

# AN OPTIMIZED MAPPING ALGORITHM FOR CLASSIFIED VIDEO TRANSMISSION WITH THE H.264 FLEXIBLE MACROBLOCK ORDERING

*Sio Kei Im and A J Pearmain*

Electronic Engineering Department, Queen Mary University of London  
Mile End Road, London, E1 4NS, UK, Email: marcusim@ipm.edu.mo

## ABSTRACT

A new mapping algorithm is proposed for categorization of frames' macroblocks into two (or more) classes in flexible macroblock ordering to give error resilient video transmission. The successful transmission of macroblock data not only enhances the quality of the associated pixels, but also improves the quality of the adjacent lost macroblocks by improving the efficiency of error concealment. Therefore, in our scheme, by carefully modeling the decoder error concealment algorithm at the encoder side, we classify the macroblocks according to their eventual influence on picture quality. Within a limited bit rate budget, we employ an optimization algorithm to select the best group of high-priority macroblocks. We show that prioritized transmission of the more important macroblock group will improve the video quality in error situations where our mapping algorithm outperforms the default mappings of the H.264 codec.

## 1. INTRODUCTION

With the advent of wideband communication systems and the high performance of electronic devices, the demand for multimedia services over a variety of transmission channels has been increasing in recent years. A wide range of different applications that support video communications is now emerging into the market. However, there is still a huge challenge to optimize these applications in order to deliver better quality video in a lower bandwidth and hence at lower cost. One of the best solutions currently available is to employ the state-of-the-art video codec, H.264 (or MPEG-4 part 10). It delivers the same video quality at considerably lower bit rates than those required for earlier video codecs.

However, the higher compression achieved by H.264 is at the price of more sensitivity to errors where a single error may destroy all the video sequence. On the other hand, many of the current communication channels still do not provide reliable transmission. Hence, it is necessary for the video codec to include error resilience provisions to spatially and temporally localize the errors [1]. In this regard, a number of tools have been included in the video

codecs [2]. Among these methods, flexible macroblock ordering (FMO) is a new feature that has been introduced in the H.264 standard, and was not included in the previous ones.

In FMO mode, macroblocks (MBs) can be flexibly assigned to any slice in a frame, and so transmitted in a non-scanning order. Since each slice is independently decodable, one can scatter MBs spatially into different slices. Hence, the missing MBs of a lost slice can be surrounded by correctly received MBs, and an appropriate error concealment method can recover the slice efficiently [2][3]. There are a number of mapping structures for macroblock distribution including using an interleaved pattern and a dispersed pattern as shown in Fig. 1(a) and (b). If either of the slice groups of Fig. 1 is lost but the other group is present, the lost macroblocks can be recovered from the available neighboring macroblocks.

The slices are normally protected by using an error protection method prior to transmission. However, protecting the whole data will introduce a high overhead. Therefore, unequal error protection (UEP) schemes have been suggested to efficiently protect the more important portions of the data [4]. In UEP schemes a video bitstream is divided into layers with different importance and the more important layer is protected better than the other layer(s). It follows that the slice groups of FMO may also be appropriate for UEP scenarios. According to [5], UEP uses FMO, where a region of interest (ROI) in the pictures has been protected more strongly than the other regions. This may improve the quality subjectively but introduces new problems such as determining the ROI itself.

In [6] we showed that if the MBs are grouped based on their contribution to the final picture quality after error concealment, a better mapping scheme can be achieved. It is also shown that distinct protection of the more important slice group improves the video quality significantly when there are transmission errors. In this paper, we further improve the algorithm by careful modeling of the decoder error concealment at the encoder, and also by reducing the bit rate overhead of the method.

In Section 2 we describe the core of the idea and briefly explain the method described in [6]. Section 3 introduces the improved algorithm, and Section 4 contains the simulation results, followed by a conclusion in Section 5.

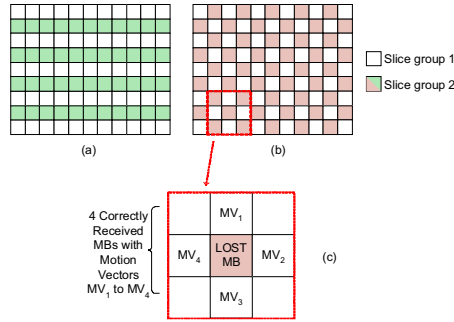


Fig. 1. FMO with two slice groups (SG), (a): interleaved, (b): dispersed, (c): dispersed when the second SG is lost.

## 2. MACROBLOCK CLASSIFICATION

We note that the improvement caused by FMO is due to the effectiveness of the error concealment. For example in the error concealment of [2], motion vectors (MVs) of the neighboring MBs (of a lost MB) are used to compensate the reference picture and the MV that results in more continuity with the adjacent MBs will be selected. Therefore, generally, the greater the number of neighbor MBs of the current MB that are received without uncorrected errors, the better the error concealment will be. In the dispersed structure of Fig. 1, for instance, a lost MB can be concealed by four adjacent MBs if the first slice group is received successfully (as shown in Fig. 1(c)). This can be achieved by protecting the first slice group better than the second one. However, there are situations in which the neighboring MBs have close similarities, such that their motion vectors may be equal to each other. That means the actual number of error concealment choices will be only one, because the four motion-compensated blocks are the same. In this situation, there will not be any difference in the error concealment result if the macroblock is surrounded by either one or more than one successfully received neighbors. That means better protection of all these four macroblocks may not provide any advantage.

A better mapping scheme can be achieved by applying the following steps: first, a frame is divided into two slice groups with a given bit-rate proportion, e.g. 50% for each group; then, assuming that only the first group can be transmitted reliably while the second group is lost, error concealment is performed for the second group. The problem will be to find the optimum mapping that maximizes the picture quality after error concealment. Unfortunately, finding the best mapping requires a very expensive full-search. For example, for a QCIF (176x144) video sequence, where there are 99 MBs in each frame, there are in the order of  $(99)^2$  different ways to divide the MBs into two groups. That means the encoding process should be carried out several times, and the bit rate of the two groups and the quality of the picture after error concealment should be measured. The best mapping that

results in the best quality and also satisfies the required group-rates-proportion could then be selected.

This ideal optimization is not practically feasible and therefore in [6] we simplified it to the following algorithm. First, one encoding pass is carried out assuming that all the MBs are in one slice group. This pass will determine the approximate bit rate ( $R$ ) and the (error free) coding distortion  $D_{\text{coded}}$  for each MB. The approximation at this stage is from the assumption that all MBs are in the same slice group. In fact in the arithmetic coding of each MB, the data in the adjacent blocks may contribute to the block coding [7], while at the end, some adjacent blocks may move to another slice group and hence are not allowed to be used in the arithmetic coding procedure. However, our experiments show that this will not significantly affect the accuracy of our rate and distortion estimation. Then an initial concealment distortion ( $D_{\text{concealed}}$ ) is calculated for each MB by moving all MBs to the second slice group. That means, for the initial concealment, there is no available neighboring MB to help the error concealment. Subsequently, an importance factors (IF) is calculated for each MB:

$$IF_i = \frac{(D_{i\text{concealed}} - D_{i\text{coded}})}{R_i}, \quad (i: \text{MB index}).$$

In the next step, the macroblock with the highest IF will be selected and moved to the first slice group. Now we have one MB in the first slice group and so it may improve the concealment of its 4 neighboring MBs. Therefore,  $D_{\text{concealed}}$  of these neighbor MBs and subsequently their IFs are updated. The process of moving macroblocks to the first slice group is again repeated until the total number of bits for the first slice group reaches the required bit rate. The following section introduces a more advanced method for this classification algorithm.

## 3. THE PROPOSED ENHANCED MACROBLOCK CLASSIFICATION ALGORITHM

The procedure of the proposed classification algorithm is shown in Fig. 2. First, as in the previous method, an encoding pass is carried out and  $R$  and  $D_{\text{coded}}$  values are extracted. In the next step the IF values are calculated for each MB. In the enhanced algorithm proposed in this paper this value is calculated as follows:

$$IF_i = \frac{(D_{i\text{concealed}} - D_{i\text{coded}}) + \sum_j (D_{j\text{concealed}} - D'_{j\text{concealed}})}{R_i}$$

The first term in the numerator represents the amount of reduction in distortion if macroblock  $i$  is transmitted and hence its missing distortion ( $D_{i\text{concealed}}$ ) is replaced by the coded distortion ( $D_{i\text{coded}}$ ), as shown in Fig. 3. The second term in the numerator represents the influence of successful transmission of this MB ( $i$ ) on improving the error concealment of its adjacent MBs. Here  $j$  can have up to four

different values depending on the situation.  $D_j'$ <sub>concealed</sub> and  $D_j$ <sub>concealed</sub> are the distortion of the adjacent MB  $j$  with and without considering MB  $i$  in the error concealment procedure, respectively.

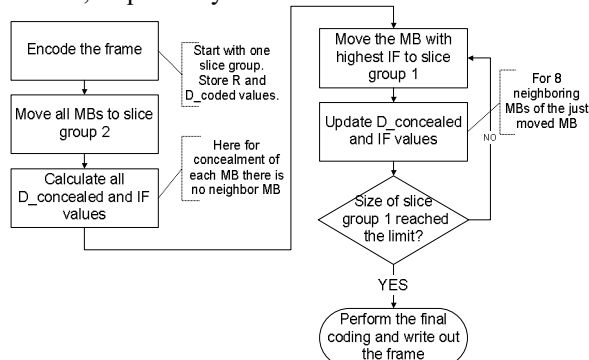


Fig. 2. The procedure for dividing the MBs into two slice groups.

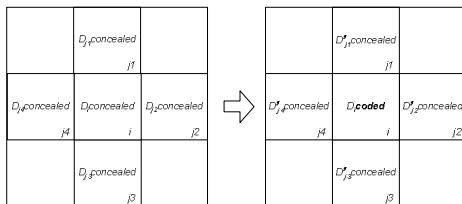


Fig. 3. On the left: all MBs (assumed lost and concealed) are in SG-2. On the right:  $i$ th MB is in SG-1 (assumed received)

The entire numerator represents the total improvement in picture quality caused by transmission of each MB. As pointed out earlier, there is a limited budget (e.g. 50% of the total) for the first slice group and hence there is an intention to accommodate those MBs in this group that have the maximum total influence. This turns out to be a knapsack optimization problem [8] which is here simply solved with a “greedy” algorithm by dividing the significance of each MB by its rate (generating IF). As in the previous algorithm the MB with the highest IF value is moved to the first slice group and  $D_{concealed}$  values for this MB and its four adjacent MBs is updated along with the IF values. Note that for the transmitted MB,  $D_{concealed}$  is now replaced by  $D_{coded}$  and in future steps no more improvement can be achieved for this particular MB. By repeating this procedure the MBs of the first and second slice groups are defined. For storing this mapping information in the H.264 standard, we use the ‘explicit’ mode of FMO in which the mapping can be manually defined by the encoder. Therefore, we may transmit an individual mapping structure for every frame by transmitting a new picture parameter set (PPS). This will add an overhead for the proposed method. In the following, we suggest a technique to reduce this overhead.

### 3.1. Temporal PPS scattering

The improvement caused by the proposed method is variable with the contents of the picture. That means that sometimes simple error concealment (i.e. a copy from the

previous frame) works satisfactorily and FMO will not add any further improvement. We detect these situations and avoid sending unnecessary PPSs in order to reduce the overhead of the proposed method. In this regard, in the proposed procedure after the first encoding pass, the sum of absolute difference between the current reconstructed picture and the previous reference picture is calculated. We normalize this sum by dividing it by the number of pixels. If this value is smaller than a threshold (in this paper the threshold is set to 10), it shows that the simple error concealment is sufficient for the current frame. Hence, we do not update the PPS. Otherwise, a PPS update is transmitted to update the FMO mapping. Our investigations show that at low bit rates (coarse quantization) fewer PPSs are transmitted than at high bit rates. This is a desired behavior because the PPS overhead at lower bit rates is more noticeable.

## 4. SIMULATION RESULTS

We compared the performance of the proposed method with the two other mapping schemes of Fig. 1 which are already included in the default mappings of the standard. Therefore, they have the advantage that there is no need to transmit an extra PPS for frames and hence have lower overhead. The result of the method in [6] is also shown for comparison. We assumed that the first slice group (SG-1) of all streams is transmitted in a reliable manner and so error free but the second group has a lower priority and so has some errors.

Fig. 4 shows the frame-by-frame luminance PSNR of the two first seconds of the Foreman sequence coded in different FMO modes. Here, the bit rate is 100 kb/s which is a rather high bit rate and hence the overhead of the proposed FMO and the method in [6] are not very significant. This is apparent from the error free graphs of these two methods (grouped into one graph for clarity) which show only about 0.2 dB less PSNR than that of the dispersed and interleaved methods. In this test, to assess the error resilience of the proposed method, we have deliberately removed the second slice group of a randomly selected frame 5. It can be seen that the new proposed method has significantly outperformed all the other methods. Fig. 5 shows the results for the same experiment but at a lower bit rate. It can be seen that at this bit rate the overhead of the proposed method has increased and resulted in around 0.5 dB drop in PSNR. Nevertheless, the proposed method has less overhead than the method in [6] because the number of transmitted PPSs is less, while it has recovered the picture quality more effectively than the other FMO modes.

Fig. 6 and Fig. 7 depict the average PSNR (of 100 simulation runs) of the methods under investigation at different bit error rate (BER) values. SG-1 in all cases is error free while SG-2 experiences the shown BER values. It can be seen that the proposed classification scheme has significantly improved the average PSNR for the high bit rate stream. For the low bit rate, the proposed method is still

superior to the others but the improvement is not very significant, partly due to fewer transmitted PPSs.

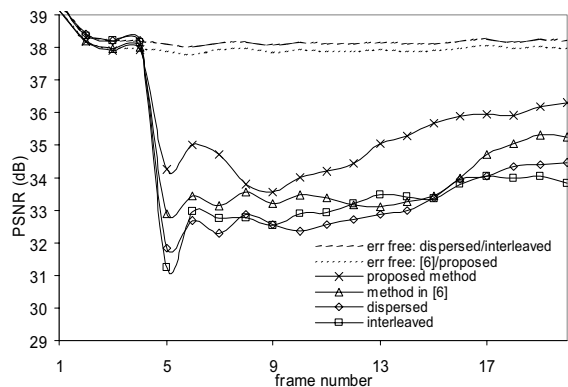


Fig. 4. SG-2 of frame 5 is lost, Foreman QCIF 10 Hz 100 kb/s.

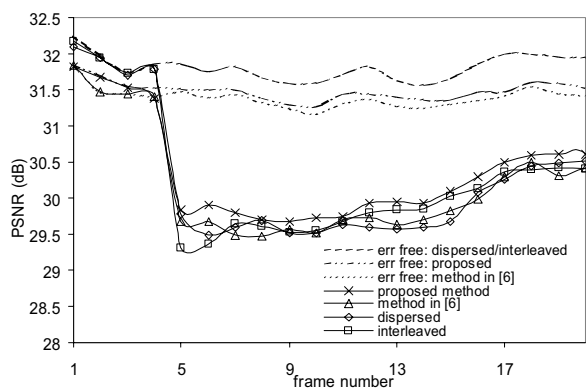


Fig. 5. SG-2 of frame 5 is lost, Foreman QCIF 10 Hz 28 kb/s.

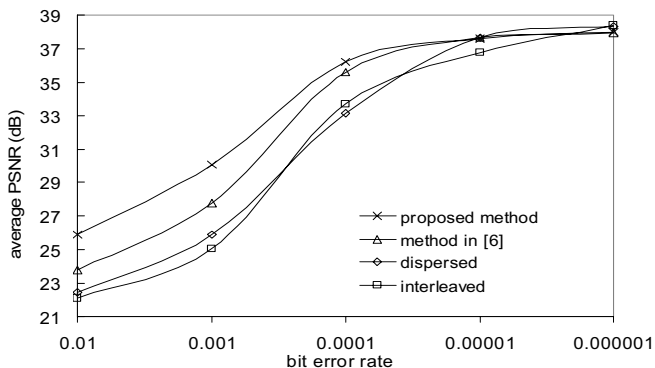


Fig. 6. Average PSNR vs. BER, Foreman QCIF 10 Hz 100 kb/s.

## 5. CONCLUSION

A method is proposed that classifies each macroblock based on the picture quality after error concealment. The classification metric considers the impact of the macroblock data on its own pixels and also the improvement that it makes to the error concealment of its adjacent macroblocks. The important macroblocks are then grouped in the first slice group where they are highly protected again error. Furthermore, to avoid sending unnecessary mapping headers at each frame, we measure the picture complexities and

update the macroblock mapping only for pictures that have more complexity and hence are more sensitive to error. The simulation results showed that this method outperforms a previously proposed method as well as the standard mappings of the H.264 codec.

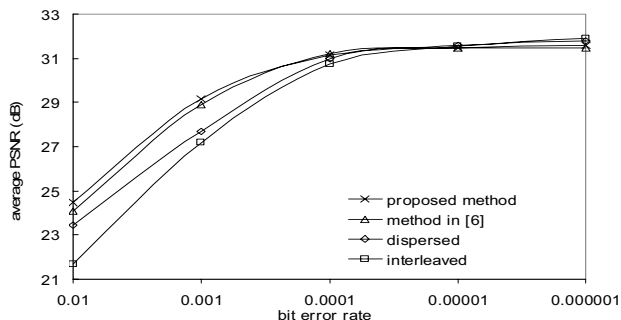


Fig. 7. Average PSNR vs. BER, Foreman QCIF 10 Hz 28 kb/s.

## 6. REFERENCES

- [1] A. Vetro, J. Xin, and Sun Huifang, "Error resilience video transcoding for wireless communications," *IEEE Wireless Communications*, Vol. 12, pp. 14-21, Aug. 2005.
- [2] T. Stockhammer, M. M. Hannuksela and T. Wiegand, "H.264/AVC in wireless environment," *IEEE Trans. Circuits Syst. Video Technol.*, v. 13, pp 657 – 673, July 2003.
- [3] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: a review" *Proc. IEEE*, Vol. 86, pp. 974-997, May 1998.
- [4] M. M. Ghandi, and M. Ghanbari, "Robust video transmission with an SNR scalable H.264 codec," *Proc. 7th IEEE Int. Conf. High Speed Networks Multimedia Commun. (HSNMC)*, pp. 932-940, June/July 2004.
- [5] M. M. Hannuksela, Y. K. Wang, M. Gabbouj, "Isolated regions in video coding," *IEEE Trans. Multimedia*, vol. 6, pp. 259-267, April 2004.
- [6] S. K. Im, and A. J. Pearmain, "Unequal error protection with the H.264 flexible macroblock ordering", *Proc. SPIE VCIP 2005*, vol. 5960, pp. 1033-1040, July 2005.
- [7] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, v. 13, pp. 620-636, July 2003.
- [8] S. Martello, and P. Toth, *Knapsack problems: algorithms and computer implementations*, John Wiley and Sons, 1990.