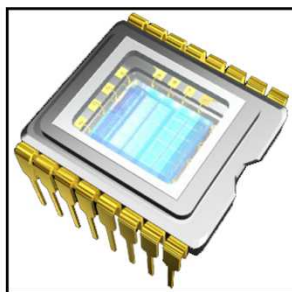
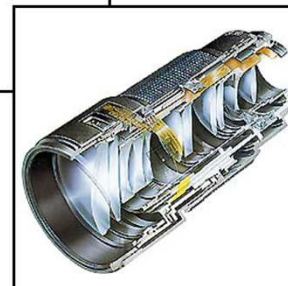


# Pořízení, komprese a segmentace videa

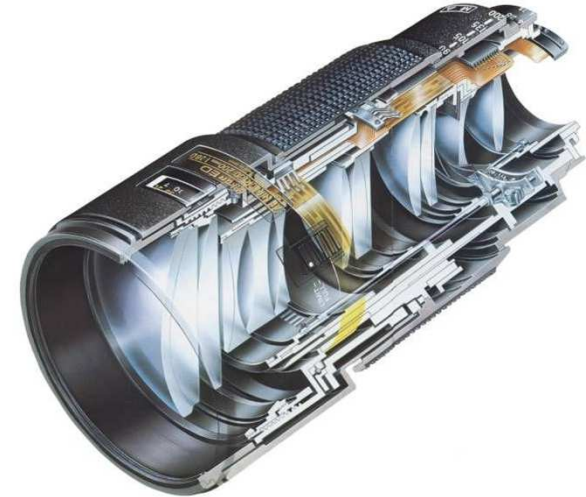


doc. Ing. Josef Chaloupka, Ph.D.



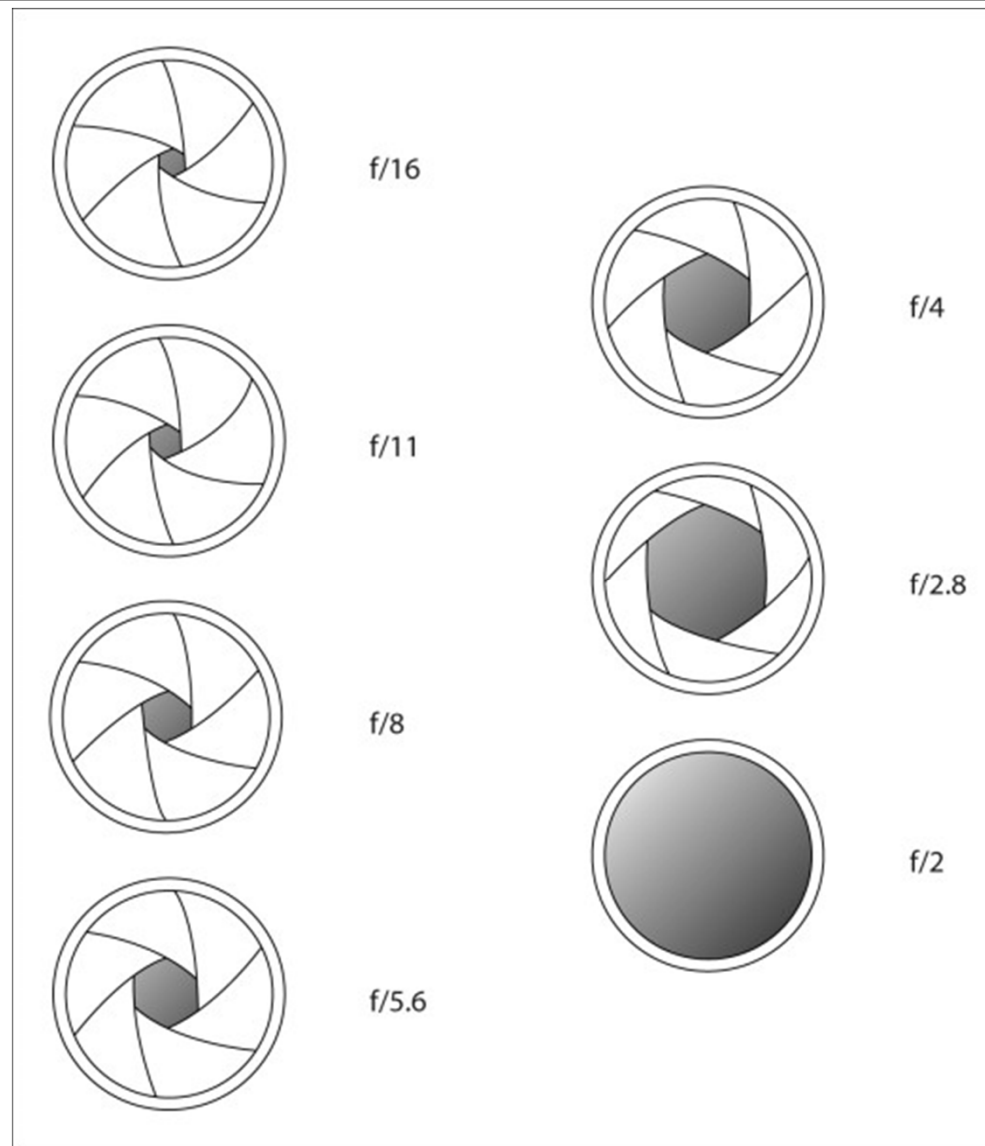
# OPTICKÁ ČÁST KAMERY

- Skutečné objektivy >>> **tlusté čočky**  
>>> přední a zadní ohnisko

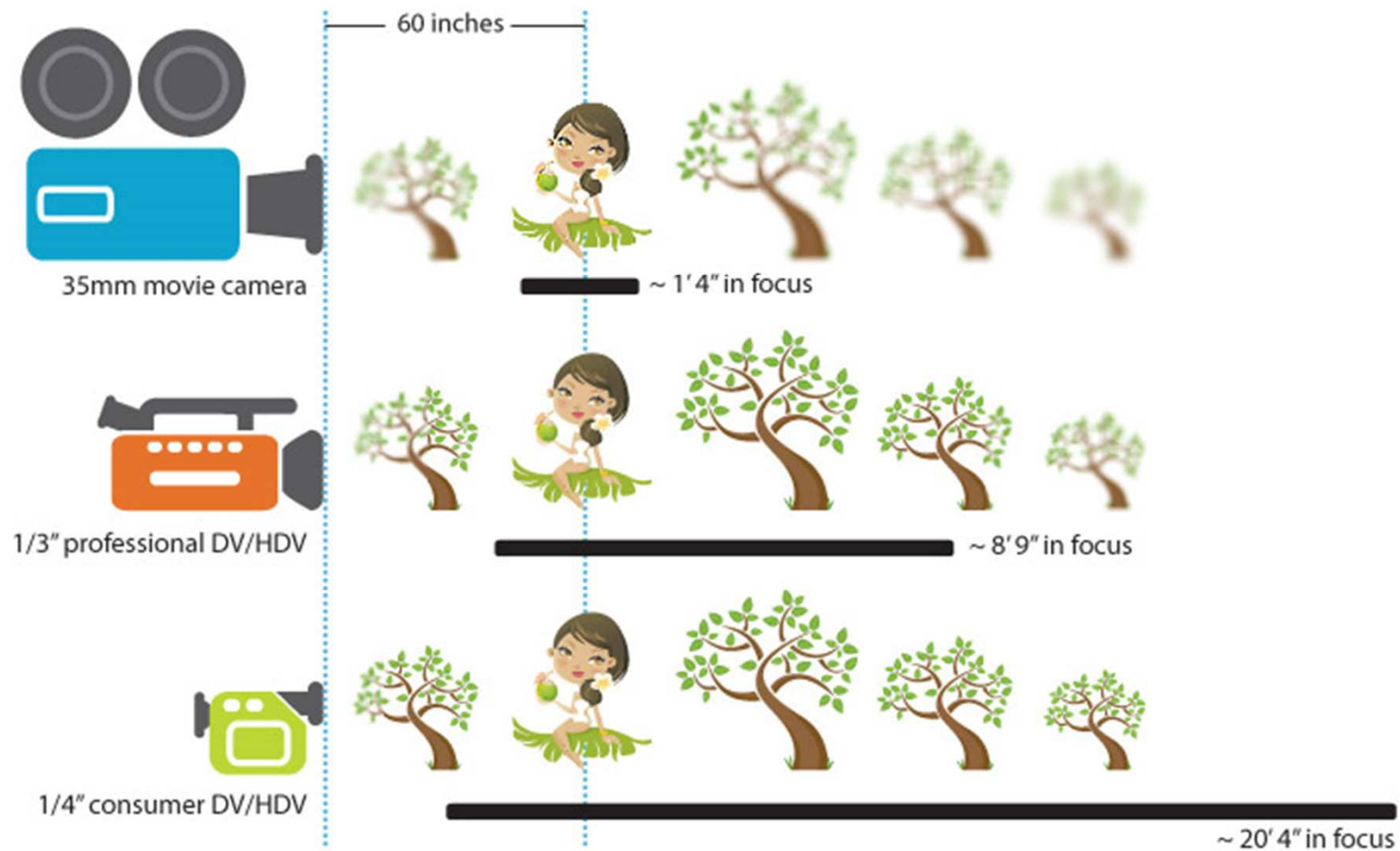


- **Optické vady** >>> z bodového předmětu nevznikne bodový obraz, distribuce intenzit v obraze – šířka měřicího stupně rozostření, posun středu distribuce (geometrické zkreslení)
- Průchod světla prostředím (vzduchem) >>> rozptyl >>> pokles vysokých frekvencí ve fourierovském spektru >>> kompenzace filtrem derivačního charakteru

# OPTICKÁ ČÁST KAMERY – clona kamery

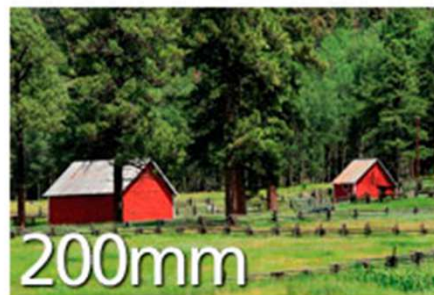


# OPTICKÁ ČÁST KAMERY – hloubka zaostření





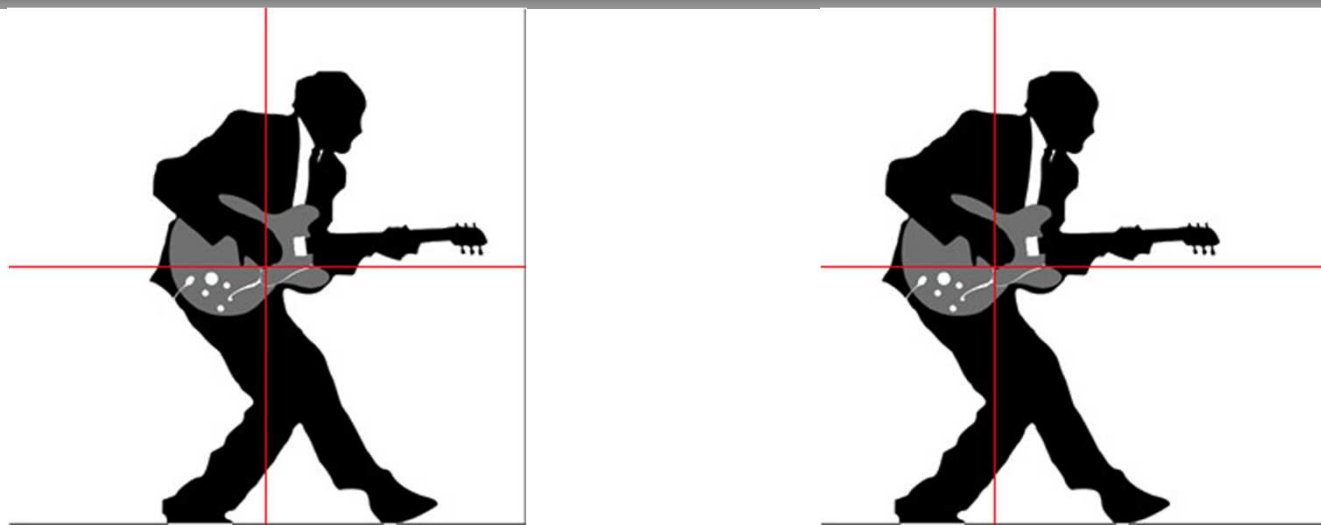
# OPTICKÁ ČÁST KAMERY – ohnisková vzdálenost



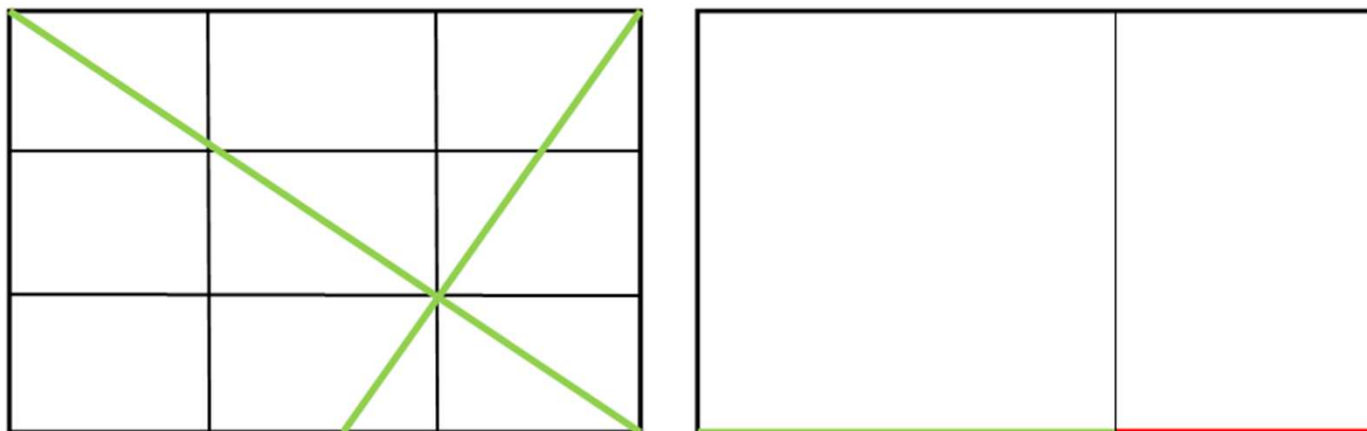
Rybí oko

Ohnisková vzdálenost (přepočteno na 35mm kinofilmové políčko): 4 – 6mm (rybí oko), 14 – 35mm (širokoúhlý objektiv), 50mm (základní objektiv), 60 – 300mm (dlouhé ohnisko) >300mm ultradlouhé ohnisko

# Kompozice obrazu - základ



Středová kompozice



Zlatý řez – poměr 1,618

Video: krátké stříhy (scénář), nepoužívat příliš zoom (+ pohyb), jízda, švenk

# Formáty video souborů

- Komprese? Video Full HD: 1920 x 1080 pixelů, 25 snímků za sekundu, RGB (24 bit.) – datový tok: 155,5 MBps, 560 GBpmin
- Komprese: ztrátová x bezztrátová
- Formáty videa:
- **VCD** (Video CD) >>> komprese MPEG-1, kvalita jako VHS, rozlišení 352x288 pixelů, datový tok >>> 1,15 Mbps, zvuk >>> stereo, 44,1 kHz, 16 bit., 1CD: 80 min.
- **VHS** >>> komprese MPEG-1, rozlišení 352x288 pixelů, datový tok >>> 2,4 Mbps, zvuk >>> stereo, 44,1 kHz, 16 bit.
- **SVCD** (Super VCD) >>> vylepšený VCD >>> komprese MPEG-2, rozlišení 480x576 pixelů, datový tok >>> 2,4 kbps, zvuk >>> stereo, 44,1 kHz, 16 bit.
- **DVD** >>> komprese MPEG-2, rozlišení 720x576 pixelů, datový tok >>> 6 Mbps, zvuk >>> stereo, 44,1 kHz, 16 bit.
- **MicroMV** >>> komprese MPEG-2, rozlišení 720x576 pixelů, datový tok >>> 12 Mbps, zvuk >>> stereo, 44,1 kHz, 16 bit., mini kazety Sony (menší než DV)

# Formáty video souborů

- **Ogg Theora**
  - Volný formát od sdružení Xiph.Org
  - Srovnatelný s formáty MPEG-4, RealVideo a Windows Media Video
  - VBR, využití DCT
- **Ogg Tarkin**
  - Využívá Diskrétní vlnkovou transformaci DWT, 3D signál – šířka, výška, čas
- **Ogg Dirac**
  - Od společnosti BBC jako základ, přenos video streamu po internetu
- **OggUVS (OggYUV, OggRGB)**
  - nekomprimované video



# Formáty video souborů

- **Windows Media Video**

- Od společnosti Microsoft
- Srovnatelný s formátem RealVideo
- Pro streamování videa na internetu, možnost vložení do kontejneru AVI nebo ASF (Advanced Systems Format), různé rozlišení 528x576, 640×480, 768x432, HD ready (1280 x 720), full HD (1920x1080)

- **Real Video**

- Od firmy RealNetwork (obdobně jako Real Audio)
- Pro streamování videa na internetu, multiplatformní

- **MOV formát**

- QuickTime aplikace, různé kodeky

# Formáty video souborů

- **AVI**

- Jeden z nejstarších formátů (kontejnerů) pro multimediální data
- Různé kodeky pro kompresi (ztrátové x bezztrátové), video může být i nekomprimované >>> pro přehrání avi souboru je nutné vlastnit příslušný kodek
- Individuální nastavování parametrů videa a zvuku: datový tok, rozlišení obrazu, vzorkovací frekvence a komprese zvuku...
- Časté využití kompresních algoritmů MPEG-4 (kodeky DivX, XviD...)

- **HD**

- HD ready (720p vertikálně)
- Full HD (1080i)
- Ultra HD (UHD) TV (3840 x 2160), kino (4096 x 2304)
- Ultra High Definition Television (UHDTV) 7680 x 4320

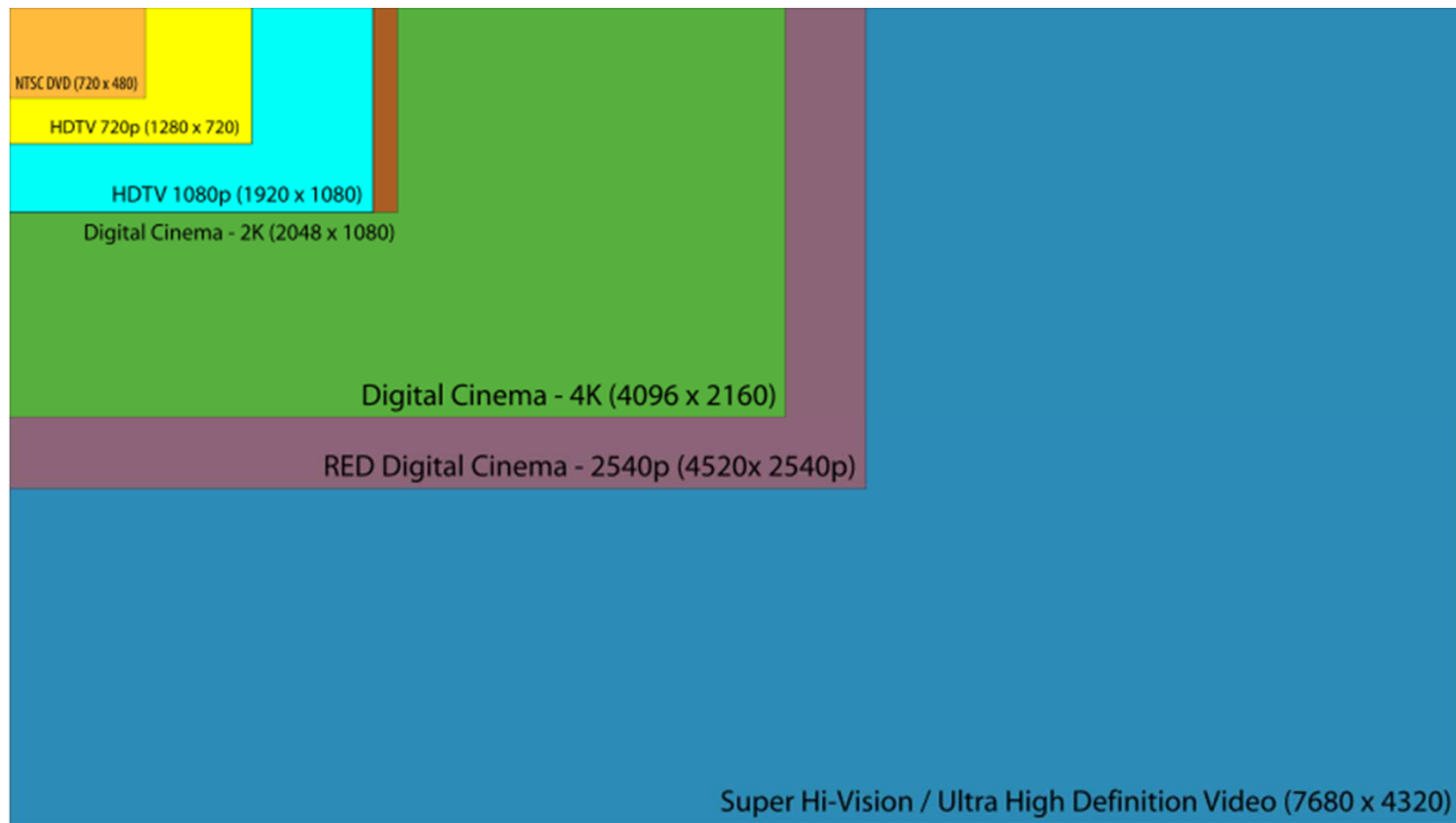
Sony FMP-X1



# TV / video – kvalita / rozlišení – normy

- h720p (1280x720p) poměr stran 16:9 – HR Ready
- h1080i (1440x1080i) poměr stran 4:3
- h1080p (1920x1080p) poměr stran 16:9 – Full HD
- 480w NTSC 720x480, poměr stran 4:3
- 576w PAL 720x576, poměr stran 4:3
- h720p60
- Parametry:
  - p - progresivní snímání (po celých snímcích)
  - i - prokládané snímání řádků
  - w - prokládané snímání řádků s potlačeným roztřepením okrajů
  - h - vysoké rozlišení
  - 720 - počet řádků v snímku
  - p60 - počet snímků za sekundu u HDTV 50, 60Hz

# TV / video – kvalita / rozlišení – normy



# TV / video – kvalita / rozlišení – normy

- DVB (Digital Video Broadcasting)

DVB-S (Satellite) satelitní vysílání

DVB-T (Terrestrial) pozemní vysílání

DVB-C (Cable) kabelové televize

DVB-H (Handhelds) pozemní vysílání pro přenosná zařízení – mobilní telefony apod..

- HDTV (High-definition television)

Digitální vysílání

16:9, Rozlišení h720p, h1080p, h1080i, ...

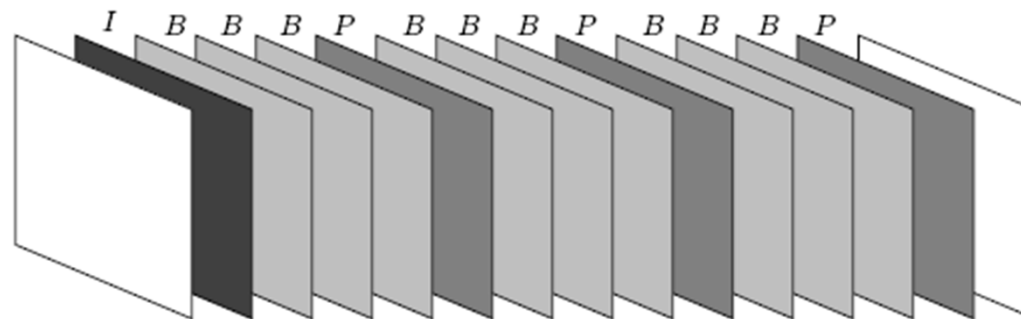
# Komprese videa

- Video >>> sled obrázků v čase, velký datový tok
- Většinou se používají ztrátové kompresní algoritmy
- Komprese jednotlivých obrazů nebo využití komprese za předpokladu podobné informace ve sledu snímků
- **Motion JPEG**
- Každý snímek zakódován jako JPEG, horší kompresní poměr než u kodeků využívajících sled snímků
- **MPEG (1-4)**
- Komprese na základě sledu snímků



# Komprese videa

## ● MPEG



- I (Intra frame) >>> key frame
- P (Predictive frame) >>> rozdíl snímku od předchozího snímku I nebo P
- B (Bi-directional frame) >>> rozdíl od průměru předchozího I nebo P snímku a následujícího I nebo P snímku
- Na počátku je nutné nalézt key frame I
- V jednom video střihu (vizuálním segmentu) se obraz příliš nemění >>> hledají se části obrazu, kde došlo k pohybu (výraznější změně)

# Kompresa videa

- **MPEG**

- Každý snímek rozdělen na bloky (8x8, 16x16 pixelů)

- V každém bloku (po dvou po sobě následujících snímcích) spočítán pohybový vektor

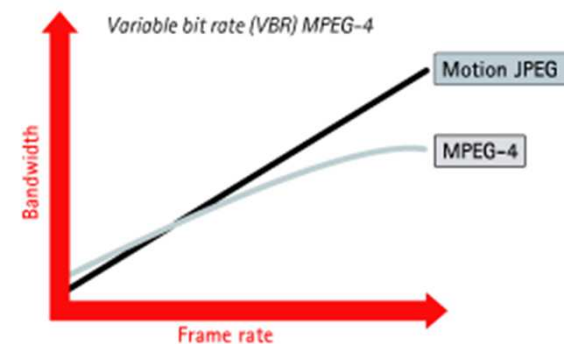
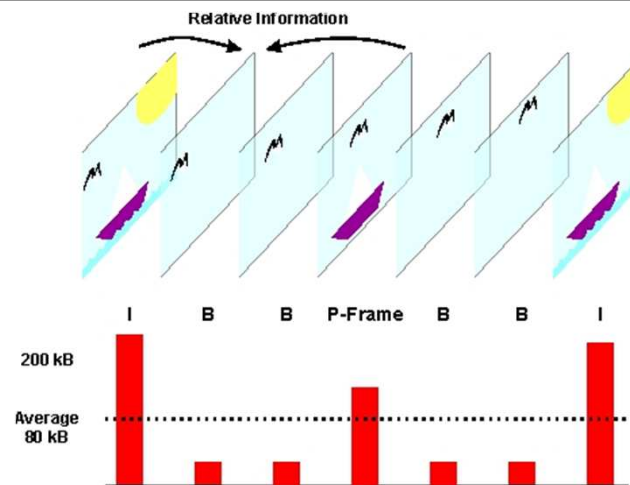
- Snímek I zabírá nejvíce místa, B nejméně

- MPEG-1 >>> rozlišení snímků 352 × 288, snímkovací frekvence 25fps, datový tok 1,5 Mbps

- MPEG-2 >>> využití VBR, lepší poměr kvalita x datový tok než MPEG-1

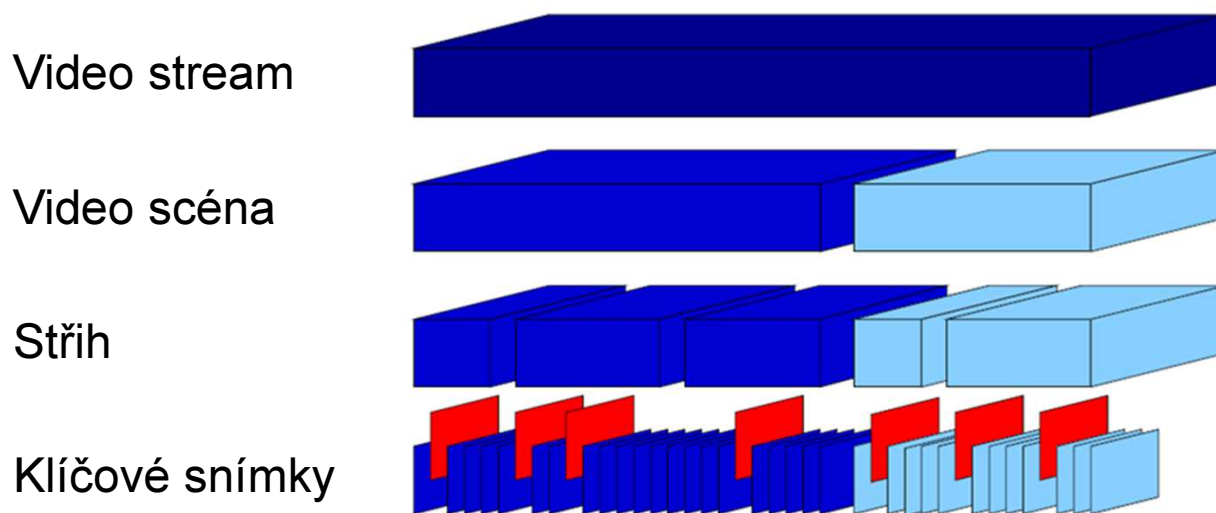
- MPEG-3 >>> původně pro HDTV, dnes součástí MPEG-2

- MPEG-4 >>> datové toky 10 kbps až 1 Mbps, u rychlých změn v obraze se zvyšuje datový tok, součástí kodeku práce s av syntézou...

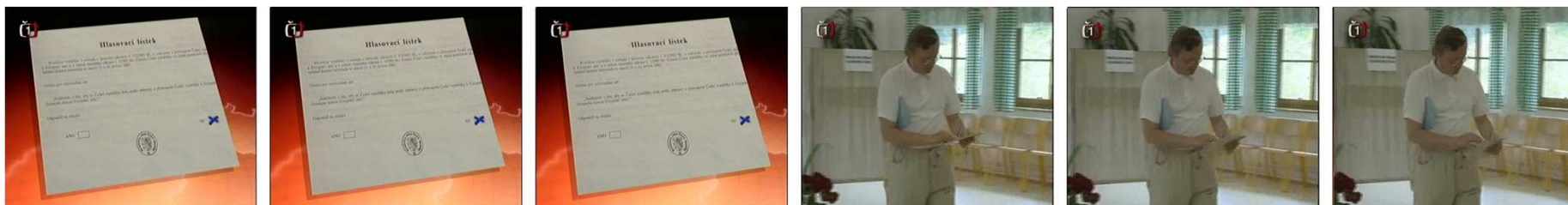


# Vizuální segmentace video nahrávek

- Detekce změn jednotlivých víceméně homogenních částí vizuálního streamu
- Homogenní část >>> sled snímků, ve kterých se příliš nemění vizuální informace
- Na místě detekované změny se nachází hranice segmentu standardně označovaná při zpracování videa jako střih
- Využití: pro indexaci audio-vizuálních dat (video nahrávek), předzpracování videa před motion-kompresí, pomocná informace při přepisu TV pořadů



# Vizuální segmentace – přechody – video efekty



ostrý střih (shot cut)



rozsvětlení (fade in)



ztmavení (fade out)

# Vizuální segmentace – přechody – video efekty



**průnik (dissolve)**



**clona (wipe)**



**změna scény pohybem (push to the left)**



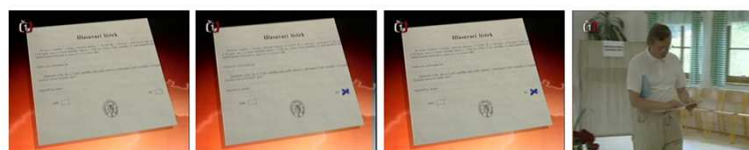
# Vizuální segmentace – metody

- **Metody založené na porovnání informace v obrazových bodech**

- Suma barevných hodnot pixelů ve dvou následujících snímcích

$$(1) \left| \sum_{x=1}^X \sum_{y=1}^Y f_i(x, y) - \sum_{x=1}^X \sum_{y=1}^Y f_{i+N}(x, y) \right| > T \quad (2) \sum_{x=1}^X \sum_{y=1}^Y |f_i(x, y) - f_{i+N}(x, y)| > T$$

- T – segmentační práh, N – posun snímků, obvykle 1



8606 (00:05:44) 8710 (00:05:48) 8814 (00:05:53) 8815 (00:05:53)



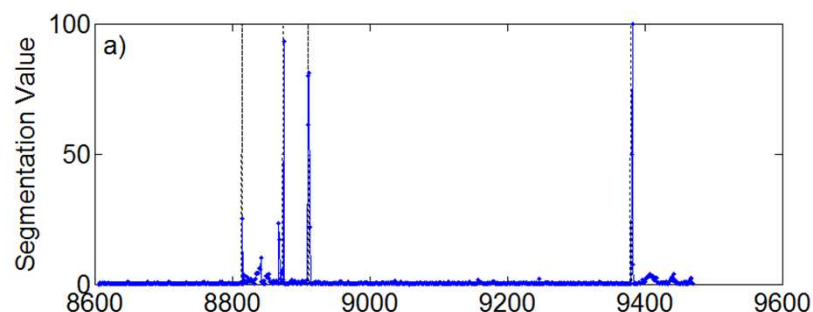
8844 (00:05:54) 8874 (00:05:55) 8875 (00:05:55) 8893 (00:05:56)



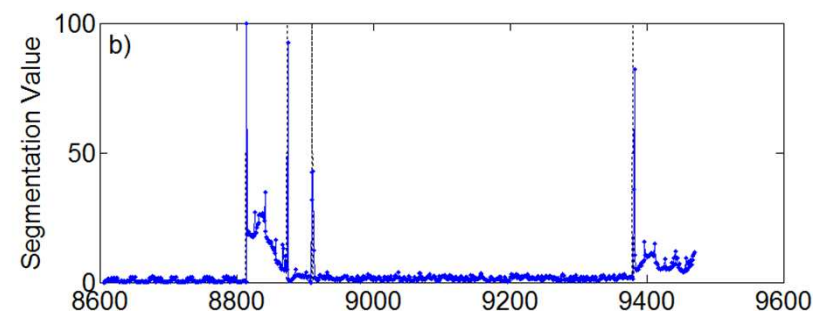
8909 (00:05:56) 8912 (00:05:56) 9146 (00:06:06) 9380 (00:06:15)

# frame (hodina:minuta:sekunda)

(1)



(2)



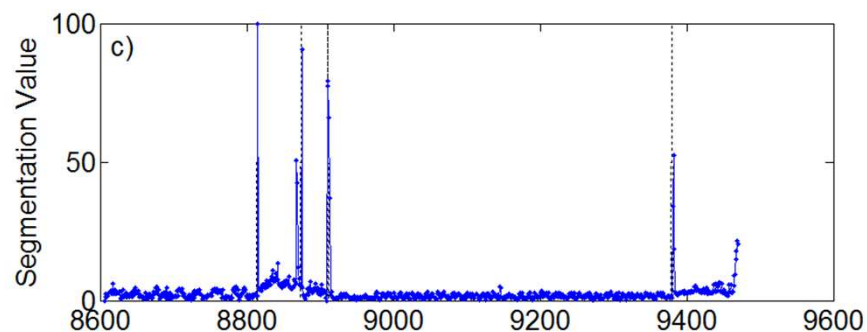
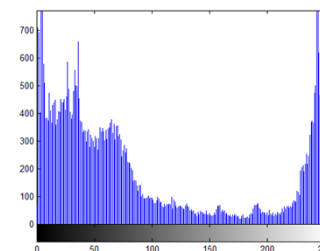
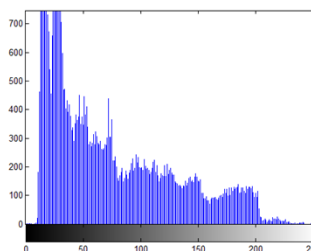
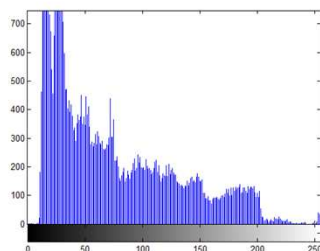
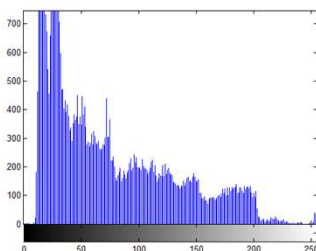


# Vizuální segmentace – metody

- Histogramově zaměřené metody**

- 1 obraz = 1 histogram, 1 histogram != 1 obraz

$$\sum_{v=1}^V |H_i(v) - H_{i+N}(v)| > T$$



# Vizuální segmentace – metody

- **Metody využívající obrazové příznaky**
- Video snímek charakterizován vektorem příznaků
- Vizuální příznaky: momenty počítané z obrazu, hrany, příznaky získané z barevných přechodů, příznaky získané na základě použití různých statistických metod a přístupů (střední hodnoty, rozptyly, kovarianční matice, Bayesův model)
- Jako příznaky lze také využít koeficienty z různých 2D transformací obrazu, např. parametry získané pomocí Houghovy transformace nebo diskrétní Fourierovy transformace DFT, vlnkové transformace DWT, diskrétní kosinové transformace DCT

# Vizuální segmentace – metody

- **Metody využívající obrazové příznaky – výpočet z DCT**

- 1) Převod snímku do barevné složky zvoleného barevného prostoru, 2) výpočet DCT:

$$F(u,v) = \frac{2c(u)c(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left(\frac{2x+1}{2N}u\pi\right) \cos\left(\frac{2y+1}{2N}v\pi\right) \quad c(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{pro } k = 0 \\ 1 & \text{jinde} \end{cases}$$

- 3) Z jednotlivých DCT koeficientů je následně spočítána mocnina:

$$E(u,v) = F(u,v)^2$$

- 4) P-nejvyšších umocněných DCT koeficientů je poté použito jako vizuální příznaky, které většinou dobře popisují charakteristickou informaci v daném video snímku.
- 5) Příznakový vektor je vhodné normalizovat, např. je počítán logaritmus z příznakového vektoru.
- 6) Pro dva po sobě následující snímky se poté počítá vzdálenost mezi jednotlivými vizuálními příznaky:

$$\sum_{p=1}^P |VP_i(p) - VP_{i+N}(p)| > T$$

# Vizuální segmentace – metody

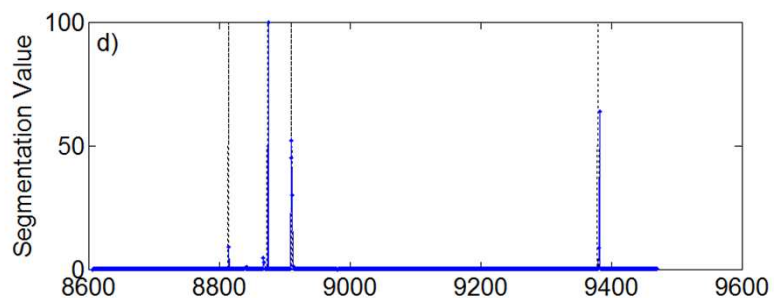
- Metody využívající obrazové příznaky – výpočet z DCT



vzdálenost = 0.274



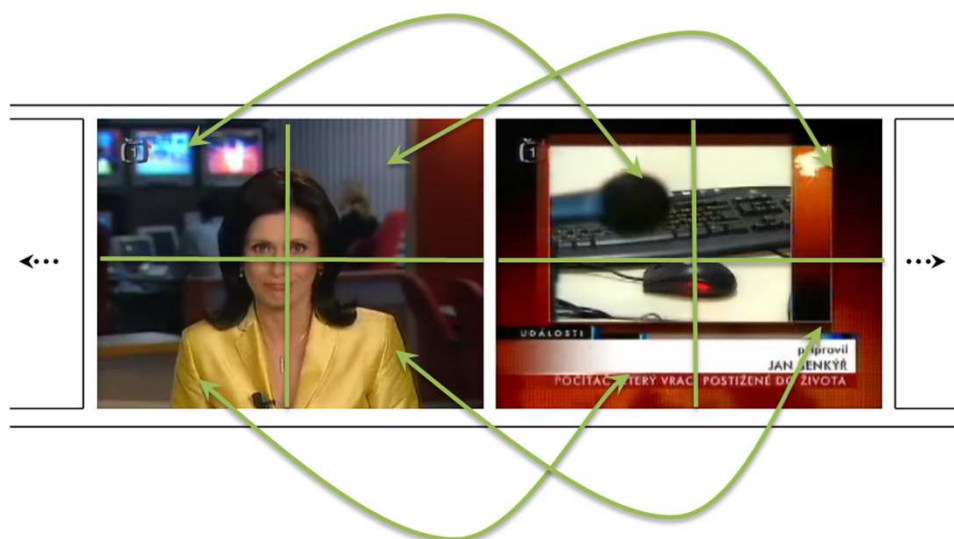
vzdálenost = 21.477



# Vizuální segmentace – metody

- **Blokově zaměřené metody**

- Obraz rozdělen na několik bloků (2x2, 4x4...), rozdíly jsou počítány ze stejných bloků dvou následujících snímků
- Bloky mohou být různě veliké, výpočet jenom z několika bloků, jednotlivým blokům lze nastavit různé váhy
- Pro zjištění rozdílu bloků: pixelově- , histogramově- , příznakově-zaměřené vizuálně segmentační metody



# Vizuální segmentace – metody

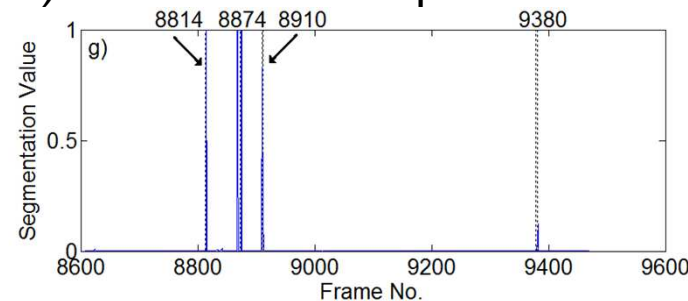
- Blokově zaměřené metody – využití statistických koeficientů**

- Výpočet statistických veličin v jednotlivých blocích, např. rozptyl a střední hodnota, pro dva korespondující bloky je pak počítána funkce  $L(i, b)$ :

$$L(i, b) = \frac{\left( (\sigma_{i,b}^2 + \sigma_{i-N,b}^2) / 2 + ((\mu_{i,b} - \mu_{i-N,b}) / 2)^2 \right)}{\sigma_{i,b}^2 \sigma_{i-N,b}^2}$$

- kde  $\sigma_{i,b}^2$  je rozptyl a  $\mu_{i,b}$  je střední hodnota barevných hodnot v jednotlivých blocích  $b$  v  $i$ -tém snímku
- Hodnota  $L(i, b)$  je následně porovnávána s určitým prahem  $T_b$ . Pokud je větší než tento práh, tak  $L(i, b) = 1$ , jinak  $L(i, b) = 0$ . Kritérium pro detekování hranice vizuálního segmentu je poté:

$$\sum_{b=1}^B w_b L(i, b) > T$$



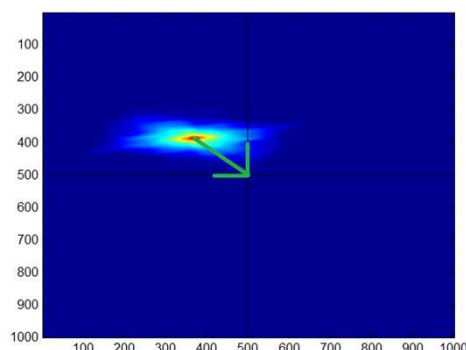
- kde  $B$  je počet bloků,  $T$  je práh a  $w_b$  je hodnota váhy pro jednotlivé bloky
- Nevýhodou u této metody je nutnost stanovení dvou prahů  $T$  a  $T_b$ , kritérium



# Vizuální segmentace – metody

## ● Další metody

- Kombinace předchozích metod, urychlení >>> segmentace hrubá (N následujících snímků) a jemná v oblastech změn
- Metody vycházející ze změny pohybu >>> 1) rozdělení obrazu na bloky, 2) výpočet pohybových vektorů v každém bloku (v dvou následujících snímcích), 3) analýza pohybových vektorů

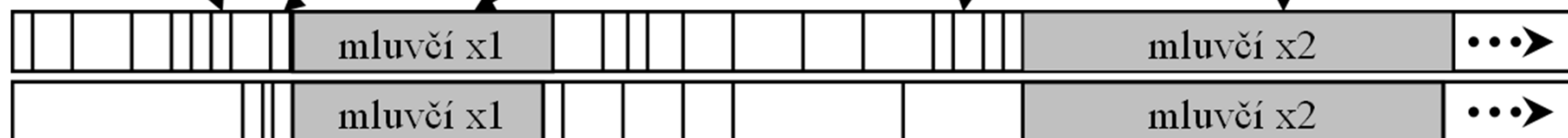


# Vizuální segmentace TV nahrávek

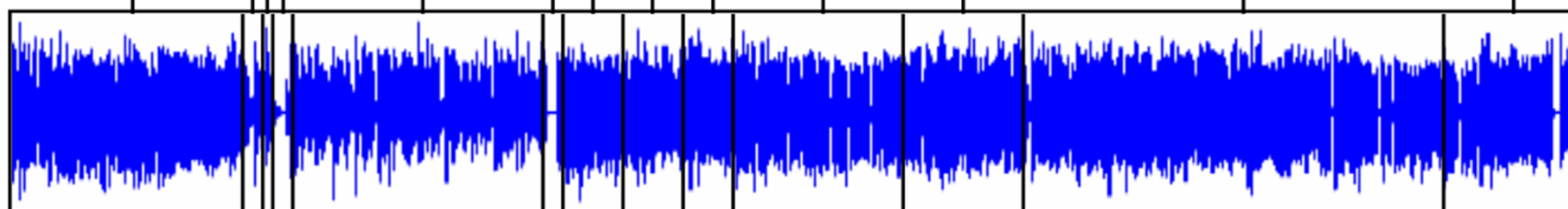
## Vizuální signál



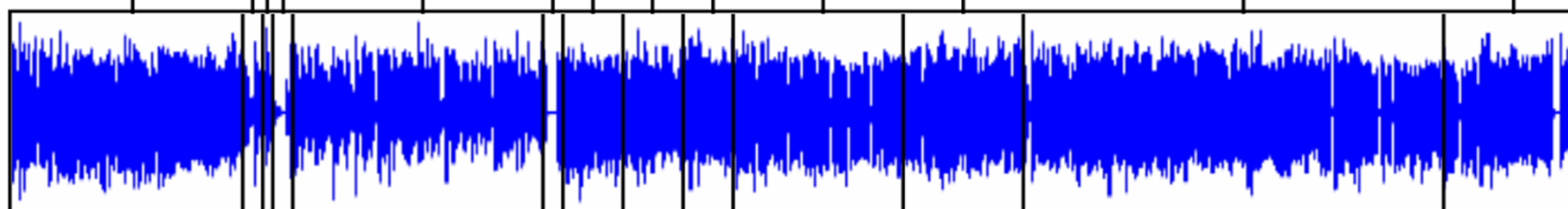
## Vizuální segmenty



## Audio segmenty



## Audio signál



# Vizuální segmentace TV nahrávek

