# A VERY LOW BIT-RATE VIDEO CODING ALGORITHM BY FOCUSING ON MOVING REGIONS

*Kwok-Wai Wong, Kin-Man Lam and Wan-Chi Siu*

Centre for Multimedia Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong

## ABSTRACT

Block-based motion estimation and compensation are the most popular techniques for video coding. As the shape and the structure of an object in a picture are arbitrary, the performance of such conventional block-based methods may not be satisfactory. In this paper, a very low bit-rate video coding algorithm by focussing on moving regions is proposed. The objective is to improve the coding performance, which gives better subjective and objective quality than that of the conventional coding methods at the same bit rate. We pre-defined eight patterns to represent the moving regions in a macroblock. The patterns are then used for motion estimation and compensation to reduce the prediction errors. Furthermore, in order to increase the compression performance, the residual errors of a macroblock are rearranged into a block with no significant increase of high order DCT coefficients. As a result, both the prediction efficiency and the compression efficiency are improved.

## 1. INTRODUCTION

Recently, the demand for the applications of digital video communication such as video conferencing, videophone, and the high-definition television has increased considerably. However, the transmission rates over public switched telephone networks (PSTN) are very limited. Therefore, very low bit-rate video coding is one of the important technologies for such applications. ITU-T recommendation H.263 [1] is one of the successful international standards for video compression using block-based techniques. However, the computational complexity of motion estimation for real-time applications may be too high. In order to reduce the complexity for motion estimation, several fast search algorithms are proposed [2-4]. In our approach, a fast search method with good performance is also devised.

The coding scheme of H.263 is based on the block-based methods [5], its prediction efficiency will not be as large as expected because the shape of a moving object is always arbitrary. Takahiro *et al.* [6] proposed a method to represent these shapes by means of four different patterns. The approach gives a better performance compared to the H.263, but the complexity of motion estimation and compensation might be too heavy for real-time applications. Also, the four patterns might not be sufficient to represent the moving objects for practical applications. In this paper, an encoding method of less complexity by using eight predefined patterns is proposed.

This work is supported by HKPolyU research grant G-V907

The basic encoding and decoding structures of our approach are based on H.263. It includes motion estimation and compensation, DCT transformation, and quantization. In our approach, moving regions in a frame are detected and then partitioned into macroblocks. One of the eight pre-defined patterns will be used to represent the moving regions for the purpose of motion estimation and compensation. The residual errors of the macroblock will be rearranged into a block without significant increase of high order DCT coefficients. However, if the patterns are not sufficiently enough to represent the moving regions in a macroblock, the conventional DCT-based coding method will be employed. As a result, both the picture quality and the runtime are improved.

## 2. LOW BIT-RATE VIDEO CODING USING PATTERNS

A frame in a sequence may consist of moving regions and static regions. It is unnecessary to encode the static regions in a frame which can be obtained from the reference frames directly. The moving regions should be encoded precisely, which are important for visual quality. In our approach, the moving regions in a frame are detected, and one of the pre-defined patterns is used to encode the moving region in a macroblock. The details of our algorithm are described as follows.

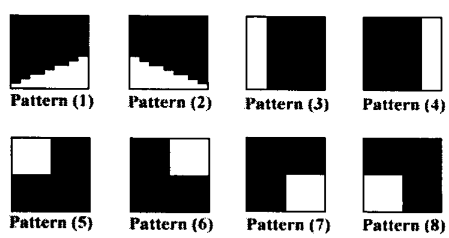


**Figure 1.** The eight patterns for moving region approximation

### 2.1 Moving Region Detection

The pre-defined eight patterns which represent the moving region in a macroblock are shown in Fig. 1. The white areas represent the moving region, while the black areas are the static region. The moving region in a frame is detected by comparing the current frame $C(x,y)$ to its previous frame $P(x,y)$. The pattern of a moving region $M(x,y)$ in a frame is obtained as follows:

$$M(x,y)=T|C(x,y)-P(x,y)| \qquad (1)$$

0-7803-6536-4/00/$10.00 (c) 2000 IEEE

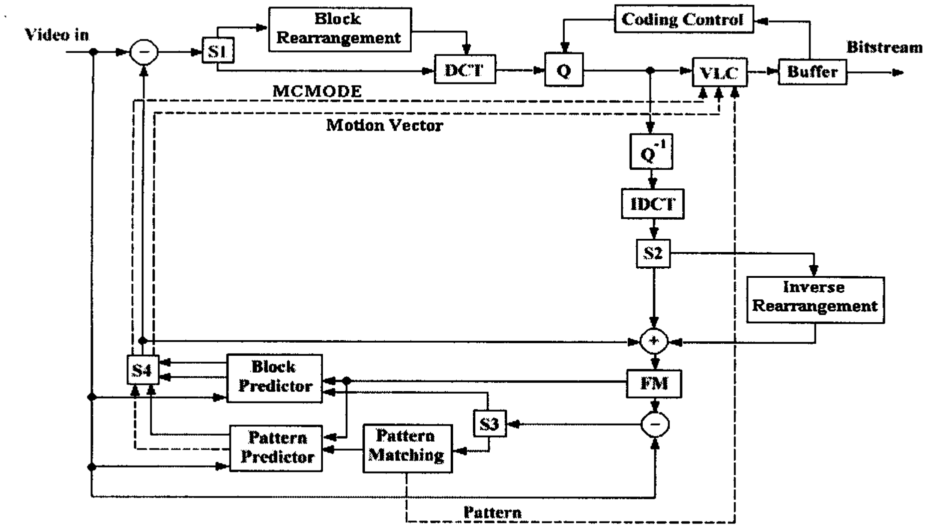


**Figure 2.** Block Diagram of Encoder

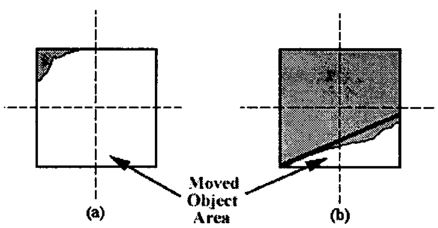


**Figure 3.** Block type selection: (a) an AMB, and (b) a RMB approximated by pattern (1)

where $T(.)$ is a thresholding function. The processed frame is then divided into macroblocks. If a macroblock contains static region only, the contents of the macroblock can be obtained directly from its previous frame. If the macroblock contains both static regions and moving regions, one of the eight pre-defined patterns will be used to represent the moving regions and the process of motion estimation and compensation will be employed by considering the moving region only. The best match pattern is obtained by finding a pattern that has the minimum function value. The function is defined as follows:

$$D_{K,N} = \frac{1}{256}\sum_{i=0}^{15}\sum_{j=0}^{15}\left|M_K(i,j) - P_N(i,j)\right| \qquad (2)$$

where $M_K(i,j)$ represents the $K^{th}$ macroblock in a frame, and $P_N(i,j)$ represents the pre-defined pattern number $N$. The intensity level of the patterns for the moving region is defined as one, while the static region is zero. However, if the patterns are not well enough to represent the moving region, the conventional motion estimation and DCT-based methods will be employed.

## 2.2 Architecture of the Encoder

The block diagram of the proposed video encoder is shown in Fig. 2. The structure is similar to a conventional video encoder, but includes additional features for encoding the moving regions:

1. Type of macroblock (MB): A MB will be classified according to the results of moving region detection. Three types of MB are defined in our approach: static MB (SMB), active MB (AMB), and active-region MB (RMB). If the contents of the MB are all zero, it is assumed to be a SMB. Otherwise, the MB will be divided into four sub-blocks for further classification. If the contents of all the sub-blocks containing non-zero value or the value of the function $D_{K,N}$ is greater than a threshold, the MB is defined as an AMB. If only a part of moving regions is detected in a MB, it is classified as a RMB. Examples of an AMB and a RMB are shown in Figure 3(a) and (b), respectively. If RMB is detected, the best match pre-defined pattern will be extracted by the "Pattern Matching " unit.

2. Block rearrangement: If the input is a RMB, the switch S1 will be connected to the "Block Rearrangement" unit to rearrange the residual error of the moving regions into a block with size of 8×8. Hence, fast DCT algorithms can be used to encode the block. Similarly, the output from the IDCT will be rearranged in an inverse manner by connecting the switch S2 to the "Inverse Rearrangement" unit.

3. Two motion predictors: The respective prediction procedures for AMB and RMB are different. The switch S3 is switched to the upper path for AMB, while it is connected to the lower path for RMB for which pattern prediction is applied. The MCMODE (motion predictor mode) is generated, and together with the motion vector, are encoded by the variable length code.

0-7803-6536-4/00/$10.00 (c) 2000 IEEE

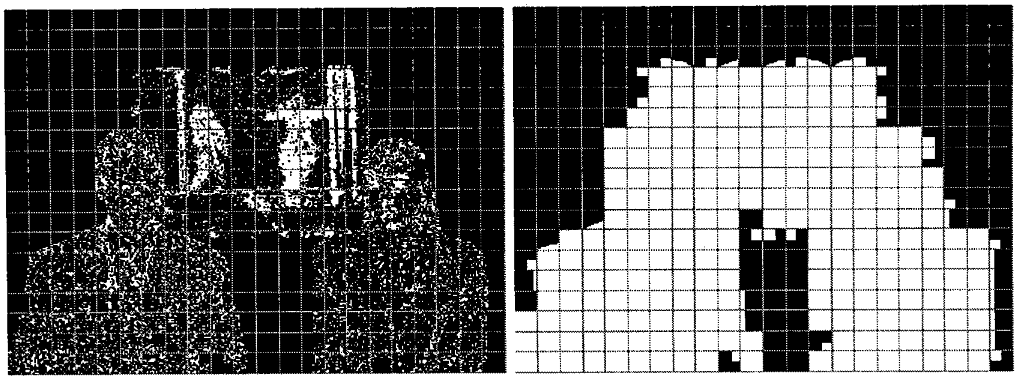

Figure 4. An example of pattern approximation from the sequence of "News"

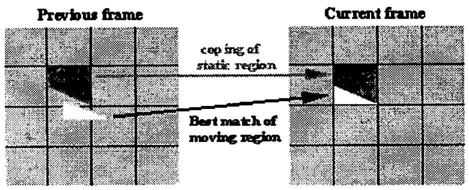


Figure 5. The principle of interframe coding for RMB

## 2.3 Motion Estimation and Compensation

The coding performance, such as the picture quality, the run-time, and the compression ratio, is affected by the motion estimation and compensation process, so it is an important process for video coding. In our approach, the motion vectors are zero if the SMB is detected. For the AMB, the conventional motion estimation and DCT-based methods are used. For RMB, both moving and static regions exist, but it is not necessary to use the static region in motion estimation. In [7], an efficient method was proposed for reducing the missing edge effect by using the edge information. In our approach, the motion vectors for the RMB are obtained by considering the moving regions only, which can reduce the complexity of the motion estimation process as well as prevent the missing edge effect. It is more precise if the static region is neglected for motion estimation.

The mean-absolute difference (MAD) criterion function is employed for motion estimation, the motion vector for the moving region in a RMB can be obtained by finding the minimum function value which is defined as:

$$MAD(dx,dy) = \frac{1}{64} \sum_{\substack{\forall P_N(i,j) \\ =1}} \left| C(i,j) - G(i+dx, j+dy) \right| \quad (3)$$

where G($i,j$) represents the MB from a reference picture, and ($dx,dy$) is a vector representing the search location. From eqn. (3), the total number of points for calculating the MAD is reduced. In this case, the motion vector of the moving regions can be obtained, while the contents of the static region can be directly copied from the previous frame. As a result, both the prediction errors and the complexity for motion estimation are reduced. The best match pattern determination in a video sequence and the principle of interframe coding for RMB are illustrated in Fig. 4 and 5, respectively.

## 2.4 Prediction Errors Encoding

The residual errors encoding for the AMB is based on the conventional DCT-based coding methods. For the RMB, the prediction error of the moving region in a MB are rearranged to a block of size 8x8 and then DCT transform is employed. Based on the distributions of the residual errors, the rearrangement method is devised without significant increase of the high order transform coefficients. As a result, only a block is needed to encode for a RMB and the compression ratio is increased. The rearrangements of the residual errors for the eight patterns are illustrated in Fig. 6. The arrows in the diagrams indicated the rearrangement order in a MB. The rearrangements of patterns 5-8 are not shown as they are already in the form of a block.

# 3. SIMULATION RESULTS

Results are obtained using four video sequences: an active video sequence, *Foreman*, a typical head-and-shoulder video sequence, *Akiyo*, and two sequences with smooth motion, *Mother-daughter* and *News*. The results of our proposed method are compared to the H.263 using the same I frame. Table 1 and Fig. 7 illustrate the first 100 frames coding results of our approach as well as the H.263 scheme. Table 2 shows the total number of RMB and the overall time saving of our approach. It can be observed that our method is not always suitable for any sequence. As the number of RMB is increased, both the compression ratio and the PSNR also will be increased. However, as fewer RMB are detected, the run-time is increased and the PSNR has a little decreased. This occurred due to less patterns being used for motion estimation, and the compression efficiency is reduced by storing the pattern information. Therefore, there is no advantage of our approach for encoding an active video sequence, such as *Foreman*, but this approach outperforms the H.263 in terms of the PSNR and runtime for sequences with smooth motion.

# 4. CONCLUSION

We have proposed a video coding algorithm of less complexity for very low-bit rate coding. The algorithm detects moving regions in a video frame, which are then divided into macroblocks for further analysis. We predefined eight patterns for representing the moving region, and performing motion estimation and compensation. In conclusion, both the computation for motion estimation and the encoding of the prediction errors are reduced. Furthermore, in order to reduce the size of a MB to be encoded, we devised a rearrangement method to compact the residual errors of a MB into a block of size 8x8. However, the total number of detected RMB is decreased in motion intensive video sequences, so the performance will be degraded and close to that of the H.263.

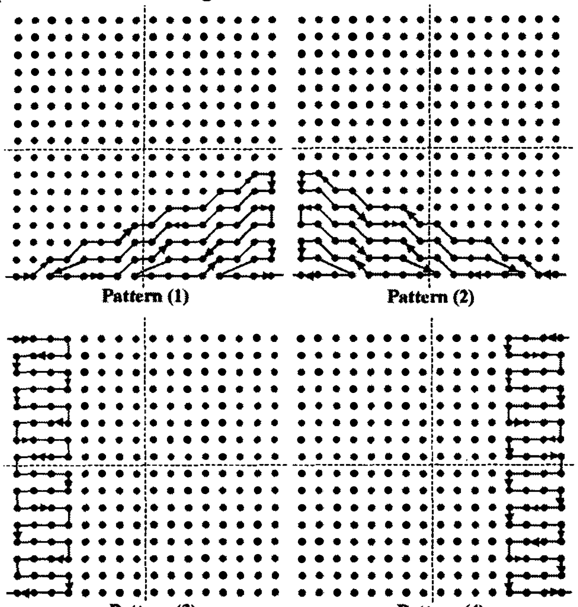


**Figure 6.** Rearrangement of the residual errors

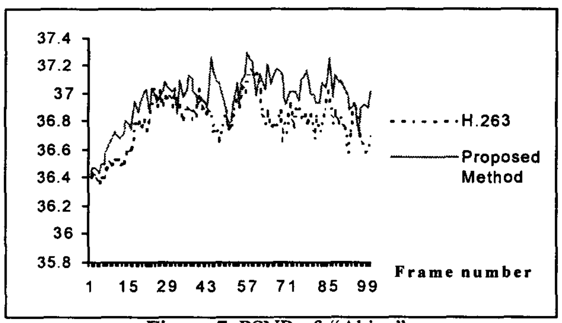


**Figure 7.** PSNR of "Akiyo"

| Sequence name | Akiyo (QCIF) QP=8 | | News (CIF) QP=10 | |
|---|---|---|---|---|
| Coder | H.263 | Proposed method | H.263 | Proposed method |
| Average PSNR(dB) | 36.79 | 36.95 | 35.39 | 35.52 |
| Average bits/frame (luminance) | 515 | 467 | 3870 | 3668 |
| MVD per frame | 161 | 172 | 1140 | 1204 |
| Total bits/frame | 1112 | 1111 | 6553 | 6552 |

**Table 1(a)** Simulation results

| Sequence name | Mother-daughter (QCIF) QP=10 | | Foreman (QCIF) QP=8 | |
|---|---|---|---|---|
| Coder | H.263 | Proposed method | H.263 | Proposed method |
| Average PSNR(dB) | 33.86 | 33.87 | 34.498 | 34.495 |
| Average bits/frame (luminance) | 1074 | 1047 | 3528 | 3521 |
| MVD per frame | 503 | 488 | 592 | 591 |
| Total bits/frame | 1986 | 1958 | 4592 | 4586 |

**Table 1(b)** Simulation results

| Sequence name | Akiyo QP=8 | News QP=10 | Mother-daughter QP=10 | Foreman QP=8 |
|---|---|---|---|---|
| Format | QCIF | CIF | QCIF | QCIF |
| Total number of RMB in the first 100 frames | 1234 | 4422 | 360 | 45 |
| Time saved per frame | 10.82% | 9.64% | 0.17% | -8.22% |

**Table 2.** Time saving

## REFERENCES


1. ITU-T Recommendation H.263, "Video coding for low bit rate communication," 1995.
2. L.M. Po and W.C. Ma, "A novel four-step search algorithm for fast block motion estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp.313-317, Jun 1996.
3. Y.L. Chan and W.C. Siu, "New adaptive pixel decimation for block motion vector estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 1, Feb. 1996.
4. C.H. Lin and J.L. Wu, "A lightweight genetic block-matching algorithm for video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 4, Feb. 1998.
5. A. Murat Tekalp, "Digital video processing," *Upper Saddle River, N.J.: Prentice Hall PTR*, 1995.
6. T. Fukuhara, K. Asai and T.Murakami, "Very low bit-rate video coding with block partitioning and adaptive selection of two time-differential frame memories," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 7, no.1, Feb.1997.
7. Yui-Lam Chan and Wan-Chi Siu, "Block Motion Vector Estimation using Edge Matching : An Approach with Better Frame Quality as Compared to Full Search Algorithm," *IEEE International Symposium on Circuits and Systems (ISCAS'97)*, pp. 1145-8, June 1997.