1. **Rezolucija čovjekovog čujnog sustava**

Rezolucija čovjekovog čujnog sustava je ograničena, a uz to i frekvencijski ovisna. Već smo vidjeli da se u pužnici nalaze osjetilne stanice koje reagiraju na audio valove. Pužnica pri tome funkcionira kao banka filtara (frekvencijskih) i to tako da filtri bliže početku pužnice imaju najveći odziv na najviše frekvencije, a filtri bliže kraju imaju najveći odziv na najniže frekvencije.

To znači da čovjek ima ograničenu sposobnost razlikovati različite frekvencije pri čemu sposobnost razlikovanja različitih frekvencija ovisi upravo o našim audio filtrima. Pojasna širina tih filtara je različita za različite frekvencije, a može se odrediti eksperimentalno. Širine audio filtara nazivaju se Kritični pojasevi (*engl. Critical Band*). Pri tome treba voditi računa da naši audio filtri nemaju oštru granicu nego se jako preklapaju, pa je podjela na Kritične pojaseve samo aproksimacija stvarnosti.

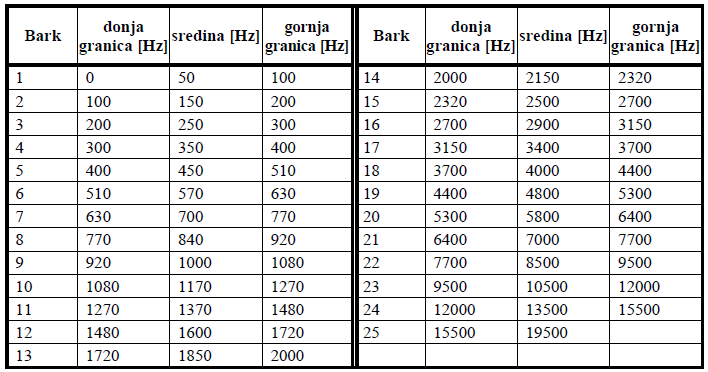
Kritični pojas širok je 100 Hz za niske frekvencije, odnosno više od 4 kHz za najviše frekvencije, pa je cijeli pojas audio frekvencija podijeljen je na 25 kritičnih pojaseva. Jedinica za mjeru kritičnog pojasa je Bark (od Barkhausen). Skala Bark je linearna do frekvencije od 500 Hz, a približno logaritamska na višim frekvencijama. U tablici 6.2 dana je podjela spektra audio frekvencija na kritične pojaseve.

Dva zvuka jednake glasnoće a male razlike u frekvenciji, zvučat će gotovo jednako glasno kao i pojedinačni zvukovi. Razlog je u činjenici da se oba nalaze unutar jednog kritičnog pojasa pa uzbuđuju iste osjetilne stanice što ne doprinosi značajnom povećanju glasnoće. S druge strane, ako dva zvuka upadaju u različite kritične pojaseve, njihovo istodobno slušanje rezultirat će značajnim povećanjem glasnoće. Stoga širokopojasni zvukovi iste razine zvuče glasnije nego uskopojasni zvukovi (oni koji upadaju u jedan kritični pojas).

Kritični pojasevi igraju značajnu ulogu i kod mogućnosti frekvencijskog razlikovanja dvaju zvukova. Minimalna primjetna razlika između dvije uzbude (dva tona) naziva se tek zamjetljiva razlika (just noticeable difference - jnd). Tek zamjetljiva razlika (jnd) za pitch (frekvenciju) ovisi o frekvenciji, razini, trajanju i vrsti zvuka. Unutar jednog kritičnog pojasa nalazi se oko 30 tek zamjetljivih razlika (jnd). To znači da smo za niske frekvencije u stanju razlikovati tonove razmaknute za oko 3 Hz, ali je za visoke frekvencije to značajno više. Tek zamjetljiva razlika (jnd) postoji i za razine zvuka. I kada je u pitanju razina zvuka jnd značajno ovisi karakteristikama zvuka, a načelno možemo kazati da je jnd oko 1 dB. Vidimo da je uistinu rezolucija čovjekovog čujnog sustava ograničena, a uz to i ovisna i to ne samo o frekvenciji nego i o drugim karakteristikama zvuka.

Kritični pojas je mjera koja je bliska mjeri mel (od melody). Mel je percepcijska skala

(definirana od slušača) takva da razlika u distanci između pojedinih jedinica percepcijski ima uvijek isto povećanje frekvencije. To znači da povećanje frekvencije od Barka 3 do Barka 4 (od 250 Hz do 350 Hz) percepcijski doživljavamo isto kao i povećanje frekvencije od Barka 14 do Baka 15 (od 2150 Hz do 2500 Hz).

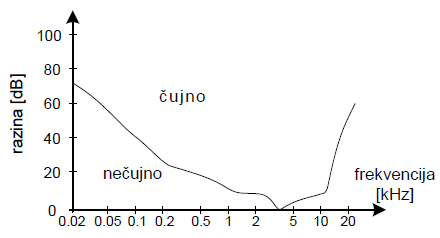


**Tablica 6.2** Podjela spektra audio frekvencija na kritične pojaseve

1. **Prag čujnosti, Frekvencijsko maskiranje, Vremensko maskiranje**

PRAG ČUJNOSTI:

Zamislimo slijedeći eksperiment. Nalazimo se u tihoj prostoriji. Generirajmo zvučni val na nekoj frekvenciji (npr. 1 kHz) vrlo male razine (npr. 0.1 dB). Na toj frekvenciji i s tom razinom nećemo čuti zvuk. Povećajmo razinu dok ne čujemo zvuk, što će na toj frekvenciji biti oko 10 dB. Ponovimo eksperiment za sve frekvencije. Time smo dobili osjetljivost ljudske percepcije zvuka u ovisnosti o frekvenciji (prag čujnosti) što je prikazano na slici 6.11. To je zapravo prag čujnosti (na tišini) koji kaže kolika je razina zvuka potrebna na pojedinim frekvencijama da bi zvučni val uopće čuli.



**Slika 6.11** Prag čujnosti (na tišini)

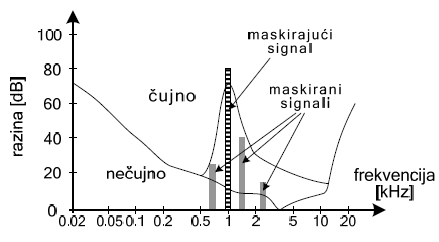
Dakle, signal na određenoj frekvenciji treba biti iznad praga kako bi ga mogli čuti. Naravno, ako je signal ispod praga čujnosti nema niti potrebe da ga prenosimo. Ovu karakteristiku koriste mnogi koderi kako bi učinkovitije komprimirali audio signal.

FREKVENCIJSKO MASKIRANJE:

Frekvencijsko maskiranje je percepcijska pojava kod koje signal manje razine (maskirani signal) može postati nečujan (maskiran) ako istovremeno postoji i signal veće razine (maskirajući signal). Frekvencijsko maskiranje je najveće u kritičnom pojasu u kojem je maskirajući signal lociran, dok je utjecaj u susjednim kritičnim pojasevima manji.

Zamislimo slijedeći eksperiment. Generirajmo zvučni val na nekoj frekvenciji (npr. 1000 Hz) s razinom koja je iznad praga čujnosti (npr. 80 dB). Paralelno s tim tonom generirajmo još jedan zvučni val bliske frekvencije (npr. 1050 Hz) malo iznad praga čujnosti. Bez obzira što je drugi ton iznad praga čujnosti nećemo ga čuti jer je frekvencijski maskiran zbog prisustva glasnijeg tona bliske frekvencije. Dakle, prag čujnosti se mijenja u ovisnosti o razini frekvencijski bliskih signala.

Prag čujnosti kod frekvencijskog maskiranja za signal frekvencije 1 kHz razine 80 dB prikazan je na slici 6.12. Krivulja praga čujnosti je manja strma prema višim frekvencijama što znači da se više frekvencije lakše frekvencijski maskiraju.



**Slika 6.12** Frekvencijsko maskiranje audio signala u ovisnosti o amplitudi

Na slici 6.13 prikazano je frekvencijsko maskiranje za slučaj iste razine, ali različitih frekvencija. Očito je da je na većim frekvencijama efekt maskiranja (pojasna širina) veći nego na nižim frekvencijama što je u skladu i sa širinom kritičnih pojaseva. Tako će većina ljudi primijetiti razliku između tonova od 1kHz i 3kHz dok je samo manji broj ljudi u stanju primijetiti razliku između tonova od 14kHz i 17kHz.

Ovdje je opisano samo maskiranje uslijed jednog maskirajućeg signala. Ako istovremeno postoji više maskirajućih signala, svaki od njih ima svoj prag maskiranja, te je u tom slučaju potrebno izračunati globalni prag maskiranja.

Efekt frekvencijskog maskiranja može se iskoristiti za kompresiju audio signala tako da se ne kodiraju oni signali koji su frekvencijski maskirani tj. oni koji su ispod praga čujnosti. Ali, efekt frekvencijskog maskiranja može se iskoristiti i tako da se za prijenos signala koji nisu maskirani potroši manje bita nego što bi to bio slučaj bez maskiranja. Distorzija (koja se može čuti) u danom podpojasu je NMR (slika 6.14) i mjeri se kao:

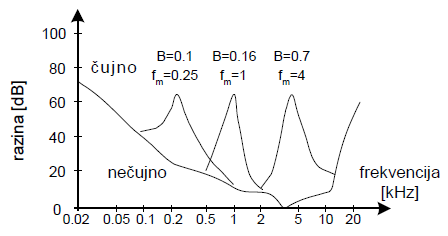
NMR(m) = SMR − SNR(m) [dB]

gdje je

NMR(m) - Omjer šuma i maskiranja za m-bitni kvantizator (*engl. noise-to-mask ratio*),

SMR - Omjer signala i maskiranja (*engl. signal-to-mask ratio*) - maksimalna vrijednost je na lijevoj granici kritičnog pojasa a minimalna je na frekvenciji maskirajućeg signala i iznosi oko 6dB,

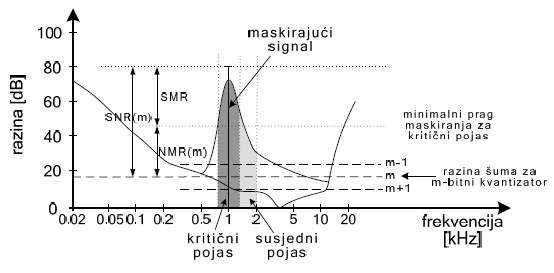
SNR(m) - Omjer signala i šuma za m-bitni kvantizator (*engl. signal-to-noise ratio*).



**Slika 6.13** Frekvencijsko maskiranje audio signala – pojasna širina u ovisnosti o frekvenciji

Šum zbog greške kvantizacije neće se čuti sve dok je ispod praga čujnosti. To znači da se distorzija, unutar kritičnog pojasa, neće čuti sve dok je SNR(m) veći od SMR, odnosno dok je NMR(m) negativan. Što je efekt maskiranja veći, to je manje bita potrebno da bi razina šuma bila ispod praga čujnosti. Šum zbog greške kvantizacije će u susjednim pojasevima također biti maskiran ali u manjoj mjeri nego unutar kritičnog pojasa.

Opisali smo slučaj samo jednog maskirajućeg signala. Kao što smo već rekli, ako se signal sastoji od više maskirajućih signala potrebno je odrediti globalni prag maskiranja.



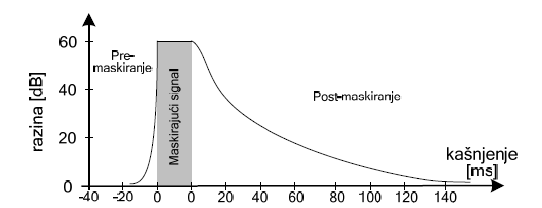
**Slika 6.14** Frekvencijsko maskiranje i šum kvantizacije

VREMENSKO MASKIRANJE:

Vremensko maskiranje je percepcijska pojava u vremenskom području koja se javlja kada u kratkom vremenskom razmaku postoje dva zvučna signala.

Zamislimo slijedeći eksperiment. Generirajmo neki zvučni val (ton) koji je iznad praga čujnosti te paralelno s tim tonom generirajmo još jedan na bliskoj frekvenciji koji je također iznad praga čujnosti (za slučaj tišine), ali koji se u ovom slučaju ne čuje jer je frekvencijski maskiran. Zatim prekinimo generiranje tona kojeg čujemo. Signal koji je do tada bio maskiran nećemo odmah čuti nego tek nakon nekog vremena. Ta pojava se zove vremensko maskiranje što je prikazano na slici 6.15.

U ovisnosti o razinama signala, jači zvuk može maskirati slabiji čak i ako maskirani signal prethodi maskirajućem (pre-maskiranje). Efekt post-maskiranja traje od oko 50 ms do oko 200 ms, dok je efekt pre-maskiranja značajno kraći, oko jedne desetine post-maskiranja. Vremensko maskiranje se, kao i frekvencijsko, može iskoristiti za kompresiju audio signala.



**Slika 6.15** Vrem. maskiranje audio signala (maskirajući signal 1 kHz, 60 dB; maskirani signal 1.1 kHz)

1. **Osnovne tehničke karakteristike MPEG-1/2 audio kodera**

MPEG-1/2 koriste sličnu audio kompresiju (MPEG-2 podržava praktično isti audio koder kao i MPEG-1. Međutim, MPEG-2 podržava još jedan koder, MPEG AAC, koji ima značajna poboljšanja u odnosu na MPEG-1 audio koder.). Algoritam je definiran kroz tri sloja (layer I, II, III) uz neke dodatne mogućnosti kod MPEG-2 (npr. surround zvuk). MPEG-1 audio podržava frekvencije uzorkovanja od 32 kHz, 44.1 kHz (CD) i 48 kHz (DAT) te četiri moda:

**1**. Monophonic – jedan audio kanal,

**2**. Dual-monophonic – dva neovisna kanala, npr. za engleski i hrvatski jezik,

**3**. Stereo – za stereo kanale,

**4**. Joint-stereo – za stereo kanale koji koriste korelaciju kako bi komprimirali signal.

MPEG-2 je, u odnosu na MPEG-1, proširio frekvencije uzorkovanja još i na 16 kHz, 22.05 kHz i 24 kHz te je još dodao i mogućnost višekanalnog moda (5+1 kanala – surround zvuk).

Algoritam kompresije koji se koristi u MPEG-1/2 ima slijedeće korake:

**1**. Audio signal transformirati u frekvencijsko područje, te dobiveni spektar podijeliti u 32 frekvencijska podpojasa koja se međusobno ne preklapaju – to je podpojasno filtriranje (*engl. subband filtering*).

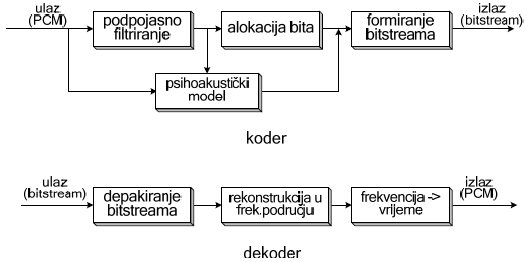
**2**. Odrediti, upotrebom psihoakustičnog modela, iznos maskiranja koje uzrokuje susjedni pojas.

**3**. Ako je snaga u trenutnom pojasu manja od praga maskiranja, pojas se ne kodira.

**4**. U protivnom, odrediti broj bita potreban za prijenos koeficijenta koristeći psihoakustički model i to tako da je šum kvantizacije manji od efekta maskiranja (jedan bit manje u kvantizatoru povećava šum za 6 dB). Ako je greška kvantiziranja manja od praga maskiranja tada se komprimirani audio signal neće (percepcijski) razlikovati od originalnog signala.

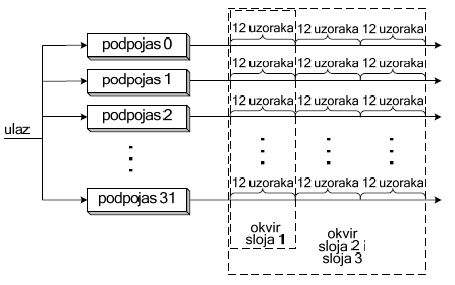
**5.** Formirati tok bitova (*engl. bitstream*).

MPEG-1/2 definiraju 3 različita sloja (metode kompresije) za audio signal. Temeljni model je isti, ali složenost raste sa svakim slojem. Na slici 7.10 dana je pojednostavljena shema MPEG-1/2 audio kodera i dekodera. U nastavku su dane detaljnije sheme za MPEG sloj I/II te III.

****

**Slika 7.10** Pojednostavljena shema MPEG-1/2 audio kodera i dekodera

MPEG-1/2 audio algoritmom podaci se dijele u okvire od kojih svaki sadržava 384 uzorka (12 uzoraka za svaki od 32 podpojasa), što je prikazano na slici 7.11. Ekvivalent 384 uzorka, uz frekvenciju uzorkovanja od 44100 Hz, je 8.71 msec.

****

**Slika 7.11** MPEG-1/2 audio: podjela na okvire

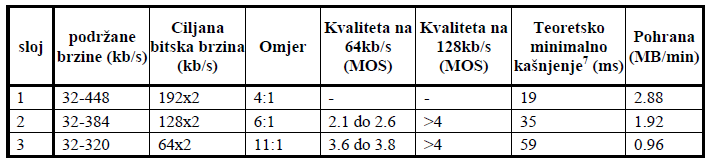
Osnovna svojstva MPEG-1/2 kodera/dekodera po pojedinim slojevima su:

**Sloj I**: Za podpojasno filtriranje koristi se Filter bank. Koristi se samo jedan okvir. Psihoakustični model koristi samo frekvencijsko maskiranje.

**Sloj II**: Kao i Sloj 1 koristi Filter bank za podpojasno filtriranje. Koristi tri okvira u filtru (prethodni, trenutni i slijedeći – ukupno 3x384=1152 uzorka). Koristi osim frekvencijskog maskiranja i jedan dio vremenskog maskiranja.

**Sloj III**:Koristi bolji filtar za podpojasno filtriranje (filter bank + MDCT (Modified DCT)), koristi tri okvira, psihoakustički model koristi osim frekvencijskog maskiranja i vremensko maskiranje, može uzeti u obzir i stereo redundanciju. Koristi Huffmanovo kodiranje.

U tablici 7.2 dane su osnovne karakteristike pojedinih slojeva MPEG-1/2 audio algoritma.

****

**Tablica 7.2** MPEG audio kodiranje: karakteristike slojeva 1, 2 i 3

Sloj 1 je bio ciljano napravljen za DCC (Digital Compact Cassette). Sloj 3 je imao cilj komprimirati audio signal što je moguće bolje (više) bez obzira na složenost algoritma. Sloj 2 je napravljen kao kompromis između složenosti sloja 3 i performansi sloja 1.

Ciljana brzina je ona koja praktično osigurava CD kvalitetu. Iskorištavajući stereo efekt, može se postići i manja brzina (to može samo sloj 3). Tipične brzine i pripadajuće kvalitete (stereo signala) za MP3 su:

192-256 CD kvaliteta,

160 uglavnom CD kvaliteta,

96-128 skoro CD kvaliteta,

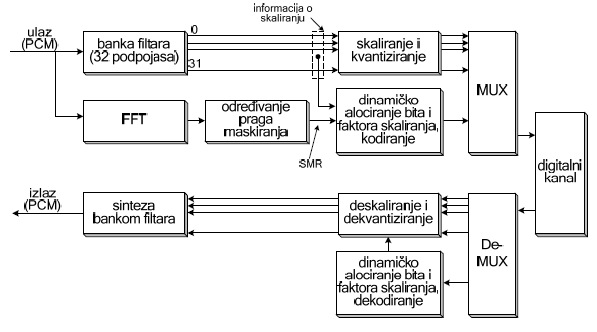
36-64 skoro FM kvaliteta.

Uz brzinu prijenosa od 2x128 kb/s niti ekspert ne može razlikovati originalni i kodirani audio.

Podsjetimo, originalni CD audio ima brzinu 1.411 Mb/s pa je to kompresija od skoro 6 puta. To znači da glazbeni CD koji inače može sadržavati oko 1 sat glazbe (točnije 650MB=74 min) sada može sadržavati skoro 6 sati glazbe CD kvalitete ili 12 sati glazbe skoro CD kvalitete.

1. **MPEG-1/2 audio koder sloj I i II**

MPEG-1/2 sloj I i II imaju sličnu strukturu (slika 7.12). Najveća razlika je nešto boljoj (finijoj) kvantizaciji koja se koristi kod sloja II u odnosu na sloj I te što sloj II koristi dio vremenskog maskiranja.

****

**Slika 7.12** Shema MPEG-1/2 audio sloja I i II

Sloj I i II koristi banku filtara (pojasne filtre) da bi podijelio ulazni signal na 32 jednako razmaknuta podpojasa. Primjerice, za frekvenciju uzorkovanja od 44.1 kHz, širina podpojaseva je 689 Hz. Jednako razmaknuti podpojasevi ne odgovaraju kritičnim pojasevima. Dapače, na nižim frekvencijama jedan podpojas obuhvaća nekoliko kritičnih pojaseva što predstavlja nedostatak ovog pristupa.

Psihoakustički model za određivanje praga maskiranja ne koristi banku filtara nego posebno i neovisno prebacivanje iz vremenskog u frekvencijsko područje jer za izračun praga maskiranja treba puno preciznija frekvencijska rezolucija. Za to se koristi Fourierova transformacija. Sloj I koristi FFT (Fast Fourier Transform) s analizom 512 uzoraka, a Sloj II i Sloj III FFT s analizom 1024 uzorka.

Prije kvantiziranja se na bloku od 12 uzoraka izvrši skaliranje i to tako da se vrijednosti uzoraka podijele s faktorom skaliranja tako da je vrijednost najvećeg uzorka jednaka jedan. Potom se potreban broj bita za svaku spektralnu komponentu određuje dinamičkim alociranjem bita na temelju psihoakustičkog modela. Pri tome algoritam za alociranje bita, osim psihoakustičkog modela uzima u obzir i raspoloživu (željenu) brzinu. Iterativna procedura minimizira NMR (noise-to-mask ratio) u svakom podpojasu. Alociranjem bita definira se broj bita kojim će se kodirati svaki uzorak. Ukoliko je taj broj bita različit od nula, dekoderu se šalje još i faktor skaliranja. Faktor skaliranja je broj kojim će dekoder pomnožiti kvantiziranu vrijednost da bi dobio stvarnu vrijednost uzorka. Kombinacijom alociranja bita i faktora skaliranja može se dobiti dinamičko područje i veće od 120dB (ali ne za sve podpojaseve).

Sloj II, za razliku od Sloja I, koristi blokove od 3x12 uzoraka, pri čemu faktor skaliranja može biti isti za sva tri bloka od 12 uzoraka, isti samo za dva ili različit za sva tri bloka uzoraka, u ovisnosti o njihovim stvarnim vrijednostima. Ako to neće prouzročiti distorzije, koristi se isti faktor skaliranja za sva tri bloka. Ovim pristupom se smanji potreban broj bita za prijenos faktora skaliranja u odnosu na Sloj I i to je jedna od većih razlika sloja I i II. U konačnici se do dekodera prenose skalirane i kvantizirane spektralne vrijednosti, faktor skaliranja i informacija o broju alociranih bita.

**MPEG-1/2 audio koder sloj III**

Shema MPEG-1/2 audio sloja III dana je na slici 7.13. Najznačajnije razlike MPEG-1/2 sloja III u

odnosu na slojeve I i II su:

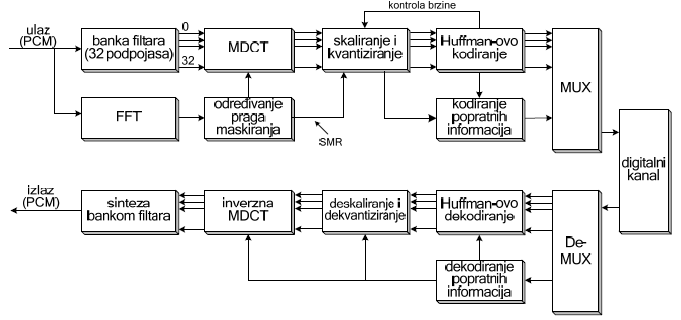
• MDCT. Kako bi postigao bolju frekvencijsku rezoluciju bližu podjeli na kritične pojaseve, Sloj III dodatno u frekvencijskom području dijeli svaki od 32 podpojasa i to pomoću modificirane diskretne kosinusne transformacije (MDCT - Modified Dicrete Cosine Transform).

• Faktor skaliranja u pojasevima. Za razliku od Sloja I i II koji svakom podpojasu pridijele posebni faktor skaliranja, Sloj III koristi isti faktor skaliranja za više MDCT koeficijenata formirajući tako pojasnu širinu otprilike jednaku kritičnim pojasevima.

• Entropijsko kodiranje podataka. Zbog boljeg sažimanja, Sloj III koristi Huffmanovo kodiranje.

• Korištenje rezervoara bitova (*engl. bit reservoir*). Svaka skupina od 1152 (384x3) uzorka ne mora se kodirati istim brojem bitova. Koder može za pojedinu skupinu uzoraka koji su zahtjevni za kodirati „posuditi“ bitove od skupine manje zahtjevnih uzoraka. Time se praktično koristi kodiranja s promjenjivom brzinom prijenosa (*engl. Variable bit rate*  *coding*)

• Korištenje vremenskog maskiranja. Sloj III koristi nešto bolji psihoakustički model od Slojeva I i II jer uključuje i vremensko maskiranje.

****

**Slika 7.13** Shema MPEG-1/2 audio sloja III

1. **MPEG-4 Advanced Audio Coding (MPEG-4 AAC)**

Osim poznatog MPEG Audio Layer 3 kodera (MP3) razvijen je i 1997. godine standardiziran MPEG Advanced Audio Coding (MPEG AAC) koder koji je poboljšana verzija u odnosu na MPEG Audio Layer 3.

MPEG-4 AAC koder (slika 7.14) ima istu temeljnu ideju kao i MPEG-1/2 koder (percepcijsko podpojasno kodiranje), ali uz neka poboljšanja od koji je dio naveden:

• Širi opseg frekvencija uzorkovanja (od 8 kHz do 96 kHz) u odnosu na MPEG-1/2 (16 kHz do 48 kHz).

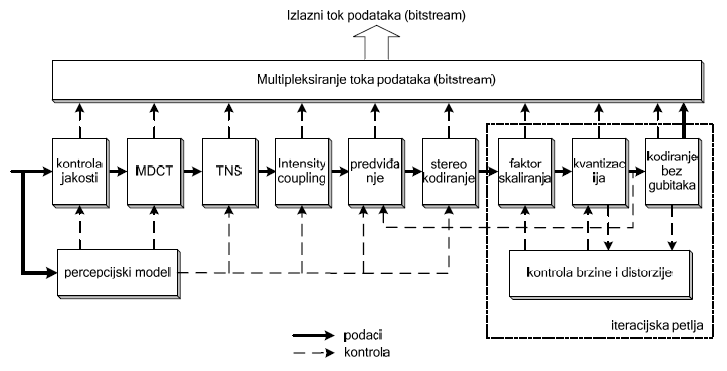
• Podržava do 48 kanala (MPEG-1/2 audio koder podržava 2 kanala u MPEG-1 modu, te maksimalno 5+1 kanala u MPEG-2 modu).

• Poboljšano i pojednostavljeno pojasno filtriranje jer je napušten hibridni model pa se koristi samo MDCT.

• Korištenje predviđanja (koristi činjenicu da su neki uzorci zvuka jednostavni za predviđanje, prediktor gleda prethodna dva uzorka kako bi analizirao zvučni uzorak).

• Temporal Noise Shapening (TNS) – vremenski ovisna kvantizacija koja koristi činjenicu da tonalni signal u vremenskom području ima tranzijentne vrhove u frekvencijskoj domeni i obrnuto, tj. signal koji je tranzijentan u vremenu je „tonalan“ u frekvencijskom području (tonalni signal je periodičko ponavljanje zvuka dok je tranzijentni signal kratkotrajni zvuk koji se jako brzo mijenja od malih do velikih amplituda ili obrnuto). Primjenom TNS moguće je u vremenskoj domeni raspršiti šum kvantizacije.

• Dinamičko mijenjanje veličine MDCT bloka (jedan blok od 1024 točaka ili 8 blokova od 128 točaka – ako se signal mijenja ili postoje tranzijenti koristi se 8 kraćih prozora od 128 točaka, u suprotnom se koristi prozor od 1024 točke jer se većim prozorom postiže bolja frekvencijska rezolucija pa time i mogućnost primjene složenijeg i boljeg psihoakustičkog modela).

****

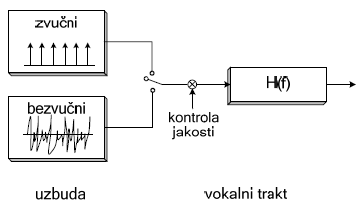
**Slika 7.14** Shema AAC kodera – neki blokovi su opcionalni

MPEG-4 AAC ima prednost u odnosu na MP3 kod niskih brzina (manje od 128 kb/s). Kod većih brzina ta prednost nije toliko izražena. MPEG-4 AAC definira nekoliko kompresijskih shema (profila) za različite primjene.

1. **Govor specifični algoritmi**

Kodiranje valnog oblika kao npr. kod DPCM, ADPCM, DM i sl. ima cilj prenijeti (kodirati) valni oblik što je moguće vjernije. MP3 i slični algoritmi koji koriste psihoakustični model također prenose valni oblik. Istina, to je valni oblik koji nije jednak originalnom jer je cilj prenijeti informaciju, a ne sami valni oblik, ali ipak ono što se prenosi još uvijek ima formu valnog oblika. Kada je u pitanju govorni signal, može se umjesto valnog oblika prenijeti informacija o tome kako čovjekov govorni sustav generira (stvara) govorni signal. Koderi koji koriste takav pristup (sintezu govora) zovu se vokoderi (voice+coder). Vokoderom se može kodirati govor s brzinom ispod 4.8 kb/s. Iako je na toj brzini govor razumljiv ipak zvuči umjetno, a prepoznavanje govornika je otežano. Podsjetimo da se standardni govorni signal (8 kHz, 8 bita/uzorku) prenosi s brzinom od 64 kb/s, te da algoritmi koji kodiraju valni oblik ne mogu uspješno kodirati s brzinom manjom od 16 kb/s.

Da bi se razumio postupak vokodiranja potrebno je znati mehanizam generiranja (stvaranja) govora što je ukratko opisano u poglavlju 6.5. Možemo kazati da govor karakterizira tip uzbude (zvučni ili bezvučni) te vokalni trakt (formanti i jakost). Shematski prikaz vokodera ilustriran je na slici 7.15. Prvi vokoder (analogni) predstavljen je još 1939. godine.

****

**Slika 7.15** Shematski prikaz vokodera

**Linear Predictive Coding (LPC)**

Prvi šire korišteni vokoderi (digitalni) temelje se na linearnoj predikciji a intenzivnije se razvijaju od kraja 70-tih godina prošlog stoljeća. Linearno prediktivno kodiranje (LPC) temelji se na modeliranju segmenata govora. Najprije se za dani segment govora izračunaju parametri linearnog filtra (LPC koeficijenti) koji definiraju govorni sustav. LPC algoritam parametre filtra određuje tako što nalazi formante signala. Potom se govorni signal filtrira inverznim filtrom - rezultat koji se dobije zove se rezidual. Rezidual u osnovi određuje uzbudu, pa je iz reziduala relativno lako odrediti da li je signal zvučan ili nije, te odrediti pitch period ukoliko je zvučan. Kod LPC vokodera, za razliku od kasnijih složenijih vokodera, uzbuda je jednostavna: to je niz impulsa iste amplitude i istog razmaka. Ovako dobiveni govor zvuči umjetno, a brzina koja se postiže je 2.4 kb/s.

Primjer vokodera koji se temelji na linearnoj predikciji je US Fedaral Standard 1015 (LPC- 10e standard). To je standard, predstavljen 1984. godine, koji definira prijenos govornog signala s 2.4 kb/s uz slijedeće parametre:

brzina uzorkovanja 8 kHz,

dužina okvira 180 uzoraka = 22.5 ms,

10 LPC koeficijenata = 42 bita,

pitch i voicing 7 bita,

jakost 5 bita.

**Residual Excited Linear Prediction (RELP)**

RELP radi slično kao i LPC. Razlika je što se iz reziduala ne traži pitch period već se rezidual kao takav koristi za uzbudu. To znači da se rezidual prenosi do dekodera pri čemu se kodira kao valni oblik. RELP postiže dobru govornu kvalitetu na 9.6 kb/s.

**Code Excited Linear Prediction (CELP)**

Za razliku od LPC, CELP dozvoljava široki izbor različitih uzbuda koja su sve pohranjene u kodnu knjigu. Najbolja uzbuda pronađe se tako da se za svaku moguću uzbudu generira govorni signal i usporedi ga se s originalnim. Izabere se ona uzbuda koja je rezultirala najmanjom razlikom sintetiziranog i originalnog signala. Dakle, ovaj koder koristi vektorsku kvantizaciju da bi kodirao rezidual. CELP postiže brzinu od 4.8 kb/s. Primjeri CELP standarda su US Federal Standard 1016 (CELP 3.2) s brzinom od 4.8 kb/s te G.728 LD-CELP (Low-Delay Code-Excited Linear Prediction).

**Multipulse Excited Coding (MPE) i Regular Pulse Excited Coding (RPE)**

Kao i kod prethodnih metoda, MPE i RPE poboljšavaju kvalitetu govora (u odnosu na LPC) boljim izborom uzbude. MPE kao uzbudu koristi niz impulsa koji za razliku od LPC mogu biti različite amplitude i razmaka. Razmak i položaj odrede se kao i kod CELP metodom pokušaja (tzv. analysisby- synthesis tj. analiza sintezom).

RPE radi na sličan način osim što su, kao što i ime sugerira, impulsi jednako razmaknuti.

GSM mobilni telefoni koriste varijantu RPE kodera koja radi na 13kb/s. To je Regular Pulse Excited

Long Term Prediction (RPE-LTP) definiran u GSM standardu 06.10 s brzinom od 13kb/s.

1. **Percepcija LMS signala kao nijanse, zasićenja i sjajnosti**

LMS signal je odziv fotosenzora na elektromagnetsko zračenje u oku. Kako se LMS signal odnosi prema našem stvarnom doživljaju elektromagnetskog zračenja? Dakle, kako se LMS signal odnosi prema bojama, jer boje su naš doživljaj elektromagnetskog zračenja iz vidljivog dijela spektra?

Kao što postoje tri veličine (LMS) kojima definiramo elektromagnetsko zračenje, tako postoje i tri veličine kojima definiramo percepciju elektromagnetskog zračenja, odnosno boje. To su:

• nijansa (*engl. hue*),

• zasićenje (*engl. saturation*),

• sjajnost (*engl. brightness*).

**Što je to nijansa?**

Nijansa (ili ton) je atribut vizualne percepcije koji je određen dominantnom valnom duljinom.

Pojednostavljeno gledano, ako se pomiče dominantna valna duljina vidljivog spektra (slika 8.2) onda se pomiče (mijenja) i nijansa promatrane boje.

**Što je to zasićenje?**

Zasićenje je atribut vizualne percepcije koji je određen čistoćom boje i ovisi o količini bijelog svjetla (drugih boja) pomiješanog s nijansom. Zasićenje se može kretati od sive boje, preko pastelnih boja do potpuno zasićene boje. Čista nijansa (100% zasićenja) je potpuno zasićena, tj. nema pomiješane niti jedne druge boje. S druge strane 0% zasićenja je siva boja. Nijansu i zasićenje za danu boju zajedničkim imenom zovemo kroma (*engl. chrominance, chroma*).

**Što je to sjajnost?**

Kako ga definira CIE (Commission Internationale de L'Eclairage - International Commission on Illumination), sjajnost je atribut vizualne percepcije prema kojem neko područje izgleda da emitira više ili manje svjetla.

Kako je percepcija sjajnosti jako kompleksna, CIE je definirao osvjetljenje (*engl. luminance*) i svjetloću (*engl. lightness*), dvije veličine koje je puno lakše analitički pratiti (računati). Osvjetljenje i svjetloća su dijelom atributi vizualne percepcije, a dijelom fizikalne veličine. Osim toga definirat ćemo i jednu čisto fizikalnu veličinu vezanu uz percepciju boja; to je intenzitet (*engl. intensity*).

**Što je to intenzitet?**

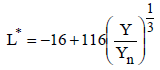
Intenzitet (*engl. intensity*) je mjera toka snage koja je izračena s nekog objekta odnosno koja upada na neki objekt. To je dakle fizikalna veličina; veća snaga odgovara većem intenzitetu. Intenzitet se mjeri u jedinicama Watt/m2.

**Što je to osvjetljenje?**

Osvjetljenje (*engl. luminance*) je tok snage (dakle intenzitet) otežan s funkcijom luminancijske efikasnosti. Funkcija luminancijske efikasnosti za tzv. standardnog promatrača, kako je definira CIE, jednaka je sumi svih triju krivulja sa slike 8.5. Kada se spektralna distribucija snage (slika 8.2) oteža funkcijom luminancijske efikasnosti i integrira u cijelom vidljivom području dobije se osvjetljenje. Dakle, osvjetljenje osim o intenzitetu ovisi i o ljudskom vizualnom sustavu i to preko funkcije luminancijske efikasnosti. Osvjetljenje se označava s Y. Strogo gledano, osvjetljenje bi se trebalo računati u candelima po metru kvadratnom, ali se obično normalizira na vrijednost od 1 do 100 u odnosu na osvjetljenje bijelog. Tako danas kvalitetni monitori imaju osvjetljenje bijele referentne točke od 80 cd/m2, pa se normalizirano osvjetljenje Y=1 uzme da je jednako toj vrijednosti.

**Što je to svjetloća?**

Ljudski vizualni sustav ima nelinearni doživljaj osvjetljenja. Izvor koji ima osvjetljenje samo 18% referentnog osvjetljenja vidjet ćemo kao upola sjajan . Ili npr., promjena normaliziranog osvjetljenja s 0.1 na 0.11 i promjena s 0.5 na 0.55 percepcijski se vidi kao ista promjena sjajnosti [Foley et all. 1990]. Stoga je CIE definirao percepcijski odziv na osvjetljenje koji se zove svjetloća (*engl. lightness*) i označava s L\*. L\* je prema CIE definiran kao modificirani treći korijen osvjetljenja.

****

gdje je Yn osvjetljenje referentne bijele boje. Ako je Y već normaliziran onda nije potrebno dijeliti s Yn. L\* poprima vrijednosti od 1 do 100. Slijedi da je percepcija osvjetljenja, koju nazivamo svjetloća, približno logaritamska. Napomenimo da čovjek ne može razlikovati promjenu svjetloće manju od 1% što odgovara promjeni L\* za jedan; to je rezolucija čovjekovog vizualnog sustava s obzirom na svjetloću. Svjetloća je dakle percepcijsko/fizikalna veličina koja bi trebala što je moguće više odgovarati percepcijskoj veličini sjajnosti. Ipak sjajnost je puno kompleksnija veličina i ovisi o još nizu drugih faktora kao što je npr. boja pozadine.

Elektromagnetsko zračenje samo po sebi ne sadrži nikakve boje. Ne postoji ništa crveno u većim valnim duljinama niti plavo u manjim. Percepcija boja je u potpunosti kreacija mozga i ni na koji drugi način nije sadržana u zračenju. U tom smislu niti gore spomenuti signali LMS nisu boje. To je jednostavno signal odziva fotosenzora na elektromagnetsko zračenje. Tek se u mozgu na temelju tih signala dobije odgovarajuća nijansa, zasićenje i sjajnost. Tako npr. različite spektralne razdiobe mogu rezultirati percepcijski istim bojama. Taj fenomen zove se metamerizam.

1. **Modeli boja za slikovni signal**

**RGB model boja**

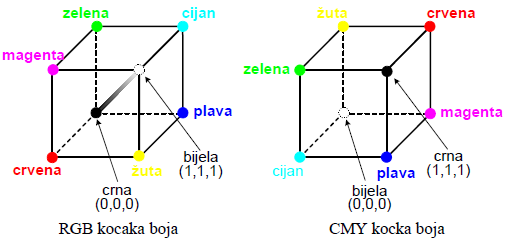
Primarne boje kod RGB modela su crvena, zelena i plava (engl. red, green, blue). Definirati određenu boju u RGB sustavu znači odrediti količinu primarnih boja (crvene, zelene i plave) potrebnu da se njihovim aditivnim miješanjem dobije tražena boja. CIE je 1931. godine odredila standardne valne duljine tih primarnih boja

• crvena 700 nm,

• zelena 546.1 nm,

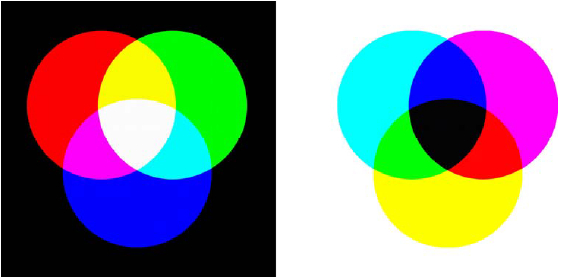
• plava 435.8 nm.

Na slici 8.9 prikazana je RGB kocka, tj. RGB model boja u Cartesianovom koordinatnom sustavu. Svaka točka unutar kocke predstavlja jednu jedinstvenu boju određenu iznosom primarnih boja. Sive boje, dakle one s jednakim iznosom svih triju primarnih boja nalaze se na spojnici crne i bijele boje (slika 8.9).

****

**Slika 8.9** RGB i CMY kocke boja

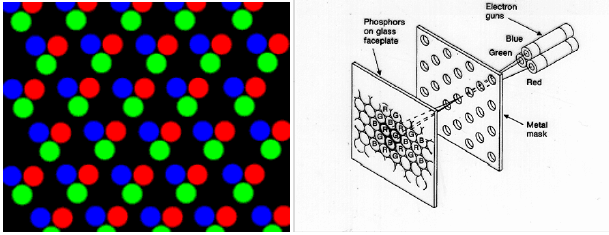
Na slici 8.10 prikazano je aditivno miješanje primarnih boja (crvene, zelene i plave) čime se dobivaju sekundarne boje: žuta (crvena+zelena), cijan (plava+zelena) i magenta (crvena+plava) te bijela boja (crvena+zelena+plava).



RGB model boja CMY model boja

**Slika 8.10** Miješanje primarnih boja za RGB i CMY modele boja

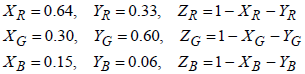
RGB model je aditivni model jer se boje dobivaju dodavanjem primarnih boja. RGB model boja koristi se za CRT (Cathode Ray Tube) monitore i TV prijemnike te općenito sve uređaje koji projiciraju boje. CRT monitori i TV prijemnici imaju tri fosforna elementa R, G i B (slika 8.11) koji na elektronsku uzbudu reagiraju elektromagnetskim zračenjem u vidljivom dijelu spektra tj. reagiraju generiranjem crvene, zelene i plave boje. Dovoljno gust raspored fosfornih elemenata osigurava da te tri boje vidimo kao jednu.



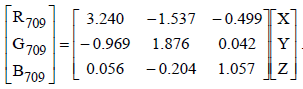
**Slika 8.11** Raspored fosfornih elemenata kod CRT monitora

Postoji više različitih setova RGB primarnih boja ovisno o definiciji triju primarnih boja i referentne bijele boje (bijele točke – *engl. white point*). Naime, RGB sustav boja prvenstveno je namijenjen upotrebi na LCD, CRT i plazma prikaznim uređajima, tipično su to monitori i TV prijemnici. Kako su se prikazni uređaji godinama razvijali i mijenjali (poboljšavali) tako se javljala i potreba za novim setovima RGB boja koji bi pratili promjene u CRT tehnologiji. Time smo zapravo, na temelju RGB modela boja, dobili niz različitih (ali ipak sličnih) RGB prostora boja.

Jedan primjer RGB prostora boja je preporuku ITU-R BT.709 koja definira prostor boja za potrebe HDTV, a odgovara i suvremenim monitorima. Primarne boje (R, G, B) kod BT.709 definirane su kao



Odnos RGB i XYZ koordinata kod BT.709 dan je s



Često se uz BT.709 primarne boje koristi bijela točka D6516. Time se, uz neke dodatne karakteristike, dobije sRGB – još jedan popularni RGB prostor boja. Bijela točka (W) je kod D65 definirana kao

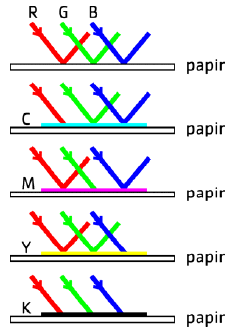


**CMY i CMYK modeli boja**

Primarne boje kod CMY modela boja su cijan, magenta i žuta (*engl. cyan, magenta, yellow*). To su sekundarne boje RGB bojama koje nastaju miješanjem crvene, zelene i plave. Na slici 8.9 prikazana je CMY kocka boja, a na slici 8.10 miješanje primarnih CMY boja. Miješanjem primarnih CMY boja dobiju se sekundarne boje koje su jednake primarnim RGB bojama (dakle crvena, zelena i plava).

CMY model se koristi u tiskarskoj industriji kod uređaja za printanje/tiskanje gdje kolor pigmenti na papiru apsorbiraju pojedine boje. Ako na papir isprintamo toner/tintu žute boje to ne znači da ćemo vidjeti žutu boju, nego da nećemo vidjeti plavu. Stoga se CMY nazivaju subtraktivne boje. Subtraktivnost CMY (CMYK) modela boja ilustrirana je na slici 8.12.

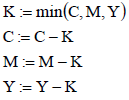
Teoretski, ako prilikom printanja koristimo 100% cijana, magente i žute boje, dobit ćemo crnu boju (cjelokupna svjetlost tj. sve boje se apsorbiraju). Međutim, u praksi se CMY boje ne mogu koristiti same. Zbog nesavršenosti tinte i drugih ograničenja u tiskarskom procesu potpuna apsorpcija svjetla nije moguća, stoga se puna crna odnosno siva boja u praksi ne mogu dobiti miješanjem CMY tinte u jednakim omjerima. Stvarni rezultat je manje ili više tamno smeđa boja. Kako bi se mogla dobiti prava crna (i siva) boja, printeri imaju i dodatnu, crnu tintu, označenu s K. Drugi razlog uvođenja crne tinte je cijena. Naime, tinta u boji je skuplja od crne tinte (crna tinta je u osnovi ugljen). I treći razlog je što je printanje samo jedne tinte puno brže od printanja tri različite tinte što pogotovo dolazi do izražaja jer većina isprintanog materijala treba samo crnu boju.



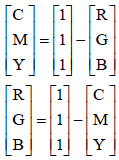
**Slika 8.12** Subtraktivnost CMYK modela boja

Tako se CMY model boja u praksi koristi kao CMYK model i nezamjenjiv u tiskarskoj industriji. Bilo koji model boja koristili, kada želimo nešto isprintati, uvijek će se u konačnici taj model konvertirati u CMYK model. Slično je i kod monitora, koji god model koristili, kad nešto želimo prikazati na zaslonu monitora, uvijek će se napraviti pretvorba u RGB model boja.

Konverzija CMY modela boja u CMYK dana je slijedećim izrazima

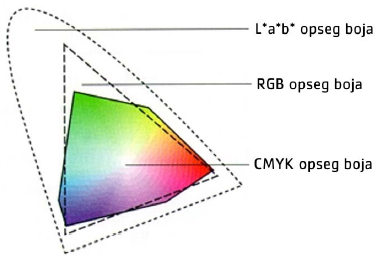


Konverzija između CMY i RGB modela dana je izrazima



**Usporedba opsega boja za L\*a\*b\*, RGB i CMYK modele boja**

Opseg boja (*engl. color gamut*) je skup svih boja koje se, za dani model boja, mogu dobiti koristeći primarne boje. Na slici 8.13 dana je usporedba opsega boja za L\*a\*b\*, RGB i CMYK modele boja. L\*a\*b\* opseg boja je najveći i on obuhvaća sve moguće boje, dakle sve vidljive valne duljine i njihove kombinacije. RGB opseg ja manji što znači da neke vidljive boje ne možemo dobiti dodavanjem RGB komponenti. Npr. čista cijan ili čista žuta boja ne mogu se prikazati na zaslonu monitora. CMYK opseg je najmanji (ali obuhvaća i neke boje koje RGB ne obuhvaća). To znači da neke od boja koje možemo vidjeti pa čak i neke od boja koje možemo prikazati na zaslonu monitora, ne možemo isprintati na printeru. To je problem s kojim se susreće tiskarska industrija i općenito problem koji se javlja prilikom bilo kakvog printanja u boji o čemu treba voditi računa.



**Slika 8.13** Usporedba opsega boja za L\*a\*b\*, RGB i CMYK modele boja

Stvarni RGB opseg boja ovisi o kvaliteti zaslona (fosfornih elementa, uzbude i sl.). Isto tako stvarni CMYK opseg boja ovisi o kvaliteti tinte i papira te o primijenjenoj tehnologiji pa svi monitori (kao i svi printeri) nemaju isti opseg boja bez obzira što svi koriste isti, osnovni model boja (CMYK odnosno RGB). Naravno, cilj proizvođača, a želja kupaca monitora i printera, uvijek je što veći opseg boja.

1. **Modeli boja za video signal**

Modeli boja za video signal temelje se na osvjetljenju i razlici boja. Dakle, jedna komponenta je uvijek osvjetljenje dok je informacija o boji sadržana u druge dvije komponente, tzv. razlike boja. Komponenta osvjetljenja predstavlja crno/bijeli dio signala i odgovara CIE Y primarnoj boji, ali joj nije u potpunosti jednaka. Naime CIE Y je suma linearnih RGB komponenti dok je u praktičnim primjenama, kod modela boja namijenjenih video signalu, Y komponenta uvijek suma nelinearnih, gama korigiranih, RGB komponenti (vidjeti poglavlje Gama korekcija). Ta se razlika ponekad javlja u imenu i oznaci pa se Y komponenta kod CIE XYZ sustava zove osvjetljenje (*engl. luminance*), a gama korigirano osvjetljenje luma (*engl. luma*) i označava se s Y'. Crtica (') označava da se radi o gama korigiranim vrijednostima te se osim Y' koristi i R', G', B' i sl.

Dakle, Y' sadrži informacije o osvjetljenju pojedinih boja, te je uz to i gama korigirana. Osvjetljenje je vezanu uz funkciju luminancijske efikasnosti. Tako npr., uz isti intenzitet u cijelom spektru percepcijski vidimo zelenu boju puno svjetliju nego crvenu, a crvenu svjetliju nego plavu. To je iskazano u odgovarajućim težinskim faktorima (preporuka ITU-R BT.601-2 definira 0.587 za zelenu, 0.299 za crvenu i 0.114 za plavu) koji se koriste prilikom računanja Y'. To su ujedno i težinski faktori koje bi trebali koristiti kada pretvaramo sliku u boji u sivu sliku kako bi osjećaj sjajnosti bio isti za obje slike. Ako crvena, zelena i plava komponenta nisu gama korigirane onda su težinski faktori (preporuka ITU-R BT.709) jednaki 0.715 za zelenu, 0.213 za crvenu i 0.072 za plavu boju.

Informacija o boji, sadržana u druge dvije komponente dobije se izbacivanjem informacije o osvjetljenju što je najlakše napraviti oduzimanjem. Kako osvjetljenje već sadrži veliki udio zelene boje (58.7%) to se druge dvije komponente dobiju oduzimanjem lume od nelinearne plave (B'-Y'), odnosno oduzimanjem lume od nelinearne crvene (R'-Y'). Te dvije komponente definiraju boju pa se poput nijanse i zasićenja zajedno zovu kroma.

Različiti faktori skaliranja se primjenjuju na komponente razlike boja za različite aplikacije. Y'UV skaliranje (model boja) optimizirano je s obzirom na kompozitni NTSC/SECAM/PAL TV signal. Y'IQ se opcionalno koristi za kompozitni NTSC TV signal, dakle ima istu namjenu kao Y'UV. Y'PbPr skaliranje (model boja) optimizirano je za komponentni analogni video. Y'CbCr koristi se za komponentni digitalni video, npr. kod MPEG-a. Kodakov PhotoYCC koristi faktor skaliranja optimiziran tako da opseg boja odgovara filmu u boji.

Iako su komponente razlike boja (UV, IQ, PbPr, ...) nelinearne, dakle gama korigirane, obično se kod njih ne koristi crtica kako bi se istakla nelinearnost. Naime te komponente su uvijek nelinearne pa se ta činjenica uglavnom posebno niti ne ističe (za razliku od Y', R', G', B' komponenti koje mogu biti i linearne i nelinearne).

**Y'UV model boja**

Y'UV model boja inicijalno je napravljen za kompozitni PAL TV kolor signal ali se sada koristi i za kompozitni SECAM i NTSC TV signal. Informacija o boji sadržana je u komponentama razlike boja koje se ovdje nazivaju U i V. Na slici 8.14 prikazano je rastavljanje slike u boji na Y', U i V komponentu. Y' sadrži informaciju o osvjetljenju, pa bi, isprintane na crno-bijelom printeru, originalna slika i slika komponente Y' trebale biti iste.

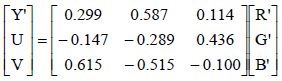
Informacija o osvjetljenju (luma) dobije se, kao i za sve druge modele, na temelju izraza

****

Kod PAL sustava za razlike boja koristi se slijedeće skaliranje

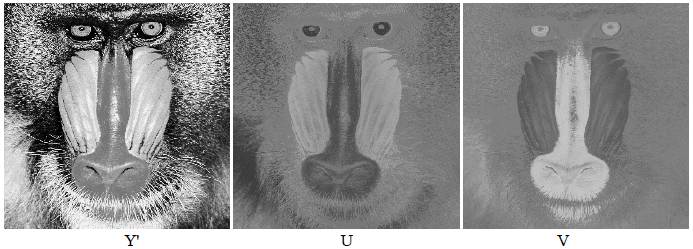
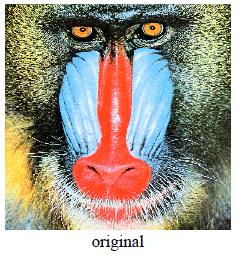
****

To znači da je odnos između R'G'B' i Y'UV modela boja dan s

****

Za sivu sliku vrijedi U=V=0, tj. nema kroma komponenti - postoji samo Y' komponenta. Y'UV model omogućava gledanje slike i na crno/bijelom TV prijemniku koji u tom slučaju koristi samo Y' komponentu, a U i V komponente jednostavno zanemari.

Ljudski vizualni sustav je najosjetljiviji na Y' komponentu. Stoga je u PAL sustavu od ukupno raspoloživog frekvencijskog pojasa 5 MHz alocirano za Y', 1.3 MHz za U i 1.3 MHz za V komponentu. Time je video signal praktično kodiran (komprimiran) koristeći se činjenicom da čovjekov vizualni sustav daje nekim informacijama veći prioritet. To nije moguće napraviti s RGB modelom boja jer je kod RGB modela boja svaka komponenta percepcijski jednako važna.

****

**Slika 8.14** Rastavljanje slike u boji na Y', U i V komponente

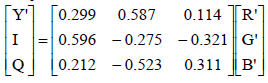
**Y'IQ model boja**

Y'IQ model boja se opcionalno upotrebljava u NTSC televizijskom sustavu. Y' komponenta je ista kao i kod Y'UV modela boja, a umjesto U i V komponenti koriste se I i Q (I - In-phase; Q - quadraturephase). I i Q su rotirane osi U i V (za 33°) te vrijedi

I = Vcos33 – Usin33 tj. I = 0.596R' – 0.275G' – 0.321B'

Q = Vsin33 + Ucos33 Q = 0.212R' – 0.523G' + 0.311B'

Dakle transformacija R'G'B' u Y'IQ je dana s

****

Slično kao i kod Y'IQ modela boja, ljudski vizualni sustav je najosjetljiviji na Y' komponentu. Potom je najosjetljivije na I i na kraju na Q. Stoga je u NTSC sustavu alocirano 4.2 MHz za Y', 1.5 MHz za I i 0.55 MHz za Q komponentu.

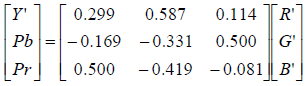
**Y'PbPr model boja**

Y'PbPr je model boja koji ima istu Y' komponentu kao i Y'UV model. Druge dvije komponente (razlike boja) su drugačije skalirane te se nazivaju Pb i Pr pri čemu vrijedi

Pb = (B'-Y')0.564

Pr = (R'-Y')0.713

odnosno



Skaliranje rezultira chrominance komponentama (Pb i Pr) koje su uvijek između -0.5 i 0.5 za R'G'B' u intervalu od 0 do 1. Y' komponenta je uvijek u intervalu od 0 do 1.

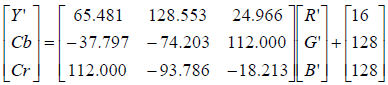
**Y'CbCr model boja**

Za aplikacije kod kojih se koristi 8-bitno kodiranje boja (256 mogućih razina – od 0 do 255) često se koristi modifikacija Y'UV modela boja sa skaliranim i pomaknutim komponentama razlike boja koje se nazivaju Cb i Cr. Pri tome se definira Y' komponenta tako da može poprimiti jednu od 220 razina (od 16 do 235), a Cb i Cr komponente tako da mogu poprimiti jednu od 224 razine (od 16 do 240, gdje je nula jednaka 128). Y'CbCr komponente moraju se skalirati i pomaknuti kako bi zadovoljile tražene uvjete pa za nove Y'CbCr komponente vrijedi  
 Y’ := 219 Y’ + 16

Cb = 224 [0.564(B’-Y’)]+128 = 126(B’-Y’)+128

Cr = 224 [0.713(R’-Y’)]+128 = 160(R’-Y’)+128

odnosno



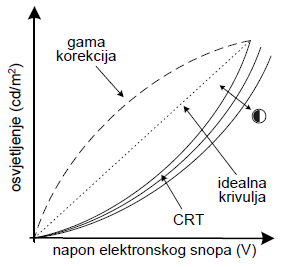
gdje su R'G'B' komponente iz intervala [0,1]. Time se dobije vektor od tri decimalna broja (Y', Cb i Cr) koja je još potrebno zaokružiti.

YCbCr model boja upotrebljava se kod većine standarda za komprimiranje video i slikovnih signala (JPEG, MPEG, H.261, H.263, H.264, itd).

**Gama korekcija**

Fosforni elementi na zaslonima monitora i TV uređaja svijetle kada ih pogodi elektronski snop. Međutim CRT nije linearan uređaj pa osvjetljenje koje uređaj generira nije proporcionalno s naponom elektronskog snopa (slika 8.15). Primjerice, ako na standardnom CRT monitoru (gama 2.2) ulaz smanjimo na polovicu (RGB vrijednosti smanjimo na 50%) rezultat će biti intenzitet osvjetljenja od samo 22% (umjesto 50%).

Da bi neki model boja mogli koristiti za uređaje koji se temelje na CRT-u potrebno je izvršiti tzv. gama korekciju. Uloga gama korekcije je da kompenzira nelinearnost CRT17 kako bi se korektno reproduciralo osvjetljenje što je ilustrirano na slici 8.15. Gama korekcija se obično radi na uređajima koji stvaraju signal (npr. kamere) kako bi se minimizirala cijena prikaznih uređaja (TV odnosno monitor).



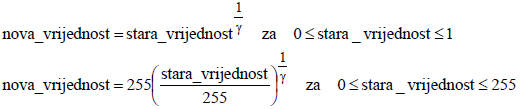
**Slika 8.15** Prikazna karakteristika CRT monitora i gama korekcija

Kod CRT vrijedi da je generirano osvjetljenje proporcionalno s naponom signala na neku potenciju tj.



gdje se eksponent (γ) zove gama faktor. Većina monitora ima gama od 1.7 do 2.7, a noviji uglavnom od 2.35 do 2.55. Tipična vrijednost gama faktora koja se korist kod modela boja za PAL/SECAM sustav je 2.8, a kod modela boja za NTSC sustav 2.2. Nažalost, implementacija gama korekcije nije jednoznačna na svim platformama.

Gama korekcija se sastoji od primjene inverznog izraza izrazu (8.2) i to na samu sliku prije njenog prikazivanja. Dakle, gama korekcija prilagodi (promijeni) vrijednosti piksela, a time i napon, kako bi omogućila percepcijska linearnost. To znači da pikseli mijenjaju svoju vrijednost prema slijedećem izrazu



Gama korekcija se ne smije pomiješati s posvjetljivanjem/potamnjivanjem što je funkcija koju imaju praktično svi programi za obradu slike a čiji je cilj posvijetliti/potamniti sve piksele za istu vrijednost.

1. **Softverski orijentirani modeli boja**

Softverski orijentirani modeli boja koriste se u raznim grafičkim programima kako bi korisnici lakše definirali boje. Nastali su u vrijeme kada su se boje morale numerički definirati što je bilo jako teško. Danas je manja potreba za ovakvim modelima jer se boje mogu birati vizualno. Postoji više softverski orijentiranih modela boja kao što su:

HSB – Hue, Saturation, Brightness (nijansa, zasićenje, sjajnost)

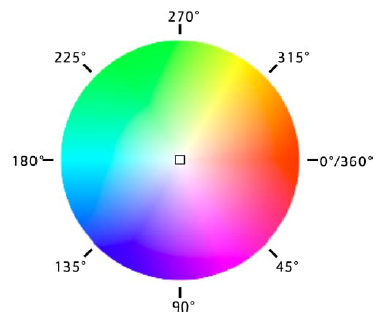
HLS - Hue, Lightness, Saturation (nijansa, svjetloća, zasićenje)

HSV – Hue, Saturation, Value (nijansa, zasićenje, iznos)

HSI – Hue, Saturation, Intensity (nijansa, zasićenje, intenzitet)

Softverski orijentirani modeli boja su intuitivni modeli jer su bliski ljudskom poimanju boja. Temelje se na definiranju nijanse i zasićenja, te treće veličine koja se odnosi na količinu emitiranog svjetla, a različito se naziva i definira u različitim modelima (svjetloća, sjajnost, iznos, intenzitet).

Nijansa definira samu boju, a kod monokromatskog svjetla odgovara valnoj duljini. Nijansa se tipično definira da bude u intervalu [0°, 360°] te opisuje kutnu razliku u odnosu na crvenu boju (slika 8.16).

****

**Slika 8.16** Nijansa kod softverski orijentiranih modela boja

Zasićenje opisuje čistoću boje a izražava se u odnosu na neutralnu sivu. Mijenja se od 0% (nema zasićenja, tj. siva boja) do 100% (puno zasićenje) i ovisi o danoj nijansi i danoj svjetloći (slika 8.17).

****

**Slika 8.17** Zasićenje kod softverski orijentiranih modela boja

Svjetloća, sjajnost, intenzitet, iznos i sl. odnose se na količinu emitiranog svjetla i uglavnom su bliski CIE definiciji svjetloće (slika 8.18).

****

**Slika 8.18** Svjetloća kod softverski orijentiranih modela boja

Napomenimo da izrazi koji se koriste u nazivima softverski orijentiranih modela boja (naročito za određivanje količine emitiranog svjetla kao što je npr. sjajnost i intenzitet), nisu uvijek u skladu s odgovarajućim CIE definicijama.