Video Compression 2: Bi-directional Coding

Multimedia Systems (Module 4 Lesson 3)

Summary:

- MPEG Coding
  - Bi-directional Motion-Compensation
- MPEG Parameters
- MPEG 2 and 4

Sources:

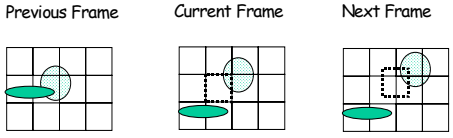
- "Digital Compression for Multimedia: Principles and Standards", Jerry D. Gibson, Toby Berger, Tom Lookabaugh, Dave Lindbergh and Richard L. Baker.
- My research notes

MPEG

- "Moving Picture Coding Experts Group", standards body for delivery of video and audio.
- MPEG-1 Target: VHS quality on a CD-ROM or Video CD (VCD) (352 x 240 + CD audio @ 1.5 Mbits/sec)
- MPEG-2 allows different levels and profiles
- Both standards have four parts:
  - Video: Defines the video compression decoder
  - Audio: Defines the audio compression decoder
  - System: Describes how various streams(video, audio or generic data) are multiplexed and synchronized.
  - Conformance: Defines a set of tests designed to aid in establishing that particular implementations conform to the design.

The Problem

- Some macroblocks need information that is not present in the previous reference frame.
  - Maybe, such information is available in a succeeding frame!

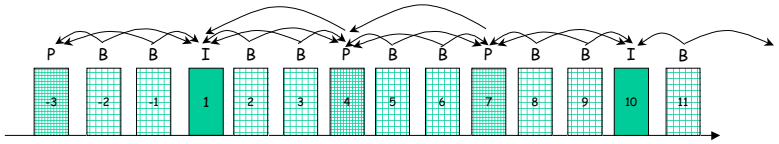


- Add a third frame type (*B-frame*): To form a B-frame, search for matching macroblocks in both *past* and *future* frames.
- Typical pattern is `IBBPBBPBB IBBPBBPBB IBBPBBPBB`
  Actual pattern is up to encoder, and need not be regular.

Bitstream order vs. Display order

Bitstream (Transmit) order:

1(I), 4(P), 2(B), 3(B), 7(P), 5(B), 6(B), 10(I), 8(B), 9(B)



Frame and Macroblock Prediction Types

Some definitions:

- **Anchor frame:** *A frame that can be used for prediction*

We now discuss the various frame types and Macroblock types

Macroblock Type	Prediction
Nonpredicted Macroblock	none
Backward-predicted macroblock	References temporally nearest subsequent anchor frame
Forward-predicted macroblock	References temporally nearest previous anchor frame
Bidirectionally predicted macroblock	Averages predictions from temporally nearest, previous and subsequent anchor frames

Frame Type	Anchor Frame	Macroblock Types
I-frame	Yes	Nonpredicted
P-frame	Yes	Nonpredicted, Forward predicted
B-frame	No	Nonpredicted, Forward predicted, Backward predicted, Bidirectionally predicted

Bidirectional Prediction

- B-frames allow effective prediction of uncovered background, areas of the current picture that were not visible in the past and visible in the future.
- B-frames can provide for interpolation equivalent to an even finer degree than half-pixel (1/4 pixel for example).
- If good prediction is available in both the previous and subsequent anchor frames, then averaging the two predictors reduces noise and hence increases efficiency.
- -ve: Motion estimation becomes more complex (look farther).
- A ratio of 5:3:1 between the number of bits spent on I, P and B frames is quite common.
- Errors in B-frames tend to limit the effect to that B-frame only.

MPEG Notation

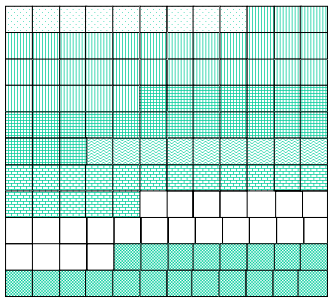
Though the standard does not dictate this, the pattern (order) of frames are commonly referred by the following notation: (*N,M*) where:

- *N* is the number of frames from one I-frame(inclusive) to the next (exclusive).
- *M* is the number of frames from one anchor(inclusive) to the next(exclusive).

The example sequence discussed before would be a (*N=9, M=3*) pattern.

Differences from H.261

- Larger gaps between I and P frames, so need to expand motion vector search range. Uni-Quant for P and Non-uniform-Quant for I.
- To get better encoding, allow motion vectors to be specified to fraction of a pixel (1/2 pixel).
- Bitstream syntax allows random access, forward/backward play, etc.
- Added notion of *slice* for synchronization after loss/corrupt data (see figure at right: 7 slices in frame).



MPEG-1 Parameter Constraints

Parameter	Constraint
Horizontal size	< 720 pixels
Veritcal size	< 576 lines
Total # of macroblocks per picture	< 396
Total # of macroblocks per second	< 396 x 25 (or 396 x 30)
Frame rate	< 30
Bit rate	< 1.86 Mbps
Decode buffer	< 376832 bits

MPEG2

- Unlike MPEG-1 which is basically a standard for storing and playing video on a single computer at low bit-rates, MPEG-2 is a standard for digital TV. It meets the requirements for HDTV and DVD (Digital Video/Versatile Disc).
- MPEG2 Supports the following levels:

Level	Size	Pixels/sec	Bit-rate(Mbps)	Application
Low	352 x 288 x 30	3 M	4	consumer tape equiv.
Main	720 x 576 x 30	12 M	15	Studio TV
High 1440	1440 x 1152 x 60	96 M	60	Consumer HDTV
High	1920 x 1152 x 60	128 M	80	Film Production

- It supports multiple profiles based on scalability
- Supports both field prediction and frame prediction.
- Besides 4:2:0, also allows 4:2:2 and 4:4:4 chroma-subsampling

MPEG4

- Originally targeted at very low bit-rate communication (4.8 to 64 Kb/sec), it now aims at the following ranges of bit-rates:
  - video -- 5 Kb to 5 Mb per second
  - audio -- 2 Kb to 64 Kb per second
- It emphasizes the concept of *Visual Objects* --> Video Object Plane (VOP)
- Objects can be of arbitrary shape, VOPs can be non-overlapped or overlapped
- Supports content-based scalability
- Supports object-based interactivity
- Individual audio channels can be associated with objects
- Good for video composition, segmentation, and compression; networked VRML, audiovisual communication systems (e.g., text-to-speech interface, facial animation), etc.
- Standards being developed for shape coding, motion coding, texture coding, etc.