

A Robust Matching Algorithm Based on Global Motion Smoothness Criterion

Mikhail Mozerov¹ and Vitaly Kober²

¹Institute for Information Transmission Problems of RAS,
19 Bolshoi Karetnii, Moscow, Russia
mozer@iitp.ru

²Department of Computer Science, Division of Applied Physics, CICESE,
Km 107 Carretera Tijuana-Ensenada, Ensenada 22860, B.C., México
vkober@cicese.mx

Abstract. A new robust matching algorithm for motion detection and computation of precise estimates of motion vectors of moving objects in a sequence of images is presented. Common matching algorithms of dynamic image analysis usually utilize local smoothness constraints. The proposed method exploits global motion smoothness. The suggested matching algorithm is robust to motion discontinuity as well as to noise degradation of a signal. Computer simulation and experimental results demonstrate an excellent performance of the method in terms of dynamic motion analysis.

1 Introduction

Extraction of motion information is an essential part of any video processing system. Such popular tasks as relative depth from motion, 3-D shape recovery, autonomous vehicle or robot navigation, and moving object detection usually involve various motion analysis techniques. Many techniques analyzing motion from optical flow computation have been proposed in the past two decades [1-7]. However, reliable optical flow estimation remains a difficult problem when smoothness constraint is violated. Furthermore, algorithms based on the optical flow concept are also very sensitive to large values of sought motion vectors (more than one pixel) and to noise degradation of a signal. Among a wide variety of approaches, there exist three main categories of motion estimation methods: gradient-based methods [1], frequency-based methods [2], and matching techniques [4]. In this work we remain in framework of matching concept that aims to solve the correspondence problem. It is well known that the correspondence problem is inherently ambiguous, and some additional information must be added to solve it. Various approaches have been suggested for solving the correspondence problem [8-12]. The identification of correspondence between the same points in consecutive images is often formulated as a local (area-based) optimization problem, or shortest-path technique [9]. On the other hand, the correspondence between the same points in neighbor images can be considered as a global optimization problem [11-12]. So the matching is carried out between 2-D arrays of images. A drawback of this approach is owing to contradictions between the

smoothness constraint of motion vectors between adjacent pixels and a real signal discontinuity at borders of object segments. In this paper, the local motion smoothness constraint is replaced by a global motion smoothness criterion. The latter yields a high performance in optical flow based techniques. We suggest a new matching algorithm, which is based on dynamic programming and global smoothness criterion. Computer simulation with various image sequences shows that the proposed algorithm is robust to motion discontinuity and to noise degradation of a signal. Experimental results with real dynamic images illustrate a very good performance of the method in terms of motion vector accuracy.

2 Optical Flow Constraint and Global Optimization Technique

A common assumption in dynamic image analysis is that the intensity of a point keeps constant value along its trajectory. More precisely, let $f(x,y,t)$ denote the intensity of the pixel at the coordinates (x,y) and time t . Starting from the point (x_0,y_0) at time t_0 , we define the trajectory of this point in time as $(x_0 + u_x\delta t, y_0 + u_y\delta t, t_0 + \delta t)$ with

$$f(x_0, y_0, t_0) = f(x_0 + u_x\delta t, y_0 + u_y\delta t, t_0 + \delta t), \tag{1}$$

where $u = (u_x(x_0,y_0,t_0), u_y(x_0,y_0,t_0))$ is the velocity vector (called the flow vector) of a point (x_0,y_0) at time t_0 and δt is called the interframe interval.

In motion analysis common algorithms usually work if some conditions are fulfilled. For instance gradient-based methods [1] are subject to that the motion vector $(\delta r_x = u_x\delta t, \delta r_y = u_y\delta t)$ and the interframe interval δt are small. Therefore Taylor’s expansion may be applied to Eq. (1),

$$f(x + \delta r_x, y + \delta r_y, t + \delta t) = f(x, y, t) + \frac{\partial f(x, y, t)}{\partial x} \delta r_x + \frac{\partial f(x, y, t)}{\partial y} \delta r_y + \frac{\partial f(x, y, t)}{\partial t} \delta t, \tag{2}$$

and finally

$$\frac{\partial f(x, y, t)}{\partial x} u_x + \frac{\partial f(x, y, t)}{\partial y} u_y + \frac{\partial f(x, y, t)}{\partial t} = 0. \tag{3}$$

Note that Eq. (3) is not sufficient for computing the components of velocity field. Another drawback of optical flow approach is a severe restriction for sought values of motion vectors. In general for a sampled signal (digital image), Eq. (3) holds only if motion vector values equal or less than one pixel (image sampling interval). Actually, the sign “=” in Eq. (3) must be replaced by “≈”. This means that the time interval in the most cases is a fixed value.

We propose a new method that is based on matching techniques. With the help of matching techniques Eq. (1) can be rewritten as

$$a(x, y) = g(x + r_x(x, y), y + r_y(x, y)), \tag{4}$$

where $a(x,y)$ is the intensity function of anchor frame (snapshot at t_0) and $g(x,y)$ is the intensity function of the target frame (snapshot at $t_0 + \delta t$).

We use the following dissimilarity function

$$E_{i,j,k,l} = |g_{i+k\delta r, j+l\delta r} - a_{i,j}|^n \tag{5}$$

as a local feature of correspondence matching as well as a local error function to compute the optical flow. In a sampled space the dissimilarity function values can be described by a 4-D array $\{E_{i,j,k,l}; i=0,\dots, I-1; j=0,\dots, J-1; k=-K,-K+1,\dots, K; l=-L,-L+1,\dots, L\}$. Here, I and J are the size of images, and K, L are reasonable values to carry out the correspondence matching, $a(i,j)$ and $g(i,j)$ are the sampled intensity functions of the anchor and the target frames, respectively. Suppose that the motion vectors possess subpixel values, n equals to 1 or 2, and, finally, the vector's sampling interval belongs to the interval $0 < \delta r \leq I$ that means a subpixel accuracy of the vector estimation in Eq. (4).

Now we need additional constraints to solve the problem. The common approach is that motion vectors possess small signal variations. So absolute differences between all adjacent elements of the motion vector field are assumed to be bounded by values δv :

$$|\Delta \mathbf{r}_{i,j}| \equiv (|k_{i,j} - k_{i\pm 1, j\pm 1}|, |l_{i,j} - l_{i\pm 1, j\pm 1}|) \leq \delta v. \tag{6}$$

Now, for a sampled signal the global optimization problem is formulated as follows: find the motion vector field $\{\mathbf{r}_{i,j}\} \equiv \{k_{i,j}, l_{i,j}\}$ with the local smoothness constraint in Eq. (6) in such a way to minimize the sum of the dissimilarity function Eq. (5) evaluated over all elements of images:

$$(\mathbf{r}_{i,j}) = \underset{\substack{r_{i,j}, |\Delta \mathbf{r}_{i,j}| \leq \delta v}}{\text{ARG MIN}} \left(\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} E_{i,j,r} \right), \tag{7}$$

where $|\Delta \mathbf{r}_{i,j}| \leq \delta v$ denote the smoothness constraint in Eq. (6).

On the other hand, the most successful methods based on optical flow concept utilize the global motion smoothness criterion. In this case the objective function is a combination of the dissimilarity function and the squared values of gradients. So Eq. (7) can be rewritten as

$$(\mathbf{r}_{i,j}) = \underset{\substack{r_{i,j}}}{\text{ARG MIN}} \left(\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} (E_{i,j,r} + w |\Delta \mathbf{r}_{i,j}|^2) \right), \tag{8}$$

where w is a regularizing parameter.

The method proposed in [11] optimizes Eq. (7) by means of modified dynamic programming. The problem in Eq. (8) can be also solved withy modified dynamic programming:

$$\begin{aligned}
 O^I(E_{i,j,r}) &= E_{i,j,r}^I = S_{i,j,r}^{I+} + S_{i,j,r}^{I-} - E_{i,j,r}, \\
 S_{i,j,r}^{I\pm} &= \text{MIN} \left(S_{i\pm k,j,r}^{I\pm} + E_{i,j,r} + w |k_{i,j} - k_{i\pm 1,j}|^2 \right), \\
 S_{0,j,r}^{I+} &= E_{0,j,r}, \quad S_{I-1,j,r}^{I-} = E_{I-1,j,r}, \\
 O^J(E_{i,j,r}^I) &= E_{i,j,r}^J = S_{i,j,r}^{J+} + S_{i,j,r}^{J-} - E_{i,j,r}^I, \\
 S_{i,j,r}^{J\pm} &= \text{MIN} \left(S_{i,j\pm l,r}^{J\pm} + E_{i,j,r}^I + w |l_{i,j} - l_{i,j\pm 1}|^2 \right), \\
 S_{i,0,r}^{J+} &= E_{i,0,r}^I, \quad S_{i,J-1,r}^{J-} = E_{i,J-1,r}^I.
 \end{aligned} \tag{9}$$

Finally, the solution can be found by simple procedure,

$$(\mathbf{r}_{i,j}) = \underset{\mathbf{r}_{i,j}}{\text{MIN}} \left(O^J \left[O^I (E_{i,j,r}) \right] \right), \tag{10}$$

where O^I and O^J are two consecutive transforms with the use of the recurrence operator in Eq. (9) along I and J axis, respectively.

Note, that after one transform (O^I , e.g.) the solution is equal to the optimal path that can be calculated with conventional dynamic programming. However the global optimization in Eq. (8) requires 2D optimization, whereas conventional dynamic programming solves only 1D optimization problem. After the second transform O^J the necessary optimization is obtained.

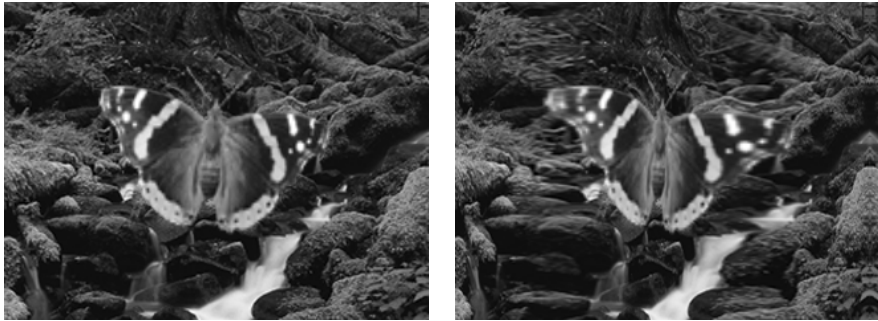
So, the proposed algorithm consists of the following steps.

- Form the initial 4D matrix $E_{i,j,r}$ of the dissimilarity function using Eq. (5).
- Perform two consecutive transforms with the use of the recurrence operator in Eq. (9) along I and J axes, respectively.
- Extract motion information with the help of Eq. (10).

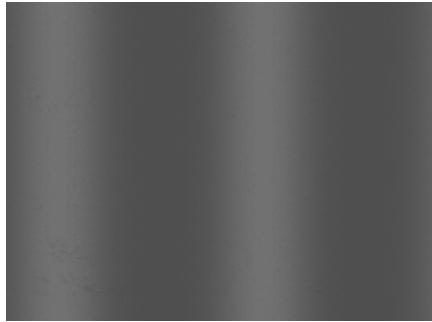
3 Computer Experiments

Computer experiments are carried out to illustrate and compare the performance of conventional matching and proposed algorithms. We are also interested in understanding how well the proposed matching behaves if a signal distorted due to additive noise. In our computer experiments matched pair of images are generated using known test motion vector fields. In our case, the conventional representation of resultant motion vectors by needle diagrams is not effective visual tool. We illustrate matching results by scalar maps. The gray-scale map presentation requires scalar values of motion vector fields. Generated test fields are also scalar (like horizontal disparity in stereo images).

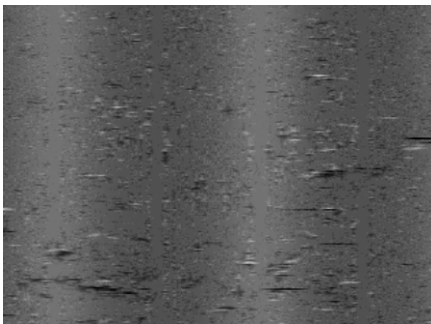
Fig. 1 (a) shows the pair of matched test images. The scalar valued map of a known vector field is given in Fig. 1 (b). Figs. 1 (c) and (d) show the scalar valued maps obtained by matching with local smoothness constraint and with global smoothness criterion, respectively. The visual comparison of the resultant maps shows that the performance of the proposed algorithm is obviously much better.



(a)



(b)



(c)



(d)

Fig. 1. (a) Test pair of matched images. (b) The scalar valued map of a known vector field. (c), (d) The scalar valued maps obtained by matching with local smoothness constraint and with global smoothness criterion, respectively.

Next we carry out experiments with test images that are degraded using additive Gaussian noise. Fig. 2 (a) shows the pair of matched test images degraded due to the noise. The map of a known vector field is given in Fig. 2 (b). Figs. 2 (c) and (d) show the resultant maps obtained by matching with local smoothness constraint and with global smoothness criterion, respectively. The visual and numerical analysis of the resultant depth maps shows that the proposed matching algorithm is robust to the noise degradation of a signal.

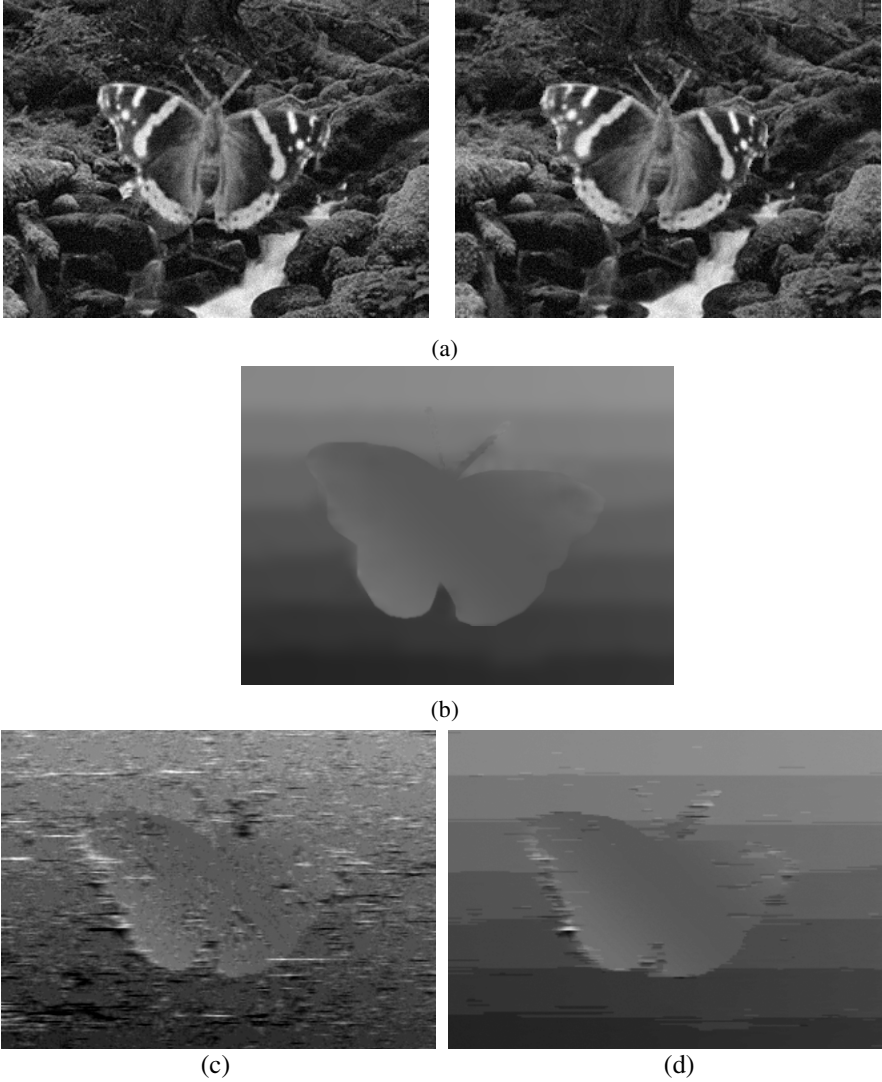


Fig. 2. (a) Test pair of matched images degraded due to additive Gaussian noise. (b) The scalar valued map of a known vector field. (c), (d) The resultant scalar valued maps obtained by matching with local smoothness constraint and with global smoothness criterion, respectively.

The proposed operator in Eq. (9) includes a smoothness parameter w that must be defined during matching process. With the help of many computer experiments we found that this parameter can be represented as follows:

$$w = 1.5 \frac{\sum_{i=0}^{I-1} |a_{i,j} - a_{i+1,j}|}{I}, \tag{11}$$

where $a_{i,j}$ is the image intensity value of the anchor frame.

In other word, the sought parameter is proportional to the mean absolute value of signal gradient along the chosen axis (in Eq. (11) it is the horizontal axis of matched images).

We carried out many computer experiments with different simulated motion vector fields, motion fields, and degradation of matched images. So numerical analysis on the base of the mean squared errors (MSE) criterion shows that the proposed algorithm has advantage over conventional matching algorithms.

4 Conclusion

In this paper, a new motion estimation method based on dynamic programming matching and global motion smoothness criterion has been proposed. The method demonstrates much better results than those obtained with the use of local smoothness constraints. The proposed method is robust to additive noise.

References

1. Horn B.K.P., Schunck B.G., Determining optical flow. *Artificial Intelligence*, Vol. 17, (1981), 85-204.
2. Watson A.B., Ahumada A.J., Motion: Perception and Representation. In: Tsotsos, J.K., (ed.), (1983), 1-10.
3. Ohta Y. and Kanade T., Stereo by intra- and inter-scanline search using dynamic programming, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.7, (1985), 139-154.
4. Anandan P., Measuring Visual Motion from Image Sequences. PhD thesis, Univ. of Massachusetts, Amherst (1987).
5. Heitz F., Boutheymy P., Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, (1993), 1213-1232.
6. Kanade T. and Okutomi M., A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, pp. (1994), 920-932.
7. Cedaras C., Shah M., Motion based recognition: A survey. *Image and Vision Computing*, Vol. 13, (1995), 129-154.
8. Anthony Y.K.H., and Pong T.C., Cooperative fusion of stereo and motion, *Pattern Recognition*, Vol. 28, (1995), 553-562.
9. Chung H.Y., Yung N.H.C., Cheung P.Y.S., Fast motion estimation with search-center prediction, *Optical Engineering*, Vol. 40, (2001), 952-963.
10. Petrakis E.G.M., Diplaros A., and Milios E., Matching and retrieval of distorted and occluded shapes using dynamic programming, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.24, (2002), 1501-1516.
11. Mozerov M., Kober V., Tchernykh A., Choi T. S., Motion estimation with a modified dynamic programming, *Optical Engineering*, Vol. 41, (2002), 2592-2598.
12. Mozerov M., Kober V., Motion Estimation Based on Hidden Segmentation, *IEICE Transaction on Fund.*, Vol. E88-A, (2005), 376-381.