Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 51(65), Fascicola 2, 2006

Feature Based 2D Image Registration using Mean Shift Parameter Estimation

Daniela Fuiorea[1], Dan Pescaru[2], Vasile Gui[1], Corneliu I. Toma[1]

Abstract – A new method of feature based 2D image robust registration is proposed. The image distortion is modeled as a similarity transform with four parameters, estimated sequentially by 1D transforms, resulting in an increased sample density as compared to 4D space processing. By adopting a mean shift estimator, advantages of RANSAC and M-estimators can be combined within a single and sound theoretical framework. Experimental results confirm the validity of the proposed approach.

Keywords: image registration, robust estimation, mean shift, similarity transform

I. INTRODUCTION

Image registration is one of the basic image processing operations in many computer vision applications, like remote sensing, biomedical imaging, surveillance, robotics, multimedia etc [1]. The goal is to overlay two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. To register two images, a transformation must be found so that each point in one image (reference image) can be mapped to a point in the second (sensed image). In other words, the transform geometrically "optimally" aligns two images. Due to the diversity of images to be registered and to various type of degradations it is impossible to design a universal method applicable to all registration tasks. Every method should take into account not only the assumed type of geometric deformation between images but also radiometric deformations and noise corruption, required registration accuracy and application-dependent data characteristics.

Registration methods consist of the following four steps:

- Feature detection
- Feature matching
- Transform model estimation
- Image resampling and transformation

Features can be specific image points are image areas. The present study concentrates on the first approach.

Suppose the feature detection and matching problems have been solved by an appropriate automatic method. It is well known that the correspondence problem is difficult in the general case and prone to errors. Even a single gross correspondence error can drive the solution far away from the real one. Therefore robust estimation methods are needed to cope with point correspondence errors. One of the first robust estimators proposed for image registration was the RANSAC estimator [2]. Recently M estimators and related kernel based estimators received much attention in the community of researchers looking for robust solutions in computer vision [3]. The two methods have complementary merits. The M-estimators find good solutions but require a good initial estimate to converge correctly. RANSAC does not need to start from an initial estimate [4], but the solution does not take into account all the available data, thus its precision is not maximized. In the present work, a mean shift [5] based solution is proposed for robust parameter estimation in image registration. Like RANSAC, the mean shift estimator does not require an initial estimate. At the same time, as the (related) M-estimators, the mean shift estimator makes a better use of the available inlier samples.

II. BRIEF REVIEW OF THE MEAN SHIFT

Given a sample of N d-dimensional data points, $\mathbf{x}_i$, drawn from a distribution with multivariate probability density function $p(\mathbf{x})$, an estimate of this density at $\mathbf{x}$ can be written as [4]:

$$\hat{p}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} K_{\mathbf{H}}(\mathbf{x}-\mathbf{x}_i) \qquad (1)$$

where

[1] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail daniela.fuiorea@etc.upt.ro
[2] Facultatea de Automatizări şi Calculatoare, Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail dan@cs.utt.ro

BUPT

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K_{\mathbf{H}}(\mathbf{H}^{-1/2}\mathbf{x}) \quad (2)$$

is the kernel function depending on the symmetric positive definite $d \times d$ matrix **H**, called bandwidth matrix. Frequently **H** has a diagonal form or even the form **H**= h^2**I**, assuming the same scale h for all dimensions, i.e. a single scale parameter and an isotropic estimator, K_h. A radially symmetric estimator can be generated starting from a 1D kernel function K_1 as:

$$K^R(\mathbf{x}) = \alpha K_1(\|\mathbf{x}\|), \quad (3)$$

with α is a strictly pozitive constant chosen such that the kernel function integrates strictly to 1. The profile of the radially symmetric kernel is defined as:

$$K^R(\mathbf{x}) = c_{k,d}\, k(\|\mathbf{x}\|^2), \quad (4)$$

with $c_{k,d}$ a normalization constant.

Starting from any location **y**, a gradient ascent *mean shift* algorithm can be used to find the location of the maxima of the estimated PDF closest to the starting location. This can be simply done by iterating the equation

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2\right)}, \quad j = 1,2,... \quad (5)$$

where

$$g(x) = -k'(x) \quad (6)$$

until convergence. The proof of the convergence can be found in [4]. More, in practice the convergence is very fast, typically only two or three iterations being needed.

III. ROBUST IMAGE REGISTRATION

A widely used 2D geometric transformation in image registration is the similarity transform, consisting of rotation, translation and scaling. The model is defined by the equations:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} s & 1 \\ 1 & s \end{bmatrix} \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad (7)$$

relating the old pixel coordinates (x,y) to the new ones. As this transform preserves the angles and curvatures, it has been named “shape-preserving mapping”. The four parameters of the transformation can be unambiguously determined from the correspondence of two pairs of points. However, in most of the cases, the number of the points available for estimating the transformation parameters s, φ, t_x and t_y is higher. By denoting the vector of parameters as

$$\mathbf{p} = \begin{bmatrix} s \\ \varphi \\ t_x \\ t_y \end{bmatrix}, \quad (8)$$

82

the problem of estimating the geometric transformation can be formulated as the problem of minimizing a measure of the matching error of the available data points:

$$\mathbf{p} = \underbrace{\arg\min}_{\mathbf{p}} \sum_i \rho(r_i), \quad (9)$$

where the residuals r_i represent estimations of the matching error between a pair of corresponding features after registration. By choosing:

$$\rho(r) = r^2, \quad (10)$$

a least-squares matching is obtained. This problem has been extensively studied in early work on point matching. See for example the frequently cited work [6] with the improvements from [7]. Because of the squaring up in equation (10), least squares fitting is notoriously sensitive to the presence of the outliers, data samples deviating widely from “typical” samples. The problem can be alleviated by using a different shape of the function in equation (10), in order to reduce the influence of the outlier samples. M-estimators [3] are one of the most notorious examples from this category. In the present work, we use a density estimation approach, based on the mean shift to obtain robust estimates of the similarity transform. The approach is closely related to the M-estimator, as pointed out in [5]. However, the interpretation of the density estimator is different. Links can be found between mean density estimators and the RANSAC as well, but this is a subject beyond the scope of the present paper.

The number of corresponding points available in different applications varies widely. We concentrate on the case where this number is relatively small and obtaining reliable estimates in the presence of outliers is more difficult. A useful step in order to obtain higher sample densities is to reduce the dimension of the search space. In this paper we propose a solution based on search in 1D spaces, as opposed to the general approach of simultaneous estimation of all parameters in a 4D space. We start from the observation that angles between line segments are not changed by translation or rescaling. Therefore, the rotation parameter, φ can be estimated based on such

BUPT

angles prior to estimating the translation or rescaling parameters. Rescaling parameter estimation can also be done prior to translation or rotation estimation, based on distances between pairs of points. On the other hand, rotation or rescaling strongly affect translation parameters, as illustrated in the example in figure 1.

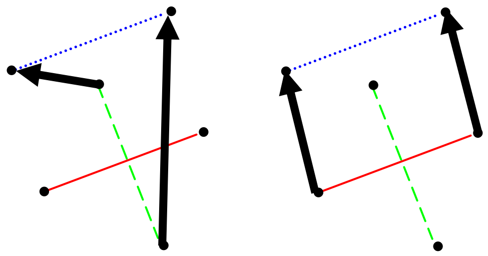

Fig. 1. Translation vectors before (left) and after (right) rotation compensation

From the example above, it is clear that to robustly estimate the translation vector components it has to be done *after* rotation and rescaling have been estimated and compensated.

Denote by $\{V_i\}$, $i = 1,2,\ldots, N$ a set of points from the reference image and $\{Q_i\}$, $i = 1,2,\ldots, N$ the corresponding points from the registered image. In the spirit of the RANSAC estimator, we form minimal sets of points, to estimate transform parameters. For rotation angle estimation, minimal set means pairs of points (Q_i, Q_j) and (V_i, V_j), with the corresponding vectors $\mathbf{q}_{ij}$ and $\mathbf{v}_{ij}$. The angle between the lines (Q_i, Q_j) and (V_i, V_j) is then given by the equation:

$$\cos(\varphi) = \frac{\mathbf{q}_{ij}^T \mathbf{v}_{ij}}{\left\|\mathbf{q}_{ij}^T\right\| \left\|\mathbf{v}_{ij}^T\right\|} \qquad (11)$$

Denote by $\{\varphi_i\}$, $i = 1,2,\ldots,M$ the set of M angles obtained by pairs of points. The rotation angle estimate is defined as the highest density location obtained by the mean shift algorithm starting from all data samples, φ_i. This is the mean shift filtered data set. Notice that the denominator in equation (5) is - up to a factor - a measure of the sample probability density estimated with the shadow kernel of $K()$. Therefore, no additional processing is needed to compute and compare probability densities. Moreover, very fast mean shift implementations can be obtained by marking all locations visited by the algorithm through iterations and associating corresponding intervals to the location of convergence.

In a similar manner, scale factor estimates can be obtained from the sets of pairs of points using the equation:

$$s = \|\mathbf{v}_{ij}\| / \|\mathbf{q}_{ij}\|. \qquad (12)$$

After performing the inverse geometrical transform to compensate for scale and rotation angle, robust translation vector component estimation is performed by point correspondences. Given a pair of points with position vectors $\mathbf{v}_i$ and $\mathbf{q}_i$, we form the translation data samples

$$\mathbf{t}_x = \mathbf{v}_{xi} - \mathbf{q}_{xi}, \qquad (13)$$
$$\mathbf{t}_y = \mathbf{v}_{yi} - \mathbf{q}_{yi}, \qquad (14)$$

then proceed to translation parameter estimation using mean shift.

IV. EXPERIMENTAL RESULTS

In order to obtain qualitative assessment of the proposed registration method, artificial image pairs have been generated with known geometrical transformation parameters. Feature points have been selected interactively in both images. For reference, a least squares estimator [6][7], was also used, mostly to validate the performances of the proposed approach for small errors, where the lest squares estimator works at its best. Results for the case of a similarity transform consisting of a translation and a 45^o rotation are shown in figure 2. The original image is shown in figure 2a, while the similarity transformed image is shown in figure 2b. In figure 2c, the results of the robust registration method proposed in this paper for image pairs from figure 2a and figure 2b are illustrated. The same results for the least-squares registration are illustrated in figure 2d. In figure 2e and figure 2f, the matching errors for the robust registration and for the lest square registration methods are displayed. Note that the errors were evaluated only within the minimum area rectangle enclosing the feature points used for registration and that the error images are displayed in negative contrast, for better visibility. A careful examination of the error images in figure 2e and figure 2f reveal the presence of significantly higher errors for the least squares estimator as compared with the robust mean shift based estimator, both in terms of translation and rotation parameters.

In order to obtain quantitative evaluation of the performances of the proposed image registration technique, in a second series of experiments, we generated 1D data sets with controlled percentage of outliers. Both inlier and outlier samples were generated as uniformly distributed random sequences. The outlier samples were generated with a standard deviation 10, while the inlier samples were generated with standard deviation 1. Comparative results for the mean shift and least square estimated parameters are given in figures 3 and 4, for 10% and respectively 33% outlier percentage. The mean shift estimator was implemented with the Epanechnikov kernel and scale parameter $h = 2$. As theoretically expected, the mean shift estimator errors are virtually unchanged and low,

BUPT

while the lest squares estimator errors are increasing with the outlier percentage.

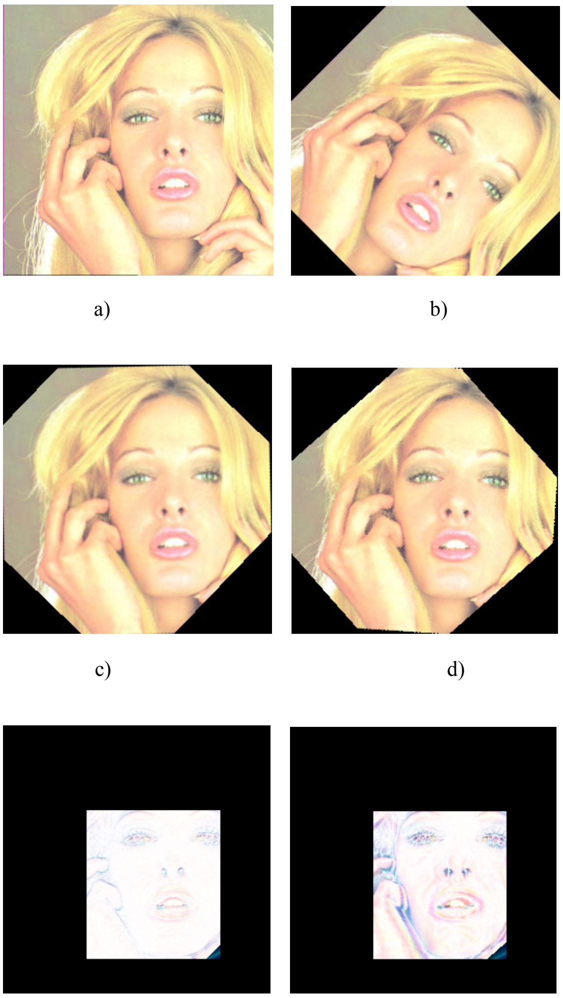

Fig. 2. a) Original image; b) Similarity transformed image; c) Mean shift registered image; d) Lest squares registered image; e) Matching error for the mean shift estimator; f) Matching error for the lest squares estimator.

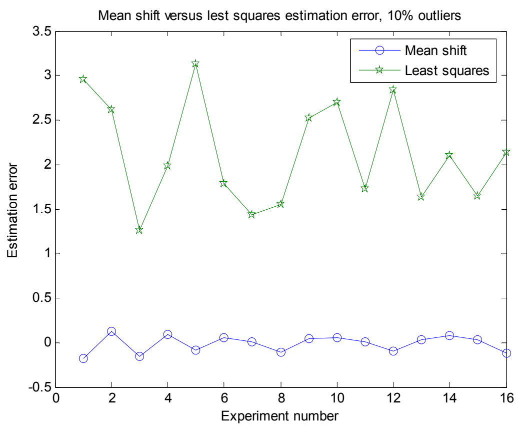

Fig.3. Mean shift versus least squares estimation errors, with outlier percentage 10%

Fig.4. Mean shift versus least squares estimation errors, with outlier percentage 33%

.

V. CONCLUSION

The mean shift based 2D image registration method proposed proved to be reliable and computationally efficient in our work. It can safely tolerate a high percentage of spurious data. Unlike for the RANSAC type robust estimators, the feature space is searched in a systematic and computationally efficient manner. No initial guess solution is needed as in the case of M-estimators. By a careful analysis, the search in a 4D space has been replaced by four 1D searches. This technique results in an increased sample density in the lower dimensional space, making kernel density estimation performances potentially less dependent on bandwidth selection.

REFERENCES

[1] B. Zitova, J. Flusser, “Image registration methods: a survey”, *Image Vision and Computing* 21*”,* Elsevier, 2003, pp. 977-1000.

[2] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography”. *Comm. Assoc. Comp.Mach*, 24(6):381–395, 1981.

[3] H. Chen, P. L. Meer, and D. E. Tyler, „Robust regression for data with multiple structures”. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1069–1075, 2001.

[4] V. Lepetit and P. Fua, “Monocular Model-Based 3D Tracking of Rigid Objects: A Survey”. In *Foundations and Trends in Computer Graphics and Vision* Vol. 1, No 1 (2005) 1–89.

[5] D.Comaniciu, P.Meer, “Mean Shift: A Robust Approach toward Feature Space Analysis”, *IEEE Trans. PAMI*, Vol.24, No.5, pp.603-619, 2002.

[6] K.S. Arun, T.S. Huang, S.D. Blostein: Least-squares fitting of two 3-D point sets. IEEE Transactions on Pattern Analysis and Machine Intelligence 9(5), 698-700, 1987.

[7] S. Umeyama: Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(4), 376-380, 1991.

BUPT