Techniques for Video Mosaicing

P.Saravanan, Narayanan .C.K., P.V.S.S Prakash, and Prabhakara Rao .G.V

***Abstract*—** Video Mosaicing is the stitching of selected frames of a video by estimating the camera motion between the frames and thereby registering successive frames of the video to arrive at the mosaic. Different techniques have been proposed in the literature for video mosaicing. Despite of the large number of papers dealing with techniques to generate mosaic, only a few authors have investigated conditions under which these techniques generate good estimate of motion parameters. In this paper, these techniques are studied under different videos, and the reasons for failures are found. We propose algorithms with incorporation of outlier removal algorithms for better estimation of motion parameters.

***Keywords*—** Motion parameters, Outlier removal algorithms, Registering , and Video Mosaicing.

I. Introduction

THE need to combine pictures into panoramic mosaics existed since the beginning of photography, as the camera's field of view is always smaller than the human field of view. Also, many times large objects could not be captured in a single picture as is the case in aerial photography. Using a wide field of view (fish-eye) lens can be a partial solution, but the images obtained with such a lens have substantial distortions, and capturing an entire scene with the limited resolution of a video camera compromises image quality. A more common solution is photo-mosaicing: aligning, and pasting, frames in a video sequence, which enables a more complete view. The most obvious applications of a mosaic representation are a means of visualization (since mosaics provide a wide and stabilized field of view) and video compression (since mosaics are efficient scene representations). However, mosaics are also useful in other applications, such as scene change detection, efficient video search and video indexing, efficient video manipulation.

Optical flow techniques are based on the idea that for most points in the image, neighboring points have approximately the same brightness [9]. If the overlap of the images is very large, (i.e. the motion is very small) it has been shown that a non-linear criterion minimization using the Levenberg--Marquardt method yields very good results [8], but it is very sensitive to the local minima and computationally expensive. Another class of techniques is based on extraction of features. Feature based methods rely on accurate detection of features.

Another strategy for video mosaicing is based on feature tracking. Feature point tracking process can be broadly divided in two steps. In the first step feature points are extracted in the first frame using a feature extractor, which are tracked in the successive frames in the second step, thereby establishing a correspondence between the frames. Thus using the feature map, the two frames can be registered. This method relies on the accuracy of the feature maps. Compared to the other gradient-based flow methods, this approach is computationally less expensive.

The rest of this paper is organized as follows. Section 2 gives an overall idea of feature tracking and optical flow based mosaicing, Section 3 deals with the problem of outlier detection and elimination. Section 4 presents the results obtained. Section 5 draws the conclusions.

Manuscript received March 15, 2005.
P.Saravanan is with the Sri Sathya Sai Institute Of Highier Learning, Puttaparthy, India (e-mail: saravananawaits@ yahoo.co.in).

II. Optical Flow And Feature Tracking

A. Optical Flow

Assuming a model of constant flow within a region of image we can combine information from neighboring gradient constraint equations to determine the best flow [*u,v*] satisfying all the equations by finding the [*u,v*] that minimizes the sum of the constraints over the neighborhood:

$$E_D(u,v) = \sum_{(x,y)\in R} \rho(I_x(x,y)u + I_y(x,y)v + I_t(x,y)),$$

where, $\rho(x) = x^2$ and R is some image region. More generally, one can assume a more complex flow model: **u**(x,y) = **u**(x,y,**a**) where **a** are the parameters of the model. Example models of image flow in a region R include constant, affine, and planar [8]. Our goal is to estimate the parameters, **a**, of the model within a region by minimizing:

$$E_D(a) = \sum_{(x,y)\in R} \rho(\nabla I^T u(a) + I_t),$$

In the constant case the model is simply the same as the equation above:

$$u(a) = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

For an affine flow model we have is:

$$u(x,y;a) = \begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix} = \begin{bmatrix} a_1 + a_2 x + a_3 y \\ a_4 + a_5 x + a_6 y \end{bmatrix}$$

Notice that when $\rho(x) = x^2$ this is a standard least squares regression.

B. Feature Tracking

Kanade Lucas Tomasi tracker with a translation or affine tracking model is used to track features in successive frames.

For translational model, *I(x, y, t+ ζ)* and *I(**x** - **d**)* represent the frames corresponding to time *t+ζ* and *t* respectively, and for affine tracking model *I(x, y, t+ ζ)* and *I(**Mx** - **d**)* represent the frames corresponding to time *t+ζ* and *t* respectively, where ***d*** is the displacement and ***M*** is a 2 x 2 matrix accounting for the affine warping. ***M*** can be written as ***M*** = ***1*** + ***D,*** with $\boldsymbol{D} = \left(\boldsymbol{d}_{ij}\right)$ a deformation matrix and ***1*** the identity matrix. The motion parameters are estimated by minimizing the sum of squared differences (SSD) [7]:

$$\varepsilon = \Sigma_w [I(\mathbf{Dx}+\mathbf{x}+\mathbf{d}, t+\zeta) - I(\mathbf{x}, t)]^2$$

Where $\boldsymbol{D} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$.

After the feature maps are obtained, a motion model is fit to them to compute the underlying transformations between the frames [5].

III. Outlier Detection

In the problem of mosaicing for only the background motion, considering pixels on a moving object or outliers is not useful. This intelligence is not added in [2] where a global least square minimization approach is adopted to estimate the affine parameters. Similarly due to the rather uniform nature of feature tracking [4], it is unavoidable that points which are not suitable for the purpose, might get tracked. These points may be from moving objects and outliers. To complicate things further, feature points may be lost due to occlusion or other factors. Methods for detecting and handling them are proposed in this section.

A. Optical Flow with Outlier detection

We use the robust estimation of multiple motions to detect outliers. In this approach we use the ρ to be Geman–McClure norm, $\rho(x,\sigma) = \dfrac{x^2}{\sigma + x^2}$ instead of least square minimization in . This approach can be used to estimate affine flow parameters for the multiple motions separately. Hence we can estimate the affine transformation parameters of background from its affine flow parameters.

B. Feature Tracking with Outlier detection

In this approach we consider the relative displacement of the feature points in both the images to arrive at the outliers. This is formulated as follows. Let the two sets of points tracked be ***Y*** and ***X.*** Determine the relative distance matrix ***D*** for the feature set ***X***, where $\left(\boldsymbol{d}_{ij}\right) = \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|$ $\boldsymbol{i}, \boldsymbol{j} \in \{1,2,...,\boldsymbol{k}\}$. Let ***D'*** represent the same relative distance for the feature set ***Y***.i.e.,

$$\left(\boldsymbol{d}_{ij}{}'\right) = \left\|\boldsymbol{y}_i - \boldsymbol{y}_j\right\|$$
$$\boldsymbol{i}, \boldsymbol{j} \in \{1,2,...,\boldsymbol{k}\}$$

Where, ***k*** represents the number of feature points tracked. If ($\boldsymbol{x}_i, \boldsymbol{y}_i$) and ($\boldsymbol{x}_j, \boldsymbol{y}_j$) are correct feature maps, then ideally $\boldsymbol{d}_{ij}$-$\boldsymbol{d}_{ij}$\` must be equal to 0. At the same time for incorrect feature maps this difference has a large non zero value. Figure 3.3 gives a better picture of this technique. Assume that ($\boldsymbol{x}_i, \boldsymbol{y}_i$), ($\boldsymbol{x}_j, \boldsymbol{y}_j$) and ($\boldsymbol{x}_k, \boldsymbol{y}_k$) are the pairs of tracked feature maps and among them ($\boldsymbol{x}_k, \boldsymbol{y}_k$) is an incorrect map. Since the other two feature maps are correct feature maps the distance $\boldsymbol{d}_{ij}$ is equal to $\boldsymbol{d}_{ij}$\`. But $\boldsymbol{d}_{ik}$ and $\boldsymbol{d}_{ik}$**'** as well as $\boldsymbol{d}_{jk}$ and $\boldsymbol{d}_{jk}$\` are not equal. Hence by finding the difference between the relative distance of $\boldsymbol{x}_k$ with respect to all other $\boldsymbol{x}_i$ and the relative distance of $\boldsymbol{y}_k$ with respect to all other $\boldsymbol{y}_i$, clearly marks ($\boldsymbol{x}_k, \boldsymbol{y}_k$) as an false map. This simple approach will also work in the presence of moving foreground objects. If points on the moving objects are tracked then, the relative distance of the point on the object with respect to the points in the background changes in the two images. This information can be easily captured by this technique. This scheme is reliable even when there is rotation between two frames.

IV. Results And Analisis

Two test sequences were used of different complexity. The aerial sequence (figure 3.1) has a simple translation motion. In this sequence we have the camera symbol '**+'** and the non overlapping region that should not be considered while extracting features or for registration in the process of mosaicing. The translation parameters obtained from tracking with outlier removal and optical flow with outlier removal are nearly equal since they incorporate outlier removal step. Optical Flow [2] performs poorly because of global minimization without any outlier removal process. Due to this limitation Mann's algorithm finds it difficult to generate mosaics when there are multiple motions. To substantiate this a road sequence (Figure 3.2 was used for comparison and it was found that Mann's algorithm could not give a reasonable homography. It is clearly evident that Optical Flow [1] is a better featureless algorithm over Optical Flow [2] algorithm for mosaicing videos with multiple motions.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:11, 2007

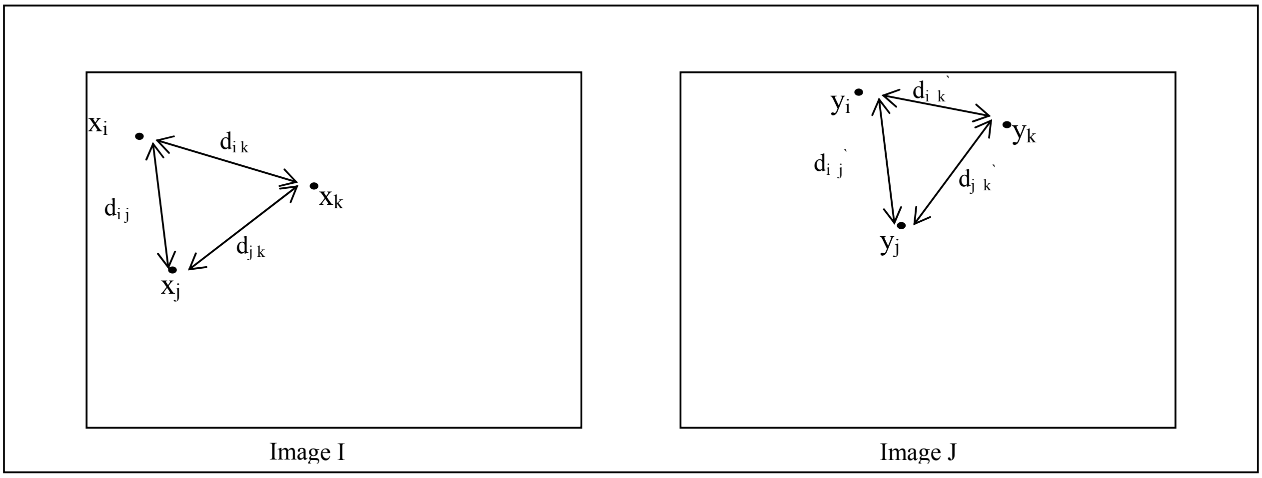

Fig. 3.1 Distance Metric

Algorithms	a_0	a_1	a_2	a_3	a_4	a_5
Tracking [3][4]	8.436678	1	0	2.169910	0	1
Optical Flow [2]	9.919314	1	0	2.060935	0	1
Optical Flow [1]	8.407109	1	0	2.180797	0	1

Fig. 3.1a 0^{th} frame in a aerial sequence

Fig. 3.1b 1^{st} frame in a aerial sequence

Fig. 3.2a 0^{th} frame in a road sequence

Fig. 3.2b 1^{st} frame in a road sequence

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:11, 2007

A few of the frames of an aerial video sequence (300 frames)

The mosaic obtained

REFERENCES

[1] Micheal.J.Black and P Anandan "The Robust Estimation Of Multiple Motions: Parametric and Piecewise Smooth Flow Fields" CVIU Vol 63 pp 77-104, 1996.

[2] S. Mann, and R. Picard, "*Video Orbits of the Projective Group: A Simple Approach to Featureless Estimation of Parameters*", IEEE Trans. on Image Processing, Vol. 6, No. 9, pg. 1281-95, 1996.

[3] Bruce D.Lucas and Takeo Kanade, "*An iterative Image registration Technique with an application to stereo vision*", from Proceedings of Image Understanding workshop, pg.121-130, 1981.

[4] J.Shi and Carlo Tomasi, "*Good Features to Track*", IEEE Conference on Computer Vision and Pattern Recognition (CVPR94), Seattle, June 1994.

[5] Shinji Umeyama, "Least-Squares Estimation of Transformation Parameters Between Two Point Patterns", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 13, No. 4, April, 1991.

[6] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In G. Sandini, editor, *Proc. of Second European Conference on Computer Vision, ECCV-92*, volume 588 of *LNCS-Series*, pages 237–252. Springer-Verlag, May 1992.

[7] P.J.Burt, C.Yen, and X.Xu, "Local Correlation Measures for Motion Analysis: A Comparative Study." ,IEEE CPRIP, pg.269-274,1982

[8] R. Szeliski, "Image Mosaicing for Tele-Reality Applications", Technical Report CRL 94/2, Digital Equipment Corporation, 1994.

[9] B. Horn, and B. Schunck, "Determining Optical Flow", Artificial Intelligence, Vol. 17, pg. 185-204, August 1981.