



A BRIEF OVERLOOK INTO AUDIO AND VIDEO CODECS AND WAYS TO IMPROVE THEM USING FILTERS

CE487 – Smart
Multimedia
Systems

Furkan
MUDANYALI

CONTAINERS

- Can hold single or multiple media streams
 - Some containers exclusively hold one type of stream
 - Other containers can hold multiple of various media streams
-

VIDEO CONTAINERS

- MKV (Matroska)
 - - An open format that can contain virtually any kind of stream
 - MP4
 - - Can hold H264, AV1 and VP9 video codecs and AAC, FLAC, Opus and MP3 audio codecs
 - WEBM
 - -Can hold AV1, VP8 and VP9 video codecs and Opus and Vorbis audio codecs
-

AUDIO CONTAINERS

- FLAC
 - - The container for the FLAC codec.
 - M4A
 - - Can hold AAC or ALAC codec streams.
 - DSF
 - - Container for DSD (Direct Stream Digital).
-

CODECS

- Compression of media streams
 - Can be lossy or lossless
 - Lossy codecs will lower the quality but in turn will have significantly smaller size
 - Lossless codecs will retain the quality of the source stream at the cost of larger size
-

VIDEO CODECS

- H.264 (AVC)
 - - Block oriented and motion-compensated coding. The most common video codec.
 - HEVC (H.265)
 - - Improvement over H264 that supports variable sized blocks
 - - Delivers around the same quality of H.264 at half the bitrate.
 - - Not widely adopted due to high licensing fees.
-

FIXED-SIZE BLOCKS VS VARIABLE-SIZE BLOCKS



VIDEO CODECS

- VP8
 - - Block based codec that shares a lot with H.264
 - - Aimed to replace the GIF format, not terribly great on higher resolutions.
 - VP9
 - - Successor to VP8 and competitor to H.265
 - - Comparably same performance with H.265 but widely adopted.
 - AV1
 - - Successor to VP9.
 - - 35% to 50% higher data compression compared to VP9.
 - - Currently the choice for most video streaming services such as YouTube, Netflix and even Facebook with more to come.
-

AUDIO CODECS: LOSSY STEREO

- AAC
 - - Successor to MP3.
 - - Widely adopted, standard format for almost all portable devices.
 - Opus
 - - Successor to Vorbis, which was an alternative to MP3.
 - - Performs well even at extremely low bitrates
 - - Lower latency, preferred by voice call services such as Whatsapp and Signal.
-

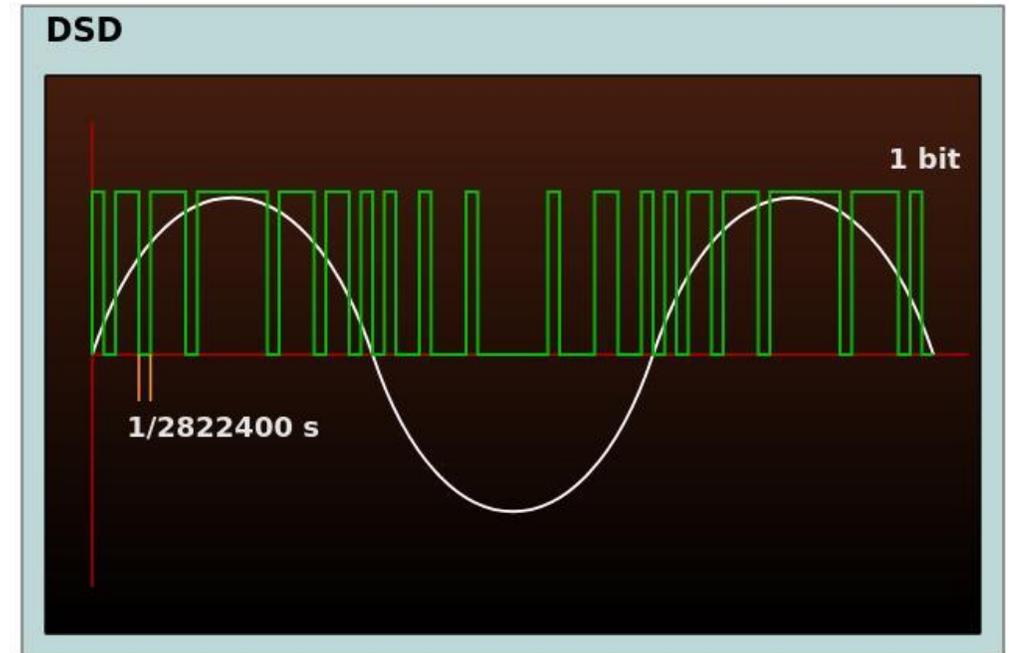
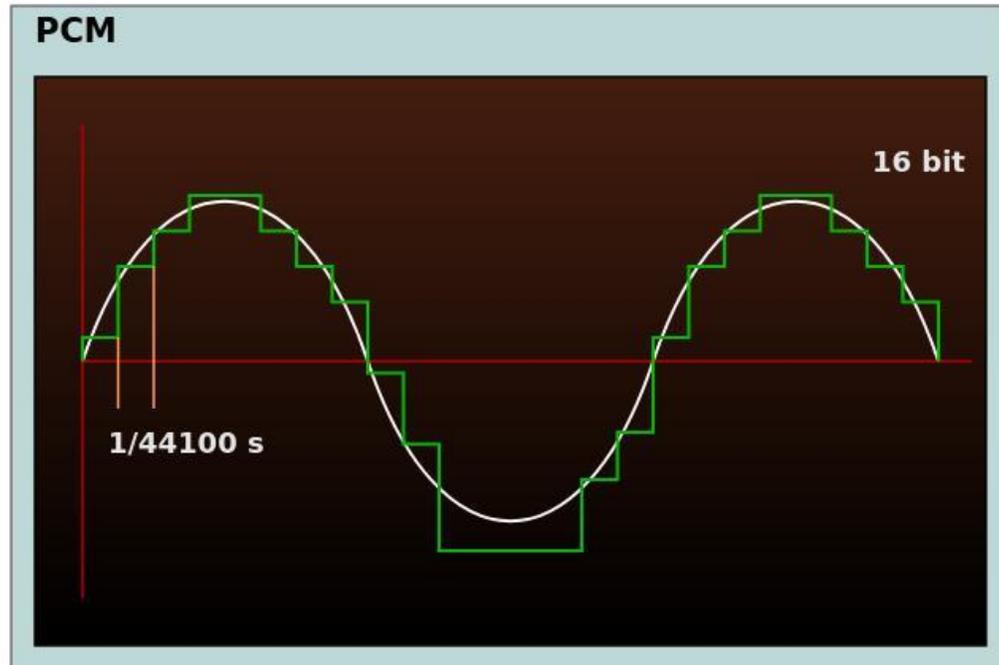
AUDIO CODECS: LOSSLESS STEREO

- **FLAC**
 - - Lossless format that can reduce the filesize from 50 to 70% depending on compression.
 - - Able to decode quickly independent from the compression level.
 - **ALAC**
 - - Lossless format by Apple that is comparable to FLAC in compression.
 - - Requires around 4 times more CPU power to decode compared to FLAC.
 - **DSD**
 - - Uses delta-sigma modulation compared to pulse-code modulation (PCM) like in conventional audio formats.
 - - Requires specialized hardware to decode.
 - - Adopted in mainly Japan.
-

AUDIO CODECS: SURROUND

- AC3
 - - Lossy format by Dolby Digital that can hold up to 6 channels (5.1) at 640 kbps.
 - DTS
 - - Lossy format by Digital Theater Systems that is comparable to AC3.
 - - Despite being able to reach 1500 kbps, it was mostly encoded in 750 kbps to make room for more audio tracks.
 - Dolby TrueHD
 - - Lossless format by Dolby Digital that can hold up to 8 channels (7.1) at 18 mbps.
 - DSD
 - - It is also able to hold 6 channels (5.1) at around 20 mbps. Was distributed along with the stereo mixes in SACDs (Super Audio CD).
-

DSD VS. PCM, A VISUAL COMPARISON



By Paweł Zdziarski from Wikimedia Commons <https://commons.wikimedia.org/wiki/File:PCM-vs-DSD.svg>

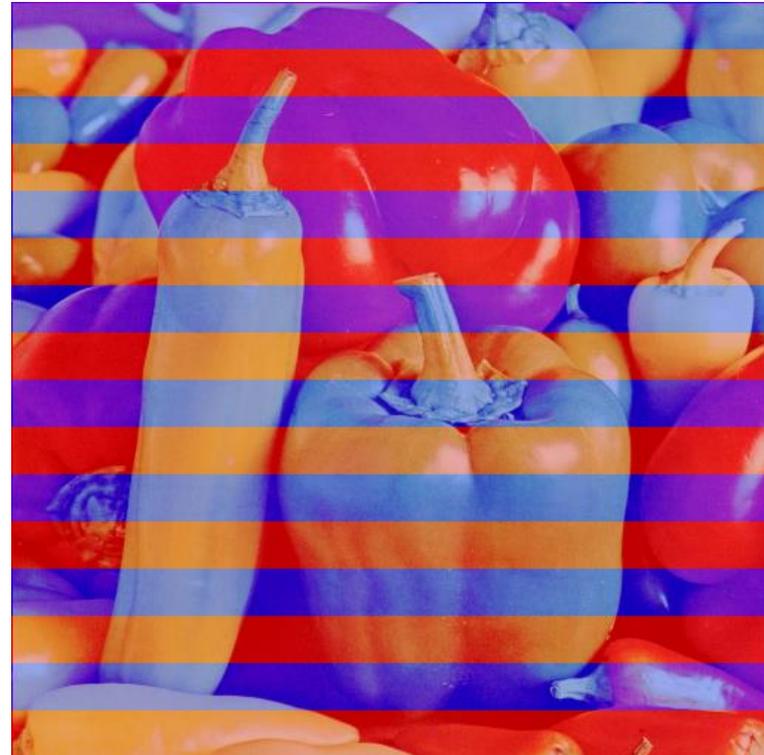
VIDEO FILTERS

- Software components to process the video stream in one way or another.
 - Deinterlacing
 - Denoising
 - Deflicking
 - Deblocking
 - Frame Interpolation
-

INTERLACED VS PROGRESSIVE



Progressive



Interlaced

Peppers by University of Southern California

<https://sipi.usc.edu/database/database.php?volume=misc&image=13>

DEINTERLACING

- Field Combination
 - - Weaving
 - - Blending
- Field Extension
 - - Half Sizing
 - - Line Doubling



Combing Effect

By Mikus from Wikimedia Commons
[https://en.wikipedia.org/wiki/File:Interlaced_video_frame_\(car_wheel\).jpg](https://en.wikipedia.org/wiki/File:Interlaced_video_frame_(car_wheel).jpg)

FIELD COMBINATION: WEAVING

- In the PS2 game "Ratchet: Deadlocked", which runs at 60 frames per second, the "combing effect" can clearly be seen.
 - Used the following script to capture: <https://furkanmudanyali.com/blog/codecs/capture.sh>
 - Link to the recording: <https://furkanmudanyali.com/blog/codecs/original.mp4>
-

FIELD EXTENSION: LINE DOUBLING

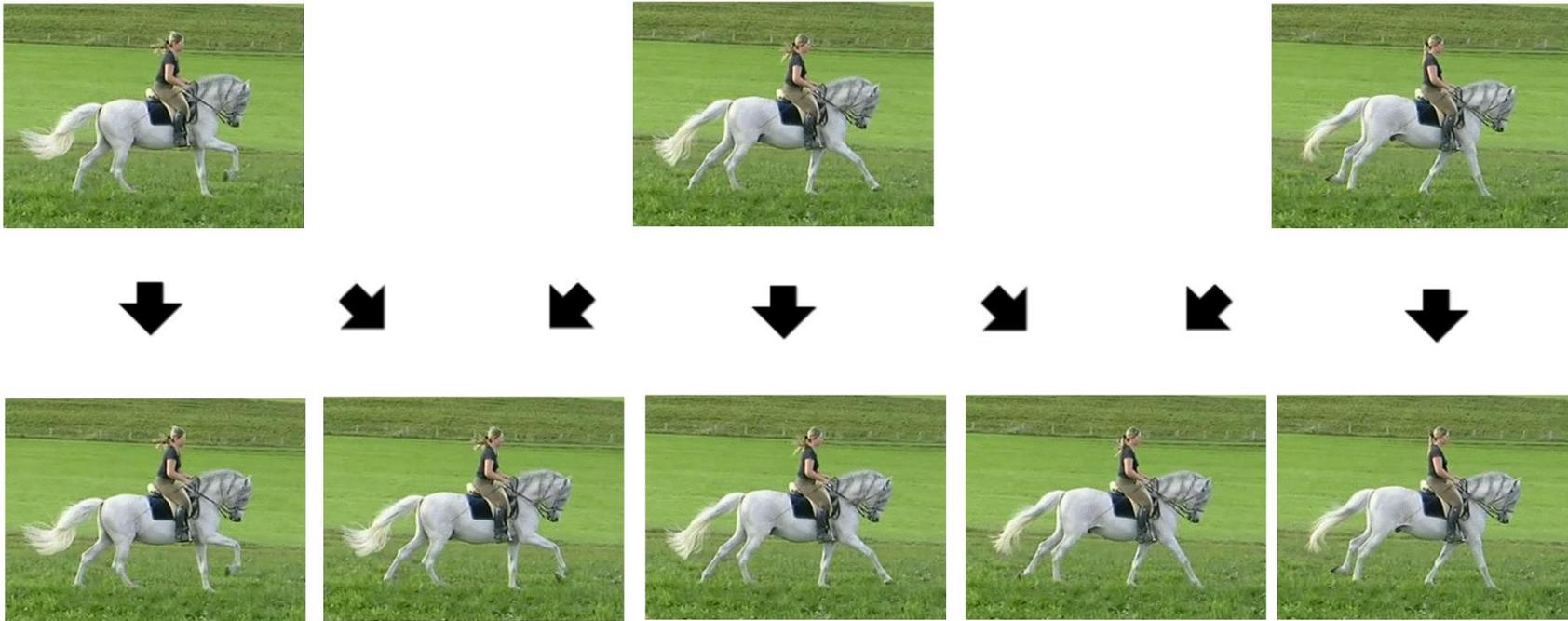
- The same PS2 game "Ratchet: Deadlocked" passed through BWDIF and YADIF deinterlacing filters.
 - YADIF: <https://furkanmudanyali.com/blog/codecs/yadif.mp4>
 - `ffmpeg -i input.mkv -vf "yadif=mode=0:parity=0:deint=0" yadif.mp4`
 - BWDIF: <https://furkanmudanyali.com/blog/codecs/bwdif.mp4>
 - `ffmpeg -i input.mkv -vf "bwdif=mode=0:parity=0:deint=0" bwdif.mp4`
-

FRAME INTERPOLATION

- Doubles the framerate by blending the 2 adjacent frames and filling the gap.
- Minterpolate (Motion Interpolation: Blend):
https://furkanmudanyali.com/blog/codecs/minterpolate_basic.mp4 (took 6 minutes to process)
- `ffmpeg -i input.avi -vf "minterpolate=fps=48:mi_mode=blend" minterpolate_basic.mp4`
- Minterpolate (Motion Interpolation: Adaptive overlapped block motion compensation with bidirectional motion estimation and variable-size block motion compensation):
https://furkanmudanyali.com/blog/codecs/minterpolate_advanced.mp4 (took 4 hours to process)
- `ffmpeg -i input.avi -vf "minterpolate=fps=48:mi_mode=mci:mc_mode=aobmc:me_mode=bidir:vsbmc=1" minterpolate_advanced.mp4`

FRAME INTERPOLATION

Motion interpolation of a horse and rider galloping.



Frame interpolation example by Waugsberg, Peregrine Fisher and nr1jack from Wikimedia Commons https://commons.wikimedia.org/wiki/File:Motion_interpolation_example.jpg

ANY QUESTIONS?

Thank you :-)
