SUB-PICTURE: ROI CODING AND UNEQUAL ERROR PROTECTION

Miska M. Hannuksela, Ye-Kui Wang**, and Moncef Gabbouj****

*Nokia Mobile Software, Tampere, Finland
**Tampere International Center for Signal Processing (TICSP), Tampere, Finland
*** Tampere University of Technology, Finland

ABSTRACT

Region-of-interest coding and unequal error protection are two important tools in video communication systems to improve received visual quality. One common property of the two techniques is that unequal coding or transmission is applied to improve the quality of the most important parts of images. The proposed sub-picture coding technique facilitates both region-of-interest coding and unequal error protection by partitioning images to regions of interest and separating the corresponding coded data units from each other. Simulation results show that the overall subjective quality is considerably improved compared to the conventional coding schemes.

1. INTRODUCTION

1.1. Slice-Based Coding

Modern video coding standards, such as ITU-T Recommendation H.263 and ISO/IEC MPEG-4 Part 2, allow division of coded pictures to slices. Slices can be regarded as a way to split a bit-stream to transport packets that can be decoded independently. While spatial and syntactical prediction is disabled across slice boundaries, motion vectors may cross slice edges. This fact causes spatio-temporal error propagation, when motion vectors point to areas that are reconstructed incorrectly. In order to prevent this phenomenon, H.263 includes the optional independent segment decoding mode (H.263 Annex R). When this optional mode is in use, slice boundaries are treated as picture boundaries, and therefore no spatio-temporal error propagation over slice boundaries occurs. Due to restricted motion prediction, compression efficiency drops compared to normal slice-based operation.

1.2. Region-Of-Interest Coding

Region-of-interest (ROI) coding refers to techniques that code some spatial pictorial regions using more refined method than the rest. One method is to apply smaller quantization parameters (QP) hence allocate more bits to the region of interest [1]. As a result, better quality can be observed in the interested region; and hence the overall subjective quality is improved. A shortcoming of ROI coding is that it may cause a boundary effect between the ROI region and the leftover region due to unequal coding.

1.3. Unequal Error Protection

Unequal error protection (UEP) refers to techniques that protect part of the transmitted bit-stream in the transport system better than the rest. Examples of applicable UEP techniques include application-layer selective retransmission [2][3], transport-layer forward error control (e.g. RFC 2733 [4]), guaranteed network Quality of Service (e.g. QoS architecture of Universal Mobile Telecommunication System [5]), and Differentiated Services (DiffServ) [6][7].

In order to apply unequal error protection, video bit-streams have to be organized in portions of different importance in terms of visual quality. Techniques achieving this goal include data partitioning and scalable video coding. Data partitioning refers to a technique where subjectively equally important codewords of all macroblocks (MBs) in a slice are partitioned into a continuous block of data. Typically, MB headers and motion information form one partition and coded prediction error blocks form another partition. Scalability refers to the capability of a compressed sequence to be decoded at different bit-rates. Scalability can be further categorized into temporal, SNR, and spatial scalability. Video coding schemes with temporal scalability is reviewed in [8], spatial and SNR scalable video coding are reviewed in [9].

1.4. Overview of the Paper

Conventional ROI coding treats the entire bit-stream equally in transmission, while data partitioning and scalable coding generally treat coding of an entire image equally in spatial domain. Arbitrarily shaped objects [10], as defined in the MPEG-4 standard, can be used to extract the regions of interest. However, the arbitrary shape coding often inherits relatively high processing requirements.

This paper proposes a simple region-based coding method, called sub-picture coding, which is based on rectangular shapes. It can be considered as a simplification of arbitrary shape coding. While sub-picture coding does not provide means to define object boundaries accurately, it suits simultaneously ROI coding and unequal error protection. In addition, a gradual ROI bit allocation method is proposed to reduce the boundary effect between the selected ROI region and the leftover region.

Section 2 of the paper presents the sub-picture coding method. The gradual ROI bit allocation method is described in Section 3. Simulation results of the proposed technique are provided in Section 4. The paper is concluded in Section 5.

2. SUB-PICTURE CODING

The proposal adds a sub-picture coding layer between picture and slice layers. Sub-pictures are rectangular except for the so-called background sub-picture, which consists of the picture area not falling to any of the rectangular sub-pictures. Rectangular sub-pictures are also referred to as foreground sub-pictures. Sub-pictures boundaries are aligned with macroblock (MB) boundaries. Sub-pictures do not overlap. A slice resides within one sub-picture only. A slice in a background sub-picture may not contain spatially adjacent MBs within a MB line, as it can be intervened by foreground sub-pictures. Figure 1 shows an example of a picture having one foreground sub-picture and one background sub-picture, both of which include two slices.

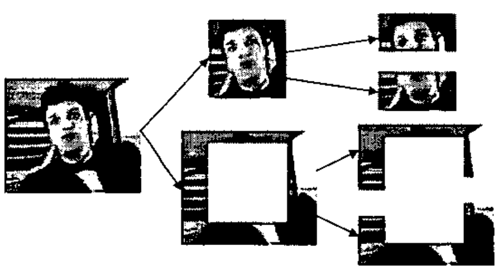

Fig. 1. Example of picture, sub-pictures, and slices.

There are two coding modes associated with sub-pictures: normal mode and independent mode. In the normal mode, sub-picture boundaries are treated as slice boundaries. In the independent mode, boundaries of foreground sub-pictures are treated as picture boundaries and sub-picture segmentation is static over a group of pictures. In other words, in the independent mode, temporal and spatial prediction over sub-picture boundaries is prevented when coding the foreground sub-pictures. No such limitation exists when coding a background sub-picture.

Coding of the foreground sub-pictures takes place before coding of the background sub-picture, therefore different coding methods can applied conveniently to different regions, and the bit-streams of different regions are separated to suit UEP.

3. GRADUAL ROI BIT ALLOCATION

Conventional ROI bit allocation methods, such as the methods proposed in [1], assign two QP values, generally with a large difference, for each picture, one for the ROI region and the other for leftover region. As a result, annoying boundary effects are caused between the two regions. To reduce the boundary effects, a gradual ROI bit allocation scheme described in the following is proposed.

0	0	0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	1	1	0	0
0	0	1	2	2	2	2	2	1	0	0
0	0	1	2	3	3	3	2	1	0	0
0	0	1	2	3	3	3	2	1	0	0
0	0	1	2	2	2	2	2	1	0	0
0	0	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Fig. 2. Gradual ROI bit allocation/quantization.

The background QP, QPb, is fixed for all background MBs. The foreground QP is a function of the distance from regional edge: if a MB is adjacent to the edge, its QP is QPb-1; if the distance between a MB and the edge is larger (e.g. one MB), the QP of the MB is QPb-2; for other foreground MBs, the QP is QPb-3. An example is shown in Figure 2, where the gray colored MBs form the foreground sub-picture.

4. SIMULATIONS

4.1. Conditions

We selected multicast Internet streaming as a target application. As interactive error concealment cannot be used in large scale with IP multicast, transport coding level forward error correction (FEC) according to RFC 2733 [4] was used. We used the ITU-T Video Coding Experts Group (VCEG) common conditions for the low-delay Internet applications [11] as much as we considered appropriate.

Sub-picture coding and the gradual ROI bit allocation were implemented in and tested with H.26L Test Model Long-Term (TML) version 8.6 [12][13]. A constant foreground sub-picture was selected, and finer QPs decided by the gradual ROI bit allocation method were used in the foreground region. The independent sub-picture decoding mode was in use. The scheme was compared to TML-8.6 with and without ROI quantization (abbreviated as Conv+ROI and Conv, respectively). The ROI quantization for the conventional codec was the same as for the sub-picture coding scheme.

The simulator consisted of the following phases. First, the TML encoder was used to encode a sequence and generated an RTP packet stream. Then, a packet loss simulator erased some of the generated packets according to the error patterns released in [14]. The resulting packets stream was decoded using the TML decoder.

Encapsulation into RTP packets was done as follows: In the sub-picture coding scheme, INTRA pictures were encapsulated into five packets. There were two packets for the foreground sub-picture: one packet contained odd MB rows and another packet contained even MB rows. This slice interleaving mechanism, introduced in [15], was used to obtain a better error concealment result. One parity FEC packet was generated for the two foreground packets according to RFC 2733. The background sub-picture was packetized into another two packets using slice interleaving method. Two consecutive INTER pictures consisted a group, and for each such group there were two foreground sub-picture packets, one parity FEC packet for the foreground packets, and two background sub-picture packets. A sub-picture packet contained data from two pictures: MBs from even rows of a certain frame and MBs from odd rows of the next frame or vice versa. When sub-picture coding was not in use, there were three packets for each INTRA and INTER frame: two packets for the entire picture (slice interleaving applied), and one parity FEC packet for the two packets.

As multicast Internet streaming was assumed, the coding and transport scheme was tailored for the worst expected case. We selected optimal coding parameters and encapsulation options by trial and error according to the PSNR performance in the 20 % packet loss rate. We tested several INTRA macroblock coding rates as well as several slice sizes to find the optimal ones. The packet stream obtaining the best PSNR performance was multicast. The PSNR performance in 0, 3, 5, and 10 % packet loss rate was then simulated with the multicast packet stream. In order to obtain statistically significant results, each packet stream was virtually transported ten times for each packet loss rate. The first loss pattern position of a run starts from where the previous run ended.

4.2. Results

The experiments were done using Carphone, Hall, Coastguard, Foreman, News and Irene sequences, with different frame rates and bit-rates. We present only part of the results due to lack of space, more results can be obtained from the authors.

Figure 3. Example snapshots, 20 % packet loss rate. From left to right, the used codecs are Conv, Conv+ROI, and the proposed codec.

Figure 3 shows some example snapshots of Foreman@64 kbps and Carphone@64kbps in 20 % packet loss rate. From left to right in the figure, the snapshots were produced by the Conv, Conv+ROI, and the proposed codec, respectively. It can be seen that in both sequences the proposed sub-picture coding scheme with gradual bit allocation maintains the best subjective image quality.

Figure 4 presents the average luminance PSNR values of the foreground area of Irene@384kbps, News@64kbps, Foreman@64kbps, and Carphone@64kbps. These objective results confirm the subjective ones. Sub-picture coding improves the average luminance PSNR of the foreground area consistently regardless of the packet loss rate. This improvement is gained at the expense of background image quality, which is degraded by coarser quantization and less error protection. In fact, the overall PSNR in the sub-picture coding case drops a little compared to the ROI quantization case and somewhat more compared to the constant quantization case. However, errors in the background are far less noticeable than errors in the foreground, and therefore the overall subjective quality is improved.

5. CONCLUSION

Many video communication systems can make use of region-of-interest coding and unequal error protection to improve received visual quality. In order to apply them, pictures should be partitioned to different regions, or video bit-streams have to be organized in portions, of different importance in terms of visual quality. Techniques achieving these goals include data partitioning, scalable video coding, and region-based video coding.

This paper introduced a simple region-based video coding method called sub-picture coding. It provides means for both region-of-interest coding and unequal error protection based on partitioned regions of interest. In sub-picture coding, a picture is partitioned to one or more rectangular foreground sub-pictures and to a hollow background sub-picture along macroblock boundaries. Sub-picture edges are treated as slice boundaries or as picture boundaries depending on the signaled coding mode.

The sub-picture coding method was simulated in multicast Internet streaming conditions and it was found to be superior to conventional coding. Due to its simplicity and the obtained results, we think that sub-picture coding has some potential to become one of the key video coding tools for region-of-interest coding and unequal error protection.

6. REFERENCES

[1] D. Chai, K. N. Ngan, and A. Bouzerdoum, "Foreground / Background Bit Allocation for Region-Of-Interest Coding", IEEE 2000 International Conference on Image Processing, vol. 2, pp. 923-926, 2000.

[2] A. Miyazaki, H. Fukushima, K. Hata, T. Wiebke, R. Hakenberg, C. Burmeister, N. Takatori, S. Okumura, and T. Ohno, "RTP Retransmission Payload Format", IETF Internet Draft draft-ietf-avt-rtp-selret-02.txt, 18 July 2001.

[3] D. Leon, V. Varsa, "RTP retransmission framework", IETF Internet Draft draft-leon-rtp-retransmission-00.txt, July 2001.

[4] J. Rosenberg, H. Schulzrinne, "An RTP Payload Format for Generic Forward Error Correction" IETF RFC 2733, December 1999.

[5] 3GPP TS 23.107, "QoS Concept and Architecture (Release 4)", V4.2.0, October 2001, available via ftp://ftp.3gpp.org/Specs/2001-09/Rel-4/23_series/23107-420.zip.

[6] K. Nichols, S. Blake, F. Baker, and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", IETF RFC 2474, December 1998.

[7] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services", IETF RFC 2475, December 1998.

[8] M. M. Hannuksela, "Simple Packet Loss Recovery Method for Video Streaming", International Packet Video Workshop PV2001, Kyongju, South Korea, 30 April – 1 May 2001.

[9] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No. 3, March 2001.

[10] N. Brady, "MPEG-4 Standardized Methods for the Compression of Arbitrarily Shaped Video Objects", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 8, December 1999.

[11] S. Wenger, "Common conditions for wire-line, low delay IP/UDP/RTP packet loss resilient testing", ITU-T Video Coding Experts Group, document VCEG-N79r1, September 2001.

[12] T. Wiegand (editor), "H.26L Test Model Long-Term Number 9 (TML-9) draft0", ITU-T Video Coding Experts Group, document VCEG-N83d1, December 2001.

[13] H.26L Test Model Software, version 8.6, October 2001.

[14] S. Wenger, "Error patterns for Internet video experiments", ITU-T SG16 document Q15-I-16-R1, October 1999.

[15] S. Wenger and G. Côté, "Using RFC2429 and H.263+ at low to medium bit-rates for low-latency applications", International Packet Video Workshop PV99, April 1999.

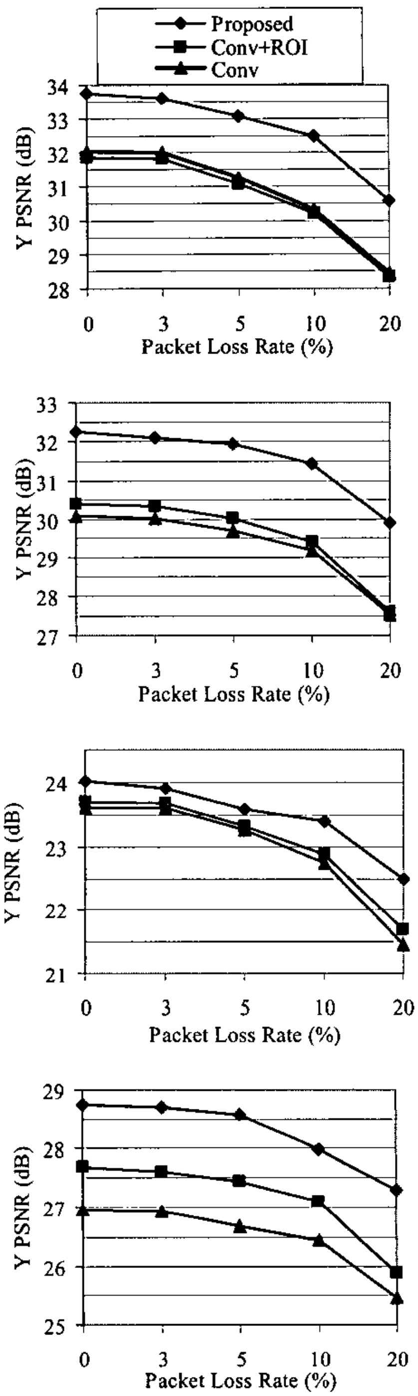

Fig. 4. Foreground average luminance PSNR values of Irene, News, Foreman, and Carphone (up to down)