TELKOMNIKA, Vol.11, No.1, January 2013, pp. 322~329
ISSN: 2302-4046

Rectification of Uncalibrated Images for Stereo Vision

Han Huiyan*, Han Xie, Yang Fengbao
School of Information and Communication Engineer, North University of China, Taiyuan 030051, Shanxi, China
*Corresponding author, e-mail: hhy980344@163.com

Abstract

This paper aims at rectification of uncalibrated stereo images. We use the direct method to compute the rectifying transformations, not to deduce from the fundamental matrix, so to reduce the uncertainty; Minimize the distance from the point to its epipolar line to ensue the uniqueness of the two rectifying transformations. We only extract a few correspondent points, make use of the condition of the correspondent points in the rectified images also meet the polar constraint, minimize distance of all of the correspondent points to the corresponding epipolar line, then obtain the system equations, through solving system of equations, thus get the rectifying transformation matrix and rectify the stereo image pairs. Experiments show that, the size of the rectified images change not so seriously, the vertical disparity is eliminated, so the subsequent stereo correspondence can be fulfilled on the horizontal scan line of the rectified images completely.

Keywords: *Stereo Vision, Rectification transform matrix, Disparity, Fundamental matrix*

1. Introduction

Stereo vision is the technique aims at inferring scene structure from two or more images taken from different viewpoints, and is used in many applications [1]. One of the important and difficult problems in the hot research area is the stereo correspondence. Stereo correspondence is the process of seeking the correspondent pixel point in the other image for each image point [2], and the two pixel points are homologous points and they are projections of the same object point. Because the constraint epipolar geometry, the corresponding point must lie on an epipolar line in the other image, this can also be termed as Fundamental matrix, simplifies the problem from 2D to 1D. If the images are caught from rectilinear stereo rig, the epipolar lines will be horizontal and paralled with image scan lines,i.e. X-axis. However, the mechanical accuracy of stereo rig is finite,and some of the artificial reasons,the stereo rig is not ideal always, and we can warp the images to make the correspondent point lie on the horizontal scan line, the process is image rectification.

The early software rectification approach must rely on the camera parameters and the necessity of camera calibration limits its range of applications. Thus many scholars have studied and proposed rectification methods without calibrating cameras. Robert [3] first proposed the uncalibrated rectification, the transformation must preserve the orthogonality of the two center lines of the rectified images. Hartley [4] thought one of the two transformations must be a rigid one. Loop and Zhang [5] decomposed the transformation into similarity and shearing factor to minimize the image distortions. Isgro and Trucco [6] did't need the epipolar geometry. Mallon and Whelan [7] relied soley on the examination of the fundamental matrix, minimized perspective distortions. Chen et al. [8] analyzed the different regions for the position of the epipole and resampled the rectification images. Most of the algorithms use the priori estimation knowledge of the fundamental matrix, because the estimation need the correspondent points and accuracy of the matrix has its uncertainty and these indirect methods might lead to some unexpected rectified images.

Isgrò and Trucco [9] adopted an entirely different method and obtained the two homographies directly without first estimating the fundamental matrix. However, in order to obtain the uniquely rectification, he minimized disparity along the x-axis, but the special

Received October 1, 2012; Revised November 19, 2012; Accepted December 9, 2012

constraint will not be appropriate for certain applications, and the approach will greatly distort the original images sometimes.

In this paper, we propose a method for rectifying two uncalibrated images with reduced geometric distortion. We rectify the two uncalibrated images without explicitly estimation of the fundamental matrix. Its advantage is to formulate a new set of parameters of the two homographies, and solves the rectification problem using the criterions that minimize the distance between the correspondent point and its corresponding epipolar line.

The remainder of the paper is organized as follows: First our rectification method is introduced. Then the test stereo image rectification is carried out and results are gave and compared with others'. Finally we conclude our paper.

2. Rectification method

Our method derives from the projective geometry, and we specify the representation of various variables in the following text as: The point in image is expressed in a 3D column vector with homogeneous coordinate. Matrix is expressed by bold upper-case letter, such as F and H. The transposed vector and matrix are expressed by adding a letter T as superscript, such as m^T and F^T. The vector and matrix after rectification are expressed by adding a line on the top of the variable, such as $\overline{\mathrm{m}}$ and $\overline{\mathrm{F}}$.

2.1. Epipolar geometry constraint

There are two images, left image I and right image I'.Point m in I and m' in I' are the projections of object point M. The optical center of left and right camera is C and C' respectively, any plane that contains any object point M and the base line (the line joint C and C') is called epipolar plane. The intersection line of any epiploar plane and image plane is called epipolar line. All the epipolar lines in each image plane intersect at a single point, we call this point the epipole.(e for I and e' for I'). The epipole e is the projection of the optical center C' of the right camera and be visible in I' and vice versa. According to epipolar geometry constraint, given a pixel point m in I,the corresponding point m' must lie on the epipolar line L' in the right image I'. The fundamental matrix F maps the point m to its corresponding epipolar line L',i.e. L'=Fm, m' is on line L', these can be expressed as:

$$\begin{cases} m'^T Fm = 0 \\ m^T F^T m' = 0 \end{cases} \qquad (1)$$

Epipoles e and e' are the null space of F and F^T, respectively, as expressed by equation (2).

$$Fe = F^T e' = 0 \qquad (2)$$

All the epipolar lines(L and L') will pass the epipoles.

$$L^T e = L'^T e' = 0 \qquad (3)$$

2.2. Epipolar geometry of the rectified image

Given a pair of stereo images, rectification determines a transformation such that conjugate epipolar lines become collinear and parallel to X-axis of the image and the corresponding points will have the same y coordinates and the epipoles are mapped to a point at infinity.The rectified epipoles $\overline{e}$ and $\overline{e'}$ will have the special coordinates , $\overline{e} = \overline{e'} = [1\ 0\ 0]^T$,and the fundamental matrix after rectification will equal to skew-symmetric matrix of the unit vector [1 0 0] ,i.e.

$$\overline{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

H and H' are the rectification transformation matrix for left and right images respectively,the correspondent points $\overline{\mathrm{m}}$ and $\overline{m'}$ in the rectified images also satisfy the epipolar geometry constraint, similar to (1) ,the constraint can be expressed as:

$$\begin{cases} \overline{m'}^T \overline{F}\,\overline{m} = 0 \\ \overline{m}^T \overline{F}^T \overline{m'} = 0 \end{cases} \tag{5}$$

$$\begin{cases} \overline{m} = \mathrm{Hm} \\ \overline{m'} = \mathrm{H'm'} \end{cases} \tag{6}$$

Combine equation (5) and (6),we can obtain equation (7) as:

$$\begin{cases} (H'm')^T \overline{F}(Hm) = 0 \\ (Hm)^T \overline{F}^T (H'm') = 0 \end{cases} \tag{7}$$

The equation (7) can be interpreted as an equation, because the other equation can be obtained by a transposition operation on the both sides of the equal sign of the one equation.

Compare equation (1) with (7),we can obtain the following equation.

$$F = H'^T \overline{F} H \tag{8}$$

Thus, some methods compute the rectifying homographies by estimating the fundamental matrix , because the solution of the fundamental matrix has certain uncertainty,the accuracy of rectifying homographies is limited and may cause great distortions sometimes.Moreover,given some corresponding points,the homographies obtained from (8) is not unique, so must impose some other constraints such as minimizing the horizontal disparity or preserving the orthogonalty and aspect ratio to ensure the uniqueness of the homographies and then reduce the rectified image distortion.

2.3. Image rectification

We follow the calibrated rectification[10], the rectification process can be thought as a linear transformation, and the rectified virtual camera parameters can be obtained by rotating the cameras around their optical centers until the two focal planes become coplanar. So the rectifying transformation is the collineation induced by the plane at infinity between the intrinsic camera parameters before and after rectification, so the two rectifying transformations can be expressed as:

$$\begin{cases} H = \overline{A}RA^{-1} \\ H' = \overline{A'}R'A'^{-1} \end{cases} \tag{9}$$

Where $\overline{A}$ and $\overline{A'}$ are the intrinsic rectified camera parameters of left and right camera respectively, A and A' are the intrinsic camera parameters of left and right camera before rectification respectively , A^{-1} is the inverse matrix of matrix A .From equation (9),we can see the rectified images are took by a pair of cameras, the location relation between them is one translation along the baseline.

Given pairs of correspondent points (m^j, m'^j) and substitute them in equation (7),we can obtain a nonlinear equation(10) of two rectifying transformations H and H':

$$(H'm'^j)^T \overline{F}(Hm^j) = 0 \quad (j = 1, 2...n) \tag{10}$$

From equation(10),we cannot see its explicit physical meaning, so we deform it and obtain a geometrically and measurable quantity, that is the distance from a pixel point to its corresponding epipolar line, similar to Sampson error, that is the first order approximation of the geometric reprojection error, moreover, that is the distance function for fitting the fundamental matrix, for our case, the error of the j-th correspondent point pair is expressed as:

$$E_S{}^j = \frac{[(m'^j)^T Fm^j]^2}{\left\|NFm^j\right\|^2 + \left\|(m'^j)^T FN\right\|^2} \tag{11}$$

Where N is a matrix as:

$$N = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

To prevent inconsistency of the epipolar geometry between the left and right images, we choose to operate simultaneously on both images and minimize the error to 0 for all the correspondent points. Hence,we obtain the system equation

$$\{ E_S{}^j = 0 \} \quad \text{(j=1,2…n)} \tag{12}$$

The two rectifying transformations are obtained by solving the system equation(12) using least-square method. From equation (9),the intrinsic camera parameters (A,A') and the rotation matrices (R,R') are unknown, but the rectified intrinsic camera parameters $(\overline{A}, \overline{A'})$ can be set arbitrarily, provided that the vertical focal length and vertical coordinate of the principal point are the same ,i.e. the data in the second row and third row in $\overline{A}$ are equal to corresponding data in $\overline{A'}$ .moreover, $(\overline{A'})^T \overline{F}\overline{A'} \approx \overline{F}$ (are equal up to a invariant factor),so we omit the $\overline{A}$ and $\overline{A'}$,thus the fundamental matrix F can be simplified as :

$$F = A'^{-1} R'^T \overline{F} R A^{-1} \tag{13}$$

Each rectifying transformation rely on five intrinsic camera parameters and three rotation parameters, however, rotation of one camera along its X-axis can be omitted, for example, we set the rotation around its X-axis of left camera to zero. For simplicity, we assume the skew of the two cameras before rectification to zero, principle point locates at the center of the image, aspect ratio equal to one. Thus, the only remaining unknown parameters are the focal length of the two cameras. We assume they are the same and equal to f, then, the intrinsic parameters of the two cameras are:

$$A = A' = \begin{bmatrix} f & 0 & w/2 \\ 0 & f & h/2 \\ 0 & 0 & 1 \end{bmatrix} \tag{14}$$

Where, w and h are the width and height of the image.

From the analysis and deduction above, we can learn, to obtain the two rectifying transformations, we must know the six unknown parameters: local length f and two rotation angles of left camera and three rotation angles of right camera. According to [11], the focal length f varies in the range $[\frac{1}{3}(w+h), 3(w+h)]$.

Finally, we can obtain rectifying transformations through equation (9). We outline our rectifying method as followings steps:

(1) Initialize the six parameters to zero.

(2) Substitue all the correspondent points obtained in advance into equation (12). use least-square method to solve the equation, limit the range of the five rotation angles to $[-\frac{\pi}{2}, \frac{\pi}{2}]$,in order to acquire the effective value of local length f, we expand the range of f to $\left[\frac{1}{3\sqrt{3}}(w+h), 3\sqrt{3}(w+h)\right]$.

(3) Optimize the value using Levenberg-Marquardt method, thus obtain the optimal solution of the six parameters.

(4) Use the six parameters in step(3) ,we compute the intrinsic camera parameters and rotation angles, A、A'、R、R'、$\overline{A}$ and $\overline{A'}$.

(5) Substitute the result of step (4) into equation (9), then we obtain the two rectifying transformation H and H'.

3. Experiments and Analysis

We perform rectifying experiments using three of authoritative testing stereo image pairs provided by SYNTIM team [12]. Firstly, given a pair of image, some feature points are extracted using SIFT [13] method, then the 128D descriptors are compared to get the putative correspondent points, in order to improve the accuracy of the correspondent points, we then use RANSAC to discard the outliers, all the inliers will fit the fundamental matrix, and are used as the known parameters for our rectifying, compute the minimum bounding box of the two rectified images, with bilinear interpolation and the constraint of the minimum bounding box, we obtained the rectified images. The results are shown in figure 1, 2 and 3. The first row are the original images, and second row rectified images. Moreover, we selected seven feature points arbitrarily, and be labeled with red cross in left image, the yellow line in the right image is the corresponding epipolar line of the feature point in left image.

3.1 Visual evaluation

We can see from Figure 1, 2 and 3, the sizes of the rectified images change not so seriously, and the epipolar lines in the right image on second row are cascaded with each other and coincident with the X-axis of the image, and that is of course we expect, and achieve the rectification goals, can simplify the subsequent stereo correspondence greatly.

3.2. Quantitative evaluation

We perform the quantitative evaluations in two forms beside the visual evaluation. One is to compute the mean value of the Sampson error of all the correspondent points before and after rectification, that is the root mean squared distance (in pixel unit) from each point to its epipolar line. The other is mean vertical disparity (in pixel unit) of each correspondent point pair before and after rectification. The two evaluation values are compared with Hartley' method.

Mean value of the Sampson error of all the correspondent points before (E_{ori}) and after (E_{rec}) rectification are computed using the following equation:

$$\begin{cases} E_{ori} = \frac{1}{n}\sqrt{\sum_j E_S{}^j(m^j, m'_j)} \\ E_{\mathrm{rec}} = \frac{1}{n}\sqrt{\sum_j E_S{}^j(\overline{m^j}, \overline{m'_j})} \end{cases} \quad (15)$$

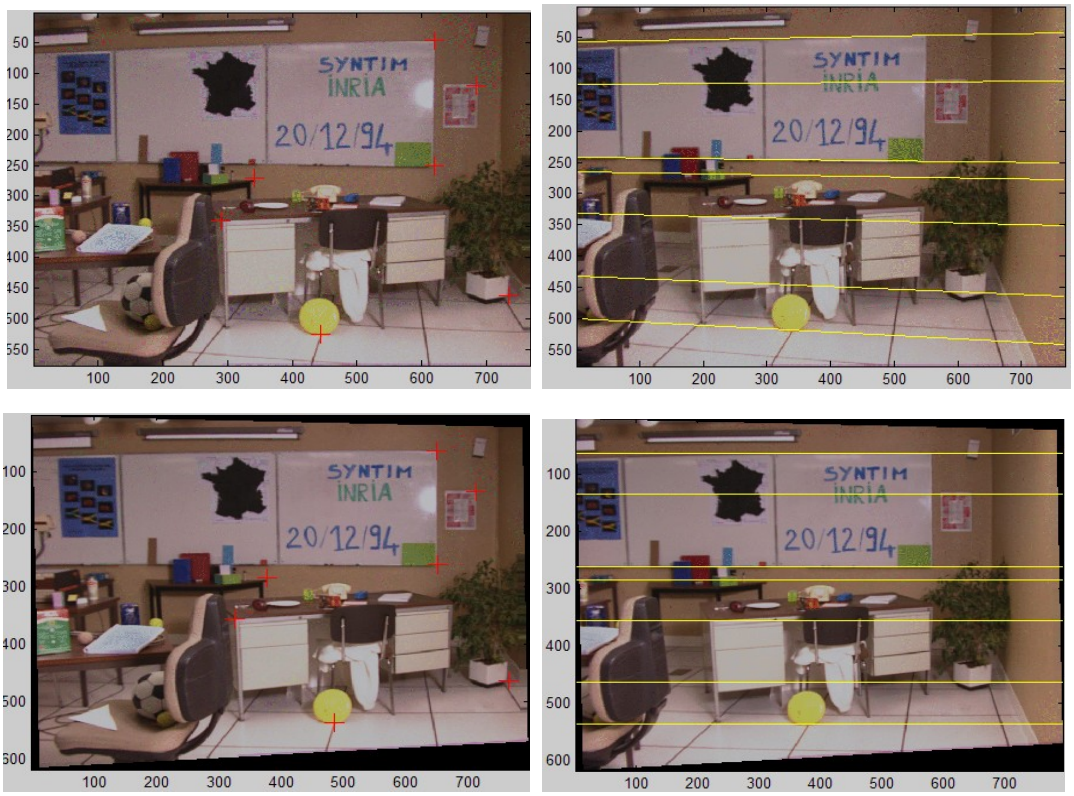

Figure 1. Original (the first row) and the rectified (the second row) of image “Sport”

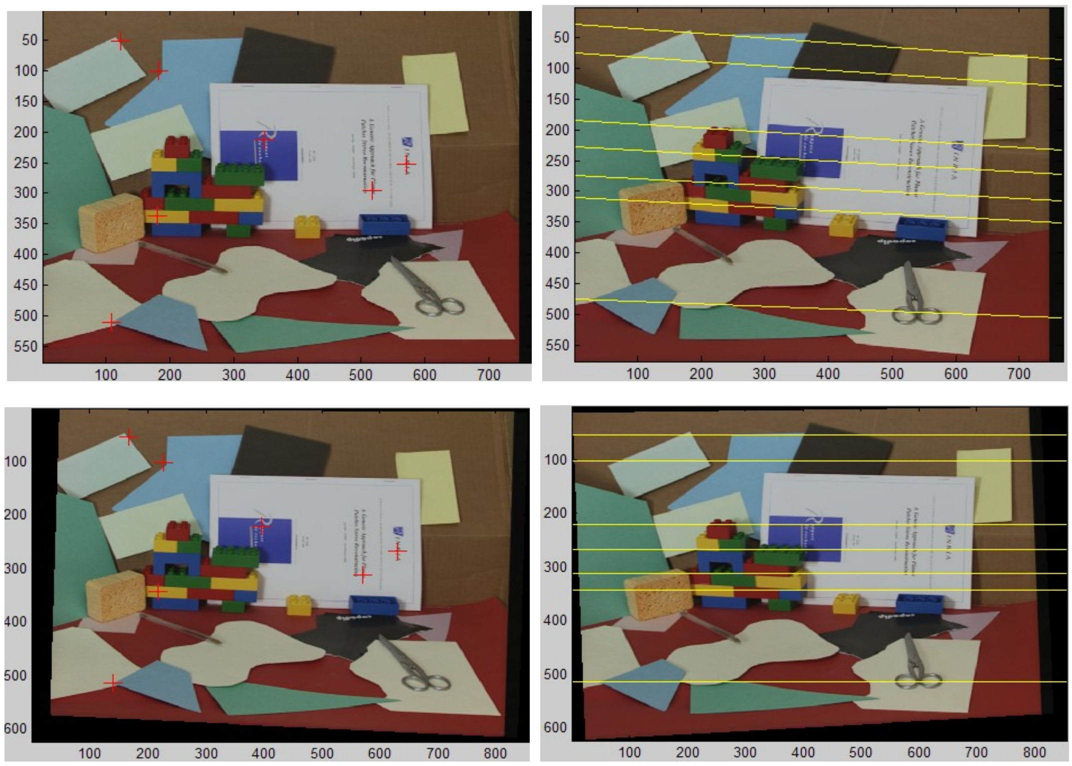

Figure 2. Original (the first row) and the rectified (the second row) of image “Tot”

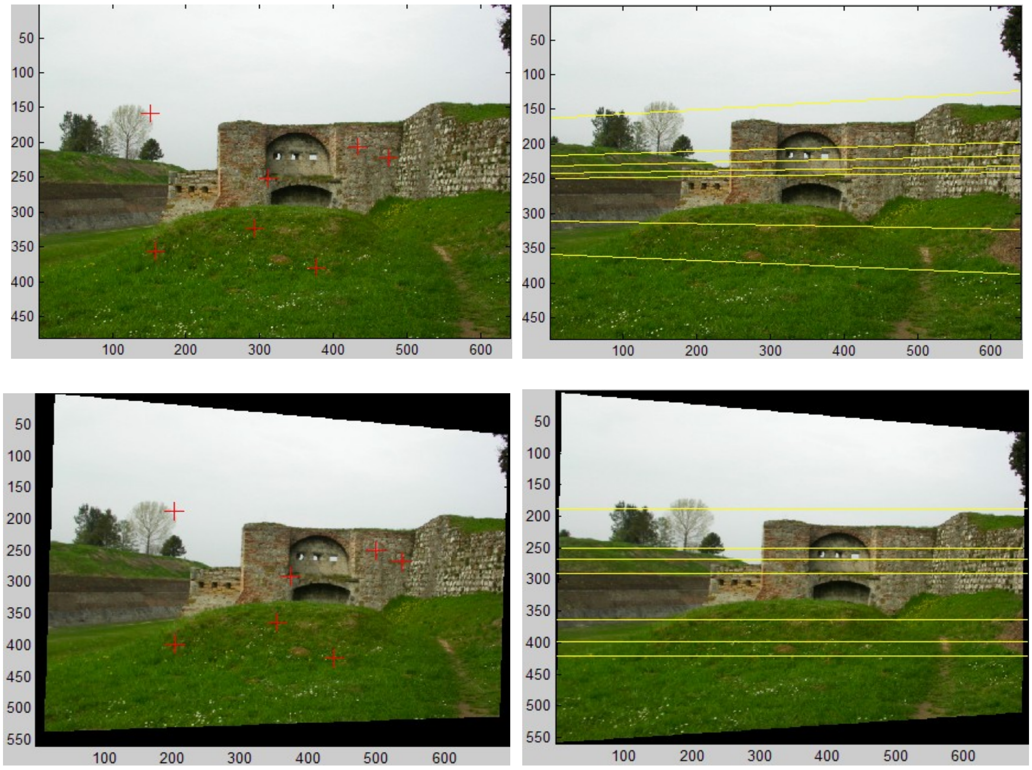

Figure 3. Original (the first row) and the rectified (the second row) of image “Cporta”

Where n is the number of the correspondent point pair, (m^j ,m'^j) is the j-th point pair, and the fundamental before rectification is computed using the improved eight points algorithm, the fundamental after rectification is computed using equation (12). Table 1 lists the mean distance of proposed method and Hartley’s. we can see from table 1, the mean error of our method is lower than Hartley’s, indicationg the effectiveness of ours.

Table 1. The distance before and after rectification

method	Hartley			Proposed		
Image	Sport	Tot	Cporta	Sport	Tot	Cporta
Original	0.31681	0.7417	0.27015	0.31681	0.7417	0.27015
Rectified	0.20463	0.26952	0.18234	0.19342	0.2151	0.17530

The vertical disparities before and after rectification are expressed as:

$$\begin{cases} \overline{\mathrm{d}_{ori}} = \dfrac{1}{n}\sqrt{\sum_j \mathrm{d_y}^j (m^j, m'^j)} \\ \overline{\mathrm{d}_{\mathrm{rec}}} = \dfrac{1}{n}\sqrt{\sum_j \mathrm{d_y}^j (\overline{m^j}, \overline{m'^j})} \end{cases} \quad (16)$$

Where n is the number of the correspondent point pair , $\mathrm{d_y}^j$ is the vertical disparity of the j-th point pair. Table 2 gives the mean vertical disparities of all the point pairs before and after rectification. We can obtain our mean vertical disparity is more close to zero than Hartley’s.

Table 2. The mean vertical disparity of point paris before and after rectification

method	Hartley			Proposed		
Image	Sport	Tot	Cporta	Sport	Tot	Cporta
Original	9.5241	12.9534	13.6204	9.5241	12.9534	13.6204
Rectified	0.7951	0.9537	0.6982	0.6582	0.7451	0.5238

4. Conclusion

This paper puts forward the rectification method of uncalibrated images. we use the direct method to obtain the two rectifying transformations. By directly computing the parameters of the transformation to reduce its uncertainty; By minimize the distance from a point to its epipolar line to ensure its uniqueness.

First, extract some correspondent points in the stereo image pair, the correpondent points in the rectified images also meet epipolar geometric constraint, minimizing distance from a correspondent point to its epipolar line, get the system equation, solve the system equations to obtain the rectifying the transformation matrices, and rectify the images. Experiments show that, the sizes of the rectified image change not so seriously, vertical disparity value is eliminated, greatly simplify the subsequent stereo correspondence.

Acknowledgements

This project is supported by the National Nature Science Foundation of China (No. 61171057) and Shan Xi Province Nature Science Foundation of China (No.2010011023-1).

References

[1] Feng Wei and Bao Wenxing. A new technology of remote sensing image fusion. *Telkomnika*. 2012; 10(3): 551-556.
[2] Sun Jun and Wang Yan. A new image segmentation algorithm and its application in lettuce object segmentation. *Telkomnika*. 2012; 10(3): 557-563.
[3] Robert L, Zeller C & Faugeras, O.et al. Applications of non-metric vision to some visually-guided robotics tasks, In: Visual Navigation: From Biological Systems to Unmanned Ground Vehicles. *Aloimonos Y*.1997:145-199.
[4] R Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision*. 1999; 35(2): 1-16.
[5] C Loop and Z Zhang. Computing rectifying homographies for stereo vision. In proceedings of the *IEEE Conferenceon*. Computer Vision and Pattern Recognition. 1999: 125–131.
[6] Isgrò F & Trucco E. On Robust Rectification for Uncalibrated Images. *IEEE International Conference on Image Analysis and Processing*. 1999; 297-302.
[7] J Mallon and PF Whelan. Projective rectification from the fundamental matrix. *Image and Vision Computing*. 2005; 23(7): 643–650.
[8] Chen Z, Shen P, Liu Y, Wu C, Quan L. Two epipole constraint of estimating fundamental matrix and its analyzing. *Proceedings of SPIE*, 3811. Denver Colorado, USA.1999; 337–344.
[9] F Isgro and E Trucco. Projective rectification without epipolar geometry. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*.1999; 1094–1099.
[10] Fusiello A. A Compact Algorithm for Rectification of Stereo Pairs. *Machine Vision and Applications*. 2000; 12(4): 1-8.
[11] A Heyden and M Pollefeys. Multiple view geometry. In G Medioni and SB Kang, editors, Emerging Topics in Computer Vision, Prentice Hall. 2005; 45-107.
[12] DG Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*. 2004; 60(2): 91–110.
[13] http://perso.lcpc.fr/tarel.jean-philippe/syntim/paires.html