International Journal of Soft Computing 1(2) : 128-132, 2006

3-D Motion Estimation of Elastic Body from Monocular Image Sequenc Using MRF with Entropic Constraints

Yunhua Zhang, Yaming Wang and Wenqing Huang,
College of Informatics and Electronics, Zhejiang Sci-Tech University, 5 Second Avenue,
Xia-Sha Higher Education Zone, Hangzhou 310018, Peoples' Republic of China

Abstract: A novel approach to 3-D motion estimation of elastic body from monocular image sequence is proposed in this paper. First, with the establishment of feature point correspondence between consecutive image frames, the affine motion model and the central projection model are presented for local elastic motion. Then, in order to obtain the global motion parameters and overcome the ill-posed 3-D estimation problem, a framework of Markov Random Field (MRF) with entropic constraints is proposed. By incorporating the motion prior constraints into the MRF, the motion smoothness feature between local regions is reflected. This converts the ill-posed problem into a well-posed one and guarantees the robust solution. Experimental results from a sequence of synthetic image sequence demonstrate the feasibility of the proposed approach.

Key words: 3-D elastic motion, motion estimation, MRF, entropic constraints, image sequence

INTRODUCTION

The estimation of elastic motion from image sequence is one of the most studied problems within computer vision. Its importance stems from the wide applications it has in virtual reality, ventricle diagnoses and teleconferencing[1]. However, most of the efforts in this computer vision area have deal with rigid objects because its simplicity[1,2]. Our surrounding environment is generally not a rigid place. Therefore, more attentions should be paid to the problem of elastic motion estimation.

Most of the existing approaches for elastic motion estimation used dynamic shape models to provide the mechanism for fitting and tracking visual data. Two typical classes of shape models, namely, parametric models and physically-based models were used. A number of parametric models, such as splines[3], superquadrics[4] and Fourier descriptor[5], have been proposed for geometric shape representation. Although parametric models concisely capture the global shape of the objects, they are able to represent only a limited class of objects and are inadequate for the analysis and representation of complex, dynamic real-world objects[1]. The developed physically-based models for elastic motion estimation includes snakes[6], deformable superquadrics[7] and finite element[8]. Although physically-based models have some advantage over parametric models in modeling the dynamically deformable objects, the robustness of motion estimation using these models will suffer when our knowledge of these physical parameters is either vague or unknown.

In this study, a MRF approach is proposed for 3-D elastic motion from monocular image sequence. First, the affine motion model and the central projection model are proposed for local elastic motion. Then, the MRF with entropic constraints is proposed for global elastic motion estimation. In order to cope with the ill-posed 3-D motion estimation problem, the motion smoothness feature is incorporated into the MRF, thus the robust solution can be achieved.

MODELS FOR LOCAL ELASTIC MOTION

In this study, we choose affine motion model for local elastic motion, because affine motion model is a general non-rigid motion model and has more power in describing local non-rigid motion[9]. Consider the jth point $P_j^i = (X_j^i, y^{i+l}_j, Z_j^i)^T$ on the elastic body at time i moving to a point $P_j^{i+L} = (X_j^{i+l}, y_j^{i+l}, Z_j^{i+l})^T$ at time i+1 after a elastic motion, the motion between time i and time i+1 can be described as

$$\begin{bmatrix} X_j^{i+1} \\ Y_j^{i+1} \\ Z_j^{i+1} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X_j^i \\ Y_j^i \\ Z_j^i \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \quad (1)$$

Corresponding Author: Yaming Wang, College of Informatics and Electronics, Zhejiang Sci-Tech University, 5 Second Avenue, Xia-Sha Higher Education Zone, Hangzhou 310018, Peoples' Republic of China

where the motion parameters that need to be estimated are

$$w = (a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3, d_1, d_2, d_3)^T \quad (2)$$

At time I+1, the measurement model for a single point $P_j^{i+L} = (X_j^{i+1}, y_j^{i+1}, Z_j^{i+1})^T$ is given by

$$p_j^{i+1} = \Re(P_j^{i+1}) + n^{i+1} = \begin{bmatrix} x_j^{i+1} \\ y_j^{i+1} \end{bmatrix} + \begin{bmatrix} n_x^{i+1} \\ n_y^{i+1} \end{bmatrix} \quad (3)$$

where $\Re$ is imaging function, n is image plane noise,X_j^{i+1} and y_j^{i+1} denote the image plane coordinate, n_y^{i+1} and are image plane noise components. In the case of central projection, Eq. 3 can be reformulated as

$$p_j^{i+1} = \begin{bmatrix} \Re_x(P_j^{i+1}) + n_x^{i+1} \\ \Re_y(P_j^{i+1}) + n_y^{i+1} \end{bmatrix} = \begin{bmatrix} f\dfrac{X_j^{i+1}}{Z_j^{i+1}} + n_x^{i+1} \\ f\dfrac{Y_j^{i+1}}{Z_j^{i+1}} + n_y^{i+1} \end{bmatrix} \quad (4)$$

where $\Re_x$ and $\Re_y$ are the imaging function for image coordinate "x" and "y", respectively, f is the focal length.

MOTION ESTIMATION BASED ON MRF

It should point out that, the affine motion model defined by Eq. 1 is only suitable for local elastic motion. If we directly estimate the each local elastic motion, the motion edges between different local motions will not be smooth. This violates the motion inertia principle. Furthermore, the 3-D motion estimation is an inverse problemand most inverse problems are ill-posed. In order to overcome these problem, some motion smoothness prior constraints should be incorporate into the estimation process.

MRF and regularization are two typical approaches for encoding priors of constraints. MRF is more general than regularization in that it can encode not only the smoothness prior but also priors of other constraints[10]. MRF-based models have to be chosen among the two when the priors are due to those other than the smoothness, e.g. in texture modeling and analysis.

In this study, the input data consists of the two-dimensional image coordinates of each feature point. The correspondence of these feature points between consecutive frames is assumed to be finished. Due to the fact that feature points in an image frame are usually not on a lattice, we cope with the estimation problem by using a MRF with irregular sites to incorporate the motion constraints. A site of the MRF indexes a given correspondence of image feature points; A label represents an admissible elastic motion transformation.

Let the feature points in a frame be given as data, d = $\{p_i | i \in S\}$,where S = {1,2,...M}and M is the total number of feature point. We can define motion parameters on a MRF where a site is a feature point (x,y) and a label w(x,y) represents an admissible vector of motion parameters defined in (2).

Likelihood energy: The noise term n_x^{i+1} and n_j^{i+1} in Eq. 4 can be assumed to be independent, identically distributed (IID),$N(0,\sigma^2)$

$$\begin{aligned} n_x^{i+1} &= x_j^{i+1} - \Re_x(P_j^{i+1}) \sim N(0,\sigma^2) \\ n_y^{i+1} &= y_j^{i+1} - \Re_y(P_j^{i+1}) \sim N(0,\sigma^2) \end{aligned} \quad (5)$$

The likelihood energy can then be defined by

$$U(d|w) = \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\{\begin{matrix}\left(x_j^{i+1} - \Re_x(P_j^{i+1})\right)^2 \\ +\left(y_j^{i+1} - \Re_y(P_j^{i+1})\right)^2\end{matrix}\right\} \quad (6)$$

where M is the number of the feature points of a frame, N is the total number of the image frames.

Prior energy: According to the Markov-Gibbs equivalence[11], The joint probability can be written in terms of a Gibbs distribution, which is define as

$$P(w) = Q^{-1} \times e^{-\frac{1}{T}U(w)} \quad (7)$$

where Q Is a normalizing constant called the partition function, T is a constant called the temperature which can be assumed to be 1 and U (w) is the prior energy.

In order to derive the prior energy of the MRF with irregular site, we define the neighborhood system to comprise nearby sites within a radius of r

$$N_i = \left\{i' \in S \,|\, [\mathrm{dist}(p_i, p_i')]^2 \le r, i' \ne i\right\} \quad (8)$$

where Dist (p_i,$p_{i'}$) is the Euclidean distance between the feature points and r is selected according to the size of the object and local affine motion feature. We consider cliques of up to order two and so the clique set $C = C_1 \cup C_1$ $C_1 = \{\{i\} | \in S\}$ where is the set of single-site cliques and $C_2 = \{\{i,i'\} | i' \in Ni, \in S\}$ pair-site cliques. Then the prior energy can be defined as

$$U(w) = \sum_{i \in S} V_1(w_i) + \sum_{i \in S}\sum_{i' \in N_i} V_2(w_i w_{i'}) \quad (9)$$

The single-site potentials $V_1(w)$ may be used to force w_i to stay in the admissible set of motion parameters. The two-site potentials $V_2(w_i\ w_{i'})$ determine interactions between the individual w_i 's.

Single-site potentials based on second order entropy: In this study, because of the motion smoothness, motion parameters a_1, b_1 and c_3 are very close to 1. On the other hand,a_2,a_3,b_1,b_3, c_1 and c_2 are very close to 0. We use second order entropy (Ent-2) to reflect this characteristic.

The entropy of a discrete random variable is defined by[12]

$$H(x) = -\sum_{i=1}^{n} q_i \log q_i \text{ with } q_i = \frac{x_i}{\sum_{j=1}^{n} x_j} \qquad (10)$$

H (x) attains its global maximum when all q_i are the same, which corresponds to a uniform distribution with a value of $H_{max} = \log$. On the other hand, the lowest entropy level, $H_{mix} = 0$, is attained when all element q_i but one are set to zero.

A generalization of the above standard entropy has been defined as[13]

$$H_\alpha(x) = -\sum_{i=1}^{n} q_i \log q_i \text{ with } q_i = \frac{p_i}{\sum_{j=1}^{n} p_j} \qquad (11)$$

and

$$p = \Delta^\alpha x \qquad (12)$$

where a = 0,1,2... and Δ is a discrete difference operator. It has been shown that, for an appropriate choice of the

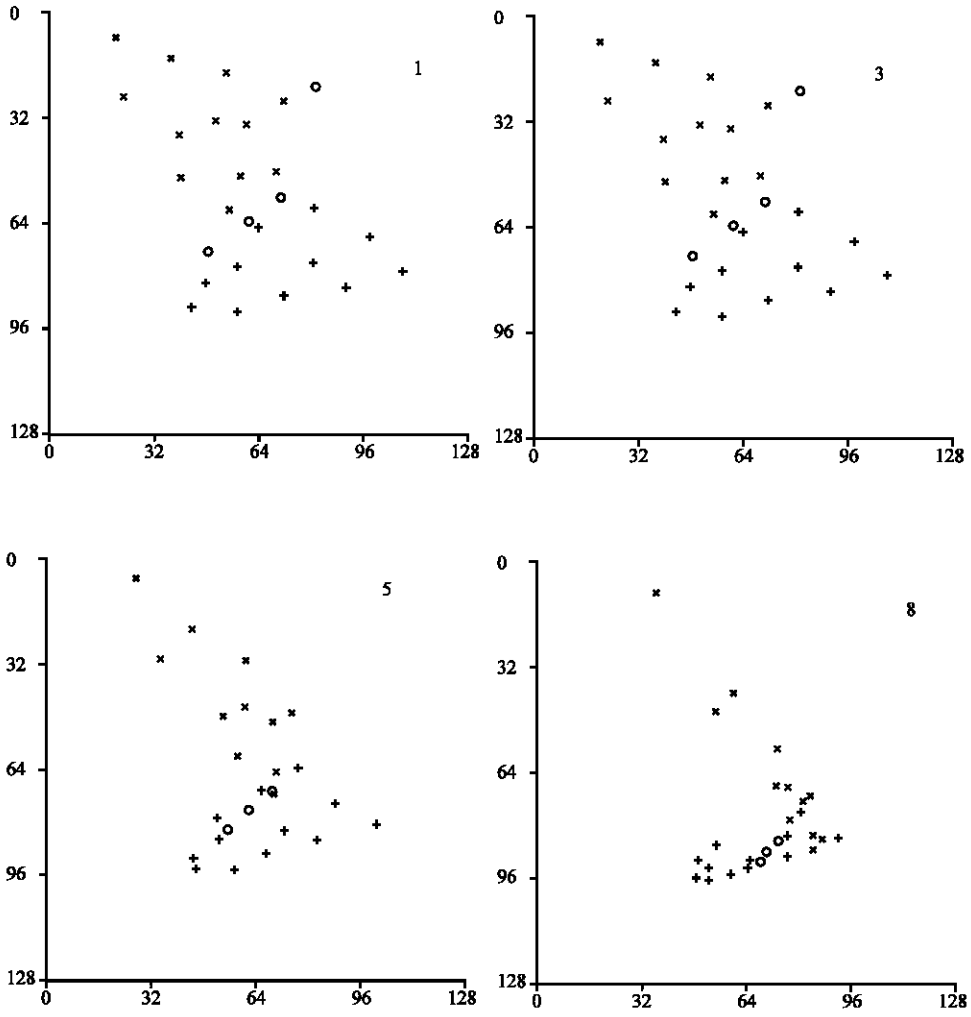

Fig. 1: The 1st, 3rd, 5th, 8th frames of a synthetic image sequence

Table1: Estimated motion parameters compared with given ones

Motion parameters	a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3	d_1	d_2	d_3
W_1	0.900000	0.100000	0.100000	0.100000	0.800000	0.100000	0.100000	0.100000	1.0000 00	0.0200 00	0.0200 00	0.0100 00
W^*_1	0.9002 3	0.1000 3	0.1000 2	0.1000 3	0.8002 1	0.1000 3	0.1000 3	0.1000 3	1.0003	0.0203 1	0.0210 6	0.0104 0
W_2	0.8000 0	0.2000 0	0.2000 0	-0.100000	0.9000 0	-0.100000	0.1000 0	-0.100000	0.9000 0	0.0200 0	0.02000 0	0.0100 0
W^*_2	0.8001 1	0.2000 3	0.2000 2	-0.090001	0.9001 1	-0.100011	0.1000 1	-0.100111	0.9001 2	0.0203 1	0.0196 5	0.0093 2
W_3	0.8500 0	0.1500 0	0.1500 0	0.0000 0	0.8500 0	0.0000 0	0.0000	0.0000 0	0.9500 0	0.0200 0	0.0200 0	0.0100 0
W^*_3	0.8541 5	0.1507 7	0.1507 6	-0.978995	0.8941 5	-0.965900	0.0205 0	0.0100 1	0.9545 5	0.0210 0	0.0231 1	0.0093 1

differential operator, an over-constrained approximation will still retain some of the information contained in the data.

Let x_1, a_1, x_2, b_2, $x_2 = c_3$, x_4 to x_9 equal to a_2, a_3, b_2, b_3 c_1 and c_2, respectively. Assuming that $x_{min} < x_1 < x_{max}, i = 1,2,...9$ and setting a = 2, we define the single-site potentials based on Ent-2 as

$$V_1(w) = -H_2(x) \tag{13}$$

here, the elements of vector p are given by

$$p_i = x_{i+1} - 2x_i + x_{i-1} + 2\left(x_{max} - x_{min}\right) + \theta i = 2,3,\cdots,8 \tag{14}$$

where is a small positive constant, it can be set to be 10^{-8}.

Two-site potentials: In order to reflect the motion smoothness between neighboring feature points, we define the two-site potentials as

$$V_2(w_i w_{i'}) = \min\left(\left\|w_i - w_{i'}\right\|^2, \alpha\right) \tag{15}$$

where a > 0 is a threshold, it is used to stop V_2 (w_i $w_{i'}$)'s increasing as $\| w_i w_{i'} \|$ becomes very large.

Minimization of posterior energy: When the likelihood energy and the prior energy $U(d|w)$ are available, the posterior energy follows immediately as $U(d|w) = U(w) + U(d|w)$ The optimal solution of motion parameters is

$$w^* = \arg\min_w U(w \mid d) \tag{16}$$

In order to find the optimal solution of motion parameters, the posterior energy must be minimized. There exist methods to solve such optimization problems. One of the most popular methods is the Simulated Annealing (SA)[14]. It is a global optimization method. Instead of performing gradient descent, a random search method is used to generate the next configuration. The random search is controlled by a sequence of decreasing temperatureand this enables the algorithm to escape from local minima into the global minimum.

Experimental results: This experiment involves a synthetic image sequence generated by using a given set of motion parameters. The sequence consists of 8 image framesand each frame contains 128×128 pixels and 26 feature points. Figure 1 shows the 1st, 3rd, 5th and 8th frames of the sequence. There are 26 feature points in each frame. The process of feature point correspondence is assumed to be finished beforehand. Three groups of feature points moving from frame to frame with three different types of local motion are marked as "×", "o" and "+", respectively. We use our approach to estimate the 3-D elastic motion parameters. The motion estimation results compared with the given parameters are shown in Table 1, where,w_1, w_2, w_3 are the given motion parameters for feature points "×", "+" and "o", respectively, w_i is the corresponding estimation results for w_i.

It is obvious that the estimation robustness for point group marked as "×" and "+" is good. The estimation error of the feature points "o" is relatively big, this is due to the fact that this point group contains only 3 feature points, the robustness of estimation is affected to some extend.

CONCLUSION

We have developed an approach to 3-D elastic motion estimation from monocular image sequence. By using the MRF framework with entropic constraints, the motion prior constraints is incorporated into the

estimation process, this guarantees robust solutions. Experimental results from a synthetic image sequence demonstrate the feasibility of the proposed approach. In addition, this method may offer significant improvement in cloud motion analysis, left ventricle motion analysisand other applications in computer vision.

However, our approach is computationally costly, this is the major problem in real applications. We will cope with this problem in our future research.

ACKNOWLEDGEMENTS

This study is supported by the National Science Foundation of China under Grant 60473038 and Zhejiang Provincial Natural Science Foundation of China under Grant RC02064.

REFERENCES

1. Aggarawal, J.K., Q. Cai and W. Liao *et al*., 1998. Nonrigid motion analysis: articulated and elastic motion. Computer Vision and Image Understanding, 70: 142-156.
2. Wang, Y. and G. Baciu, 2003. Human motion estimation from monocular image sequence based on cross-entropy regularization. Pattern Recognition Letters, 24: 315-325.
3. Bookstein, F.L, 1989. Thin-plate splines and the decomposition of deformations. IEEE Trans. PAMI, 11: 567-585.
4. Solina, F. and R. Bajcsy,1990. Recovery of Parametric models from range images: The case for superquadrics with global deformations. IEEE Trans. PAMI, 12: 131-147.
5. Staib, L.H. and J.S. Duncan, 1992. Boundary finding with parametrically deformable models. IEEE Trans. PAMI, 14: 1061-1075, 1992.
6. Ranganath, S., 1994. Contour extraction from cardiac MRI studies using snakes. IEEE Trans. Medical Imaging, 14: 275-286.
7. Terzopoulos, D. and D. Metaxas, 1991. Dynamic 3D models with local and global deformations: Deformable superquadrics. IEEE Trans. PAMI, 13: 703-714.
8. Pentland, A. and B. Horwitz, 1991. Recovery of non-rigid motion and structures. IEEE Trans. PAMI, 13: 730-742.
9. Zhou, L., C. Kambhamettu and D.B. Goldof, 1998. Structure and nonrigid motion analysis of satellite cloud images. ICVGIP'98, New Delhi, India, pp: 285-291.
10. Li, S.Z., 1995. Markov Random Field Modeling in Computer Vision, Springer-Verlag, London, UK.
11. Kindermann, R. and J.L. Snell, 1980. Markov Random Fields and Their Applications, Providence, R.I.: Am. Mathematical Society.
12. Shannon, C.E. and W. Weaver, 1949. The Mathematical Theory of Communication. Urbana, IL: Univ. Illinois Press.
13. Ramos, F.M., H.F.C.Velho, J.C. Carvalho and N.J. Ferreira, 1999. Novel approaches to entropic regularization. Inverse Problems, 15: 1139-1148.
14. Kirkpatrick, S., C.D. Gellat and M.P. Vecchi, 1983. Optimization by simulated annealing. Sci., 220: 671-680.