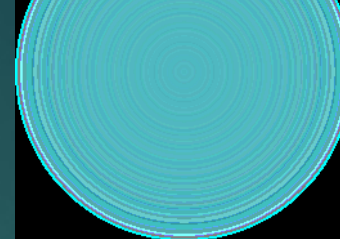




# Parametry, kontejnery, kodeky a formáty multimédií

LABORATOŘ VZDĚLÁVACÍCH TECHNOLOGIÍ

MICHAL ČERNÝ



# Jak zaznamenat zvuk v počítači

- ▶ MIDI klávesy – ideální varianta – každý stisk klávesy se přenese do notového zápisu, metadata rozliší případně rozliší nástroj. Pro většinu případů ne úplně použitelné.
- ▶ Z analogového zdroje: je potřeba provést převod A/D a při výstupu D/A. Tato cesta se používá prakticky vždy.

# Princip převodu A/D

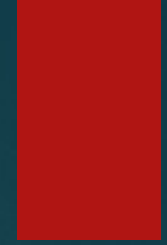
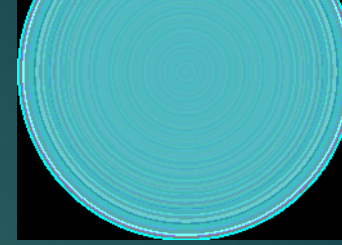
- ▶ Zvuk je spojitě složitě vlnění
- ▶ Máme zařízení, které převede tlakové vlny (zvuk) do elektrický signál (mikrofon). Zde už je první zkreslení – omezené možnosti kmitání, membrány, zpoždění, interference, šumy,...
- ▶ Změny napěťové úrovně jdou do zvukové karty (některé mikrofony zpracování provádějí sami), která je převede na posloupnost bitů
- ▶ Druhy převodníků:
  - ▶ Paralelní
  - ▶ Kompenzační
  - ▶ Integrační
  - ▶ Sigma-delta



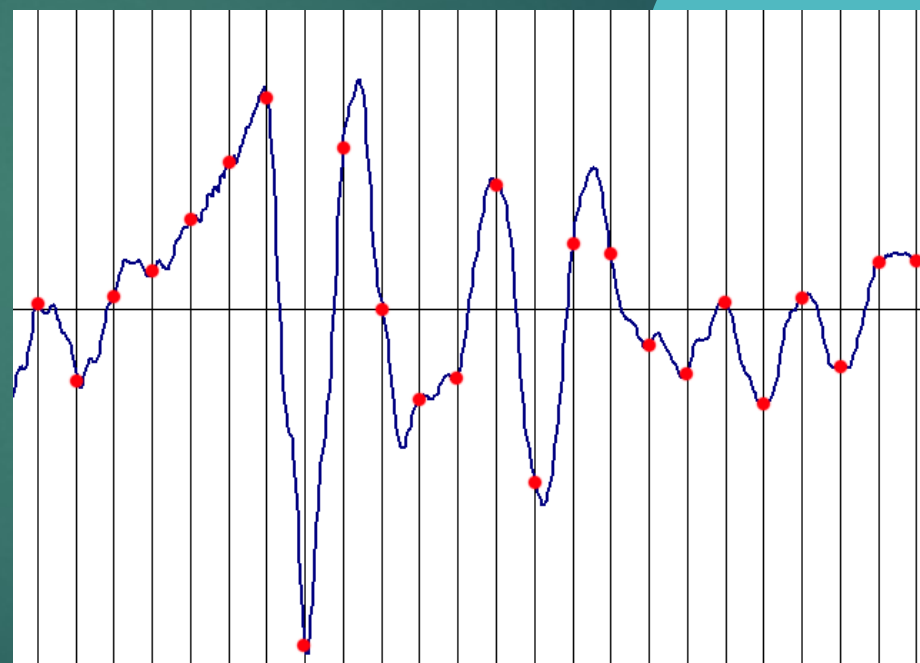
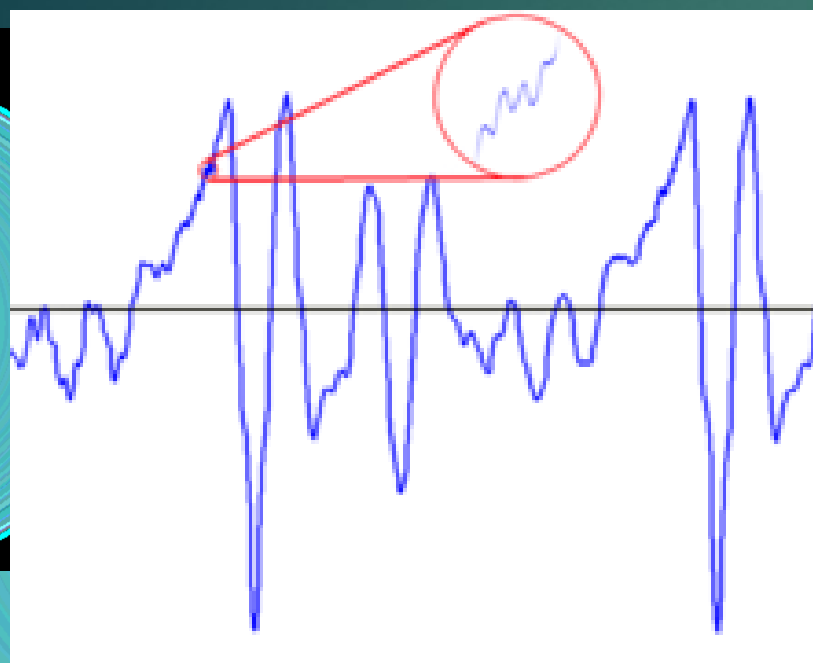
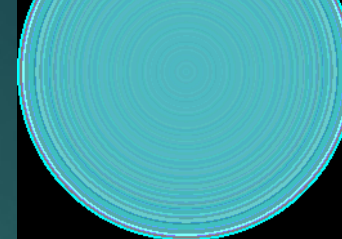


# Parametry

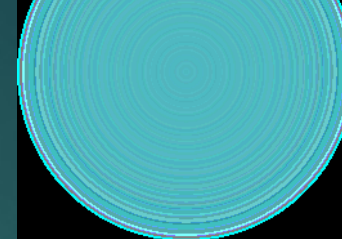
- ▶ Vzorovací frekvence (jak často jsou čtena data)
- ▶ Kvantifikace zvukového signálu (hloubka)
- ▶ Kvantizační šum: poměr chyby při kvantifikaci od jednotlivých vzorků a užitečného signálu. Měří se v decibelech.
- ▶ Počet kanálů (1,2, 6 (5.1))



# Vzorkování

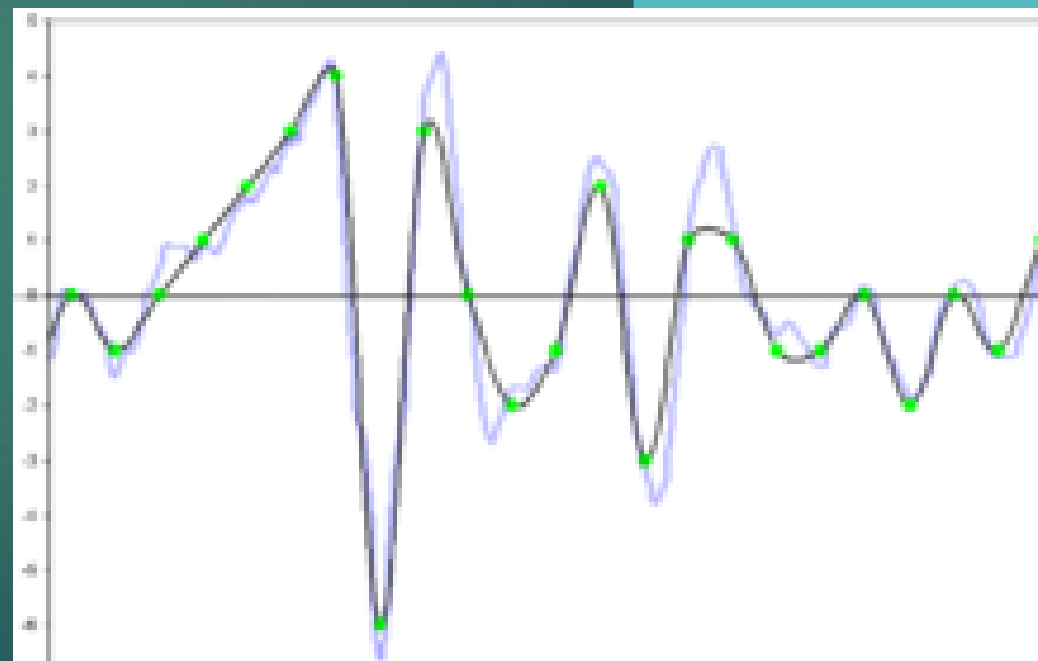
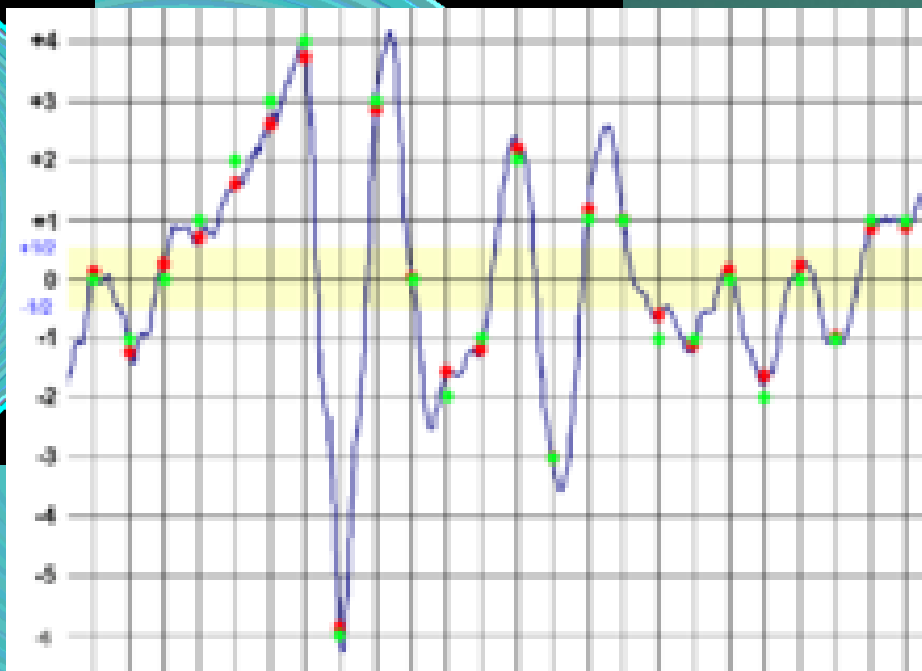


# Kvantifikace



► kvantifikace

► rekonstruovaný signál



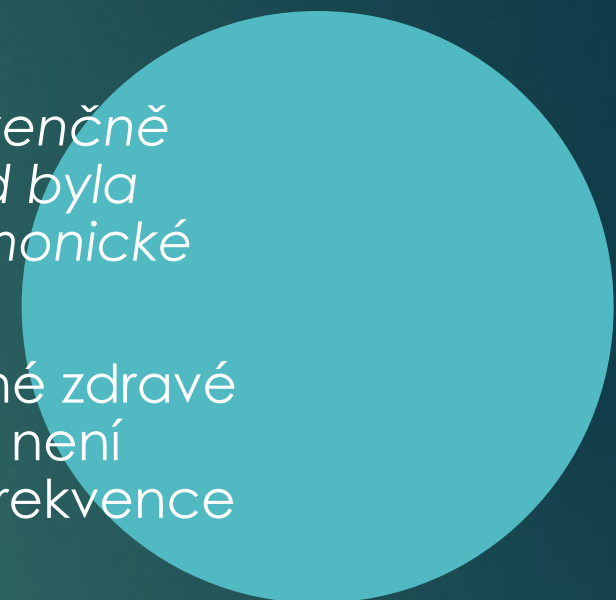
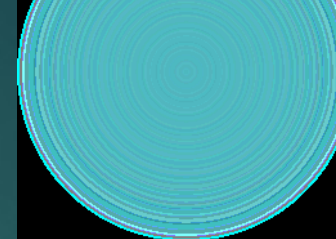
# Aliasing

- ▶ Shannonův teorém: „Přesná rekonstrukce spojitého, frekvenčně omezeného, signálu z jeho vzorků je možná tehdy, pokud byla vzorkovací frekvence vyšší než dvojnásobek nejvyšší harmonické složky vzorkovaného signálu.“

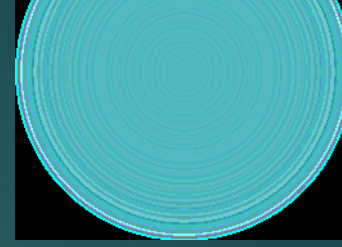
- ▶ Například u záznamu na CD je to 44,1 kHz neboť průměrné zdravé lidské ucho slyší maximálně cca do 20 kHz (což už dávno není pravda – platí to spíše pro malé děti) a tudíž vzorkovací frekvence 44,1 kHz byla zvolena s určitou rezervou.

- ▶ <http://onlinetonegenerator.com/>

- ▶ Aliasing: naprosté zkreslení signálu často zcela měnící smysl sdělení (rotující kotouč stojí, hodiny jdou opačně,...).







# Jak zaznamenat (pohyblivý) obraz v počítači

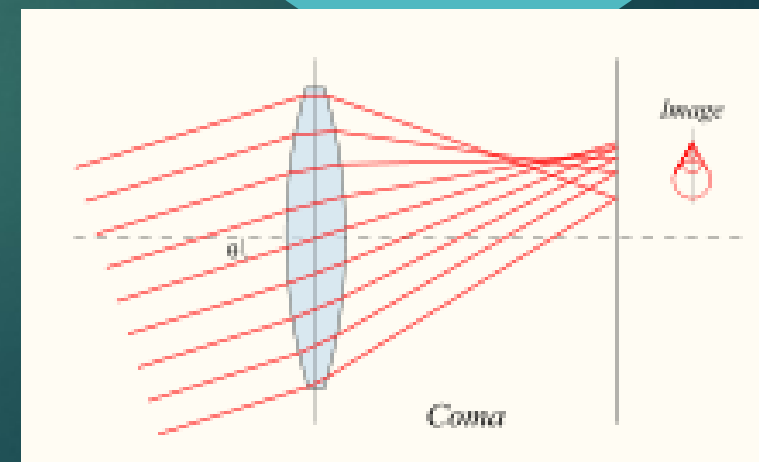
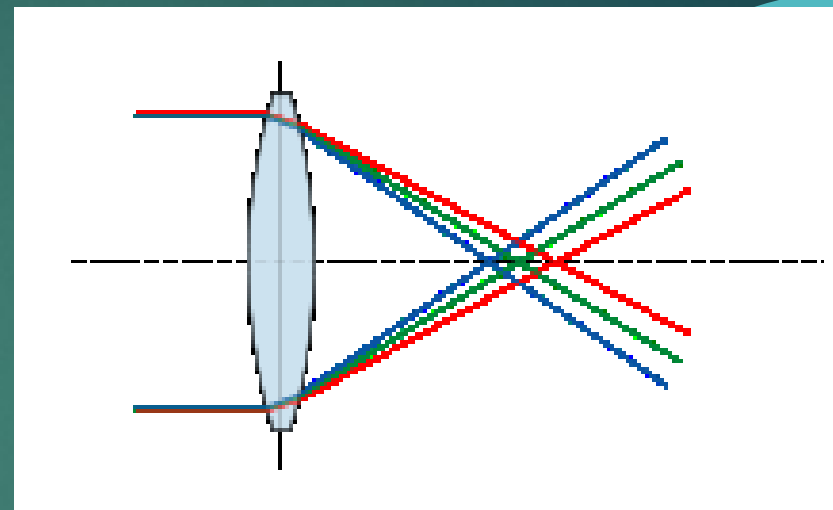
- ▶ Vytvoříme jej přímo digitálně – například můžeme modelovat 3D scény a záznam z nich. Jediné omezení jsou:

- ▶ Výkon počítače
- ▶ Objem dat
- ▶ Kvalita textur
- ▶ Případné softwarové limity

- ▶ To ale většinou nestačí a musíme použít nějakou kameru. To přináší podobné problémy jako se zvukem

# Problémy s obrazem

- ▶ Optické vady:
  - ▶ Barevná vada
  - ▶ Sférická vada
  - ▶ Astigmatická vada či koma
  - ▶ Zkreslení obrazu
  - ▶ Zklenutí
- ▶ Vzorkovací frekvence
- ▶ Rozišení
- ▶ Barevná hloubka



# Parametry videa

- ▶ Rozlišení (problematická standardizace):

- ▶ PAL: 720i 720x576

- ▶ HDTV: 1080i 1920x1080

- ▶ Digital cinema 4K: 4096x1714 nebo 3996x2160

- ▶ ...

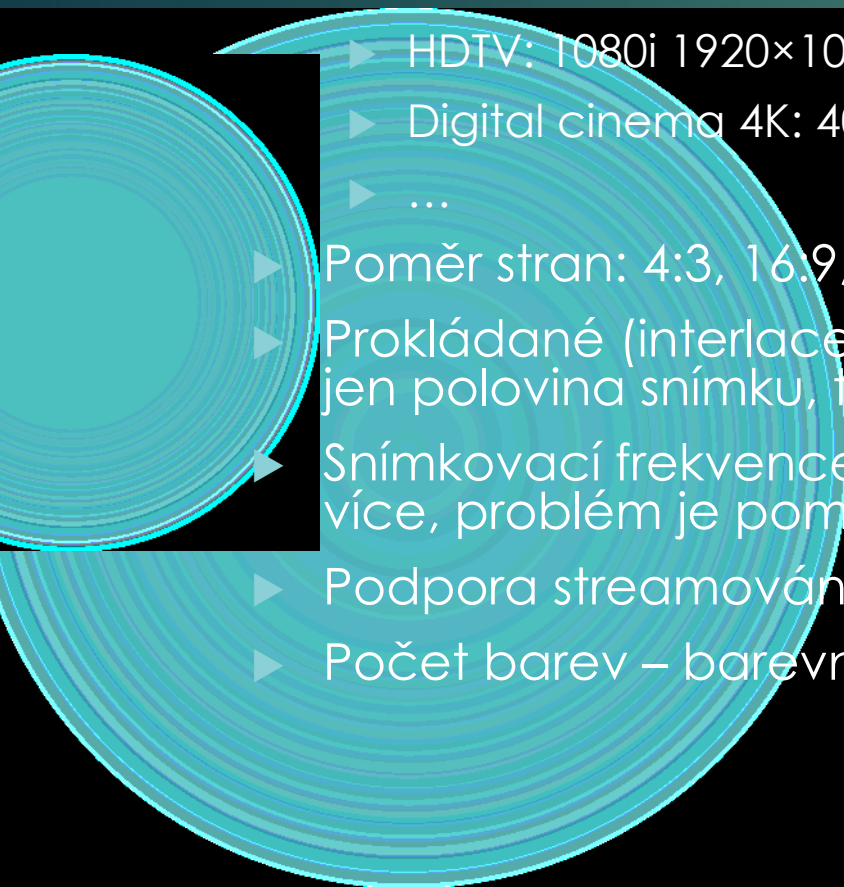
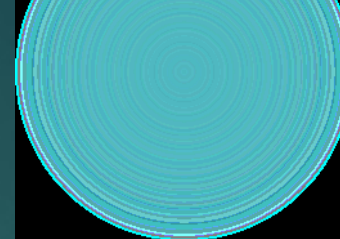
- ▶ Poměr stran: 4:3, 16:9, ...

- ▶ Prokládané (interlace) nebo úplné. U prokládaného se zobrazuje vždy jen polovina snímku, takže se střídají sudé a liché řádky)

- ▶ Snímkovací frekvence (25, 30, 50 fps) vysokorychlostní kamery 1000 i více, problém je poměr rozlišení a frekvence (omezený datový tok)

- ▶ Podpora streamování

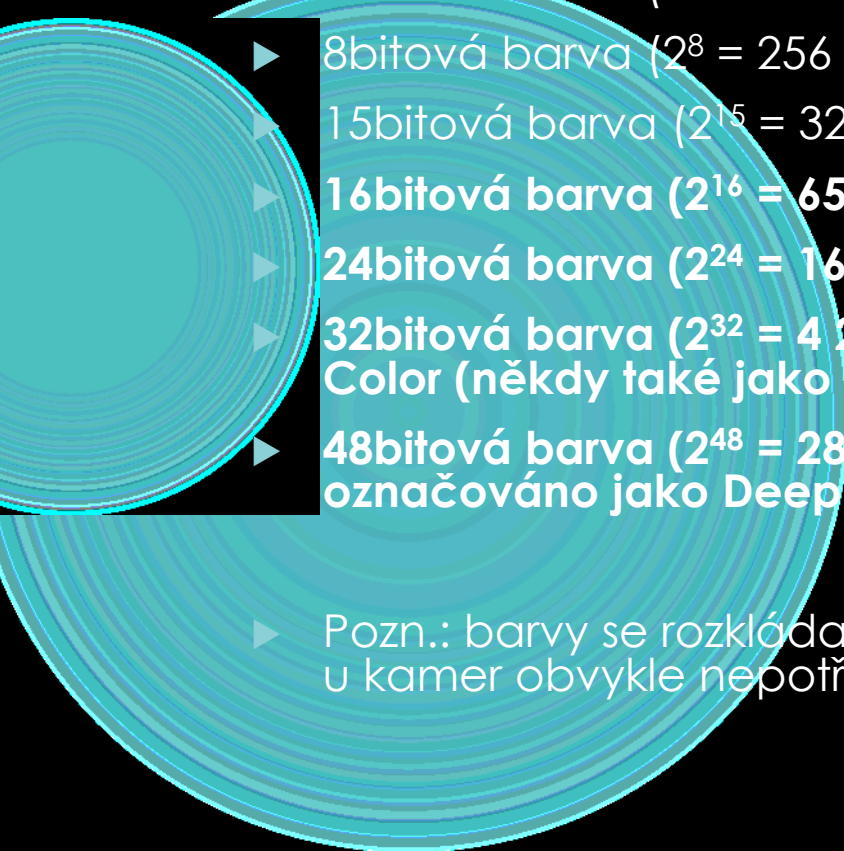
- ▶ Počet barev – barevná hloubka



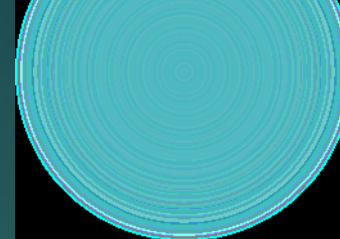
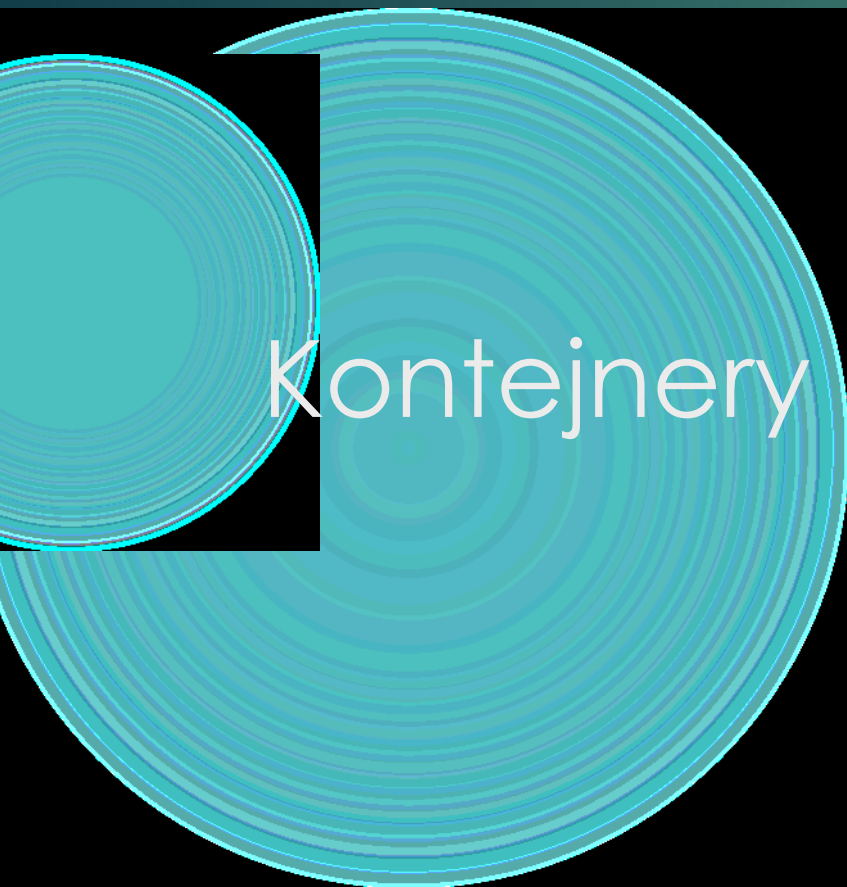


# Používané barevné hloubky

- ▶ 1bitová barva ( $2^1 = 2$  barvy) také označováno jako Mono Color (nejpoužívanější je, že bit 0 = černá = a bit 1 = bílá )
  - ▶ 4bitová barva ( $2^4 = 16$  barev)
  - ▶ 8bitová barva ( $2^8 = 256$  barev)
  - ▶ 15bitová barva ( $2^{15} = 32\,768$  barev) také označováno jako Low Color
  - ▶ **16bitová barva ( $2^{16} = 65\,536$  barev) také označováno jako High Color**
  - ▶ **24bitová barva ( $2^{24} = 16\,777\,216$  barev) také označováno jako True Color**
  - ▶ **32bitová barva ( $2^{32} = 4\,294\,967\,296$  barev) také označováno jako Super True Color (někdy také jako True Color)**
  - ▶ **48bitová barva ( $2^{48} = 281\,474\,976\,710\,656 = 281,5$  biliónů barev) také označováno jako Deep Color**
- ▶ Pozn.: barvy se rozkládají do různých kanálů – nejčastěji RGB + Alfa kanál (to ale u kamer obvykle nepotřebujeme), někdy RGBY (přidává se žlutý kanál).







# Co jsou to kontejnery

- ▶ Médium je zabalené v rámci, kterému se říká kontejner.
- ▶ Pro některé aplikace je kontejner uzavřený a zajímavá je jen hlavička, což přináší řadu výhod.
- ▶ Typicky má kontejner hlavičku a patičku. Patička není principiálně povinná.
- ▶ Umožňují do jednoho balíčku (či toku) spojit audio, video, titulky, ... a řadu dalších informací.



# Kontejnerové formáty pro video

- ▶ 3GP (3GP2) – pro mobilní telefony, .3gp, zjednodušený MP4
- ▶ ASF (WMV) – Microsoft, speciálně pro streamování, .asf, uzavřený, index na konci
- ▶ **AVI (Audio Video Interleave) – Microsoft, založen na RIFF, .avi, Matroska (MKV) – .mkv, otevřený standard, knihovna pod LGPL, binární XML**
- ▶ **MPEG-TS (Transport stream) – MPEG-2 Part 1, DVB, pro streamování**
- ▶ **MPEG-PS (Program stream) – DVD (VOB – podpora DVD menu), specifikován v MPEG-1 a MPEG-2 Part 1**
- ▶ MP4 – MPEG-4 Part 14, .mp4, založen na MOV (Apple)
- ▶ MOV – framework QuickTime, .mov, .qt
- ▶ **Ogg – Xiph.Org Foundation, není chráněn patenty, .ogv, .ogg**
- ▶ OGM – hack Ogg, aby podporoval jakékoliv kodeky, .ogm
- ▶ RealMedia – .rm, pro streamování, kodeky RealVideo a RealAudio, může se měnit bitrate.
- ▶ Flash Video (FLV) – .flv, Adobe Flash Player, formát sice otevřený, ale používá patentované kodeky (H.264, VP6, MPEG-4 ASP, Sorenson H.263)



# AVI (Audio Video Interleave)

- ▶ Zřejmě nejstarší kontejner od Microsoftu, proprietární.
- ▶ Na začátku je hlavička s informacemi o obrazu (framerate, rozlišení, kodek,...) a o zvuku (samplerate, kvantizace, kodek,...).
- ▶ Na konci index, tedy seznam jednotlivých snímků:
  - ▶ Nedá se streamovat (potřebujeme poslední část souboru pro vlastní zpuštění)
  - ▶ Zvukové a obrazové stopy nejsou spojené časem, dochází k rozsynchronizování.
- ▶ Typy AVI souborů:
  - ▶ **AVI 1.0** - umožňuje nahrávat pouze do velikosti 1 GB, maximální počet snímků je 22500, tedy asi čtvrt hodiny záznamu pro 25sn/s, používal se ve Windows 3.1, již se nepoužívá
  - ▶ **AVI 1.1** - rozšířeno nahrávání a indexování do velikosti souborů 2GB, některé programy řeší překonání této bariéry pomocí nahrávání do více souborů
  - ▶ **AVI 2.0** - označuje se také OpenDML, má neomezenou velikost souboru, ale FAT32 omezuje maximální velikost na 4GB, na NTFS je (téměř) neomezeno (velikost disku)

# AVI (Audio Video Interleave) II.

- ▶ Používá se ještě DV AVI, které je dvou typů:

- ▶ Type 1: jeden stream se zvukem i videm

- ▶ Type 2: zvuk a obraz je zvlášť

- ▶ Hlavička začíná RIFF, velikostí souboru a typem, dále jsou informace o typu obsažených streamů a dále multiplexovaná data. Na konci je již zmíněný index

- ▶ AVI je dobré pro stříh nebo obecně zpracovávání videa, ale rozhodně ne pro streamování nebo přenos po síti.



# MPEG stream

- ▶ ISO/IEC standard – vše je zabaleno do jednoho streamu
- ▶ Spojení obrazu, zvuku a případných dalších dat je časové
- ▶ Důraz kladen na jednoduchost, neboť jde o formát i pro zábavní elektroniku
- ▶ Více úrovněový model:
  - ▶ **Elementary stream (ES)** – vytvoření streamu z jednoho zdroje. Data jsou dohromady spojena do *access units*.
  - ▶ **Packetised Elementary Stream (PES)** tvoří posloupnost paketů, obsahuje několik *access units*. Maximální velikost takového paketu je 65536 bytů, obsahuje hlavičku s informacemi o streamu a data (ES stream). PES můžeme buď ukládat nebo vysílat.

# MPEG stream II.

- ▶ Takto vytvořené video či audio je možné uložit do souborů MPV (MPEG video), .M2V (MPEG-2 video), .MPA (MPEG Audio), .MP2 (MPEG Audio Layer 2), .MP3 (MPEG Audio Layer 3)

- ▶ **MPEG Program Stream (PS)** – řadí za sebe jednotlivé pakety a vytváří tak souvislý tok dat, který je časově synchronizován.

- ▶ **MPEG Transport Stream (TS)** – je určen do prostředí s velkou chybovostí (DVB). Rozdělí pakety na menší části a souvislým datovým tokem zajišťuje přenos. Umožňuje multiplexovat více programů do jednoho streamu.

# Advanced Systems Format (ASF) a WMA

- ▶ Uzavřený formát Microsoftu, je potřeba mít konkrétní proprietární kodeky
- ▶ Umožňuje streamování, nahrazuje v tomto ohledu AVI
- ▶ Je založený na objektech

# Další

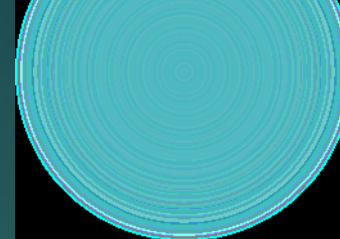
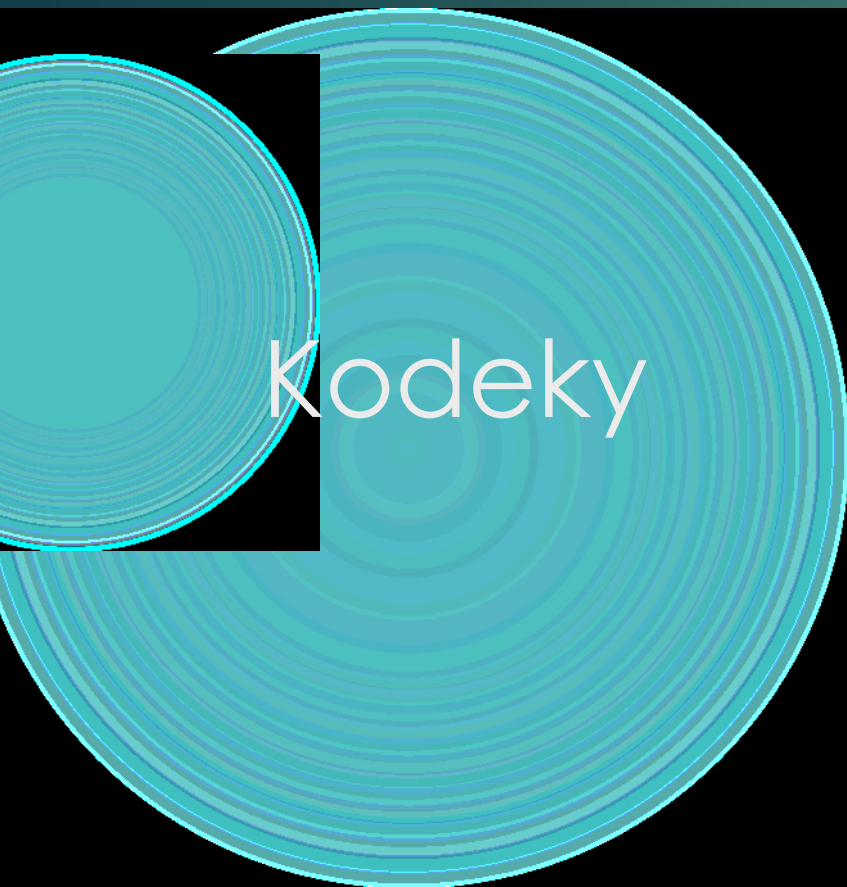
- ▶ Ogg Media (OGM) je navázán na MPEG-4 kodek. Měl nahrazovat AVI je otevřený, ale ne úplně rozšířený.

Quicktime (MOV) od Apple je založený na objektové struktuře, poměrně robustní, může nést různé druhy informací, takže je například použitelný pro flash.

- ▶ Matroška je založený na XML struktuře - na začátku je hlavička, která identifikuje typ EBML souboru a verzi, následuje segment(y), který tvoří datovou funkční část, typicky je jen jeden

- ▶ Segment se dále dělí na menší části (sekce), každá obsahuje jiný typ informace
- ▶ Podpora řady dat připojených k videu či audio
- ▶ Velké možnosti, zatím nepříliš velké rozšíření
- ▶ Podrobněji na: <http://www.matroska.org/technical/specs/index.html#track>







# Co, k čemu a proč?

- ▶ Kodek – kodér a dekodér. Tedy popis toho, jak se má analogová data zapsat do digitální podoby a zase nazpátek.
- ▶ Různé druhy médií se hodí různě kódovat, neboť se liší:
  - ▶ V druhu – jinak pracujeme s hudbou, jinak s textem či videm
  - ▶ Ve ztrátovost – chceme zachovat úplnou informaci (většinou do poměru 1:2) nebo snížit velikost a náročnost na přehrání a zpracování (1:4 až 1:50)
  - ▶ Ve způsobu komprese, nastavených algoritmech
  - ▶ V otevřenosti zdrojového kódu
  - ▶ ...

# Některé přístupy ke kompresi

- ▶ DCT (diskrétní kosinová transformace), FFT (rychlá Fourierova transformace) nebo DWT (diskrétní vlnková transformace) – v zásadě matematický trik, založený na **převodu dat z jedné domény druhé** (například časové do frekvenční).
- ▶ Potlačení některých dat – například u zvuku frekvence které již člověk neslyší snížíme na 44,1 kHz, v případě obrazu vytváříme oblasti, které jsou spojitě a mají (přibližně) jednu barvu,...

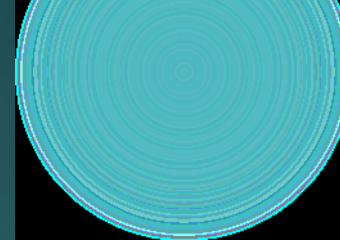
# DivX

- ▶ Pracuje s formátem videa v MPEG-4 ASP
- ▶ Populární formát, dobrá komprese, široká podpora
- ▶ Proprietární licence, návaznost na vlastní přehrávač, spojení s kontejnerem (například DivX Plus Streaming™)
- ▶ Umožňuje hardwarovou akceleraci
- ▶ Relativně zajímavé experimenty v Labs (4K video, práce s různými kontejnery, streamování,...)

# Xvid

- ▶ Pracuje s formátem videa v MPEG-4 ASP.
- ▶ Pracuje se na podpoře maximální kvality obrazu, nových algoritmů pro kompresi videa.
- ▶ Důležitá je podpora procesorů (SSE2/SSE3 SIMD) a operačních systémů. Kodek by měl využívat všechny jejich možnosti co možná nejvíce. Problémem je, že každý operační systém je jiný, všechny procesory nejsou intel...
- ▶ Projekt je open source (GPL)





# Zvuk

- ▶ MP3 – masivní rozšíření, problémy s mluveným slovem a licemi.
- ▶ AC-3 – ztrátový formát, podporuje i 5.1, 6.1 či 7.1 zvukové kanály, maximální tok 640 kbit/s. Dnes používaný zvukový formát pro DVD, kina atp.
- ▶ AAC – různorodá skupina formátů nahrazující mp3 (založená na MPEG-4 standardu), řada různorodých kóderů určujících kvalitu záznamu. Může být součástí MOV, MPEG, 3GP, AVI a MKV kontejnerů.
- ▶ WAV obecný formát pro komprimovaná i bezztrátová data od MS.

# Video

- ▶ Dirac – ztrátový formát, technologicky podobný jpeg2000, cílem je PAL kvalita (720i) v reálném čase.
- ▶ H265 – nový formát pro video ve vysokém rozlišení (8K) a oproti H264 má mít poloviční datový tok.
- ▶ H264 (MPEG-4 part 10 a MPEG-4 AVC)
- ▶ MPEG4 – kolekce audio a video kompresních metod, velice komplexní formát. Hojně užívaný pro stream, vysílání i ukládání dat.
- ▶ Theora – jediný formát videa podporovaný Wikipedií, otevřený, navázaný na Ogg.