Genetic-based Image Mosaicing+

Tao FENG , Changbo HU and Songde MA

National Laboratory of Pattern Recognition, Chinese Academy of Science, P.O.Box 2728, Beijing, China

Email: {fengt, cbhu, masd}@nlpr.ia.ac.cn

Abstract:*An automatic image mosaicing technique that generates panorama images from images sequence or video is described in this paper. The approach is based on genetic algorithms (GAs) to match feature-points and to estimate homography parameter simultaneously. There are three main points about our approach: (1) an automated method for image mosaic and feature correspondence; (2) robust estimation of parameters based on genetic algorithm and MLE refining under certain feature position noise; (3) compensation of global illumination change. We described our approach for image mosaicing in detail. Experiment results on real image are presented.*

KeyWords: *image mosaicing, genetic algorithm, homography*

1. Introduction

Image mosaicing is to generate a uniform panorama image from image sequence or video of static 3D scene. The premise of image mosaicing is inter-image global transformations such as affine transformation and homography (planar perspective transformation). A number of techniques have been developed for image mosaicing ([1], [2]). We developed an approach for corner correspondence and homography estimation simultaneously and automatically. The overall flowchart is in Figure 1. First, SUSAN ([3]) corner detector is applied to detect corner points automatically. Second, we setup initial corner point matches by computing local intensity correlation. There are many outliers in set of the initial matches that can not be modeled as Gaussian. GAs is applied to the data set to estimate homography and find out the good matches. The estimate of homography is refined by MLE under the assumption of independent, isotropic Gaussian noise on corner point position. Finally, after illumination compensation, merged panorama is outputted.

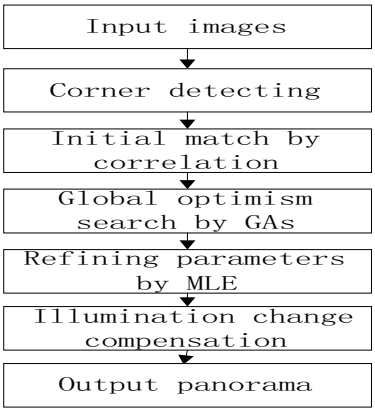


Fig.1.Flowchart of genetic-based image mosaicing

2.Homography Matrix Estimation

2.1. Homography

Given two images taken from the same viewpoint but in different directions, it is well known that the mapping between these two images is completely specified by homography or so-called planar perspective transformation. Two corresponding points are related by an equation of the form

$$\mathbf{x}' = \lambda \cdot H \cdot \mathbf{x} = \lambda \cdot \begin{pmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{pmatrix} \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \qquad (1)$$

where $\mathbf{x} = (u, v, 1)^T$ and $\mathbf{x}' = (u', v', 1)^T$ are projective coordinates (or homogeneous coordinates) of point pairs which are images of the same point in 3D space. λ is an arbitrary non-zero scalar. H is the homography matrix. The Degrees of Freedom (DOF) of matrix H is 8 up to scale.

From (1) we know each matched point pair provides two linear constraints on elements of homography matrix H. They are

$$\begin{pmatrix} u & v & 1 & 0 & 0 & 0 & -u\cdot u' & -v\cdot u' & -u' \\ 0 & 0 & 0 & u & v & 1 & -u\cdot v' & -v\cdot v' & -v' \end{pmatrix} \cdot \mathbf{h} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (2)$$

where $\mathbf{h} = (h_{00} \;\; h_{01} \;\; h_{02} \;\; h_{10} \;\; h_{11} \;\; h_{12} \;\; h_{20} \;\; h_{21} \;\; h_{22})^T$.

The group of homography matrices is called the planar projective group. Members of the group have 8 DOF. For n point correspondences we obtain $2 \times n$ equations in 8 unknowns. In general, H matrix can be computed linearly by a minimal subset of four pairs of matched points without any three of them are co-linear.

+ This work is funded by research grants from the NSFC （No.69865005） and the 973 research project（G1998030500）

Equation (2) for 4 point correspondences can be written as

$$A \cdot h = 0 \qquad (3)$$

The vector h that minimizes algebraic residuals $\|A \cdot h\|$, subject to $\|h\| = 1$, is given by the eigenvector of least eigenvalue of $A^T \cdot A$. The eigenvectors can be obtained directly from the SVD of matrix A. h is the null-vector of A.

2.2. Initialize corner points matching

Let $I_0(u,v)$ and $I_1(u,v)$ be any two input images, Cl and Cr be the corner points sets detected by SUSAN corner detector correspondingly. We can define the cross-correlation of two corner points $x = (u,v) \in I_0$ and $x' = (u',v') \in I_1$ as

$$S(u,v,u',v') = \frac{\sum_{i=-w/2}^{w/2}\sum_{j=-w/2}^{w/2}\{I_0(u+i,v+j)\cdot I_1(u'+i,v'+j)\}}{\sqrt{\sum_{i=-w/2}^{w/2}\sum_{j=-w/2}^{w/2} I_0(u+i,v+j)^2}\cdot\sqrt{\sum_{i=-w/2}^{w/2}\sum_{j=-w/2}^{w/2} I_1(u+i,v+j)^2}} \qquad (4)$$

where $I_0(u,v)$ and $I_1(u,v)$ is the intensity of two images at given point (u,v). $S(u,v,u',v') \in [0,1]$ is the value of cross-correlation within a $w \times w$ window. Any two points with the cross-correlation value above the given threshold S_t is considered as potential match.

3. Feature Correspondence by Genetic Algorithms

Genetic algorithms ([4]) are a stochastic optimization technique. GAs can be considered as a multi-point search, the result is less likely to be fall into a local optimal solution. The initial set of candidate point pairs contains many outliers and the search space has many local optical solutions. We developed an approach to apply GAs to search global optimal solution under homography assumption. The application of GAs is carried out in following steps:

Initialization: Initialize the values for parameters N (size of chromosome pool), P_{c0} (initial crossover rate) and P_{m0} (initial mutation rate). Takes potential point correspondences as genes and create N initial chromosomes Randomly from the genes.

Step 1: Evaluate the fitness of each chromosomes in the chromosome pool.

Step 2: Compute P_c and P_m with different statistics attributes of chromosome pool. Selects chromosomes to perform crossover and mutation operator from chromosome pool according to the fitness, the crossover ratio and mutation ratio.

Step 3: Crossover and mutation operators are applied to the selected chromosomes.

Step 4: New chromosomes are evaluated by their fitness function value.

Step 5: New chromosomes are inserted into chromosome pool and the bad chromosomes in fitness are eliminated through selection.

Step 6: If the chromosome is not convergence then goto Step 2, otherwise output the best chromosome and the optimum solution.

3.1 Chromosome Coding

We use potential point correspondences as genes. Any gene can be described as a 4-vector $g = (\mathbf{u_0}, \mathbf{u_1})$ with $\mathbf{u_0} = (u_0, v_0)$ and $\mathbf{u_1} = (u_1, v_1)$ is inhomogeneous coordinates of corresponding corner points. A set of 4 genes, which is the minimal subset for estimating homography, is taken as a chromosome. Let G be the set of genes. A Chromosome is arrayed as $g_0 g_1 g_2 g_3$, where g_0, g_1, g_2, $g_3 \in G$

3.2 Evaluation of Chromosome

The position error of corner detecting is assumed to be Gaussian. That is

$$p(\mathrm{u} \mid \mathrm{u}_0) = \frac{1}{2\pi}\|\Sigma\|^{-1/2}\exp\left[-\tfrac{1}{2}(\mathrm{u}-\mathrm{u}_0)^T \Sigma^{-1}(\mathrm{u}-\mathrm{u}_0)\right] \qquad (5)$$

where $\mathbf{u} = (u \quad v)^T$ and $\mathbf{u}_0 = (u_0 \quad v_0)^T$ is the position with noise and the exact position of the corner. $\Sigma = \begin{pmatrix} \sigma_u & \sigma_{uv} \\ \sigma_{uv} & \sigma_v \end{pmatrix}$ is the correlation matrix, which is additionally assumed to be isotropic distribution ($\sigma_u = \sigma_v = \sigma$) and independent ($\sigma_{uv} = 0$). The value of σ is decided by the corner detector.

To estimate the variance of mapped image coordinates, we expand (2) in a Taylor series and truncate it to first order so that the transformation is linearised. The variance of mapped image coordinate $\mathbf{u_0}' = (u_0', v_0')^T$ is:

$$\Sigma' = Var(\mathbf{u_0}') = \frac{\partial \mathbf{u_0}'^T}{\partial \mathbf{u_0}} \cdot \Sigma \cdot \frac{\partial \mathbf{u_0}'}{\partial \mathbf{u_0}} \qquad (6)$$

where $\mathbf{u_0} = (u_0, v_0)^T$ is the detected corner point coordinate, Σ' denotes the variance of linearly mapped point coordinate.

After H matrix is estimated for the given chromosome, to test whether a gene satisfies the homography the difference between the mapped coordinate of one point and the other point must obey the condition:

$$(\mathbf{u_1} - \mathbf{u_0}')^T \Sigma'^{-1} (\mathbf{u_1} - \mathbf{u_0}') < 2 \qquad (7)$$

where $\mathbf{u_0}'$ denotes the coordinate of mapped first point, $\mathbf{u_1}$ denotes the coordinate of second point in the correspondence. This condition means that the gene that

satisfies the homography will pass the test with the probability $P > 0.9$.

Each chromosome is evaluated by a fitness function. For any given chromosome, we can compute a homography matrix H. The fitness function is defined as fellow,

$$f(chromosome) = \sum_{i=1}^{N} match(g_i, H)$$
$$match(g_i, H) = \begin{cases} 1 & , \quad (\mathrm{u}_{i1} - \mathrm{u}_{i0}')^T \Sigma'^{-1} (\mathrm{u}_{i1} - \mathrm{u}_{i0}') < 2 \\ 0 & , \quad others \end{cases} \tag{8}$$

Where $g_i \in G, i \in [1, n]$ and H is estimated linearly by the chromosome using (3), $\mathbf{u_{i1}}$ denotes the image coordinate of second point of g_i, $\mathbf{u_{i0}}'$ denotes the mapped coordinate of first point of g_i.

3.3. Refining homography

The estimate homography is refined using MLE (Maxium-Likehood Estimate) under the Gaussian noise assumption of corner point position. For isotropic, independent Gaussian noise (3) can be written as

$$p(\mathbf{u_0}' | \mathbf{u}_1) = \frac{1}{2\pi \cdot \sigma^2} \exp\left[- \frac{\|\mathbf{u_0}' - \mathbf{u_1}\|^2}{2 \cdot \sigma^2} \right] \tag{9}$$

where σ is the standard deviation, $\mathbf{u_0}' = (u_0' \quad v_0')^T$ is the mapped position of $\mathbf{u_0} = (u_0 \quad v_0)^T$ by (2), $\mathbf{u_0}$ and $\mathbf{u_1} = (u_1 \quad v_1)^T$ is the correct point correspondence . $\|\mathbf{u_1} - \mathbf{u_0}\|$ is the Euclidean distance between $\mathbf{u_1}$ and $\mathbf{u_0}$ .The MLE of H matrix is equivalent to minimize the sum of the distance between all correct correspondences. That is

$$\mathbf{min}\left\{ \sum_{i=1}^{M} \|\mathbf{u_0}' - \mathbf{u_1}\|^2 \right\} \tag{10}$$

where $\mathbf{u_0}'$ and $\mathbf{u_0}$ satisfy (2).

We use Levenberg-Marquardt algorithm to search the solution to (8).

4. Image Mosaicing with Illumination Adjustment

In this section we describe the method we used to merge images into one image and compensation the global illumination change.

First we taken any one of the input images as the reference image. Then other images are warped into the reference image coordinate with (2).

We model this change as a linear transformation with

$$I_i(u \quad v) = a_i \cdot I_0(u \quad v) + b_i \tag{11}$$

constant coefficients a_i and b_i can be estimated by

$$a_i = \frac{E\{(I_i(u,v) - E\{I_i(u,v)\})^2\}}{E\{(I_0(u,v) - E\{I_0(u,v)\})^2\}} \text{ with } (u \quad v) \in D \tag{12}$$
$$b_i = E\{I_i(u,v)\} - a_i \cdot E\{I_0(u,v)\}$$

where $E\{I_0(u \quad v)\}$ is the mathematical expectation of $I_0(u \quad v)$, D is the overlapped area between two images. At a certain position $(u \quad v)$ there are different intensities for different warped image. We use the average value as the intensity of the merged image.

5. Experimental Results

We present an experiment on image mosaicing under pure rotation of the camera in Figure 2. Two of original images are shown in Fig 2(a). There is significant illumination change between two images. In Fig 2(b) initial correspondences is superposed on the intensity images of original images. There are much wrong correspondences (outliers), about 60%. The good matched point pairs as the result of GAs are superposed on the intensity images shown in Fig.2(c). Fig 2(d) is the merged image with homography refined by MLE and compensation to global illumination change. There is no significant illumination difference in Fig 2(d). Fig.2 (e) shows the convergence curve of the chromosome pool.

6. Conclusion

An automatic genetic-based image mosaicing approach is discussed in this paper. Experiments on real image demonstrate that our approach is effective and robust for image mosaicing under homography model. The main advantages of our approach are

- It needs fewer images than the intensity-based approaches.
- Our method is robust to outliers.

Image-based rendering and walking through 3D scene are our future works based on this paper.

REFERENCES

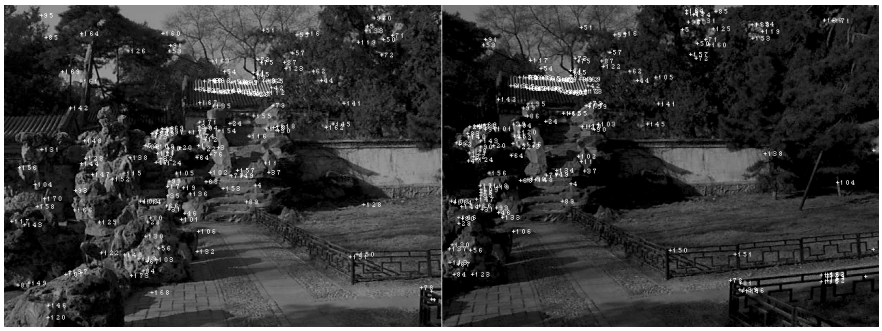
[1] R.Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and texture-mapped models. In *Computer Graphics Proceedings, Annual Conference Series*, pages 251--258, Proc. SIGGRAPH'97 (Los Angeles), August 1997. ACM SIGGRAPH.

[2] S.Chen. QuickTime VR – an image-based approach to virtual environment navigation. Computer Graphics (SIGGRAPH'95), pages 29-38,August 1995.

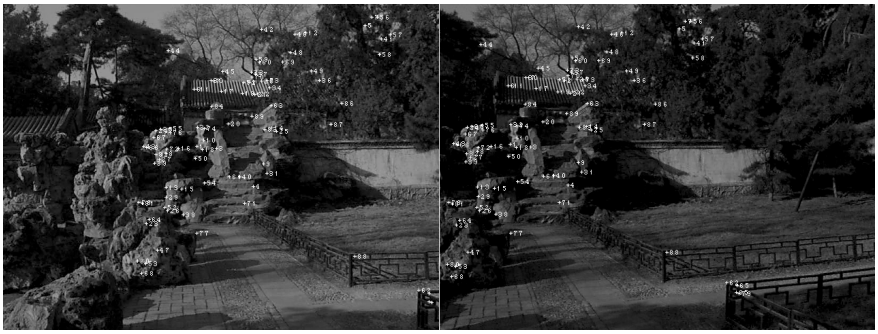
[3] S.M. Smith and J.M. Brady. SUSAN - a new approach to low level image processing. *Int. Journal of Computer Vision*, 23(1):45--78, May 1997.

[4] Holland, J. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor.1975

(a) Original images

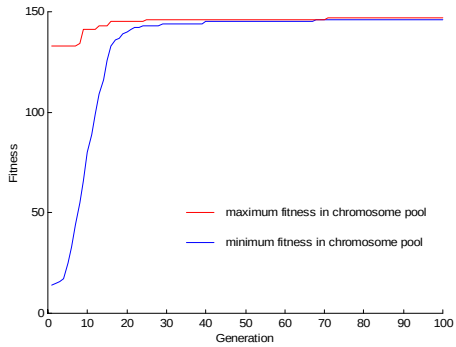


(b) Initial matched point pairs superposed on the original images



(c) Good matched point pairs selected by GAs

(d) Combined image warped onto a cylinder



(e) Convergence curve of the chromosome pool

Figure 2. Image mosaicing for a pair of images taken in the Summer Palace