SOURCE : Japan
TITLE   : Considerations on Processing Delay with Hybrid Coding
PURPOSE : Discussions

1. Introduction

This document[1] discusses processing delay of hybrid type codecs. A block diagram of a typical codec is shown, and processing delay is estimated for each block.

Related previous documents are: AVC-109 (Status Report on ATM Video Coding Standardization, Issue 2, September 24, 1991), Section 4.3.3, and MPEG91/100 (Proposal Package Description for MPEG Phase 2, Issue 3, August 23, 1991), Section 3.4, (3), both of which states the processing delay target of 150mS.

2. Assumptions

A hybrid type codec using motion compensated prediction and DCT is assumed. An MPEG like bidirectional prediction method is also assumed to include more delay demanding situations than with H.261. A picture format is assumed to be SCIF (Super CIF, 720x576x59.94 progressive format) as a common base for both 525 and 625 line systems. The concept of SCIF is first proposed in AVC-29, and actual methods of format conversion into/from SCIF is discussed in AVC-80.

3. Codec configurations

Figure 1 shows a typical block diagram of a hybrid codec. Each element of the block is classified into three categories according to its processing delay:

(DDD)   more than 1 frame time,

(DD)   one severalth of 1 frame time, typically in the order of 16 line interval.

(D)   considerably smaller than 1 frame time, typically several times a macroblock time.

(null) element (may) exist, but not to be included to consist processing delay, or to be discussed jointly with other part of the system.

---

[1] Also submitted to the joint part of the session with MPEG.

4. Estimates of processing delay

- C1: Pre-processing (from camera to source coder)

  - C1a: Format conversion (camera output format to SCIF) → (DDD)
  - C1b: Prefiltering such as for noise reduction → (D)
  - C1c: Input buffer → (DD),
    assuming camera/coder synchronization and post-buffer implementation.
  - C1d: Frame re-ordering for bidirectional prediction → (DDD)

- C2: Source coding

  - C2a: Prediction (MC, loop filter) → (D)
  - C2b: DCT → (D)
  - C2c: Q → (D)
  - C2d: Local decode including IQ and IDCT → (null),
    not to be included for encoders without frame skips.

- C3: Video multiplex coding (Q index to video multiplex out)

  - C3a: VLC → (D)
  - C3b: Video multiplex → (D)

- C4: Transmission coding (video multiplex out to channel input)

  - C4a: Transmitter output buffer → (DDD)
  - C4b: Error correction coding → (D or DD)
  - C4c: AAL coding → (D or DD)
  - C4d: AAL cell assembly delay → (D)

- D1: Transmission decoding (channel output to video de-multiplex input)

  - D1a: AAL cell de-assembly delay → (D)
  - D1b: AAL decoding → (D or DD)
  - D1c: Error correction decoding → (D or DD)
  - D1d: Transmitter input buffer → (null), to be discussed with C4a.

- D2: Video de-multiplex (video de-multiplex output to IQ input)

  - D2a: Video de-multiplex → (D)
  - D2b: VLD → (D)

- D3: Source decoding (Q index to video output)

Doc.AVC–124

- D3a: IQ → (D)
- D3b: IDCT → (D)
- D3c: Prediction (MC, loop filter) → (D)

- D4: Post processing (source decoder output to monitor input)

  - D4a: Frame re-ordering for bidirectional prediction → (DDD)
  - D4b: Postfiltering for noise reduction → (D)
  - D4c: Format conversion (SCIF to monitor input) → (DD)
  - D4c: Display buffer when decoding and display not synchronized → (DDD)

5. Discussions

- Format conversions (C1a and D4c) require one field to several frame delay, depending on the complexity and necessity of each conversion process such as interlace-to-noninterlace conversion, line number conversion, and frame rate conversion.

- Frame re-ordering for bidirectional prediction (C1d and D4a) causes processing delay as is shown in Fig. 2. Total amount of this delay is equal to core frame interval, i.e., the number of consecutive B-pictures plus one.

- Transmission buffer delay (C4a+D1d) is a constant in CBR mode operation, the actual amount of which to be determined by the fluctuation of the bit amount for each frame. It becomes variable in VBR, but can be upperbounded by Fmax/Peak as is discussed in AVC-90, where Fmax is the maximum allowable amount of encoded data for one frame (usually occurs at I-pictures) and Peak is the peak bit rate.

- Codec realization with almost no delay at transmission buffer is also discussed in AVC-89. Realization of such codec must further be pursued, considering the actual implementation of coding parameter control method that both meets the image quality requirement and an imposed UPC scheme by a network.

- Display buffer delay (D4c) can be eliminated by using a possible AAL function for source video clock recovery at a decoder.

6. Conclusion

Processing delay with hybrid coding is discussed. "Format conversion" and "Frame re-ordering" are the two main factors for processing delay.

Doc.AVC–124

References

AVC-29: "Common Progressive Picture Format for High Quality Applications",
B, F, FRG, I, N, NL, UK, Paris, May, 1991.

AVC-80: "Conversion Simulation between 525 line Interlaced Format and SCIF",
Japan, Santa Clara, August 1, 1991.

AVC-89: "Considerations on the window size of traffic descriptor",
Japan, Santa Clara, August 1, 1991.

AVC-90: "Considerations on delay with VBR coding",
Japan, Santa Clara, August 1, 1991.

AVC-109: "Status Report on ATM Video Coding Standardization, Issue 2",
September 24, 1991.

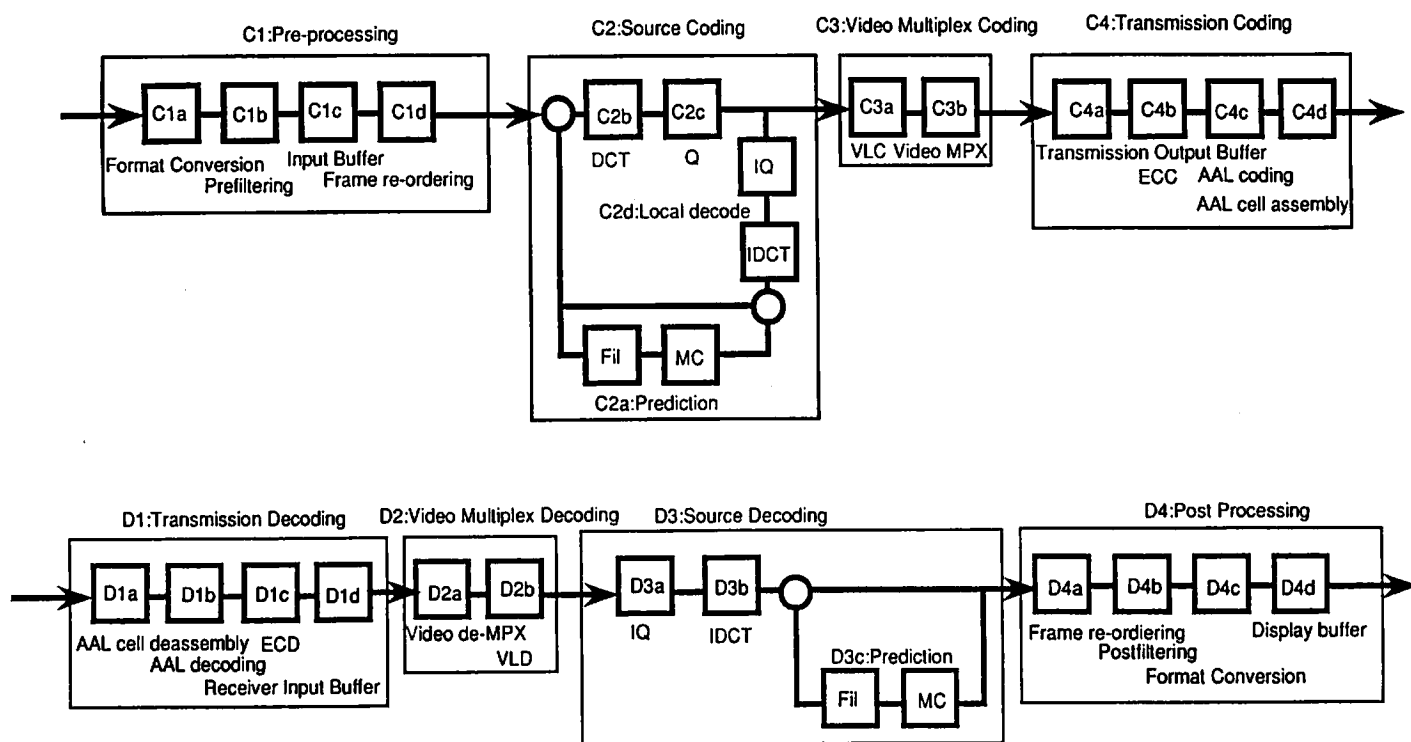MPEG91/100: "Proposal Package Description for MPEG Phase 2, Issue 3",
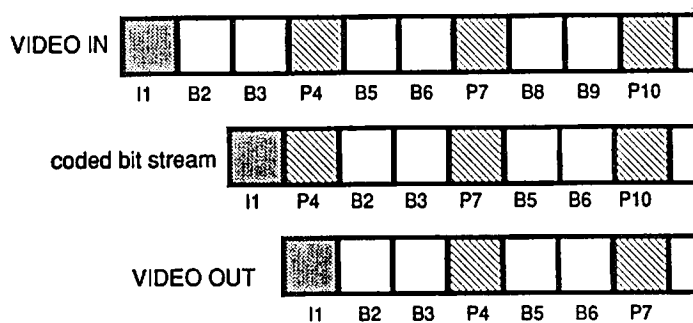August 23, 1991.

Fig. 1 Block diagram of a hybrid codec

Fig. 2 Frame re-ordering

Doc.AVC–124