Compression and File Formats

Compressing Moving Images

Moving versus still images:

- *Temporal* as well as *spatial* compression
- Data transfer rates are critical
  - How much data is needed to maintain 24 frames/sec?

Methods:

- Motion JPEG, Cinepak, Indeo, MPEG, many others
- Known as *CODECs*
  - compression / decompression algorithms
  - hardware and software implementations
  - symmetrical / asymmetrical

Why and where to compress

Consider video as displayed on a television screen

- Raw data rate is huge:
  - 768 (across) by 576 (vertical) by 25 (frames/sec)
    - 31Mbytes/sec or 1.85Gbytes/minute

Can we compress in hardware and/or in software to provide reasonable data rates and storage requirements?

The DV standard

Digital video (DV) standard includes compression on camera

CCIR601 standard specifies signal sampling

- Subsamples chrominance
- Eyes are less sensitive to variations in colour than to variations in brightness
- Uses luminance, Y (=0.2125R + 0.7154G + 0.0721B), B-Y, and R-Y
- Luminance is sent each pixel, but others are subsampled e.g. colour (B-Y & R-Y) sent every 4th pixel

Data rate is 166Mbits (just over 20Mbytes)/second

The DV standard

166Mbits/second is inconveniently high

DV standard uses further compression to bring the data rate down to 25Mbits/second (just over 3Mbytes/sec)

- Compression is on-camera
- Widely used

Still too high a data rate for storage

- 25Mbits/sec = 210Mbytes/min = about 12Gbytes per hour
- High for transmission, or for storage on a DVD

Spatial Compression

Compress each frame (still image) independently

Animated GIFs

- single file containing multiple images compressed using LZW compression
- only suitable for simple, cartoon-like animations

Motion JPEG (*MJPEG*)

- compresses each frame using JPEG
- suitable for videos of real-world action
- implemented in hardware for real-time capture and compression of video
- no standard file format! (*MJPEG-A* emerging)

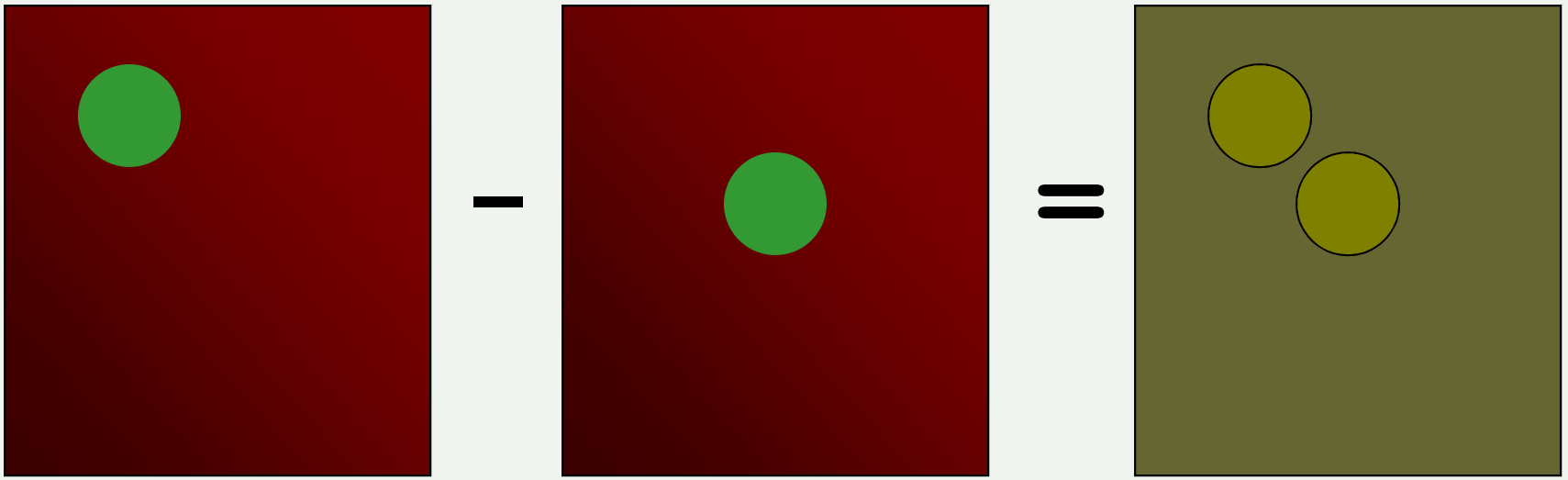
Temporal Compression

Relies on only small changes between frames:

Difference Frames

Naïve approach subtracts each pixel in current frame from corresponding pixel in previous frame

- result is 0 for pixels that do not change between frames
- difference frames will contain large numbers of zeros and so will compress well, using e.g. run-length encoding (RLE).

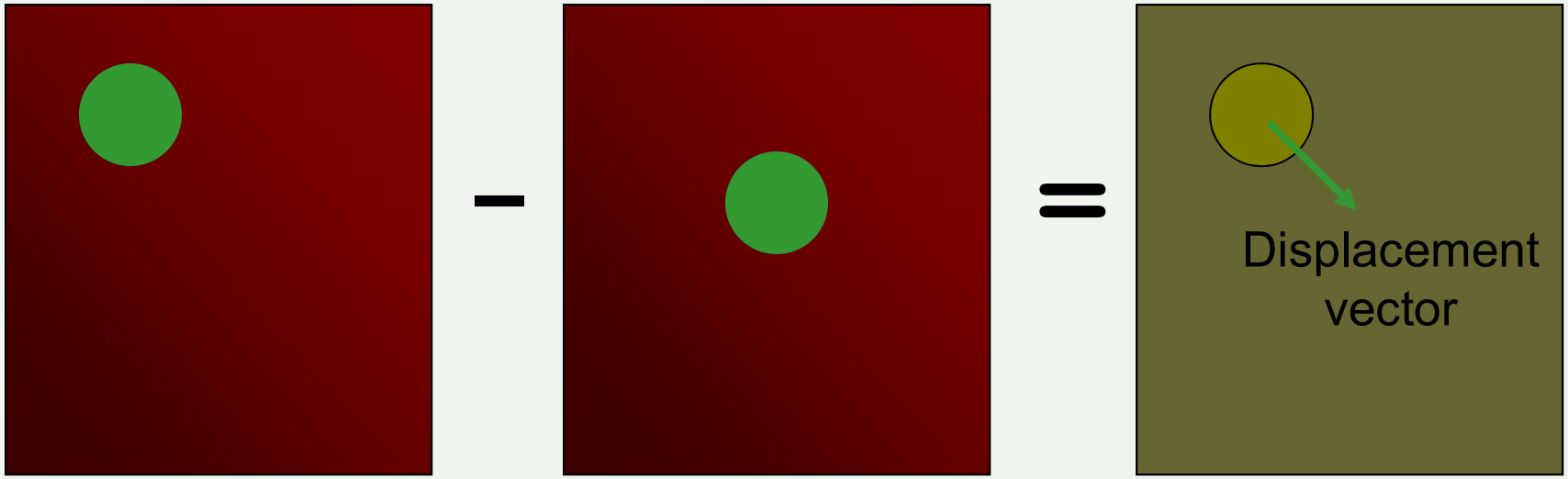


Motion Compensation

Can possibly do better when there is movement between frames

- whole objects (e.g. a walking person) move between frames
- or camera moves

Difference frame would be better if we could match pixels from the moving object between frames



Temporal Compression

Basic idea is this:

- Certain frames designated as *key frames*
  - e.g. every sixth frame (may be user-settable)
- Key frames either are left uncompressed or only spatially compressed
- Each frame between key frames is replaced by a *difference frame*
  - records only differences between frame and preceding frame or preceding key frame
- Difference frames readily compressed with RLE or dictionary-based schemes

Software CODECs for Multimedia

Multimedia is often delivered on CD/DVD or over the Internet

- low to very low data transfer rates

High compression required for video

Software implementations of a variety of *codecs*

- Cinepak
- Intel Indeo
- MPEG-1, MPEG-2, MPEG-4.

Processing power needed for compression and (real-time) decompression

- limits frame size and rate

Cinepak and Indeo

Similar methods that use *vector quantization.* Each frame divided into small rectangular blocks of pixels (the *vectors)*

Use a *code book* consisting of typical vectors

- areas of flat colour
- sharp or soft edges
- different textures
- (the *dictionary* of still image compression)

Quantization is the process of matching each image vector with its closest vector in the code book. Each image vector is replaced by its index in the code book. Note there may not be an exact match resulting in loss of quality.

Cinepak and Indeo (2)

Highly asymmetric: compression takes much longer than decompression -> 150:1 for Cinepak

Also use temporal compression based on key and difference frames

- Automatic keyframe selection
- Good enough for 320x240 pixels at 12 frames/sec from CD-ROM on a basic PC
- not great viewing...

MPEG-1

Motion **P**icture **E**xperts **G**roup

- Audiovisual compression
- See http://www.mpeg.org/MPEG/starting-points.html

Defines a data stream and a decompressor

- Compression is left to the software developer
  - "competitive advantage in the marketplace"
- Standard does define compression implicitly
- Spatial compression similar to JPEG
- Temporal compression via difference frames and motion compensation.

Motion Compensation

Motion compensation in MPEG does not attempt to identify objects in a scene

- A hard computer vision / artificial intelligence problem

Divides each frame into 16 by 16 pixel blocks, known as *macroblocks*

- Different from the smaller blocks used in spatial compression
- Attempts to predict where the corresponding macroblock is in the next frame via simple searching in a limited area around the macroblock
- Difference frame constructed by subtracting matching macroblocks
- BUT also need to store displacement vectors

Frame Types in MPEG

Key frames called *I-pictures* (*I* for *intra*)

- Difference frames based on previous frames called *P-pictures* (predictive pictures)
- MPEG also allows difference frames based on later frames: *B-pictures* (bi-directional prediction)

Display order:

IBBPBBIBBPBBI

Bitstream order:

IPBBIBBPBBIBB

- decoder needs to reconstruct display order from bitstream

Group of Pictures (GOP): IBBPBBPBBPBBIBBPBBPB

MPEG Variations

MPEG-1: Delivers around 2Mbps of video and 250Kbps of 2-channel audio from CD-ROM

- 352x240 pixels at 30 fps on standard PC
- 40 minutes of video on CD-ROM (9 hours on DVD)
- best for multimedia applications

MPEG-2: Needs higher data rates of 3-15Mbps, but delivers better picture quality, interlaced video formats and multichannel audio

- designed for broadcast digital video (e.g. DVD)

MPEG-4

Newer standard offering multiple data rates

- H.264 or MPEG-4 Advanced Video Coding (MPEG-4 AVC)
- Based on Profiles
  - 5Kbps to 4Mbyte/s, with interactive multimedia as well as video
  - Low rate profiles for video-phones.
  - Simple profile, Advanced simple profile
- Allow different compression techniques for different rectangular areas

Widely used:

- HDTV, Blu-ray
- Web: HTML5 video, YouTube, Flash Player, Silverlight

Other CODECs

Low bitrate applications (e.g. 28kbps over a phoneline)

- More sophisticated motion estimation
- Wavelets, fractals etc
- H.261 and H.263
  - designed for video conferencing
  - use multiple compression techniques
  - *"better than nothing"* performance over a phoneline

High bitrate applications

- VC-1
  - Microsoft's alternative to MPEG 4
  - Windows Media Video 9
  - HD and Blu-Ray

Summary of Bit Rates

(from https://en.wikipedia.org/wiki/Bit_rate)

16 kbit/s – videophone quality (minimum necessary for a consumer-acceptable "talking head" picture using various video compression schemes)

128–384 kbit/s – business-oriented videoconferencing quality using video compression

400 kbit/s YouTube 240p videos (using H.264)[21]

1 Mbit/s YouTube 480p videos (using H.264)[21]

1.15 Mbit/s max – VCD quality (using MPEG1 compression)[22]

2.5 Mbit/s YouTube 720p videos (using H.264)[21]

3.5 Mbit/s typ – Standard-definition television quality (with bit-rate reduction from MPEG-2 compression)

4.5 Mbit/s YouTube 1080p videos (using H.264)[21]

Summary of Bit Rates (contd)

9.8 Mbit/s max – DVD (using MPEG2 compression)[23]

8 to 15 Mbit/s typ – HDTV quality (with bit-rate reduction from MPEG-4 AVC compression)

19 Mbit/s approximate – HDV 720p (using MPEG2 compression)[24]

24 Mbit/s max – AVCHD (using MPEG4 AVC compression)[25]

25 Mbit/s approximate – HDV 1080i (using MPEG2 compression)[24]

29.4 Mbit/s max – HD DVD

40 Mbit/s max – 1080p Blu-ray Disc (using MPEG2, MPEG4 AVC or VC-1 compression)[26]

File Formats for Video

Like file formats for still images, they provide a packaging that includes details of the contents of the file, such as the codec used. Various formats, such as...

- Microsoft's AVI & WMV
- QuickTime
- MPEG
- Ogg & WebM: open source (support on some HTML5 browsers)

You need appropriate player and CODECs on your computer to play files.

- Windows Media Player, QuickTime, HTML5 <video>

Video for Windows

AVI: **A**udio and **V**ideo **I**nterleaved

- Video and synchronized audio in a small window at around 15 frames-per-second without specialized hardware
- Video is interleaved with the audio
- Variety of *codecs* (identified by 4-byte code, FOURCC):
  - Cinepak (CVID)
  - Intel Indeo
  - Microsoft Video
  - Microsoft RLE (MRLE)
  - Others

ASF (Advanced Systems Format)

- WMV (Windows Media Video)
- streaming

QuickTime

QuickTime *movie (".mov")* is a framework for organizing and manipulating the data that make up the actual video frames

- many video *codecs* supported
- can contain media other than video
- media not necessarily stored in the *.mov* file

A movie plays according to a time base

- frames may be dropped to keep in time
- enables constant playback speed on any system
- maintains synchronization between video and audio

QuickTime

Component-based architecture

- Plug in new components to deal with new formats or provide new operations
- Components include:
  - Compressors (e.g. Cinepak, Indeo, Sorenson)
  - Sequence grabbers: digitize video
  - Movie player: will handle AVI and MPEG
  - Transcoders: translate between data formats

Open Formats

OGG: open container format

- high quality digital multimedia
- streaming
- multiplex a number of independent streams for audio, video, text (such as subtitles), and metadata.
- CODECS: Theora
  - But OGG files can contain MPEG-4 video and others

WebM: open format for HTML5 <video>

- sponsored by Google
- CODECS: VP8 and VP9

Streamed Video

Playback *as it arrives* over the Internet

Progressive download

- download movie to hard disk
- start playing back when playback time of downloaded content equals download time of remainder of movie
- still a large delay before playback begins and not suitable for "live" video

True streaming

- never stored on user's disk
- must arrive at playback rate
- different movie formats depending on connection speed
- *Streaming QuickTime* and *RealVideo*

File Formats for Animation

Adobe's Shockwave Flash

- Native file format for *Flash* movies
- Vector images and animations
- Metafile: can incorporate bitmap images and sound
- *De facto* standard for animation on the Web
- Browser plugins freely available

SVG / SMIL / X3D

- XML languages for use on web pages

Microsoft's Silverlight

- Programmable web browser plugin
- .NET languages and development tools

Shockwave Flash (SWF)

File contains two broad classes of items…

- Definitions: dictionary of symbols that appear in the animation
- Control items: instructions to place, remove or move a symbol
  - movement specified using transformation matrices

Rather like a program (cf. FIG, POVRay, SVG, X3D)

- stored in compressed binary format: SMALL!