

Boja i zvuk u video kompresiji

Prijenos podataka tj. bitrate nekomprimiranog video signala je u rasponu od 270 MB/sec za SDTV, 1,5 GB za HDTV i 3 GB/sec ili više ultra high definiton TV. Takva količina podataka je prevelika za efikasnu pohranu ili transmisiju. Zato svi podatci moraju biti sažeti tj. komprimirani. Za komprimiranje koristimo perceptualno kodiranje da bi smo smanjili količinu podataka. Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa za primanje signala (oči i uši) da bi se ustanovili limiti naše percepcije i u skladu tome reducirali podatci koje mi ne možemo primijetiti. Te se operacije odvijaju prilikom transmisije video signala ili prilikom snimanja. Kod transmisije („stremanja“) moramo voditi računa o tome koliko podataka šaljemo kako bi primatelj signala mogao pratiti video nesmetano.

Kodiranje izvora (source-coding)- sažimanje podataka direktno prilikom snimanja. Sažima se analogni signal koji dolazi na senzore uređaja za snimanje za razliku od post-procesiranja.

Nakon kodiranja izvora količina podataka će biti svedena na 1 do 15 MB ovisno o video kodeku i željenoj rezoluciji videa/zvuka.

Taj proces se izvodi na dva načina:

Reduciranje suvišnih (ponavljanja) i nevažnih podataka (što ljudsko oko ne primijeti da nedostaju)

1. Kompresija bez gubitaka (losless)- suvišni podatci se ponavljaju u podatkovnom prijenosu ili podatci koji se mogu lako i bez gubitka izračunati nekim matematičkim algoritmima prilikom dekodiranja. Znači dijelovi koda se zamjenjuju kraćim kodovima koji ih matematički opisuju.
2. Kompresija s gubitcima (lossy)- nevažni podatci se brišu kako bi se signal smanjio. Pošto oko više primjećuje izmjenu svijetlo/sjena nego boje, preciznost boje može biti reducirana. Takve izbrisane podatke je nemoguće vratiti, pa se zato kompresija naziva kompresija s gubitcima.

Redukcija podataka za boju

Video kamera nam daje izlazne signale u RGB kolor sustavu. Ti signali se matematički pretvaraju u luminantne (Y) i krominantne komponente (Cb i Cr), takav sustav boja označava se sa YCbCr.

Konverzija se radi jednostavnim matematičkim radnjama:

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

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

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

Ljudsko uho ima dinamički raspon od oko 140 decibelai raspon frekvencija od 20 herca do 20 000 herca. Valovi veće frekvencije od 20 000 herca nazivaju se ultrazvuk, a frekvencije manje od 20 herca nazivaju se infrazvuk.

Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha.

Analogni signal snima se različitim uređajima te se putem filtera limitira njegov raspon prije nego što se digitalizira. Taj proces zove se sempliranje (uzorkovanje signala). Time analogni signal činimo mjerljivim.

Sampling rate (brzina uzorkovanja) se mjeri kao broj snimljenih audio uzoraka unutar jedne sekunde.

Sampling depth (preciznost uzorka) se mjeri u bitovima po uzorku. Ta mjera određuje koliko mogućih razina amplitude može postići zvučni signal.