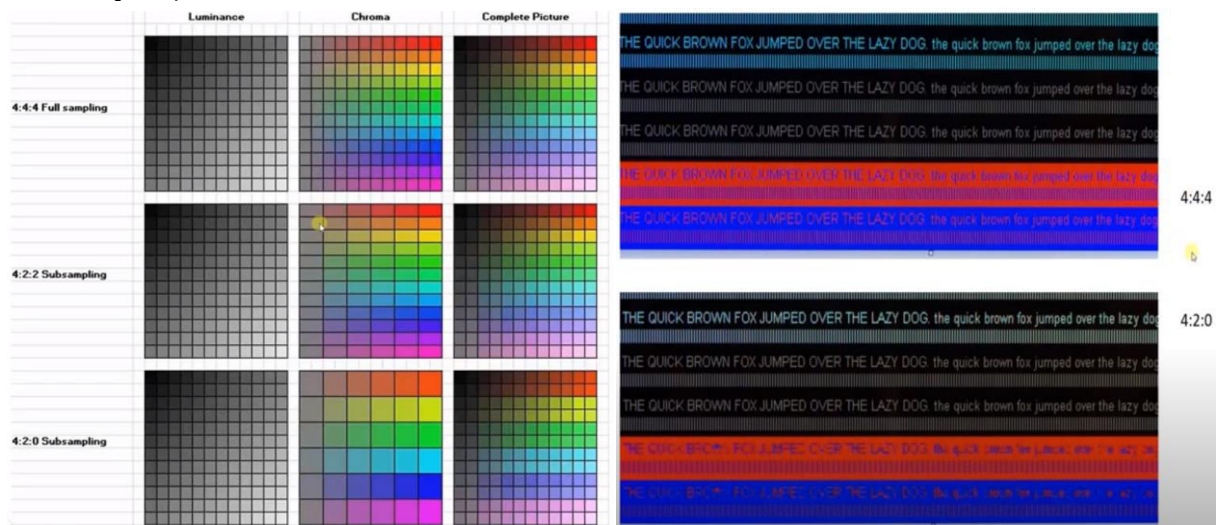


Na ovoj slici vidimo da se video signal razlučuje na luminantni dio (signal koji doživljavamo kao crno bijelu sliku) i na chroma dio koji nam predstavlja podatke o boji odnosno tonu.

U video rječniku često možemo čuti o omjerima određenih brojeva, ti brojevi predstavljaju odnos frekvencija uzrokovanja za luminantnu i dvije krominantne komponente video signala. Taj proces nazivamo uzrokovanje boje odnosno chromasubsampling. Kada govorimo o omjeru 4:4:4 to je originalna rezolucija slike nakon preračunavanja iz RGB u Y Cb Cr sustav. Za sva četiri uzorka Y komponente imamo četiri uzorka Cb i Cr. Četvorka se u ovom slučaju odnosi na veličinu utorka a obično su to četiri piksela. Druga dva broja se odnose na Chroma komponente i oba su relativna prema prvom broju i oni definiraju horizontalno i vertikalno uzrokovanje.



Video signal može podnijeti puno veću kompresiju slike zbog tromosti našeg oka da primijeti neke sitne razlike u detaljima dok se dešava pokret tj. više smo usredotočeni na pokret nego na detalje na slici. Ako gledamo statičnu sliku onda ćemo primijetiti razlike kao što možemo vidjeti na ovim primjerima:



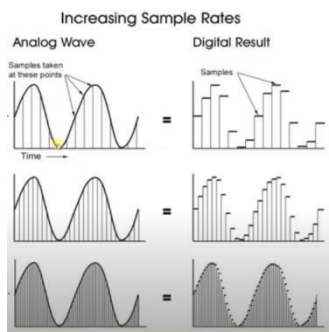
Kada se želimo koncentrirati na sitne detalje kao što je tekst onda ćemo koristiti 4:4:4 omjer, za razliku od 4:2:0 gdje vidimo da su se ti sitni detalji izgubili i tekst je zapravo ostao nečitljiv.

Kada govorimo o video jako malo imamo takvih primjera gdje stvarno moramo gledati detalje, kod pokretne slike je ovakav omjer sasvim zadovoljavajući.

## Digitalni audio signal

Analogni signal se snima različitim uređajima te se putem filtera emitira njegov raspon prije nego što se on digitalizira, taj proces se zove sempliranje ili uzorkovanje signala. Analogni zvuk koji je kontinuiran pretvaramo u uzorke koji su mjerljivi, analogni signal nije mjerljiv ali njega moramo kroz filtere učiniti mjerljivim te na kraju dobivamo nekakav konačan broj uzoraka koji još nazivamo i indiskretnim brojem uzoraka.

Brzina uzorkovanja se mjeri kao sample per second odnosno broj snimljenih audio uzoraka unutar jedne sekunde. To je mjera slična frame per second odnosno brzina izmjenjivanja slika, kada govorimo o videu. Za audio signal najmanji broj uzoraka po sekundi je 8 tisuća Hz odnosno 8Kz. Ako koristimo manju frekvenciju zvuk će se činiti isprekidanim slično kao što će slika trzati pri manjim frame rateu, tj. Nećemo primjećivati kontinuitet nego ćemo primjećivati zasebne slike odnosno zasebne zvukove.

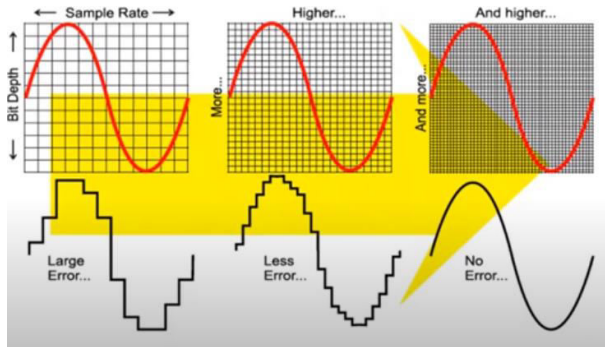


Sampling rate je mjerna na x-osi zvučnog vala koja nam kaže koliko je finoće uzoraka u jedinici vremena.

Na ovoj slici vidimo različit sample rate u odnosu na analogni signal zvuka. S lijeve strane imamo analogni val a s desne strane vidimo prema broju uzoraka kakav rezultat smo dobili. Ako je uzorkovanje premalo nećemo dobiti sličan val analognom i zvuk će biti lošiji. Kvaliteta zvuka neće biti jednaka kao što je kvaliteta analognog zvuka. S većim sample rateom dobivamo sličniji val onome koji smo

snimili analogno ali moramo paziti da taj sample rate ne bude prevelik zbog toga što kada idemo u veće veličine sample ratea nećemo dobiti puno kvalitetniji zvuk jer naše uho neće čuti tu kvalitetu.

Druga mjera o kojoj govorimo kod zvučnog signala je sampling depth odnosno preciznost uzorka, ona se mjeri u bitovima po uzorku a jednako je tako važna kada pričamo o kvaliteti zvuka. Ta mjera nam određuje koliko razina amplitude može postići zvučni signal. To je finoća podjele jačine zvuka na y-osi.



Na ovoj slici možemo vidjeti na y-osi sampling depth kako se na kraju producira val kada imamo veći bit ili manji. Ako govorimo o 8-bitnom zvuku to znači da možemo postići 256 različitih amplituda odnosno jačina zvuka, a utorak o 16 bita može postići 65,5 tisuća različitih amplituda zvuka.

Treća celičina o kojoj pričamo je bit rate koju zapravo imamo i kod video signala, a bit rate kod audio signala je produkt preciznosti uzorka odnosno sample depth pomnožen sa brzinom sampling rate a odnosno brzinom uzrokovanja.

$$\text{Bit rate} = \text{Sample rate} * \text{Sample depth}$$

$$= \text{uzorak / sek} * \text{bit / uzorku} = \text{bit / sek}$$

Komprimirani audio signal uvijek ima bit rate koji je manji od umnoška sample ratea i sample deptha tj

frekvencije i amplitude koju ljudsko oko ne može čuti se režu i komprimiraju. Zvuk se također kao i video signal moram komprimirati jer je njegov ne komprimirani kapacitet vrlo velik. Kod zvuka se također služimo tehnikama kompresije odnosno redukcije nevažnih i suvišnih podataka kao što je to i kod slike. Sve metode redukcije koriste nesavršenost ljudskog uha kako bi maknuli nepotrebne signale iz prijenosa. Neki zvukovi većih amplituda će maskirati zvukove odmah pored njih koji su manjih amplituda te ih uho neće ni čuti što znači da se te manje amplitude mogu ukloniti bez straha da će se primijetiti razlika. Redukcija nevažnih i suvišnih podataka može smanjiti prijenos zvučnog signala čak i do 90% od originalnog snimljenog signala a da se značajno ne naruši kvaliteta zvuka.

- Najčešće korišteni sampling rate:
- 8 kHz – telefonija
- 44.1/48 kHz – TV / CD
- 96/192 kHz – blu-ray
- > 300 kHz

Nyquist-Shannonov teorem kaže da se kontinuirani izvor signala može uzrokovati i savršeno rekonstruirati iz tih uzoraka najmanje dvostrukom frekvencijom od najveće moguće frekvencije u signalu. Tj. Tvrdi da bi se spriječio gubitak informacije kada se signal digitalno semplira. Sample rate mora biti najmanje dvostruke

veličine od najveće očekivane frekvencije signala. Veća frekvencija osigurava bolju kvalitetu zvuka nakon obrade.