Auto-Correlation Probabilistic Relaxation Matching Method

Xin Wang[1], Masanori Sugisaka[2], Jin Wang[3]

[1]Graduate School of Engineering,
The University of Tokushima, Tokushima 770-8506, Japan

[2]Department of Electrical and Electronic Engineering,
Oita University, Oita 870-1192 Japan

[3]College of Textile and Garments,
Hebei University of Science and Technology, Shijiazhuang 050031 China

Abstract

The statistics of the neighborhood gradients in an image can yield useful descriptions of target features. This paper presents a new approach, which is iterative and begins with the detection of all potential correspondence pairs, on auto-correlation relaxation algorithm by using those features to match image clips. Each pair of matching points is assigned a number representing the probability of its correspondence. The probabilities are iteratively recomputed to get global optimum sets of pairwise relations. This algorithm could be found wide applications of matching video frames, industrial detection recognition and so on.

1. Introduction

Image matching is the problem of finding correspondence points in two or more images of the same scene, traditionally. Recently, many researches also show their interest in image stream, which is large amount of information for 3-D image reconstruction from 2-D images.

Two image points p and p' match if they result from the projection of the same physical point P in the scene, a property that is often approximated by a similarity constraint requiring, for example, p and p' to have similar intensity or color. The desired output of an image correspondence algorithm is a disparity map, specifying the relative displacement of matching points between images. The image correspondence problem is inherently under constrained and further complicated in image stream and also by the fact that the images typically contain noise. Traditional approaches thus either try to only recover a subset of matches, or make additional assumptions.

The feature based matching problem can generally be fall into three categories: matching points, curves, and areas. Point based matching is, for a location in one image, to find the displacement that aligns this location with a matching location in the other image [5][7][12]; curve-based matching is by analysis of the similarity and compatibility between curves in different images [2][3][4]; area based method yield a dense disparity map by matching small image patches as whole with respect to geometry, textures [6][8][10][11]. Theoretically, point correspondences, which matches points with a certain amount of local information, are the robust way. Traditional point based approaches, however, have two foundational difficulties when applied to more general scenes. First, they usually assumed known camera geometries for stereo matching, so less point's relation in one image considered. Second, the similarity constrains are seriously required.

We depict the attributes of a target by its significant energy points. Our idea to understand a target is from its points of this target to its lines, from its points to its lines and areas. This is related to manipulate our knowledge database in our further works. If the proposed auto-correlation approach is used alone, it also gives very satisfied results [14].

2. Auto-correlation model

Let $I(X)$(also denoted as I), $X=\{x_i|_{x_i \in X}, X \subset \mathcal{R}^n\}$, be the image function in an image frame. Given the shift of X as ΔX, the auto-correlation function is defined as:

$$f(X) = \sum_{w} \big(I(X) - I(X + \Delta X)\big)^2 \quad (1)$$

where X presents the global position in the working window w . According to the Taylor expansion, in the case of $X \subset \mathcal{R}^2$, we rewrite the elements in X as x, y, the item $I(X + \Delta X)$ has:

$$
\begin{aligned}
&I(X+\Delta X) \\
&\quad = I(X) + I_x(X)\Delta x + I_y(X)\Delta y +, \cdots \\
&\quad \approx I(X) + \nabla I(X)(\Delta X)^t
\end{aligned} \tag{2}
$$

where $I_x{=}\partial I(X)/\partial x$, $I_y{=}\partial I(X)/\partial y$. Substituting the above approximation (2) into Eq.(1), then

$$
\begin{aligned}
&f(X) \\
&= \sum_w \big(\nabla I(X)(\Delta X)^t\big)^2 \\
&= \sum_w (\Delta X) \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} (\Delta X)^t \\
&= (\Delta X) \left(\sum_w \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \right) (\Delta X)^t \qquad (3) \\
&= (\Delta X)\big(\mathcal{Q}(X)\big)(\Delta X)^t \qquad (4)
\end{aligned}
$$

$I(X)$ is smoothed as $\mathcal{I}(X)$ by a Gaussian G. Then we write $\mathcal{Q}(X)$ as $\nabla I (\nabla I)^t$, build up a transform relation $\mathcal{H}(X)$ in a local window about X.

$$
\begin{aligned}
\mathcal{H}(X) &= \mathcal{T}(X) * \sum \left\{ \nabla I (\nabla I)^t \right\} \\
&= \mathcal{T}(X) * \sum \begin{Bmatrix} (G_x * \mathcal{I})^2 & (G_x * \mathcal{I})(G_y * \mathcal{I}) \\ (G_x * \mathcal{I})(G_y * \mathcal{I}) & (G_y * \mathcal{I})^2 \end{Bmatrix}
\end{aligned} \tag{5}
$$

where G is with standard deviation one, $G_x{=}\partial G/\partial x$, $G_y{=}\partial G/\partial y$. $\mathcal{T}(X)$ is a weight mask to weight the derivatives over the window. In (5), $\mathcal{I}{=}\mathcal{I}$(X), there are relations $\partial I/\partial x{=}\partial/\partial x{*}I$, $\partial/\partial x{*}(G{*}\mathcal{I}){=}(\partial/\partial x{*} G){*}\mathcal{I}{=}\partial G/\partial x{*}\mathcal{I}$. $\mathcal{H}(X)$ captures the local structure. The eigenvectors of this matrix are the principal curvatures of the auto-correlation function. We consider a cost function $\mathcal{M}$(X) :

$$
\mathcal{M}(X) \propto \mathcal{H}_\infty + \mathcal{H}_\in \tag{6}
$$

where $\mathcal{H}_\infty$, $\mathcal{H}_\in$ are the determinant and trace of $\mathcal{H}(X)$ respectively. We can get some image energy points, called interest points, by (5) and (6). An example of $\mathcal{M}(X)$ is illustrated in Fig.1 and Fig.2.

3. Probabilistic relaxation algorithm

Assume that there be two local windows $A'_m \subset I(X)$ and $A'_n \subset I(X+\Delta X)$. Let $A_m{=}\{\mathbf{x}_m\}$ be the set of all interest points in the first starting image that is input state space, and $A_n = \{\mathbf{y}_n\}$ the interest points in the second image that is output state space. Let $\mathbf{c}_{mn}$ be a vector connecting points $\mathbf{x}_m$ and $\mathbf{y}_n$ (thus $\mathbf{y}_n{=}\mathbf{x}_m{+}\mathbf{c}_{mn}$). Let the probability of correspondences of two points $\mathbf{x}_m$ and $\mathbf{y}_n$ be P_{mn}. Two

Fig.1: An original image.

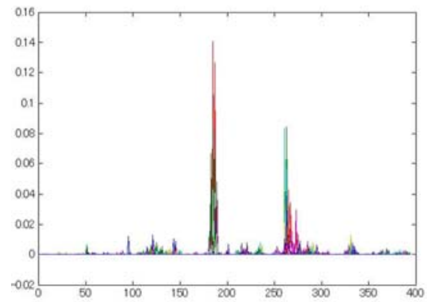

Fig.2: Distribution of the global correlation about Fig.1.

points $\mathbf{x}_m$ and $\mathbf{y}_n$ can be considered potentially corresponding if their distance satisfies the assumption of maximum velocity,

$$
|\mathbf{x}_m - \mathbf{y}_n| \le D_{max} \tag{7}
$$

where D_{max} is the maximum distance a point may move in the time interval between two consecutive images. Two correspondences of points $\mathbf{x}_m\mathbf{y}_n$ and $\mathbf{x}_k\mathbf{y}_l$ are termed consistent of

$$
|\mathbf{c}_{mn} - \mathbf{c}_{kl}| \le D_{dif} \tag{8}
$$

where D_{dif} is a pixel distance deviation. Consistency of corresponding point pairs will increases the probability that a correspondence pair is correct. We Determine the sets of interest points $A_m \subset A'_m \subset I_i(\mathbf{x})$, $A_n \subset A'_n \subset I_j(\mathbf{x})$, and construct a data structure as follows:

$$
[\mathbf{x}_m, (\mathbf{c}_{m_1}, P_{m_1}), (\mathbf{c}_{m_2}, P_{m_2}), \ldots, (NoV^*, NoP^*)]
$$

where P_{mn} is the probability of correspondence of points $\mathbf{x}_m$ and $\mathbf{y}_n$, NoV^*, and NoP^* are special symbols indicating that no potential correspondence was found.

We initialize the probabilities $P^{(0)}_{mn}$ of correspondence based on local similarity –if two points correspond, their neighborhood should correspond as well:

$$
P^{(0)}_{mn} = \frac{1}{1 + k_p \omega_{mn}} \left(1 - P^{(0)}_{(NoV^*, NoP^*)} \right) \tag{9}
$$

where $P^{(0)}_{(NoV^*,NoP^*)}$ is the initialized probability of no correspondence, k_p is a constant and

$$\omega_{mn} \;=\; \sum_{\mathbf{\Delta x}}[I_m(\mathbf{x}_m \pm \mathbf{\Delta x}) - I_n(\mathbf{y}_n \pm \mathbf{\Delta x})]^2 \tag{10}$$

$\mathbf{\Delta x}$ defines a neighborhood for image match testing – a neighborhood consists of all points $(\mathbf{x} + \mathbf{\Delta x})$, where $\mathbf{\Delta x}$ is defined as a symmetric neighborhood around $\mathbf{x}$.

We iteratively determine the probability of correspondence of a point $\mathbf{x}_m$ with all potential points $\mathbf{y}_n$ as a weighted sum of probabilities of correspondence of all consistent pairs $\mathbf{x}_k\mathbf{y}_l$, $\mathbf{x}_k$ are neighbors of $\mathbf{x}_m$ and the consistency of $\mathbf{x}_k\mathbf{y}_l$ is evaluated according to $\mathbf{x}_m, \mathbf{y}_n$. A quality q_{mn} of the correspondence pair is defined as

$$q^{(s-1)}_{mn} = \sum_k \sum_l P^{(s-1)}_{kl} \tag{11}$$

where s denotes an iteration step, k refers to all points $\mathbf{x}_k$ that are neighbors of $\mathbf{x}_m$, and l refers to all points $\mathbf{y}_l \subset A_n$ that form pairs $\mathbf{x}_k\mathbf{y}_l$ consistent with the pair $\mathbf{x}_m\mathbf{y}_n$.

The probabilities of correspondence are updated for each point pair $\mathbf{x}_m$, $\mathbf{y}_n$.

$$\hat{P}^{(s)}_{mn} = P^{(s-1)}_{mn}(k_a + k_b q^{(s-1)}_{mn}) \tag{12}$$

where k_a and k_b are preset constants. They deal with the convergent speed of P_{mn}. Normalize

$$P^{(s)}_{mn} = \frac{\hat{P}^{(s)}_{mn}}{\sum_j \hat{P}^{(s)}_{mj}} \tag{13}$$

Those interest points that hold high probabilities that obviously differ from those interest points without correspondences. Repeat (11) (12) and (13) until the $P^{(s)}_{mn}{>}P_{thr}$ (threshold) is found for all points $\mathbf{x}_m$, $\mathbf{y}_n$.

4. Experiments and discussions

To show availability of the presented approach, experiments are executed by matching image stream. A starting image is given in Fig.3. To illustrate the method easily, we have made an 11-points-based contour model as shown in Fig.3. Details how to make this contour can be found in [7][15][16], but beyond scope of this paper. As shown in Fig.4, one small window is selected in our experiment. Because the experiment is to match image stream, the second frame for matching should take into account this displacement with the respect to time t, and

Fig.3: Stating clip from image stream.

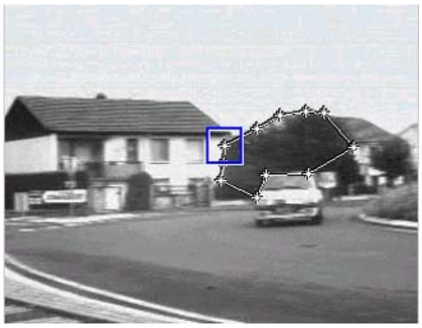

Fig.4: Local window selected.

this is the reason why Fig.6 gives the larger window in it. The frames should be referred to the same center. Comparing Fig.5 and Fig.6, we can find 8 pairwise relations are detected from A'_m and A'_n. Fig.7 and Fig.8, the images zoomed in, show us surprising good results.

From Fig.3 to Fig.8, the entire experimental images are in the form of gray. In the case of color image, color images can bring us more detail image energy distribution and less similarity than gray ones. So this method is also recommended to the color case.

5. Conclusions

This paper presents a new matching approach without camera calibrations. The matching principle is based on the probabilistic relaxation under local image energy constraints. This algorithm can find contributions to image classification [14] and object tracking [7][15][16]. We also recommend this technique to visual based intelligent navigation.

Acknowledgements

Part of this research was done while Xin Wang was at the Department of Electronics, Oita Institute of Technology.

Fig.5: Matching results of Fig.4.

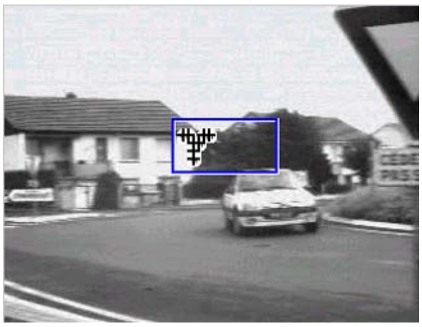

Fig.6: Matching results of the clip following Fig.4.

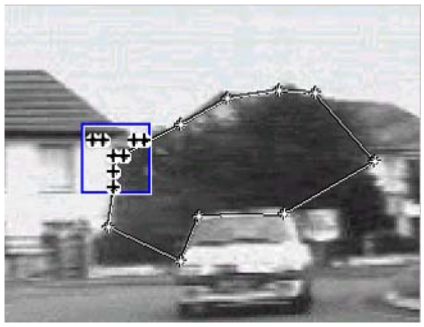

Fig.7: Fig.5 is zoomed in.

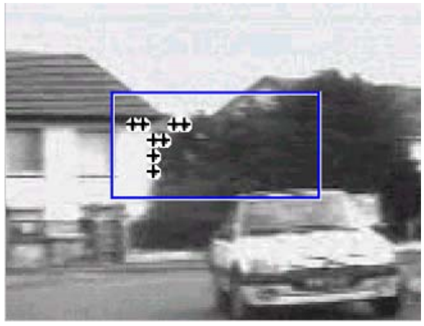

Fig.8: Fig.6 is zoomed in.

References

[1] D. Scharstien, "Stereo Matching with Nonlinear Diffusion," *International Journal of Computer Vision*, vol.28, no.2, pp. 155-174, 1998.

[2] S. B., Pollard, J. Mayhew, and J. P. Frisby, "Implementation details of the PMF stereo algorithm," *3D Model Recognition from Stereoscopic Cues, MIT Press, Cambridge*, MA, pp. 33-41, 1991.

[3] L. Robert, and O. Faugeras, "Curve-based ster -eo: Figural continuity and curvature," *Proc. CVPR'91*, pp. 57-62, 1991.

[4] A. Kumar, A. Tannenbaum, and G. Balas, "Optical flow: a curve evolution approach," *IEEE Trans. Image Processing*, vol.5, no.5, pp. 598-610, 1996.

[5] R. Green, D. Gray, and J. Powers, *Artificial Vision*, Academic Press, London, 1997.

[6] W. Yu, K. Daniilidis, and G. Sommer, "Approximate orientation steer ability based on angular Gaussians," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 193-205, 2001.

[7] X. Wang, M. Sugisaka, W. Xu, "Active contour model based tracking analysis" *Proc. of the 12th Int. Conference of Systems Engineering, UK,* vol, 3, pp. 347-352, 2003.

[8] D. A. Forsyth, J. Ponce, *Computer Vision*, Prentice Hill, USA, 2003.

[9] B. Jähne, *Image processing for scientific applications*, CRC Press, New York, 1997.

[10] M. Brown, D. Terzopoulos, *Real time computer vision*, Cambridge University Press, 1994.

[11] H. Zhang, S. Roth, it Camera aided robot calibration, CRC Press, Boca Raton, 1996.

[12] L. Torresani, C. Bregler, "Space-time tracking," *Computer Vision - ECCV 2002, Part I*, pp. 801-812, 2002.

[13] W. Kropatsch, R. Klette, F. Solina, and R. Albrecht, *Theoretical foundations of computer vision*, Springer Wien, New York, 1996.

[14] X. Wang, M. Sugisaka, W. Xu, "Recognizing static target in video frames taken from moving platform" *Proc. of the Int. Conference on Control, Automation, and System,* Korea, pp. 223-227, 2003.

[15] X. Wang, M. Sugisaka, W. Xu, "Image energy pilot, FA soil-salt recognition Approach," *The 4th Int. Conference on Advanced Mechatronics Toward Evolutionary Fusion of IT and Mechatronics*, pp. 373-376, 2004.

[16] X. Wang, M. Sugisaka, W. Xu, "Model based recognition for artificial brain," *Proc. of the 9th Int. Symposium on Artificial Life and Robotics,* vol, 2, pp. 605-609, 2004.