A NEW REAL-TIME PATTERN SELECTION ALGORITHM FOR VERY LOW BIT-RATE VIDEO CODING FOCUSING ON MOVING REGIONS

Manoranjan Paul, Manzur Murshed, and Laurence Dooley

Gippsland School of Computing and Info. Tech., Monash University, Churchill Vic 3842, Australia

E-mail: {Manoranjan.Paul,Manzur.Murshed,Laurence.Dooley}@infotech.monash.edu.au

ABSTRACT

Very low bit-rate video coding, using regular shaped patterns to focus on moving regions in macroblocks, has gained significant attention recently. This paper presents a new *real-time pattern selection* (RTPS) algorithm using a large codebook of thirty two patterns. The algorithm uses a *relevance* measurement for all the patterns and a moving region, to eliminate a large number of irrelevant patterns prior to the actual best likelihood pattern selection procedure. Both theoretically and empirically it is proven that not only is the computational complexity of the new algorithm comparable to the contemporary algorithm that use a pattern codebook size of only eight patterns but also the new algorithm reduces the bit-rate significantly, while maintaining comparable subjective quality.

1. INTRODUCTION

Reducing the transmission bit-rate while concomitantly retaining image quality is the most daunting challenge to overcome in the area of very low bit-rate video coding, e.g., H.26X standards [6]–[8]. Recently the MPEG-4 [5] video standard successfully introduced content-based coding, by dividing video frames into separate segments comprising a background and one or more moving objects. This idea is exploited in several low bit-rate macroblock-based video coding algorithms [1][14] through a simplified segmentation process that avoids the handling of arbitrary shaped objects, and therefore can use popular macroblock-based motion estimation. These algorithms innovatively focus on moving regions through the use of regular pattern templates, from a *pattern codebook* (see Figure 1), on non-overlapping rectangular blocks of 16×16 pixels each, known as *macroblocks* (MBs).

In [14], *macroblocks* were classified according to the following three mutually exclusive classes: 1) *Static MB* (SMB): Blocks that contain little or no motion; 2) *Active MB* (AMB): Blocks that contain moving object(s) with little static background; and 3) *Active-Region MB* (RMB): Blocks that contain both static background and some part(s) of moving object(s). In [1] a pattern codebook of four 128-pixel patterns was used. Further improvements were obtained in [14] using a pattern codebook of eight 64-pixel patterns (P_1–P_8 in Figure 1). In [10] Paul *et al.* presented a *variable pattern selection* (VPS) algorithm to select the λ best-matched patterns from a codebook of patterns P_1–P_{24} in Figure 1 using a "greedy" algorithm, where $\lambda \in \{4,8,16,24\}$. The idea was further extended in [11] with reduced computational complexity using the 32-pattern codebook in Figure 1. In [12] a new parametric approach of macroblock classification was proposed, which clearly outperforms the previous definition in [14] for certain parameter values.

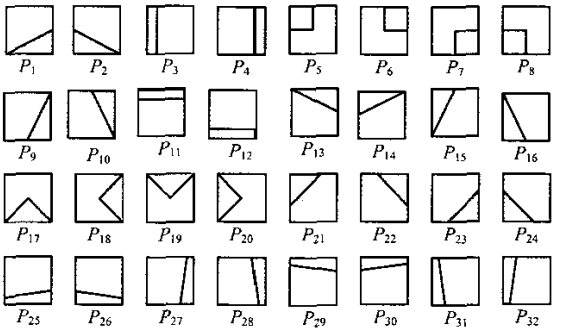

Figure 1: The pattern codebook of 32 regular shaped 64-pixel patterns, defined in 16×16 blocks, where the shaded region represents 1's and the white region represents 0's.

Variable pattern selection approach is not readily applicable to real time video coding, as the coding process must be preceded by the selection of the λ best-matched pattern set. It has also been reported in [14] that using eight instead of four patterns improved the *peak signal to noise ratio* (PSNR) and coding efficiency significantly. A similar, but diminishing trend was also observed in [10][11], when the pattern codebook size was further extended. In this paper, we present for the first time a new real-time, low bit-rate video coding algorithm focusing on moving regions using the 32-pattern codebook in Figure 1 and an extended parametric definition of MB classifications in [12].

The computational complexity of this new approach is kept within the real time threshold by eliminating a large number of irrelevant patterns. A pattern is considered irrelevant to a moving region if the distance between their respective gravitational centers exceed a prescribed threshold. For example, if a moving region is well represented by pattern P_{12} then patterns P_5, P_6, P_{11}, P_{19} etc. may well be considered irrelevant for some thresholds. The exact condition for a pattern to be considered *relevant* is discussed in the next section.

Both theoretically and empirically it is proven that the computational complexity of the new RTPS algorithm is comparable to the original algorithm presented in [14]. However, experimental results also reveal that RTPS reduces the bit-rate by as much as 5.5% without losing any subjective quality (i.e. the change in PSNR is bounded by 0.5 dB).

This paper is organized as follows. The *relevance* of a particular pattern to a moving region is defined in Section 2. Section 3 presents the RTPS algorithm and the coding technique is further elaborated in Section 4. In Section 5, the computational complexity of the algorithm is analyzed and compared with that of the algorithm in [14]. Some experimental results are presented in Section 6, while Section 7 concludes the paper.

2. PATTERN RELEVANCE MEASURE

Let $C_k(x,y)$ and $R_k(x,y)$, $0 \le x,y \le 15$, denote the k^{th} macroblock of the current and reference frames respectively, where the frame dimension is W pixels $\times$ H lines. The moving region $M_k(x,y)$ in the k^{th} macroblock of the current frame is obtained as follows:

$$M_k(x,y) = T(|\, C_k(x,y) \bullet B - R_k(x,y) \bullet B \,|) \qquad (1)$$

where B, of size 3×3, is the structuring element of a morphological closing operation • [2][9], $|v|$ = the absolute value of v, $T(v) = 1$ if $v > 2$ or 0 otherwise, $0 \le x,y \le 15$, and $0 \le k < W/16 \times H/16$.

Let G(A) denote the gravitational center of the 16×16 matrix A of bits (0 or 1), such that

$$\mathrm{G}(A) = \left[\frac{1}{256} \sum_{x=0}^{15} \sum_{y=0}^{15} x \cdot A(x,y), \frac{1}{256} \sum_{x=0}^{15} \sum_{y=0}^{15} y \cdot A(x,y) \right] \qquad (2)$$

Lemma 1: *Without losing any generality, it can be assumed that the gravitational center of a moving region will never be on the boundary of the macroblock.*

Proof: A moving region can have its gravitational center on the boundary of the macroblock if and only if the region itself is part of either a horizontal or a vertical boundary line. Such a moving region should never be classified as an RMB. ■

Table I: Values of $\eta_{\max}(\eta_{\min})$ and $\Delta(\eta_{\min})$ for possible $\eta_{\min}$ values

$\eta_{\min}$	$\eta_{\max}(\eta_{\min})$	$\Delta(\eta_{\min})$	$\eta_{\min}$	$\eta_{\max}(\eta_{\min})$	$\Delta(\eta_{\min})$
2	11	5.81	15	32	12.72
4	11	6.00	17	32	13.00
6	11	6.19	18	32	13.28
8	17	7.00	20	32	16.38
9	26	8.19	22	32	17.44
10	27	8.44	24	32	18.81
11	31	9.38	26	32	19.00
12	32	10.38	28	32	20.19
13	32	10.81	31	32	21.00
14	32	11.44	32	32	21.44

Let the *relevance* of the k^{th} macroblock with pattern P_n be calculated as:-

$$\nabla_{k,n} = \mathrm{dist}(\mathrm{G}(M_k), \mathrm{G}(P_n)), \qquad (3)$$

using the Manhattan distance:-

$$\mathrm{dist}(a,b) = |x(a) - x(b)| + |y(a) - y(b)| \qquad (4)$$

where, $x(\cdot)$ and $y(\cdot)$ denote the x- and y-coordinates respectively. Manhattan distance is preferred to Euclidian distance because of its reduced computational time.

Let $\Delta(\eta_{\min})$ be the minimum value, which guarantees that for at least $\eta_{\min}$ patterns, the *relevance* measure $\nabla_{k,.} \le \Delta(\eta_{\min})$ for any k. $\Delta(\eta_{\min})$ can be calculated as follows:

$$\Delta(\eta_{\min}) = \max_{1 \le x,y \le 14} \eta_{\min}^{\text{th}} \min_{1 \le n \le 32} (\mathrm{dist}((x,y), \mathrm{G}(P_n))) \qquad (5)$$

In the above calculation, the gravitational center of all moving regions is assumed to be never on the border (see Lemma 1).

$\Delta(\eta_{\min})$ also leads to an upper limit in the number of possible *relevant* patterns, calculated as follows:

$$\eta_{\max}(\eta_{\min}) = \max_{1 \le x,y \le 14} \sum_{1 \le n \le 32} \begin{cases} 1, & \text{if } \mathrm{dist}((x,y), G(P_n)) \\ & \le \Delta(\eta_{\min}); \\ 0, & \text{Otherwise.} \end{cases} \qquad (6)$$

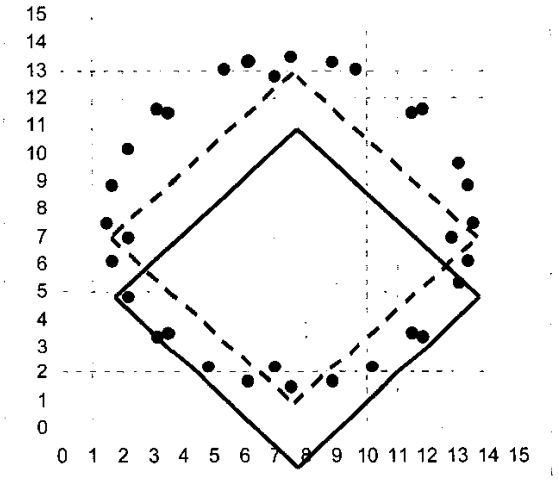

Figure 2: An example supporting the calculated values of $\eta_{\max}(4) = 11$ and $\Delta(4) = 6$.

Pattern P_n is considered to be *relevant* to the moving region in the k^{th} macroblock if and only if $\nabla_{k,n} \le \Delta(\eta_{\min})$, for all k and n. The average number of *relevant* patterns can be approximated by $(\eta_{\min} + \eta_{\max}(\eta_{\min}))/2$. Values of $\eta_{\max}(\eta_{\min})$ and $\Delta(\eta_{\min})$ for all possible $\eta_{\min}$ values are given in Table I. It is interesting to note that there exist some consecutive values of $\eta_{\min}$ for which the same $\Delta(\eta_{\min})$ value is obtained, e.g., $\Delta(3) = \Delta(4) = 6.00$. In such cases, only the maximum $\eta_{\min}$ value is tabled. Figure 2 clearly proves the validity of the aforementioned calculations, where each square represents an area bound by the

Manhattan distance 6 from its center. If the gravitational center of a moving region is exactly the same as the center of the dotted square, there exist only four *relevant* patterns; while as many as eleven patterns can be *relevant* when the gravitational center of a moving region is exactly the same as the center of the solid square.

3. THE RTPS ALGORITHM

Let the *likelihood* of the k^{th} macroblock with pattern P_n be calculated as

$$D_{k,n} = \frac{1}{256}\sum_{x=0}^{15}\sum_{y=0}^{15} |M_k(x,y) - P_n(x,y)|. \quad (7)$$

The k^{th} macroblock is then classified as follows, for all k:

1) If $\left| M_k \right| < 8$ then the k^{th} macroblock is classified as an SMB.
2) Else if
$$\left| M_k \right| < \delta, \quad (8)$$
where $\delta \in \{64,96,128\}$, AND
$$\min_{\forall n \left| \nabla_{k,n} \le \Delta(\eta_{\min}) \right.} (D_{k,n}) < 0.25 \quad (9)$$
then the k^{th} macroblock is classified as an RMB whose moving region is represented by the first pattern P_i in the codebook where $D_{k,i} = \min_{\forall n \left| \nabla_{k,n} \le \Delta(\eta_{\min}) \right.} (D_{k,n})$.
3) Else the block is classified as an AMB.

Besides using this extended definition of SMB, RMB, and AMB, the *real-time pattern selection* (RTPS) algorithm also calculates the $D_{k,n}$ value partially, quadrant-by-quadrant. Let τ be the *speed-up factor* of this technique compared to calculating the $D_{k,n}$ value as a whole. It has been empirically found that τ increases as $\eta_{\min}$ increases. This observation is presented in Figure 3 for the Miss America video sequence.

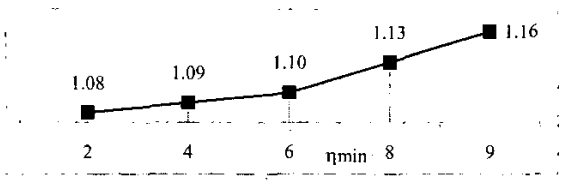

Figure 3: Values of τ for different $\eta_{\min}$ values on the Miss America video sequence.

4. CODING TECHNIQUE

SMBs and the static regions of RMBs are skipped from coding and transmission as they can be obtained from the reference frame. For each AMB, as well as the moving region of each RMB, motion vector and residual errors are calculated using conventional block-based methods, with the obvious difference in having the shape of the blocks for the moving regions of RMBs as that of the best-match pattern, rather than being square. To avoid multiple 8×8 blocks of DCT calculations for only 64 residual error values per moving region of RMB, these 64 values are *rearranged* into an 8×8 block. An *inverse rearrangement* is performed when decoding.

Pattern identification numbers are coded using variable length Huffman codes as given in Table II. These codes are obtained using the average pattern frequencies over a large number of standard and non-standard video sequences.

Table II: Pattern ID number ◊ variable length code

1◊ 1010	9◊ 010011	17◊ 001100	25◊ 111011
2◊ 00101	10◊ 11100	18◊ 010010	26◊ 0011101
3◊ 00001	11◊ 01000	19◊ 001101	27◊ 0101011
4◊ 00000	12◊ 01011	20◊ 001111	28◊ 0010000
5◊ 011	13◊ 001001	21◊ 0010001	29◊ 111010
6◊ 100	14◊ 10110	22◊ 101111	30◊ 11100
7◊ 110	15◊ 101110	23◊ 010100	31◊ 0011100
8◊ 0001	16◊ 11111	24◊ 111101	32◊ 0101010

5. COMPUTATIONAL COMPLEXITY

Let β be the total number of *candidate* RMBs, meeting condition (8).

For each candidate RMB:

i) The moving region consists of $(\delta + 8)/2$ number of 1's on average. So, the average number of operations required to calculate the gravitational center of a moving region, based on (2), is $256 + (\delta + 8) + 2 = \delta + 266$.

ii) The *relevance* measure, in (3), takes $3 \times 32 = 96$ operations in total for all 32 patterns.

iii) The likelihood measure in (7), is calculated on average for $(\eta_{\min} + \eta_{\max}(\eta_{\min}))/2$ patterns, each taking 512 operations.

So, the total number of operations required by the RTPS algorithm for pattern searching is:-

$$\begin{aligned} &\text{OP(RTPS)} \\ &= \beta(\delta + 266 + 96 + (\eta_{\min} + \eta_{\max}(\eta_{\min}))/2 \times 512/\tau) \quad (10) \\ &= \beta(\delta + 362 + (\eta_{\min} + \eta_{\max}(\eta_{\min})) \times 256/\tau). \end{aligned}$$

In contrast, for each candidate RMB, the algorithm in [14], computes only eight likelihood measurements (7) and so for the same video sequence, the number of operations required is:-

$$\text{OP}([14]) = \beta(8 \times 512) = 4{,}096\beta. \quad (11)$$

For $\eta_{\min} = 4$, the average number of *relevant* patterns per candidate RMB becomes $(4+11)/2 = 7.5$, which is close to the pattern codebook size of algorithm [14]. To keep the PSNR comparable or even better, the RTPS algorithm must consider $\eta_{\min} \ge 4$.

Assume that the RTPS algorithm is using $\eta_{\min} = 4$ and $\delta = 128$. If OP(RTPS) ≤ OP([14]), τ must be ≥ 1.065. Figure 4 shows that the average τ, for $\eta_{\min} = 4$, is 1.10, which makes $\text{OP(RTPS)} = 3{,}981\beta \cong \text{OP}([14])$. It can, therefore, be claimed that the computational complexity of the RTPS algorithm is comparable to algorithm [14] while keeping the PSNR comparable.

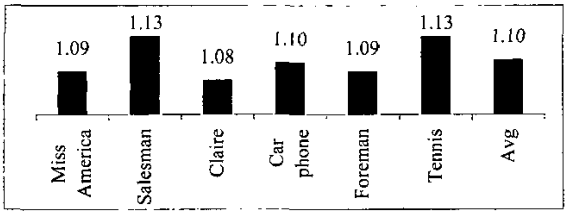

Figure 4: Speed-up factors of calculating the $D_{k,n}$ value partially quadrant-by-quadrant for $\eta_{min} = 4$ on six standard sequences.

6. EXPERIMENTAL RESULTS

Both the algorithms along with the H.263 standard have been tested on a large number of standard and non-standard video sequences of CIF and QCIF digital video formats [13] with different degrees of object and camera motions. However, for the sake of brevity, experimental results are presented using the first 100 frames of six standard video sequences. Table III shows that the RTPS algorithm outperforms both the algorithm in [14] and the H.263 standard in terms of lower bit-rate and higher PSNR for $\eta_{min} = 4$ and $\delta = 64$. However, the RTPS algorithm with $\eta_{min} = 4$ and $\delta = 128$ reduces the bit-rate by as much as 5.5% without losing any subjective quality.

7. CONCLUSIONS

Recently several studies on pattern representation of moving regions in blocked-based video motion estimation and compensation have been reported. In this paper, a new *real-time pattern selection* (RTPS) algorithm has been developed using a 32-pattern codebook. The RTPS algorithm uses a *relevance* measurement, in the form of the Manhattan distance between two gravitational centers, among all the patterns and a moving region to eliminate a large number of irrelevant patterns prior to the actual best likelihood pattern selection procedure. The algorithm uses a novel technique in guaranteeing lower and upper limit of *relevant* patterns. It has been established that not only the computational complexity of the RTPS algorithm is comparable to the previous algorithm in [14] but also the RTPS algorithm reduces the bit-rate by as much as 5.5%, while maintaining comparable subjective quality.

8. REFERENCES

[1] Fukuhara, T., K. Asai, and T. Murakami, "Very low bit-rate video coding with block partitioning and adaptive selection of two time-differential frame memories," *IEEE Trans. Circuits Syst. Video Technol.*, 7, 212–220, 1997.

[2] Gonzalez, R.C. and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

[3] ISO/IEC 11172, International Standard, 1992.

[4] ISO/IEC 13818, MPEG-2 International Standard, 1995.

[5] ISO/IEC N4030, MPEG-4 International Standard, 2001.

[6] ITU-T Recommendation H.261, "Video codec for audio-visual services at p×64kbits/s," 1993.

[7] Draft ITU-T Recommendation H.263, "Video coding for low bit-rate communication," 1996.

[8] Draft ITU-T Recommendation H.263, "Video coding for low bitrate communication," Version 2, 1998.

[9] Maragos, P., "Tutorial on advances in morphological image processing and analysis," *Opt. Eng.*, 26(7), 623–632, 1987.

[10] Paul, M., M. Murshed, and L. Dooley, "A Low Bit-Rate Video-Coding Algorithm Based Upon Variable Pattern Selection," Proc. of 6th Int. Conf. on Signal Processing, Beijing, China, August 26–30, Vol-2, 933–936, 2002.

[11] Paul, M., M. Murshed, and L. Dooley, "A Variable Pattern Selection Algorithm with Improved Pattern Selection Technique for Low Bit-Rate Video-Coding Focusing on Moving Objects," to appear in the Proc. of Int. Workshop. on Knowledge Management Technique, Crema, Italy, September 16–18, 2002.

[12] Paul, M., M. Murshed, and L. Dooley, "Impact of Macroblock classification on low bit rate video coding focusing on moving region," To appear in the Proc. of Int. Conference of Computer and Information Technology, Dhaka, Bangladesh, December 26–27, 2002

[13] Shi, Y.Q. and H. Sun, *Image and Video Compression for Multimedia Engineering Fundamentals, Algorithms, and Standards*, CRC Press, 1999.

[14] Wong, K.-W., K.-M. Lam, and W.-C. Siu, "An Efficient Low Bit-Rate Video-Coding Algorithm Focusing on Moving regions," *IEEE trans. circuits and systems for video technology*, 11(10), 1128–1134, 2001.

Table III: Relative bit-rate and PSNR for six standard sequences using H.263 standard, the algorithm in [14], and the RTPS algorithm

Video sequences	Video format	H.263		[14]		RTPS ($\eta_{min} = 4$)					
						$\delta = 64$		$\delta = 96$		$\delta = 128$	
		Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)	Bit-rate (kbps)	PSNR (dB)
Miss America	QCIF (176 × 144)	53.8	44.8	53.2	45.0	52.0	45.1	50.4	44.9	50.3	44.9
Car phone	QCIF (176 × 144)	230.3	39.8	228.2	39.5	228.3	39.7	222.8	39.3	221.0	39.2
Foreman	QCIF (176 × 144)	290.1	38.3	288.2	37.7	290.6	37.9	283.2	37.4	281.2	37.3
Salesman	CIF (352 × 288)	725.9	40.0	716.1	40.1	702.2	40.0	686.0	39.8	683.0	39.7
Tennis	CIF (352 × 240)	1,630.4	36.6	1,614.5	36.3	1,613.3	36.5	1,604.7	36.2	1,600.1	36.1
Claire	CIF (352 × 288)	139.9	44.7	131.9	44.9	130.5	44.8	127.4	44.9	126.7	44.7