MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

MOVEMENT VECTOR DETECTION WITH RELIABILITY INDICES

Naoya Ohta
C&C Information Technology Research Laboratories
NEC Corporation

1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki,
Kanagawa 213, Japan

ABSTRACT

There is one method, called the *Block Gradient Method* here, whereby to detect image movements from an image sequence. This method has an advantage over the computing costs in a movement vector. However, this method has a general problem, in that the reliability of the detected movements varies with varying conditions, such as the image texture. Therefore, it is important to obtain some indices which indicate the reliability of the detected movements for sequential tasks which use them. In this paper, the reliability of the detected movement by this method is discussed and indices indicating it are proposed. The effectiveness of the indices is shown by experiments. A movement vector stabilizing process, using the indices, is also described. The algorithm proposed in this paper will offer the suitable movement information for the information integration, which is processed in the successional tasks.

INTRODUCTION

Tasks involved in the vision system, as in a moving robot, which deals with a dynamic scene, can be generally classified into two categories. The first tasks derive primitive or low level image information, for example, movements, edges and regions, from an image sequence. The other tasks integrate such information, in order to detect or recognize the objects. There is no problem, if the tasks in the first category can offer complete information which the tasks should offer. However, so long as they cannot, and the reliability of the information varies with varying conditions, it is as important as the information accuracy to offer some values indicating the information reliability. The reason is that the offered information cannot be made good use of, without them. In this paper, the reliability of the movement detected by the *Block Gradient Method*[1][2], which is one kind of gradient based method, is discussed. There are two factors, which cause incorrect detection results. One relates to the assumptions entertained in this method, and the other relates to the image texture or the aperture problem. Corresponding to them, three indices are proposed. One index represents the first factor, and the remaining two indices represent the latter factor.

BLOCK GRADIENT METHOD AND AN INDEX

In this section, the *Block Gradient Method* is explained, from the viewpoint of a model for movement detection and its matching. The first index is also defined.

Equation (1) shows the basic relation between a movement and image derivatives, which is used in gradient based method[1][2].

$$E_x\, u + E_y\, v + E_t = 0 \qquad (1)$$

In this equation, E_x, E_y, and E_t are image derivatives, with respect to space x and y, and time t, respectively. The u and v are x and y components of the movement vector. Although this relation is obtained at each point on an image, one relation is insufficient to determine movement components u and v. For this reason, in the *Block Gradient Method*, a small area (block) S is set on an image, and it is assumed that every point in the area has the same movements. The viewpoint described above can be concluded as a model for the movement detection.

[Model for Movement Detection]

* The relation represented by Eq.(1), is correct everywhere in area S.
* Movements in area S have the same value.

Interpreting this model mathematically, E_x, E and E_t are functions of x and y, defined in area S. This model asserts that E_t can be represented as a linear combination of E_x and E_y.

$$E_t(x,y) = -\,u\, E_x(x,y) - v\, E_y(x,y) \qquad (2)$$

However, because the model is not flawless, it doesn't completely fit the functions obtained from a real image sequence. It is necessary to add an error function E_r to Eq.(3).

$$E_t(x,y) = -\,u\, E_x(x,y) - v\, E_y(x,y) + E_r(x,y) \qquad (3)$$

Matching the model is accomplished by minimizing the norm for the error function E_r, namely $|E_r|$.

[Criterion for Model Fitting]

* Minimizing the norm for the error function $E_r(x,y)$.

The movement vectors are determined as u and v, which give the minimum value of $|E_r|$. To find u and v values which satisfy this condition, the function $E_r(x,y)$ should be selected to be orthogonal to functions $E_x(x,y)$ and $E_y(x,y)$.

$$(E_r \cdot E_x) = 0, \qquad (E_r \cdot E_y) = 0 \qquad (4)$$

The notation $(f_1 \cdot f_2)$ denotes the inner product of functions f_1 and f_2.

$$(f_1 \cdot f_2) = \iint_S f_1(x,y)\, f_2(x,y)\, dx\, dy \qquad (5)$$

Creating inner products of Eq.(3) with E_x and E_y, and applying the condition represented by Eq.(4), the next equation can be derived.

$$\begin{pmatrix} -g \\ -f \end{pmatrix} = \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \qquad (6)$$

The (u_0,v_0) denotes the movement vector, which should be calculated. The symbols appearing in Eq.(6) are as follows.

$$a=(E_x \cdot E_x),\ b=(E_y \cdot E_y),\ c=(E_t \cdot E_t), \qquad (7)$$
$$f=(E_y \cdot E_t),\ g=(E_t \cdot E_x),\ f=(E_x \cdot E_y)$$

Thus, the movement vector is calculated by the following equation.

$$\begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \frac{1}{ab - h^2} \begin{pmatrix} b & -h \\ -h & a \end{pmatrix} \begin{pmatrix} -g \\ -f \end{pmatrix} \qquad (8)$$

Now, consider the meaning of the minimum value of $|E_r|$. An image pattern movement generates the norm for the time derivative function E_t. This value is divided to two parts, in the manner described above. One is the part explained by the model and the other is the part not explained. The latter is E_r, when its norm is minimum. It is denoted as E_{r0}. The power of E_{r0}, namely $|E_{r0}|^2$, is the remaining power of E_t, which can't be explained by the model. Thus, $|E_{r0}|$ represents the model fitness grade. The first index r_e is defined as this value.

$$r_e = |E_{r0}| = \min(|E_r|) \tag{9}$$

The r_e value is calculated by the next equation, using values shown in Eq.(7) and Eq.(8).

$$r_e = \sqrt{g\ u_0 + f\ v_0 + c} \tag{10}$$

MOVEMENT RELIABILITY AND TWO INDICES

The reliability of a detected movement vector, which depends on an image pattern, is discussed in this section. The indices which represent it are also defined.

The movement of an image is detected as the norm of E_t. The greater the $|E_t|$ generated by a unit movement is, the more easily detectable the movement is, because of the robustness versus noise. Therefore, the reliability of a detected movement can be estimated by the ratio of $|E_t|$ against the length of the movement.

A movement $(\delta u, \delta v)$, whose direction is θ and whose length is δr, can be represented as follows.

$$\begin{pmatrix} \delta u \\ \delta v \end{pmatrix} = \delta r \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \tag{11}$$

Symbolizing the norm of E_t generated by this movement as $\delta|E_t|$, it is represented by Eq.(12).

$$\delta|E_t| = \delta r\ |\ E_x \cos\theta + E_y \sin\theta\ | \tag{12}$$

Symbolizing the ratio by R, R^2 results in Eq.(14).

$$R = \delta|E_t| \ /\ \delta r \tag{13}$$

$$R^2 = |\ E_x \cos\theta + E_y \sin\theta\ |^2 \tag{14}$$

$$= (\cos\theta\ \sin\theta) \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

Equation (14) can be transformed into the next equation by the principal axis transformation.

$$R^2 = (\cos\theta\ \sin\theta)\ (\ \mathbf{e}_1\ \mathbf{e}_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \mathbf{e}_1^t \\ \mathbf{e}_2^t \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \tag{15}$$

In Eq.(15), λ_1 and λ_2 are the eigenvalues for the coefficients matrix ($\lambda_1 \geq \lambda_2$), and $\mathbf{e}_1$ and $\mathbf{e}_2$ are the unit length eigenvectors corresponding to them. From Eq.(15), it is known that the ratio R becomes maximum ($\sqrt{\lambda_1}$), when the movement angle θ coincides with the $\mathbf{e}_1$ direction, and becomes minimum ($\sqrt{\lambda_2}$), when in the $\mathbf{e}_2$ direction.

The indices, which represent the reliability depending on the image patten, are defined as follows.

$$\mathbf{r}_1 = \sqrt{\lambda_1}\ \mathbf{e}_1, \quad \mathbf{r}_2 = \sqrt{\lambda_2}\ \mathbf{e}_2 \tag{16}$$

The $\mathbf{r}_1$ value gives the most reliable direction for the detected movement and its grade. The $\mathbf{r}_2$ gives the least reliable direction and grade. The formulae for a numeric calculation of the indices are shown below. The definitions of symbols appearing in the next equation are found in Eq.(6) and Eq.(7).

$$\lambda_1 = (\ a + b + \sqrt{(a-b)^2+4h^2}\ )\ /\ 2 \tag{17}$$

$$\lambda_2 = (\ a + b - \sqrt{(a-b)^2+4h^2}\ )\ /\ 2$$

$$\mathbf{e}_1 = (\ h\ ,\ \lambda_1 - a\ )\ /\ \sqrt{h^2+(\lambda_1-a)^2}$$

$$\mathbf{e}_2 = (\ h\ ,\ \lambda_2 - a\ )\ /\ \sqrt{h^2+(\lambda_2-a)^2}$$

EXPERIMENTS

Experiments were carried out in order to investigate the effectiveness of the indices. A pseudo image sequence for the first experiment was created by the procedures described below. First, two shots of the same laboratory scene were taken. Figure 1 shows one of the two. Next, one of the two frames was moved, by 1 pixel distance, along both x and y axes. Each frame for the pseudo image sequence was smoothed by 5×5 rectangular mask.

Using this pseudo image sequence, the movement vectors and their reliability indices are calculated. The block size used is 16×16 versus the 512×512 whole image size. Detected movements are displayed in Fig.2. The norm or value for individual indices is displayed in Fig.3 with gray level.

While the detected movements are correct, where an image pattern has variation, those include errors, where it has no variation. Furthermore, movement vectors on an edge pattern, such as the window frame, have a large error along the direction parallel to the edge. Corresponding to these facts, while both $|\mathbf{r}_1|$ and $|\mathbf{r}_2|$ for a correct vector are large, those for on incorrect vector are small. With regard to a vector on an edge, its $|\mathbf{r}_1|$ value is large versus a small $|\mathbf{r}_2|$. As the $|\mathbf{r}_2|$ is large in significant areas on an image, it has a characteristic like the *Interest Operator*[3].

As the $|\mathbf{r}_2|$ represents the least reliable direction and grade of the movement, it is possible to estimate the vector's accuracy, using only the $|\mathbf{r}_2|$ value, for simplicity. Needles in Fig.4 express their $|\mathbf{r}_2|$ grade by the display density. The lightest gray expresses less than 15, the medium gray indicates from 15 to 30, and the darkest gray shows more then 30. It was found that the correct vectors can be selected by thresholding about $|\mathbf{r}_2|$. The averages of the error included in the movement, are respectively 1.01, 0.30 and 0.15 pixel. Figure 5 shows the relation between the error ($|\delta v|$) and $|\mathbf{r}_2|$ by a distribution graph. Where the $|\mathbf{r}_2|$ is small, the error is widely distributed. However, as the $|\mathbf{r}_2|$ becomes large, the error becomes small.

The r_e represents the model fitness grade. Model unfitness is mainly caused by the change in the image pattern, which occurs in a movement boundary area. Because the used image sequence doesn't include a movements boundary, the r_e value doesn't exhibit significant features. Another pseudo image sequence with a movement boundary was created. A rectangular area was moved by 1 pixel along both x and y axes, in an opposite direction to that for the whole image movement. The r_e values on this image are displayed in Fig.6(a). Large value is seen on the boundary. This effect is emphasized by applying the *Iterative Gradient Method*[4]. In this method, in the case that the detected movement is greater than 1 pixel, the window on one frame is reset with a displacement. The displacement is selected to be the nearest integer value to the detected movement. The movement is iteratively calculated according to pre-determined times or until convergence. The result with one iteration is displayed in Fig.6(b).

An experiment was also carried out on a real image sequence. One frame for the input image sequence is shown in Fig.7. The movement detection result is shown in Fig.8. In this case, the *Iterative Gradient Method* was used with one iteration. The vectors, whose $|\mathbf{r}_2|$ values are greater than 17, are displayed in Fig.9. It was found that obviously incorrect vectors, such as on the road, were abandoned.

MOVEMENT VECTORS STABILIZATION

As a sample of procedures succeeding the movement detection, a vector stabilization process using the relaxation

Figure 1: Original image.

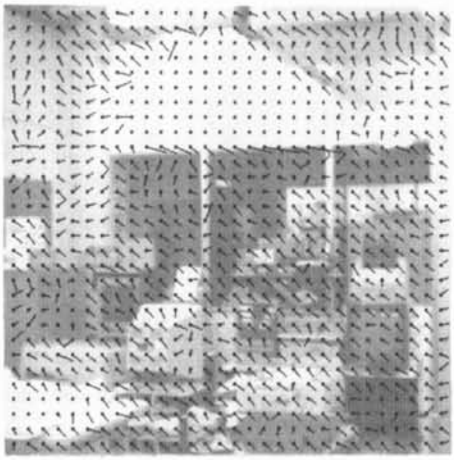
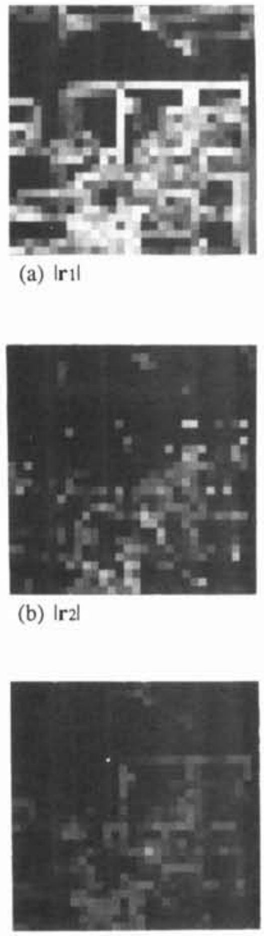


Figure 2: Detected movements. Correct movements.



(a) $|\mathbf{r}_1|$

(b) $|\mathbf{r}_2|$

(c) r_e

Figure 3: Values of the indices.

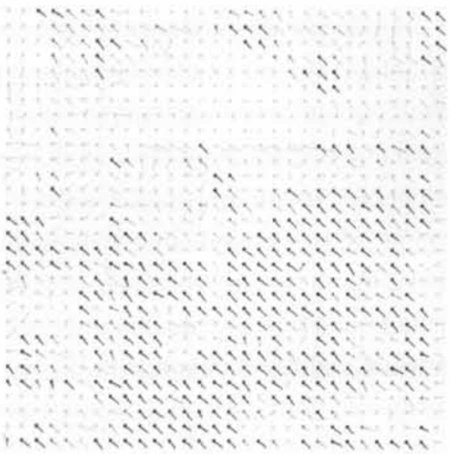


Figure 4: Detected movements vs. $|\mathbf{r}_2|$.

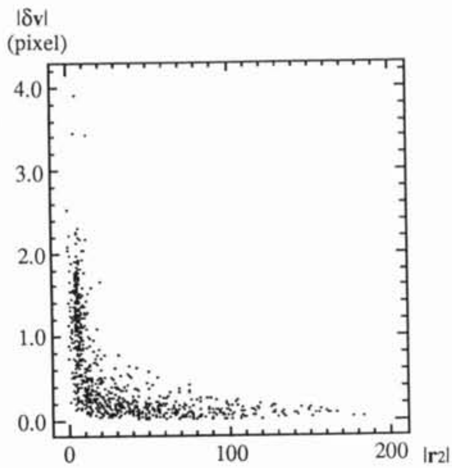


Figure 5: Relation between the error and $|\mathbf{r}_2|$.

method is described. This process uses the two reliability indices, $\mathbf{r}_1$ and $\mathbf{r}_2$, and modifies vectors easier in the less reliable direction, than in the more reliable one. The energy functional, which should be minimized, is defined by the next formula[5][6].

$$\Phi(u,v) = \alpha\Psi(u,v) + \Omega(u,v) \tag{18}$$

$$\Psi = \iint \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 dxdy \tag{19}$$

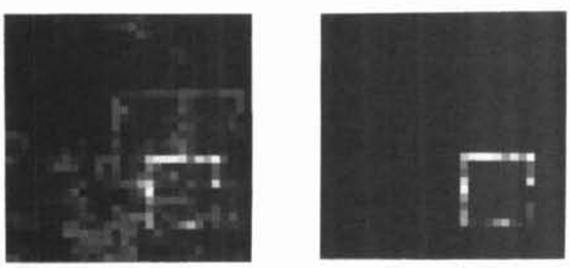
$$\Omega = \iint (\mathbf{r}_1\cdot\Delta\mathbf{v})^2 + (\mathbf{r}_2\cdot\Delta\mathbf{v})^2 dxdy \tag{20}$$

$$\Delta\mathbf{v} = (\ u-u_0\ ,\ v-v_0\ )$$

Functional ψ is the stabilizing functional and functional Ω is the penalty functional. The α is a constant balancing two constraints. The Ω is defined as large energy is required to modify a vector forward in the highly reliable direction.

The *Euler-Lagrange equation* for the energy functional Φ were calculated, and differentials in the equation are replaced with finite differences. The SOR scheme for numerical calculation results in the next formulas.

$$u_{ij}^{k+1} = \xi\{\ \alpha(\ u_{i+1\,j}^{k+1} + u_{i-1\,j}^{k+1} + u_{i\,j+1}^{k} + u_{i\,j-1}^{k}\ ) + a_{ij}u_{0ij} + h_{ij}(v_{0ij}-v_{ij}^{k}\ )\ \}\ /\ \{\ a_{ij} + 4\alpha\ \} - (\xi-1)u_{ij}^{k} \tag{21}$$



(a) No iteration.

(b) One iteration.

Figure 6: r_e value.

$$v_{ij}^{k+1} = \xi\{\ \alpha(\ v_{i+1\,j}^{k+1} + v_{i-1\,j}^{k+1} + v_{i\,j+1}^{k} + v_{i\,j-1}^{k}\ ) + b_{ij}v_{0ij} + h_{ij}(u_{0ij}-u_{ij}^{k+1})\ \}\ /\ \{\ b_{ij} + 4\alpha\ \} - (\xi-1)v_{ij}^{k}$$

The superscript k represents the iteration step, ξ is the acceleration parameter. Practical values for the parameters are 1.8 and 1000 as ξ and α, in this experiment.

The calculation result is shown in Fig.10. Although the movement vectors were stabilized, all didn't become ac-

Figure 7: Original image.

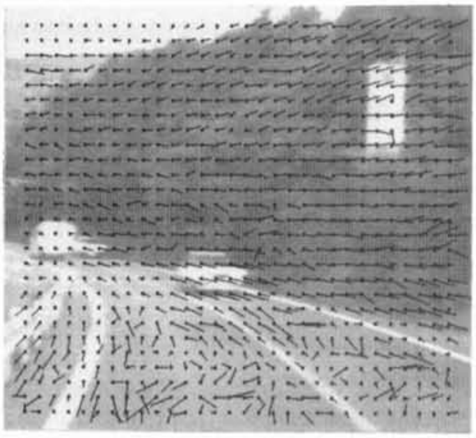


Figure 8: Detected movements.

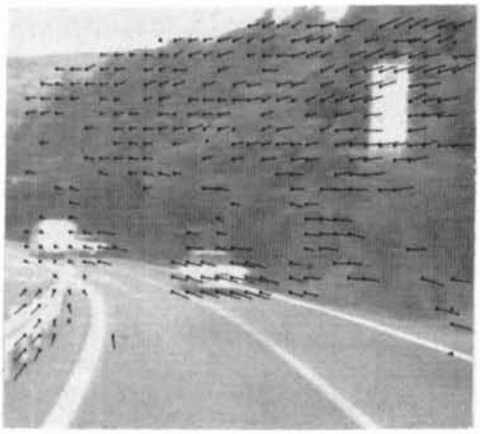


Figure 9: Detected movements ($|r_2| \geq 17$).

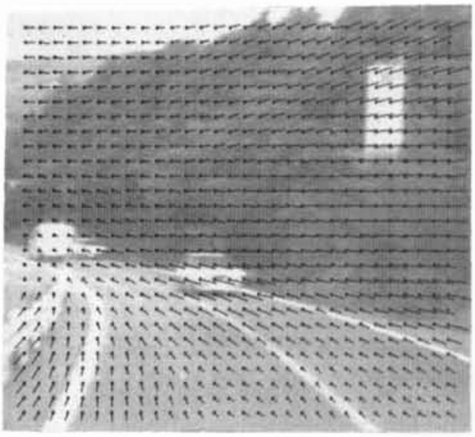


Figure 10: Stabilized movements.

curate. Especially, vectors in low reliability areas, like on the road, are obviously incorrect. However, vectors in rather reliable areas, such as on the woods, surely became more accurate. Thus, used with the reliability indices, this process is useful in a lot of cases.

CONCLUSION

The reliability of the detected movement vectors by the *Block Gradient Method* has been discussed, and indices representing the reliability have been proposed. The effectiveness of the indices was shown by experiments. Moreover, a vector stabilizing process, which uses the indices, has been shown. The proposed algorithms are suitable for practical use, regarding their effectiveness and the required computing cost.

The remaining problem is to use the index r_e in the vector stabilizing process. As this index has a large value at a movement boundary, it is planned to control the smoothness constraint by this index.

ACKNOWLEDGMENTS

The author thanks members in the Pattern Recognition Research Laboratory, especially Mr. Tsutomu Temma and Dr. Johji Tajima for considering for a good research environment. The author also thanks Mr. Hajimu Kawakami for profitable discussions and suggestions.

REFERENCES

[1] J.O.Limb and J.A.Murphy, "Estimating the Velocity of Moving Images in Television Signals", Computer Graphics and Image Processing, 4, pp.311-327, 1975.

[2] T.Fukinuki, "Measurement of Movement and Velocity of Moving Objects with Picture Signals", IEICE Technical Report, IE78-67, pp.35-41, 1978. (Japanese)

[3] H.P.Moravec, "Visual Mapping by Robot Rover", Proc. 6th IJCAI, pp.598-600, 1979.

[4] M.Wada and H.yamaguchi, "Iterative Gradient Method for Movement Estimation", Trans. IEICE, Vol.J68-D, No.4, pp.663-670, 1985. (Japanese)

[5] D.Terzopoulos, "The Computation of Visible-Surface Representation", IEEE Trans. Pattern Anal. Machine Intell., Vol.10, No.4, pp.417-438, 1988.

[6] B.K.P.Horn and B.G.Schunk, "Determining Optical Flow", Artif. Intell., Vol.17 pp.185-203, 1981.