Gradient-based motion vector estimation using variable block shape

Masatoshi HINO Nozomu HAMADA

Signal Processing Lab. Dep. of System Design Engineering, Keio University

ABSTRACT

Motion vector estimation of image seqence is one of the important process in image coding, robot vision etc. A useful method of motion vector estimation in the gradient-based method which uses the relationship between the spatial intensity gradient of image and the intensity difference between frames. We propose two block gradient-based methods using variable block shape and size in order to realize accurate estimation of motion vector. The first one uses variable block shape, and estimation is performed at every pixels. The second method uses square block with variable size, and estimation is done for each block. Simulation results show that both algorithms give more accurate estimation than conventional block gradient-based method.

1 Introduction

Motion vetor estimation methods are divided broadly into three classes such as block matching method[1], gradient-based (optical flow) method[2], and energy method[3]. The matching method gives the displacement minimizing the intensity difference between two blocks in frames as a motion vector. The energy method extracts a plane in which power spectrum of image sequence concentrates. With respect to computational time, energy method is usually used for off-line implementation. Although gradient-based method is sensitive to noise, it is easier to implement and more profitable at computational cost than matching methods.[4]

In general, gradient-based method is applicable to image sequence whose intensity is assumed to be smooth with respect to spatial and temporal variables. The fundamental equation used in the method, called by optical flow constraint (OFC) equation, indicates the relationship between spatial and temporal gradient magnitude (E_x,E_y,E_t) of image intensity $E(x, y, t)$ and motion vector (u,v), such as

$$E_x \cdot u + E_y \cdot v + E_t = 0 \tag{1}$$

at each pixel where $E_x \stackrel{\triangle}{=} \partial E/\partial x$, and same definition is applied for E_y and E_t.

The motion vector (u,v) at a pixel is estimated by solving OFC equation (1) with other conditions. The well-known condition is the spatial smoothness of u and v. The other condition, which is exactly concerned in this study, is the uniformity (i.e, constant values) of (u,v) inside a given block surrounding the pixel. The method using latter condition, which is called block gradient-based method[5], estimates motion vector (u,v) to minimize the least mean squared OFC error within a given block, such as

$$J = \sum_{i=1}^{N} (E_{xi} \cdot u + E_{yi} \cdot v + E_{ti})^2 \tag{2}$$

, where N is the number of pixels in the block. The obtained (u,v) in this manner is usually considered as a motion vector at the central pixel of the processed block. On the other hand, for less computational cost, obtained (u,v) is used for the motion vectors at every pixels in the block. According to the least mean squared method, motion vector (u, v) for this block is calculated by the following equation.

$$\begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} = R^{-1} \begin{bmatrix} -\sum_{i=1}^{N} E_{xi}E_{ti} \\ -\sum_{i=1}^{N} E_{ti}E_{ti} \end{bmatrix} \tag{3}$$

$$R = \begin{bmatrix} \sum_{i=1}^{N} E_{xi}^2 & \sum_{i=1}^{N} E_{xi}E_{yi} \\ \sum_{i=1}^{N} E_{xi}E_{yi} & \sum_{i=1}^{N} E_{yi}^2 \end{bmatrix} \tag{4}$$

In adopting block gradient-based method, determination of block size and shape is the most crucial task. These factors may greatly influence the accuracy of the estimation, especially at motion boundary area. There are some studies addressed the block size determination problem[6]. In these studies, variable size of square block is utilized.

In the proposition of eigenvalue-based method, Kaneko et al.[6] suggest that there are several causes of motion estimation error and it is difficult to estimate motion vector accurately by using fixed block. They define the estimation reliability using eigenvalue distribution which classifies block area into (a)plane, (b)ball, (c)straight line types. They propose a method to change the square block size for low reliable block, which has ball or straight type eigenvalue distribution, until the block is classified into plane type. This method measures the reliability by the eigenvalue not measuring estimation error directly and block shape is restricted within squared.

This paper proposes a motion vector estimation with variable block shape, which is not always limited to square block. The adaptively changing rule of shape and size of block is performed by estimating the error defined by eq.(2) at the successive block shape selection stage discussed later. We will show two proposed methods give more accurate results than the conventional one through simulations.

2 Proposed Method

The main idea underlying two proposed methods is to continue to change both size and shape of the block until the consistency error J becomes less than the prescribed threshold value.

2.1 Variable Block Shape Algorithm

One example of the proposed shape-transforming scheme is depicted in Fig.1. At the beginning of the process, we use an equal appropriate square block surrounding each pixel. That is, for each pixel in each frame, block gradient-based method is applied using same square block size. From the initial fixed size block, we can transform the block by each of two manners, namely, block-growing or block-reducing. For these two possibilities, we adopts reducing scheme. The overall procedure of the first method is summarized as follows.

Estimation Process

(A) For a pixel where motion vector should be estimated, we set the block with fixed square of initial size $(2L_{int}+1)\times(2L_{int}+1)$ surrounding the pixel as its block center. Then, conventional block gradient-based method is applied, then the estimated motion vector at the pixel is obtained.

(B) Compute the consistency error J for each pixel, using the motion vector obtained in STEP(A), in order to check whether the estimation by the initial square block gives accurate estimation or not. All block is divided into two groups, depending on whether the value J is less than the threshold J_{int} or not. In the blocks which are classified in the group satisfying $J < J_{int}$, the initial block size is considered to be suitable for assuming consistent motion vector setting throughout their blocks. Therefore, the estimated motion vectors in their blocks are fixed as the solution. For the blocks classified in another group are treated in the following step. The discussion for determining J_{int} is described at later section.

(C) Make four corner-subregion deleted blocks from the initial square block. Then, compute the least square motion vector solutions and J values for these four reduced blocks to determine the block with minimum J among them. The block with smallest J is called "the selected region at STEP1 " as shown in Fig.1. If the J-value corresponding to the selected region at this step is less than the threshold J_{STEP1}, then the estimation for the corresponding block is finished. Otherwise, go to next step.

(D) Make two corner-deleted blocks as shown in the STEP2 of Fig.1, then, compute motion vectors and Js. Use a new threshold J_{STEP2} for these two blocks as same manner as in the previous procedure.

(E) Successive reduction is performed to the block until the block size becomes the predetermined size(3×3 is selected in Fig.1) or J_{STEPx} is less than the threshold.

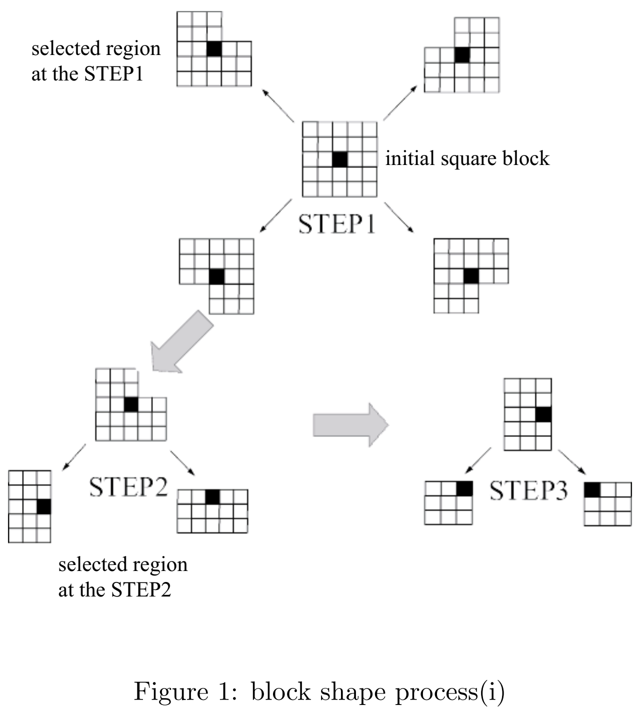

Figure 1: block shape process(i)

Determination of Threshold Values

In the above process, threshold values, such as J_{int}, J_{STEPx} (x=1,2,$\cdots$), should be determined respectively. We adopt the following values.

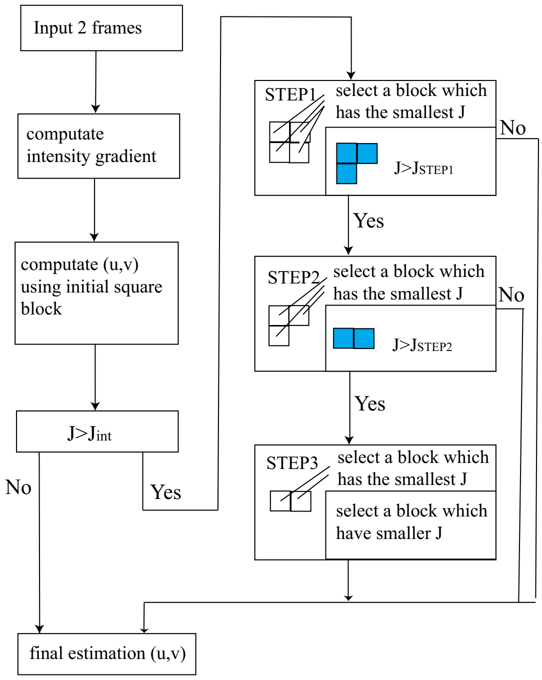

Figure 2: variable block shape process

J_{int} : the median value of J computed for initial blocks satisfying $J <$ prescribed value
J_{STEP1}: $J_{int} \times \{(2L_{int}+1)^2 - {L_{int}}^2\}/(2L_{int}+1)^2$
J_{STEP2}: $J_{int} \times \{(2L_{int}+1)(L_{int}+1)\}/(2L_{int}+1)^2$

2.2 Variable Block Size Algorithm

We can apply same basic ideas for the variable shape scheme described in 2.1 to the variable square size algorithm, which is the topic of this section. The variable-size square block is adopted for block-matching method using dynamic programming technique.[7] The different points from the variable shape algorithm are summarized as follows.

For the initial block size is selected as power of 2 for making ease to divide the block size in two at any stage. At each step, four equally divided square sub-blocks are used to compute motion vector and J from it. The sub-block with $J > J_{int}$ is the selected region at this step. Then, division of block size in two is continued for the selected sub-region.

The flow chart of variable block size process is shown in Fig.3. The threshold values in this case are set as follows.

J_{int} : the median value of J computed for initial blocks satisfying $J <$ prescribed value

J_{STEP1}: $J_{int} \times (1/4) = 0.25\ J_{int}$

2.3 Deletion of outer value of E_t

The accuracy of the above estimation methods is seriously influenced by discontinuous motion. Therefore, in the least square computation in eq.(3) should be done by deleting the pixels having larger gradient with respect to t, i.e. pixels with larger E_t values.

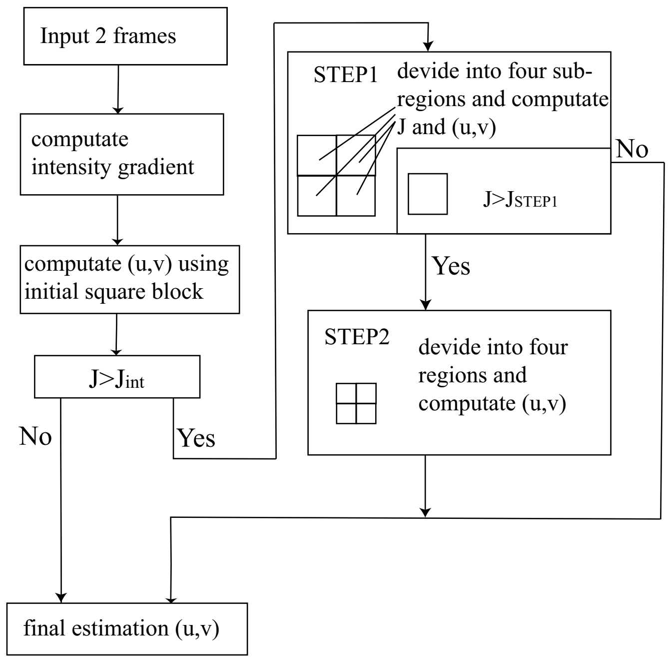

Figure 3: variable block size process

3 Simulation

We applied our methods to an artificial moving image whose one frame is shown in Fig.4. In the sequence of images, two moving objects exist. One object (left side object in Fig.4) moves toward right side with speed 1pixcel/frame, the other object move toward left side with speed 2pixcels/frame. Additionally, right side object is set to be in front of the left object.

The performance index used to measure the estimation error per pixel is given as follows.

$$P = \frac{1}{N^2} \sum_{y=1}^{N} \sum_{x=1}^{N} \sqrt{(\hat{u}-u)^2 + (\hat{v}-v)^2} \qquad (5)$$

, where, $(\hat{u},\hat{v})$ is a estimation value, (u,v) is a true value, N is the number of pixels in the block.

The proposed methods are compared with conventional block gradient-based algorithm with fixed size

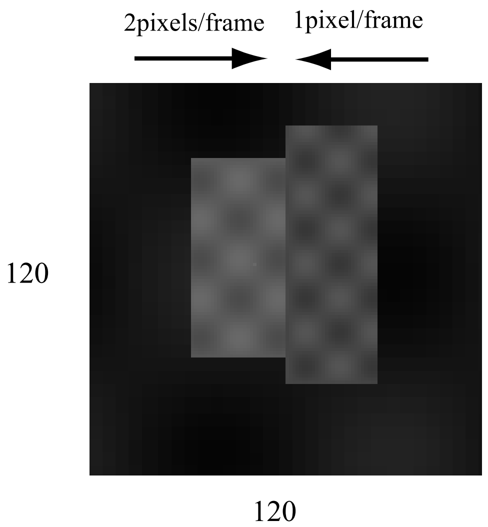

Figure 4: Applied image sequence

square block. The estimation results are shown in Table.1. The result shows that proposed methods are more accurate than conventional method. We will also apply to real moving images to investigate the effectiveness of the proposed method.

Table 1: Simulation result

method	P
conventional method	0.540
variable shape	0.138
variable size	0.129

4 Conclusion

In this paper, we proposed gradient-based motion vector estimation methods using variable block shape and size. The proposed methods give more accurate result than conventional method through a simulation. Besides, deletion of the pixels which have large temporal intensity gradient gave more accurate estimation. Future studies will be an application to real image sequence, and comparison of computational cost.

Acknowledgements

We greatly appreciate to two members in Hamada Lab. One is Mr. Akihiro Kohama who worked with one of the authors at early stage of this research. The other is Mr. Tokiro Satio for his valuable comments.

References

[1] J.R.Jain and A.K.Jain, "Displacement measurement and its application in interframe image coding", IEEE Trans. Commun, vol.COM-29, pp1799-1808, Dec. 1981.

[2] B.K.P.Horn and B.G.Schunk, "Determining optical flow", Artif. Intell. 17, pp.185-203, 1981.

[3] D.J.Heeger, "Optical Flow from Spatialtemporal Filters", IEEE 1st Int. Conf. on Computer Vision, London, pp.181-190, June 1987.

[4] J.L.Barron, D.J.Fleet, S.S.Beauchemin and T.A.Burkitt, "Performance of optical flow techniques", Proc. of IEEE Conf. on CVPR'92, pp.236-242, 1992.

[5] J.K.Kearney, W.B.Thompson and D.L.Boley, "Optical Flow Estimation: An Error Analysis of Gradient-Based Methods with Local Optimization", IEEE Trans, Pattern Anal. Machine Intell., PAMI-9, pp.229-244, 1987.

[6] Y.Kaneko, Y.Shishikui, Y.Tanaka, and F.Okano, "Improvement of motion estimation accuracy using gradient vector distribution in advanced picture coding", IEEE Trans. Broadcast., vol.42, no.3, pp237-244, 1996.

[7] I.Rhee, G.R.Martin, S.Muthukrishnan, and R.A.Packwood,"Quadtree-Structured Variable-Size Block Matching Motion Estimation with Minimal Error", IEEE Trans. CAS for Video Technology, vol.10, no.1, Feb. 2000.