COMBINING MULTIPLE-CANDIDATE SEARCH WITH ONE-BIT TRANSFORM FOR MOTION ESTIMATION

Peter H. W. Wong *, Oscar C. Au **, A. Tourapis***

Department of Electrical and Electronic Engineering

The Hong Kong University of Science and Technology

Clear Water Bay, Kowloon, Hong Kong.

Email: eepeter@ee.ust.hk *, eeau@ee.ust.hk **, alexis@ee.ust.hk ***

Tel.: +852 2358-7053 **

ABSTRACT

Motion estimation using one-bit information (1BT) to represent a pixel or a group of pixels becomes popular in recently years [1-2] as the distortion function between the reference block and the search block can be computed very efficiently. In [1], each frame is filtered and compared with the original frame to extract the edge information. However, its performance is not good enough as it degrades the predicted image by 1dB on the average comparing with full search. In this paper, we improve the performance using the results of one-bit transform as the initial estimate of the motion vector, and use re-examination strategies to find the final motion vector. Experiment results show that the proposed method improves the performance of the one-bit transform by up to 0.9 dB and outperforms many other motion estimation algorithms such as the Three-Step Search (3SS), New Three-Step Search (N3SS) and the Four-Step Search (4SS).

1. REVIEW OF MOTION ESTIMATION USING ONE-BIT TRANSFORM

For convenience, we use similar symbols as in [1]. Each 8-bit frame F is filter with a 17×17 kernel K which is a bandpass filter and given by (1). The filtered frame F' is then compared with the original frame F to create a 1-bit frame G . The creation of G is formulated in (2).

$$K_{i,j} = \begin{cases} \frac{1}{25}, & \text{if } i, j \in [0, 4, 8, 12, 16] \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$G_{i,j} = \begin{cases} 1, & \text{if } F_{i,j} \geq F'_{i,j} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Sum of exclusive OR (SXOR) is used as the distance measure instead of Sum Absolute Difference (SAD). Suppose each block consists of $N \times N$ pixels. Let $u_{i,j}$ and $v_{i,j}$ denote the values of the pixels at row i and column j of the blocks u and v respectively. The distance measure between two blocks is formulated in (3).

$$\|u, v\|_1 = \sum_{i,j} u_{i,j} \otimes v_{i,j} \quad (3)$$

where $\otimes$ denotes the exclusive-or operation. For each reference block in a frame to be predicted, it searches within a searching window to find the best match with the least distortion from the reference block.

2. MULTIPLE-CANDIDATE RE-EXAMINATION

We observe that the performance of 1BT can be significantly improved if we re-examine the best N vectors from the 1BT using other approaches. This can be explained that the 1BT limits the variation of small searching region. We propose the multiple-candidate one-bit transform (M1BT) and the M1BT with full search (M1BTFS) approach to improve the performance of 1BT by a re-examination process.

The M1BT is simply to re-examine the best N candidate motion vectors from the 1BT. The M1BTFS consists of two stages. In the first stage, the distortion for the best motion vector from 1BT is first compared with a threshold. If the distortion is greater than the threshold, we examine the second best vector. If the distortion is still above the threshold, we examine the next best vector and so on. This stage will stop when either the distortion is below the threshold or the best N motion vectors have been examined. If the minimum distortion obtained after the best N candidates are examined is still above a certain threshold, the process proceeds to the second stage in which Full Search is employed. The threshold value is the average distortion of the last frame.

2.1. Distortion Measure

In this paper, we use two block distortion measures: Sum of Absolute Difference (SAD) and Sub-sampled SAD (SSAD). In SSAD, we sub-sample all blocks using the checker board pattern to reduce the computational complexity by a factor of 2. For two blocks u and v , let u_{ij} and v_{ij} denote the intensity of the pixel in the i -th row and j -th

column of u and v respectively. The distortion D between the two blocks is computed as.

$$D = \sum_{(i \bmod 2)=(j \bmod 2)} |u_{ij} - v_{ij}| \quad (4)$$

where $(k \bmod 2)$ is the remainder when k is divided by 2.

3. SIMULATION RESULTS

We applied the M1BT and M1BTFS on two testing sequences 'Tennis' and 'Football'. Each frame consists of 360×240 pixels. The first 88 frames of 'Tennis' and 99 frames of 'Football' are used. The search window are of size ± 7 pixels. The average Mean Square Error (MSE) against N is shown in Fig. 1, where N is the number of re-examined motion vectors from 1BT. It can be seen that re-examination can indeed reduce the MSE significantly. Most of the reduction is achieved by N in the range of 4 to 6. Although SSAD can reduce complexity by half when compared with SAD, the achievable MSE reduction is lowered.

Table 1 and 2 list the average MSE, the relative complexity and the decrease in PSNR of various algorithms when compared with Full Search (FS), for the case of $N=6$. M1BT1 is M1BT using SAD and M1BT2 is M1BT using SSAD. Note that, if the relative complexity of FS is 225, that of 1BT is about 1[1]. This computation is included in the values for M1BT1, M1BT2 and M1BTFS. The tables show that various computation and performance tradeoff can be achieved with the proposed algorithms. In particular, M1BTFS can achieve a PSNR within 0.18dB of Full Search while achieving a computation reduction factor of 18. When larger distortion is allowed, M1BT2 can achieve a computation reduction factor of 56.

Figure 1 Ave. MSE against number of re-examined candidates (Top: Tennis; Bottom: Football)

4. CONCLUSIONS

In this paper, we proposed a fast motion estimation algorithm by combining re-examination with 1-bit transform. The motion vectors are estimated first using 1-bit transform and then refined by a re-examination process. Experimental results show that the proposed algorithms can achieve good estimation while maintaining low complexity. Various performance tradeoff can be achieved by the proposed algorithms. In particular, M1BTFS can achieve close-to-optimal estimation and significant complexity reduction at the same time.

5. ACKNOWLEDGEMENT

This work was funded by the RGC CERG grant # HKUST6009/97E.

Method	MSE	Relative complexity	PSNR to FS (dB)
FS	148.47	225	0
1BT	190.85	1 (approx.)	1.09
3SS	208.28	25	1.47
N3SS	183.57	22.53	0.92
4SS	177.32	21.38	0.77
M1BT1	166.38	7	0.49
M1BT2	176.47	4	0.75
M1BTFS	154.62	12.43	0.18

Table 1 Comparison between different methods for the sequence 'Tennis'

Method	MSE	Relative complexity	PSNR to FS (dB)
FS	201.91	225	0
1BT	244.77	1 (approx.)	0.84
3SS	225.75	25	0.48
N3SS	215.63	20.79	0.29
4SS	222.11	19.38	0.41
M1BT1	219.15	7	0.31
M1BT2	227.70	4	0.52
M1BTFS	211.15	14.29	0.19

Table 2 Comparison between different methods for the sequence 'Football'

REFERENCE

- [1] B. Natarajan, V. Bhaskaran, K. Konstantinides, "Low-Complexity Block-Based Motion Estimation via One-Bit Transforms," *IEEE Trans. on Circuits System for Video Technology*, Vol. 7, No 4, pp. 702-706, Aug. 1997
- [2] Xiaobing Lee, Ya-Qin Zhang, "A Fast Hierarchical Motion-Compensation Scheme for Video Coding Using Block Feature Matching" *IEEE Trans. on Circuits System for Video Technology*, Vol. 6, No. 6, pp. 627-635, Dec. 1996.