



Chiew, T-K., Chung-How, J. T. H., Bull, D. R., & Canagarajah, C. N. (2002). Rapid block-based global motion estimation and its applications. In *International Conference on Consumer Electronics, 2002 (ICCE 2002)*. (pp. 228 - 229). Institute of Electrical and Electronics Engineers (IEEE).
10.1109/ICCE.2002.1014005

Peer reviewed version

Link to published version (if available):
10.1109/ICCE.2002.1014005

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact open-access@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

RAPID BLOCK-BASED GLOBAL MOTION ESTIMATION AND ITS APPLICATIONS

Tuan-Kiang Chiew, James T.H. Chung-How, David R. Bull and C. Nishan Canagarajah

Image Communications Group, Centre for Communications Research
University of Bristol, U.K.

Abstract

This paper proposes a novel low-complexity block-based global motion estimation algorithm for real-time digital video processing applications. The algorithm involves an 8x8 block-based local motion estimation followed by a least square method to obtain the global motion vector parameters. The translation-zoom global motion model with 3 parameters was adopted and least square method computation is done after removing outliers. Blocks are considered outliers if: (i) the block is over-cluttered or (ii) the reference block has too little activity for an accurate motion vector to be estimated. The algorithm is used to estimate the global motion of 3 QCIF test streams at 10 fps, and the parameters are used in 5 typical applications. Respective simulations have produced favorable results emphasizing the usefulness of the proposed algorithm.

Introduction

Global motion replaces motion vectors of background regions with single set of parameters. Previous attempts in estimating global motion ([1], [2], [3], [4]) are computationally intensive, which makes them unsuitable for real-time implementation. This paper proposes a block-based global motion estimation algorithm which involves local block matching and estimation of parameters by least square methods. Outliers are eliminated from the estimation process by identifying flat reference areas and cluttered areas. The parameters found are then used in the following applications:

1. H.263+ reference picture resampling,
2. Foreground/background segmentation,
3. H.263+ foreground quality enhancement by modified TMN8 rate control.[6]
4. Error resilience: H.263+ sliced structured mode with unequal error protection and H.223 segmentation [7]for transmission over Bluetooth 1 channels [8].
5. Automatic background sprite formation.

Due to the block-based nature of the algorithm, it can be readily used in real-time applications.

Theory

Evaluation of global motion parameters begins by dividing the picture into blocks of 8x8 pixels, each indexed by the ordered pair (r, c) where r is the row index and c is the column index.. Block matching algorithm depicted in TMN8 [6] is used which yields the $t(r, c) = \{m, u, v\}(r, c)$ tuple, where (u, v) forms the motion vector which produces the minimum sum-of-absolute-difference (SAD), m . The process is modified, however, to obtain 3 tuples of the least 3 SAD: $T(r, c) = \{t_1, t_2, t_3\}(r, c)$, instead of 1 tuple: t_1 is the tuple $\{m_1, r_1, c_1\}$ with the least SAD m_1 , and t_3 is the tuple $\{m_3, r_3, c_3\}$ with the 3rd least SAD m_3 . Next, the variances of blocks from the reference picture which

are used in the motion estimation of the current block (r, c) is evaluated and denoted by $Var(r, c)$. Block (r, c) is considered an outlier if either of the 2 conditions is met:

- $\min_{\substack{i=1,2,3 \\ j=1,2,3 \\ i \neq j}} \left(|u_i - u_j| + |v_i - v_j| \right) > d_{Threshold}$
- $Var(r, c) < Var_{Threshold}$

The first criterion eliminates cluttered regions containing dispersed local minima; the second criterion eliminates 'flat' regions which are sensitive to noise. Each remaining block is used to produce a tuple: $B(r, c) = \{x, y, v_x, v_y\} = \{8c+4, 8r+4, u_1(r, c), v_1(r, c)\}$, where (x, y) is the center of the block and $\langle u_1, v_1 \rangle (r, c)$ is the motion vector.

Assuming the (z, a, b) model where z, a, b denotes zoom factor, horizontal displacement and vertical displacement respectively, we form the equation for each valid block:

$$\begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} x & 1 & 0 \\ y & 0 & 1 \end{pmatrix} \begin{pmatrix} z-1 \\ a \\ b \end{pmatrix} \quad \text{Eq 1}$$

By applying Eq 1 to the $\{x, y, v_x, v_y\}$ tuples from each valid blocks and concatenate them, we obtain an over-determined equation system:

$$\begin{pmatrix} v_{x1} \\ v_{y1} \\ v_{x2} \\ v_{y2} \\ \vdots \end{pmatrix} = \begin{pmatrix} x_1 & 1 & 0 \\ y_1 & 0 & 1 \\ x_2 & 1 & 0 \\ y_2 & 0 & 1 \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} z-1 \\ a \\ b \end{pmatrix} \Rightarrow \begin{pmatrix} z-1 \\ a \\ b \end{pmatrix} = (A^T A)^{-1} A^T B \quad \text{Eq 2}$$

and by using least square methods, z, a and b can be found:

$$B = A \begin{pmatrix} z-1 \\ a \\ b \end{pmatrix} \Rightarrow \begin{pmatrix} z-1 \\ a \\ b \end{pmatrix} = (A^T A)^{-1} A^T B \quad \text{Eq 3}$$

The global motion parameters can then regenerate an 8x8 block-based vector field. Large discrepancy between this vector with the original vector renders the block a foreground block. Repeating this decision criterion on every block yields foreground/background segmentation. This method, along with a simple clustering operation is used to generate the foreground and background slices used in applications 2, 3, 4 and 5.

Results

1. Reference Picture Resampling, H.263.(Annex Q)

The QCIF coast-guard is encoded at 64 kbps, 10 fps with H.263+ Annex Q [5] using warping parameters from our algorithm; results from the base-line setting is used as reference. There is a 0.15 dB improvement in PSNR over base-line H.263 encoder.

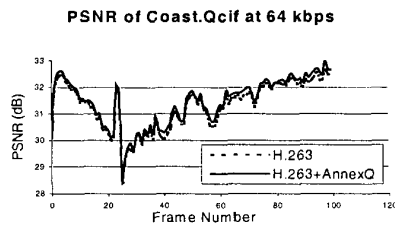


Figure 1. PSNR of reference picture Resampling mode.

2. Automatic foreground/background segmentation.

3 test streams are simulated to extract the foreground from the moving background. The figure below shows the extracted blocks at (a) coast guard, frame 84; (b) Hall, frame 39; and (c) Mobile, Frame 93. In (a) and (b) Foreground for 75% of the frames have been extracted successfully; the extraction in (b) is less successful because the moving man stops occasionally and is hence 'absorbed' into the background.

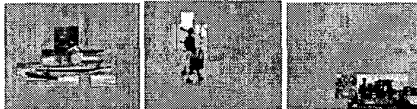


Figure 2. Segmented foreground blocks from selected frame.

3. Foreground quality improvements of H.263 sequences

The extracted foreground in Mobile was encoded at 64 kbps with TMN 8 rate control [6] where a weightage of 0.7 was given to the foreground; the PSNR of the foreground blocks is shown to have improved by 3 dB while keeping the overall PSNR within 0.1 dB deterioration.

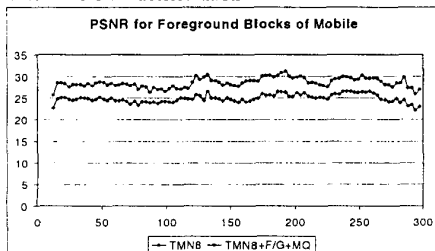


Figure 3. PSNR comparison of foreground blocks.

4. Error resilience: Sliced structured mode with unequal error protection

Simulations with H.223 Annex D [7] over Bluetooth 1 [8] are carried out and the results show an improvement of error-resilience. A total of 100 Bluetooth channels 1 are simulated and with E_b/N_0 of 25 dB ($BER \sim 0.5 \times 10^{-3}$) and there is an average improvement of PSNR of 2 dB. A typical frame in a sequence is shown:



Figure 4. Frame 299, left-GOB coded; right-slice-coded coded both transmitted through DHI, $E_b/N_0 = 25$ dB ($PER = 4.6e^{-3}$)

5. Background sprite formation

The Coastguard sequence is used to generate a background sprite by absorbing only background blocks into the long-term memory, the results after 300 frames is shown. Clearly, remnants of the boat are still fairly visible. Further improvements needs to be done to obtain sprites of a higher quality.



Figure 5. Sprite generated from QCIF coastguard at 10 fps.

Conclusions

Simulation results from the 5 applications shows that our algorithm can be used to:

1. Improve compression efficiency of H.263+ encoder.
2. Extract foreground from moving backgrounds.
3. Improve error resilience via slice-structured mode in H.263+.
4. Form long-term background sprites.

All the applications mentioned have been tested with a Pentium III 850 MHz machine and all algorithms are performed off-line. The longest time taken is 4 times that of the real-time requirements. With increasing processor speed, these applications will be able to run in real-time in the near future.

Detailed descriptions of this technique and results will be presented in the full paper.

References

- [1] T. Vlachos, "Simple method for estimation of global motion parameters using sparse translational motion vector fields," *Electronic Letters*, 34(1), pp. 60-62, Jan. 1998.
- [2] F. Moscheni, F. Dufaux, M. Kunt, "A new two-stage global/local motion estimation based on a background/foreground segmentation," *IEEE Proc ICASSP*, pp. 2261-2264, vol. 4, May 9-12, 1995.
- [3] C.T. Chu, D. Anastassiou, S.F. Chang, "Hierarchical Global Motion Estimation/Compensation in Low Bitrate Video Coding," *IEEE Proc ISCS*, pp 1149-1152, June 9-12, 1997.
- [4] F. Dufaux and J. Konrad, "Efficient, Robust, and Fast Global Motion Estimation for Video Coding," *IEEE Trans on Image Processing*, 9(3) March 2000.
- [5] ITU-T Rec. "H.263 version 2 (H.263+), Video Coding for Low Bit Rate Communication," 1998.
- [6] Thomas R Gardos, "Video Codec Test Model, Near-Term Version 8".
- [7] ITU-T Rec. "H.223, Multiplexing Protocol for Low Bitrate Multimedia Communication," 1998.
- [8] Bluetooth Consortium, Bluetooth Specifications Version 1.1, 2001.