

A multimédiatechnológiák alapjai

Dr. Firtha Gergely

April 1, 2020

Contents

1	Az emberi látás tulajdonságai	3
2	Fény- és színmérés	5
3	Videókomponensek	7
3.1	Eszközfüggő színterek	7
3.2	A TV-technika luma és chroma komponensei	14
3.3	Videójel komponensek	14
3.4	A digitalizálás kérdései	14
4	Videóformátumok	15
4.1	SD formátumok	15
4.2	A HD formátum	28
4.3	Az UHD formátum	36
5	A videótömörítés alapjai	41
6	Az MPEG tömörítés	43

Fejezet 1

Az emberi látás tulajdonságai

Fejezet 2

Fény- és színmérés

Fejezet 3

Videókomponensek

Az előző fejezet bemutatta az emberi látás képi reprodukció szempontjából legfontosabb tulajdonságait és részletesen tárgyalta a fény- és színmérés alapjait, bevezetve a világosság fogalmát és a CIE XYZ színteret. Ez a fejezet ezekre az ismeretekre építve bemutatja a televíziós-technikában használt színes-képpont ábrázolás módját, ez alapján bevezetve a jelenleg is alkalmazott analóg és digitális videójel komponenseket.

Videotechnika szempontjából az XYZ színteret ritkán alkalmazzák képpontok színkoordinátáinak tárolására, kivétel ez alól a digitális mozi és mozifilm-archiválási alkalmazások¹. Ugyanakkor az XYZ tér lehetővé teszi a különböző megjelenítők és kamerák által reprodukálható színek halmazának egyszerű vizsgálatát. A következő szakasz ezeket a konkrét videóeszközökre jellemző, ún. **eszközfüggő színtereket** mutatja be.

3.1 Eszközfüggő színterek

Az előző fejezetben láthattuk, hogy az emberi látás trikromatikus jellegének, valamint linearitásának (illetve az egyszerű lineáris modellnek) köszönhetően a látható színek egy lineáris 3D vektortérben ábrázolhatóak, amelyben a vektorok összegzési szabálya érvényes: Két tetszőleges szín keverékéből származó eredő színinger meghatározható a két színbe mutató helyvektor összegeként (függetlenül az eredeti színingereket létrehozó fény spektrumától). Az xy -színpatkón ennek megfelelően két szín összege a két színpontot összekötő szakasz mentén fog elhelyezkedni.

Ebből következik, hogy az emberi látás metamerizmusát kihasználva, a látható színek nagy része előállítható mesterségesen, megfelelően megválasztott alapsínek összegeként. Ez általánosan véve a színes képreprodukció alapja. Természetesen nem

¹Ennek oka, hogy egyrészt reprodukcióra közvetlenül nem használható, hiszen az XYZ alapsínek nem valós színek (az X,Y,Z bázisvektorok helyén nem található látható szín), másrészről a teljes látható színek tartománya igen nagy bitmélységet igényel, ráadásul feleslegesen: Az XYZ tér pozitív térfogatát a látható színek csak részben töltik ki (sok olyan kód lenne, amihez nem tartozik látható szín), ráadásul a ezen belül is a megjelenítők a látható színeknek csak egy részét képesek reprodukálni.

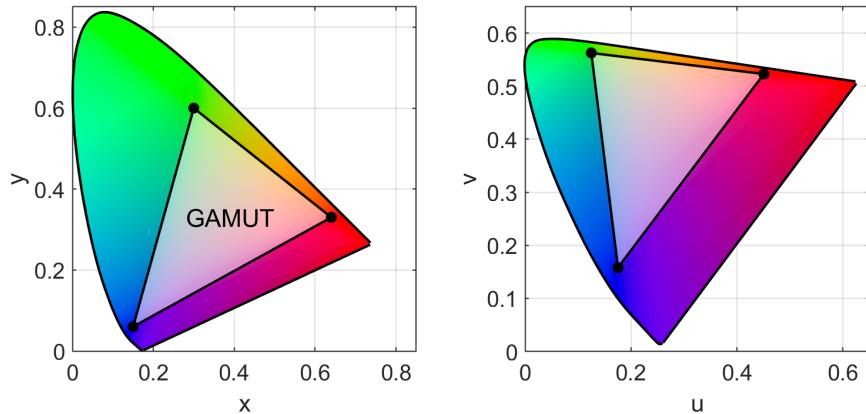


Figure 3.1. Az azonos alapszínekkel dolgozó SD formátum, HD formátum és az sRGB színtér gamutja xy és uv diagramon ábrázolva.

lehet célunk az összes látható szín visszaállítása: Minthogy a színpatkót a spektrál-színek határolják, így elvben végtelen számú spektrál-színt kéne alapszínként alkalmazni az összes látható szín kikeveréséhez. Felmerül tehát a kérdés, hány alapszín szükséges a színpatkó megfelelő lefedéséhez.

A színdiagramban könnyen felvehető 4 színpont úgy, hogy a négy szín keverékeit lefedő négyzet (azaz a reprodukálható színek területe) csaknem azonos területű legyen a látható színek területével. Ugyanakkor az *Luv* színtér színpatkójából láthattuk, hogy az emberi felbontás zöld árnyalatokra vonatkozó felbontása rossz, és az perceptuálisan egyenletes színdiagram inkább háromszög alakú. Ez azt jelenti, hogy három megfelelően megválasztott alapszínnel—amelynek különböző arányú keverékeinek színezete egy háromszögön belül helyezkedik el—az egyenletes színezetű (*uv*) színpatkó jelentős része lefedhető. Ebből kifolyólag az additív színkeverésen alapuló képreprodukciós eszközök szinte kizárolag három megfelelően megválasztott piros, zöld és kék alapszínnel dolgozik. Az ezekből a színekből pozitív együtthatókkal (RGB intenzitásokkal) kikeverhető színek összességét egy adott **eszközfüggő színtérnek** nevezzük. Ezzel ellentétben a kolorimetrikus, abszolút színterek, mint pl. a CIE XYZ, vagy Luv, Lab színterek ún. **eszközfüggetlen színtereknek**. Továbbá az adott eszközfüggő színtérben reprodukálható különböző színezetű színek az *xy*-színpatkóban felvett háromszögét a színtér **gamutjának** nevezik. Egy egyszerű példa színterek gamutjára a 3.1 ábrán látható.

Amennyiben egy RGB színtér teljesen ismert², tetszőleges C színe meghatározhatóak azok az RGB intenzitások, amelyekkel az RGB alapszíneket súlyozva megkaphatjuk a C színt (amennyiben az RGB értékek pozitívak). Ezek az adott C szín RGB koordinátái.

Vizsgáljuk most, hogyan szokás egy adott eszközfüggő (RGB) színtér definiálni

²Természetesen nem csak RGB színterek léteznek, nyomdatechnikában pl a CMYK eszközfüggő színterek elterjedtek, amelyek esetében a négy alapszín a nyomdában alkalmazott tinták színét jelzi. A következőkben a vizsgálatunkat kizárolag RGB színterekre végezzük el.

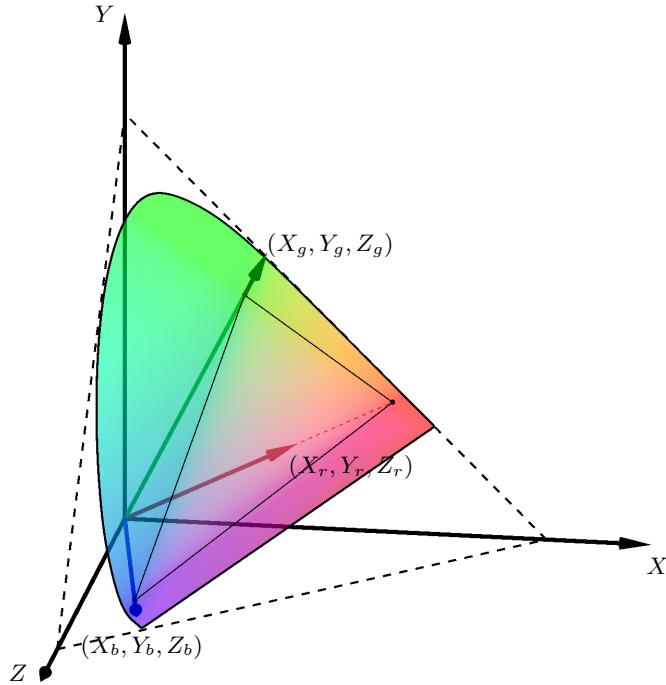


Figure 3.2. RGB színtér alapszíneinek helye, és metszéspontja az egységsíkkal az XYZ színtérben.

a gyakorlatban, azaz hogyan kell megadni a színteret ahhoz, hogy ezután tetszőleges szín RGB koordinátái számíthatók legyenek.

Eszközfüggő színterek definíciója:

Tekintsünk egy három alapszínt alkalmazó RGB színteret. Az R, G és B alapszínek természetesen egy-egy vektorként találhatóak meg az XYZ koordinátarendszerben, és vetületük/metszéspontjuk az egységsíkkal adja meg a színpatkón vett xy -koordinátáikat. Ezt illusztrálja a 3.3 ábra. Az alapszin-vektorok XYZ koordinátáit jelölje rendre

$$\mathbf{r}_{XYZ} = \begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix}, \quad \mathbf{g}_{XYZ} = \begin{bmatrix} X_g \\ Y_g \\ Z_g \end{bmatrix}, \quad \mathbf{b}_{XYZ} = \begin{bmatrix} X_b \\ Y_b \\ Z_b \end{bmatrix} \quad (3.1)$$

Amennyiben a három alapszin XYZ koordinátái ismertek, úgy a színtér teljesen definiálva van: tetszőleges \mathbf{c}_{XYZ} színvektor koordinátái meghatározhatóak az adott eszközfüggő RGB térben, amely \mathbf{c}_{RGB} vektor tehát azt írja le, milyen súlyozással

keverhető ki az adott C szín az RGB alapszínekből:

$$\underbrace{\begin{bmatrix} R_c \\ B_c \\ G_c \end{bmatrix}}_{\mathbf{c}_{RGB}} = \mathbf{M}_{XYZ \rightarrow RGB} \underbrace{\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}}_{\mathbf{c}_{XYZ}}, \quad (3.2)$$

ahol $\mathbf{M}_{XYZ \rightarrow RGB}$ egy bázistranszformációs mátrix. Vice versa, az RGB színtérben adott szín XYZ koordinátái meghatározhatóak

$$\underbrace{\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}}_{\mathbf{c}_{XYZ}} = \mathbf{M}_{RGB \rightarrow XYZ} \underbrace{\begin{bmatrix} R_c \\ B_c \\ G_c \end{bmatrix}}_{\mathbf{c}_{RGB}} \quad (3.3)$$

egyenletből. Természetesen fennáll a $\mathbf{M}_{RGB \rightarrow XYZ} = \mathbf{M}_{XYZ \rightarrow RGB}^{-1}$ összefüggés.

Utóbbi transzformációs mátrix egyszerűen meghatározható elemi lineáris algebra ismereteinkkel: $\mathbf{M}_{RGB \rightarrow XYZ}$ mátrix oszlopai egyszerűen az RGB színtér bázisainak XYZ -ben vett reprezentációja, azaz általánosan igaz

$$\underbrace{\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}}_{\mathbf{M}_{RGB \rightarrow XYZ}} = \underbrace{\begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix}}_{\mathbf{M}_{RGB \rightarrow XYZ}} \cdot \underbrace{\begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix}}_{\mathbf{c}_{RGB}} \quad (3.4)$$

összefüggés³.

A transzformációs mátrixok több szempontból jelentősek: egrészt lehetővé teszik a különböző színtérkonverziókat (lásd köv. bekezdés), valamint egy adott RGB téren ábrázolt képpont c_Y koordinátája megadja az adott szín relatív fénysűrűségét, azaz világosságát.

Itt jegyezzük meg, hogy az XYZ téren vizsgálva adott RGB bázisvektorokkal a pozitív együtthatókkal kikeverhető színek halmaza egy paralelepipedont feszít ki, azaz adott eszközfüggő RGB színtér az XYZ egy paralelepipedonként ábrázolható. Az RGB együtthatók definíció szerint 0 és 1 között vehetnek fel értékeket. Ennek megfelelően egy adott RGB téren az ebben a színtérben reprodukálható színek egy kockában helyezkednek el, ahol a kocka 3 origóból induló éle mentén az alkalma-zott RGB alapszínek helyezkednek el. Emiatt az RGB színtereket gyakran RGB kockaként említik. A transzformációs mátrixok tehát gyakorlatilag olyan lineáris transzformációt valósítanak meg, amelyek a paralelepipedont kockába, és a kockát paralelepipedonba viszik.

³Az összefüggés érvényessége könnyen tesztelhető pl. $\mathbf{c}_{RGB} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ helyettesítéssel, amely az R

alapszín RGB -ben vett reprezentációja, és (3.4) egyenletben a transzformációs mátrix első oszlopát választja ki.

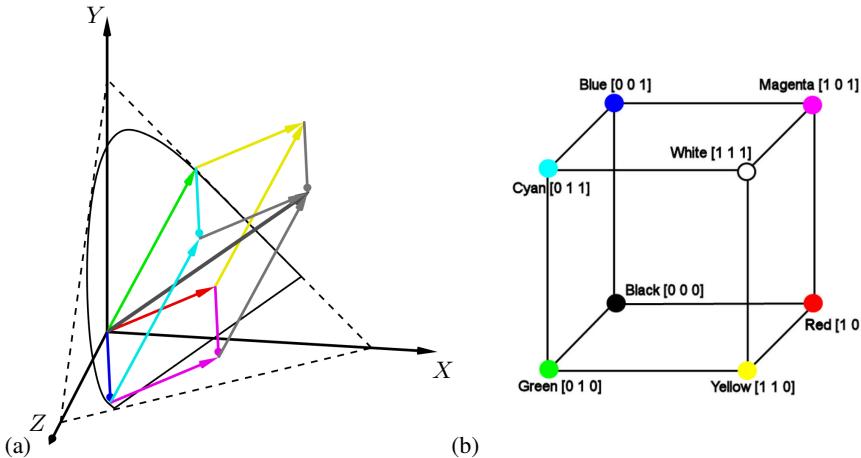


Figure 3.3. Egy adott *RGB* színtér ábrázolása az XYZ térben (a) és az *RGB* kockában (b). Az (a) ábrán szereplő vektorok színe a végpontjukban található szín határozza meg.

Egy *RGB* színtér tehát teljes egészében adott, amennyiben az alapszín-vektorok XYZ koordinátái ismertek (ez tehát 9 koordináta ismeretét jelenti). A gyakorlatban az ilyen definíció helyett az alkalmazott alapszínek színezetét, azaz xy -koordinátáit adják meg. Ennek az oka egyrészt a színtér gamutjának egyszerű ábrázolása (lásd 3.1 ábra), másrészt a fehér szín konzisztens, *RGB* színtértől független relatív fénysűrűsége (Y koordinátája).

Definíció szerint egy adott színtér ún. **fehérpontja** az adott téren elérhető legvilágosabb (legnagyobb fénysűrűségű) pontja, amelyet az alapszínek egyenlő arányú keverékkel érhetünk el. Mivel adott téren a 100%-os fehér a legvilágosabb elérhető szín, ezért definíció szerint a relatív fénysűrűsége (Y koordinátája) egységnyi. A 100%-os fehér tehát hasonlóan az XYZ -hez, definíció szerint

$$\mathbf{w}_{RGB} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{és} \quad Y_w = 1. \quad (3.5)$$

A 3.3 ábrán látható példában a fehér szín vektora a paralelepipedon szürkével jelölt föatlója, ezen vonal mentén helyezkednek el a különböző világosságértékű (árnyalatú) fehér színek. A fehér szín színezete, azaz x_w és y_w koordinátái ezen vektor egységsíkkal vett döfespontja határozza meg. Általánosan tehát, definiáltuk az adott *RGB* tér fehérpontját, amelynek érzékelt színezetét az adott alapszínek határozzák meg. Ez más szóval a szín akromatikus pontja, amely kijelzőről kijelzőre változhat, az alkalmazott pl. LCD elemek függvényében.

A három alapszín xy -koordinátái mellett a fehérpont x_w és y_w koordinátái és $Y_w = 1$ fénysűrűsége már elegendő információ szükség esetén a transzformációs mátrixok meghatározásához.

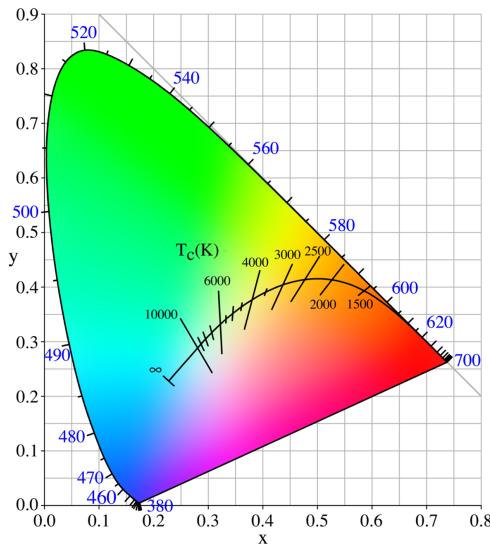


Figure 3.4. Különböző hőmérsékletű feketetest sugárzók által kellett színek összessége, azaz a Planck görbe.

A fehér színről általában:

Látható tehát, hogy a fehér szín önmagában szubjektív fogalom: adott környezetben a leginkább akromatikus fényingert nevezzük fehérnek, amelynek spektrális sűrűségfüggvénye minél inkább egységes (azaz minél több spektrális komponens tartalmaz), és ezzel analóg módon *RGB* színtér esetén a színvektora minél közelebb van a csupa-egy vektorhoz. A fehér fogalom egységesítéséhez vezettek ún. szabványos megvilágításokat (standard illuminants), amelyet szabványosított *RGB* színterek esetén előírnak, mint fehérpont. Ezeknek a szabványos megvilágításoknak a spektrális sűrűségfüggvénye adott (és persze az általa keltett színinger *xy*-koordinátái). Ilyen szabványos megvilágítások a következők:

- E fehér: egyenlő energiájú fehér, a CIE XYZ színtér fehérpontja. Kolorimetria szempontjából jelentős, videotechnikában kevésbé fontos a szerepe, mivel a gyakorlatban nem fordul elő olyan fényforrás, amely minden hullámhosszon azonos energiával sugároz.
- A fehér: a CIE által szabványosított, egyszerű háztartási wolfram-szálas izzó fényét (azzal azonos színérzetet keltő) fényforrás spektruma és színe, 2856 K korrelált színhőmérséklette⁴.
- B és C fehér: Az A fehérből egyszerű szűréssel nyerhető, napfényt szimuláló megvilágítások. A B fehér a déli napsütést modellez 4874 K színhőmérséklettel, míg a C fehér a teljes napra vett átlagos fény színét (spektrumát) modellez 6774 K színőmérséklettel.

⁴A korrelált színőmérséklet (correlated color temperature, CCT, T_C) azon feketetest sugárzó hőmérsékletét jelzi, amely az emberi szemben a minősítendő fényforrással azonos színérzetet kelt. A feketetest (hőmérsékleti) sugárzó által keltett színingerek az *xy* színdiagramon az ún. Planck-görbét járják be, amely a 3.4 ábrán látható.

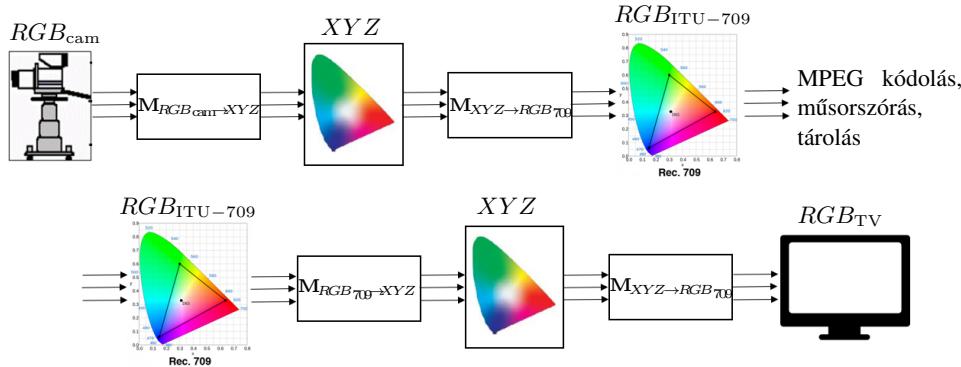


Figure 3.5. Színtér-konverzió folyamatábrája.

- D fehér: színtén a napfény közelítésére alkalmazott megvilágítások sora. Videotechnika szempontjából a legfontosabb a D65 fehér, amely jelenleg is az UHD formátumok színterének szabványos fehérpontja.

Színtér konverziók:

Az eddigiekben látható volt, hogyan definiálható egy eszközfüggő színtér az alapszíneivel. Ahogy az elnevezés is mutatja, ezek a színterek jellegzetesen adott eszközre érvényesek, pl. egy kamera a beépített *RGB* szenzorok, egy kijelző az alkalmazott *RGB* kristályok által meghatározott *RGB* színtérben dolgoznak. Emellett léteznek szabványos *RGB* színterek amely a képi és videotartalom tárolására, továbbítására szolgálnak egységesített, szabványos módon. A következő bekezdés ezeket a szabványos videószíntereket tárgyalja részletesebben. Felmerül tehát a természetes igény az egyes színterek közti átájárásra, amelyet **színtér konverzióknak** nevezünk.

A színtérkonverziót az *XYZ* színtér teszi lehetővé, amely egy eszközfüggetlen, abszolút színtér: egyes színterek közti konverzió a forrás által létrehozott jelek *XYZ* színtérbe való transzformációjával, majd ezen reprezentáció a nyelő színterébe való transzformációval történik. Az *XYZ* színtér így tehát színterek közti átájárást biztosít, ún. Profile Connection Space-ként működik (hasonlóan pl. a gyakran azonos célra alkalmazott *Lab* színtérhez).

Egy tipikus színtér konverziós folyamatot az 3.5 ábra mutat. Tegyük fel, hogy adott egy HD kamera által rögzített képanyag, ahol a kamera színterét *RGB_{cam}* jelöli. A HD formátum szabványos színteret alkalmaz, amelyet az ITU-709 ajánlásban rögzítettek (lásd később). A kamera *RGB* jeleit tehát az esetleges kódolás és tárolás előtt ebbe a HD színtérbe kell konvertálni. Ez a konverzió a kamerajelek *XYZ* térbe, majd innen az ITU-709 színtérbe való konverzióval oldható meg, amely konverziók a megfelelő transzformációsmátrixszal való szorzással valósítható meg:

$$\begin{bmatrix} R_{ITU-709} \\ B_{ITU-709} \\ G_{ITU-709} \end{bmatrix} = \mathbf{M}_{XYZ \rightarrow RGB_{709}} \cdot \left(\mathbf{M}_{RGB_{cam} \rightarrow XYZ} \cdot \begin{bmatrix} R_{cam} \\ B_{cam} \\ G_{cam} \end{bmatrix} \right) \quad (3.6)$$

Természetesen az egymás utáni két mátrixszorzás összevonható, így a két *RGB* színtér között közvetlen lineáris leképzés határozható meg. Ez a transzformáció jellegzetesen már a kamerán belül megvalósul. Hasonlóképp, megjelenítőoldalon a

$$\begin{bmatrix} R_{\text{cam}} \\ B_{\text{cam}} \\ G_{\text{cam}} \end{bmatrix} = \mathbf{M}_{XYZ \rightarrow RGB_{\text{TV}}} \cdot \left(\mathbf{M}_{RGB_{709} \rightarrow XYZ} \cdot \begin{bmatrix} R_{\text{ITU-709}} \\ B_{\text{ITU-709}} \\ G_{\text{ITU-709}} \end{bmatrix} \right) \quad (3.7)$$

transzformációt kell elvégezni.

Ez az egyszerű transzformációs módszer lehetővé teszi egy adott színtérben mért színpontok másik térbe való ábrázolását. Ugyanakkor felmerül a probléma, hogy nagyobb gamuttal rendelkező színterből kisebbre való áttérés esetén az új színtérben nem ábrázolható, gamuton kívüli színek negatív *RGB* koordinátákkal jelennek meg, míg kisebb gamutú térből való áttérés esetén a nagyobb gamutú tér egy része kihasználatlan marad. A probléma megoldására a fenti transzformációk mellett az egyes színterek gamutját valamilyen nemlineáris leképzés segítségével lehet egymásrailleszteni (ex-pandálással, kompresszállással). Ezek az ún. gamut-mapping technikák.

A következőkben a konkrét kamerák, illetve megjelenítők által alkalmazott színtereket vizsgálatától eltekintve, a szabványos SD, HD és UHD szabványokban alkalmazott színtereket vizsgáljuk.

A videotechnika színterei:

3.2 A TV-technika luma és chroma komponensei

3.3 Videójel komponensek

3.4 A digitalizálás kérdései

Fejezet 4

Videóformátumok

Az előző fejezetben láthattuk, milyen komponensekből épül fel a videóadat. A következő fejezet azt taglalja, hogy az eddig tárgyalt komponensekből hogyan állnak elő a különböző analóg, valamint digitális videoformátumok, illetve ezen formátumok paraméterválasztásának kérdéseivel foglalkozik. Láthatjuk, milyen irányelvek mentén került megválasztásra az egyes formátumok képmérete, térbeli felbontása (pixelszáma), képfrissítési frekvenciája. Emellett láthatjuk, hogy immár konkrét paraméterek mellett hogyan épül fel az analóg és digitális videójel.

4.1 SD formátumok

Elsőként a korai, normál felbontású NTSC és PAL analóg televíziós rendszerek képformátumát és paramétereinek megválasztását tárgyaljuk. Bár ezen analóg rendszerek már csak elvétve vannak használatban világszerte—Magyarországon például több éves digitális átállásra való előkészülés után 2013-ban szűnt meg az analóg műsorszórás—, mégis fontos tárgyalni főbb jellemzőit. Ennek oka történelmi jelentőségük mellett az, hogy a jelenlegi digitális műsorszórásban (a HD adás mellett) legelterjedtebb **normál felbontású (Standard Definition, SD)** digitális formátumokat közvetlenül az NTSC és PAL videojelek digitalizálásával kapjuk meg.¹

Képarány és képméret:

Elsőként fontos leszögezni, mekkora képméretre kell optimális formátumparamétereket választani. Az 1 fejezetben látható volt, hogy az emberi szemben a színlátás helye a sárgafolt, ezen belül is az éleslátásért a látogödörben (fovea centralis) elhelyezkedő receptorok felelnek. A látogödör mérete alapján az éleslátásunk a teljes ≈ 200 fokos látószögünkiből kb. 10-15 fokot fed le a horizontális irányban.² A normál felbontású televíziós szabvány megalkotása során a cél ezen fő látószög tartalommal való kitöltése volt, vagyis a normál felbontású televízió kb. a látótérből 10

¹Pontosabban az NTSC és PAL kompozit jeleket alkotó chroma és luma komponensek digitalizálásával.

²<http://hyperphysics.phy-astr.gsu.edu/hbase/vision/retina.html>

fokot kell, hogy kitöltsön (azaz a periférius látásnak a képalkotásban nem volt szerepe). Természetesen a konkrét képméret ezek után a nézőtávolság függvénye. Adott pixelméret/sortávolság mellett az optimális nézőtávolság megválasztásával a későbbiekben foglalkozunk.

A kép mérete mellett fontos térbeli jellemző a kijelző horizontális és vertikális dimenziójának aránya, azaz az ún. **képarány**. Az SD formátum alapjául szolgáló NTSC szabvány létrehozása az 1940-es évekig nyúlik vissza, és kidolgozása során nyilvánvaló törekvés volt a korabeli mozifilmek megjelenítésével való kompatibilitás biztosítása. A mozi korai korszaka, így a teljes némafilm korszak (az anamorf lencsék megjelenése előtt) kizárolag 4:3 képarányt alkalmazott, azaz a horizontális és vertikális képhosszak aránya 1.33 volt³. Habár az 50-es években megjelentek az első szélesvásznú mozinformátumok, az NTSC szabvány ezt a 4:3 képarányt fogadta el a televízió szabványos képarányának.

Képfrissítési frekvencia:

Következő kérdésként vizsgáljuk a mozgókép temporális mintavételi frekvenciájának, azaz a másodpercenként felvillantott képelemek számának megválasztási szempontjait. A továbbiakban ezt a frekvenciát **képfrissítési frekvenciának** nevezzük. Ennek meghatározásához két szempontot szükséges figyelembe vennünk. Egyrészt mozgó objektumok képi reprodukciója során fontos, hogy elegendő mozgási fázist tároljunk ahoz, hogy a megfigyelő folytonosnak érzékelje a képtartalom változását. Emellett elegendően magas képfrissítési frekvenciát kell választanunk a **villogás (flickering)** elkerüléséhez, azaz a képfrissítési frekvenciának a **fúziós frekvencia (flicker fusion threshold)** fölé kell, hogy essen.

Mint látni fogjuk, az utóbbi igényt szigorúbb követelményt a képfrissítési frekvencia megválasztásánál. Ennek oka az ún. béta mozgás (beta movement) nevű optikai illúzió, amely a látás azon jellemzője, hogy egymás után vetített statikus képek sorozatát 10 – 12 kép/másodperc (vagy frame-per-sec, fps) változás fölött az emberi szem már folytonos, látszólagos mozgásként érzékeli. A béta mozgás magyarázata máig sem teljesen tisztázott, leggyakrabban a látóidegen terjedő ingerület létrejöttének gyakoriságával, terjedési tulajdonságaival magyarázzák. A béta mozgás miatt tehát a folytonos mozgás biztosításához 20Hz képfrissítési frekvencia már elegendő lenne⁴, ezt a képfrissítési frekvenciát azonban az átlagos néző még villogónak érzékelné.

³A 4:3 képarány létrejötte egészen Thomas Alva Edison munkájáig vezethető vissza, aki az általa használt 35 mm széles filmen egy képkockát 4 perforációyi magasságúra (19 mm) definiált. A perforációk közötti kihasználható szélességből (25.375 mm) így a hasznos terület épp 4:3-hoz képarányúra adódik. A 35 mm-es filmen 4 perforációyi képméretet 1909-ben fogadták el általános szabványnak ("4-perf negative pulldown"), lehetővé téve a szabványos mozikamerák, mozigépek és így a mozi térhódítását.

⁴Érdemes megjegyezni, hogy ez a képfrissítési frekvencia csak ahoz elegendő, hogy ténylegesen mozgásnak érzékeljük a képsorozatot, ettől még a mozgás gyakran „darabos”: a nagyobb—pl. 60 kép/másodperccel rögzített és vetített képek folytonosabbnak, „simábbnak” fognak tűnni. Épp ezért számos modern kijelző, illetve számítógépes szoftver képes időbeli interpolációra, amely során az MPEG kódolóban is használatos mozgásbecslés alkalmazásával megpróbálják „kitálálni” az egyes képkockák közötti tartalmat. Érdekes tény azonban, hogy a néző szeme már kellően hozzászokott a mozin 24 fps rögzítési frekvenciához, emiatt a magasabb fps-el rögzített, vagy interpolált videó természetellenesen hat. Ennek a hatásnak a neve a szappanopera effektus (soap opera effect), amely elnevezés onnan származik, hogy a TV-s szappanoperákat—a klasszikus filmhez képest olcsón—közvetlenül digitális videóra rögzítették jellemzően 60fps-el.

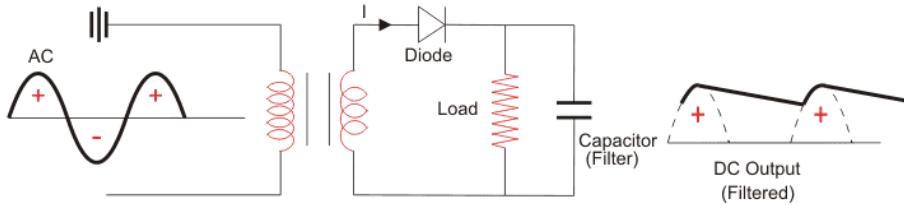


Figure 4.1. Periodikus hálózati brumm megjelenése az egyenirányított tápfeszültségen egyutas egyenirányítás esetén

Ahhoz, hogy ezt elkerüljük, a képfrissítési frekvenciának tehát magasabbnak kell lennie a fúziós frekvenciánál. A fúziós frekvencia fényingerek változásának azon frekvenciája, amely fölött a fényinger változását az emberi szem már nem képes követni. Különböző világosságú felületek váltakozása esetén a gyakorlatban efölött a megfigyelő csak egy „összeolvadt” átlagos világosságot érzékelni. A fúziós frekvencia értéke számtalan tényezőtől függ. Többek között emberről emberre változik, függ az átlagos megvilágítási szinttől és színhőmérséklettől, az adaptációs állapottól, a váltakozó fényinger színétől (a frekvencia növelésével jellemzően 15-20 Hz környékén a színezetbeli fluktuáció megszűnik, és csak a világosságszintek közötti vibrálás érzékelhető) amplitúdójától, és a gerjesztés helytől a retinán: azaz, hogy a villogást a fő látóterünkben, vagy a periférius látásunkkal érzékeljük-e.

Általánosan elmondható, hogy a fúziós frekvencia az embereknél 50-90 Hz közé esik: a fő látótérben, amelyet a csapok dominálnak a látás lassabb, itt a fúziós frekvencia 50 Hz, míg a periférius látás jóval gyorsabb, itt a fúziós frekvencia magasabb. Mivel az NTSC bevezetésénél a cél a fő látótér tartalommal való kitöltése volt, így célszerűen a képfrissítési frekvenciát 50-60 Hz környékére kellett választani. A konkrét érték megválasztását azonban már a katódsugárcsöves TV technológia egy hátránya határozta meg: a katódsugárcső tápfeszültségére rákerülő hálózati „brumm”.

A brumm (angolul ripple) a hálózati váltófeszültség egyenirányításának tökéletlenségeből származó periodikus zavarjel, ahogy az a 4.1 ábrán látható. A zavarjel frekvenciája a hálózati frekvenciával egyezik meg (egyutas egyenirányítás), vagy annak két-szerese (kétutas egyenirányítás esetén). Magyar elnevezése a hangerősítők kimenetén megszólaló jellemzően 50 Hz-es mélyfrekvenciás zugából származik. Televízió esetében mivel ez a zavarjel közvetlenül hozzáadódik a katódsugárcső vezérlőjeléhez, ezért a zavarjel kirajzolódik a kijelzőn, így látható hibát okoz.

Vizsgáljuk meg, mi rajzolódik ki a képernyón, ha a katódsugárcső vezérlője, azaz maga a videó jel (az egyszerűség kedvéért fekete-fehér esetben, azaz a jel a kirajzolandó fénysűrűség) periodikus, legegyszerűbb esetben 0 és 1 között oszcilláló szinuszos, azaz

$$Y(t) = \frac{1}{2} \sin 2\pi ft + \frac{1}{2}, \quad (4.1)$$

ahol f a vezérlőjel frekvenciája. A képernyőre ekkora sorról sorra kirajzolódik ez a szinuszos vezérlőjel. A kérdés, hogy mi a képernyő tartalma a f vezérlőfrekvencia

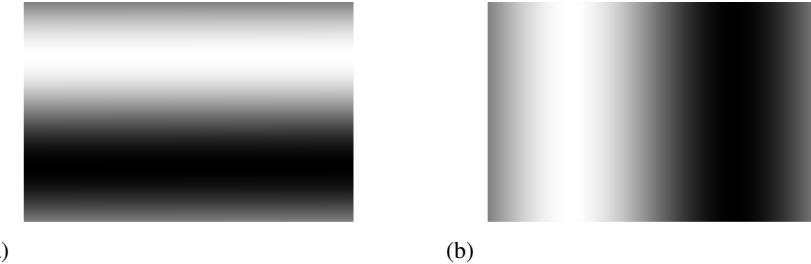


Figure 4.2. Periodikus jel képernyőn megjelenítve $f = f_V$ (a) és $f = f_H$ (b) választás-sal

függvényében.

Jelölje a képfrekvenciát f_V (mint vertikális frekvencia), és a sorfrekvenciát f_H (mint horizontális frekvencia), köztünk természetesen fennáll az

$$f_H = N_V \cdot f_V \quad (4.2)$$

összefüggés, ahol N_V a képernyő sorainak száma. Továbbá a későbbiekben jelölj $T_H = \frac{1}{f_H}$ a soridőt és $T_V = \frac{1}{f_V}$ a képidőt. Könnyen belátható, hogy

- $f = f_H$ választással minden sor tartalma ugyanazon szinuszhullám, a hullám kezdőfázisa minden sor elején és minden kép elején azonos, így egy álló, horizontális hullámforma jelenik meg a képernyőn, ahogy az a 4.2 (a) ábrán látható.
- $f > f_H$ választással a szinuszos jel fázisa sorról sorra lassan növekszik (mivel a periódushossza rövidebb, mint egy TV sor), így a hullámforma a horizontálishez képest enyhe dőlést mutat. Emellett már az első sorban is a hullám kezdőfázisa képről képre változik, így a teljes képtartalom lassan balra mozog. Hasonlóképp a sorfrekvencia alatti választással lassan jobbra mozgó képet kapunk.
- $f = f_V$ választással a teljes szinuszhullám egy teljes kép kirajzolásának ideje alatt rajzolódik ki. Mivel egy sor ideje alatt (megfelelően nagy N_V sorszám esetén) a jel értéke alig változik, ezért soronként állandónak tekintheő a tartalom. Így tehát a teljes képidő alatt egy álló, vertikális szinuszhullám jelenik meg a kijelzőn, ahogy az az 4.2 (b) ábrán látszik.
- $f > F_V$ választással a jelalak kezdőfázisa képről képre nő, így a hullámalak lassan felfelé mozdul. Hasonlóképp $f < F_V$ esetén a hullámalak lefelé mozog.

Az eszmefuttatás eredményeképp beláttuk, hogy periodikus jelek megjelenítése során megfelelő választással álló rajzolatot jeleníthetünk meg a kijelzőn. Márpédig a hálózati brumm épp ilyen periodikus zavarjelként jelenik meg a képernyőn, frekven-ciája pedig az adott régió hálózati frekven-ciája. A korai, fekete-fehér televíziós rend-szer megalkotása során végzett megfigyelési tesztek egyértelműen kimutatták, hogy elektromos zavar esetén az álló zavarkép jóval kevésbé zavarja a nézőt, mintha a zavar mozgó rajzolatként jelenne meg. Ennek megfelelően mind az amerikai, mind

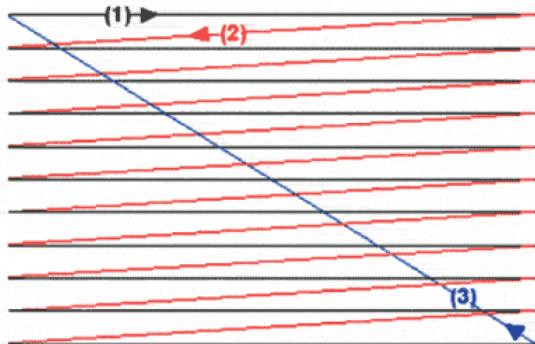


Figure 4.3. Progresszív letapogatás szemléltetése (az egyszerűség kedvéért 11 sorral ábrázolva), az aktív sortartalommal (1), sorvisszafutással (2) és képvisszafutással (3)

később, az európai rendszer esetében a képfrekvenciát a hálózati frekvenciának választották meg, így biztosítva, hogy az esetleges hálózati brumm a kijelzőn egy vertikális állóképként jelenik meg, amely a nézők számára alig észrevehető. Így tehát az amerikai rendszerben a képfrekvencia értéke $f_{V,USA} = 60$ Hz, az európai rendszerben $f_{V,EU} = 50$ Hz lett.⁵

Progresszív és váltott soros letapogatás:

A következőkben a különböző képernyő-letapogatási (scanning) módok vizsgáljuk. Az elnevezés a CRT kijelzőkhöz köthetők, ahol a katódsugár ténylegesen végigpásztázta valamelyen trajektória mentén.

A legkézenfekvőbb képernyő bejárási mód az ún. **progresszív letapogatás (progressive scanning)**, amely során a katódsugár egy képidő alatt sorról-sorra bejárja a képernyő összes sorát. A letapogatás módját a 4.3 ábra szemlélteti. Természetesen a jelenlegi LCD kijelzők esetében értelmetlen letapogatásról beszélni, ezek progresszív megjelenítési módban egyszerre változtatják az összes pixelsor tartalmát. Átviteltechnika szempontjából hasonlóan, ez azt jelenti, hogy az adott interface-en (konzumer berendezések esetében jellegzetesen HDMI-n keresztül) a kijelzőn megjelenítendő adat sorról sorra érkezik, és természetesen a teljes kép adatait egy soridő alatt továbbítani kell. A progresszív formátumot az alkalmazott sorszám utáni „p” jelölés mutatja, lásd HD esetében 1080p.

Bár a progresszív letapogatás tűnik a legegyértelműbb, legkézenfekvőbb megoldásnak, mégis, egészen az UHDTV szabvány megjelenéséig nem ez volt az általánosan elfogadott megoldás. Ennek okait a következőkben tárgyaljuk.

Az előző fejezetben láthattuk, hogy a folytonos mozgás biztosításához már 20 – 25 Hz képfrekvencia elegendő lenne, míg a villogás elkerüléséhez legalább 50 – 60 Hz képfürészési frekvencia szükséges. Ez már bizonyos szintű tömörítést tesz lehetővé,

⁵A helyzet a színes TV bevezetésével, azaz az NTSC megjelenésével Amerikában bonyolódott, mivel a színségérvivő frekvenciáját nem lehetett megfelelően megválasztani. Részletek nélkül: ennek eredményeképp mind a képfrekvenciát, mind a sorfrekvenciát 0.1 %-al csökkentették, így az amerikai rendszer képfrekvenciája $f_V = 60 \cdot \frac{1000}{1001} = 59.94$ Hz lett végül. Ezt a változás szerencsére a megfelelő szinkronjeleknek köszönhetően a már létező TV vevőkészüléket nem befolyásolta.

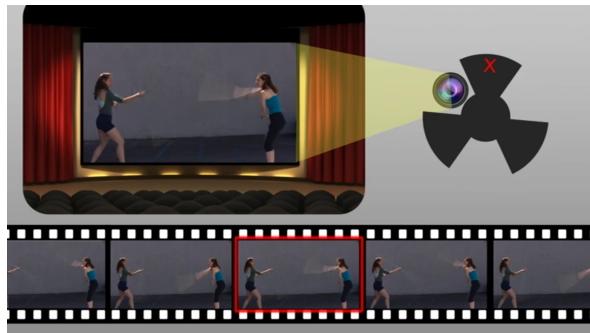


Figure 4.4. Triple blade shutter működése: <https://www.youtube.com/watch?v=jrSzRAch930>

hiszen a teljes képtartalom elegendő, ha lassabban változik, mint kijelző rajzolási frekvenciája.

Ez a tömörítés már a korai mozitechnikában is megjelent: A korai, némafilmes korszakban számos képfrekvencia volt használatban 16 – 24 Hz között. Manapság mozitechnikában a szabványos rögzítési frekvenciát 24 fps-re rögzítették. A villogás elkerüléséhez (tehát a képfrissítési frekvencia növeléséhez) speciális rekesszel látták el a vetítőgépet. A fénynyaláb útjában forgó rekesz, amelyen kettő, vagy három rés volt található (az ún. "two", vagy "three blade shutter") egy képkocka megjelenítése során tett meg egy teljes fordulatot, így a vetítőgép ugyanazt a képkockát kétszer, vagy háromszor villantja fel, mielőtt továbbhúzza a mozigép a szalagot. Ezzel az egyszerű trükkkel a 24 fps-en rögzített tartalmat 48fps, illetve manapság jellemzően 72 fps-en lehet megjeleníteni a mozikban.

Hasonlóan elven, a modern megjelenítők esetében a kijelző képfrissítési frekvenciája (pl. amivel egy LCD kijelző esetében a háttérvilágítás villog 200 Hz körül) jóval a tényleges képtartalom frissítési frekvenciája fölött van. A TV műsorszórás bevezetésének idején azonban a vevőkészülékek nem voltak képesek a képtartalom tárolására, a vett jel közvetlenül, valós időben rajzolódott ki a kijelzőre. A feladat megoldásául, azaz a másodpercenként átvivendő képek számának csökkentésére, és így sávszélesség-takarékoságra az ún. **váltott-soros letapogatást (interlaced scanning)** vezették be.

A megoldás alapötlete—ahogy a 4.5 ábrán is látható—a következő: Ahelyett, hogy a kijelző egy teljes képidő alatt az összes sort egymás után végigpásztázná, bontsuk a képernyőt páros és páratlan sorszámú sorokra, amelyek így egy páratlan és egy páros félképet alkotnak. A teljes képet (angolul **frame**) tehát két **félképre** (angolul **field**) bontjuk. A kijelző ezután a teljes képidő első felében a páratlan, a második felében a páros sorokat pásztázza végig. A váltottsoros formátum jelzése a sorszám mögé illesztett „i” jelzés (pl. 1080i).

Természetesen a képernyő tartalma a félképek frissítési frekvenciájával, az ún. **félképfrekvenciával** frissül, tehát ahhoz, hogy elkerüljük a villogást a félképfrekvenciának kell a fúziós frekvencia fölé esnie. Így váltottsoros letapogatás esetén a félképfrekvencia lett az európai rendszerben 50, valamint az amerikaiban 60 Hz-re (pontosabban 59.94 Hz-re) választva. A teljes, effektív képfrekvencia pedig ezek felére, tehát 25 Hz, illetve 30 Hz-re (29.97 Hz-re) adódik. A technikával tehát a mozin technikához hasonlóan, a képfrissítési frekvenciát elegendően magasra emelték, míg a

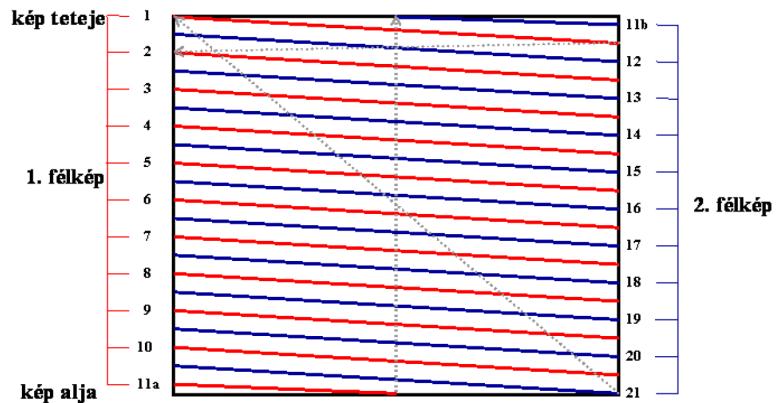


Figure 4.5. Váltott-soros képbontás (TV sorok közbeszövése), a jobb áttekinthetőség kedvéért 21 sorral. A teljes képernyő pásztázásához (feltéve, hogy az elektronnyaláb az első félkép első sorának elejéről indul) az első félképnek fél sorban kell végződni, míg a második félképnek félsorral kell kezdődni. Ez csak páratlan teljes sorszám esetén teljesül (mindkét félkép $N_{\frac{V}{2}} + \frac{1}{2}$ sorból áll, a teljes sorszám $2N_{\frac{V}{2}} + 1$, ami szükségszerűen páratlan)

tényleges, teljes „felbontású” képtartalom ehhez képest fele sebességgel érkezik.

Fontos megjegyezni, hogy a félképek (field-ek) különböző időpillanatokban készülnek, azaz nem ugyanazon teljes képhez tartoznak (nem állítható elő egy teljes kép páros-páratlan sorra való felbontásával). Ez eltérés a mozis rendszerhez képest, amely ugyanazt a képkockát mutatta be többször. Ennek eredményeként a következő módhatóak el a váltottsoros videóról:

- A váltott-soros letapogatás a progresszívhez képest 2:1 arányú tömörítést valósít meg, azaz a továbbítandó adatmennyiséget (és így a szükséges sávszélességet) lefelezi
- Álló képtartalomnál a progresszív letapogatással megegyező vertikális felbontást valósít meg (hiszen a páros és páratlan félképek ugyanazt a képet egészítik ki)
- Gyorsan mozgó képtartalom mellett a függőleges felbontás gyakorlatilag a progresszív formátum fele (hiszen a félkép tartalma folyamatosan változik)

Általánosan elmondhatjuk, hogy lassan változó képtartalom esetén (pl. filmek) a váltott-soros letapogatás megfelelően nagy vertikális felbontást és a progresszívnál folytonosabb mozgásreprodukciót biztosít, megfelelő tömörítés (sávszélességhatékonyság) mellett. Gyors kameramozgások esetén, pl. sporttartalom már láthatóvá válhatnak a felezett vertikális felbontásból származó hatások.

Az elmondottak alapján a normál felbontású SD formátum kizártlag váltottsoros letapogatási módot alkalmaz. A HD szabvány bevezetésével már mind interlaced, mind progresszív formátumok léteznek, míg UHDTV esetén a szabványok már kizártlag progresszív formátumokat definiálnak.



Figure 4.6. Megfelelő deinterlacing technika nélkül váltottSOROS formátum megjelenítése progresszív kijelzőn.

Érdekességképp elmondható, hogy az interlaced technika számos kérdést, nehézséget is felvet egyszerűsége mellett.

Egyik példaképp: korábban (előadáson) láthattuk, hogy a térbeli mintavételi frekvencia megsértése térbeli átlapolódási jelenségekhez vezet, amelyek jellegzetesen térben periodikus képek esetében (pl. téglafal, „kockás” ing) jól látható Moiré ábrák megjelenését okozza. Mivel interlaced esetben a vertikális mintavételi frekvenciát lefelezzük, ezért félképeken ezek a Moiré ábrák erőteljesen megjelenhetnek, az egymás utáni átlapolódó félképek váltakozása pedig igen zavaró átlapolódási jelenségekhez, ún. interline twitter jelensége vezet már állókép megjelenítése esetén is. A jelenségre egy szemléltető példa [itt](#) található. Minthogy az összes SD formátum interlaced letapogatást alkalmazott, épp az interline twitter jelensége volt a fő oka a TV felvételek során a négyzetrácsos, csíkos öltözékek elkerülésének.

További érdekes kérdést vet fel az interlaced és progresszív formátum közötti konverzió. Progresszívről interlaced formátumba a feladat viszonylag egyértelmű, a teljes kép páros és páratlan sorokra bontásával megoldható. A váltottSOROS formátumról progresszívre történő konverzió konzumer felhasználási szempontból gyakoribb, gondoljunk csak egy jellegzetesen váltottSOROS formátumban rögzített DVD lemez jellemzően progresszív számítógép monitoron történő megjelenítésére. Legegyszerűbb stratégiaként a monitor a szomszédos félképeket összeszűve alakít ki egy teljes felbontású képet. Ez azonban gyors mozgások esetén ún. fésűsdési jelenségekhez vezet, amelyet az [4.6](#) ábra szemléltet. Épp ezért, a konverzióhoz kifinomultabb **deinterlacing** eljárás szükséges a félképek sorai közötti adatok interpolációjához. A feladat létjogosultsága manapság is nagy, hiszen a jelenlegi LCD TV és számítógép monitorok már nem támogatnak natív interlaced megjelenítést, még a HD műsorszórás még napjainkban is váltottSOROS formátumot alkalmaz (jellemzően 1080i-t).

Analóg SD formátumok, az analóg videójel:

Az előzőek alapján bevezethetjük a normál felbontású analóg televíziós műsorszórás képformátumát:

Ahogy azt már korábban láthattuk két analóg képformátum terjedt el a világon a

színes műsorszórás kezdetével:

- Az Egyesült Államokban és Japánban alkalmazott NTSC képformátumot az FCC vezette be 1953-ban. Az NTSC formátum a korabeli technológiának megfelelően 525 TV-sorból áll, és váltott-soros letapogatást alkalmaz. A korábban tárgyalt okokból kifolyólag a rendszer kéfrissítési frekvenciája, azaz a félképfrekvencia 60 Hz (59.94 Hz), amelyből természetesen a képfrekvencia 30 Hz-re (29.97 Hz-re) adódik.
- Európában, Ausztráliában és Ázsiában a PAL rendszer került bevezetésre 1967-ben. A PAL formátum sorszáma $N_V = 625$, váltott-soros letapogatással, míg a helyi hálózati frekvenciának megfelelően a félképfrekvencia 50 Hz és így a képfrekvencia 25 Hz.

Az NTSC és PAL formátum fő jellemzőit a 4.1 táblázat tartalmazza.

Table 4.1

	NTSC	PAL
Összes sorok száma:	525	625
Aktív sorok száma:	480	576
Képfrekvencia:	30 Hz (29.97 Hz)	25 Hz
Félképfrekvencia	60 Hz (59.94 Hz)	30 Hz
Sorfrekvencia:	$525 \cdot 30 = 15750$ Hz (15734 Hz)	15625 Hz
Soridő:	$63.49 \mu\text{s}$ ($63.55 \mu\text{s}$)	$64 \mu\text{s}$

Fontos megjegyezni, hogy a CRT kijelző letapogatása során minden sor végén, minden félkép végén az elektronnyaláb kioltásra került, míg visszatérítették a következő sor, illetve félkép elejére. Ez a visszatérítés természetesen véges időbe telik. Ennek eredményeképp minden soridő, illetve félképidő tartalmaz inaktív, kioltási időintervallumokat (**blanking interval**), amelyben hasznos videójel nem található. Ezeket az időintervallumokat nevezük **sorkioltási időnek** (**horizontal blanking**) valamint **félképkoltási időnek** (**vertical blanking**).

Ennek megfelelően a teljes soridő felbontható ún. **aktív soridőre**, amely a tényleges videojelet tartalmazza és a sorkioltási időre, amely a sorszinkron (horizontális szinkron) jeleket, és egyéb jelzéseket tartalmaz.

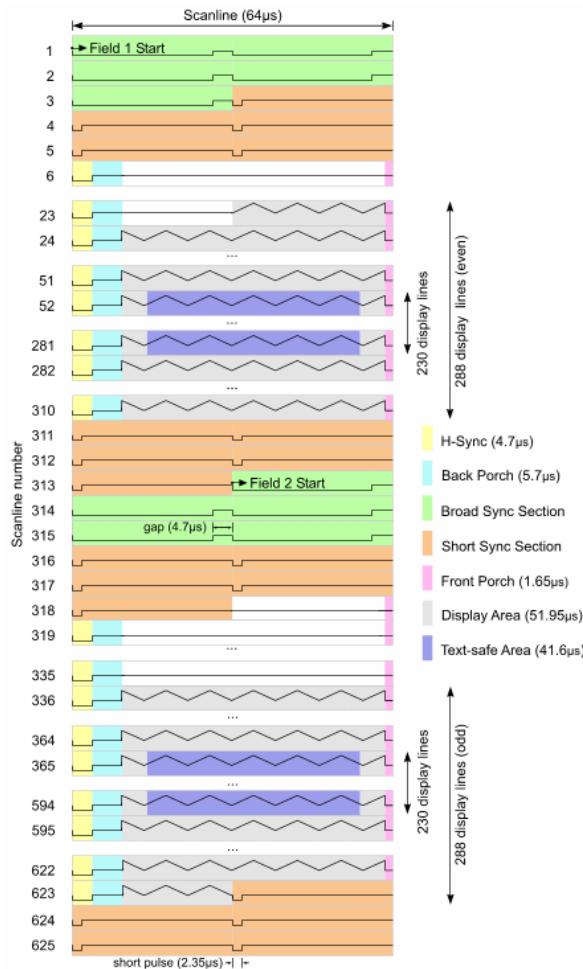


Figure 4.7. Egy teljes kép felépítése váltottosoros letapogatás esetén egyetlen videokomponensre ábrázolva:

- aktív soridő: szürke
- sorkioltási idő: magenta, cián és sárga
- félképkoltási idő: zöld, narancssárga, fehér

Hasonlóan, a teljes félképidő felbontható aktív sorokra, amelyek a megjelenítésre kerülő TV sorok, valamint a félképkoltási időre, amely vertikális szinkronjeleket és egyéb járulékos adatokat tartalmaz. Az aktív sorok számát szintén a 4.1 táblázat tartalmazza.

Egy teljes TV kép (azaz két egymás utáni félkép) felépítése a PAL rendszerben a ?? ábrán látható. Az ábra természetesen csak egyetlen video komponens felépítését szemlélteti. Az ábrán jól megfigyelhetőek a a félkép és képkoltási idők, bennük pedig az ún. **félképszinkron jelek (VSYNC)** (a zöld tartományban) és **sorszinkron jelek (HSYNC)** (sárga tartomány). Ezek a jelek a TV vevő (vagy általában a megjelenítőeszköz) szinkronizációját biztosítják a megfelelő megjelenítés érdekében. A szinkronjelek hibás vétele esetén a kép vertikálisan (félképszinkron hiányában), vagy horizontális (sorszinkron hiányába) elmozdul. Ezeket a jelenségeket „jitter”-nek, il-

letve „rolling”-nak nevezzük.

A jelenleg elterjedt megjelenítők esetén ezek a kioltási idők természetesen okafigyottá váltak: a modern, főként stúdió célú CRT megjelenítők már jóval kisebb kioltási idő mellett is működőképesek, míg LCD megjelenítők esetén egyáltalán nincs szükség kioltási időre. Ennek ellenére a kioltási idők a jelenlegi digitális szabványok esetén is ugyanúgy jelen vannak, így pl. a HDMI szabvány esetében is. Ennek egyik, természetes oka az, hogy a technika fejlődésével megjelenő újabb és újabb szabványok minden a már létező, korábbi szabványokra épülnek. Másrészt a kioltási idők lehetővé tettek egyéb, kiegészítő adatok tárolását is ezekben az időszegmensekben. Így a kioltási időkben továbbítható pl. a teletext adat, feliratok, és digitális esetben a video kísérő audio adat is. Ezen adatok helyét az ITU-R BT.1364 és az SMPTE 291M szabványok definiálják. Megjegyezhető, hogy a szabványok a digitális hang átvitelét (pl. a HDMI szabvány esetén is) a sorkioltási időben írják elő⁶.

Térbeli felbontás és az SD formátum:

Az analóg videojel tárgyalása után a továbbiakban rátérhetünk a videojel digitális reprezentációjának tárgyalására. Az első digitális videoformátumot a normálfelbontású, SD videót az ITU (akkoriban CCIR) alkotta meg 1982-ben a ITU-601 szabvány formájában⁷.

Az SD formátum gyakorlatilag az eddig tárgyalt videojel komponensek digitális reprezentációjának tekinthető, azaz a ?? látható videojel teljes egészében digitalizációra került kioltási intervallumokkal együtt, minden a luma és chroma komponensekre (más szóval az $Y'P'_bP'_r$ jelek közvetlen digitalizációjával kaphatjuk). A digitalizált videojelek neve–ahogy arról már szó volt– $Y'C'_bC'_r$ jelek. A jelek elvi előállítása az 4.8 ábrán látható.

A digitalizáció egyes kérdéseit már a korábbiakban érintettük. Nyitott kérdés még a soronkénti mintaszám meghatározása, amely a sorok számával együtt megadja az SD formátum felbontását (pixelszámát). A feladat tehát az analóg videojel mintavételi frekvenciájának meghatározása.

A mintavételi frekvencia megválasztásánál a következő szempontokat vették figyelembe:

- Természetes törekvés volt, hogy a több évtizede egymás mellett létező NTSC és PAL rendszerre egyszerre alkalmazható legyen, azaz mind PAL, mind NTSC video digitális ábrázolását lehetővé tegye. Emellett nyilvánvalóan a mintavételezést úgy kell végrehajtani, hogy minden rendszerben egy sorba egész számú mintavételi periódus (azaz pixel) férjen bele. Ebből következik, hogy a mintavételi frekvencia a sorfrekvencia egész számú többszöröse kell, hogy legyen minden az NTSC, minden a PAL rendszerben, azaz

$$f_s = n \cdot f_H^{\text{PAL}} = m \cdot f_H^{\text{NTSC}}, \quad (4.3)$$

⁶Egy egyszerű példaként HDMI audio átvitelre: 1080p HD formátum esetén (összes sor: 1125, ld. később) 60 Hz-es képfrekvencia mellett a sorfrekvencia $f_V = 67.5 \text{ kHz}$. 192 kHz mintavételi frekvenciájú 8 csatornás audioanyag átvitele esetén az egy kép alatt átvivendő audiominták száma: $\frac{8 \cdot 192000}{60} = 25600$ minta, azaz soronként kb. 23 minta átvitele szükséges. Ez a HDMI 1.0 szabvány által megengedett audiosebesség. (Példa folytatása itt található.)

⁷Munkájáért a CCIR 1983-ban tehenikai Emmy díjat is kapott

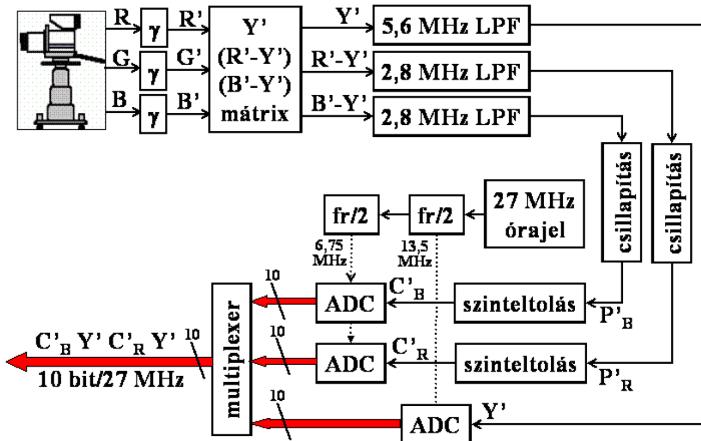


Figure 4.8. Digitális SD jel előállítása a videokomponensek digitalizálásával

ahol n, m egész számok. Minthogy a sorfrekvenciák

$$f_H^{\text{PAL}} = 25 \cdot 625 = 15625 \text{ Hz} \quad (4.4)$$

$$f_H^{\text{NTSC}} = 30 \cdot \frac{1000}{1001} \cdot 525 = 15734.2 \text{ Hz}, \quad (4.5)$$

ezek legkisebb közös többszöröse

$$144 \cdot f_H^{\text{PAL}} = 143 \cdot f_H^{\text{NTSC}} = 2.25 \text{ MHz}. \quad (4.6)$$

A mintavételi frekvencia tehát 2.25 MHz egész számú többszöröse.

- Emellett a mintavételi tételek értelmében a mintavételi frekvencia az átlapolódás elkerülésének érdekében legalább a mintavett jel sávszélességének kétszerese kell, hogy legyen. Korábban láttuk, hogy a luma jel sávszélessége 6 MHz, a chroma jeleké pedig ennek a fele.

A legkisebb frekvencia amire a két előbbi feltétel teljesül 13.5 MHz. Ezt választották tehát a világosságjel mintavételi frekvenciájának, miközben a színkülönbségi jelek számára, figyelembe véve az emberi látás tulajdonságait, felezett mintavételi frekvenciát (6.5 MHz) választottak. Ez az európai rendszerben 1 sorra 864, az amerikaiban 858 teljes mintaszámot eredményez, amely a sorkioltási időt is tartalmazza.

A két rendszer további egységesítésének érdekében egy soron belül az aktív pixelek számát közösen 720 pixelre választották (amelyből csak 704 pixel tartalmaz tényleges képi adatot, a digitalizálás előtti analóg jel kezdetének bizonytalansága, szélekhez közel torzításai, elmosódásai miatt). Ezzel tehát megkaptuk az SD formátum tényleges képméretét, ahogy az az 4.9 ábrán látható. Az aktív sorok száma alapján, és mivel minden rendszerben kizártlag interlaced video definiált, a két formátum megjelölése **480i** és **576i**. Könnyen belátható, hogy szabványos 4:3 képarány azonos horizontális pixelszám de különböző sorszám mellett csak úgy érhető el, ha az egyes képelemek (pixelek) nem négyzetalakúak (azaz a **pixel aspect ratio (PAR)** értéke 1-től

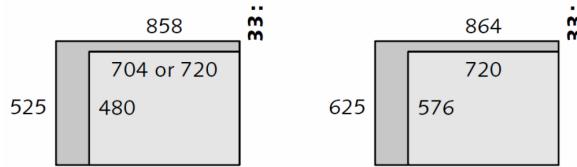


Figure 4.9. Az SD formátum képmérete az amerikai és az európai rendszerben

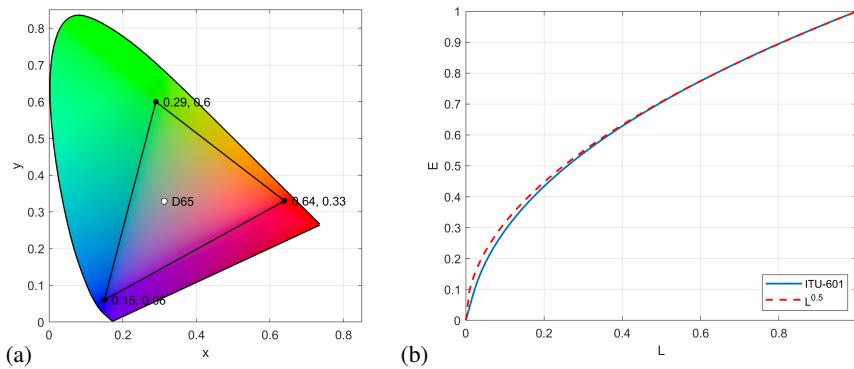


Figure 4.10. Az SD formátum gamutja(a) és Gamma-függvénye (b)

különböző). Számítástechnikában a monitorok ezzel szemben négyzetes pixelméretet definiáltak, így az elterjedt számítógépes SD formátum a jól ismert 640x480 pixelszám.

Röviden összefoglalva a jelen, és előző fejezetet a két SD formátum létrehozásának lépései és főbb tulajdonságai:

- A formátum primary színei és a színtér gamutja a 4.10 (a) ábrán látható. A színtér fehérpontja D65 fehér. A luma komponens számításának módja⁸

$$Y' = 0.299R' + 0.587G' + 0.112B' \quad (4.7)$$

- A forrás RGB jelei a perceptuális kvantálás megvalósításának érdekében Gamma-torzításon mennek keresztül, ahol a Gamma-függvény, vagy Optoelectronic Transfer Function:

$$E = \begin{cases} 4.500L, & \text{ha } L < 0.018 \\ 1.099L^{0.45} - 0.099, & \text{ha } L \geq 0.018, \end{cases} \quad (4.8)$$

ahol $L \in \{R, G, B\}$. A teljes görbe jól közelíthető egy $L^{0.5}$ függvénnyel

⁸Itt jegyezzük meg, hogy „matematikaiatlanul”, ezek a luma együtthatók az NTSC szabvány együtthatókból származnak, tradíció miatt a luma jel számítási módját nem változtatták meg az SD szabvány bevezetésével annak ellenére, hogy az alapszínek megváltoztak. Emiatt az $XYZ \rightarrow RGB$ mátrix második sora jelen esetben nem az itt bemutatott luma együtthatókat eredményezné. Ez egy újabb példa arra, hogy a hasonló matematikai következetlenségek nem állnak távol a gyakorlatban alkalmazott videoteknológiától.

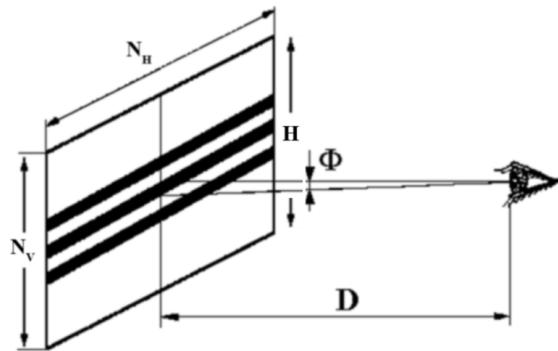


Figure 4.11. Geometria az optimális nézőtávolság származtatásához

- A formátum képaránya 4:3. A későbbiekben a HD megjelenése után ezt kiegészítették 16:9 képarányú formátummal is.
- Kizárolag interlaced formátum definiált
- A videókomponensek megfelelő sávkorlátozás után mintavételezésen és kvantáláson esnek át. A kvantálás 8, vagy 10 biten történik.
- A világosságjel mintavételi frekvenciája $f_s = 13.5$ MHz. Az 576i (625 soros, 50 félkép/s) rendszerben az aktív felbontás így 720x576 pixel, a 480i (525 soros, 60 félkép/s) rendszerben 720x480 px.
- A színkülönbségi jel az eredeti, stúdióformátumban 4:2:2, tehát a horizontális színcsatolás a világosságjelének a fele. Ezt később kiegészítették 4:2:0 struktúrával is konzumer célokra.

4.2 A HD formátum

Az előzőekben részletesen tárgyaltuk az SD digitális videoformátum megalkotásának alapelveit. A részletes vizsgálat oka, hogy ugyanezek az alapelvek, jelfeldolgozási lépések érvényesek a jelenlegi HD és UHD formátumok esetén is, valamint a jelenlegi SD műsorszórás Magyarországon is 576p (azaz már progresszív) formátumban történik.

Láthattuk, hogy az SD formátum megalkotásánál az egyes paramétereket úgy választották meg, hogy a kitűzött kb. 10 fokos látószögben minél élethűbb képi reprodukción lehessen megvalósítani. A HD és UHD formátumok tárgyalása előtt vizsgáljuk meg, hogy adott felbontás (pixelméret) mellett mekkora távolságból kell az adott kijelzőt megfigyelni, rávilágítva ezzel a HD formátum létrehozásának fő motivációjára.

Optimális nézőtávolság:

Általánosan elmondható, hogy pixel alapú képi reprodukció során a fő szempont,

hogy a szomszédos pixelekből érkező fény sugarak által bezárt szög az emberi szem felbontóképessége alá essen. Ezzel biztosítva van, hogy a kijelző pixelstruktúrája nem látható (a kép nem „pixeles”), valamint az RGB alapszínek alkalmazó reprodukció is lehetővé válik, hiszen az egyes alapszínek érzékelése helyett az additív színkeverés a szemben megvalósul.

Korábban láthattuk, hogy az emberi szem felbontása 1 szögperc (azaz $\frac{1}{60}^\circ$) (legalábbis a világosságjelre véve, segítségünkre van, hogy színezetre ennél is rosszabb). Adott pixelméretre természetesen ebből már meghatározható az a minimális nézőtávolság, amelyre az előbbi feltétel teljesül. Mivel jellemzően a kijelzőknek nem a pixelmérete van megadva, hanem a kijelző mérete és a vertikális, ill. horizontális pixelszám, ezért célszerű a fenti minimális nézőtávolságot ezek függvényében kifejezni.

Vizsgáljuk az 4.11 ábrán látható geometriát adott H magasságú, N_V sorszámú kijelző esetén. A pixelméret ekkor természetesen $\frac{H}{N_V}$. A kijelző a megfigyelőtől D távolságra helyezkedik el. A szomszédos (szemközti) pixelekből a megfigyelő szemébe érkező fény sugarakra felírható ekkor a

$$\tan \frac{\Phi}{2} = \frac{H}{2N_V D} \quad (4.9)$$

egyenlőség. Alkalmazzuk a tangens függvény kisargumentumú lineáris közelítését, azaz $\tan x \approx x$, ha $x \ll 1$. Ekkor

$$\Phi = \frac{H}{N_V D} \rightarrow D = \frac{H}{N_V \Phi} \quad (4.10)$$

érvényes. Az emberi szem felbontását $\frac{1}{60} \cdot \frac{\pi}{180}$ rad = $2.9 \cdot 10^{-4}$ behelyettesítve adott felbontású és méretű kijelző esetén az optimális (minimális) nézőtávolságra

$$D = H \frac{1}{N_V 2.9 \cdot 10^{-4}} \quad (4.11)$$

adódik. Ez a távolság az ún. Lechner-távolság, amely tehát megadja, hogy a tervezés során figyelembe vett képméret és felbontás mellett mekkora az optimális nézőtávolság adott képformátum esetén.

Amennyiben az a_r képátló ismert (SD esetén 4:3, HD esetén 16:9), az összefüggés kifejezhető a képszélesség függvényében is. SD esetén ez

$$D = \frac{W}{a_r} \frac{1}{N_V 2.9 \cdot 10^{-4}}, \quad (4.12)$$

ahol W a kép szélessége. Ekkor ha a kijelzőt az így kapott optimális távolságról nézik, meghatározható a képernyő által bezárt vízszintes látószög:

$$\tan \frac{\Phi_H}{2} = \frac{W}{2D} \rightarrow D = \frac{W}{2 \tan \frac{\Phi_H}{2}}, \quad (4.13)$$

és így adott felbontás mellett a horizontális látószög:

$$\Phi_H = 2 \arctan \left(\frac{a_r N_V 2.9 \cdot 10^{-4}}{2} \right). \quad (4.14)$$

Table 4.2. Fontosabb SD és HD formátumok ideális nézőtávolsága és az így kitöltött horizontális látószög

	Amerikai	Európai	HDTV
TV-sor/képmagasság:	480	576	1080
Nézőtávolság:	7-szeres képmagasság	6-szoros képmagasság	3-szoros képmagasság
Nézőtávolság:	4.25-szörös képátló	3.6-szoros képátló	1.5-szörös képátló
Vízszintes látószög	kb. 11 fok	kb. 13 fok	kb. 32 fok

Az eredményeket az eddig bemutatott SD és a következőkben tárgyalt HD formátumokra kiszámítva a 4.3 táblázat foglalja össze. Láthatjuk, hogy az SD felbontást ideálisan a képmagasság 6-7-szereséről célszerű nézni. Ekkor valóban, a formátum tervezésének kiindulási pontjába érünk vissza, azaz a kijelző a fő látóterünket, kb. 10-13 fokot tölti ki horizontálisan.

Ez már előreveti a HD formátum megalkotásának fő célját: a vizuális élmény fokozását nagyobb kitöltött látószög alkalmazásával. A HD formátum célja tehát—ellenértében a közhiedelemmel—nem a pixelben kifejezett felbontás növelése, és így azonos felületre minél nagyobb számú képpont belezsúfolása, hanem az otthoni vizuális élmény növelése a tartalommal lefedett látótér megnövelésével.

Rövid HD történelem:

A nagyfelbontású (High Definition) formátum létrehozása gyakorlatilag a televíziózás megjelenése óta a teljes XX. századon átívelt. Noha manapság a kifejezetten jellemzően digitális formátumra utal, már a korai analóg technika korában is léteztek HD kezdeményezések. Olyannyira, hogy már 1949-ben, Franciaországban kísérleti műsorszórást kezdtek monokromatikus (fekete-fehér), de 819 sort alkalmazó analóg rendszerben (a műsorszórás ebben a formátumban egészen 1983-ig tartott). A Szovjetunióban kísérleti jelleggel 1958-ban kifejlesztettek egy színes, 1125 sorból álló analóg rendszert, a hadászati célra létrehozott technikát azonban végül a gyakorlatban nem hasznosították. Végül a japán NHK cég vezette be az első mai értelemben vett HD műsorszórást 1989-ben. Rendszerük, az ún. Hi-Vision, vagy MUSE (Multiple sub-Nyquist Sampling Encoding) 5:3-as képarányú, 1125 soros analóg interlaced videó műsorszórására volt képes. A japán HD műsorszórás erőteljesen ösztönözte a HD formátum szabványosítását, amely azonban az analóg rendszer hatalmas sávszélessége miatt egészen a 90-es évek elejéig nem valósult meg.

A HD szabvány létrejötte végül a digitális tömörítési módszerek, főként az MPEG-1 és MPEG-2 tömörítések megjelenésének köszönhető. A szabványt 1990-ben tették



Figure 4.12. A 16:9 oldalarány, mint sok gyakori képarányú téglalap határoló-síkidoma és metszete

közre az [ITU-709](#) (Rec. 709) ajánlásban.

HD paraméterek:

Az ITU-709 szabvány a következő HD paramétereket határozta meg:

- Képarány:** az első szabványosított formátumjellemző a HD rendszer képaránya volt, amelyet 16:9 értékűre választottak. A választás nem kézenfekvő, mivel a formátum létrehozásakor gyakorlatilag nem állt rendelkezésre ebben a képarányban nyersanyag: mind a mozifilmek, mind a korábbi SD formátumú anyagok ettől eltérő képformátumban kerültek rögzítésre. A korabeli nyersanyagok jellegzetesen 4:3 (SD videó), 15:9, 1.85:1, 2.2:1, illetve 2,35:1 (mozis szabványok) voltak. Azt találták, hogy amennyiben azonos területű téglalapokat vetítünk egymásra a fenti, gyakori oldalarányokkal, akkor az így kapott téglalap sokaság éppen egy 16:9 oldalarányú téglalapba rajzolható bele. Továbbá az ezen téglalapok metszete szintén éppen egy 16:9 arányú téglalapot határoz meg. A geometria a [4.12](#) ábrán látható. A gondolatmenet eredményeképp, annak érdekében, hogy a legtöbb létező képarányú korabeli nyersanyag optimálisan megjeleníthető legyen az új formátumban esett a választás a ma már jól ismert 16:9 képarányra.
- Képfrissítési frekvencia:** Az SD rendszer hagyatékaként a HD szabvány számos képfrissítési frekvenciát támogat, így a 24 Hz, 25 Hz, 30 Hz, 50 Hz és 60 Hz képfrekvenciákat, valamint és ezek $\frac{1000}{1001}$ -szeres módosításait. Ezen kívül, ezek a frissítési frekvenciák minden progresszív, minden váltottsoros módban alkalmazhatók.
- Mintavételi frekvencia:** A mintavételi frekvencia esetében kiindulásul az SD mintavételi frekvencia szolgált. A HD formátum célja SD-hez képest minden

függőleges ($\times 2$), mind vízszintes ($\times 2$) felbontásduplázás, ezen felül a 4:3 képarány helyett 16:9 alkalmazása ($\times 4/3$), amely együttesen a mintaszám–és így a mintavételi frekvencia—legalább $2 \times 2 \times \frac{4}{3} = 5.33$ -szorozódását jelenti. A legtöbb képfrekvenciához való kompatibilitás biztosítása érdekében (azaz egy sorba egész számú minta férjen) a mintavételi frekvencia $5.5 \cdot 13.5 = 74.25$ MHz lett, illetve törtszámú képfrekvencia esetén ennek az $\frac{1000}{1001}$ -szerese. Progresszív esetben 50 – 60 Hz képfrekvencia esetén a mintavételi frekvencia ennek a duplája: $f_s = 148.5$ MHz.

- **Felbontás:** Hosszas egyeztetések után⁹ a szabvány 1080 aktív sort definiált 1125 teljes sorszámmal mind progresszív, mind váltottsoros letapogatás mellett (SMPTE 274M szabvány). Emellett a különböző képfrekvenciák egységes kezelése érdekében a soronkénti mintaszámot fixen 1920 mintára választották, így a HD formátum felbontása 1920x1080 pixel. A különböző képfrekvenciájú HD formátumok így tehát kizárolag a soronkénti inaktív pixelek számában különböznek egymástól (azaz a sorkioltási idő hosszában).

Table 4.3. Néhány HD formátum mintavételi frekvenciája, összes sor és oszlopszáma, felbontása. A soronkénti pixelek száma $S_{LT} =$

alapján számítható.

Rendszer	f_s [MHz]	S_{TL}	L_T	Aktív pixelszám
720p50	74.25 MHz	1980	750	1280×720
720p59.94	$74.25 \cdot \frac{1000}{1001}$ MHz	1650	750	1280×720
1080i25	74.25 MHz	2640	1125	1920×1080
1080i30	74.25 MHz	2200	1125	1920×1080
1080p50	148.5 MHz	2640	1125	1920×1080
1080p59.94	$148.5 \cdot \frac{1000}{1001}$ MHz	2200	1125	1920×1080

A különböző HD formátumok jelölése a következő:

- Teljes aktív felbontás pixelben kifejezve. Gyakran rövidítésképp csak a vertikális méretet jelölik meg.
- képletapogatás módja: p a progresszív, i az interlaced letapogatást jelöli

⁹Az első HD szabvány a Japán rendszer nyomán 1035 sort definiált. Ezt később visszavonták

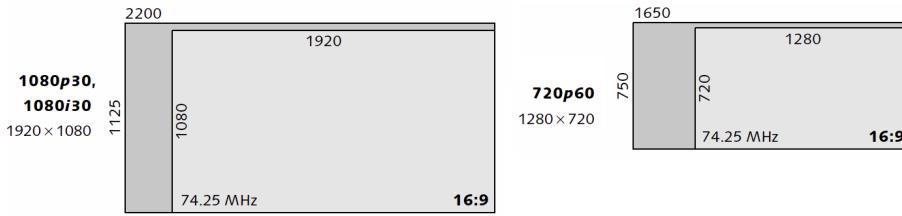


Figure 4.13. Az 1080 soros és 720 soros HD formátum szemléltetése

- képfrekvencia (frame rate) mind *p*, mind *i* esetben. Interlaced videó esetén gyakran a képfrekvencia helyett—hibásan—a félképfrekvenciát jelölik.

Így pl. az 1080*i*25 a 1080 soros, váltottsoros 25 Hz képfrekvenciájú, és 50 Hz félképfrekvenciájú formátumot jelöli.

A progresszív HD formátumok nagy adatsebesség-igénye miatt az ITU-709-et néhány évvvel később kibővítették egy alacsonyabb felbontású formátummal, ez 720 aktív sort és 1280 aktív pixelt alkalmaz, és kizártlag progresszív letapogatással definiálták¹⁰. Jelölése: **720p**. A két HD formátumra egy-egy példa a 4.13 ábrán látható.

- A HD formátum alapszínei az SD ITU-601-el megegyezők, így a gamutja is azonos. A luma komponens számításának módja

$$Y' = 0.2126R' + 0.7152G' + 0.0722B'. \quad (4.15)$$

Az együtthatók az SD-vel ellentétben már kolorimetriailag is helyesek, azaz a színtér alapszíneiből (és fehérpontjából) felírható $RGB \rightarrow XYZ$ transzformációs mátrix Y sorából ugyanezeket a világosság-együtthatókat kapnánk.

- A szabvány Gamma-karakterisztikája az SD-vel megegyező:

$$E = \begin{cases} 4.500L, & \text{ha } L < 0.018 \\ 1.099L^{0.45} - 0.099, & \text{ha } L \geq 0.018, \end{cases} \quad (4.16)$$

ahol $L \in \{R, G, B\}$.

- A szabvány az SD-vel azonos 8 és 10 bites digitális reprezentációt ír elő.
- A stúdió szabvány alapvetően 4:2:2 szín-mintavételezési struktúrát definiál.

A szabványos HD videójel ezek mellett teljesen az SD-vel azonos felépítésű, egy egyszerű példa a 4.14 ábrán látható. Az egyetlen különbség gyakorlati megvalósítás szempontjából, hogy a kioltási időszakokban a szinkron impulzusok ún. 3 állapotúak ($0, \pm 300$ mV).

¹⁰Az 1080p bevezetése során az egyik kitűzött cél a legalább duplázott sorszám volt. A 720p az SD és HD között félúton: másfélszerves sort alkalmaz így a sorszáma $\frac{3}{2} \cdot 480 = 720$ -ból adódik, míg az oszlopszám a 16:9 képarányból, négyzetes pixelek mellett

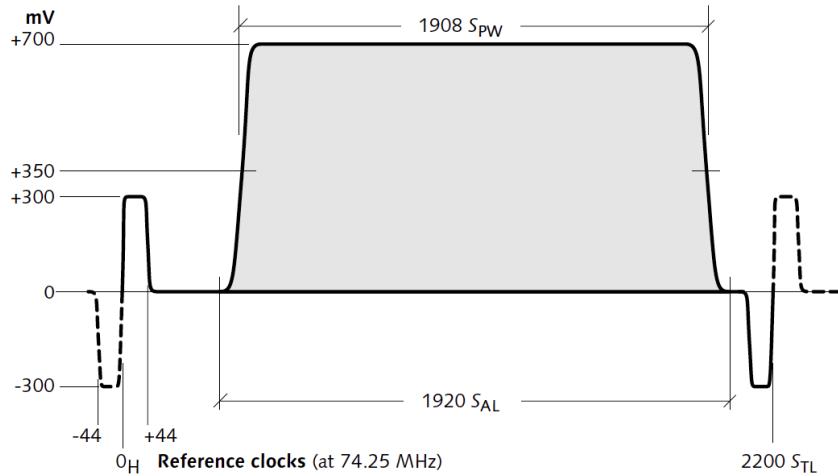


Figure 4.14. 1080 soros HD formátum egy sorának felépítése

Tömörítetlen videó adatsebesség:

Vizsgáljuk most néhány alapvető képformátum esetén a videó tömörítetlen adatsebességét! Az aktív és teljes pixelszámokat a 4.15 ábrán látható jelölésekkel jelölve a teljes adatsebesség

$$BR_T = \underbrace{S_{TL} \cdot L_T \cdot f_f \cdot n_{bit}}_{\text{mintánkénti bitrate}} \cdot n_{CS}, \quad (4.17)$$

ahol f_f a képfrekvencia, n_{bit} a mintánkénti bitszám és n_{CS} a színkülönbségi jelek alulmintavételezését jelöli (komponens/minta). Utóbbi értéke 4:4:4 struktúra esetén $n_{CS} = 3$, 4:2:2 esetén $n_{CS} = 2$, 4:2:0 esetén $n_{CS} = 1$.

Hasonlóan, az aktív tartalom bitsebessége

$$BR_A = S_{TA} \cdot L_A \cdot f_f \cdot n_{bit} \cdot n_{CS} \quad (4.18)$$

alapján számítható.

Néhány gyakran alkalmazott videóformátum teljes és aktív videósebessége a ?? táblázatban látható¹¹. Látható, hogy a 720p és 1080i formátumok tömörítetlenül azonos adatmennyiséget generálnak, ugyanakkor nagy előnye a progresszív formátumnak jóval hatékonyabban tömöríthetősége. Ahogy korábban a váltott-soros formátum előnye ezzel szemben, hogy állóképekre a progresszívvel azonos vertikális felbontást biztosít, bár gyors mozgásokra ez a felbontás romlik. Épp ezért, azon műsorszolgáltatók, amelyeknél a sport-tartalom elsődleges jellemzően a 720p50/60 formátumot használ, míg a főként filmeket, hírműsorokat sugárzó operátorok jellemzően 1080i-t alkalmazznak. Magyarországon jelenleg a HD adások mellett minden szolgáltató biztosít SD felbontású verziót is, amely jellemzően 576p50 formátumot alkalmaz.

¹¹ Az UHD formátum kioltási idejei szabványosan 90 inaktív sor és 560 inaktív pixel/sor

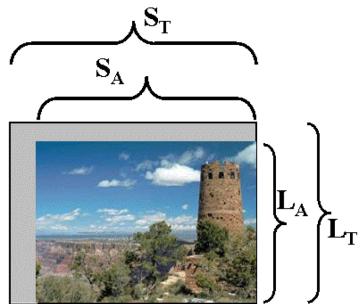


Figure 4.15. Jelölés-konvenció az adatsebesség számításához

- S_{TL} : Összes mintaszám/sor
- S_{TA} : Aktív mintaszám/sor
- L_T : Összes sor/kép
- L_A : Aktív sor/kép

Table 4.4. Videóformátum mintavételi frekvenciája, összes sor és oszlopszáma, felbontása, $n_{\text{bit}} = 10$ bitmélység esetén

Rendszer	Mintavételi frekvencia	Teljes bitrate 4:2:2	Aktív bitrate 4:2:2	Aktív bitrate 4:4:4
576p50	13.5 MHz	0.54 Gbit/s	0.41 Gbit/s	0.62 Gbit/s
720p60	74.25 MHz	1.49 Gbit/s	1.11 Gbit/s	1.67 Gbit/s
1080i30	74.25 MHz	1.49 Gbit/s	1.24 Gbit/s	1.86 Gbit/s
1080p60	148.5 MHz	2.97 Gbit/s	2.49 Gbit/s	3.73 Gbit/s
2160p60	297 MHz	11.88 Gbit/s	9.96 Gbit/s	14.93 Gbit/s

Jelenleg stúdiótechnikában a legelterjedtebb digitális videóinterface az SDI (Serial Digital Interface), konzumer elektronikában pedig a HDMI (High-Definition Multimedia Interface). A fenti adatsebességek kézzelfoghatóvá-tételéhez, néhány jelenleg is használt HDMI verzió a következő adatmennyiségek továbbítását teszi lehetővé:

- HDMI 1.0-1.2: 4.95 Gbit/s (3.96 Gbit/s tényleges)¹²
- HDMI 2.0: 18 Gbit/s (14.4 Gbit/s tényleges)
- HDMI 2.1: 48 Gbit/s (38.4 Gbit/s tényleges)

Természetesen a HDMI méretezéshez a fenti táblázatból a teljes adatsebességet kell tekinteni, hiszen a HDMI kábelben terjedő HD jel tartalmazza a kioltási időket is: mint

¹²A HDMI szabvány itt nem részletezett okokból ún. 8b/10b csatornakódolást alkalmaz, amely során 8 bitnyi adatot 10 biten visz át. A 4.95 Gbit/s teljes sávszélességnek tehát csak $\frac{8}{10}$ része használható ki tényleges adatra.

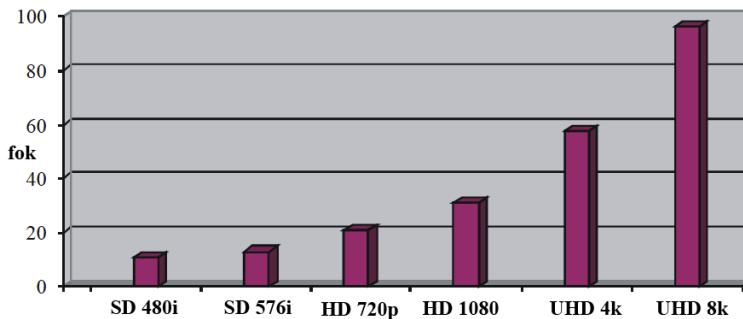


Figure 4.16. Az egyes formátumok által biztosított horizontális látószög, a kijelzőt ideális nézőtávolságból nézve.

korábban tárgyaltuk ezekben az időrésekben kerülnek az audio csatornák és egyéb járulékos adatok elhelyezésre. Látható, hogy a HDMI 1.0 szabványt főként 1080p videó továbbítására fejlesztették ki 4:2:2 formátum 10, vagy 12 bit, míg 4:4:4 formátum már csak 8 biten továbbítható. A HDMI 2.0 szabványt már 4k míg a 2.1 szabvány 8k UHD videó továbbítására fejlesztették ki.

4.3 Az UHD formátum

Láthattuk, hogy a HD formátum megalkotása során a fő cél a néző látóterének—az SD-hez képest—nagyobb részének tartalommal való kitöltése volt. A HD-hez hasonlóan az UHD (Ultra High Definition) formátum a vizuális élmény továbbfokozását tűzte ki fő céljául a pixel-szám—és így a kitöltött látószög—további növelésével.

Hasonlónak a HD-hoz, a továbbemelt felbontású formátum fejlesztése a japán NHK nevéhez kötődik, akik már 2003-ban kísérleti UHDTV felvételt készítettek 16 darab HDTV rögzítővel és 4 darab 3840x2048 felbontású CCD szenzorral. Az első UHDTV szabvány ezek után már 2007-ben megjelent (SMPTE 2036), majd a jelenleg is érvényben lévő ITU-R BT.2020 szabványt 2012-ben fogadták. A szabvány két új formátum paramétereit kodifikálja, a 4k formátumot, amely az 1080 soros HD formátumhoz képest minden dimenzió mentén duplázott felbontást (így 4-szeres teljes pixelszámot) alkalmaz, és a 8k-t, amely a 4k-hoz képest kétszeres felbontást definiál.

A szabvány célja szerint ideálisan a 4k formátum a néző horizontális látószögéből kb. 58°-ot, a 8k formátum 96°-ot tölt ki, azaz már a perifériás látás jelentős részét is tartalommal tölti ki. Az egyes formátumok által ideálisan kitöltött látószög változását a [4.16](#) ábra szemlélteti.

Ahhoz, hogy az ekkora látószögben megfelelő minőségű képi reprodukció valósuljon meg, az ITU-2020 szabvány számos szempontból javította a HD formátum alapparamétereit:

- Felbontás:** A szabvány két felbontást definíál, ezek a 3840x2160 (4k) és a 7680x4320 (8k) aktív képméretek. A HD szabványnak megfelelően minden

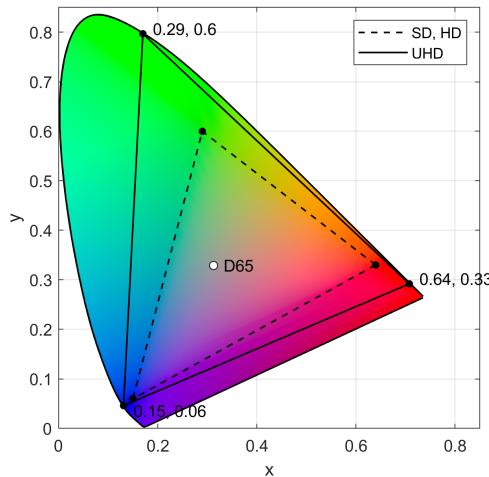


Figure 4.17. Az ITU-2020 szabvány gamutja, összehasonlítva a HD színtérrel

formátum képaránya 16:9 és négyzetes pixelekből áll.

- A HD-vel ellentétben az ITU-2020 szabvány már kizártlag csak progresszív letapogatást támogat (így a p/i formátummegjelölés UHD esetén okafogyott. Fontos, hogy a kijelző már optimális esetben a néző perifériás látásának jelentős részét is kitölți, ahol az érzékelést már a szem pálcikái dominálják. Mivel ezek fúziós frekvenciája jóval nagyobb, mint a fő látóterünké, ezért a villogás elkerülése érdekében a képfrekvenciát az UHD esetében emelni kellett. Ennek megfelelően a szabvány a 120, 119.88, 100, 60, 59.94, 50, 30, 29.97, 25, 24, 23.976 Hz képfrekvenciákat támogatja).
- **Színtér:** Az ITU-2020 szabvány az NTSC óta először új alapszíneket vezetett be a nagyobb ábrázolható színtartomány érdekében (és a világosság dinamikatartományának növelése érdekében, ld. köv. pont). A szabvány színterének gamutja a 4.17 ábrán látható. Megfigyelhető, hogy a választott alapszínek spektrális színek (azaz a színpatkó határán helyezkednek el), az RGB hullámhosszok rendre 630 nm, 532 nm és 467 nm¹³. Ezzel a CIE színpatkó 75.8%-át lefedi (összehasonlításképp, a HD esetében ez 35.9%), és a Pointer féle valós felületi színek

¹³Természetesen ez nem azt jelenti, hogy az UHD kijelzők RGB fényforrásai spektrális színek lennének, a szabvány csak a tároláshoz és továbbításhoz használt színteret kodifikálja. A megjelenítés már az egyes kijelzők saját színterében történik, amit a ténylegesen alkalmazott alapszínek korlátoznak. Szakmai körökön belül napjainkban is fontos hírnék számít, ha egy kijelző az ITU-2020 szabványos színterét közel egészében képes megjeleníteni. Így pl. 2018-ban a JDI cég bemutatott egy 17.3"-es kijelzőt, amely lézeres háttérvilágítással az ITU-2020-as szabvány gamutjának 97%-át képes megjeleníteni.

egészét¹⁴. Az új alapszínekből a világosság a következőképp számítható:

$$Y = 0.2627 R + 0.6780 G + 0.0593 B. \quad (4.19)$$

- **Mintánkénti bitszám:** A tárolt színek tartományának növelése természetesen magával vonja a világosságjel dinamikatartományának növekedését is (nyilván, hiszen adott szín világosságtartalma 4.19 alapján az RGB értékekkel kiszámítható). UHD esetében tehát már nem csak az SD esetében ökolszabályként tárgyalt 100:1 dinamikatartomány ábrázolása volt a cél. Ennek oka, hogy a 100:1 dinamikatartomány a látás fő látóterében, rögzített tekintet mellett értelmezendő. Amennyiben egy jeleneten belül a nézőnek lehetősége van körülözni, úgy a pupilla tágulása és összeszűkülése ezt a dinamikatartományt kb. 5-szörösre emeli. Mivel rendeltetésszerűen egy UHD kijelző akkora részét tölt ki a látómezőnek, hogy ez a lokális adaptációs mechanizmus végbe tud menni, így a megfelelő reprodukált dinamikatartomány biztosításához a reprezentálandó dinamikatartományt is növelni kellett. Részben ezt valósítja meg a szabvány növelt gamutja, illetve erre szolgál a jelenleg elterjedőben levő HDR kiegészítés is. Az ábrázolt dinamikatartomány növelése természetesen magával vonja az ábrázolásra használt bitek számának növelését is, hogy az egyes szintekhez tarozó világosságértékek közti különbség továbbra is az emberi látás modulációs küszöbje alatt maradjon. Kísérletek kimutatták, hogy a ITU-2020 színterének perceptuális kvantálásához 11 bit elegendő, így a szabvány a HD szabvánnyal ellentétben már kizárálag 10 és 12 bit reprezentációt ír elő.
- **Gamma függvény:** Az ITU-2020 szabvány Gamma-korrektív függvénye az SD és HD szabványokkal megegyezik:

$$E = \begin{cases} 4.500L, & \text{ha } L < \beta \\ \alpha L^{0.45} - (\alpha - 1), & \text{ha } L \geq \beta, \end{cases} \quad (4.20)$$

ahol $\alpha = 1.09929682680944$ és $\beta = 0.018053968510807$. Különbségeként látható: a szabvány a függvény paramétereit nagyobb pontossággal definiálja, és 12 bites ábrázolás esetén a paraméterek 5 tizedesjegy pontossággal számolandók.

- **Mintavételi struktúra:** A szabvány 4:2:0, 4:2:2 és 4:4:4 chromamintavételi struktúrákat engedélyez, utóbbi esetben természetesen luma/chroma jelek helyett közvetlenül RGB jelek kerülnek tárolásra és átvitelre. 4:2:0 és 4:2:2 esetében az YC_bC_r ábrázolás mellett lehetőség van ún. konstans fénysűrűségű luma-chroma reprezentációra is.

A fejezet zárógondolataként térdünk vissza a különböző sorszámu kijelzők ideális nézőtávolságához, az ún. Lechner távolsághoz. Az optimális nézőtávolságra

¹⁴A Pointer féle valós felületi színek azon színek halmaza, amelyek a természetben előfordulnak, mint valamely felületről visszavert fény által kellett színerzeti (ellenében a természetben elő nem forduló színekkel, pl. neon, vagy monitor által kikeveri színek). A Pointer gamutot egy 4089 mintából álló mérési adatbázisból állították össze, és publikálták az eredményeket 1980-ban, azóta a Pointer gamut-lefedés az egyes színterek minősítésének de facto szabványa. A Pointer féle gamut itt található, illetve itt található egy összehasonlítás a gyakran alkalmazott színterek gamutjával.

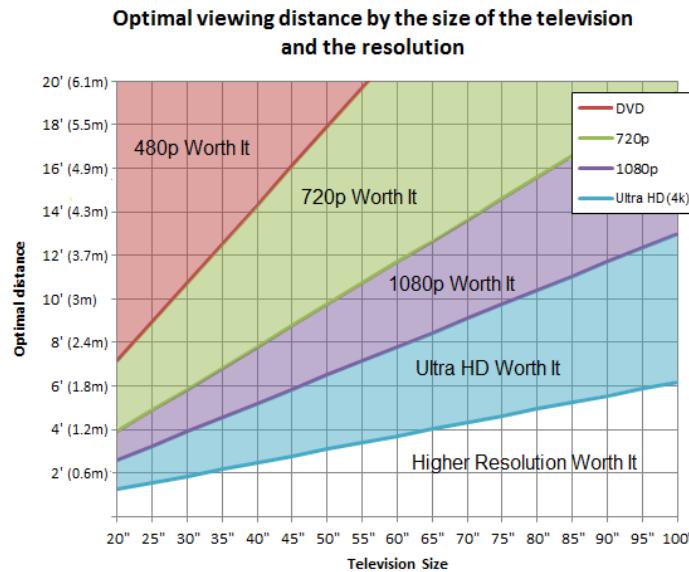


Figure 4.18. Optimális nézőtávolság a kijelzőátmérő függvényében

vonatkozóan számos ajánlás létezik, így léteznek gyártói, forgalmazói, illetve THX ajánlások.

- SMPTE 30: a már tárgyalt, Lechner tavolságon alapuló ajánlás, amely HD felbontás esetén 1.6 x képátló nézőtávolságot ír elő, így a horizontális látószög kb. 30 fok. Ez a házimozi közösségen belül legelfogadottabb, általános nézőtávolság
- Gyártói és forgalmazói ajánlás: HD felbontás esetén 2.5 x képátló nézőtávolság, így a horizontális látószög kb. 20 fok.
- THX ajánlás: HD felbontás esetén 1.2 x képátló nézőtávolság, így a horizontális látószög kb. 40 fok, amely a THX szerint a legjobban közelíti a mozikban biztosított élményt.

Bár a fenti ajánlások jelentősen eltérnek egymástól, abban azonban egyetértenek, hogy "minél közelebb, annál jobb".

Az optimális nézőtávolság ábrázolható a képernyőátmérő függvényében különböző videóformátumok mellett. Az így kapott grafikont 4.18 mutatja. Ekkor egy adott képernyő méret mellett a nézőtávolságot növelte természetesen a pixelstruktúra nem válik láthatóvá, így pl. az 1080p vonala fölött a HD kijelzők használhatók. A gondolatmenet alapján a grafikon területekre oszthatók, amelyek megadják, egy adott kijelzőméret és nézőtávolság mellett milyen felbontású kijelző az optimális választás.

Statisztikák szerint (egy szintén Bernard J. Lechner nevéhez köthető felmérés szerint) a háztartásokban az átlagos nézőtávolság kb. 2.7 méterre adódott. Ekkora nézőtávolság mellett 50" képátló felett már 1080p felbontású TV-t érdemes venni, míg az UHD tartalom adta előnyök kiélezéséhez legalább 75"-es (~ 1.9 m) kijelzőre lenne

szükség. Jelenleg az ekkora kijelzők természetesen még nem elterjedtek, így ez az adat leginkább azt mutatja, hogy jelenleg a 4k tartalmak előnyeit sem használják ki teljes egészében a kijelzők, és a vásárlók nagyrésze. Ennek ellenére mára sorra jelennek meg a konzumer felhasználásra szánt 8k felbontású kijelzők is, és a 8k műsorszórás is kísérleti jelleggel megkezdődött. Többek között 2017-ben lőttek fel az első dedikáltan 8k tartalom közvetítésére szánt műholdat (BSAT-4a), amely tervek szerint a 2020-as Japánban tartandó nyári olimpiát hivatott 8k felbontással az ITU-2020 szabvány szerint közvetíteni (amelyet végül a koronavírus járvány miatt 2021 nyarára ütemeztek át).

Ellenőrző kérdések

- Mi volt az oka a váltottsoros formátum bevezetésének? Mi volt a váltottsoros megoldás lényege?
- Mik voltak az SD formátum mintavételi frekvenciájának szempontjai? Hogy következett ebből a HD formátum mintavételi frekvenciája?
- Határozza meg egy 65" átmérőjű 2160p kijelző (16:9 képarányú) ideális nézőtávolságát!
- Sorolja fel az UHD szabvány néhány újdonságát az SD-hez és HD-hez képest!
- Határozza meg 2160p60 formátum esetén az aktív és teljes pixelszámot! Az aktív sorok száma a szabvány szerint 90 sor. A mintavételi frekvencia 297 MHz.
- Határozza meg az előző feladat formátumára a teljes adatsebességet 4:2:2 mintavételi struktúra esetén, 12 bit/minta ábrázolás mellett! Hányas HDMI verzió képes a videóadat továbbítására, ha a HDMI interface szabványosan 8 bitnyi adatot 10 biten ábrázol és visz át, és a különböző verziók sebességei a következők:
 - HDMI 1.0-1.2: 4.95 Gbit/s
 - HDMI 1.3-1.4: 10.2 Gbit/s
 - HDMI 2.0-1.2: 18 Gbit/s
 - HDMI 2.1: 48 Gbit/s

Fejezet 5

A videótömörítés alapjai

Fejezet 6

Az MPEG tömörítés