Question(s):	VQEG	**Meeting, date:**	May 7-11, 2007
Study Group:	**Working Party:**	**Intended type of document** (R-C-D-TD):TD	
Source:	NTT (Nippon Telegraph and Telephone Corporation), Japan		
Title:	Proposal of scope of hybrid - perceptual / bitstream project		
Contact:	Keishiro Watanabe NTT Japan		Tel: +81-422-59-7204 Fax: +81-422-59-5671 Email: watanabe.keishiro@lab.ntt.co.jp
Contact:	Jun Okamoto NTT Japan		Tel: +81-422-59-6526 Fax: +81-422-59-5671 Email: okamoto.jun@lab.ntt.co.jp

Please don't change the structure of this table, just insert the necessary information.

Summary

This contribution investigates the scope of the hybrid – perceptual / bitstream project. Then, it proposes to split the scope into two different areas taking into account several application scenarios and consequently the complexity of the quality assessment method. That is, one is specifically for non-intrusive quality monitoring with light-weight implementation, and the other is for scenarios that do not require low complexity, e.g., quality benchmarking in an intrusive manner.

1. Introduction

The hybrid – perceptual / bitstream project was started at the last VQEG meeting in Tokyo in September 2006. Accordingly, ITU-T SG9 launched a new study item called J.bitvqm.

We think that the aim of this project should include the standardization of a quality assessment method that enables in-service non-intrusive video quality monitoring. However, the scope of J.bitvqm is not necessarily clear yet. Therefore, in this contribution, we show the quality monitoring scenarios assumed for speech communication services in ITU-T Recommendation P.564, which provides a means for monitoring speech quality in IP telephony services, and propose to split the scope of our project into two different areas.

2. Discussion

One of the applications of the hybrid - perceptual / bitstream model can be non-intrusive quality monitoring, which is also studied in WG2 of ITU-T Focus Group – IPTV

(FG-IPTV) [1]. A similar study has already been conducted for IP telephony services in ITU-T SG12 and resulted in the creation of Recommendation P.564 [2]. The quality monitoring scenarios described in P.564 are introduced in Figures 1, 2, and 3.

In these scenarios, it is often difficult to obtain decoded media signals and to calculate perceptual degradation based on them due to, for example, unavailability of computational power. This is especially true in the case of the embedded scenario in Figure 3. In addition, at the mid-point in the network, it is impossible to obtain the decoded signal from the end device. Therefore, in these scenarios, it is not necessarily appropriate to adopt a model utilizing decoded video media signals.

This is the rationale for creating Recommendation P.564, which estimates subjective quality solely from packet-header information. To provide such a method for multimedia services including IPTV, Q.14 of SG12 launched a new study item called P.NAMS [3], whose framework is depicted in Figure 4 in comparison with J.bitvqm.

Although P.NAMS will be a very useful tool to monitor the quality of multimedia services in a very light-weight way, it cannot take into account the effects of video content on the subjective quality. That is, video quality cannot be estimated accurately without looking into the payload information.

3. Proposal

For the reasons mentioned in the previous section, we propose to develop a model that utilizes the payload information without an access to decoded media signals. To show the relationship among the current scope of J.bitvqm, P.NAMS, and the proposed new approach, we present Figure 5, which shows the information used in models in relation to the protocol stack. The proposed approach is something in between P.NAMS and current J.bitvqm.

Because the motivation of the proposed approach is to provide a model with light computational load, we propose to evaluate the validity of such models in terms not only of the accuracy of quality estimation but also of computational complexity.

Reference:

[1] ITU-T FG IPTV-DOC-0066, “Working document - Performance monitoring for IPTV,” Jan. 2007.

[2] ITU-T Recommendation P.564, “Conformance testing for narrowband voice over IP transmission quality assessment models,” July 2006.

[3] ITU-T SG12 TD202 (GEN/12), “Liaison Statement to VQEG on SG 12 activities on multimedia streaming quality performance modeling,” Jan. 2007.

(1) Dynamic operation scenario: perceived quality is estimated using RTP stream information received at the quality monitoring device and parameters of network elements collected using RTCP-XR.

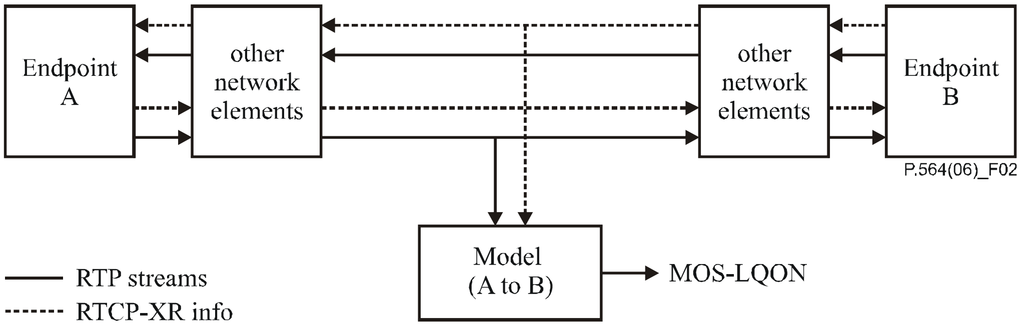


Figure 1. Dynamic operation scenario (cited from P.564).

(2) Static operation scenario: perceived quality is estimated using RTP stream information received at the quality monitoring device and network element parameters (endpoint/reference information).

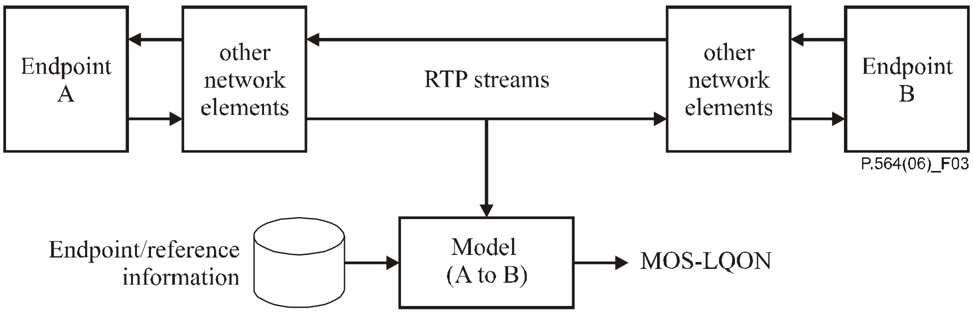


Figure 2. Static operation scenario (cited from P.564).

(3) Embedded scenario: perceived quality is estimated using RTP stream information received at the endpoint and endpoint parameters (not using analog output).

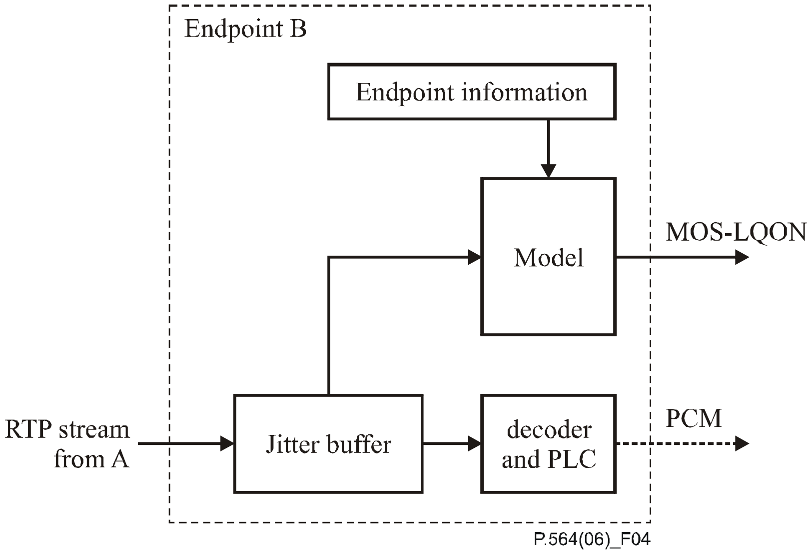
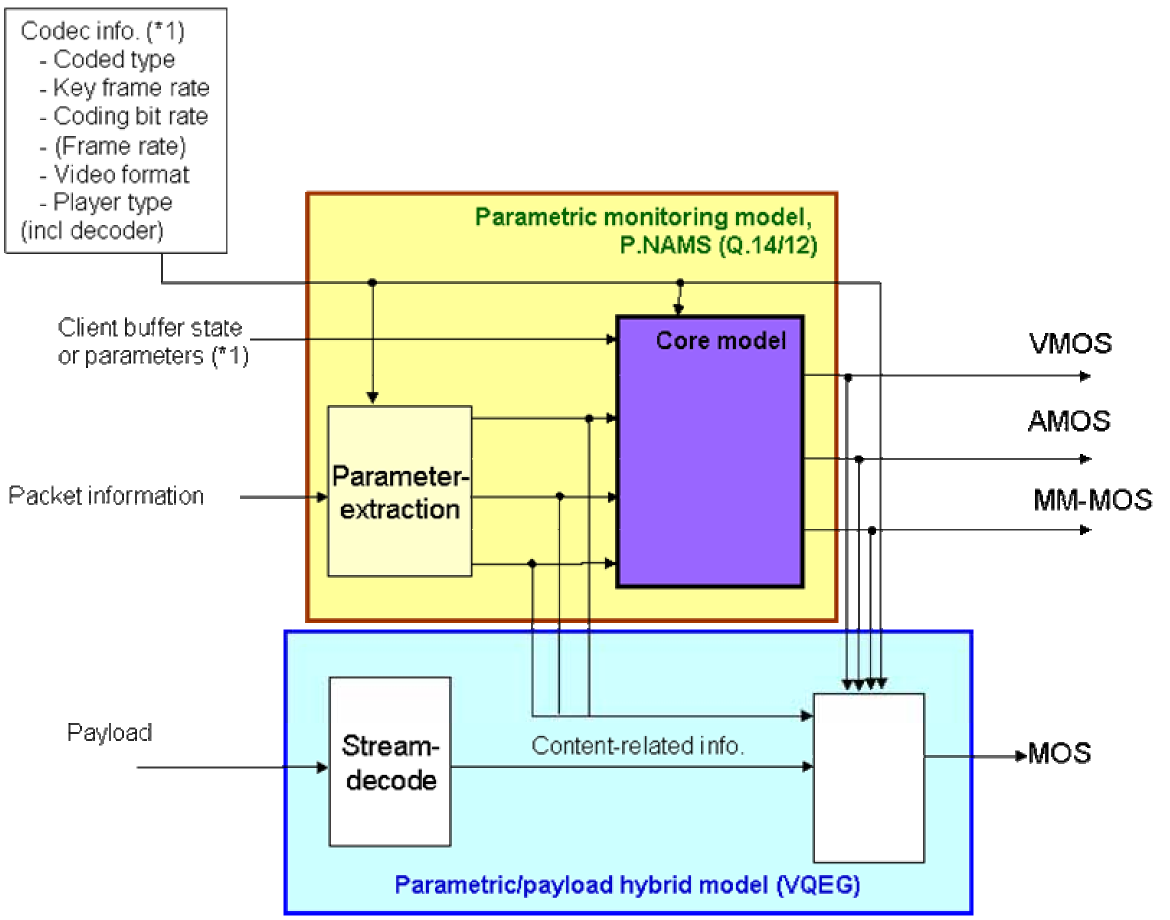
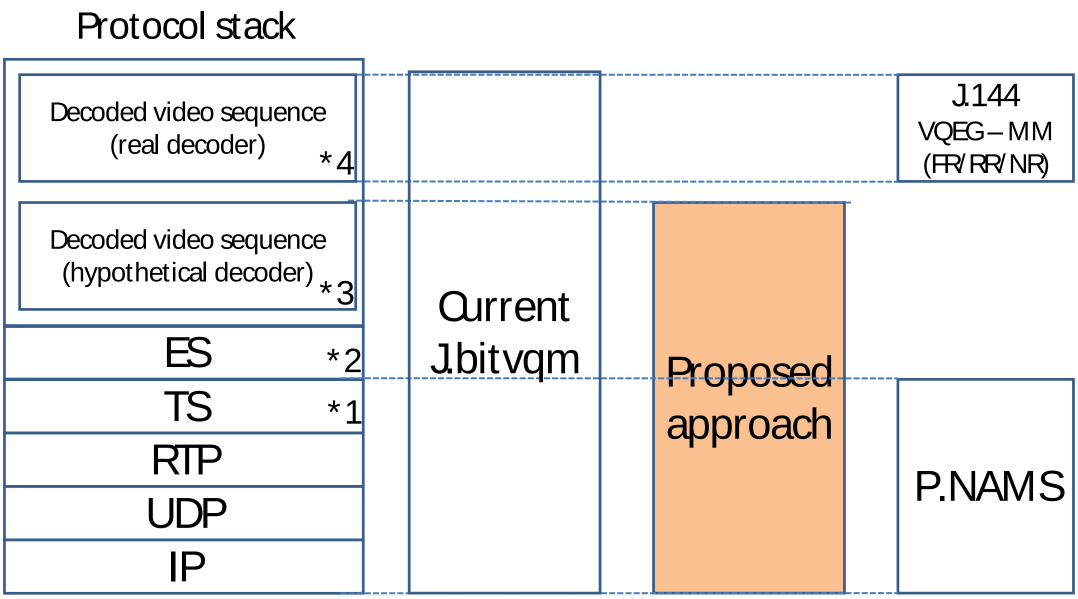


Figure 3. Embedded scenario (cited from P.564).



*1 State of client (playing/buffering), Estimation of buffer behaviour based on for example buffer size/ initial buffering time, rebuffering timers

Figure 4 Relationship between P.NAMS and J.bitvqm.



*1: TS is used for broadcasting applications.
*2: ES is composed of encoding parameters, for example, bitrate, frame rate, resolution, intra refresh rate, DCT coefficients, motion vectors and any other syntax parameters.
*3: Decoded video sequence (hypothetical decoder) Is a video signal virtually decoded by using a reference decoder when the decoded video sequence (real decoder) is not available.
*4: Decoded video sequence (real decoder) is the video signal actually decoded by the user's decoder, for example, a set-top box.

Figure 5. Proposed approach.