Image Rectification Using Affine Epipolar Geometric Constraint

Sui Liansheng
Xi'an University of Technology /School of Computer Science and Engineering, Xi'an, China
Email: slsitrd@hotmail.com

Zhang Jiulong and Cui Duwu
Xi'an University of Technology /School of Computer Science and Engineering, Xi'an, China
Email: {chinajiulong, cuidw}@xaut.edu.cn

Abstract—**To rapidly and accurately search the corresponding points along scan-lines, rectification of stereo pairs are performed so that corresponding epipolar lines are parallel to the horizontal scan-lines and the difference in vertical direction is zero. In this paper, the method to rectify image pairs could be divided into three steps including projective transformation, affine transformation and shearing transformation. The projective transformation matrix is computed under the affine epipolar geometry constraint, and the values of unknown parameters are searched by an algorithm which does not require the relative matrix be positive definite. In this paper, an optimization function is presented to remove the difference in vertical direction and an algorithm is developed to estimate initial values of some parameters such as scale weights and vertical offset.**

Index Terms—**image rectification, affine epipolar geometric constraint, epipolar geometry**

I. Introduction

In stereo computer vision system, stereo image pairs are two images taken from two distinct viewpoints for the same scene. The projection of a point in the scene onto the left and right mages constitute corresponding points, the disparity of which can be used to calculate the depth of space point. There exists epipolar geometry constraint in the stereo image pairs, i.e. for a given point in left image, we can search for its corresponding points in right image along an epipolar line. When the position of camera only differ horizontally and optical axis are parallel, the corresponding epipolar line from the two images will lie on the same horizontal scan-line, and we only need to search for corresponding points along the scan-line. In this case, we can avoid the computation of epipolar line and eliminate the ambiguity of matching, which will then accelerates the searching process. Furthermore, the algorithms for 3D constructions and motion analysis are based on image sequence, which are valid only in ideal case. So it is necessary to rectify the stereo images in advance.

The process of image rectification is as follows: projective transformation is performed on the left and right images; the image is then projected onto a common plane parallel to the baseline of the camera, where the epipolar lines of the image plane lie on the same scan line, and there is no vertical difference between the corresponding epipolar lines. Different projective transform can rectify the image onto the different space plane, with different rate of distortion. So projective transform should reduce the distortion as far as possible after the image rectification and avoid over-sampling and sub-sampling [1]. Over-sampling occurs when new pixels are produced because of the enlarged image scale, whereas sub-sampling is caused when some pixels are lost due to reduced image scale.

The rectification of stereo image pairs can be performed under the condition of calibrated camera [2], but generally the rectification is under the uncalibrated condition, which has become an important research field of stereovision. Therefore, scholars all over the world present various algorithms for rectifying epipolar line. Hartley [3] determines the projective matrix through the constraint that disparity between the corresponding points is minimum. Francesco [4] presents an algorithm of image rectification without computation of fundamental matrix and only dependent on coordinates of corresponding points, but the initial value computed by nonlinear optimization in the rectification lacks credit. Al-Zahrani [5] defines a reference plane by using arbitrary three groups of corresponding points in the stereo image pairs to determine the projective matrix, where the distortion after the rectification directly depends on the selection of three groups of corresponding points. Hsien-Huang P.Wu [6] determines and optimizes some parameters of projective matrix by minimizing the square sum of distance between image point and corresponding epipolar line, and estimates other parameters by shearing transform. Charles Loop [7] divides the transformation matrix into a projective and

Sui Liansheng is with the school of Computer Science and Engineering, Xi'an University of Technology, 5th Block, Jinhua South Road, XI'an City, Shannxi Province, China.

affine component, and effectively reduce the distortion after image transform during the process of projective transformation, which is a high precise rectification algorithm of stereo image pairs. But the involving matrix must be positive-definite so as to compute projective matrix. Nevertheless these matrices may not be positive-definite because of the noise disturbance on corresponding image points. Therefore we cannot compute projective matrix and thus the rectification process fails.

The paper presents a multi-stage effective rectification algorithm of stereo image pairs under the uncalibrated condition, which is based on the idea of multi-stage rectification presented by Charles Loop [7]. In the rectification process of image pairs, we use the affine epipolar geometry constraint to compute the projective transformation matrix of image pairs, thus making all epipolar lines parallel to each other in the image pairs. We then use the constraint of minimized vertical difference of corresponding epipolar lines to compute proportional factor and vertical offset, which completes the optimization process of affine transformation matrix. This algorithm does not require the involving matrix be positive-definite when computing the projective matrix, and ensure the completion of rectification given arbitrary corresponding image points. This algorithm also describes the calculation process of initial value of optimal parameters when optimizing the affine transform matrix.

The paper is structured as follows. In Section II we interpret the epipolar geometry and relative properties of stereovision system, and describe the geometric meaning of projective transformation matrix used to rectify image pairs. In Section III we divide the process of image rectification into three stages that the projective transformation matrix, affine transformation matrix and shearing transformation matrix are computed respectively, and discuss the ability that given arbitrary corresponding image points which may be destroyed by noise enormously, the projective transformation can be computed with the affine epipolar geometry constraint. In Section IV we present experiments on the representative image pairs, and analyze the results. In Section V we make concluding remarks and hint the future work.

II. Epipolar Geometric Relation of Stereo Image Pairs

A. Epipolar geometry

Epipolar geometry constraint has an important role in the stereovision system [8]. Let m_l and m_r be the projections of a 3D point M in images I_l and I_r, and F is a fundamental matrix of rank 2, then m_l and m_r satisfy the epipolar constraint equation

$$m_r^T F m_l = 0 . \tag{1}$$

Let $l_r = Fm_l$, l_r is an epipolar line corresponding to m_l in the image I_r, then m_r must lie on l_r; Let $l_l = F^T m_r$, l_l is an epipolar line corresponding to m_r in the image I_l, then m_l must lie on the l_l. All epipolar lines in I_l pass through the epipole e_l, all epipolar lines in I_r pass through the epipole e_r. The points e_l and e_r satisfy the following equation

$$Fe_l = 0, F^T e_r = 0 .$$

All epipolar lines in image I_l and I_r are parallel to horizontal direction; besides, they are parallel among each other. The corresponding epipolar lines lie on the same scan-line in I_l and I_r, and there is no vertical difference between the corresponding epipolar lines. The epipoles e_l and e_r of images I_l and I_r are at infinite point in the horizontal direction, which is denoted as

$$e_l = e_r = (1 \quad 0 \quad 0)^T .$$

The fundamental matrix for a rectified image pair is defined

$$\hat{F} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} .$$

B. Geometric meaning of projective transformation matrix

Let H_l and H_r be the projective transformation matrices of images I_l and I_r, and let I_l' and I_r' be the image pairs after rectification, and m_l' and m_r' be points after transformation on m_l and m_r, then m_l' and m_l, m_r' and m_r satisfy the following two equations,

$$m_l' = H_l m_l, m_r' = H_r m_r .$$

The images I_l' and I_r' satisfy the epipolar constraint. Therefore there exist

$$m_r'^T \hat{F} m_l' = 0$$

$$m_r^T \underbrace{H_r^T \hat{F} H_l}_{F} m_l = 0 .$$

i.e.

$$F = H_r^T \hat{F} H_l \quad . \tag{2}$$

Note the H_l and H_r satisfying (2) is not unique. The aim of image rectification is to determine the optimal H_l and H_r to minimize the distortion after image transformation.

As shown in the Fig.1, the projective transformation matrices used for rectification have the definite geometric meaning. let H_l be of the form

$$H_l = \begin{pmatrix} u_l^T \\ v_l^T \\ w_l^T \end{pmatrix} \begin{pmatrix} u_{la} & u_{lb} & u_{lc} \\ v_{la} & v_{lb} & v_{lc} \\ w_{la} & w_{lb} & w_{lc} \end{pmatrix} .$$

Because H_l maps epipole e_l to an infinite point, it follows that

$$H_l e_l = (u_l^T e_l \quad v_l^T e_l \quad w_l^T e_l)^T = (1 \quad 0 \quad 0)^T .$$

So, the epipole e_l lays on the lines v_l and w_l. With regard to H_r, the epipole e_r lies on the lines v_r and w_r. Furthermore, lines v_l and v_r, lines w_l and w_r must be the corresponding epipolar lines in the images I_l and I_r.

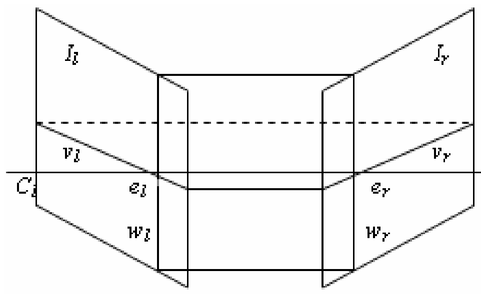

Figure 1. Geometric meaning of projective matrix.

III. STRATIFIED RECTIFICATION OF STEREO PAIRS

In this study, we divide the process of image rectification into three stages. First, we compute projective transformation P_l and P_r such that all epipolar lines are parallel in rectified images and the epipole is mapped to an infinite point. Secondly, we compute affine transformation A_l and A_r such that all epipolar lines are parallel to horizontal direction and no vertical difference exist between corresponding epipolar lines of image pairs. Thirdly, we compute shearing transformation S_l and S_r to minimize the horizontal difference between corresponding points.

A. *Projective transformation matrix*

From Fig.1, we know that projective transformation matrix P_l and P_r can be determined by corresponding epipolar lines w_l and w_r in image pairs, and have the form

$$P_l = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_{la} & w_{lb} & 1 \end{pmatrix}, P_r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_{ra} & w_{rb} & 1 \end{pmatrix}. \quad (3)$$

Given an infinite point z in image I_l, the following relation exists between epipolar lines w_l and w_r

$$w_l = [e_l]_x z, w_r = Fz \quad z = (\lambda \quad \mu \quad 0)^T . \quad (4)$$

By determining infinite point z, P_l and P_r satisfy the constraint, i.e. mapping the epipoles e_l and e_r to points at infinity; and effectively reducing distortion in rectified images. Therefore P_l and P_r should be as close as possible to affine transformation matrix. Supposing w and h are the width and height of image, Charles Loop uses the nonlinear least square method to solve the equation, determining the infinite point z. The objective function is

$$\frac{z^T \overbrace{[e_l]_x^T PP^T [e_l]_x}^{A} z}{z^T \underbrace{[e_l]_x^T P_c P_c^T [e_l]_x}_{B} z} + \frac{z^T \overbrace{F^T P' P'^T F}^{A'} z}{z^T \underbrace{F^T P_c' P_c'^T F}_{B'} z} \quad (5)$$

where

$$PP^T = \frac{wh}{12}\begin{pmatrix} w^2-1 & 0 & 0 \\ 0 & h^2-1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$P_c P_c^T = \frac{1}{4}\begin{pmatrix} (w-1)^2 & (w-1)(h-1) & 2(w-1) \\ (w-1)(h-1) & (h-1)^2 & 2(h-1) \\ 2(w-1) & 2(h-1) & 4 \end{pmatrix}.$$

The initial values of λ and u can be obtained as follows [7]. Firstly, the Quantities

$$z^T Az \big/ z^T Bz, \quad z^T A'z \big/ z^T B'z$$

of (5) are minimized separately, which give two different estimations of z, denoted by z_1 and z_2. Secondly, their average

$$(z_1 / \|z_1\| + z_2 / \|z_2\|) / 2$$

is used as the initial value of z, which is very close to the optimal solution. Minimizing $z^T Az/z^T Bz$ is equivalent to maximizing $z^T Bz \,/\, z^T Az$, denoted by $f(z)$. As A is symmetric and positive-definite, it can be decomposed as $A = D^T D$. Let $y = Dz$. Then, $f(z)$ becomes

$$\hat{f}(y) = \frac{y^T D^{-T} BD^{-1} y}{y^T y}.$$

Imposing $\| y \| = 1$, and $\hat{f}(y)$ is maximized when y is equal to the eigenvector of $D^{-T}BD^{-1}$ associated with the largest eigenvalue. So, the solution z_1 is given by following equation

$$z = D^{-1} y .$$

Similarly, the same procedure can be applied to find the estimation z_2.

From above description, it is known that we need to use the Cholesky factorization to decompose matrices A and A', which should be positive-definite as required, when computing the initial value of infinite point z. But A and A' may not be positive-definite because of noise effect on corresponding image points. Thus we cannot determine the initial value of z and cannot compute matrices P_l and P_r.

The epipolar lines in each image are parallel after we use P_l and P_r to transform the images, and the epipolar plane consisting of corresponding epipolar lines are parallel; the intersecting line at infinity among these epipolar planes is base line. Then the image pairs satisfy affine epipolar geometry constraint, as shown in Fig. 2. So we can use affine epipolar geometry constraint to solve infinite point z and determine P_l and P_r.

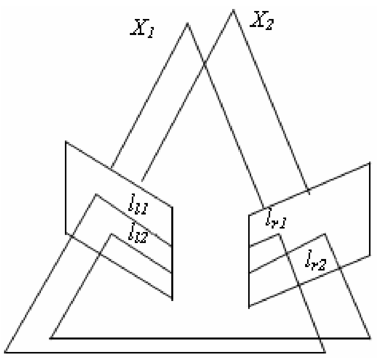

Figure 2. Affine epipolar geometry.

The rectified image pairs satisfy affine epipolar geometry constraint after transformation with P_l and P_r, the form of fundamental matrix is as follows

$$F_a = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{pmatrix} \quad (6)$$

where epipoles ela and era satisfy

$$e_{la} = \begin{pmatrix} -d & c & 0 \end{pmatrix}^T, e_{ra} = \begin{pmatrix} -b & a & 0 \end{pmatrix}^T .$$

Epipole e_l in the I_l can be computed by corresponding points in image pairs, e_l is denoted as e_l= （e_{lx}、 e_{ly}、 1） ^{T}. Since P_l maps e_l to e_{la}, we have

$$P_l e_l = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_{la} & w_{lb} & 1 \end{pmatrix}^T \begin{pmatrix} e_{lx} \\ e_{ly} \\ 1 \end{pmatrix} = \begin{pmatrix} -d \\ c \\ 0 \end{pmatrix} .$$

It follows that d=-e_{lx}, c=e_{ly}. Similarly, we denote e_r= （e_{rx}、 e_{ry}、 1） ^{T}, then a=e_{ry} ,b=-e_{rx}. Substituting these into(6) follows

$$F_a = \begin{pmatrix} 0 & 0 & e_{ry} \\ 0 & 0 & -e_{rx} \\ e_{ly} & -e_{lx} & e \end{pmatrix} \quad (7)$$

where e is an unknown parameter. The corresponding points satisfy epipolar geometry constraint after the transformation with P_l and P_r, i.e.

$$m_r{}^T P_r{}^T F_a P_l m_l = 0 .$$

Therefore we build optimization objective function as follows

$$\sum_{i-1}^{N} m_{r,i}{}^T P_r{}^T F_a P_l m_{l,i} = 0 \quad (8)$$

where $m_{l,i}$ and $m_{r,i}$ are corresponding points in images I_l and I_r. By substituting (3), (4), (7) into (8), we can acquire objective function related to λ, u and e parameters.

If we use nonlinear least square to solve object function, the selection of initial parameters should be close to global optimal solution. Otherwise, the process only results in local optimal solution. We directly and arbitrarily set the initial value of parameters λ, u and e only if λ and u is not zero simultaneously, and then perform optimization by Levenberg-Marquardt nonlinear optimization algorithm to complete the rectification process. In this paper, the initial value is selected to be (1,1,1), and the distortion of rectified images is comparatively small.

When estimating infinite points with above algorithm, it is unnecessary for relative matrix to be positive-definite, and the computation of projective matrix is completed with arbitrarily given corresponding image points.

B. Affine transformation matrix

Affine transformation matrix A_l and A_r have the form

$$A_i = \begin{pmatrix} F_{32} - w_{lb}F_{33} & w_{la}F_{33} - F_{31} & 0 \\ F_{31} - w_{la}F_{33} & F_{32} - w_{lb}F_{33} & F_{33} + v_{rc} \\ 0 & 0 & 1 \end{pmatrix} \quad (9)$$

$$A_i = \begin{pmatrix} w_{rb}F_{33} - F_{23} & F_{13} - w_{ra}F_{33} & 0 \\ w_{ra}F_{33} - F_{13} & w_{rb}F_{33} - F_{23} & v_{rc} \\ 0 & 0 & 1 \end{pmatrix} \quad (10)$$

where v_{rc} is unknown parameter in matrices A_l and A_r which needs to be solved. The matrices A_l and A_r include rotation transformation and can change the parallel epipolar lines in images into the horizontal scan-line; offset F_{33} is used to regulate corresponding epipolar lines between images so that they lie on the same scan-line. If we directly use A_l and A_r to perform rotation and vertical translation transformation on the image, then sub-sampling will occur, resulting in severe contour effect.

To avoid above case, we decompose A_l into scale, rotation and translation transformation. θ_l is angle of rotation which changes epipolar lines into horizontal scan-line and scale factor s_l decides the scaling of images which is the main reason of sub-sampling. The process of decomposition is as follows

$$A_l = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & F_{33} + v_{rc} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_l & -\sin\theta_l & 0 \\ \sin\theta_l & \cos\theta_l & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} s_l & 0 & 0 \\ 0 & s_l & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The parameters θ_l, s_l and F_{33} are known, the matrix A_l is rebuilt and has the form as follows

$$A_l' = A_l \begin{pmatrix} s_l' & 0 & 0 \\ 0 & s_l' & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (11)$$

where s_l' is unknown scale factor used to adjust horizontal and vertical scale of images, and it needs to be recalculated. Similarly, we rebuild matrix A_r, and it has the form as follows

$$A_r' = A_r \begin{pmatrix} s_r' & 0 & 0 \\ 0 & s_r' & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (12)$$

where s_r' is unknown scale factor needs to be recalculated. In summation, the matrices A_l and A_r include three unknown parameters: s_l', s_r', v_{rc}, which can be solved through nonlinear least square method.

After transforming with matrices A_l' and A_r', stereo image pairs are basically in the ideal condition, i.e. corresponding epipolar lines lie on the same scan-line, and no vertical difference exists between corresponding epipolar lines. Therefore the objective function

$$\sum_{i=1}^{N}\left((A_l'\hat{m}_{l,i})_y-(A_r'\hat{m}_{r,i})_y\right)^2 \qquad (13)$$

can be solved by nonlinear least square to determine unknown parameters s_l', s_r' and v_{rc}. $\hat{m}_{l,i},\hat{m}_{r,i}$ denote corresponding points of image pairs after the projective transformation, and $(\cdot)_y$ denote y coordinate of a point. By solving the object function, the vertical difference of corresponding points is minimized after affine transformation on images, which then minimizes the vertical difference of corresponding epipolar lines between image pairs.

When A_l' and A_r' take the form in (11) and (12), the initial value of s_l', s_r' and v_{rc} is as follows

$$\left(\frac{1}{s_l} \quad \frac{1}{s_r} \quad \frac{1}{N}\sum_{i=1}^{N}\left|(\hat{m}_{l,i})_y-(\hat{m}_{r,i})_y\right|\right).$$

Scale factor s_l and s_r of matrices A_l and A_r are used for scaling transformation. In order to control the horizontal distortion when adjusting the vertical difference of corresponding epipolar lines, we use inverse scaling transformation, so that the initial value of adjustment scale factor is selected as $\left(\frac{1}{s_l} \quad \frac{1}{s_r}\right)$. The initial value of v_{rc} denotes mean value of vertical difference of corresponding points in images.

By solving object function (13), the matrices A_l' and A_r' have the following advantages

(1) minimizing the vertical difference of corresponding epipolar lines in images;

(2) minimizing the horizontal difference of corresponding points in images.

Besides the above optimization algorithm, we present another method for optimizing the affine transformation matrix. When decomposing A_l into scale, rotation and translation transformation, we only use rotation and translation matrix, and rebuild the affine matrix that has the following form

$$A_l'=\begin{pmatrix}1&0&0\\0&1&F_{33}+v_{rc}\\0&0&1\end{pmatrix}\begin{pmatrix}\cos\theta_l&-\sin\theta_l&0\\\sin\theta_l&\cos\theta_l&0\\0&0&1\end{pmatrix}\begin{pmatrix}s_l'&0&0\\0&s_l'&0\\0&0&1\end{pmatrix}$$

where s_l' is unknown scale factor, and needs to be calculated. Similarly, we rebuild matrix A_r that has the following form

$$A_r'=\begin{pmatrix}1&0&0\\0&1&v_{rc}\\0&0&1\end{pmatrix}\begin{pmatrix}\cos\theta_r&-\sin\theta_r&0\\\sin\theta_r&\cos\theta_r&0\\0&0&1\end{pmatrix}\begin{pmatrix}s_r'&0&0\\0&s_r'&0\\0&0&1\end{pmatrix}$$

where s_r' is unknown scale factor needs to be calculated. In order to solve the unknown parameters: s_l', s_r', v_{rc}, the objective function (13) is also used, but the initial value of s_l', s_r' and v_{rc} is as follows

$$\left(1 \quad 1 \quad \frac{1}{N}\sum_{i=1}^{N}\left|(\hat{m}_{l,i})_y-(\hat{m}_{r,i})_y\right|\right).$$

Because the effect of scaling transformations in A_l and A_r are eliminated, so the initial value of adjustment scale factor (s_l' s_r') is selected as (1 1).

C. Shearing transformation matrix

After projective and affine transformation, the rectification of stereo image pairs is completed. But in order to ensure horizontal image distortion to be small enough, shearing transformation can be carried out as follows

$$S_l=\begin{pmatrix}a_l&b_l&0\\0&1&0\\0&0&1\end{pmatrix},S_r=\begin{pmatrix}a_r&b_r&0\\0&1&0\\0&0&1\end{pmatrix}. \qquad (14)$$

After projective and affine transformation on image I_l, the homogeneous coordinate of horizontal and vertical midlines can be denoted as

$$\hat{x}_l=\hat{b}_l-\hat{d}_l=\begin{pmatrix}x_{lu}&x_{lv}&0\end{pmatrix}^T$$
$$\hat{y}_l=\hat{c}_l-\hat{a}_l=\begin{pmatrix}y_{lu}&y_{lv}&0\end{pmatrix}^T.$$

where $\hat{a}_l,\hat{b}_l,\hat{c}_l,\hat{d}_l$ are the homogenous coordinates of midpoints at edges of the image respectively. By the properties of horizontal and vertical midlines

(1) perpendicularity: $(S_l\hat{x})^T(S_l\hat{y})=0$

(2) aspect ratio invariability: $\dfrac{(S_l\hat{x})^T(S_l\hat{x})}{(S_l\hat{y})^T(S_l\hat{y})}=\dfrac{w^2}{h^2}$

we can solve a_l and b_l, i.e.

$$a_l=\frac{h^2{x_{lv}}^2+w^2{y_{lv}}^2}{hw(x_{lv}y_{lu}-x_{lu}y_{lv})},b_l=\frac{h^2x_{lu}x_{lv}+w^2y_{lu}y_{lv}}{hw(x_{lu}y_{lv}-x_{lv}y_{lu})}. \qquad (15)$$

The sign of a_l is required to be positive. If it is negative, a_l and b_l are simultaneously multiplied by -1. Similarly, a_r and b_r can be solved with respect to I_r.

IV. EXPERIMENTAL RESULTS

This paper considers two aspects to evaluate the distortion of before and after rectification. Firstly, we compute the mean of difference in y coordinate of corresponding points in images to evaluate rectification accuracy of vertical difference between corresponding epipolar lines of images. The formula of computation is as follows

$$\begin{cases} \Delta y_{org} = \dfrac{1}{N}\sum_{i=1}^{N}\left|(m_{l,i})_y - (m_{r,i})_y\right| \\ \Delta y_{rec} = \dfrac{1}{N}\sum_{i=1}^{N}\left|(m_{l,i}')_y - (m_{r,i}')_y\right| \end{cases} \quad (16)$$

where $\Delta y_{org}, \Delta y_{rec}$ are denoted as the mean of difference in y coordinate of corresponding points before and after image rectification [6]. Secondly, we compute the mean of difference in x coordinate of corresponding points to evaluate the distortion of horizontal difference after rectification. The formula of computation is as follows

$$\begin{cases} \Delta x_{org} = \dfrac{1}{N}\sum_{i=1}^{N}\left|(m_{l,i})_x - (m_{r,i})_x\right| \\ \Delta x_{rec} = \dfrac{1}{N}\sum_{i=1}^{N}\left|(m_{l,i}')_x - (m_{r,i}')_x\right| \end{cases} \quad (17)$$

where $\Delta x_{org}, \Delta x_{rec}$ are denoted as the mean of difference in x coordinate of corresponding points before and after image rectification; $(\cdot)_x$ indicates x coordinates.

Through optimization of affine transformation matrices, the algorithm not only make corresponding epipolar lines become the same scan-line, eliminating the vertical difference but also make horizontal difference of corresponding points approach to minimization. To verify the function of affine transformation matrices, we use two methods to rectify input stereo image pairs. The first method only performs projective and affine transformation on images, while the second method includes shearing transformation. The matrices of transformation in these two methods are defined as follows

$$H_l = A_l' P_l, H_r = A_r' P_r .$$

and

$$H_l = S_l A_l' P_l, H_r = S_r A_r' P_r$$

where A_l' and A_r' use the form of (11) and(12). After the rectification is accomplished, we can compute the mean of difference in y, x coordinate of corresponding points by (16) and (17) to realize the evaluation on the distortion of horizontal and vertical direction in images.

In experiment, selected stereo image pairs are representative and reflect different imaging situation of vision system. The figures of experimental results are divided into three parts: the upper part are original image pairs, the middle part are result of rectification used by the first method, the lower part are result of rectification used by the second method. In these figures, the decussating points express corresponding points, the beeline express epipolar lines.

Fig.3 are the original image pairs and the result of rectification related to the office gallery. The epipoles of original image pairs approach infinite respectively. Deflection occurs in the imaging course of the left image, and scaling transformation has a weak effect on the left and right images. Fig.4 are the original image pairs and the result of rectification related to the school park. The epipoles of original image pairs converge to a point, which belongs to the imaging condition of convergent stereovision. Contrast to left image, the right image is affected more strongly by scaling transformation.

Figure 3. The office gallery.

Figure 4. The school park.

Fig.5 are the original image pairs and the result of rectification with respect to the office environment. The two epipoles approach a point respectively, and deflections occur in the imaging course of the left and right image. Fig.6 are another image pairs and the rectification results. The left epipole approaches to an infinite point, the right epipole converges to a point that is comparatively close to original image. The reason for this condition is that the camera has a relatively large forward motion in the imaging course.

Figure 5. The office environment.

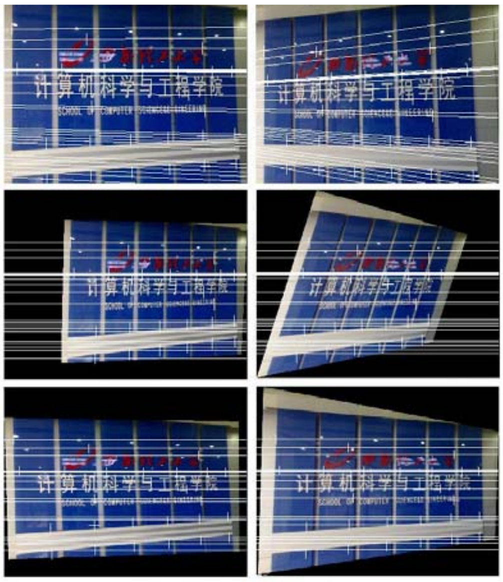

Figure 6. The name board.

The mean of difference in y, x coordinate of corresponding points in original image pairs and rectified pairs computed by two methods is given by table 1. The fourth and fifth rows are experimental results computed with the first method while the sixth and seventh rows are results computed with the second method. From the table, we can see that for all imaging conditions, the vertical difference of these two rectification methods are the same since shearing transformation has no effect on y coordinate of corresponding points. For Fig.3-5, the difference of horizontal difference between the rectification result used by the first method and original image pairs are comparatively small, so the first method is superior to the second method, especially for the condition of Fig. 5. For Fig.6, the second method is better than the first method in horizontal difference. In this case, if camera has comparatively strong forward motion in the imaging course, shearing transformation must be used to rectify the horizontal distortion of image in the process of rectification.

TABLE I.
THE MEAN COMPUTED WITH EQUATION 16-17

image / method		**Office gallery**	**Scholl park**	**Office**	**Name board**
I	Δx_{org}	56.269	201.590	131.670	140.480
	Δy_{org}	13.000	20.118	60.800	40.556
II	Δx_{rec}	53.682	211.330	202.530	312.720
	Δy_{rec}	0.977	0.837	0.717	0.422
III	Δx_{rec}	59.401	214.680	372.260	23.244
	Δy_{rec}	0.977	0.837	0.717	0.422

V. CONCLUSION

The paper divides the rectification algorithm of stereo pairs under uncalibrated camera into three phases. Firstly, we determine projective matrices of images with affine epipolar geometry constraint so that all epipolar lines in image are parallel. Secondly, we optimize the affine transformation and use the constraint of minimum vertical difference between corresponding epipolar lines to adjust scale factor and parameter of vertical offset. Finally, we carry out scaling, rotation and translation transformation on image pairs to complete the rectification process. The paper can obtain the global optimal solution for projective matrices under the condition of arbitrarily given corresponding image points. At the same time, this paper gives not only the optimization parameters to eliminate vertical difference of corresponding epipolar lines, but also presents a calculation method of initial value of optimization parameters such as scale factor and vertical offset. By analyzing the distribution of epipolar lines of stereo pairs, we conclude that if camera has comparatively large forward motion in the imaging process, shearing transformation must be performed in the rectification process.

When using affine epipolar geometry constraint to compute the projective transformation matrix, we can arbitrarily set the initial value of parameters λ, u and e only if λ and u is not zero simultaneously. In most situations, the distortion of rectified images is comparatively small after performing optimization with Levenberg-Marquardt nonlinear optimization algorithm. But the selection of initial parameters cannot be close to the global optimal solution always. In order to solve this problem, we can utilize global optimization of genetic algorithm to perform a rough estimation of unknown parameters for all corresponding points in images. The obtained parameters can be used as initial value in nonlinear optimization algorithm, namely Levenberg-Marquardt algorithm, for accurate estimation. When applying genetic algorithm, we can obtain the global optimal solution by adjusting some parameters, such as the initial range of species that can be selected to increase the distance between individuals for resulting in a wide variety of the species, the scale of species that can be selected to increase the number of species for obtaining more points searched by genetic algorithm, etc.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant No.60743009 and the Natural Science Foundation of Shannxi Province under Grant No.2006F43. The authors would like to thank for their help.

REFERENCES

[1] J. Gluckman and S. K. Nayar. “Rectifying transformations that minimize resampling effects”, In: Proceedings of the International Conference on Computer Vision and Pattern Recognition, Los Angeles, USA, 2001, 1, pp.111-117.
[2] A. Fusiello, E. Trucco, and A. Verri. “A compact algorithm for rectification of stereo pairs”, Machine Vision and Applications, 2000, 12(1), pp.16-22.
[3] R. Hartley, “Theory and practice of projective rectification”, International Journal of Computer Vision, 1999, 35(2), pp.115-127.
[4] Francesco Isgro, Trucco, “Projective rectification without epipolar geometry”, In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, 1999, 1, pp. 94-99.
[5] Al-Zahrani, S. S. Ipson, and J. G. B. Haigh, “Applications of a direct algorithm for the rectification of uncalibrated images”, Information Sciences – Informatics and Computer Science, 2004, 160(1-4), pp. 53-71.
[6] Hsien-Huang P. Wu, and Yu-Hua Yu, “Projective rectification with reduced geometric distortion for stereo vision and stereoscopic video”, Journal of Intelligent and Robotic Systems, 2005, 42(1), pp. 71-94.
[7] C. Loop and Z. Zhang, “Computing rectifying homographies for stereo vision”, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA, 1999, 1, pp. 125-131.
[8] Hartley R, Zisserman A, Multiple View Geometry in Computer Vision, Cambridge, network Britain: Cambridge University Press, 2000.

Sui Liansheng received his B.S. in computer software from Harbin Institue of Technology, China, in 1995, and M.S. in compter application from Xi’an University of Technology, China, in 2000, and the Ph.D in instrument scienece and technology from Xi’an Jiaotong University, China, in 2003.

He is now with Xi’an University of Technology as an associate professor. His research interests include stereo vision, 3D modeling, view synthesis and motion analysis.

Zhang Jiulong received his Bachelor degree in computer science, from Northwestern Polytechnic University, China, in 1993, and his Doctor degree in automatic control theories from Northwestern Polytechnic University in 2000.

He is now with Xi’an University of Technology as a lecturer. He has published nine papers regarding face recognition, and pattern recognition. His research interests include face recognition, visual tracking, and machine learning.

Dr. Zhang is member of Chinese computer Federation.

Cui Duwu received his Bachelor's degree and Master's degree in electric engineering field from Xi’an Jiaotong University, China, in 1967 and 1983 respectively. His major field of study is electric engineering.

He has worked at Qinghai power tools plant, Xi’ning, China since job assignment on graduation as a technician. He has been a teacher in Xi’an University of Technology Xi’an China since 1983. He has written several books and papers about electric circuit theory and intelligent computing. Now he is Professor and Ph.D Supervisor of Xi’an University of Technology Xi’an China. His current research interests are intelligent computing and developing computer software .