

Improved Block Motion Estimation Using Block Frequency Warping

Orhan Akbulut, Oguzhan Urhan, *Member, IEEE*, and Sarp Ertürk, *Member, IEEE*

Abstract—In this letter, an enhanced motion estimation scheme based on block frequency warping is presented. In the proposed approach, the best candidate block is initially obtained using the classical full search scheme, and then frequency warping is applied to this block to improve the matching accuracy. The proposed method outperforms the conventional matching strategy, as shown in the experimental results.

Index Terms—Frequency warping, motion estimation.

I. INTRODUCTION

MOTION estimation (ME) is one of the most important tasks in the video coding process since it exploits temporal correlation between neighboring frames to enable efficient compression of video. Motion estimation can provide significant compression efficiency by removing temporal redundancy. The compression performance is typically increased with improved prediction capabilities. Improving temporal prediction accuracy is also an important aspect in H.264/AVC [1] for improved performance.

Research in the area of motion estimation is mostly concentrated on two aspects: reducing the high computational load of motion estimation possibly at the cost of estimation accuracy; or increasing motion estimation accuracy possibly at the cost of higher computational load. Fast search techniques [2]–[5] and low bit-depth representation-based matching methods [6]–[10] are examples for approaches proposed to reduce the computational load of motion estimation.

Several approaches try to improve the compression performance of the overall video coding system by exploiting properties of the transform coder. The method presented in [11] uses a smooth constrained matching (SCM) criterion instead of the conventional minimum absolute difference (MAD) matching to improve the overall video coding performance. Another recently proposed method in [12] introduces the high-frequency component matching (HFCM) concept to exploit properties of transform coding for efficient video compression.

The number of approaches that aim to improve the estimation accuracy is much more limited compared to the many ap-

proaches proposed to reduce the computational load. For example, sub-pixel motion estimation approaches are typical examples for techniques that improve the prediction accuracy at the cost of higher computational load [1]. In a similar point of view, motion estimation with block frequency warping is proposed in this letter. In the proposed approach, initially classical full search with MAD matching is performed. Then, frequency warping is applied to the best candidate block to further improve the prediction accuracy.

II. BLOCK FREQUENCY WARPING-BASED ME

Frequency warping of signals has been proposed in the literature to make them more appropriate for processing and particularly compression. For example, the warped DCT (WDCT) is proposed in [13] and [14] to improve coding efficiency of DCT compression by changing properties of the input signal. It is shown that this approach outperforms the standard DCT-based scheme. Frequency warping has also applications in speech processing [15]. Furthermore, the application of frequency warping in the wavelet domain has been presented in [16].

Frequency warping can be carried out using FIR-type all-pass filters. The filter used in the frequency warping process can be formulated as

$$A(z) = \left(\frac{-\alpha + z^{-1}}{1 - \alpha z^{-1}} \right) \quad (1)$$

where α shows the frequency warping parameter. The filter $A(z)$ is also referred to as Laguerre filter and is used in various applications [17]–[19].

The warping parameter typically controls the amount of frequency warping. In practice, it is possible to try out different warping parameter values and use the one that gives the best result. In [13], it is suggested to vary α within the range of $[-0.1, 0.875]$ so that the warping parameter can be formulated as

$$\alpha = \left(\frac{n}{(10N)} \right), \quad -N \leq n < N - 1 \quad (2)$$

for a total of $2N$ different warping parameters. Hence, a total of 16 different warping parameters are obtained for $N = 8$. Note that the case for which the warping parameter is zero corresponds to no frequency warping, so that the classical case is also included.

In the proposed approach, it is proposed to use frequency warping to improve the prediction for block motion estimation, by modifying the frequency distribution of the candidate

Manuscript received June 20, 2007; revised October 18, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mauro Barni.

The authors are with the Kocaeli University Laboratory of Image and Signal Processing (KULIS), Electronics and Telecommunications Engineering Department, University of Kocaeli, Kocaeli 41040, Turkey (e-mail: orhanakbulut@gmail.com; urhano@kou.edu.tr; urhano@ieee.org; sertur@kou.edu.tr; sarp@ieee.org).

Digital Object Identifier 10.1109/LSP.2007.914795

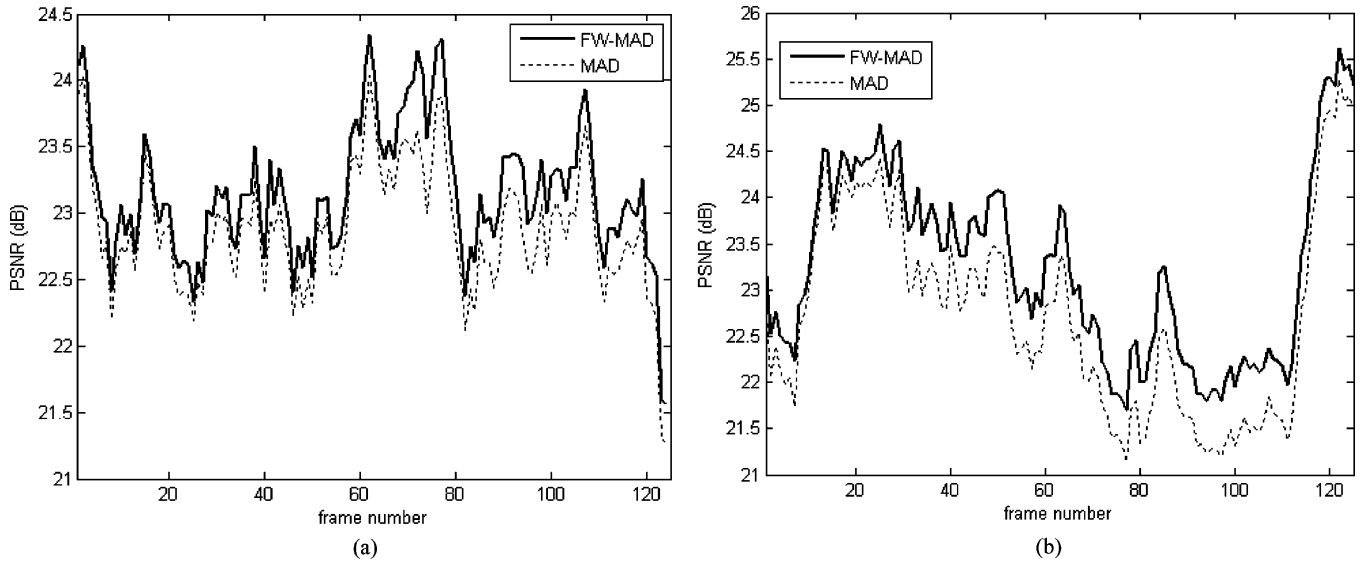


Fig. 1. PSNR results for classical MAD and proposed FW-MAD ME approaches. (a) Football sequence. (b) Mobile sequence.

TABLE I
AVERAGE PSNR (dB) OF SEVERAL SEQUENCES RECONSTRUCTED BY MAD AND FW-MAD MOTION ESTIMATION METHODS

Method	Video Sequences					
	Football (352x240) (125 frames)	Flowergarden (352x240) (115 frames)	Mobile (352x240) (140 frames)	Tennis (352x240) (112 frames)	Coastguard (352x288) (299 frames)	Foreman (352x288) (299 frames)
MAD (bs=16)	22.88	23.79	22.99	29.87	30.48	32.11
FW-MAD (bs=16)	23.16	24.26	23.47	30.11	30.82	32.43
MAD (bs=8)	24.73	25.22	23.88	31.25	31.59	32.90
FW-MAD (bs=8)	25.11	25.80	24.44	31.55	32.06	33.31

block. For this purpose, initially the motion vector and the corresponding best matching candidate block is obtained by classical MAD-based motion estimation. Then, frequency warping is applied to the best matching block with different warping parameters to further improve the prediction. The MAD is computed between all frequency warped candidate blocks and the original block, and the frequency warped candidate block giving the lowest error is used as final prediction.

III. EXPERIMENTAL RESULTS

The Football, Flowergarden, Mobile, Tennis, Foreman, and Coastguard sequences are used to assess the performance of the proposed approach. Both 16×16 and 8×8 pixel block sizes are used for evaluation. Fig. 1 shows the peak signal-to-noise ratio (PSNR) results with respect to the frame number for the 125 frames of the Football and Mobile sequences for frequency warping MAD (FW-MAD) as well as classical MAD using a block size of 16×16 pixels and a search range of 16 pixels. In this case, each frame is reconstructed using best matching blocks from the previous frame and the PSNR is computed between the original and predicted frames. It is seen from this figure that the proposed approach improves the temporal prediction performance for all frames.

Table I shows the average PSNR results for the Football, Flowergarden, Mobile, Tennis, Foreman, and Coastguard sequences obtained using full search for block sizes of 16×16 and 8×8 pixels, with search ranges of 16 and 8 pixels, respec-

tively. As seen in Table I, the proposed approach can improve the estimation performance by up to 0.6 dB on average compared to conventional MAD. Note that bs is used to refer to the block size in this table.

Fig. 2 shows sample frames for the Football sequence where a block size of 16×16 pixels and a search range of 16 pixels are used. As seen from the estimation errors in Fig. 2(c) and (d), the proposed frequency warping-based motion estimation approach provides a lower residual.

The improvement in estimation accuracy comes at the cost of additional computational load. After the candidate block is obtained using conventional motion estimation, frequency warping is applied to improve prediction. Sixteen different warps ($N = 8$) are conducted on the candidate block, and the best matching case evaluated with MAD is used as final prediction. Hence, the additional computational load sums up to 16 warps and 16 additional MAD computations per block. Since warping is only applied to the candidate block and not all search blocks, the introduced complexity is relatively limited. The warps can be regarded as linear filtering operations, while the additional MAD computations can be regarded as an increase in search locations as if 16 additional searches are carried out per block.

IV. CONCLUSION

In this letter, a novel motion estimation approach which uses frequency warping of candidate blocks is presented. Frequency

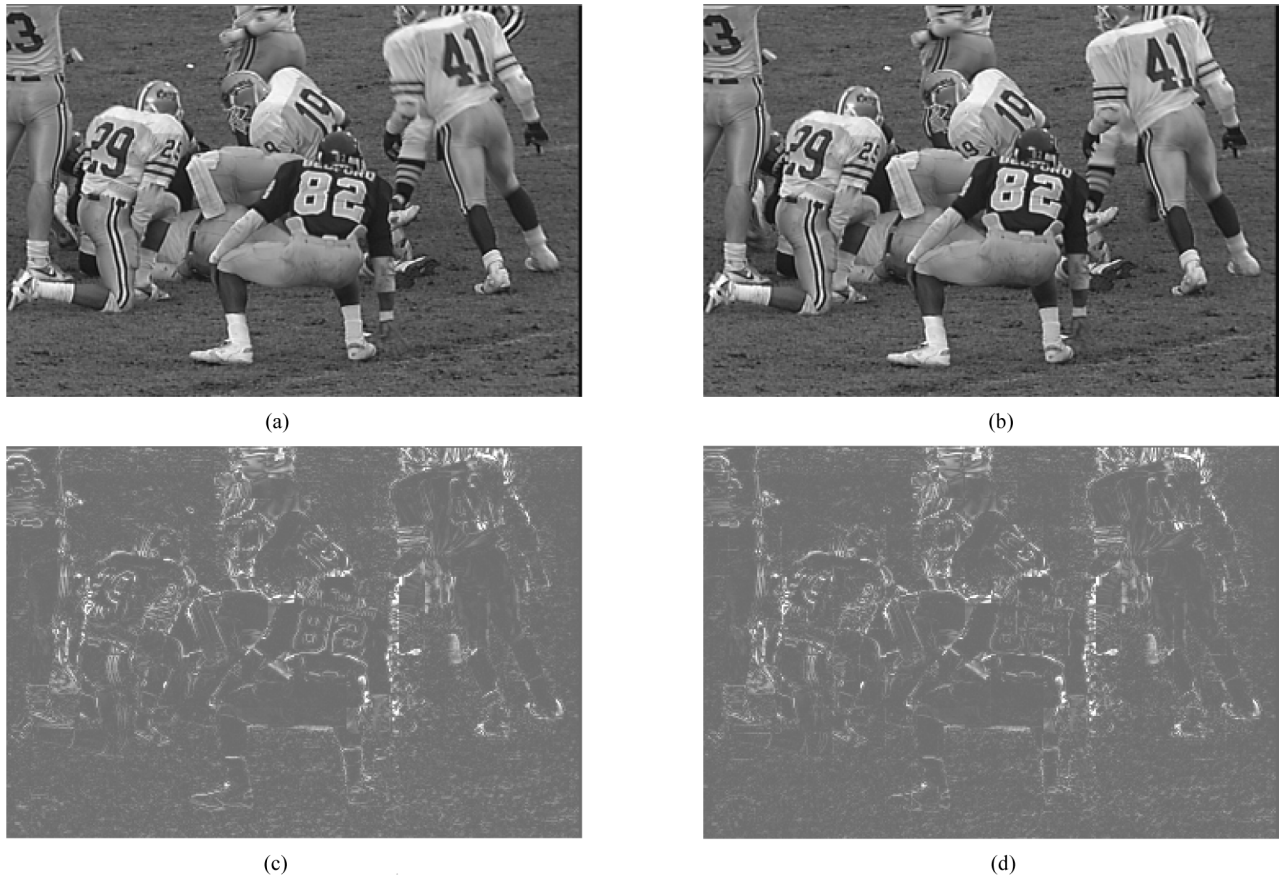


Fig. 2. Sample frames and reconstruction residuals for MAD and proposed FW-MAD approaches. (a) Football frame #95. (b) Football frame #96. (c) FS-MAD estimation error (offset = 128) PSNR = 22.58 dB. (d) FS-FW-MAD estimation error (offset = 128) PSNR = 22.93 dB.

warping is simply applied to the best matching block found by conventional MAD motion estimation to improve prediction accuracy. Experimental results show performance improvement with respect to standard MAD-based motion estimation.

REFERENCES

- [1] J. V. Team, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, 2003, ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC.
- [2] S. Zhu and K. K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 287–290, Feb. 2000.
- [3] C. Zhu, X. Lin, and L.-P. Chau, "Hexagon-based search pattern for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 5, pp. 349–355, May 2002.
- [4] H. So, J. Kim, W.-K. Cho, and Y.-S. Kim, "Fast motion estimation using modified diamond search patterns," *Electron. Lett.*, vol. 41, no. 2, pp. 62–63, Jan. 2005.
- [5] C. Zhu, X. Lin, L. Chau, and L.-M. Po, "Enhanced hexagonal search for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 10, pp. 1210–1214, Oct. 2004.
- [6] B. Natarajan, V. Bhaskaran, and K. Konstantinides, "Low-complexity block-based motion estimation via one-bit transforms," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 4, pp. 702–706, Aug. 1997.
- [7] A. Ertürk and S. Ertürk, "Two-bit transform for binary block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 938–946, Jul. 2005.
- [8] S. Ertürk, "Multiplication-free one-bit transform for low-complexity block-based motion estimation," *IEEE Signal Process. Lett.*, vol. 14, no. 2, pp. 109–112, Feb. 2007.
- [9] O. Urhan and S. Ertürk, "Constrained one-bit transform for low-complexity block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 4, pp. 478–482, Apr. 2007.
- [10] O. Urhan, "Constrained one-bit transform based motion estimation using predictive hexagonal pattern," *J. Electron. Imaging*, vol. 61, no. 3, Jul.–Sep. 2007, Article no. 033019.
- [11] X. Jing, C. Zhu, and L. P. Chau, "Smooth constrained motion estimation for video coding," *Signal Process.*, vol. 83, no. 3, pp. 677–680, Mar. 2003.
- [12] S. Ertürk, "A new perspective to block motion estimation for video compression: High-frequency component matching," *IEEE Signal Process. Lett.*, vol. 14, no. 2, pp. 113–116, Feb. 2007.
- [13] N. I. Cho and S. K. Mitra, "Warped discrete cosine transform and its application in image compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 8, pp. 1364–1373, Dec. 2000.
- [14] K. Kim, N. I. Cho, and S. K. Mitra, "Rate distortion optimization of the image compression algorithm based on the warped discrete cosine transform," *Signal Process.*, vol. 83, no. 9, pp. 1919–1928, Sep. 2003.
- [15] S. Wabnik, G. Schuller, U. Kramer, and J. Hirschfeld, "Frequency warping in low delay audio coding," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 2005 (ICASSP '05)*, Mar. 2005, vol. 3, pp. 181–184.
- [16] G. Evangelista and S. Cavaliere, "Frequency-warped filter banks and wavelet transforms: A discrete-time approach via Laguerre expansion," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2638–2650, Oct. 1998.
- [17] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Autom. Control*, vol. 36, no. 5, pp. 551–562, May 1991.
- [18] M. A. Masnadi-Shirazi, "Optimum Laguerre networks for a class of discrete time systems," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 2104–2108, Sep. 1991.
- [19] T. O. E. Silva, "Optimality conditions for truncated Laguerre networks," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2528–2530, Sep. 1994.