

A MULTIMÉDIA TECHNOLÓGIÁK ALAPJAI

Előadásvázlat

Mócsai Tamás

BME Hálózati Rendszerek és Szolgáltatások
Tanszék
2015.



- Az emberi látás pszichofizikai alapjai és legfontosabb jellemzői
- ITU-601 (SD), és ITU-709 (HD) szintér, színkülönbségi jel komponensek, Video tartalom kvantálása
- SD és HD raszter formátumok, mintavételi frekvenciák, UHDTV ajánlások, elvárások, változások
- Jeltömörítési alapok: prediktív kódolás (térbeli), lineáris predikció, transzformációs kódolás
- JPEG (DCT transzformáció), JPEG2000 (Wavelet transzformáció)
- Mozgókép tömörítés: mozgásbecslés és mozgáskompenzált predikció, keresési módszerek, és célfüggvények, MPEG-2,
- H-264/MPEG-4 AVC: különbségek az MPEG-2 höz képest, HEVC: fontosabb változások a H264/AVC-hez képest



Mit akarunk, és miért ?

- Mozgóképfarmátumok összetevői
 - képméret (felbontás, raszter)
 - képváltási (frissítési) frekvencia
 - színes képpont ábrázolásának módja (színtér, színterkepek)
- A két fő feladat:
 - Olyan képfarmátum létrehozása, amellyel biztosítható, hogy a néző számára nyújtott képminőség az eredetivel megkülönböztethetetlen legyen
 - Ehhez olyan forráskódolás, amely csak "észrevehetetlen" hibákat eredményez a dekódolt képen (**kompromisszum**)
- **Az emberi látás** számára kell "hibátlan" látványt biztosítani: ehhez ismerni kell az emberi látás legfontosabb tulajdonságait



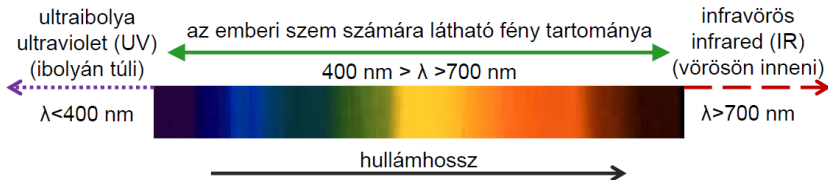
A látható fény és a színek

- Az emberi látórendszer (HVS: Human Visual System) az elektromágneses spektrumnak csak egy egészen keskeny részét képes érzékelni:
 - az ultraibolya (ibolyán túli) és az infravörös (vörösön „innen”) sugárzás közötti kb. 400 és 700 nm közé eső hullámhossz tartományt.
- A különböző hullámhosszúságú fénysugarakat a HVS különböző színekként érzékeli.
- Több fősínt különböztetünk meg:
 - **sötétkék, enciánkék, zöld, sárga, narancs, vörös, bíbor.**
 - A fehér (és a szürke) fény ezeknek a színeknek a keveréke.



A látható fény és a színek

- a legrövidebb hullámhosszú fény a kék,
- a legnagyobb hullámhosszú a vörös fény.



Emberi látás különböző embereknél

Cél: egységes (az átlagos emberi szemet) leíró rendszer meghatározása

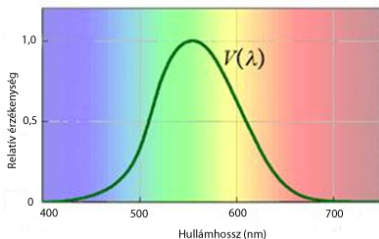
- A HVS (Human Visual System) felépítése: szinte teljesen azonos.
- A HVS működési elve: teljesen azonos.
- A látásérzékelés: egyénenként kis mértékben eltérő.
- Így minden ember kis mértékben eltérő színben látja a világot.
- A képátviteli eszközök tervezését alapvetően meghatározó HVS tulajdonság a **láthatósági függvény**ben fejeződik ki.



Láthatósági függvény

Értelmezése:

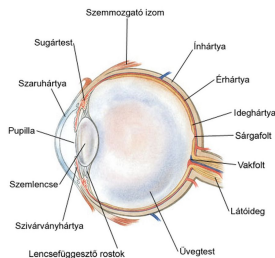
- **Láthatósági függvény** az azonos intenzitású, de eltérő hullámhosszú fény hatására a szemben keletkezett fényérzet a láthatósági függvény szerint változik.
- A láthatósági függvény nagyon sok ember látásának átlagos érzékenységét tükrözi.
- A görbe maximuma kb. 550 nm hullámhossznál van.





Szem felépítése

- A látószerv gyűjtőnév, beleértendő a szem, a látóidegek és az agy azon részei, melyek a fényingert ingerületté alakítják át.
- A szemgolyó igen bonyolult felépítésű, itt most csak a fényérzékeléssel kapcsolatos alkotóelemeivel foglalkozunk.
- A pupillán áthaladó fénysugarakat a szemlencse gyűjti össze és fókuszálja a retinára (ideghártyára), ahol fordított állású, kicsinyített, valódi kép keletkezik.

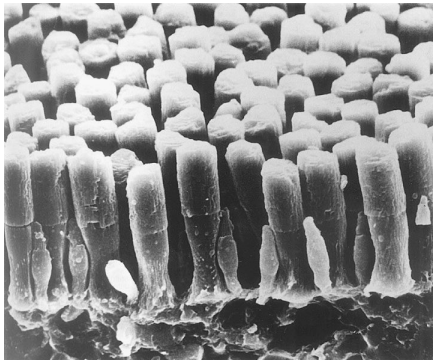




Csapok és pálcikák szerepe

Nappali és éjszakai látás

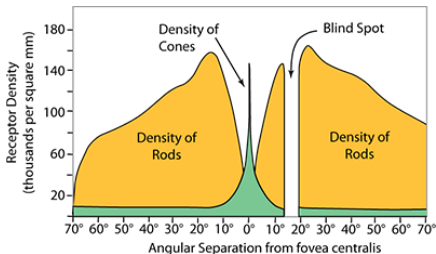
- Nappali látás (photopic):
csapokkal, ezek 3 típusa (L,M,S) biztosítja a színérzékelést
- Éjszakai látás (scotopic):
pálcikákkal - a nappali látás láthatósági függvényének maximuma eltolódik a kék árnyalatok irányába



Csapok és pálcikák szerepe

Csapok és pálcikák elhelyezkedése

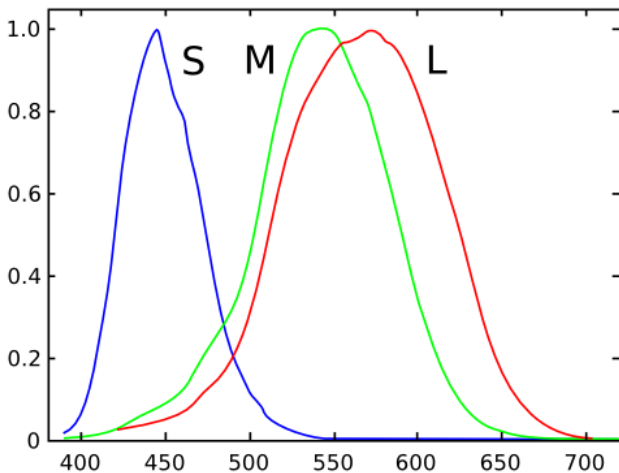
- Csapok: A retina közepén nagy sűrűségben
- Pálcikák: A retina szélén - a periférikus látásért is felelnek, érzékenyebbek, gyorsabbak, ezért a fúziós frekvencia a perifériás látás esetében magasabb





Csapok és pálcikák szerepe

L,M,S csapok érzékenysége





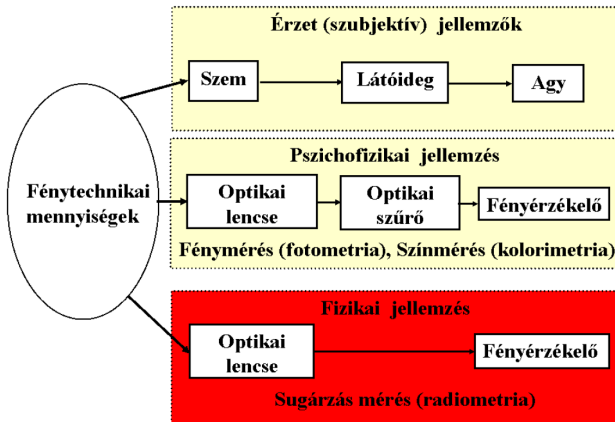
Világosság adaptáció

A látás fénysűrűség tartománya

- Teljes fénysűrűség tartomány átfogás (éjszaka sötétje - direkt napfény): kb. 10 000 000 : 1
- Ezen belül adaptáció:
 - A pupilla zárásával (gyors)
 - A látóreceptorok fotokémiai adaptálódásával (lassú)
- Egy adott megvilágítási körülményre adaptált látórendszer statikus fénysűrűség érzékelési tartománya kb 100:1 (egy adott érzékelt jeleneten belüli szemmozgás során a pupilla összehúzódása/kitágulása kb. 5x-ére növelheti ezt az értéket)
- Egy adott, nappali megvilágítási körülményre az adaptált állapotban tehát egy képen belüli maximális diffúz fehér fénysűrűséghez képesti 1% fénysűrűség még érzékelhetőnek tekinthető (100:1 tartomány), 1% fénysűrűség alatt pedig már nem érzékelhető.



Érzeti és pszichofizikai jellemzők





Érzeti és pszichofizikai jellemzők

Érzeti jellemzők

- Érzékszerveinkkel érzékeljük, a "mérőeszköz" a HVS
- Tisztán szubjektív mennyiségek

Pszichofizikai jellemzők

- Optoelektromos érzékelőkkel mérjük
- A fényforrás és az érzékelő közötti fényátviteli (optoelektromos átviteli) függvény megfelel a láthatósági függvénynek

Fizikai jellemzők

- Optoelektromos érzékelőkkel mérjük
- Tisztán radiometriai mennyiségek, függetlenül az emberi látás tulajdonságaitól

Radiometria

- Optoelektromos érzékelőkkel mérjük
- Az optikai sugárzás mérésével foglalkozik
- Optikai sugárzás: Az 1 nm hullámhosszú röntgensugárzás és a kb. 1 mm hullámhosszú rádióhullámok közé eső spektrum

Radiometriai mennyiségek

- Sugárzott teljesítmény (flux) [W]
- Sugárerősség [Wsr^{-1}] - egységnyi térszögbe (1 szteradián) sugárzott fényteliesség
- Besugárzott felületi teljesítmény [Wm^{-2}] - egységnyi felületre eső fényteliesség
- Sugársűrűség [$Wsr^{-1}m^{-2}$] - egységnyi térszögbe kisugárzott, egységnyi felületre eső fényteliesség

Radiometriai mérések

- Radiometriai mennyiségek mérése: radiométer
- Egységnyi hullámhossztartományba eső radiometriai mennyiségek mérése: spektroradiométer

Fotometriai mérőeszközök

- Fotometriai mennyiségek (HVS által észlelt tartomány) mérése: fotométer
- Egységnyi hullámhossztartományba eső fotometriai mennyiségek mérése: spektrofotométer



Pszichofizikai jellemzők

Fotometria

- A látható (kb 400-700 nm) hullámhossztartomány mérésével foglalkozik
- Az optoelektromos átalakító átviteli függvénye megfelel a láthatósági függvénynek

Radiometriai és fotometriai mennyiségek kapcsolata

- Fotometriai mennyiségek: Fényáram (luminous flux), fényerősség (luminous intensity), megvilágítás (illuminance), **fényssűrűség (luminance)**

Radiometriai mennyiség	Mértékegység	Fotometriai mennyiség	Mértékegység
Sugárzott teljesítmény (Φ_e)	W	Fényáram (Φ_v)	lm (lumen)
Sugárerősség (I_e)	W_{sr}^{-1}	Fényerősség (I_v)	cd (candela)
Besugárzott felületi teljesítmény (E_e)	W_{m}^{-2}	Megvilágítás (E_v)	lx (lux)
Sugársűrűség (L_e)	$W_{sr}^{-1}m^{-2}$	Fényssűrűség (L_v)	cdm^{-2} (nit)

Videotechnikai szempontból a legfontosabb fotometriai jellemző a **fényssűrűség (luminance)**.



Érzeti jellemzők

Világosság (lightness/brightness)

Brightness: a látásérzet azon jellemzője, amelyet az vált ki, hogy egy felület több, vagy kevesebb fényt bocsát ki (vagy úgy tűnik, mintha több, vagy kevesebb fényt bocsátana ki) - a világosság, mint érzeti jellemző szoros kapcsolatban áll a fénysűrűséggel (lásd később) Lightness: Azonos megvilágítási körülmények között egy adott felület világossága egy fehér felülethez képest.

Színezet (hue)

A látásérzet azon jellemzője, melynek eredménye végső soron a színek megnevezése (piros, sárga, stb.)

Színezettség (colorfulness)

Colorfulness: A látásérzet azon jellemzője, mely azt jellemzi, hogy egy adott színinger többé vagy kevésbé kromatikus (színezettel rendelkezik) [Akromatikus színinger: fekete-szürke-fehér]

Chroma

Egy felület színezettsége egy azonosan megvilágított fehér felület világosságához képest.

Telítettség (saturation)

Egy felület színezettsége a saját világosságához viszonyítva.



Érzeti jellemzők

$$\text{Chroma} = \frac{\text{Colorfulness}}{\text{Brightness(White)}}$$

$$\text{Saturation} = \frac{\text{Colorfulness}}{\text{Brightness}}$$

$$\text{Lightness} = \frac{\text{Brightness}}{\text{Brightness(White)}}$$

$$\text{Saturation} = \frac{\text{Chroma}}{\text{Lightness}}$$

$$\text{Saturation} = \frac{\text{Colorfulness}}{\text{Brightness(White)}} \cdot \frac{\text{Brightness(White)}}{\text{Brightness}}$$



Fúziós frekvencia és a villogás

- Fúziós frekvencia: színingerek változásának az a frekvenciája, amely felett a világosság, ill. az érzékelt szín változása már nem észlelhető
- Villogás: a fény-, vagy színinger gyors változása által keltett, általában kellemetlen benyomást értjük, ha a váltakozás frekvenciája néhány Hz-nél nagyobb, de a fúziós frekvencia értékét még nem éri el



Befolyásoló hatások

- Szem stabilitása, például:
 - 1 helyben áll (fix)
 - normál (szem)mozgást végez
 - 2 pont között ugrál
- Eltérő világosságra és színezetre (utóbbira rosszabb)
- A kontraszt időbeli változása [Hz]
- Függ az átlagos fényerősségtől, világosnál érzékenyebb (20-30-szoros a sötéthez képest)



A képtartalom mozgásának függvényében

- A legnagyobb az érzékenység az álló képtartalomra, vagy
- olyan lassú mozgásra, ahol a szem követni tudja az objektumot (normál mozgás);
- ha nem lehet követni a mozgást, akkor a gyorsabb mozgás rosszabb érzékenységséget okoz



Kontraszt érzékenység

Kontrasztérzékenység a környezeti fénysűrűség függvényében

- A megfigyelő látóterének nagy részét az Y_0 fénysűrűségű mező tölti ki, mely meghatározza a HVS adaptációját.
- A látótér közepén lévő két felület fénysűrűsége kissé eltérő, Y és $Y + \Delta Y$.
- A $\frac{Y}{Y_0}$ fénysűrűség tartomány széles határok közötti változtatása mellett a kísérlettel meghatározható a HVS $\frac{\Delta Y}{Y}$ kontrasztérzékenysége, vagyis, hogy a környezet fénysűrűségének (Y_0) függvényében mely legkisebb $\frac{\Delta Y}{Y}$ hányados esetében különböztethető még meg a két szomszédos felület egymástól (**kontrasztérzékenységi küszöb**).





Kontraszt érzékenység

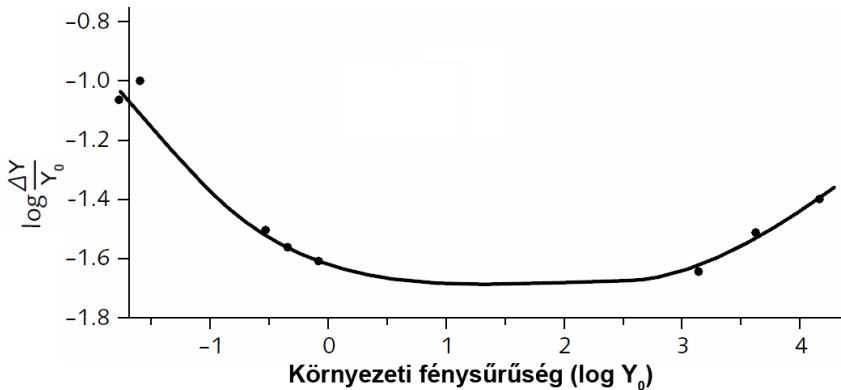
Kontrasztérzékenység a környezeti fénysűrűség függvényében

- Két dekádnyi fénysűrűség tartományban a HVS kontrasztérzékenysége 1%
- Ebben a tartományban két felület nem különböztethető meg, ha a róluk visszaverődő fény fénysűrűségeinek aránya kisebb mint 1.01
- Ez a konstans arány azt jelenti, hogy ebben a tartományban (tehát közepes megvilágítási szintek esetén) a fénysűrűség érzékelés közel logaritmikus
- A normál megvilágítási tartományokon kívül a kontrasztérzékelés romlik



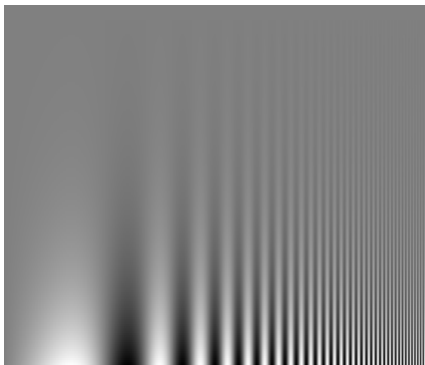
Kontraszt érzékenység

Kontrasztérzékenység a környezeti fénysűrűség függvényében



Kontraszt érzékenység térbeli frekvencia függése

Kontrasztérzékenység a térbeli frekvencia függvényében

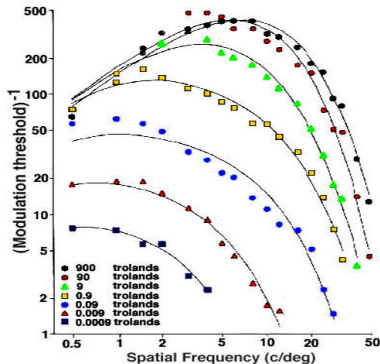


A fénysűrűség adott térbeli frekvenciával szinuszosan változik L_{min} és L_{max} között. Ekkor a kontraszt $C = \frac{L_{max} - L_{min}}{\bar{L}}$, ahol $\bar{L} = \frac{L_{max} + L_{min}}{2}$ az átlagos fénysűrűség. Az ábrán $\bar{L}$ állandó, a kontraszt felülről lefelé nő, a térbeli frekvencia balról jobbra nő.



Kontraszt érzékenység

Kontrasztérzékenység a térbeli frekvencia függvényében



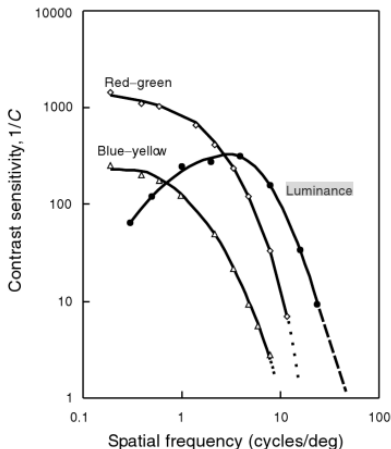
A retina átlagos megvilágítottsági szintje trolandban van kifejezve. A modulációs küszöb (modulation threshold) az a legkisebb kontraszt érték, ahol a moduláció még észlelhető. Ennek reciproka fejezi ki az érzékenységet.

A kontraszt érzékenység változik a környezeti fénysűrűség és a térbeli frekvencia függvényében. A görbesereg 60 ciklus/fok fölött gyakorlatilag érzéketlenséget mutat: ez a térbeli felbontóképesség határa.



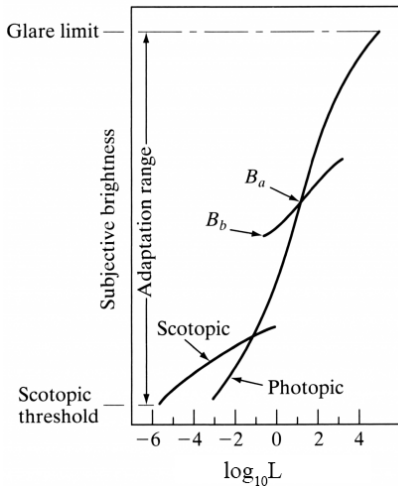
Kontraszt érzékenység

Kontrasztérzékenység a térbeli frekvencia függvényében, szín és világosságkomponensekre



A kontraszt érzékenység térbeli frekvencia határa alacsonyabb színes (piros-zöld, illetve kék-sárga modulációs modulációs tartalomra, mint a fekete-fehér modulációs tartalomra: a színekre vonatkozó térbeli felbontóképességünk alacsonyabb.

Világosság adaptáció, és szubjektív világosság



- B_a : adaptációt meghatározó közepes fénysűrűség, környezeti megvilágítás (adaptációs szint)
- B_b : adaptált állapotban legkisebb észlelhető fénysűrűség (kb. 100:1 arány !)



A szubjektív világosság közelítése a fénysűrűség függvényében

- Az előzőek alapján a világosságérzékelés a fénysűrűség függvényében logaritmikus jellemzőket mutat (logaritmikus X-tengely mentén közel lineáris függvény)
- A pontos közelítést megfelelő hatványkitevőjű hatványfüggvények adják



A CIE világosság definíciója (L^*)

A CIE világosság definíciója (L^*)

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 Y/Y_n, & Y/Y_n \leq \left(\frac{6}{29}\right)^3 \\ 116(Y/Y_n)^{1/3} - 16, & Y/Y_n > \left(\frac{6}{29}\right)^3 \end{cases}$$

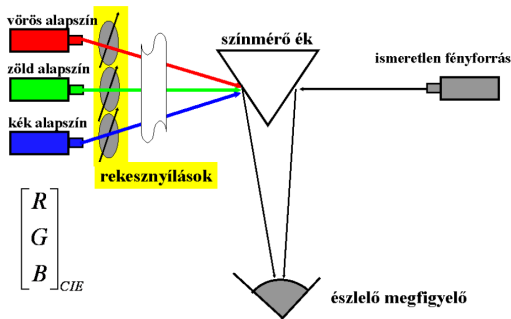
- L^* a fénysűrűség (Y) hatványfüggvénye, a fekete szint közelében (az $L^* = 8$ érték alatt) egy módosított lineáris szakasszal.
- L^* értékei 0 és 100 között vannak, egy szabadon választható Y_n referencia fehér fénysűrűséghez képest
- A $\frac{Y}{Y_n}$ hányados a **relatív fénysűrűség** (képi reprodukció természetesen során nem a rögzített jelenet eredeti fénysűrűség értékeit próbáljuk reprodukálni, hanem egy azzal közel arányos fénysűrűség eloszlást, melynek maximális fénysűrűségét a megjelenítő eszköz képességei határozzák meg
- A teljes görbe jól közelíthető egy $Y^{0.4}$ hatványfüggvénnyel



CIE RGB színmérő

CIE RGB összehasonlító színmérő kísérlet (1931)

Az ék fehér, nem tükröző felületű, melyet jobbról ismeretlen, balról ismert fény-források világítanak meg. Az éket szemből nézi az összehasonlítást végző személy, aki a három fényforrás sugárerősségét szabályozza. Egyezés: az ék éle nem látszik.





CIE RGB összehasonlító színmérő kísérlet (1931)

- A megfigyelő személy a színmérő indikátora is, hiszen az ő szubjektív megítélésére van bízva, hogy az ék két oldaláról a szemébe jutó két sugárzás egyforma, vagy különböző színérzetet kelt-e.
- Az ismert (hitelesített) fényforrások sugárerősségét változtatva kell megkeresnie azt a beállítást, amelynél az egyezés az ék két oldalán fennáll.
- Ekkor leolvasva az ismert fényforrások sugárerősségét, azokkal a pszichofizikai szín már objektíven jellemezhető.
- Az etalon fényforrások jellemzői:
 $\lambda_R = 700\text{nm}$, $\lambda_G = 546,1\text{nm}$, $\lambda_B = 435,8\text{nm}$
- Realizáció: kisnyomású higanygőzlámpa + folyadékszűrő

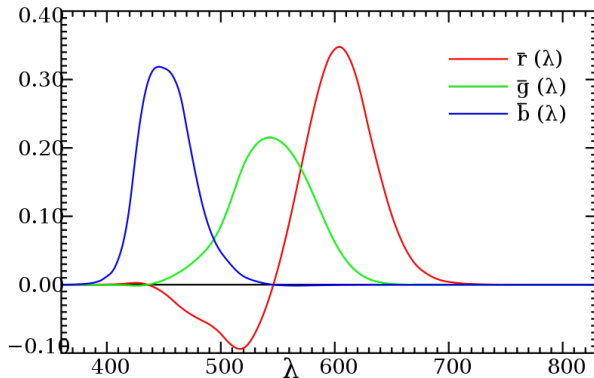


CIE RGB összehasonlító színmérő kísérlet (1931)

- Az 546,1 nm hullámhossz alatti spektrálszíneket nem lehetett kikeverni a három fényforrással.
- a kikevert szín még $R=0$ érték mellett is vörösebb volt, mint az ismeretlen spektrálszín.
- A megoldás egy módosított elrendezés ezen színekre: itt úgy lehetett egyezést elérni (az ék élét eltüntetni), hogy az R fényforrást az ék másik oldalára, az ismeretlen színhez keverték hozzá. Ilyenkor az R sugársűrűségét negatívnak vesszük.
- Megállapították azt is, hogy az egyenlő energiájú fehér (E fehér) esetén az CIE RGB fényforrások aránya :
$$L_R : L_G : L_B = 1 : 4,59 : 0,06$$



A CIE RGB színösszetevő függvények





Grassman törvénye:

- A színérzékelés 3 dimenziós lineáris tér
- Adottak az $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ spektrális színösszetevőfüggvények.
- Meghatározható pl. tetszés szerinti $\phi(\lambda)$ spektrális eloszlású folytonos spektrumú színinger három alapszín összetevője:
 - $R = \int \phi(\lambda) \bar{r}(\lambda) d\lambda$
 - $G = \int \phi(\lambda) \bar{g}(\lambda) d\lambda$
 - $B = \int \phi(\lambda) \bar{b}(\lambda) d\lambda$
- Fenáll, hogy ha $\phi_1(\lambda) \rightarrow R_1, G_1, B_1$ és $\phi_2(\lambda) \rightarrow R_2, G_2, B_2$, akkor $\phi_1(\lambda) + \phi_2(\lambda) \rightarrow R_1 + R_2, G_1 + G_2, B_1 + B_2$



Színek spektrális és érzeti azonossága

Izokrom, metamer, és heterokrom színek

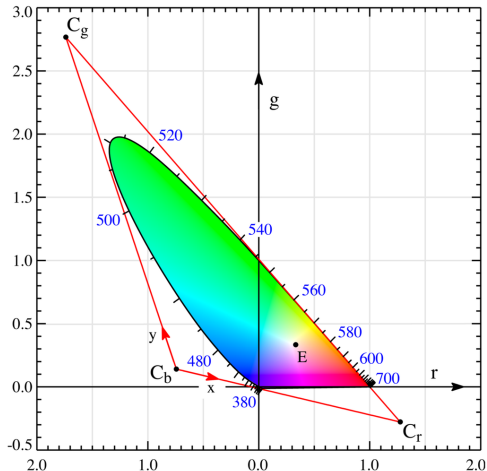
- Izokrom színek: azonos a spektrális eloszlásuk, így azonos színingert is keltenek.
- Metamer színek: bár a spektrális eloszlásuk nem azonos, ennek ellenére azonos színingert keltenek.
- Heterokrom színek: különböző színingert keltenek, így a spektrális eloszlásuk sem azonos.



CIE RGB síkbeli leképezés

CIE RGB színpatkó

- $r = \frac{R}{R+G+B}$
- $g = \frac{G}{R+G+B}$
- $b = \frac{B}{R+G+B}$
- Az RGB 3D
színtér
projekciója az
 $r + g + b = 1$
egyenletű síkra





A CIE XYZ színtér

A CIE XYZ megalkotásának követelményei

- Ne legyenek negatív értékek
- A $\bar{z}(\lambda)$ összetevő nulla értékű 650 nm fölött
- A $\bar{y}(\lambda)$ megegyezik a láthatósági függvénnyel ($V(\lambda)$) - így az Y koordináta a fénysűrűség összetevő
- Az $X = Y = Z$ szíinkordináta az egyenlő energiájú fehéret jelenti

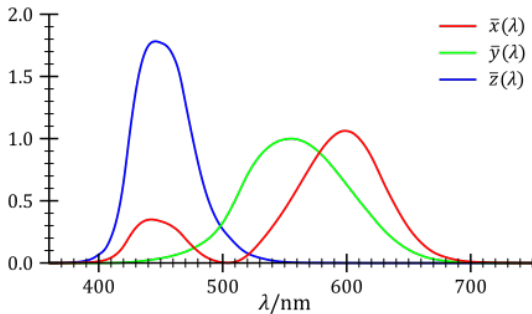
A CIE XYZ koordináták számítása a CIE RGB koordinátákból

- Lineáris transzformáció útján

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \underline{\underline{K}} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{CIE} = \begin{bmatrix} 2,77 & 1,75 & 1,13 \\ 1 & 4,59 & 0,06 \\ 0 & 0,06 & 5,6 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{CIE}$$



A CIE XYZ színösszetevő függvények



Színkoordináták számítása:

- $X = \int \phi(\lambda) \bar{x}(\lambda) d\lambda$
- $Y = \int \phi(\lambda) \bar{y}(\lambda) d\lambda$
- $Z = \int \phi(\lambda) \bar{z}(\lambda) d\lambda$

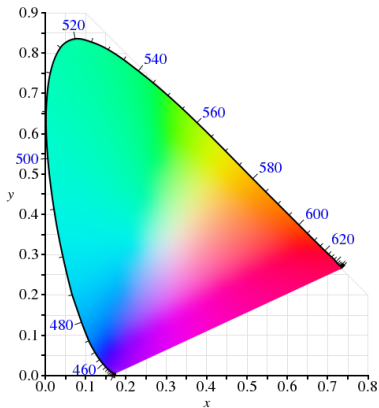


A CIE xy koordináta rendszer

A CIE xy koordináták származtatása

$$x = \frac{X}{X+Y+Z} \text{ és } y = \frac{Y}{X+Y+Z}$$

A CIE xy színháromszög



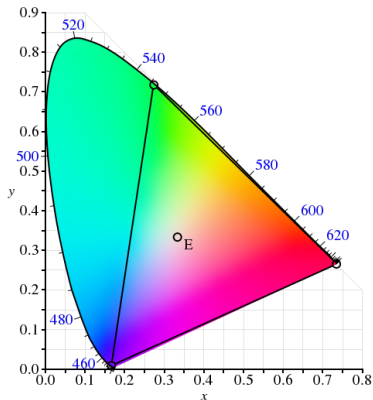


A CIE xy koordináta rendszer

A CIE xy koordináták származtatása

$$x = \frac{X}{X+Y+Z} \text{ és } y = \frac{Y}{X+Y+Z}$$

Az E fehér ($x = \frac{1}{3}, y = \frac{1}{3}$) és a CIE RGB alapszínek helye





Látható színek tartománya a CIE xy diagramon

- Látható színek tartománya: spektrálszínek és bíborszínek határolják
- Akromatikus (színtelen) tartomány: fehérnek tekinthető színek, a látható színek tartományának belsejében
- Spektrálszínek: patkó formájú görbe
- Bíborszínek: a patkó két végét összekötő szakasz



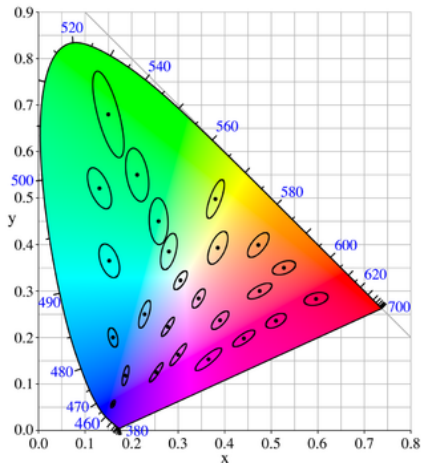
MacAdam ellipszisek, CIE uv, CIE LAB

- A CIE XYZ koordináta rendszer, és az xy színdiagram *nem uniform* a színek megkülönböztethetőségére nézve. A színdiagramon a színpontok egymástól való távolsága nem arányos a színezet eltérésének mértékével.
- A színezetváltozásnak az a minimális értéke, melyet a HVS már észrevesz, az éppen megkülönböztethető különbség (Just Noticable Difference: JND).
- Az ún. Mac-Adam ellipsziseken belül lévő színeket az átlagos emberi szem nem tudja megkülönböztetni egymástól
- A CIE 1960-ban bevezette, majd több alkalommal módosította az egyenletes színdiagramot (CIE-UCS), amelynek jelenlegi formáját 1974-ben fogadták el.
- A színérzékelést még jobban leíró színterek pl. CIE LAB (opponens színelmélet), és a CIE CAM (Color Appearance Model) színterek.

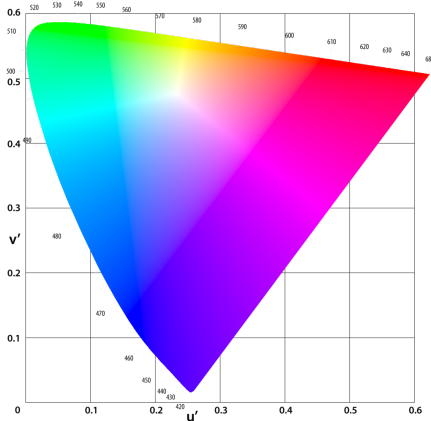


MacAdam ellipszisek, CIE Luv

MacAdam ellipszisek



CIE Luv (UCS) színdiagram





- A színlátás mechanizmusa egyenértékűnek tekinti a metamer színeket.
- A természetben előforduló színekkel azonos színérzetet keltő színingereket mesterségesen elő lehet állítani néhány alkalmasan megválasztott alapszín különböző arányú keverékével.
- Ez a a színes képreprodukáló eljárások (tvtechnika, nyomtatás, színes fotó, dia, stb.) alapja.
- A színdiagramban felvehető 4 színpont úgy, hogy a keletkező négyszög (a reprodukálható színek területe) csaknem azonos területű a látható színek területével.
- Ha figyelembe vesszük, hogy a kieső területek főleg a zöld színek tartományában vannak, ahol a szemünk színekre vonatkozó felbontóképessége igen kicsiny, jó közelítéssel 3 szín használata is elég a jó színes kép reprodukálásához.



TV technika alapszínei

- Kézenfekvő lenne a CIE három alapszínét alkalmazni a színes tv kép reprodukálásához. Gyakorlati akadályok:
- Nem lehet a képcső technikában alkalmazott fényporfajtákkal ezeket az alapszíneket létrehozni.
- A használt fénypor hatásfoka kicsi lenne ezekre a pontokra.
- Kompromisszum: egyidejűleg biztosítani a megvalósíthatóságot és a maximálisan lefedett színterületet.



TV/video RGB alapszínek

RGB színtér definiálása

Egy additív RGB színteret egyértelműen meghatároznak a választott RGB alapszínek CIE XYZ (vagy xy) koordinátái, és a fehérpontja. A színtérben kikeverhető színek területe (gamut) egy háromszög a CIE xy koordináta rendszer belsejében.

RGB színterek

Név	Fehérpont	x_R	y_R	x_G	y_G	x_B	y_B
CIE (1931) RGB	E	0.7347	0.2653	0.2738	0.7174	0.1666	0.0089
ITU-R BT.709 / sRGB	D65	0.64	0.33	0.30	0.60	0.15	0.06
NTSC (1953) / FCC 1953	C	0.67	0.33	0.21	0.71	0.14	0.08
NTSC (1987) / (SMPTE C / SMPTE 170M)	D65	0.63	0.34	0.31	0.595	0.155	0.07
PAL/SECAM (1970) / (EBU / ITU 601)	D65	0.64	0.33	0.29	0.60	0.15	0.06
Apple RGB	D65	0.625	0.34	0.28	0.595	0.155	0.07
ROMM RGB	D50	0.7347	0.2653	0.1596	0.8404	0.0366	0.0001



TV/video RGB színterek I.

RGB alapszínek

A CRT TV foszfortekhnológia fejlődésével a bevezetett RGB alapszínek (először NTSC/FCC 1953) is több változáson mentek át. Jelenleg a SMPTE C (USA), az EBU (Európa) (ITU-601) (SDTV), illetve az ITU-709 (HDTV) ajánlások szerinti alapszínek a legelterjedtebbek.

Az világosságjel (Y) az RGB alapszínekből

Az RGB alapszínek megfelelő intenzitású additív keveréséből az adott színtér választott fehérpontja előállítható.

$$Y_{601} = 0.299R + 0.587G + 0.114B$$

$$Y_{709} = 0.212R + 0.715G + 0.072B$$

Megjegyzés: Az NTSC/FCC együtthatókat az ITU-601-es ajánlás, valamint az EBU ajánlás is átvette (de új RGB alapszíneket, és fehérpontot definiáltak), így a további diákon ezen együtthatók szerepelnek, kerekítve.



A videotechnika színkülönbségi jelei

Az világosságjel (Y - fénysűrűség összetevő) az RGB alapszínekből

Az RGB alapszínek megfelelő intenzitású additív keveréséből az adott színtér választott fehérpontja előállítható.

$$Y = 0.3R + 0.6G + 0.11B$$

Az egyenlet átrendezve

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

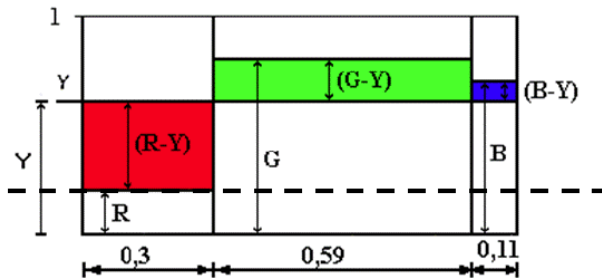


A videotechnika színskülönbségi jelei

Grafikus értelmezés

A két egyenlet együttesen egy területdiagramon ábrázolható egy adott RGB értékhármassal. Belátható, hogy az Y alatti terület egyenlő az Y feletti területtel.

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$





A videotechnika színkülönbségi jelei

Színkülönbségi jelek

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

- Az (R-Y), (G-Y) és (B-Y) értékhármasok a TV technikai színkülönbségi jelei, mert világosságinformációt nem hordoznak
- Mivel látásunk térbeli felbontása kevésbé érzékeny a képi információ "színtartalmára", célszerű a világosságinformációt és a színinformációt külön-külön kezelni
- A 3D színtérben (RGB), ha az egyik komponensnek a világosságjelet (Y) választjuk, akkor még két színkülönbségi jelet kell átvinnünk



A videotechnika színekülönbségi jelei

- Előjeles mennyiségek, a területdiagrammon az Y vízszintes vonala fölött pozitív, alatta pedig negatív a színekülönbségek előjele
- Ha két színekülönbség zérus, akkor a harmadik is nulla
- Ilyenkor $R=G=B$, azaz a fehérpontot kapjuk
- Ha nem a fehérpontban vagyunk, akkor a színekülönbségek közül legalább kettő zérustól eltérő



A videotechnika színekülönbségi jelei

$$0 = 0.3(R - Y) + 0.6(G - Y) + 0.11(B - Y)$$

- Szintén a fenti egyenletből adódóan a színekülönbségek értéktartománya:
 - $-0,7 \leq (R - Y) \leq +0,7$
 - $-0,89 \leq (B - Y) \leq +0,89$
 - $-0,41 \leq (G - Y) \leq +0,41$
- A zöld színekülönbségi jel értéktartománya a legkisebb
- Mindig a nagyobb értéktartományú jelet érdekesebb továbbítani (jel/zaj viszony!)
- Ezért a színinformáció továbbítására a vörös (R-Y), és a kék (B-Y) színekülönbségi jeleket továbbítják
- E két jelet koordinátarendszerben ábrázolhatjuk (Pl.: vektorszóp)



TV-fénysűrűség, -színezet, és -telítettség

Hogyan definiálhatjuk a fenti szubjektív mennyiségeket ?

- Az Y, vagyis a fénysűrűség összetevő értéktartománya 0 ..1 : fekete szintnél 0, a csúcs fehérenél 1
- Az RGB együtthatókat tartalmazó egyenletből adódóan ugyanez az értéktartomány érvényes az R, G és B komponensekre is



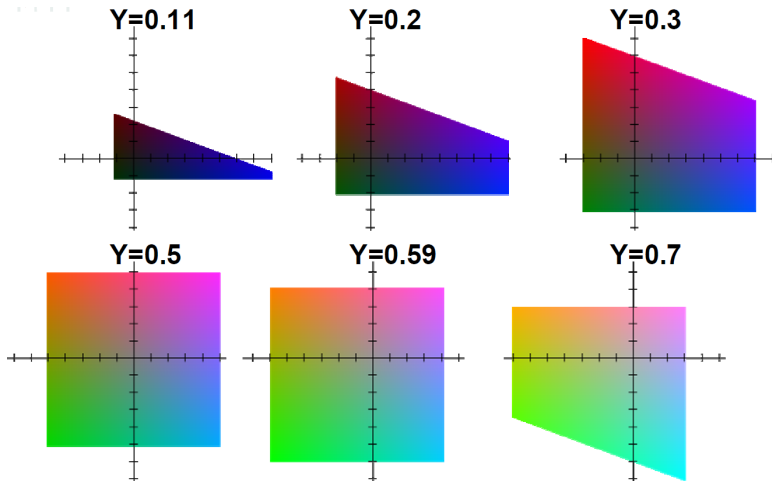
TV-fénysűrűség, -színezet, és -telítettség

Kvázi-spektrálszínek

- A kvázi-spektrálszínek a spektrál színek, illetve a bíborvonal megfelelői a televízió technikában.
- A spektrálszínek és a bíborvonal analógiájára a kvázi spektrálszínek jellemzői, hogy egyszerre nem tartalmazzák mind a három alapszínt.
- Ezért a kvázi-spektrál színek a televízió technika által visszaadható színeket megadó háromszög oldalaira vagy csúcsaira esnek



Példák B-Y,R-Y diagramokra





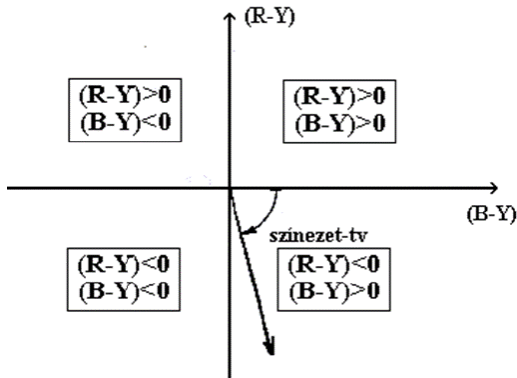
Példák B-Y,R-Y diagramokra

- Az B-Y,R-Y diagramon ábrázolhatjuk a színeket egy rögzített Y érték mellett. Más szóval minden $0 < Y \leq 1$ érték mellett rajzolható egy színes B-Y,R-Y diagram.
- Nyilván rögzített Y mellett nem biztos, hogy minden szín 100%-os telítettséggel van jelen a B-Y,R-Y diagramon. Például: teljesen telített kékre ($R,G,B=0,0,1$) $Y=0.11$, tehát pl. $Y=0.5$ értékhez tartozó B-Y,R-Y diagramon csak fehérrel hígított kék található.
- A diagramon az origóban van a fehér (szürke), vagyis ahol a színekülönbségi jelek értéke nulla.
- A diagramon az origóból kiinduló félegyenesen azok a színek vannak, amelyek egymásból kinyerhetők fehér szín hozzáadásával.
- Tehát az origóból kiinduló félegyenesen az azonos színezetű, de eltérő telítettségű színek vannak. A diagramon tehát színezetet (B-Y,R-Y) helyvektor szöge írja le. Így a színezet könnyen megadható egyetlen számmal

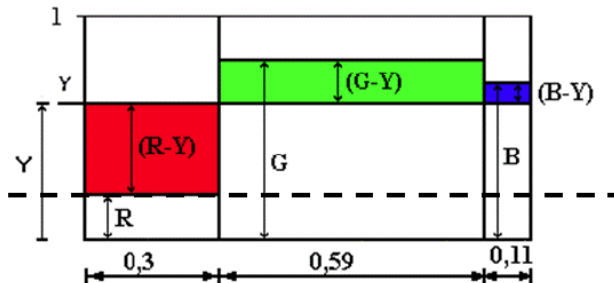
TV-fénysűrűség, -színezet, és -telítettség

- A Színezet_{TV}:

$$\arctan \frac{(R - Y)}{(B - Y)}$$



A Telítettség_{TV}



- A Telítettség_{TV} értéke legyen a fehérpontban 0, míg az RGB alapszínek által határolt háromszög élein 1
- Eszerint ha az RGB komponensek közül egy vagy kettő nulla értékű, akkor az RGB egyik élén, illetve a háromszög egyik csúcsában helyezkedik el az adott szín: ebben az esetben lehet a telítettség 1
- Ha nincs nulla értékű komponens, akkor nem lehet a telítettség 1, mert a színkoordináta elmozdult az RGB háromszög belseje felé (fehérpont felé)



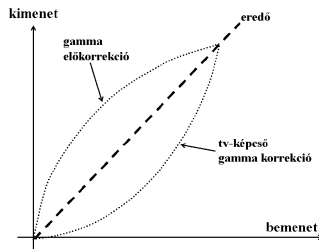
- Értelmezés a területdiagramon:
 - Keressük meg a legkisebb alapszín összetevőt (jelen esetben R)
 - Ennek magasságában húzzunk egy vízszintes egyenest, mellyel az kérdéses alapszínt kettéosztjuk
 - Így kapunk egyrészt R magasságnyi fehéret, és egy kvázi-spektrálszínt, mely nem tartalmaz R-t
 - Ezen kvázi-spektrálszín fénysűrűsége meghatározható: A fehér rész fénysűrűsége R (az előbbiek alapján), míg a maradék (kvázi-spektrálszín) fénysűrűsége az eredeti fénysűrűség (Y) mínusz a fehér komponens fénysűrűsége (jelen esetben R), ami a példában nem más, mint a vörös színkülönbségi jel (R-Y) abszolút értéke
 - ezek alapján a Telítettség_{TV} a kvázi-spektrálszín fénysűrűsége osztva az eredeti szín fénysűrűségével:
 - $$\text{Telítettség}_{TV} = \frac{|\min(R,G,B) - Y|}{Y}$$



Gamma-korrekción

Képcsövek gamma-torzítása

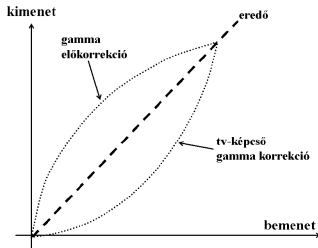
- A katódsugárcsőes (CRT) képcsövek fénysűrűsége a katódsugárcső elektronágyújának vezérlőfeszültségének nem-lineáris függvénye. $L = k U^\gamma$. Ezt a nem-lineáris átviteli karakterisztikát korrigálni kell(ett).
- A korrigálás elve az, hogy a továbbítandó jelet egy inverz hatványfüggvénnyel előtorzítják
- A teljes átviteli úton ez a gamma-korrigált jel halad végig





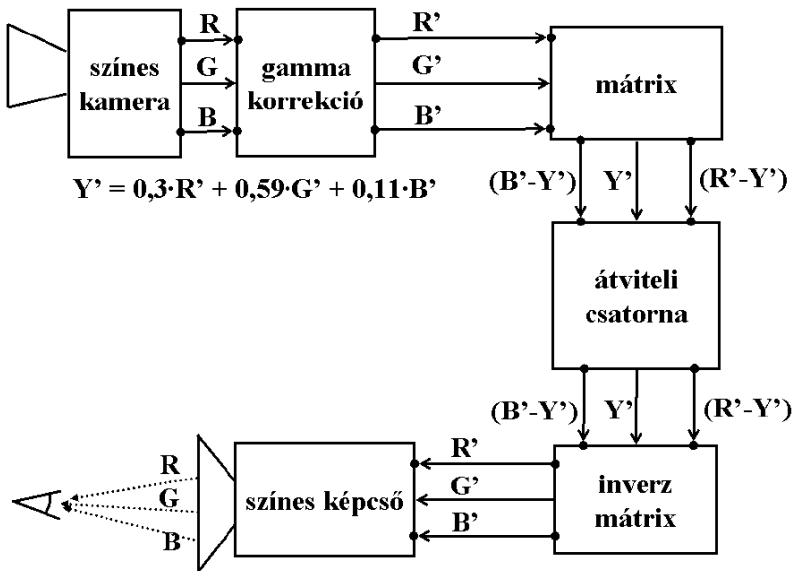
Képcsövek gamma-torzítása

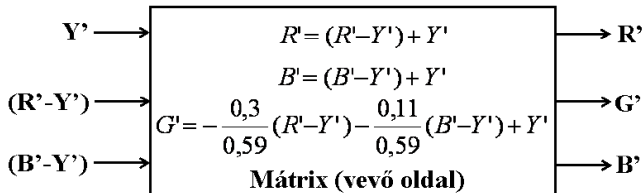
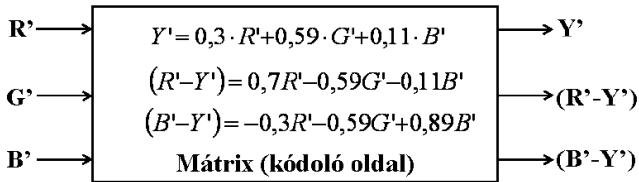
- A gyakorlatban először képezni kell az RGB alapszínek gamma-korrigált megfelelőit ($R' = R^{\frac{1}{\gamma}}$, $G' = G^{\frac{1}{\gamma}}$ és $B' = B^{\frac{1}{\gamma}}$)
- majd azokból (mátrixolással) a világosságjel (Y), és a színkülönbségi jelek gamma korrigált megfelelőit (rendre Y' , $R' - Y'$, és $B' - Y'$ - a ' a gamma-előkorrigált jelet jelenti)





Gamma-korrektció rendszertechnikája

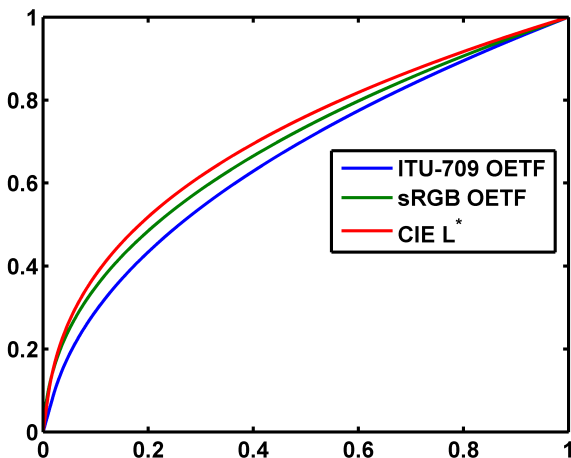






Gamma-korrekción, ismét

Nem-lineáris transzfer karakterisztikák (optoelectronic transfer functions), lineáris skálázás mellett



"Gamma-korrekción" szerepe jelenleg

Nem-lineáris transzfer karakterisztika: összefoglaló és előzetes

1. A CRT típusú megjelenítők nem-lineáris feszültség-fénysűrűség függvényének korrekciója
2. Megjelenítési körülmények kompenzációja: kontraszt növelése, mely az 1-nél nagyobb eredő hatványkitevő eredménye (részleteivel nem foglalkozunk)
3. Nem-lineáris (perceptuális) kvantálás - erről később



Videojel-komponensek

- Az Y' jel tehát a nem-lineáris transzfer karakterisztikával gamma-előkorrigált alapszínjelek súlyozott összege
- Az Y' jelet a magyar szóhasználatban általában változatlanul, világosságjelként, angolul luma-ként (nem luminance !) nevezik
- A komponens video rendszerekben 3 komponenst kell egymástól függetlenül kezelni és továbbítani/tárolni
- Kihasználva az emberi látórendszer színekre vonatkozó kisebb érzékenységét (elsősorban térbeli felbontásigény), általában nem az R' , G' és B' alapszínjeleket, hanem a teljes sáv szélességű Y' jelet, és a csökkentett sáv szélességű két színekülönbségi jelet ($R'-Y'$, és $B'-Y'$) továbbítjuk/tároljuk

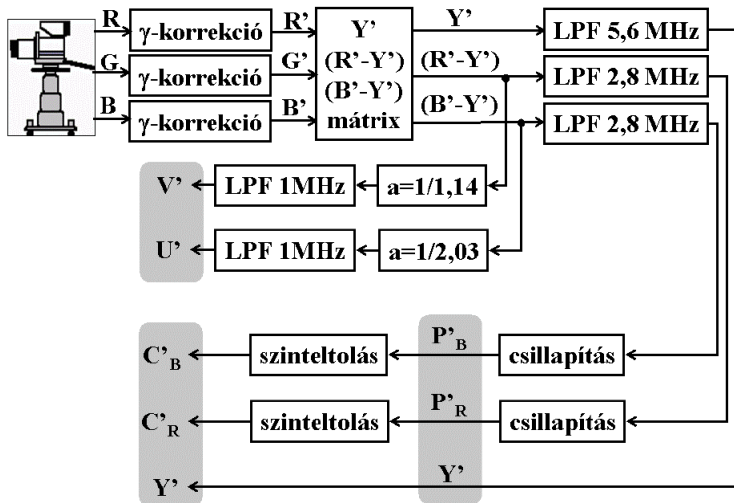


Videojel-komponens formátumok

- $Y'U'V'$: kompozit PAL és NTSC ($Y'I'Q'$) rendszerekben. Az U' és V' jelek a ($B'-Y'$) és ($R'-Y'$) jelek megfelelően skálázott, korrigált megfelelői, abból a célból, hogy a világosságjelre ültetett modulált színsegédvívó beleférjen a világosságjel dinamikatartományába. A mai analóg és digitális komponens (stúdió)-rendszerekben már nem használják, csak a hagyományos analóg műsorszórás esetén a kisugárzás céljára
- $Y' P'B P'R$: komponens analóg rendszer, a két színkülönbségi jelet előre definiált módon skálázzák/korrigálják és analóg szűrőkkel a világosságjel sáv szélességének felére korlátozzák. Az interfész konkrét megvalósításai különbözőek az EBU/SMPTE, a Sony és a Panasonic specifikációi szerint
- $Y' C'B C'R$: komponens digitális rendszer, a két színkülönbségi jelet az A/D átalakítás kivezérlési tartományához illesztés céljából skálázzák/korrigálják, majd a mintavételi struktúrának (pl. 4:2:2, 4:2:0, stb.) megfelelően alul-mintavételezik



Videojel-komponens előállítás





Videojel-komponens tartományok I.

Komponens	Összefüggés	Dinamika
$Y' =$	$0,299 \cdot R' + 0,587 \cdot G' + 0,114 \cdot B'$	$0 \dots 1$
$V' =$	$(R' - Y')/1,14 = 0,877 \cdot (R' - Y')$	$-0,62 \dots +0,62$
$U' =$	$(B' - Y')/2,03 = 0,493 \cdot (B' - Y')$	$-0,44 \dots +0,44$

Az Y' 0 - 1 értéktartománya feszültségben = 0 - 700 mV



Videojel-komponens tartományok II.

Komponens	Összefüggés	Dinamika
Y'	$0,299 \cdot R' + 0,587 \cdot G' + 0,114 \cdot B'$	0 – 1
P'_R	$(0,5/0,7) \cdot (R' - Y') = 0,713 \cdot (R' - Y')$	-0,5 ... +0,5
P'_B	$(0,5/0,89) \cdot (B' - Y') = 0,564 \cdot (B' - Y')$	-0,5 ... +0,5
C'_R	$P'_R + 0,5 = 0,713 \cdot (R' - Y') + 0,5$	0 – 1
C'_B	$P'_B + 0,5 = 0,564 \cdot (B' - Y') + 0,5$	0 – 1

Az Y' értéktartománya feszültségben: 0 - 700 mV

A P'_B/P'_R értéktartománya feszültségben: ± 350 mV

A C'_B/C'_R értéktartománya feszültségben: 0 - 700 mV



Y'CbCr digitális jelszintek (8 bit/komponens)

- 8 bites komponens ábrázolás esetén célszerű lenne a 0..255 tartományt megfeleltetni az Y' jel 0..1 tartományának - ezt szinte csak a JPEG/JFIF szabványban használják
- A lehetséges túllövések/alullovések (pl. analóg szűrők kimenetén megjelenő oszcillációk) DSP kezelésére headroom/footroom tartomány biztosított
- Az Y' jel feketeszintje: 16 (offset), fehérszintje $16+219=235$
- A Cb/Cr jelek 0 szintje: 128 (offset), max szintje $128+112=240$, min. szintje: $128-112=16$
- Így az Y' jel tartománya: $[0..219]+16$, a Cb és Cr jelek tartománya: $[-112..+112]+128$
- A 0 és a 255 szintek tiltottak, szinkronizálási célokat szolgálnak
- 10 bites interfész esetén (stúdió) megtartott a 8 bites szintjeit, és tiltott kódjait, kiegészítve 2 LSB bittel



Mintánkénti bitszám és kvantálási karakterisztika I.

Perceptuális vonatkozások (ismétlés)

- Kb. két dekádnyi fénysűrűség tartományban a HVS kontrasztérzékenysége 1%
- Ebben a tartományban két felület nem különböztethető meg, ha a róluk visszaverődő fény fénysűrűségeinek aránya kisebb mint 1.01
- Az előzőek alapján a világosságérzékelés a fénysűrűség függvényében logaritmikus jellemzőket mutat
- Az alkalmazott közelítések (pl. CIE L^*) megfelelő hatványkitevőjű hatványfüggvényeket használnak

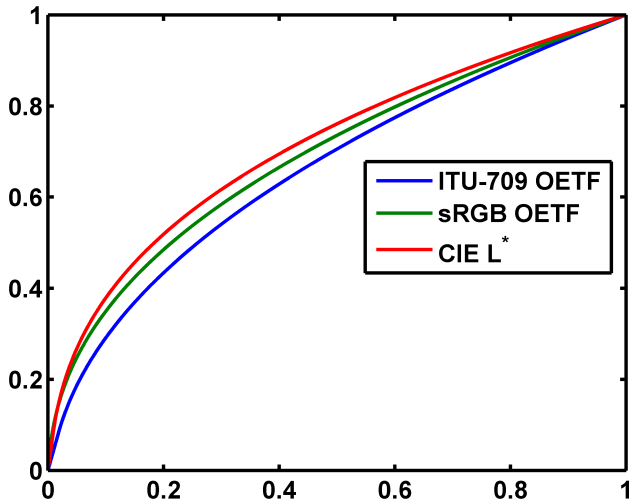
Mintánkénti bitszám és kvantálási karakterisztika I.

A CIE világosság definíciója (L^*)

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 Y/Y_n, & Y/Y_n \leq \left(\frac{6}{29}\right)^3 \\ 116(Y/Y_n)^{1/3} - 16, & Y/Y_n > \left(\frac{6}{29}\right)^3 \end{cases}$$

A teljes nem-lineáris karakterisztikát tekintve, ez leginkább egy $Y^{0.4}$ -es hatványfüggvénynek felel meg.

Mintánkénti bitszám és kvantálási karakterisztika II.



Mintánkénti bitszám és kvantálási karakterisztika III.

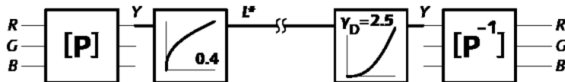
Perceptuális vonatkozások, folytatás

- Tekintsünk egy 8 bites *lineáris* kvantálást, ahol a 0 kód a fekete, a 255 a fehér szintet jelöli
- A 100-as kódhoz tartozó fénysűrűség mellett a 101-es kódhoz tartozó fénysűrűség még éppen megkülönböztethető lenne (1% eltérés)
- A 100-as kódnál kisebb értékek esetén (a feketeszinthez közel) egy lineáris kódérték eltérés jelentős világosságérzet eltérést okozna (pl: $21/20 = 105\%$), ami egy folytonos szürkeskálán pl. sávosodást eredményez a feketeszintek közelében
- A 100-as kódnál nagyobb értékek esetén (a fehérszinthez közel) egy kódérték eltérés megkülönböztethetetlen lenne, vagyis a kvantálási lépcsők kihasználtsága rossz lenne ebben a tartományban (pl. $201/200 = 0.01\%$)
- Célszerű lenne tehát a fénysűrűség/világosságérzet nem-lineáris karakterisztikát (pl. CIE L) figyelembe vevő kvantálási karakterisztikát használni

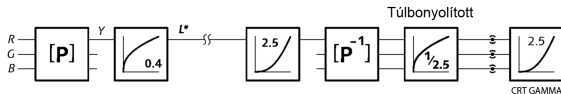
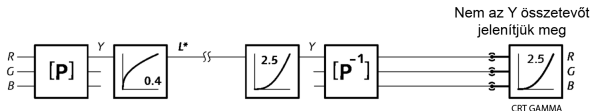
Mintánkénti bitszám és kvantálási karakterisztika IV.

Perceptuális kvantálás

- Elvi lehetőség a perceptuális kvantálásra



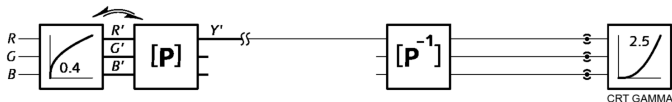
- A CRT képcsövek gamma-karakterisztikájának kompenzálása ekkor problémát jelentene



Mintánkénti bitszám és kvantálási karakterisztika V.

Perceptuális kvantálás

- Szerencsés egybeesés, hogy a HVS fényssűrűség-szubjektív világosság karakterisztikája jól közelíthető 0.4-es hatványfüggvénnyel, ami pont a gamma-karakterisztika (2.5-ös hatványfüggvény) inverze ($1/2.5=0.4$)
- A módosított rendszertechnika tehát (amiről eddig is szó volt a gamma-korrekciónak esetében)





Mintánkénti bitszám és kvantálási karakterisztika V.

Perceptuális kvantálás

- A gamma-előkorrekció jelentősége ma már elsősorban nem a CRT képcsövek karakterisztikájának kompenzálása, hanem az érzékeléshez illeszkedő kvantálás használata a digitális képi ábrázolás (mind állókép, mind mozgóképek esetében) céljára
- Nem CRT megjelenítők esetében (LCD-TFT, stb) a vezérlőfeszültség-fénysűrűség összefüggés nem is egyezik meg a CRT gammával, így a gamma-előkorrekciót egy mesterségesen előállított 2.5-es hatványfüggvénnyel (pl. LUT - Look Up Table) 'semlegesíteni' kell

Mintánkénti bitszám és kvantálási karakterisztika VI.

Perceptuális kvantálás

- A világosság érzékeléshez jobban illeszkedő kvantálás hatékonysága jobb.
- Láttuk, hogy míg a 8 bites komponensenkénti lineáris kvantálás nem elegendő
- Ha a kérdéses kontraszt tartomány 100:1 és biztosítani kell az 1%-os felbontást, $100 \times 100 = 10000$ kódszóra lenne szükség (ez 14 bites lineáris kvantálás)
- Ha a kódszavak elhelyezkedése megfelel az 1,01 relatív lépésköznek, akkor csak 463-ra lenne szükség, ami csak 9 bitet igényel ($1,01^x = 100$, ebből $x = 463$)
- A 0.4-es hatványfüggvénnyel való közelítés miatt +1 bit kell
- Kezdetben a stúdiók 8 bites nem-lineáris kvantálást alkalmaztak
- Ma általános a 10 bites nem-lineáris kvantálás



Mintánkénti bitszám és kvantálási karakterisztika VII.

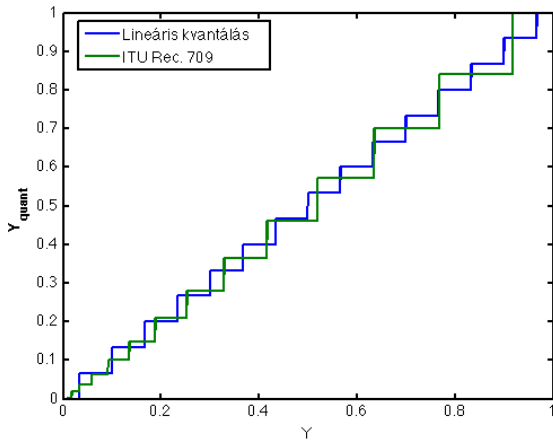
ITU-709-es nem lineáris transzfer karakterisztika

$$V' = \begin{cases} 4.500 V & V < 0.018 \\ 1.099 V^{0.45} - 0.099 & V \geq 0.018 \end{cases}$$

ahol V az R,G,B komponenseket jelöli

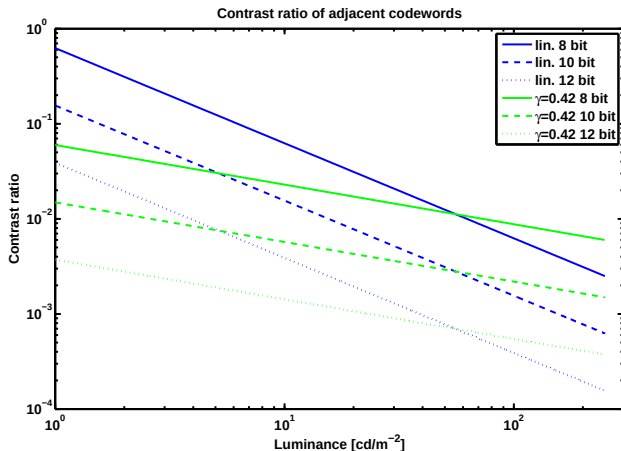
Mintánkénti bitszám és kvantálási karakterisztika VII.

ITU-709-es kvantálás példa (szemléltetés céljából "extrém" 4 bites esetre)



Szomszédos kódszavak kontrasztaránya

Lineáris, és $1/2.4=0.42$ kitevőjű kvantálási karakterisztika mellett



Szürkeskála 6 bites lineáris, és nem-lineáris kvantálási karakterisztika

Linearly quantised ramp



Non-linearly quantised ramp



Szürkeskála 8 bites lineáris, és nem-lineáris kvantálási karakterisztika

Linearly quantised ramp



Non-linearly quantised ramp





A színkülönbségi jelek alul-mintavételezése

Elvei

Az analóg (Y'UV, Y'PbPr) komponens videónál megvalósított színkülönbségi jel sávszélességcsökkentés a digitális videotartalom esetében a színkülönbségi jelek térbeli alul-mintavételezésével (decimálás) valósul meg.

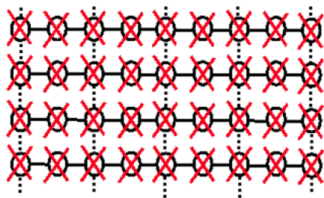
A mintastruktúra jelzése

- Pl. 4:2:0 vagy 4:2:2
 - Az első szám a világosságjel (Y') mintáinak referencia-számát adja meg (4)
 - A második szám, a színkülönbségi jelek vízszintes irányú alul-mintavételezését adja meg, a referenciához képest (4:2=2:1 alul-mintavételezés vízszintes irányban)
 - A harmadik szám vagy 0, vagy megegyezik az előzővel. Ha nulla, akkor függőleges irányban 2:1-es alul-mintavételezés van (minden második sorból történik mintavételezés). Ha megegyezik a második számjeggyel, akkor függőleges irányban nincs alul-mintavételezés

4:4:4 mintastruktúra

X : C'_R és C'_B színminták

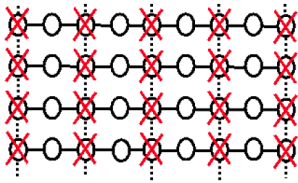
O : világosság minták



- A mintastruktúra minden tv-sorban azonos
- Alapegység: egy pixel
- 8, 10 bit/minta/komponens
- 8 bites esetben: 3 bájt = 24 bit
- 10 bites esetben: 3×10 bit = 30 bit

4:2:2 ITU-601

4:2:2 progresszív



✕ : C'_R és C'_B színminták

○ : világosság minták

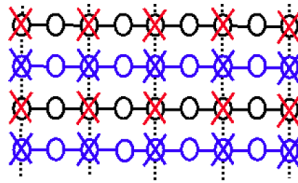
8, 10 bit/minta/komponens

8 bites: $2+1+1$ bájt = 4 bájt = 32 bit / 2 pixel

10 bites: $(2+1+1) \cdot 10$ bit = 40 bit / 2 pixel

Ha a bemenő kép formátuma 4:4:4, akkor a 4:2:2-re történő átalakítás kompressziós tényezője 3:2

4:2:2 váltott-soros



✕ : színminták 1. félkép

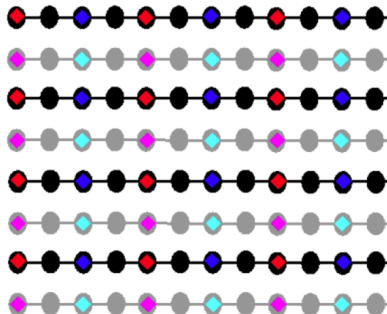
○ : világosság minták 1. félkép

✕ : színminták 2. félkép

○ : világosság minták 2. félkép

4:1:1 DV-USA

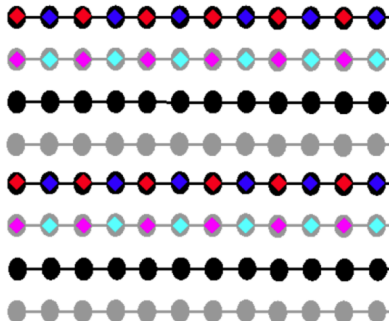
- Minden sor azonos szerkezetű
- Alapegység egy sorban 4 pixel
- A színjelek vízszintesen 4-vel alul-mintavételezettek
- A maradó színminták egybe esnek egy-egy Y' mintával
- C'_R , C'_B soron belüli offszet
- 8bit: $4+1+1=6$ bájt=48bit/4 pixel
- 10bit: $(4+1+1) \times 10=60$ bit/4 pixel
- Ha a bemenő formátum 4:2:2, akkor a kompresszió: 4:3



- ◆ : C'_R minták 1 félkép
- ◆ : C'_B minták 1 félkép
- : Y' minták 1 félkép
- ◆ : C'_R minták 2. félkép
- ◆ : C'_B minták 2. félkép
- : Y' minták 2. félkép



4:2:0 DV-EU (váltott soros)



◆ : C'_R minták 1 félkép

◆ : C'_B minták 1 félkép

● : világosság minták 1 félkép

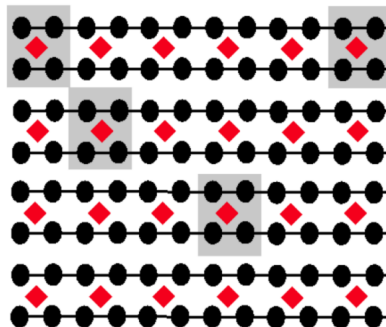
◆ : C'_R minták 2. félkép

◆ : C'_B minták 2. félkép

● : világosság minták 2. félkép

4:2:0 MPEG-1 (progresszív) és JPEG

- Az alapegység: 2×2 pixel
- A színkülönbségi jeleket vízszintesen és függőlegesen is 2-vel alul-mintavételezzük
- A színminták függőlegesen 2 sor, vízszintesen 2 szomszédos minta közé kerülnek
- 8bit: $(4+1+1)$ bájt=48bit/4pixel
- 10bit: $(4+1+1) \cdot 10 = 60$ bit/4pixel
- Ha a bemenő formátum 4:2:2, akkor a kompresszió: 4:3

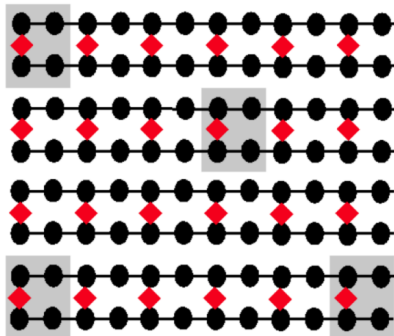


◆ : C'_R, C'_B színminták

● : világosság minták

■ : alap egység

4:2:0 MPEG-2



◆ : C'_R, C'_B színminták

● : világosság minták

■ : alapegység

- Az alapegység: 2×2 pixel
- Színjeleket függőlegesen 2-vel alul-mintavételezzük
- A színmintákat úgy számítjuk ki, hogy azok két egymás alatti sor közepére kerüljenek
- A keletkező színminták vízszintesen egybe esnek minden 2. Y mintával
- 8bit: $(4+1+1)$ bájt = 48bit/4pixel
- 10bit: $(4+1+1) \times 10 = 60$ bit/4pixel
- Ha a bemeneti formátum 4:2:2, akkor a kompresszió: 4:3

129



TV raszter formátumok megalkotásának alapelvei

- A látás a fő látótérben függőlegesen kb. 9-10 fokos szögtartományt ölel fel
- Vízszintesen (a **periférikus tér nélkül**) kb. 4:3 arányban nagyobb a látótér
- Az átlagos néző HVS-nek térbeli felbontóképessége korlátozott.



TV raszter formátumok megalkotásának alapelvei

- A televíziós kijelző sor (pixel) struktúrája ne legyen észrevehető, és zavaró,
- Ehhez a szomszédos pixelekből, illetve az egymás alatti sorokból a szemünkbe érkező fénysugarak bezárt szöge ne legyen nagyobb, mint a látás szögfelbontás határa (1 szögperc)
- Színes kép esetén az additív színkeverést nem kell elvégezni a képernyőben, ha a képpont mérete a fenti gondolatmenetet figyelembe véve elegendően kicsi.
- A szögfelbontás 1 szögpercnél kisebbre csökkentéséhez a sortávolságból számítható minimális nézőtávolságra kell legalább ülni
- Ez a kiindulópontja minden (SDTV, HDTV, UHD TV) televíziós szabvány, illetve képfelbontás tervezésének



Képváltási gyakoriság (képfrekvencia) és letapogatás módja

- Ahhoz, hogy egy folyamatosan változó képelem ne villogjon (fúziós frekvencia!), a képelem váltása nem lehet 50-60 Hz-nél kisebb.
- A fúziós frekvencia alatt a képelem villogni fog.
- Ahhoz azonban, hogy egy folyamatosan mozgó képelemet folyamatosan mozgónak lássunk, elegendő csupán 20-30 Hz gyakorisággal felrajzolni a mozgás fázisait. Ekkor ugyanazt a képet 2-3-szor felrajzolva nem látható villogás és a mozgás is folytonosnak hat (mozi).
- Ugyanazt 2-szer elküldeni sem sávszélesség hatékony módszer, és a teljes kép analóg tárolása nem megoldható, ezért a 2 félképes felbontást választották.

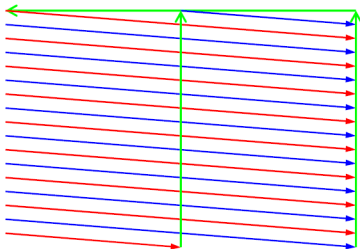


Két félképes - váltottsoros (interlaced) letapogatás

- A két félképes felbontás lényege:
- A teljes képek 25 Hz (30 Hz) gyakorisággal követik egymást.
- Egy kép két félképből áll:
 - páros sorok (páros félkép)
 - páratlan sorok (páratlan félkép)
- Az azonos képhez tartozó félképeket egymás után küldjük el, először a páratlant, aztán a párosat. A félképváltás frekvenciája így 50 Hz (60 Hz) lesz, ami által a képernyő tartalma már nem fog villogni.

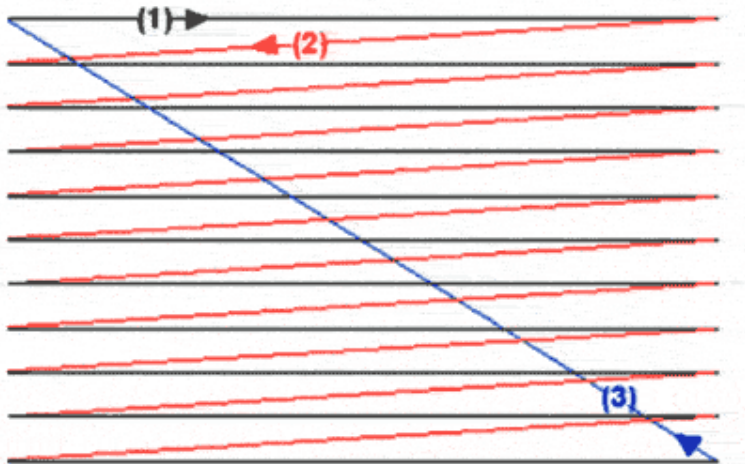
Páratlan sorszámú váltott soros letapogatás

- Csak egyfajta függőleges és vízszintes visszafutás kell.
- A teljes képen páratlan sorszám van, ezért:
 - a legalsó sor csak félig van meg, a sor felénél a függőleges visszafutás lép életbe,
 - emiatt az ezt követő legfelső sor is csak fél sor.





Progresszív letapogatás





SD formátumok

- Kezdetben a félképfrekvencia megválasztásánál az erősáramú hálózat frekvenciáját vették figyelembe
- A hálózati brumm a képen mozgó zajként jelenik meg, ha a két frekvencia jelentősen eltér (az álló zavarkép kevésbé zavaró, mint a mozgó).
- Európa, Ázsia: 625 sor (aktív sorok száma: 576), 50 Hz félkép, 25 Hz kép, $625 \times 25 = 15625$ Hz sorfrekvencia.
Formátumjelölés: 576i
- Amerika, Japán: 525 sor (aktív sorok száma: 480), 60 Hz félkép, 30 Hz kép, $525 \times 30 = 15750$ Hz sorfrekvencia.
Formátumjelölés: 480i
- Aktív oszlopok száma: 720 (704). **Nem négyzetes pixelek !**

ITU 601 SD video mintavételi frekvenciája I.

Követelmények

- A mintavételi frekvencia legyen közös az Európai (PAL) és az USA (NTSC) rendszerre
- A mintavételi struktúra legyen ortogonális, tehát egy TV-sorba egész számú mintavételi periódus férjen bele:
- $f_s = n f_h^{EU}$ és $f_s = k f_h^{USA}$, ahol n és k egész számok, f_h a sorfrekvencia
- $f_h^{EU} = 625 \times \frac{50}{2} \text{ Hz} = 15625 \text{ Hz} = 5^6 \text{ Hz}$
- $f_h^{USA} = 525 \times \frac{60}{2} \times \frac{1000}{1001} \text{ Hz} = 15734,2 \dots \text{ Hz}$
- $1001 = 7 \times 11 \times 13$

$$\begin{array}{rcl}
 & f_{H,480i} & : f_{H,576i} \\
 \hline
 & 525 \times \frac{60}{2} \times \frac{1000}{1001} & : 625 \times \frac{50}{2} \\
 7 \times 5^2 \times 3 \times \frac{5 \cdot 3 \cdot 2^2}{2} \times \frac{5^3 \cdot 2^3}{13 \cdot 11 \cdot 7} & : & 5^4 \times \frac{5^2 \times 2}{2} \\
 3 \times 3 \times 2^4 & : & 13 \times 11 \\
 144 & : & 143
 \end{array}$$



ITU 601 SD video mintavételi frekvenciája II.

- A legkisebb közös többszörös tehát
$$144f_h^{USA} = 143f_h^{EU} = 2,25\text{MHz}$$
- A mintavételi frekvenciának nagyobbnak kell lennie, mint a videojel maximális frekvenciájának (6 MHz) kétszerese
- Keressük tehát a 2,25 MHz legkisebb többszörösét, ami nagyobb mint a 12 MHz
- A 12 MHz-nél nagyobb legkisebb többszörös: 13,5 MHz
- Ezt választották a világosságjel mintavételi frekvenciájának, miközben a színkülönbségi jelek számára, figyelemmel az emberi látás tulajdonságaira, felezett mintavételi frekvenciát (6,75MHz) választottak



ITU 601 SD video mintavételi frekvenciája III.

$$f_{mv}^Y = 13,5 \text{ MHz} \xrightarrow{/2} f_{mv}^{U,V} = 6,75 \text{ MHz}$$



$$f_B^Y = 5,6 \text{ MHz}$$



$$f_B^{U,V} = 2,8 \text{ MHz}$$

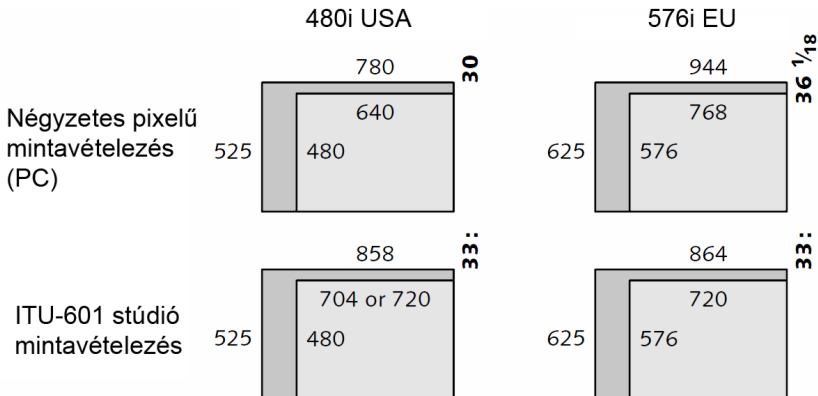
$$n_H^{USA} = \frac{f_{mv}^Y}{f_H^{USA}} = 858$$

$$n_H^{EU} = \frac{f_{mv}^Y}{f_H^{EU}} = 864$$

- A TV-sor aktív idejét 53,3 μsec -re választották, mely soronként 720 mintát eredményez
- A ténylegesen aktív időtartamra azonban csak 704 minta jut (8-8 minta fel- és lefutási tranziensek miatt)
- Mivel ez azonos az EU és az USA rendszerre, formátum konverzió esetén nincs szükség vízszintes irányú térbeli konverzióra

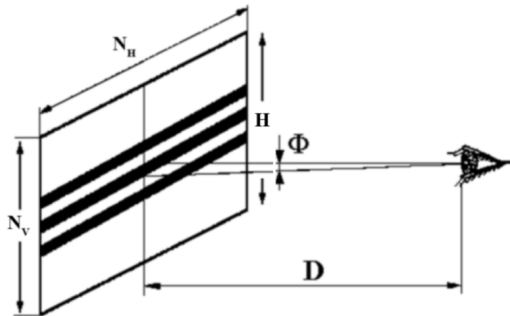


ITU 601 vs. PC mintavételi raszter





Ideális nézőtávolság I.



Jelöljük a kép magasságát pixelben: N_V -vel, a kép szélességet pixelben N_H -val, a szomszédos pixelekből a szemünkbe érkező fénysugarak által bezárt szöget Φ -vel. A nézőtávolságot jelöljük D -vel, míg a képmagasságot H -val.

Célunk, hogy $\tan\left(\frac{1}{60}\right) = \frac{H/N_V}{D}$



Ideális nézőtávolság II.

Ha $\tan(\frac{1}{60}) = \frac{H/N_V}{D}$, akkor pl.

- az EU SD 576-soros rendszerre:

$$D_{576} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{576 \times 2.9 \times 10^{-4}} H \approx 6H$$

- A 720 soros HD rendszerre:

$$D_{720} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{720 \times 2.9 \times 10^{-4}} H \approx 5H$$

- Az 1080 soros HD rendszerre:

$$D_{1080} = \frac{1}{N_V \tan(\frac{1}{60})} H = \frac{1}{1080 \times 2.9 \times 10^{-4}} H \approx 3H$$

A képhez tartozó vízszintes látószög (Θ)

$$\tan(\frac{\Theta}{2}) = \frac{\frac{A}{B}/2}{D}, \text{ és ebből}$$

$$\Theta = 2 \arctan(\frac{A}{2BD}), \text{ ahol } \frac{A}{B} \text{ a képarány (SD: 4:3, HD: 16:9)}$$



Ideális nézőtávolság III.

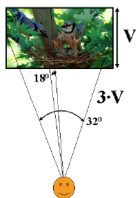
	Amerikai	Európai	HDTV
Tv-sor / képmagasság	480	576	1080
Nézőtávolság	7-szeres képmagasság	6-szoros képmagasság	3-szoros képmagasság
Nézőtávolság	4,25-szeres képátló	3,6-szeres képátló	1,5-szeres képátló
Vízszintes látószög	kb. 11 fok	kb. 13 fok	kb. 32 fok
Függőleges látószög	kb. 8 fok	kb. 9,5 fok	kb. 18 fok

A HDTV céljai és a HDTV paraméterek megalkotásának alapelvei I.

- A HDTV elsődleges célja az otthoni TV nézés élményének fokozása az által, hogy az SDTV-nél nagyobb látószöget tölt ki képtartalommal: ezáltal a TV kép "valóságszerűsége" nő
- Az elsődleges cél tehát nem a pixelben kifejezett "felbontás" növelése, nem az azonos képméretbe még több pixel "belesűrítése"

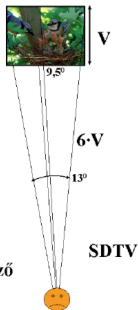


HDTV / SDTV összehasonlítás I.

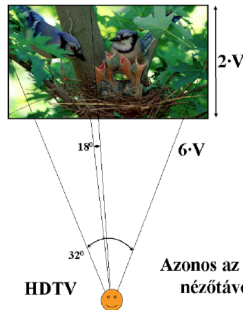


HDTV

Azonos az SD/HD kijelző
képmagassága

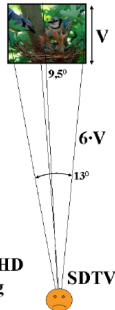


SDTV



HDTV

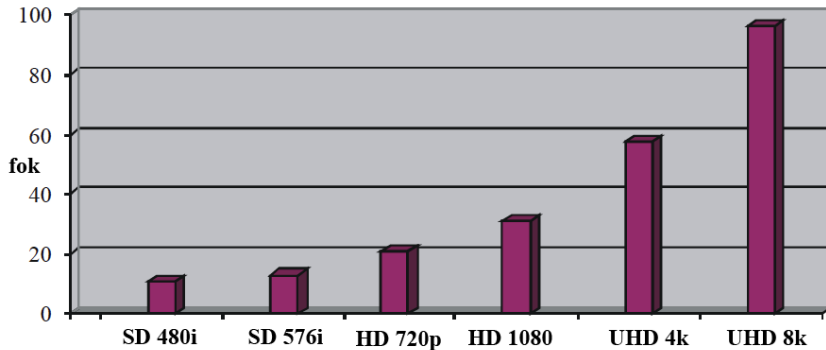
Azonos az SD/HD
nézőtávolság



SDTV

A HD tehát nem önmagában a nagyobb felbontást (kb. 6x teljes pixelszám az SD-hez képest) jelenti, hanem azt, hogy azonos szögfelbontást alapul véve a látóterünk nagyobb hányadát (periférikus látást is !) kitölti képtartalommal

Ideális nézőtávolság esetén a vízszintes látószög





Ideális nézőtávolság a szükséges képméret

Legyen a nézőtávolság fix: 3m (tipikus nézőtávolság, nappaliban, pl TV előtt)

A képmagasság: $H = DN_V \tan \frac{1}{60}$

A képszélesség: $W = H \times \frac{A}{B}$,

ahol $\frac{A}{B}$ a képarány (SD: 4:3, HD: 16:9)

- az EU SD 576-soros rendszerre: $H=50\text{cm}$, $W=67\text{cm}$
- A 720 soros HD rendszerre: $H=62\text{cm}$, $W=110\text{cm}$
- Az 1080 soros HD rendszerre: $H=94\text{ cm}$, $W=168\text{cm}$



A HD formátumtól elvárt követelmények

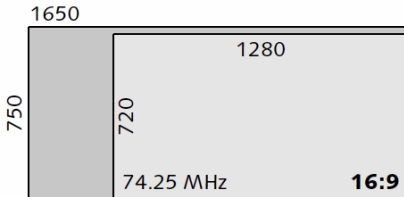
- A képméretarány (Aspect Ratio: AR) minimum 16:9
- A váltott-soros, vagy progresszív kijelzés nem definiált
- Független világosságjel és színinformáció kezelés
- Növelt fényerejű és felületű kijelző
- Minimálisan CD minőségű és sokcsatornás hangátvitel



HDTV formátumok

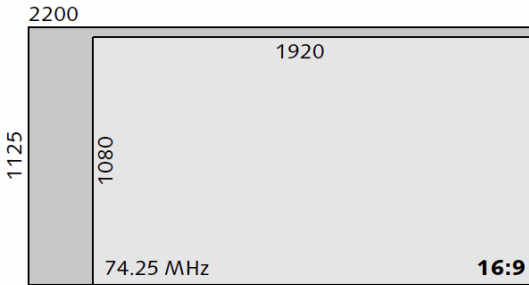
720p60

1280 × 720



1080p30, 1080i30

1920 × 1080





A HDTV RGB formátuma

A HDTV RGB formátumát az ITU 709-es ajánlása határozza meg

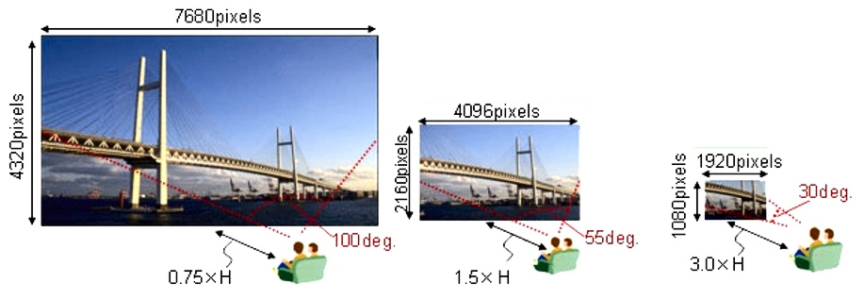
- Luma együtthatók: $Y_{709} = 0.212R + 0.715G + 0.072B$
- Nemlineáris (gamma-korrekciós) transzfer karakterisztika:

$$V' = \begin{cases} 4.500V & V < 0.018 \\ 1.099V^{0.45} - 0.099 & V \geq 0.018 \end{cases}$$



UHDTV - célkitűzések

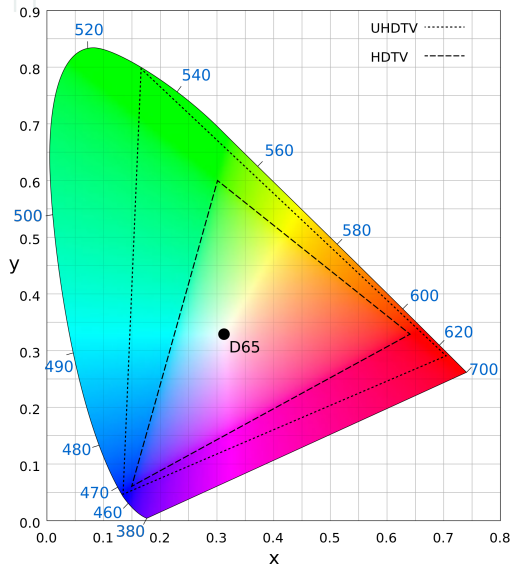
- A HDTV-hez képest tovább "fokozott" realitás-élmény:
 - még nagyobb képi látószög biztosítása.
 - javított mozgásleképzés (magasabb képsebesség támogatása)
 - kiterjesztett színtér
- A korábbiak alapján: térbeli felbontásigény 1 szögperc - változatlan
- Formátumok:
 - UHD-1 (4K): 3840 x 2160 - Látószög: 58 fok, ideális képméret 3m nézőtávolságból: 1.88 m x 3.35 m
 - UHD-2 (8k): 7680 x 4320 - Látószög: 96 fok, ideális képméret 3m nézőtávolságból: 3.76 m x 6.70 m (!!!)





ITU Rec. 2020

- Új RGB alapszínek: monokromatikus RGB alapszínek
 - R (CIE $x=0.708, y=0.292$), 630 nm
 - G (CIE $x=0.170, y=0.797$), 532 nm
 - B (CIE $x=0.131, y=0.046$), 467 nm
- Új luma együtthatók: $Y' = 0.2627 R' + 0.6780 G' + 0.0593 B'$
- Nem-lineáris transzfer karakterisztika megegyezik a Rec. 709-el
- YCbCr transzformáció és digitális reprezentáció megegyezik a Rec. 709-el
- 10 bit/minta, és 12 bit/minta formátumok
- 4:2:0, 4:2:2, és 4:4:4 (RGB) formátumok
- 60 Hz és 120 Hz (!) progresszív formátumok támogatása





UHDTV műsorszórási kísérletek

- Londoni Olimpia (2012) alatt BBC kísérleti adások 8k-ban, közösségi vetítések céljára London szerte
- Korea (2013): DVB-T2 és HEVC alapú földi 4k műsorszórási kísérletek
- Eutelsat (2013) DVB-S2 4k kísérleti adás (50 Hz, progresszív, 4 HD csatorna sávszélességben, H.264/AVC tömörítéssel)
- Japán (2014): DVB-S2 4k kísérleti adás tervezése a FIFA alatt

Aktív képtartalomra vonatkozó tömörítetlen video bitsebesség igények

- 1080i50/4:2:2 $(1080 \times 1920 + 2 \times 1080 \times 960) \times 8 \times 25 = 829 \text{ Mbit/s}$
- 1080i50/4:2:0 $(1080 \times 1920 + 2 \times 540 \times 960) \times 8 \times 25 = 622 \text{ Mbit/s}$
- 720p50/4:2:2 $(720 \times 1280 + 2 \times 720 \times 640) \times 8 \times 50 = 737 \text{ Mbit/s}$
- 720p50/4:2:0 $(720 \times 1280 + 2 \times 360 \times 640) \times 8 \times 50 = 552 \text{ Mbit/s}$
- ITU-601/4:2:2 (SD) (576i50) $(576 \times 720 + 2 \times 576 \times 360) \times 8 \times 25 = 166 \text{ Mbit/s}$
- ITU-601/4:2:0 (SD) (576i50) $(576 \times 720 + 2 \times 288 \times 360) \times 8 \times 25 = 124 \text{ Mbit/s}$
- CIF $(288 \times 360 + 2 \times 144 \times 180) \times 8 \times 30 = 37,3 \text{ Mbit/s}$
- QCIF $(144 \times 180 + 2 \times 72 \times 90) \times 8 \times 30 = 9,3 \text{ Mbit/s}$
- SIF $(288 \times 352 + 2 \times 144 \times 176) \times 8 \times 25 = 30,4 \text{ Mbit/s}$



1 órányi műsoranyag tárolási igénye

- 1080i50/4:2:2 829 Mbit/s : 347 Gbyte
- 1080i50/4:2:0 622 Mbit/s : 260 Gbyte
- 720p50/4:2:2 737 Mbit/s : 308 Gbyte
- 720p50/4:2:0 552 Mbit/s : 231 Gbyte
- ITU-601/4:2:2 (SD) (576i50) 166 Mbit/s : 70 Gbyte
- ITU-601/4:2:0 (SD) (576i50) 124 Mbit/s : 51 Gbyte
- CIF 37,3 Mbit/s : 15 Gbyte
- QCIF 9,3 Mbit/s : 3.8 Gbyte