Multimedia Software Systems
CS4551

Video Compression Basic - II

Types of predictions

- Types of predictions
  - I frame (no prediction)
  - P frame (prediction based on a previous frame)
  - B frame (bidirectional prediction)
  - Multiframe

1
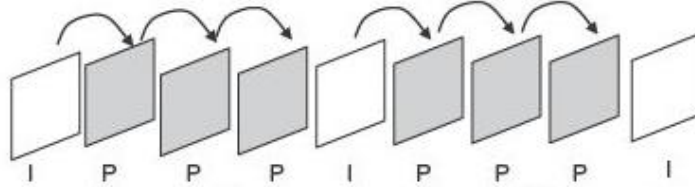
Types of Predictions – I frame

- I frame
  - Refers to a video frame compressed by *intraframe* coding.
  - Is encoded with no reference to any past or future frames.
  - Is encoded using the standard JPEG pipeline.
  - Is compressed using only spatial redundancy.
  - Requires more bits than other types of frames in general.

Types of Predictions – P frame

- P frame
  - Is compressed by using *interframe* coding.
  - Is encoded with reference to a preceding (I or P) frame.
    - The motion vectors are estimated to create a predicted frame. The motion vectors are encoded losslessly.
    - The error frame created by the difference between the reconstructed frame and the current frame. This error frame is bound to have lower entropy. The error frame is coded using the JPEG pipeline. Because the error frame has lower entropy, the coding uses higher- valued quantization tables.
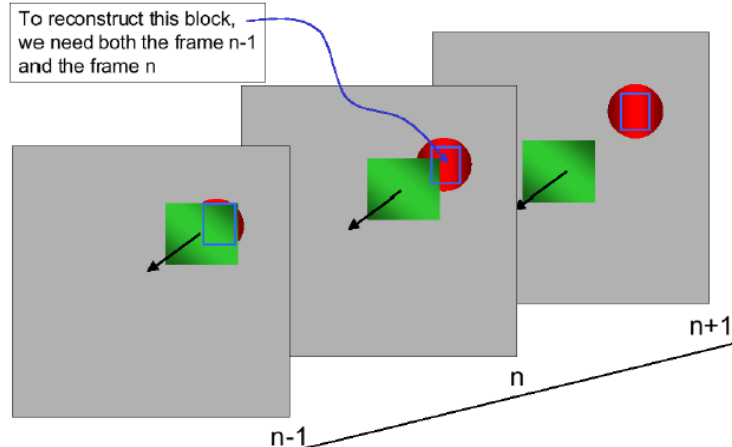
Types of Predictions – B frame

- B frame
  - B frames, also known as bidirectionally coded frames, are coded using the *interframe* coding and also exploit temporal redundancy.
  - To predict a B frame, the previous or past (I or P) frame and the next or future (I or P) frame are used. Using both frames, therefore, increases the correctness in prediction during motion compensation.

Bidirectional Motion Example

To reconstruct this block, we need both the frame n-1 and the frame n

Types of Predictions – B frame

- B frame
  - The coding of B frames is more complex compared with I or P frames with the encoder having to make more decisions.
  - To compute a matching macroblock, the encoder needs to search for the best motion vector in the past reference frame and also for the best motion vector in the future reference frame. Up to two motion vectors are computed for each macroblock.
  - The macroblock gets coded in one of three modes:
    - Prediction using only the past frame
    - Prediction using only the future frame
    - Interpolated, using both by averaging the two predicted blocks
  - Frames, therefore, might need to encode up to *two motion vectors per macroblock*.
  - Searching in two frames, although useful for compression, needs more computation time.

CSULA CS4551 Multimedia Software Systems by Eun-Young Kang

time

Prediction for B-frame based on a previous I or P frame and a next I or P frame.

CSULA CS4551 Multimedia Software Systems by Eun-Young Kang
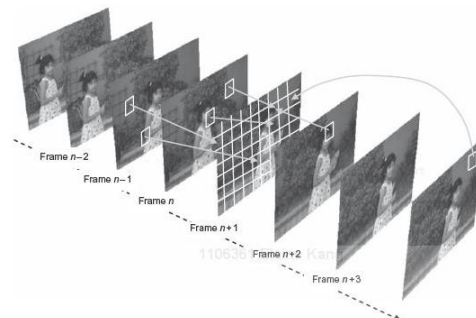
4
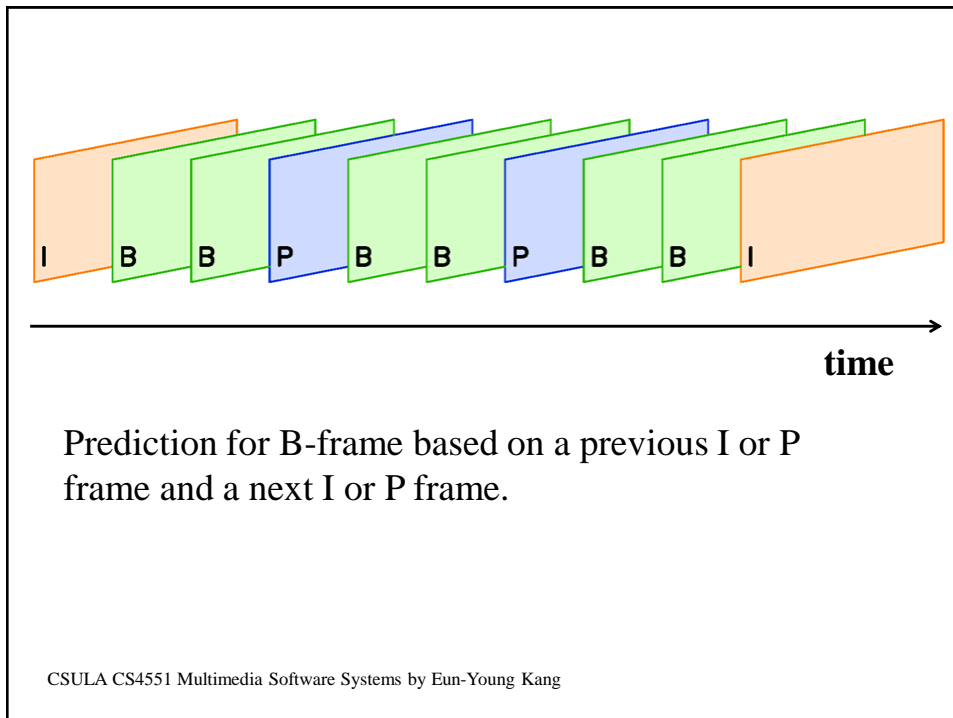
Bidirectional Motion Compensation

- A block of a bidirectionally coded frame is predicted from both a previous frame and a later frame
- Bidirectional Motion Compensation implies:
  - A *delay* (to decode the current frame, we need to have received and decoded a later frame)
  - *Different order* between the sequence of acquired and displayed frames and the sequence of encoded frames

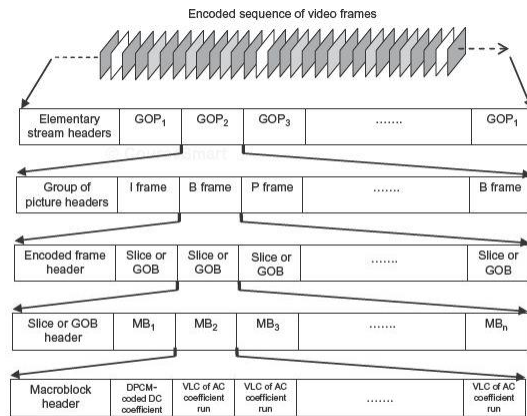| 1	2	3	4	5	6	*Frame numbers in sequential (capture) order*
I	P	P	B	B	P	*prediction type per frame*
1	2	3	6	4	5	*frame numbers in encoding/decoding order*

Types of Predictions – Multiframe

- Multiframe
  - A macroblock is multipredicted using two or more frames from the past or future.
  - This increases the search time and space substantially, but can give better predictions and, hence, lower bit rates.
  - With the availability of faster computers, multiframe predictions are now gaining popularity in recent standards.

Frame $n-2$
Frame $n-1$
Frame n
Frame $n+1$
Frame $n+2$
Frame $n+3$

Video Structure – A Sequence

Encoded sequence of video frames

Elementary stream headers	GOP$_1$	GOP$_2$	GOP$_3$	....	GOP$_1$

Group of picture headers	I frame	B frame	P frame	....	B frame

Encoded frame header	Slice or GOB	Slice or GOB	Slice or GOB	....	Slice or GOB

Slice or GOB header	MB$_1$	MB$_2$	MB$_3$	....	MB$_n$

Macroblock header	DPCM-coded DC coefficient	VLC of AC coefficient run	VLC of AC coefficient run	....	VLC of AC coefficient run

*Video structure. A typical bit stream structure is **hierarchically encoded** with GOPs, frames, slices or GOBs, macroblocks, and runs of coefficients.*

• The encoded video bit stream can be looked at **hierarchically** with the video frames divided into GOPs.
• Each picture is further divided into slices or GOBs.
• A GOB consists of macroblocks.

CSULA CS4551 Multimedia Software Systems by Eun-Young Kang

Video Structure - GOP

- A variety of video coding standards, such as MPEG, have adopted a hierarchical description to coded sequences known as a **GOP** (*Group Of Pictures)*.

- A GOP is a series of one or more frames to assist in random access of the frame sequence.

- The first coded frame in the group is an I frame, which is followed by an arrangement of P and B frames.

- The I frames and P frames are known as *anchor frames* and are used to predict B frames.

CSULA CS4551 Multimedia Software Systems by Eun-Young Kang
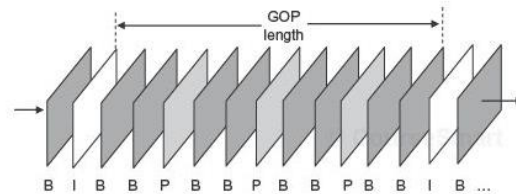
Video Structure – GOP (2)
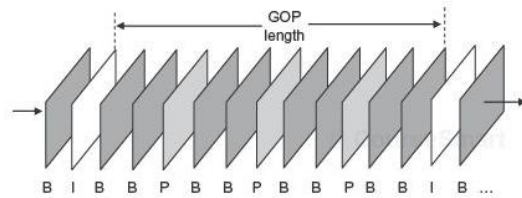
Group of Pictures (GOP). An example of a GOP used in
MPEG-1 is shown. Here, the GOP length is 12 with three P frames
between I frames and two B frames between P frames.

- The GOP length is defined as the distance between I frames.
- Standard codecs represent a GOP with two parameters N and M.
  - N is the size of the GOP that gives the number of frames in the GOP.
  - M specifies the distance between the anchor frames (I and/ or P).
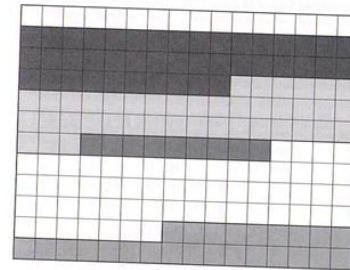 - Sown in the Figure, N 12, M 3.

Video Structure – GOP (3)

Group of Pictures (GOP). An example of a GOP used in
MPEG-1 is shown. Here, the GOP length is 12 with three P frames
between I frames and two B frames between P frames.

- Alternatively The group of pictures can be of any length, but must have at least one I frame in any GOP.
- GOPs are very useful for applications that require *random access* to the bit stream— for example, fast forward or fast rewind, where the application can randomly access the I frame, decode it, and display it, not worrying about B and P frames. Applications that need support for such features normally have small GOP sizes.

Video Structure - GOB

- A frame-level hierarchy: GOB or Slices
  - Additionally, each frame might also be divided hierarchically into groups of macroblocks called slices or GOBs.
  - The earlier standards, such as H. 261, use the term GOBs, whereas the later MPEG standards use the term slice.
  - A frame- level hierarchy allows us to have the ability to distribute bits more flexibly within a frame. Some slices might be given more bits because of higher entropy, compared with other slices. This allows for variable-length code resetting to prevent error propagation and, consequently, better quality encoding.

Slices in an MPEG-1 picture.

CSULA CS4551 Multimedia Software Systems by Eun-Young Kang