

Proposal of solution for problems of hybrid - perceptual bitstream project

NTT

Keishiro Watanabe

Jun Okamoto

2007/9/13

Introduction

2007/9/13

The demand for high-reliability, light-weight objective video quality assessment is increasing.

Therefore, the result of the hybrid perceptual / bitstream project is urgently required for providing rapid standardization in the industry.



However, there are specific problems caused by using bitstream information in the hybrid perceptual / bitstream project.



In our documents, NTT shows two proposals for those problems.

Discussion (1)

2007/9/13

Logic structure of our three documents

overview of two problems and solutions

“Proposal of solution for problems of Hybrid – Perceptual Bitstream Project” **(Document 1)**

Proposal 1:

Input information of P.NAMS **(Document 2)**

Proposal 2:

Output information of bitstream analyzer
(Document 3)

Discussion (2 - 1/2)

2007/9/13

Proposal 1

The scope of J.bitvqm studied in ITU-T SG9 should include the perceptual / bitstream model (FR, RR, NR) and the scope of P.NAMS studied in ITU-T SG12 should include the parametric packet layer and bitstream (w/o PVS) model.

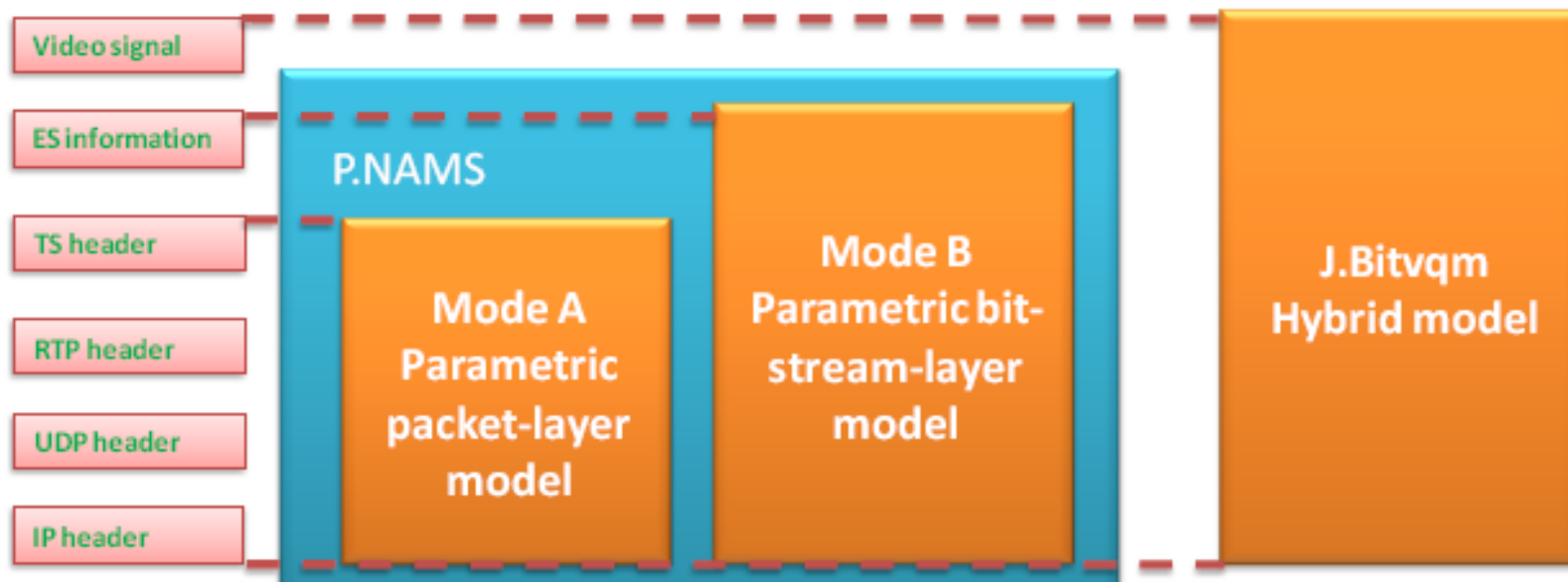


- Main purpose of P.NAMS (under consideration) will be to assess video quality for in-service quality monitoring by using parametric bitstream information without decoding.
- the scope of J.bitvqm assumes to use both PVS and bitstream information

Discussion (2 - 2/2)

2007/9/13

The detail of proposal 1



	Parametric packet-layer model	Parametric bit-stream-layer model	Hybrid model
Average quality	○	○	○
Packet-loss pattern	○	○	○
Content dependence	-	○	○
Decoder characteristics	-	-	○
Display characteristics	-	-	△

Discussion (3 - 1/6)

2007/9/13

Proposal 2

We propose to allow using all syntax information included in an elementary stream as the output information of a bitstream analyzer (assumed to use P. NAMS mode B in previous proposal), which is defined in the standardized recommendation of a coding method.



- The estimation accuracies of objective video quality assessments, especially the NR perceptual / bitstream model and parametric bitstream (w/o PVS) model, will be not enough because these models cannot use the information of a source video
- In principle, we can get syntax information in an elementary stream of a coded video sequence from a bitstream analyzer. That information has a strong correlation with the characteristics of a source video scene and will have a large influence on subjective video quality

Discussion (3 - 2/6)

2007/9/13

The verification of solution 2

Overview

- Subjective video quality assessment of coding distortion of HD H.264
- Extraction of some ES syntax information from coded H.264 bitstream
- Comparison of subjective video quality, ES layer syntax information, and bit-rate.



Construction of two parametric bitstream models for H.264 coding distortion and verification of the level of estimation accuracy of parametric bitstream model

Discussion (3 - 3/6)

2007/9/13

Creation of coded video sequences

- H. 264 high profile
 - HD: 1080i, 59.94 fields/s
- Bit-rate: 20, 18, 16, 14, 12, 10, 8, 6, 4, 2 (Mbps)
- 8 video sequences from ITU-R BT.1210 (10 seconds)

Discussion (3 - 4/6)

2007/9/13

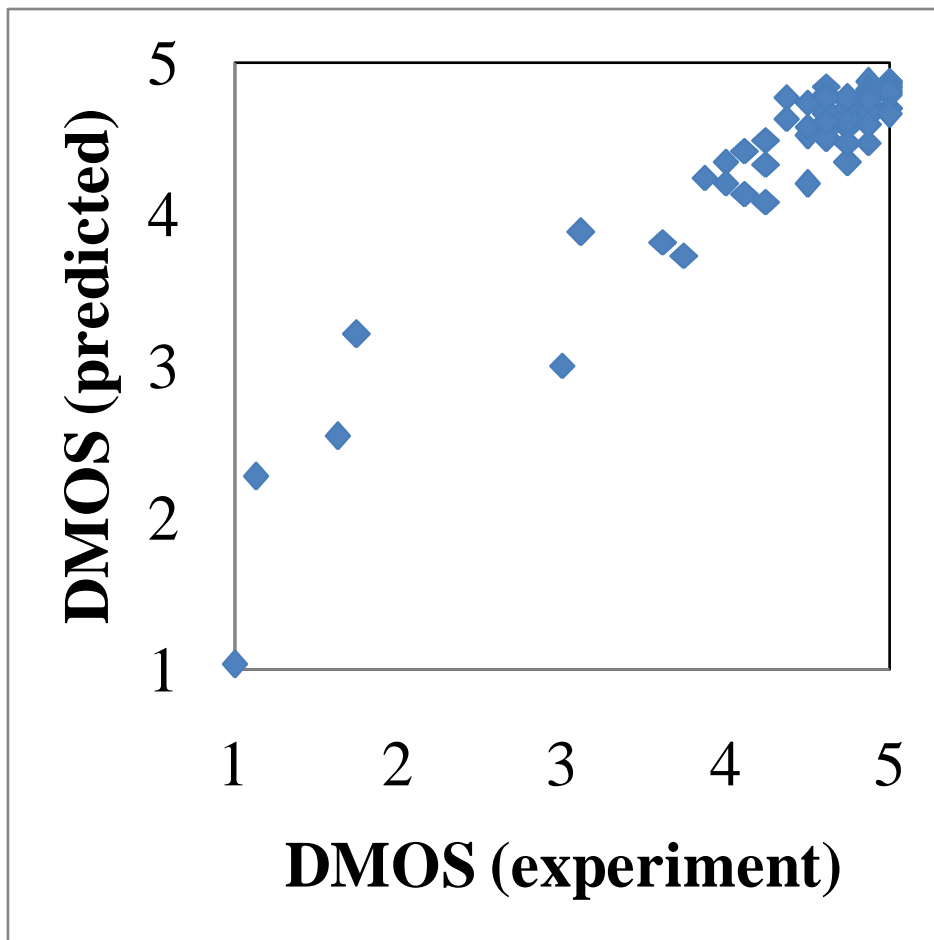
Condition of subjective video quality assessment

- Grading scale: DMOS (DSIS method)
- Number of subjects: 24 (12 men and 12 women)
- Viewing condition:
 - Monitor: Sony 32 inch CRT
 - Viewing distance: 3H
 - Illumination Intensity: 200lx

Discussion (3 - 5/6)

2007/9/13

DMOS (predicted by using ES layer parameters) vs DMOS (experiment)



Correlation coefficient = 0.97
RMSE = 0.32

Conclusion

2007/9/13

We propose to use all syntax information in elementary stream, which can consider scene characteristics (coding difficulty) as the output of bitstream analyzer.

Our proposals

2007/9/13

- (1) We proposed that the scope of P.NAMS mode A and B. We also showed that tentative relationship of the scope of P.NAMS, J.bitvqm, and hybrid perceptual / bitstream project.**

- (2) All syntax information included in an elementary stream should be used as the output of a bitstream analyzer because it contains information that can consider scene dependent factors.**

Appendix

2007/9/13

Extraction of ES layer syntax information from coded H.264 bitstream

Captured IP packet stream



Extract

MPEG2 transport stream



Demultiplex

H.264 Elementary stream (ES)



Extract

H.264 syntax information