

Multimediální systémy

Jan Outrata



KATEDRA INFORMATIKY
UNIVERZITA PALACKÉHO V OLMOUCI

přednášky



Reprezentace a formáty

Reprezentace

- = 2D pole obrazových bodů (**pixelů**) – atribut **rozlišení** v počtu pixelů na jednotku délky = **velikost pixelu**
- atributy **šířka** a **výška** v počtu pixelů a **hloubka** v bitech na pixel
- **aspect ratio** = poměr šířky a výšky, např. 3:2 (fotografie), 4:3 (TV), 16:9 (širokoúhlý, HD), 47:20 (anamorphic, kino), změna mění velikost pixelu (tzv. pixel aspect ratio, typicky čtvercové 1:1) – vnímáno jako roztažení/stlačení obrazu → převzorkování
- přechody mezi aspect ratio: **pan/scan** = oříznutí bočních okrajů, např. z 16:9 na 4:3, nebo horních okrajů, např. z 4:3 na 16:9, **letterbox** = přidání prázdné (černé) plochy nad a pod, např. z 16:9 na 4:3, nebo na strany, např. z 4:3 na 16:9

Obrázek: Ilustrace změny aspect ratio



- hloubka závisí na reprezentaci barev (barevného prostoru) – rozdělena do **barevných kanálů**, např. pro odstíny šedé 1 kanál 8 b = hloubka 8 b, pro barevnou RGB 3 kanály R (červená), G (zelená), B (modrá) po 3, 3, 2 b = hloubka 8 b (8-bit color), 5, 5/6, 5 b = hloubka 15/16 b (tzv. high color), 8, 8, 8 b = hloubka 24 b (tzv. true color) nebo více bitů na kanál (tzv. deep color)
- **alfa kanál** – míra průhlednosti pixelu při skládání obrazů (z vrstev), 0 = neprůhledný, např. 8 b (true color) nebo 1 b (high color), proměnné hodnoty např. na hranicích složených obrazů pro potlačení aliasingu

Obrázek: Využití alfa kanálu při skládání obrazů

- tzv. **halftone obraz** – technika simulace barvy nebo plynulého přechodu mezi barvami pomocí pravidelného rozmístění bodů jiných barev a proměnné velikosti místo pixelů, velikost v počtu bodech na jednotku délky, používané při tisku, barevné kanály C (azurová), M (fialová), Y (žlutá), K (černá)

Obrázek: Halftone obraz



Background

+

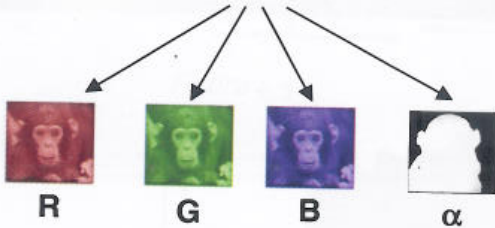


Foreground

=



Final



$$Final[i][j] = Foreground[i][j] * \alpha [i][j] + Background[i][j] * (1 - \alpha [i][j])$$



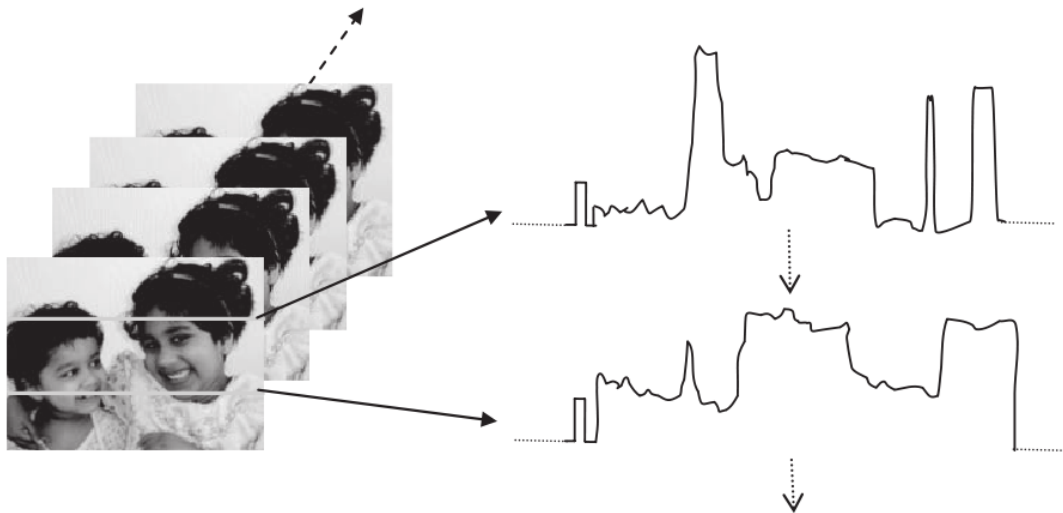
Formáty

- nekomprimované (**raw**) např. Tagged Image File Format (TIFF, i komprimovaný), Windows bitmap (BMP, hloubka až 24 b), Adobe Photoshop (PSD, více barevných modelů, vrstvy), proprietární např. z fotoaparátu
- komprimované např. JPEG (ne alfa), Portable Network Graphics (PNG, hloubka až 48 b, alfa), Graphics Interchange Format (GIF, 256 barev z palety, alfa, LZW komprese)

Reprezentace

- = posloupnost obrazů, tzv. **snímků (frame)**, se stejnými atributy, zobrazovaných rychle po sobě
- **frame rate** = rychlost zobrazování obrazů v obrazech/s (frames per second, fps), např. 24 (film), 25/30 (TV, PAL/NTSC), při nedostatečné vnímáno „blikání“ (flickering)
 - z historických důvodů jsou standardy digitálního videa těsně svázány se standardy videa analogového – přenos a zobrazování z analogového signálu na digitálních zařízeních nebo digitálního na analogových
- převod na jednorozměrný signál jako funkci hodnot bodů (pixelů) řádků obrazů v čase, tzv. **skenování (scanning)** – kvůli distribuci (přenosu) analogové TV a zobrazování na analogových CRT obrazovkách

Obrázek: Analogový video signál

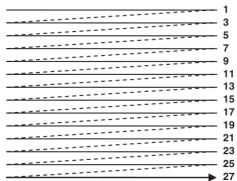


- **skenovací formát** – body na řádku zleva doprava, řádky shora dolů, např. prokládaný, progresivní
- **prokládané (interlaced)** skenování: snímek skenován (a zobrazován) po dvou částech, tzv. **půlsnícímích** = liché a sudé řádky – pohybové artefakty, pokud jsou půlsnímky složeny a zobrazeny zároveň, progresivně (např. na digitálním zařízení)

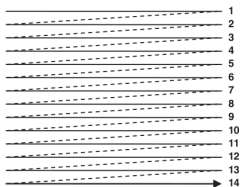
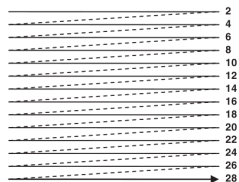
Obrázek: Prokládané skenování

- **progresivní** skenování: snímek skenován (a zobrazován) celý, tzn. všechny řádky

Obrázek: Progresivní skenování



Upper field

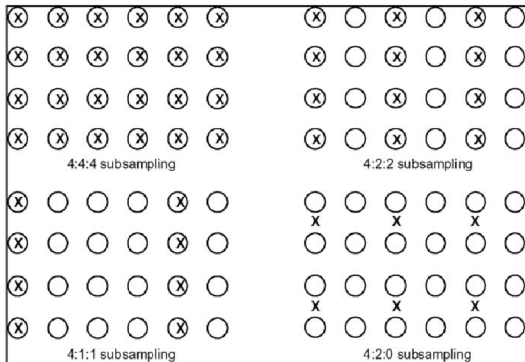
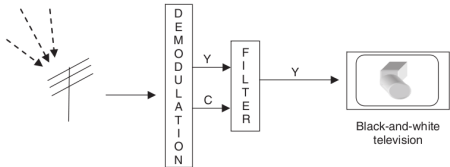
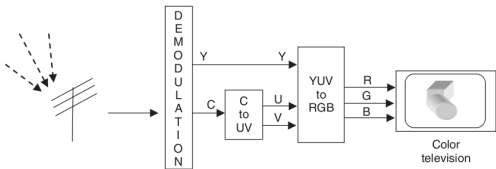
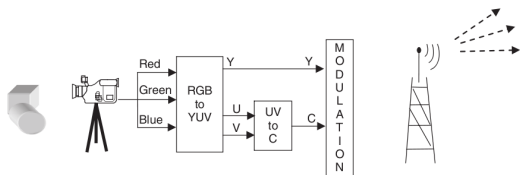


- kvůli přenosu TV signálu (synchronizaci a omezenému pásmu): prokládané skenování (synchronizační impulsy, vertikální pro další snímek 50/60-krát/s) a převod barevného prostoru RGB na YUV s podvzorkováním

Obrázek: Přenos TV signálu

- **barevný prostor YUV**: jasová Y a barevné U, V (C) složky, stejné hloubky jako u RGB
- **YUV podvzorkování** = podvzorkování barevných U, V složek pro redukci pásma, redukce barevného rozlišení méně ovlivňuje vnímání obrazu než redukce jasového rozlišení
 - u analogového signálu: polovina pásma pro jasovou složku
 - u digitálního signálu: schémata uložení pro ne každý pixel, např. 4:2:2 pro každý druhý, 4:1:1 každý čtvrtý na řádku, 4:2:0 průměr pro oblast 2×2 pixelů, 4:4:4 bez podvzorkování

Obrázek: Schémata YUV podvzorkování



- typy signálu:
 - **kompozitní (baseband, RCA)** – kombinace jasové a barevných složek do jednoho signálu, pro redukci pásma, problém interference při slabé úrovni signálu (fluktuace jasu a barev, falešné barvy), na analogových zařízeních (koaxiální vstup/výstup)
 - **S-video (Super, Y/C)** a **komponentní** – samostatné signály pro jasovou a barevné složky, na digitálních zařízeních

Formáty

- analogové: PAL, SECAM, NTSC

Obrázek: Analogové video formáty

- digitální (**Digital Video, DV**): komprimované CIF (Common Interchange Format), SIF, CCIR 601, HDTV (High Definition TV), Digital cinema
 - kompatibilita s analogovou TV – podpora prokládání
- oproti analogové TV vysoká rozlišení – **HDTV**: 720p = 1280×720 progresivní, 1080i/p = 1920×1080 prokládané/progresivní
- komprimované např. ITU H.26x , ISO MPEG x, HEVC, Theora, Google VPx, Windows Media Video (WMV)

Obrázek: Digitální video formáty

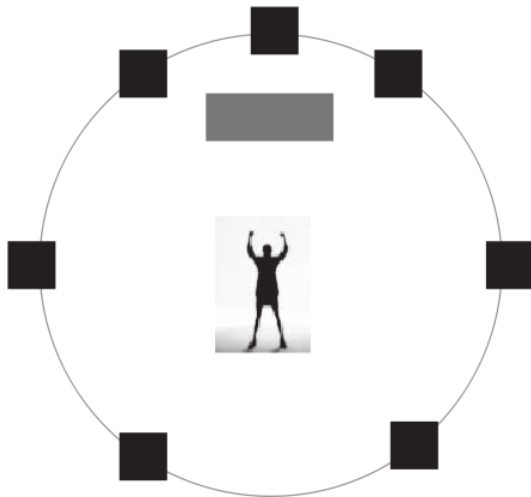
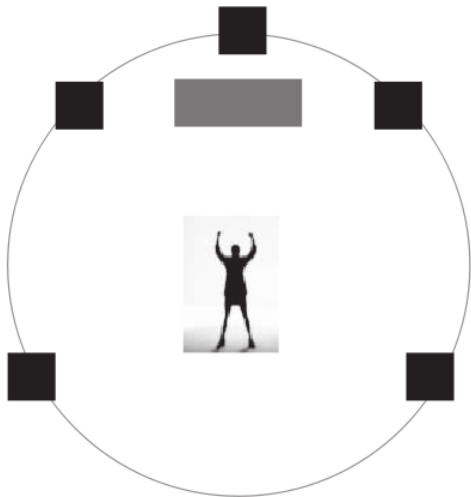
Property	NTSC	PAL	SECAM
Frame rate	30	25	25
Number of scan lines	525	625	625
Number of active lines	480	576	576
Aspect ratio	4:3	4:3	4:3
Color model	YIQ	YUV	YDbDr
Primary area of usage	North America (USA and Canada), Japan	Asia	France and Russia

Format name	Lines per frame	Pixels per line	Frames per second	Support for interlaced format	Subsampling scheme	Image aspect ratio
CIF	288	352		N	4:2:0	4:3
QCIF	144	176		N	4:2:0	4:3
SQCIF	96	128		N	4:2:0	4:3
4CIF	576	704		N	4:2:0	4:3
SIF-525	240	352	30	N	4:2:0	4:3
SIF-625	288	352	25	N	4:2:0	4:3
CCIR 601 NTSC (DV, DVB, DTV)	480	720	29.97	Y	4:2:2	4:3
CCIR 601 PAL/SECAM	576	720	25	Y	4:2:0	4:3
EDTV (576p)	480/576	720	29.97	N	4:2:0	4:3/16:9
HDTV (720p)	720	1280	59.94	N	4:2:0	16:9
HDTV (1080i)	1080	1920	29.97	Y	4:2:0	16:9
HDTV (1080p)	1080	1920	29.97	N	4:2:0	16:9
Digital cinema (2K)	1080	2048	24	N	4:4:4	47:20
Digital cinema (4K)	2160	4096	24	N	4:4:4	47:20

Reprezentace

- = posloupnost úrovní analogového zvukového signálu (**vzorků**)
- digitalizace (vzorkování + kvantizace) = **pulsní kódová modulace (PCM)**
- **frekvence vzorkování** v počtu vzorků/s (Hz), **velikost vzorku** (počet kvantizačních úrovní) v bitech, rozložení kvantizačních úrovní
- **dimenze (počet kanálů)**: mono (1), stereo (2, binaural, 1920), vícekanálové (> 2, stereophonic, 1940)
- **surround sound** = vícekanálový prostorový zvuk, posluchač obklopen reproduktory, např. 5.1 = přední středový (dialog), levý a pravý (soundtrack), boční/zadní levý a pravý (efekty), subwoofer (efekty na nízkých frekvencích, LFE, kdekoliv), 7.1 = zadní levý a pravý navíc

Obrázek: Rozmístění reproduktorů pro surround sound



- **spatial audio** („virtual surround sound“) = iluze vícekanálového prostorového zvuku s méně kanály (typicky dvěma), psychoakustické efekty pro simulaci lokalizace zvuku (posunutí kanálů vůči sobě využívající tzv. HRTF funkcí, dozvuků apod.)

Formáty

- nekomprimované např. WAV (Audio CD), G.711 (A/ μ -law, telefon, logaritmická kvantizace), G.729 (IP telefonie)
- komprimované např. MPEG 1 Layer III (MP3, psychoakustika pro kompresi), MPEG 4 Advanced Audio Codec (AAC), Dolby Digital (AC-3), Vorbis, Windows Media Audio (WMA), DTS, Speex (řeč), Opus, GSM 06.10 (AMR, mobily, komprese predikcí), Free Lossless Audio Codec (FLAC, bezztrátová)
- spolu s videem (a titulky apod.) uloženo pomocí tzv. **kontejnerů**, např. Audio Video Interleave (AVI), MPEG (PS, TS), MP4, Ogg, Matroska, WebM, QuickTime, FLV, 3GP (mobily)

Reprezentace

- = množina 2D a 3D **geometrických objektů** (body, čáry, křivky, polygony, plochy) popsaných **vektory**
- atributy délka, typ, tloušťka, barva, zakřivení atd. čáry, barva výplně apod.
 - neomezené rozlišení, ale pro zobrazení **rasterizování** = převod na rastrový obraz

Animace

- = posloupnost transformované grafiky zobrazované rychle po sobě
- **transformace** posunu, rotace, změna měřítka, zkosení (afinní) aj. – reprezentace vektorů a matic transformací pomocí homogenních souřadnic

Obrázek: Transformace pomocí homogenních souřadnic

Translations:

$$\begin{bmatrix} x_i + t_x \\ y_i + t_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

Rotations:

$$\begin{bmatrix} x_i \cos\alpha - y_i \sin\alpha \\ x_i \sin\alpha + y_i \cos\alpha \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

Scaling:

$$\begin{bmatrix} s_x x_i \\ s_y y_i \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}.$$

Formáty

- = vektorové – popis objektů, textové, binární nebo kombinované
- kombinované s rastrovým obrazem (tzv. **metafile**) – platformově nezávislé, např. Computer Graphics Metafile (CGM), WebCGM, Scalable Vector Graphics (SVG), (Encapsulated) Postscript ((E)PS, nekomprimovaný), Adobe Portable Document Format (PDF, komprimovaný)

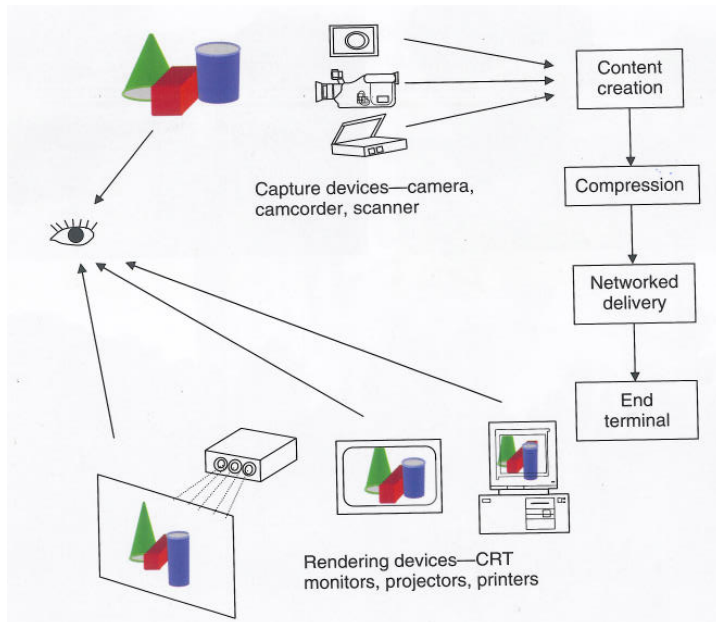


Barvy

- barvy jsou součástí obrazu, videa, grafiky
- = barvy zobrazené pomocí zobrazovacích zařízení (TV, počítače, projektory, tiskárny) by měly „**vypadat stejně**“ (= **být vnímány stejně**) jako barvy vzoru zaznamenaného pomocí záznamových zařízení (fotoaparáty, skenery, kamery), nezávisle na zařízeních

Obrázek: Ilustrace problému barev

- *co je barva?*: fyzikální veličina vs. smyslové vnímání okem interpretované mozkiem, **teorie barev** (17. st, Newton, separace světla na barvy duhy pomocí hranolu a složení zpět, později frekvence elektromagnetického záření)

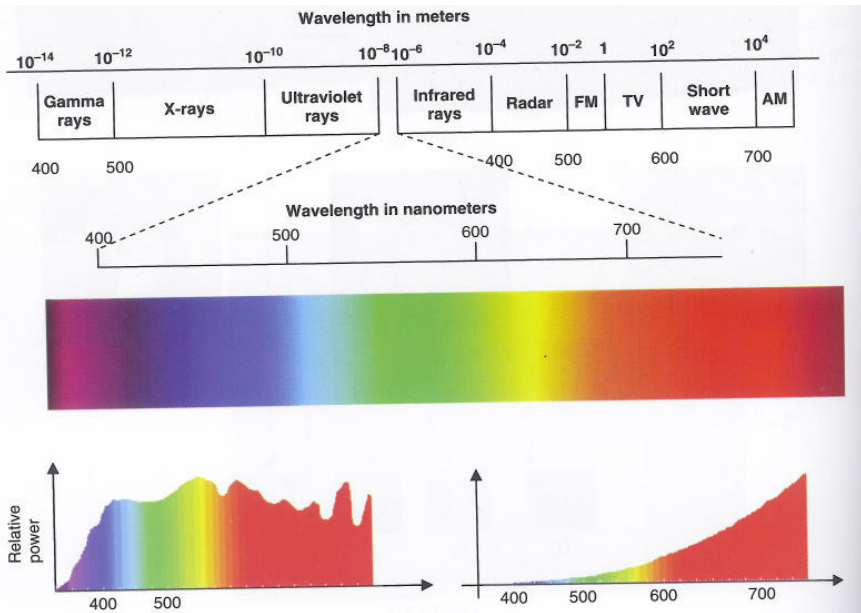


- = **kombinace frekvencí**, s různými amplitudami (energiemi), z viditelného spektra elektromagnetického vlnění (**světla**) s vlnovou délkou od 400 do 700 nm, < 400 = ultrafialové a > 700 = infračervené neviditelné
- různé zdroje světla mají různá **spektrální rozložení** (**spektrum**, „barva světla“), např. bílé (denní), žluté (žárovkové), červené (laser)

Obrázek: Viditelné spektrum světla a spektra zdrojů světla

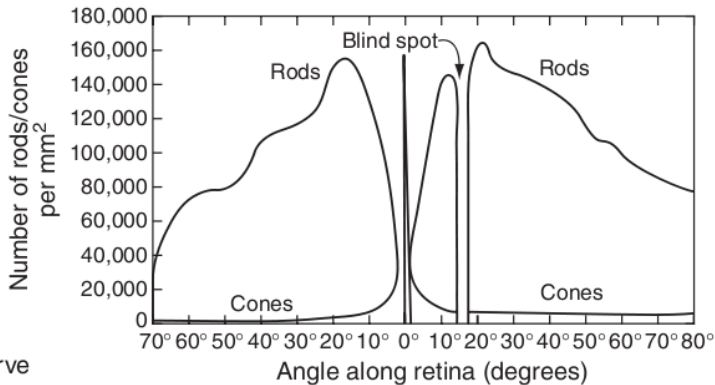
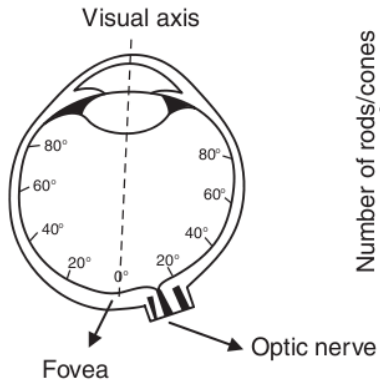
- **barva světla** = vnímaná kombinace frekvencí ze spektra světla
- interakce světla s materiály vzoru a obrazu – odrazy, absorpce, lomy → odražené světlo (kombinace frekvencí ze spektra) vnímané jako **barva materiálu**

Barva jako fyzikální veličina



- Young-Helmholtz, 1802, 1850, Maxwell
- = **trojité selektivní**: oko má tři druhy barevných receptorů (**trichromacita**), pro **červenou (R)**, **zelenou (G)** a **modrou (B) primární barvu** (absorpci spektra) – potvrzeno 1965:
- receptory oka (na sítnici): tyčinky pro jas (citlivější, větší počet, uniforměji rozložené) a čípky pro barvu (nejvíce u ohniska) – tři druhy, „modré“ (mimo ohnisko) nejcitlivější

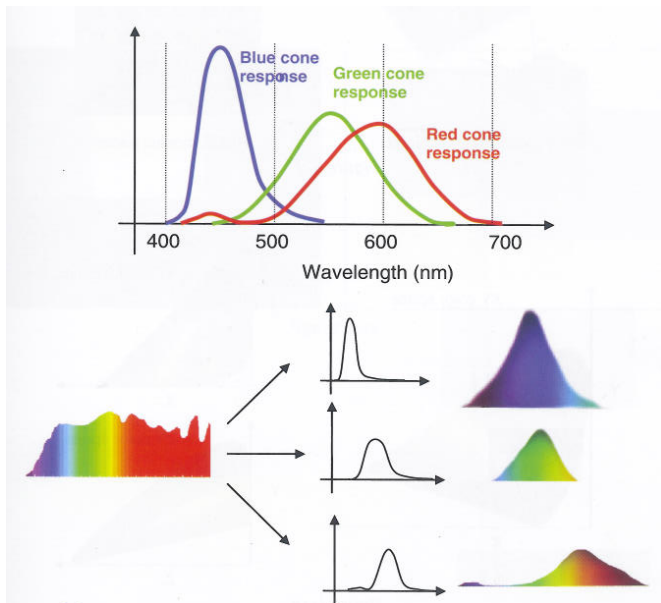
Obrázek: Sítnice oka a distribuce tyčinek a čípků



- citlivosti receptorů (tzv. **color-matching funkce**): funkce $s_i(\lambda)$, $i = 1, 2, 3$

Obrázek: Citlivost barevných receptorů oka

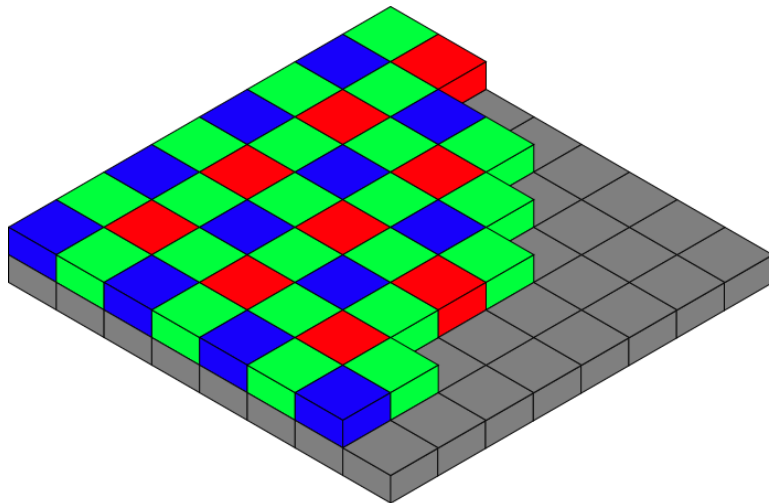
- vnímaná intenzita spektra: $c_i = \int_{\lambda=400}^{\lambda=700} s_i(\lambda) f(\lambda) d\lambda$, $f(\lambda)$ je spektrum světla, nebo digitalizované $c_i = \sum_{j=400}^{j=700} s_i(\lambda_j) f(\lambda_j)$, vektorově $c = S^T f$, kde S je matice se sloupci s_i
 - vnímaná barva = (c_1, c_2, c_3) – tzv. **tristimulus vektor**
 - **řešení problému barev**: pro stejné vnímání barev ze dvou spekter f a g , tj. $c_f = S^T f = S^T g = c_g$, nemusí být nutně $f = g$
- ⇒ jestliže záznamové zařízení zaznamenává f a zobrazovací zařízení zobrazuje g , g nemusí být rovno f , stačí takové, aby bylo $S^T f = S^T g \rightarrow$ **kalibrace zařízení**



- zaznamenaná barva (pixelu, digitalizovaná): $a = M^T f$, kde M je matice se sloupci m_i , charakterizující **odevzu barevných filtrů** pro primární barvy, více pro zelenou pro vyšší citlivost (řešené jako tzv. Bayer filtr)

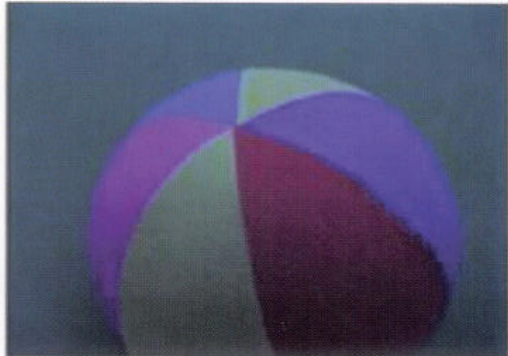
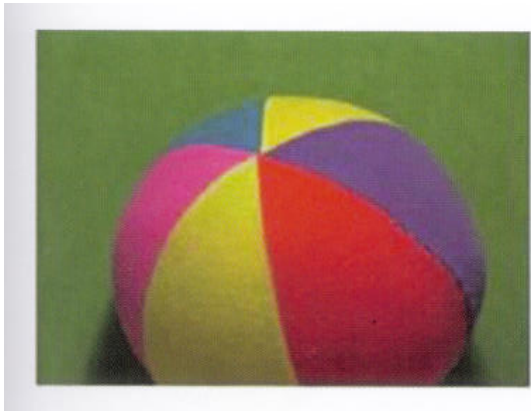
Obrázek: Bayer filtr

- zobrazené spektrum (pixelu): $g = vp$, kde p je vektor frekvencí **zobrazovaných primárních barev** a v je vektor zobrazované barvy
- stejné vnímání zaznamenané a zobrazované barvy: $S^T f = S^T g = S^T ap$,
 $M^T f = a = (S^T p)^{-1} S^T f$ – zobrazované primární barvy p a záznamové barevné filtry m_i musí korespondovat
- vztah není jednoznačně definovaný → CIE definované **standardní primární barvy**
 $p = (700, 546.1, 435.8)nm \Rightarrow$ výpočet m_i pro neznámou S – člověk ji „provádí“
- postupem **porovnávání vybraných skutečných a zobrazených barev** ze spektra f a úprav a lze vypočítat filtry m_i pro automatické úpravy a pro stejné vnímání ostatních zobrazovaných barev



- **gamma** CRT zobrazovacího zařízení: faktor (exponenciálního) zesílení (jasu) spektra zobrazované barvy – **gamma korekce**, $g = \left(\frac{v}{v_{max}}\right)^\gamma p$, např. standardní 2.2, LCD (a Plasmy) mají vyhledávací tabulku hodnot gamma pro primární barvy
- **zbývající problém**: různé světelné podmínky („barva“ světla) při záznamu a při zobrazování
- **stálost barev** = mozek automaticky kompenzuje „barvu“ světla (!) – např. bílý papír je vnímám bílý za různých světel, např. fotoaparát nebo kameru je potřeba nastavit \Rightarrow vnímání barev není lokální

Obrázek: Stálost barev



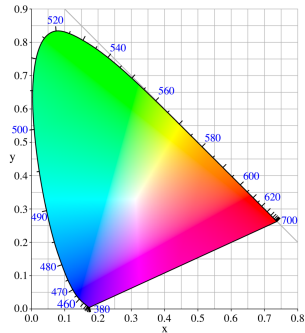
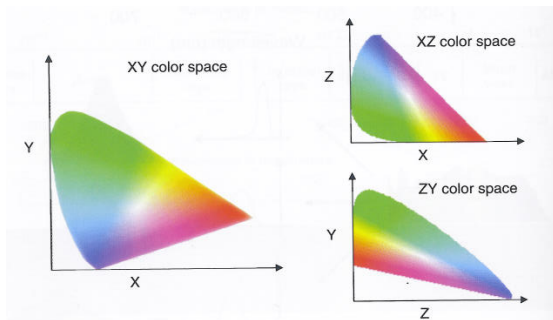
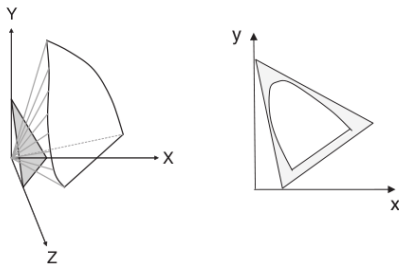


- = modely pro reprezentaci barev na základě složek (souřadnic) = **hodnot primárních barev**, typicky trojrozměrné
- **vektor barvy** = lineární kombinace vektorů hodnot složek
- oblast všech reprezentovatelných barev = **barevný gamut**

- základní, schopný reprezentovat **všechny barvy viditelného spektra** = **gamut vidění**
- **teoretický**: používaný pro studium jiných prostorů, jejich barevných rozsahů (gamutů), dominantních frekvencí barev, mixování a komplementárních barev apod.
- osy = projekce **tristimulus vektoru** na základě citlivosti receptorů oka do „barevných“ složek, osa Y pro jasovou složku barvy, osy nerealizovatelné
- odvozené prostory, např. nejpoužívanější **CIE xyY (xy)**: projekce prostoru XYZ na rovinu $X + Y + Z = 1$: $(x, y, z) = \left(\frac{X}{X+Y+Z}, \frac{Y}{X+Y+Z}, \frac{Z}{X+Y+Z}\right)$, osy x a y pro barevné složky barvy, tzv. **barevné (chromacity) diagramy**

Obrázek: Barevné prostory CIE XYZ a xyY

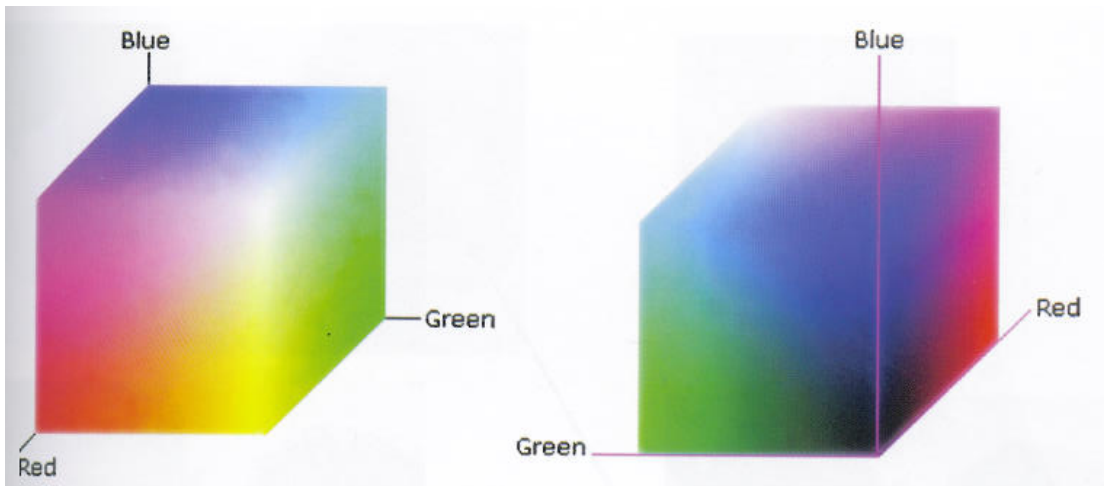
Barevný prostor CIE XYZ

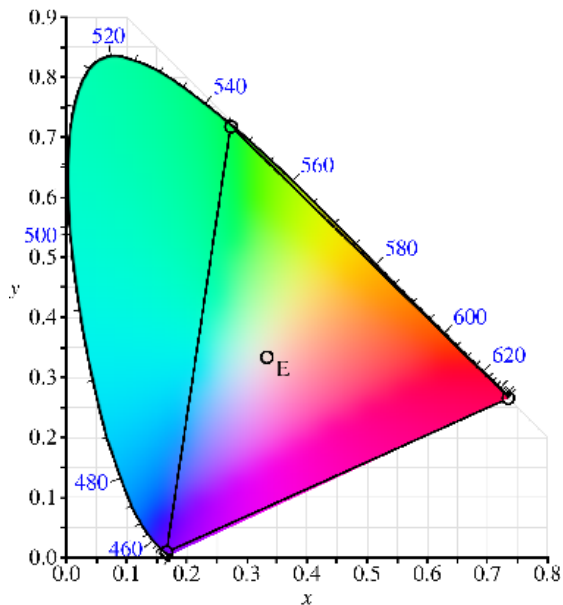


- osy = **primární barvy červená R 645.16 nm, zelená G 526.32 nm, modrá B 444.44 nm**
- normalizovaný do jednotkové „**RGB kostky**“
- **aditivní**: vektor barvy je součtem vektorů hodnot primárních barev, $(0, 0, 0) =$ černá, $(1, 1, 1) =$ bílá
- používaný záznamovými a zobrazovacími zařízeními – závislý na zařízení (nemají všechna stejná primární barvy)
- nereprezentuje všechny barvy viditelného spektra, zejména zelené

Obrázek: Barevný prostor RGB

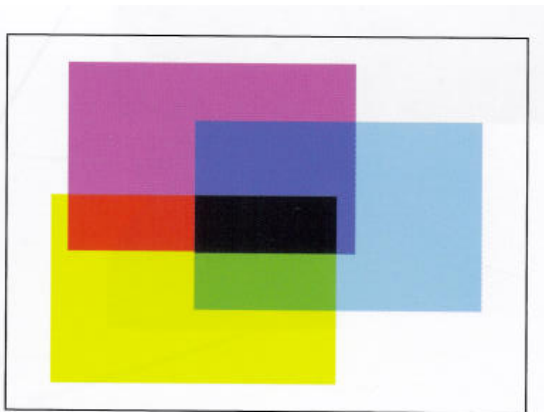
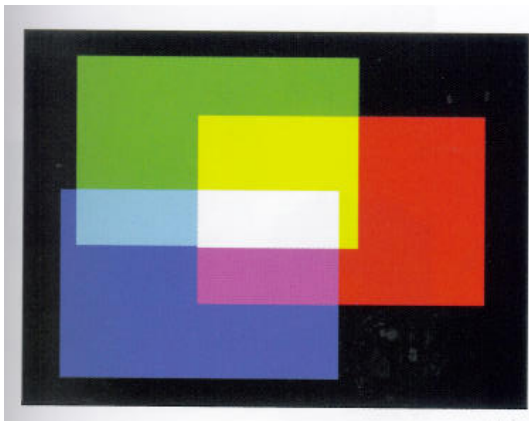
Barevný prostor RGB





- = komplementární k prostoru RGB – primární barvy **azurová C**, **fialová M**, **žlutá Y**
 - **subtraktivní** – vzhledem k prostoru RGB
 - používaný tiskovými zařízeními – tisk primárních barev (kanálů) postupně přes sebe
 - černá (1, 1, 1) – často používaná, neekonomická reprezentace, v praxi nikdy černá
- **CMYK** – navíc **černá K primární barva**, převod z CMY: $C_{CMY} = (C, M, Y)$, jestliže $\min(C, M, Y) = 1$, pak $C_{CMYK} = (0, 0, 0, 1)$, jinak $K = \min(C, M, Y)$ a $C_{CMYK} = (\frac{C-K}{1-K}, \frac{M-K}{1-K}, \frac{Y-K}{1-K}, K)$

Obrázek: Aditivní a subtraktivní slučování primárních barev v prostorech RGB a CMY





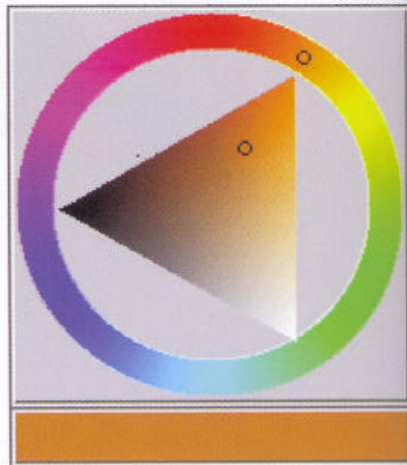
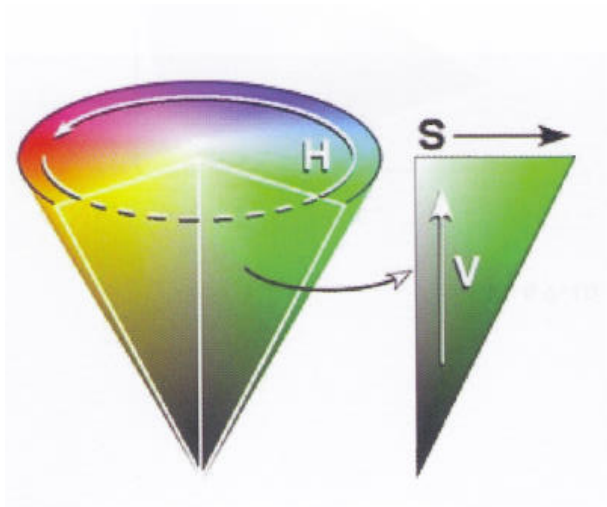
- osa Y pro **jasovou složku** barvy, osy U/I/Cb, V/Q/Cr pro **barevné složky** – šetření pásma s podvzorkováním barevných složek (redukce barevného rozlišení méně ovlivňuje vnímání obrazu než redukce jasového rozlišení)
- používaný u analogové TV (YUV v PAL a SECAM, YIQ v NTSC) a kompresních formátů JPEG a MPEG (YCbCr)
- převod z RGB:

$$T_{RGBtoYUV} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & 0.312 \end{pmatrix}$$

- **nelineární**, použití v (digitálním) zpracování obrazu pro výběr a modifikaci barev – lépe zachycuje **lidské chápání barev** (sousední odstíny)
- **hue** ... **odstín barvy** (dominantní frekvence spektra)
- **saturation** ... **nasycení barvy** od bílé k plné barvě
- **brightness (lightness, value)** ... **světlost barvy** od černé k bílé

Obrázek: Barevný prostor HSV

Obrázek: Převod z barevného prostoru RGB do HSV




```
if (R > G) then
    Max = R; Min = G; position = 0;
else
    Max = G; Min = R; position = 1;
end
if (Max < B) then
    Max = B; position = 2
end
if (Min > B) then
    Min = B;
end
V = Max;
if (Max != 0) then
    S = (Max - Min)/Max ;
else
    S = 0;
end
if (S != 0) then
    if (position = 0) then
        H = (1+G-B)/(Max-Min);
    else if (position = 1) then
        H = (3+B-R)/(Max-Min);
    else
        H = (5+R-G)/(Max-Min) ;
    end
end
end
```

- **uniformní** = rozdíl souřadnic odráží vnímaný **rozdíl mezi barvami** – třídy vnímáním nerozlišitelných barev jsou zobrazeny jako vyplněné kruhy, narozdíl od různě natočených elips v CIE xy prostoru

Obrázek: Nerozlišitelné barvy v CIE xy prostoru (tzv. MacAdamovy elipsy, zvětšené)

- LUV neodráží rozdíl jasů, LAB ano
- pro zkoumání rozdílů barev tak, jak jsou vnímány člověkem

