ERROR-ROBUST INTER/INTRA MACROBLOCK MODE SELECTION USING ISOLATED REGIONS

Ye-Kui Wang [1], *Miska M. Hannuksela* [2] *and Moncef Gabbouj* [3]

[1] Tampere International Center for Signal Processing (TICSP), Tampere, Finland
[2] Nokia Mobile Software, Tampere, Finland
[3] Tampere University of Technology, Finland

ABSTRACT

The technique isolated regions, which is based on constraining inter and in-picture prediction jointly, was adopted into the emerging video coding standard H.264 (also know as MPEG-4 part 10 or AVC), as a tool to provide gradual random access and error resilience among other things. This paper addresses the error resiliency performance of isolated regions in the aspect of error-robust inter/intra macroblock mode selection. In addition to improved error resilience, random accessibility from non-intra pictures is provided as a side product by using isolated regions. It is shown by simulations that coding of isolated regions has competitive error resilience to loss-aware rate-distortion optimized macroblock mode selection algorithm (LA-RDO) with low complexity, and has consistently better performance than the latter when employed simultaneously with LA-RDO.

1. INTRODUCTION

Video encoders can help decoders to stop spatio-temporal propagation of transmission errors by coding macroblocks (MBs) in intra mode. Various types of MB mode selection algorithms exist. The algorithms can be categorized into non-adaptive and adaptive classes, and adaptive methods can be further classified to cost-function-based and rate-distortion optimized ones. The family of non-adaptive algorithms includes the circular intra refresh algorithm that scans the picture area in a pre-defined order and codes a certain number of intra MBs per picture in the pre-defined scan order. Another example of a non-adaptive algorithm is to intra code a certain number of MBs at randomly selected MB locations. Adaptive MB mode decision methods select the intra-coded MB locations so that the content of the pictures is taken into account. For example, a static background area needs not be refreshed in intra mode as often as moving objects. Cost-function-based methods, such as [1] and [2], calculate a cost for each MB with a certain function that may take into account the amount of prediction error data after motion compensation, for example. A certain number of MBs having the highest cost are coded in intra mode. Rate-distortion optimized MB mode selection algorithms use a Lagrangian cost function that linearly combines terms “rate” and “distortion”. The mode of each MB is selected so that the combined cost is minimized. An estimate of the expected distortion caused by transmission errors and losses may be taken into account in the cost function. A number of distortion estimating algorithms have been proposed and one of them, herein referred to as the loss-aware rate-distortion-optimized MB mode selection algorithm (LA-RDO), has been selected into H.264 reference implementation [3]. The LA-RDO method has significantly better performance than non-adaptive and cost-function-based algorithms. However, its computational complexity is typically multifold, which limits the use in practical implementations.

This paper proposes to use the isolated regions coding technique [4] for error robust MB mode selection. Section 2 presents the fundamentals of the isolated regions technique. Coding of isolated regions for MB mode decision is described in Section 3. Simulation results with comparisons are provided in Section 4 and finally Section 5 concludes the paper.

2. FUNDAMENTALS OF ISOLATED REGIONS

The isolated regions technique is based on constraining inter and in-picture prediction jointly. In-picture prediction includes intra prediction of samples and motion vector prediction within a certain picture. Inter prediction is means motion compensated prediction from previously coded pictures.

An isolated region is a solid area of MBs, defining the shape of the border across which in-picture prediction is limited. In addition, when coding an isolated region, temporal prediction outside its corresponding isolated region in a reference frame should be disallowed. The corresponding isolated region of an isolated region refers to that in a previous picture from which it is evolved. A picture can contain zero or more isolated regions that do not overlap. The area of the picture that is not covered by any isolated region forms the leftover region. There is no prediction limit for the leftover region.

An isolated region may start from any picture and evolve to following pictures. Generally, an isolated region ends after it evolves to be the entire picture. However, it may also ends even when it does not cover the entire picture area. The period from the picture where an isolated region starts to the picture where it ends is defined as the period of the isolated region.

There maybe unlimited amount of evolving patterns for isolated regions, some of them were described in [4]. One evolving pattern, which is called clockwise box-out, was applied in the simulations for this paper. The first MB to be allocated in the isolated region is in the center of the picture, and the growth order from the first MB is defined as shown in Fig.1. The growth rate in terms of MBs per picture may be either constant or variable.

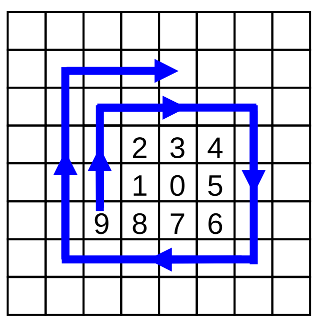

Fig.1. Growth order of box-out clockwise evolution.

The isolated regions technique provides an elegant solution for many applications, such as gradual random access, error resiliency and recovery, region-of-interest coding and prioritization, picture in picture functionality, and coding of masked video scene transitions [4-7]. With gradual random access based on isolated regions, media channel switching for receivers, bit-stream switching for the server, and allowing newcomers for multicast streaming will be as easy as conventional intra picture coding with smooth bit-rate.

3. ISOLATED REGIONS IN MACROBLOCK MODE SELECTION

The method to code isolated regions for robust MB mode selection is as follows. An isolated region starts from a picture and evolves non-decreasingly in size. If the background is static hence does not need to be intra updated, the isolated region may end before it evolves to cover the entire picture. Another isolated region repeats in the similar way after the previous one ends. The MBs in the isolated region of the first picture in the isolated region period are intra coded. The MBs in the isolated region in each of the later pictures may be either intra coded or inter coded with inter prediction from the corresponding isolated regions within the same isolated region period.

The above method can be used as either a non-adaptive or an adaptive MB mode selection algorithm. If the region growth rate is constant and all the MBs newly added into the isolated region in a picture are intra coded, the method is used as a non-adaptive algorithm. Otherwise, the algorithm falls into the category of adaptive MB mode selection algorithms.

The encoder can select a proper constant growth rate of the isolated region according to the picture size and the assumed transmission error rate. Generally, a good growth rate is equivalent to the expected loss rate of MBs. For example, for a CIF sequence that has 396 MBs per picture, if the packet loss rate is 20%, a growth rate of about 80 MBs per picture is appropriate. However, due to the possible large differences in sequence characteristics and different coding options, a content-adaptive growth rate may perform better and is under investigation.

4. SIMULATIONS

Four algorithms were compared: conventional circular intra refresh at a rate of one MB row per picture (CIR), LA-RDO, isolated regions based method (IREG), and a combination of IREG and LA-RDO. Real-time multicast/broadcast to users with different network conditions was assumed. Therefore, the coding options were selected to have best error resiliency performance. 20 % was assumed to be the largest packet loss rate. The coded sequences were decoded after packet loss simulation under packet loss rates 0, 3, 5, 10 and 20%.

In the simulations, isolated regions are coded with constant growth rate according to the clockwise box-out evolving pattern. The algorithms were implemented based on JM-1.4 [8]. Other simulation conditions are specified as follows.

Bitrate and PSNR Calculation: According to the error resilient testing conditions specified in [9], coding parameters such as quantization parameter were chosen to make the resulting bitrate as close as possible to the channel bitrate, taking into account the 40 bytes of IP/UDP/RTP headers per packet. PSNR values were calculated using each and every frame in the source sequence including skipped and lost frames. To reduce the effect imposed on the overall result by the first frame (the first encoded frame have a larger average size than the average size of the entire sequence), bitrate and average PSNR value were calculated from the sixth coded frame. This method allows coding short sequences with fair results. 300-400 frames of each designated sequence were used, to ensure that at least 100 frames are coded and at least 300 frames of a sequence are used.

Packet Loss Simulation: We assumed that the packet containing the parameter set is conveyed reliably (possibly out-of-band during the session setup), and therefore no error pattern was read from the error pattern file for it. At least one packet of the first frame should be received to avoid decoder crash. For that purpose, the first packet of the first frame was always received regardless of the corresponding error pattern.

Multiple Decoding Runs: The coded bitstream was decoded multiple times, and each time is called a decoding run. The beginning loss position of the run with order n+1 continuously follows the ending loss position of the nth run. The number of decoding runs was selected so that there are totally at least 8000 packets. The overall average PSNR was obtained by averaging the average PSNR values of all decoding runs.

Six coded sequences for each intra refresh algorithm were generated: Hall Monitor QCIF at 32 kbits/s, Foreman QCIF at 64 kbits/s, Foreman QCIF at 144 kbits/s, Paris CIF at 144 kbits/s, Paris CIF at 384 kbits/s, and Irene CIF at 384 kbits/s. The average Y-PSNR values of the coded sequences under different packet loss rates are shown in Fig.2-7.

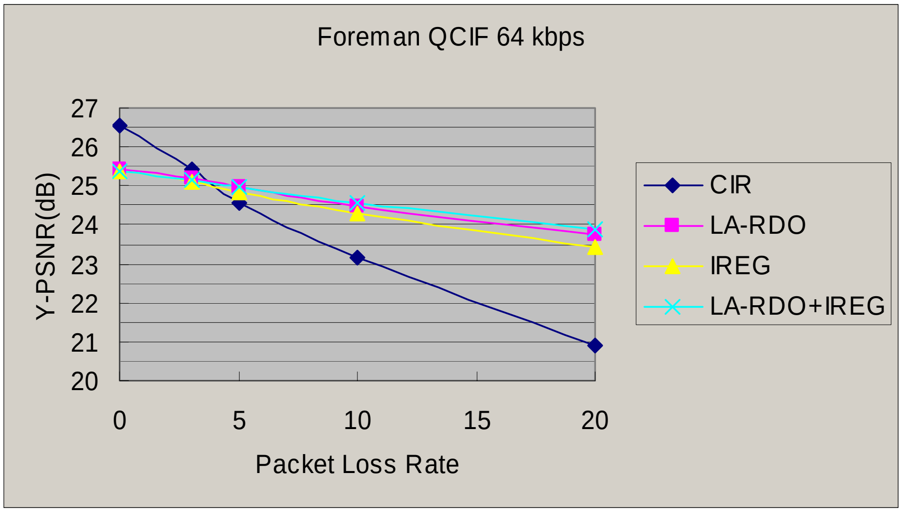

Fig.2. MB mode selection simulation results of Foreman QCIF at 64 kbits/s.

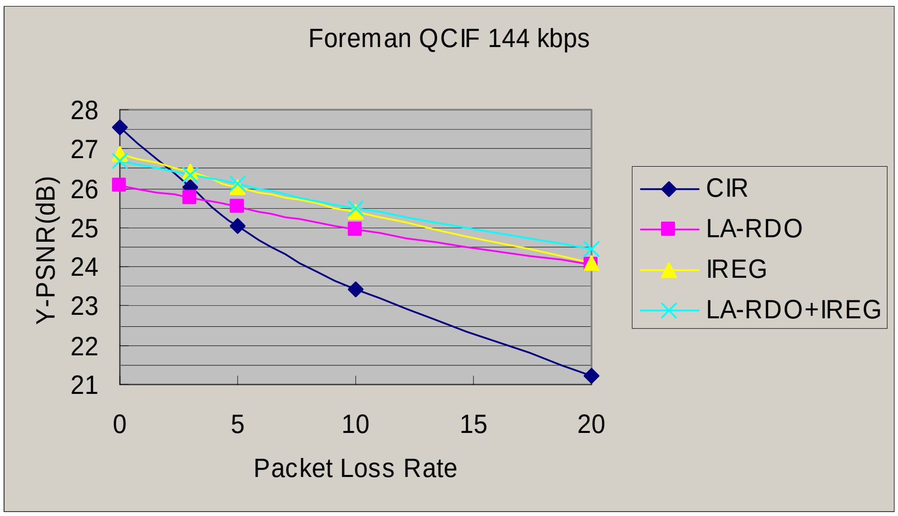

Fig.3. MB mode selection simulation results of Foreman QCIF at 144 kbits/s.

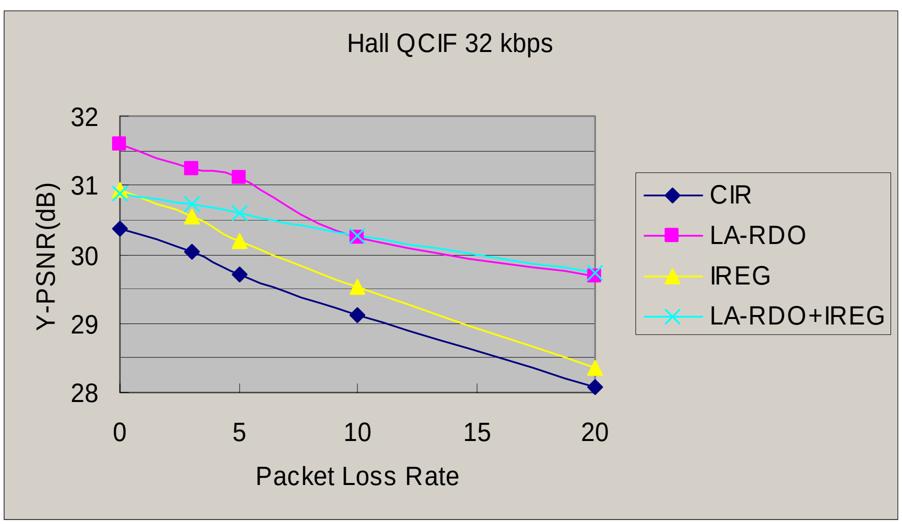

Fig.4. MB mode selection simulation results of Hall Monitor QCIF at 32 kbits/s.

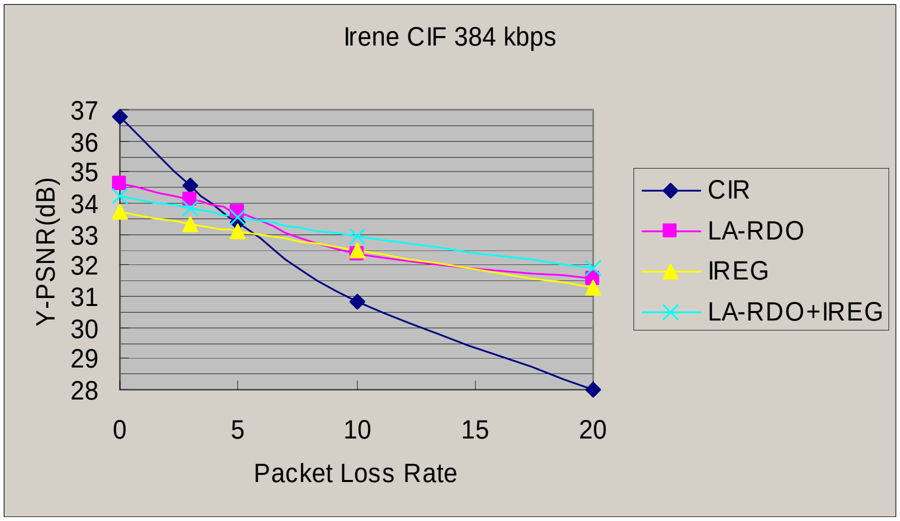

Fig.5. MB mode selection simulation results of Irene CIF at 384 kbits/s.

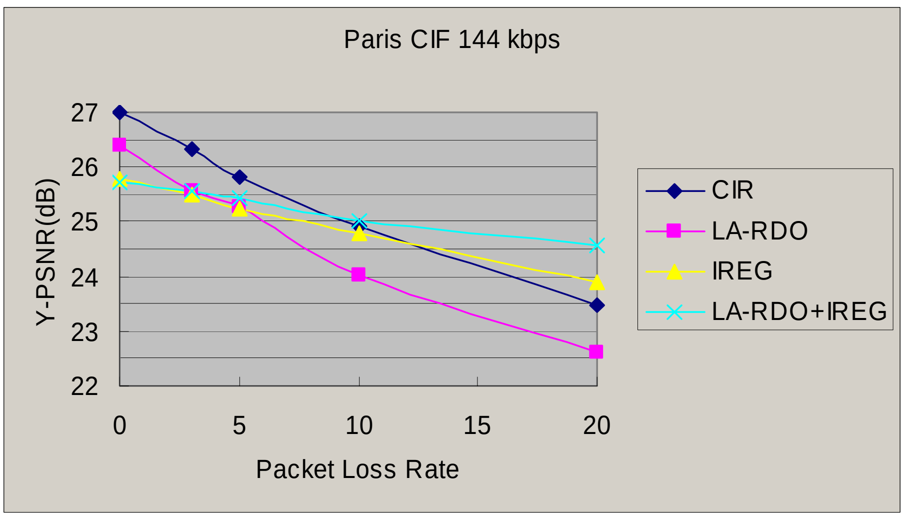

Fig.6. MB mode selection simulation results of Paris CIF at 144 kbits/s.

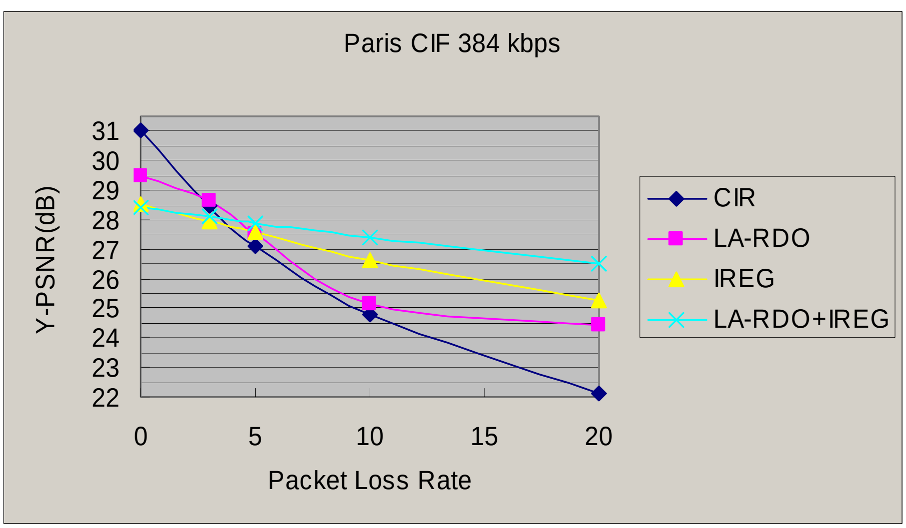

Fig.7. MB mode selection simulation results of Paris CIF at 384 kbits/s.

The simulation results show that IREG outperforms CIR significantly in all cases. Compared to LA-RDO, the performance of IREG is better in three cases while worse in the left three cases. Therefore, the performances of the two methods are competitive. However, the complexity of IREG is similar to that of CIR, which is a fractional of LA-RDO's complexity. In addition, use of IREG is not exclusive to the use of LA-RDO. When combined with LA-RDO, the error resiliency performance of IREG is can be further improved. As can be seen from Fig.2-7, the combined method outperforms LA-RDO consistently in all the six cases.

5. CONCLUSION

The error resiliency performance of the isolated regions technique in error-robust macroblock mode selection is investigated in this paper. Compared to conventional methods, the performance of the proposed method is significantly better, either in error resiliency or in computational complexity. This profit, together with many others, makes the isolated regions technique a highly valuable video coding tool.

REFERENCES

[1] ISO/IEC International Standard 14496-2:2001, Annex E, "Features supported, by the algorithm".

[2] J. Y. Liao and J. D. Villasenor, "Adaptive intra update for video coding over noisy channels," in Proc. IEEE International Conference on Image Processing, Oct. 1996.

[3] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC Draft Text, Joint Video Team document JVT-E146d37, Oct. 2002.

[4] Ye-Kui Wang and Miska M. Hannuksela, "Isolated regions: motivation, problems and solutions," Joint Video Team document JVT-C072, May 2002.

[5] Ye-Kui Wang and Miska M. Hannuksela, "Error-robust video coding using isolated regions," Joint Video Team document JVT-C073, May 2002.

[6] Ye-Kui Wang and Miska M. Hannuksela, "Gradual decoder refresh using isolated regions," Joint Video Team document JVT-C074, May 2002.

[7] Ye-Kui Wang and Miska M. Hannuksela, "Coding of masked scene transitions using isolated regions," Joint Video Team document JVT-C075, May 2002.

[8] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC Joint Model, version JM-1.4, Apr. 2002.

[9] Stephan Wenger, "Common conditions for wire-line, low delay IP/UDP/RTP packet loss resilient testing," ITU-T VCEG document VCEG-N79r1, Sep. 2001.