

## **Fejezet 1**

# **Az emberi látás tulajdonságai**



## **Fejezet 2**

# **Fény- és színmérés**



## Fejezet 3

# A TV-technika videóformátumai

### 3.1 Videóformátumok

A következő szakasz a különböző analóg, valamint digitális videóformátumok paraméterválasztásának kérdéseivel foglalkozik. Láthatjuk, milyen irányelvek mentén került megválasztásra az egyes formátumok képmérete, térbeli felbontása (pixel-száma), képfriességi frekvenciája.

#### 3.1.1 SD formátumok

Elsőként a korai, normál felbontású NTSC és PAL analóg televíziós rendszerek képfomátumát 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.<sup>1</sup>

#### Képarány és képméret:

Elsőként fontos leszögezni, mekkora képméretre kell optimális formátum-paramétereket választani. Az 1 fejezetben látható volt, hogy az emberi szemben a színlátás helye a sárgafolt, ezen belül is az éleslátásért a látógödörben (fovea centralis) elhelyezkedő receptorok felelnek. A látógödör mérete alapján az éleslátásunk a teljes  $\approx 200$  fokos látószögünk kb. 10-15 fokot fed le a horizontális irányban.<sup>2</sup> 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 fokot kell, hogy kitöltsön (azaz a periférikus látásnak a képalkotásban nem volt szerepe). Természetesen a konkrét képméret ezek után a nézőtávolság függvénye. Adott

<sup>1</sup> Pontosabban az NTSC és PAL kompozit jeleket alkotó chroma és luma komponensek digitalizálásával.

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

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) kizárólag 4:3 képarányt alkalmazott, azaz a horizontális és vertikális képhosszak aránya 1.33 volt<sup>3</sup>. Habár az 50-es években megjelentek az első szélesvásznú mozi formátumok, 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 ahhoz, 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 igény támaszt 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 lenne<sup>4</sup>, ezt a képfrissítési frekvenciát azonban az átlagos néző még villogónak érzékelné.

Ahhoz, hogy ezt elkerüljük, a képfrissítési frekvenciának tehát magasabbnak kell lennie a fúziós frekvenciánál. A fúziós frekvencia fényingerek változásának azon

<sup>3</sup> A 4:3 képarány létrejötté egészen Thomas Alva Edison munkájáig vezethető vissza, aki az általa használt 35 mm széles filmen egy képkockát 4 perforációnyi 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 perforációnyi képméretet 1909-ben fogadták el általános szabványnak ("4-perf negative pulldown"), lehetővé téve a szabványos mozikamerák, mozigépek és így a mozi térhódítását.

<sup>4</sup> Érdemes megjegyezni, hogy ez a képfrissítési frekvencia csak ahhoz 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 "kitalálni" az egyes képkockák közötti tartalmat. Érdekes tény azonban, hogy a néző szeme már kellően hozzászokott a mozi 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.

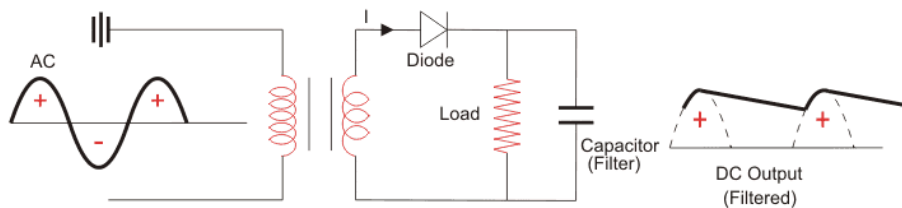


Figure 3.1. Periodikus hálózati brumm megjelenése az egyenirányított tápfeszültségen egyutas egyenirányítás esetén

frekvenciája, amely fölött a fényinger változását az emberi szem már nem képes követni. Különböző világosságú felületek váltakozása esetén a gyakorlatban előfordul 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 világosságsszintek közötti vibrálás érzékelhető) amplitúdójától, és a gerjesztés helytől a retinán: azaz, hogy a villogást a fő látóterünkben, vagy a periférikus látásunkkal érzékeljük-e.

Általánosan elmondható, hogy a fúziós frekvencia az embereknél 50-90 Hz közé esik: a fő látóterben, amelyet a csapok dominálnak a látás lassabb, itt a fúziós frekvencia 50 Hz, míg a periférikus látás jóval gyorsabb, itt a fúziós frekvencia magasabb. Mivel az NTSC bevezetésénél a cél a fő látóter 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 tökéletlenségéből származó periodikus zavarjel, ahogy az a 3.1 ábrán látható. A zavarjel frekvenciája a hálózati frekvenciával egyezik meg (egyutas egyenirányítás), vagy annak kétszerese (kétutas egyenirányítás esetén). Magyar elnevezése a hangerősítők kimenetén megszólaló jellemzően 50 Hz-es mélyfrekvenciás zugásból származik. Televízió esetében mivel ez a zavarjel közvetlenül hozzáadódik a katódsugárcső vezérlőjéhez, ezért a zavarjel kirajzolódik a kijelzőn, így látható hibát okoz.

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

$$Y(t) = \frac{1}{2} \sin 2\pi ft + \frac{1}{2}, \quad (3.1)$$

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 függvényében.

Jelölje a képfrekvenciát  $f_V$  (mint vertikális frekvencia), és a sorfrekvenciát  $f_H$

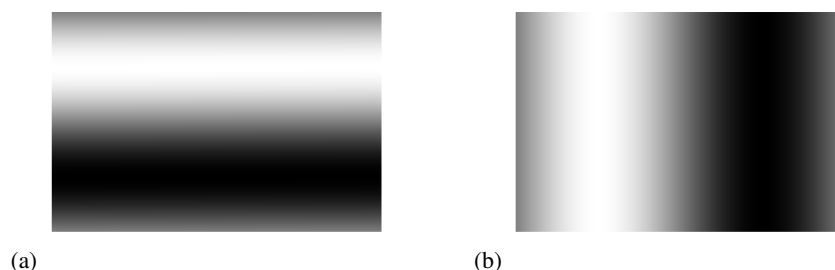


Figure 3.2. Periodikus jel képernyőn megjelenítve  $f = f_V$  (a) és  $f = f_H$  (b) választással

(mint horizontális frekvencia), köztünk természetesen fennáll az

$$f_H = N_V \cdot f_V \quad (3.2)$$

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 3.2 (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 horizontálishoz 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 tekinthető a tartalom. Így tehát a teljes képidő alatt egy álló, vertikális szinuszhullám jelenik meg a kijelzőn, ahogy az az 3.2 (b) ábrán látszik.
- $f > f_V$  választással a jelalak kezdőfázisa képről képre nő, így a hullámalak lassan felfelé mozog. 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. Márpedig a hálózati brumm épp ilyen periodikus zavarjelként jelenik meg a képernyőn, frekvenciája pedig az adott régió hálózati frekvenciája. A korai, fekete-fehér televíziós rendszer 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 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



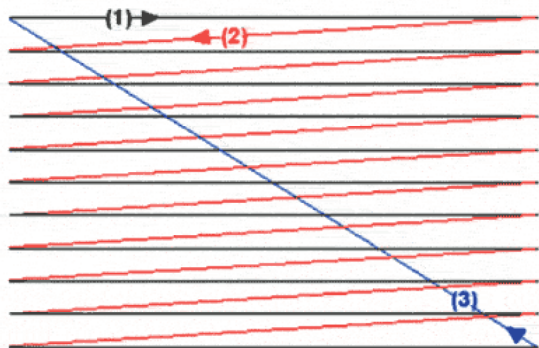


Figure 3.3. Progresszív letapogatás szemléltetése (az egyszerűség kedvéért 11 sorral ábrázolva), az aktív sortartalommal (1), sorvisszafutással (2) és képvisszafutással (3)

amerikai rendszerben a képfrekvencia értéke  $f_{V,USA} = 60 \text{ Hz}$ , az európai rendszerben  $f_{V,Eu} = 50 \text{ Hz}$  lett <sup>5</sup>.

#### Progresszív és váltott soros letapogatás:

A következőekben a különböző képernyő-letapogatási (scanning) módok vizsgáljuk. Az elnevezés a CRT kijelzőkhöz köthető, ahol a katódsugár ténylegesen végigpásztázta valamilyen trajektória mentén.

A legkézenfekvőbb képernyő bejárési 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 3.3 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 legegyszerűbb, legkézenfekvőbb megoldásnak, mégis, egészen az UHDTV szabvány megjelenéséig nem ez volt az általánosan elfogadott megoldás. Ennek okait a következőekben 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 képfrekvencia szükséges. Ez már bizonyos szintű tömörítést tesz lehetővé, hiszen a teljes képtartalom elegendő, ha lassabban változik, mint a kijelző rajzolási frekvenciája.

Ez a tömörítés már a korai mozitechnikában is megjelent: A korai, némafilmes

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



Figure 3.4. Triple blade shutter működése: <https://www.youtube.com/watch?v=jrSzRAch930>

korszakban számos képfrekvencia volt használatban 16 – 24 Hz között. Manapság mozi technikában a szabványos rögzítési frekvenciát 24 fps-re rögzítették. A villogás elkerüléséhez (tehát a képfrissítési frekvencia növeléséhez) speciális rekesszel látták el a vetítógépet. A fénynyaláb útjában forgó rekész, 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 vetítógép ugyanazt a képkockát kétszer, vagy háromszor villantja fel, mielőtt továbbhúzza a mozigép a szalagot. Ezzel az egyszerű trükkel 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-takarékosságra az ún. **váltott-soros letapogatást (interlaced scanning)** vezették be.

A megoldás alapötlete—ahogy a 3.5 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 mozis technikához hasonlóan, a képfrissítési frekvenciát elegendően magasra emelték, míg a 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

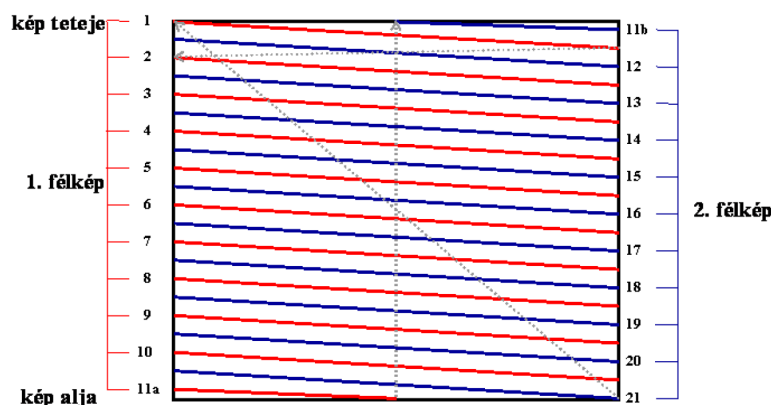


Figure 3.5. Változott-soros képbontás (TV sorok közbeszövése), a jobb áttekinthetőség kedvéért 21 sorral. A teljes képernyő pásztázásához (feltéve, hogy az elektronnaláb az első félkép első sorának elejétől indul) az első félképnek fél sorban kell végződnie, míg a második félképnek félsorral kell kezdődnie. 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)

kép páros-páratlan sorra való felbontásával). Ez eltérés a mozis rendszerhez képest, amely ugyanazt a képkockát mutatta be többször. Ennek eredményeként a következő módhatók 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 (sávszélesség-hatékonyság) mellett. Gyors kameramozgások esetén, pl. sporttartalom már láthatóvá válhatnak a felezett vertikális felbontásból származó hatások.

Az elmondottak alapján a normálfelbontású SD formátum kizárólag váltottsoros letapogatási módot alkalmaz. A HD szabvány bevezetésével már mind interlaced, mind progresszív formátumok léteznek, míg UHD TV esetén a szabványok már kizárólag progresszív formátumokat definiálnak.

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



Figure 3.6. Megfelelő deinterlacing technika nélkül váltottsoros formátum megjelenítése progresszív kijelzőn.

Egyik példaképp: korábban (előadáson) láhattuk, 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 jelenséghez 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 négyzettrácsos, 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 progresszívra 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 összeszőve alakít ki egy teljes felbontású képet. Ez azonban gyors mozgások esetén ún. fésűsödési jelenségekhez vezet, amelyet az 3.6 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, míg a HD műsorszórás még napjainkban is váltottsoros formátumot alkalmaz (jellemzően 1080i-t).

#### Analóg SD formátumok, az analóg videójel:

#### Térbeli felbontás, pixelméret:

### **3.1.2 HD formátumok**

Whitaker 592. oldal

### **3.1.3 UHD formátumok**

Periférikus látás: <https://www.quora.com/What-is-the-aspect-ratio-of-human-vision>