울산대학교 공학연구 제34권 1호 pp. 27~40, 울산대학교.
Engineering Research Vol. 34, No. 1, pp. 27~40, 2003
University of Ulsan.

A Study on the Planar Rectification of Self-Calibrated Stereo Images

Jong Soo Lee · Paul Y. Oh
School of Computer Engineering and Information Technologies ·
Dept. of Mechanical Engineering, Drexel University

<ABSTRACT>

In this paper, we study a planar rectification technique that takes perspective projection matrices obtained by self-calibration. The self-calibration takes a nonlinear estimation called the Levenberg-Marquardt method on constraints imposed by the fundamental matrix to estimate the projection matrices. The rectification technique determines new perspective projection matrices by solving four 3×4 linear homogeneous systems obtained from rectifying plane constraints. This paper shows experimental results obtained by applying the technique to controlled stereo data and real stereo images. We make the comparison based on the average vertical coordinate difference of the rectified stereo image pair.

Keywords; Camera calibration, epipole, fundamental matrix, self-calibration, rectification

1. INTRODUCTION

To model real world objects, various approaches are proposed. One of them is the stereo image based rendering, which extract features points from right and left images and determines corresponding pairs of points for the 3D structure of the model. The accuracy of 3D models obtained from this technique depends on the correct correspondences of right and left feature points. For a point in one image, the technique searches for a corresponding point from its epipolar line on the other image. Planar rectification makes epipolar lines parallel to the image rows and requires the perspective projection matrices for the stereo images. It is possible to estimate the matrices precisely with camera calibration but arbitrary images cannot be used. One must perform a self-calibration to estimate their perspective projection matrices [1,2].

In this paper, stereo images are planar rectified using the linear transform matrix obtained from self-calibrated perspective projection matrices. The self-calibration [2] is described in detail and estimates the projection matrices from the fundamental matrix. Planar rectification is described in Section 2. Rectification experiments for controlled stereo data and real stereo image pairs are detailed in Section 3and conclusions are given in Section 4.

2. PLANAR RECTIFICATION

Stereo planar rectification corrects the left and right image planes such that every epipolar line is parallel to the image rows [5][6][7][8][9]. Instead of correcting the image planes, one can rectify stereo images by aligning epipolar pairs directly [5][10][11][12]. This rectification process requires complicated computations with higher complexities and restricts the size of stereo images.

2.1 Perspective projection matrices and epipoles

2.1.1 Perspective projection matrices

Let c denote the camera focal point and R denote the image plane. Let $w=[x,y,z]^T$ be a 3-dimensional (3D) point in the real world coordinate system and $m=[u,v]^T$ be its 2-dimensional point projected on the image plane. Here, the superscript T means the transposition of vectors. The point m is on the plane R and on the 3D line passing through the two points w and c. In the projective or homogeneous coordinate systems, w and m are represented by $\widetilde{w}$ and $\widetilde{m}$and they are related by the perspective linear transformation matrix $\widetilde{P}$ as follows:

$$\widetilde{m}=\widetilde{P}\widetilde{w} \tag{2.1}$$

$$\widetilde{w}=\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \qquad \widetilde{m}=\begin{bmatrix} U \\ V \\ S \end{bmatrix} \tag{2,2}$$

$$m=\begin{bmatrix} U/S \\ V/S \end{bmatrix}, \qquad S\neq 0 \tag{2.3}$$

When $S=0$, w is defined to be on the focal plane of the camera and $z=0$. Assuming the pinhole camera model, we have the perspective projection matrix P as follows

$$\widetilde{P} = A(I\,|\,0)G \tag{2.4}$$

In the above, A is the intrinsic camera parameter matrix given by

$$A = \begin{bmatrix} a_u & 0 & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.5}$$

where a_u, and a_v are the dimension of each pixel and (u_0, v_0) is the principle coordinate. G denotes the camera extrinsic parameter matrix composed of the 3×3 rotation matrix R and the translation vector t

$$G = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{2.6}$$

For the planar rectification, the perspective projection matrix P can be defined as follows

$$\widetilde{P} = \begin{bmatrix} q_1^T & | & q_{14} \\ q_2^T & | & q_{24} \\ q_3^T & | & q_{34} \end{bmatrix} = [P\,|\,\widetilde{p}] \tag{2.7}$$

In the above, the plane $q_3^T w + q_{34} = 0$ ($S=0$) is the focal plane and the two planes, $q_1^T w + q_{14} = 0$ and $q_2^T w + q_{24} = 0$, intersect with the retinal plane forming the vertical axis ($U=0$) and the horizontal axis ($V=0$), respectively. The focal point is the intersection of the three planes. Thus, we have

$$\widetilde{P}\begin{bmatrix} c \\ 1 \end{bmatrix} = 0, \tag{2.8}$$

$$c = -P^{-1}\widetilde{p} \tag{2.9}$$

The optical ray through the point m is the line cm and the points on the line compose the set $\{w \mid \tilde{m} = \tilde{P}\tilde{w}\}$. The line also can be represented by

$$w = c + \lambda P^{-1}\tilde{m} \ (\lambda\text{: a real number}). \tag{2.10}$$

2.1.2 Epipolar lines

Let c_1 and C_2 be the two focal points of the left and right pinhole cameras as shown in Fig. 2.1. Then, a point w in the real world is perspective projected to m_1 on the left image plane and m_2 on the right plane.

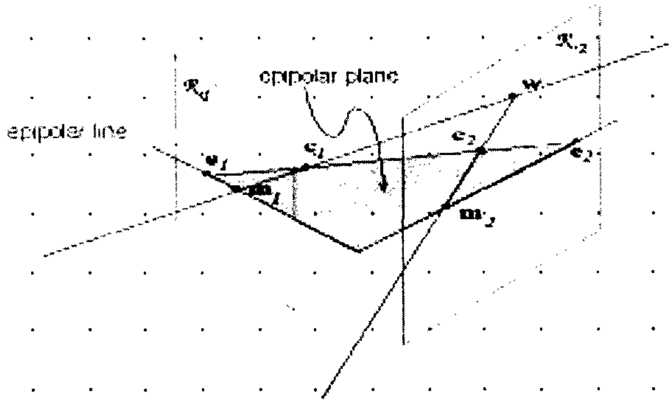

Fig. 2.1 Epipole Geometry

Then, the point m_2 on the right plane has its corresponding point on the epipolar line on the left plane. All the epipolar lines on one image plane pass through the epipole, which is the projection of the focal point c_2 on the left image plane

$$\tilde{e}_1 = \tilde{P}_1\begin{bmatrix} c_2 \\ 1 \end{bmatrix} \tag{2.11}$$

A point $\tilde{n}_1$ on the epipolar line for m_2 can be written by

$$\tilde{n}_1 = \tilde{e}_1 + \lambda P_1^{-1}\tilde{m}_2 \tag{2.12}$$

2.2 Planar Rectifications

Fig. 2.2 shows a planar rectification of a 3D point w. The point m_{01} projected on the image plane R_{01} is the point m_{n1} in the rectifying plane. And the two corresponding points m_{n1} and m_{n2} have the same vertical coordinate and $v_1 = v_2$.

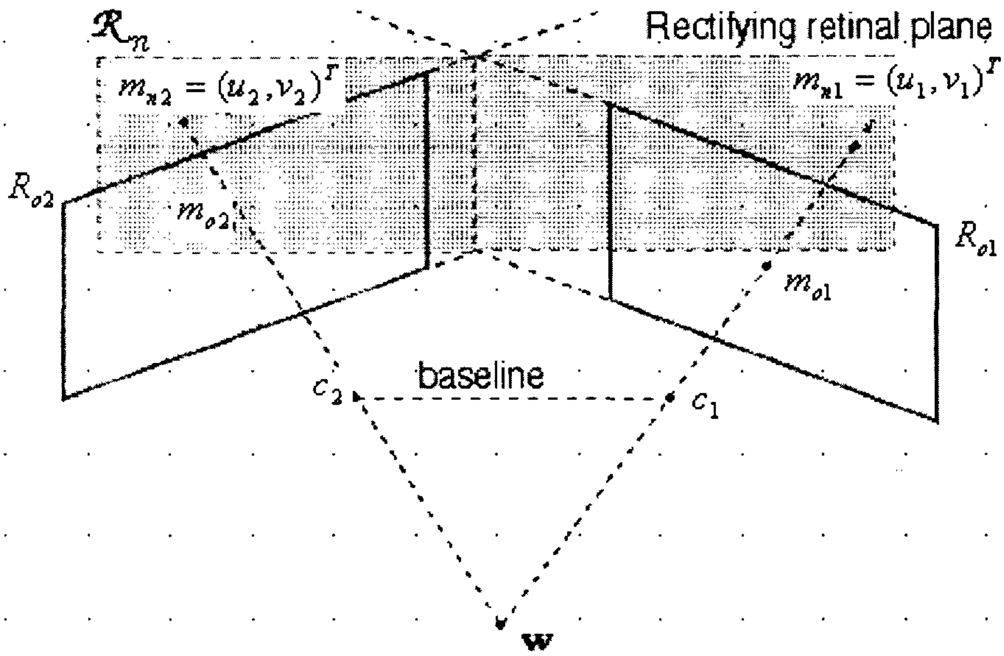

Fig.2.2 A Planar Stereo Rectification

2.2.1 Constraints for Rectification

Let the two perspective projection matrices for the rectification $\widetilde{P}_{n1}$ and $\widetilde{P}_{n2}$ be given by

$$\widetilde{P}_{n1} = \begin{bmatrix} a_1^T & | & a_{14} \\ a_2^T & | & a_{24} \\ a_3^T & | & a_{34} \end{bmatrix}, \quad \widetilde{P}_{n2} = \begin{bmatrix} b_1^T & | & b_{14} \\ b_2^T & | & b_{24} \\ b_3^T & | & b_{34} \end{bmatrix} \tag{2.13}$$

Then, two stereo cameras having $\widetilde{P}_{n1}$ and $\widetilde{P}_{n2}$ should have the same retinal plane and the same focal points of $\widetilde{P}_{o1}$ and $\widetilde{P}_{o2}$. This yields

$$a_3 = b_3\,,\quad a_{34} = b_{34} \tag{2.14}$$

$$\widetilde{P}_{n1}\begin{bmatrix} c_1 \\ 1 \end{bmatrix}=0, \quad \widetilde{P}_{n2}\begin{bmatrix} c2 \\ 1 \end{bmatrix}=0 \tag{2.15}$$

$$c_1=-P_{o1}^{-1}\widetilde{p}_{o1}, \quad c_2=-P_{o2}^{-1}\widetilde{p}_{o2} \tag{2.16}$$

From Equation (2.15) one derives

$$\begin{aligned} a_1^T c_1 + a_{14} &= 0, \quad b_1^T c_1 + b_{14} = 0 \\ a_2^T c_1 + a_{24} &= 0 \quad\ b_2^T c_1 + b_{24} = 0 \\ a_3^T c_1 + a_{34} &= 0, \quad b_3^T c_1 + b_{34} = 0 \end{aligned} \tag{2.17}$$

To satisfy $v_1 = v_2$, the equation below must hold.

$$\frac{a_2^T w + a_{24}}{a_3^T w + a_{34}} = \frac{b_2^T w + b_{24}}{b_3^T w + b_{34}} \tag{2.18}$$

Equation (2.18) yields

$$a_2 = b_2, \quad a_{24} = b_{24}. \tag{2.19}$$

The epipoles are obtained by projecting the camera focal points on the image plane as follows

$$\widetilde{e}_1 = \widetilde{P}_{n1}\begin{bmatrix} c_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a_1^T c_2 + a_{14} \\ a_2^T c_2 + a_{24} \\ a_3^T c_2 + a_{34} \end{bmatrix} \tag{2.20}$$

$$\widetilde{e}_2 = \widetilde{P}_{n2}\begin{bmatrix} c_1 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1^T c_1 + b_{14} \\ b_2^T c_1 + b_{24} \\ b_3^T c_1 + b_{34} \end{bmatrix}. \tag{2.21}$$

To make the epipolar lines parallel horizontally, the following constraints are imposed

$$\begin{aligned} &a_1^T c_2 + a_{14} \neq 0 \\ &a_2^T c_2 + a_{24} = 0 \\ &a_3^T c_2 + a_{34} = 0 \\ &b_1^T c_1 + b_{14} \neq 0 \\ &b_2^T c_1 + b_{24} = 0 \\ &b_3^T c_1 + b_{34} = 0 \end{aligned} \tag{2.22}$$

2.2.2 Rectifying Perspective Projection Matrices

The rectifying plane must be in parallel with the intersection of the two original retinal planes

$$a_3^T (f_1 \Lambda f_2) = 0 \tag{2.23}$$

where f_1 and f_2 are the third rows of $\widetilde{P}_{o1}$ and $\widetilde{P}_{o2}$ respectively. Equation (2.14) also yields $b_3^T (f_1 \Lambda f_2) = 0$.

For the rectifying plane to have a perpendicular coordinate system requires

$$a_1^T a_2 = 0, \quad ba_1^T a_2 = 0. \tag{2.24}$$

The origin (u_0, v_0) is obtained from

$$u_0 = a_1^T a_3, \quad v_0 = a_2^T a_3. \tag{2.25}$$

Setting $(u_0, v_0) = (0,0)$ yields

$$\begin{aligned} &a_1^T a_3 = 0 \\ &a_2^T a_3 = 0 \\ &b_1^T a_3 = 0 \end{aligned} \tag{2.26}$$

The width and length of each pixel in the rectifying plane can be given as follows

$$a_u = \|a_1 \Lambda a_2\|, \quad a_v = \|a_2 \Lambda a_3\| \tag{2.27}$$

$$\|a_1 \Lambda a_3\|^2 = a_u^2$$
$$\|a_2 \Lambda a_3\|^2 = a_v^2$$
$$\|b_1 \Lambda a_3\|^2 = a_u^2 \quad (2.28)$$

Using the equality $\|x \Lambda y\|^2 = \|x\|^2 \|y\|^2 - (x^T y)^2$, one can derive the following constraints.

$$\|a_1\|^2 \|a_3\|^2 = a_u^2$$
$$\|a_2\|^2 \|a_3\|^2 = a_v^2$$
$$\|b_1\|^2 \|a_3\|^2 = a_u^2 \quad (2.29)$$

The matrix can be normalized such that

$$\|a_3\| = 1 \text{ and } \|b_3\| = 1 \quad (2.30)$$

The above constraints can be regrouped into the following four categories

$$a_3^T c_1 + a_{34} = 0$$
$$a_3^T c_2 + a_{34} = 0$$
$$a_3^T (f_1 \Lambda f_2) = 0$$
$$\|a_3\| = 1 \quad (2.31)$$

$$a_2^T c_1 + a_{24} = 0$$
$$a_2^T c_2 + a_{24} = 0$$
$$a_2^T a_3 = 0$$
$$\|a_2\| = a_v \quad (2.32)$$

$$a_1^T c_1 + a_{14} = 0$$
$$a_1^T a_2 = 0$$
$$a_1^T a_3 = 0$$
$$\|a_1\| = a_u \quad (2.33)$$

$$\begin{aligned} & b_1^T c_2 + b_{14} = 0 \\ & b_1^T b_2 = 0 \\ & b_1^T b_3 = 0 \\ & \|b_1\| = a_u \end{aligned} \tag{2.34}$$

Each group in the above can be represented by a 3×4 linear homogeneous system written by

$$\begin{aligned} & Ax = 0 \\ & \|x'\| = k \end{aligned} \tag{2.35}$$

In Equation (2.35) x' is the vector composed of the first three entries of x and k is a given real number. One can solve the above system in top-to-bottom order. The solution has the form of the one parameter family $x = \alpha x_0$ where x_0 is a non-trivial solution and $\alpha = k / \|x_0'\|$

2.2.3 Rectifying Image Transformations

Two perspective projection matrices are given by

$$\tilde{P}_o = [P_o \mid \tilde{p}_o], \quad \tilde{P}_n = [P_n \mid \tilde{p}_n] \tag{2.36}$$

The 3D point is projected by the following equations.

$$\begin{aligned} & \tilde{m}_o = \tilde{P}_o \tilde{w} \\ & \tilde{m}_n = \tilde{P}_n \tilde{w} \end{aligned} \tag{2.37}$$

Equation (2.10) yields

$$w = c_o + \lambda P_o^{-1} \tilde{m}_o \tag{2.38}$$

From Equations (2.37) and (2.38) one has

$$\tilde{m}_n = P_n P_o^{-1} \tilde{m}_o \tag{2.39}$$

Using the linear transformation matrix $T = P_n P_o^{-1}$ in Equation (2.39), the image planes are rectified.

3. EXPERIMENTS AND CONSIDERATIONS

To rectify stereo images in our experiments, the fundamental matrix is first estimated, next self-calibration is performed, and finally the transform matrix to rectify the images is applied. To calculate the accuracy of these experimental rectification results, the vertical coordinate differences are computed.

Fundamental matrix estimation is performed using the technique described in [2][3]. Using this fundamental matrix the perspective projection matrices are estimated to derive the image transformation matrices for the stereo rectification.

The performances of self-calibrated rectification and camera-calibrated rectification are compared. The rectification for the controlled stereo data obtained by projecting random 3D points given is also performed.

3.1 Rectification Experiments on Controlled Stereo Data

Fifty randomly generated 3D points are projected on stereo image planes. The left projection matrix $P_{perspective1}$ is defined and the right projection matrix $P_{perspective2}$ is constructed by rotating and translating the left matrix. With respect to X, Y and Z axes, the rotations are 5, 32, and 19 degrees with translation -100mm, -20mm, and -30mm respectively. The pixel size is 0.02 mm, the focal length is 12.1mm and $(u_0, v_0) = (250,264)$.

$$Perspective1 = \begin{bmatrix} 605 & 0 & 250 & 0 \\ 0 & 605 & 264 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$Perspective2 = \begin{bmatrix} 3.5264e+002 & -1.5132e+002 & 5.3035e+004 & -6.8000e+004 \\ 2.7140e+001 & 5.9847e+002 & 2.7716e+002 & -2.0020e+004 \\ -5.2992e-001 & 7.3912e-002 & 8.4482e-001 & -3.0000e+001 \end{bmatrix}$$

Solving the system of Equation (2.35) yields

$$P_{rectify1} = P_{rectify2} =$$

$$\begin{bmatrix} -3.9731e+002 & 3.8186e+001 & -4.5466e+002 & 7.8081e-012 \\ -7.2398e+001 & -6.0052e+002 & 1.2829e+001 & 7.1010e-013 \\ -7.4459e-001 & 1.0385e-001 & 6.5939e-001 & 8.1868e-015 \end{bmatrix}$$

Using Equation (2.40) one can compute the rectified image coordinates for the left and right image points and the average difference of the vertical coordinates is 7.0415e-014.

Estimating the fundamental matrix from the perspective projected corresponding points yields

$$F = \begin{bmatrix} 7.4259e-002 & -3.0302e-001 & -1.6167e-001 \\ 1.0710e-001 & 6.2184e-002 & -7.3651e-001 \\ -1.7737e-001 & 7.8718e-001 & 3.2220e-001 \end{bmatrix}$$

Estimating the two perspective matrices $P_{denormal1}$ and $P_{denormal2}$ from F yields

$$P_{denormal1} =$$

$$\begin{bmatrix} 3.8564e+002 & -1.5544e+002 & -3.3180e+002 & 1.3878e-014 \\ 4.9341e-015 & 6.3084e+002 & 4.4800e+002 & -1.1185e-014 \\ -2.3138e-016 & -2.5870e-017 & 1.0000e+000 & -5.5511e-017 \end{bmatrix}$$

$$P_{denormal2} =$$

$$\begin{bmatrix} -1.8821e+002 & -2.2122e+002 & -6.1678e+002 & -4.6049e+002 \\ -1.1205e+002 & 7.3462e+002 & 2.7069e+002 & -1.3557e+002 \\ -5.8934e-001 & 2.7644e-001 & 1.2491e+000 & -2.0316e-001 \end{bmatrix}$$

From the above perspective, we can compute the following rectification matrices.

$$P_{rectift1} = P_{rrectify2} =$$

$$\begin{bmatrix} -3.8725e+002 & 5.2806e+001 & -1.4187e+002 & -8.8802e-012 \\ -1.0738e+002 & -6.1844e+002 & 6.2914e+001 & 0 \\ -3.2184e-001 & 1.5097e-001 & 9.3468e-001 & -5.2558e-017 \end{bmatrix}$$

In this self-calibration, the average vertical difference is 4.2011e-001. This increase is due to noise inserted in the estimations of fundamental matrix and perspective matrices.

3.2 Rectification Experiments on Real Stereo Images

For real stereo image data, pictures of a cube with a regular pattern were taken as shown in Fig. 3.1. The size of the image is 640×480. 72 corresponding feature points were manually selected to

(a) left image

(b) right image

Fig.3.1 Real Stereo Images

estimate the following fundamental matrix

$$F = \begin{bmatrix} -2.4908e-001 & 7.3327e-001 & 4.6246e+000 \\ -1.1954e+000 & -7.5577e-002 & 5.8587e+000 \\ -5.0593e+000 & -5.5484e+000 & 1.0067e+000 \end{bmatrix}$$

From the matrix, the following two perspective matrices may be estimated

$$p_{perspective1} =$$
$$\begin{bmatrix} 2.3863e+000 & -2.1148e+000 & -2.0986e+000 & -2.8103e-016 \\ 1.0249e-017 & 1.0119e+000 & 1.2064e+000 & -2.7756e-017 \\ -5.6285e-017 & -2.8722e-016 & 1.0000e+000 & -1.1102e-016 \end{bmatrix}$$

$$P_{perspective2} =$$
$$\begin{bmatrix} -4.6143e+000 & -1.1537e+000 & -3.6241e+000 & -5.3436e+000 \\ 2.9283e-001 & 9.5588e-001 & 1.1697e+000 & 6.2478e-002 \\ -3.3858e-016 & -6.9721e-016 & 1.000e+000 & -3.3307e-016 \end{bmatrix}$$

The above two matrices lead to the following rectifying matrices

$$P_{rectify1} = P_{Rectify2} =$$
$$\begin{bmatrix} -3.2800e+000 & 7.4819e-001 & -4.7434e+000 & -1.5596e-015 \\ -1.3184e+000 & 5.3532e+000 & 1.7560e+000 & -4.8866e-016 \\ -7.9370e-001 & -3.5704e-001 & 4.9251e-001 & 3.3768e-017 \end{bmatrix}$$

Calculating the average vertical coordinate difference in the way described in the previous section yields 2.7449e-001 that is comparable to the difference in the case of self-calibration given in Section 3.1.

4. CONCLUSIONS

In this paper, a planar rectification technique for stereo images using the fundamental matrix for the self-calibration was described. Applying this technique to controlled stereo data and real stereo images led to comparable rectification results with average vertical coordinate differences of 4.2011e-001 versus 2.7449e-001. However, the average difference is extremely small when precise perspective projection matrices are given by camera calibration. In the case of controlled stereo data, the average vertical coordinate difference was 7.0145e-014. This suggests that each stage of this technique be investigated for its proper numerical processing. Error propagation between stages could lead to high average vertical coordinate differences.

REFERENCES

1. A. Fusiello, E. Trucco and A. Verri. "Rectification with unconstrained stereo geometry", in A.F. Clark, editor, *Proceedings of the British Machine Vision Conference*, pp. 400-409, BMVA Press, September 1997.
2. Hyo Rim Jeong and Jong Soo Lee. "A Study on Planar Rectification", *Proceedings of KIEE*, August 2002.
3. Sung Hoon Eom and Jong Soo Lee. "A Study on the Fundamental Matrix Estimation using Nonlinear Minimization", *KIEE SP* No. 6 Vol. 38. pp. 59-66, 2001.
4. Stefan Carlsson, "Geometry and Algebra of Projective Views", Lecture Notes, Nov. 1997.
5. Q. Chen and G. Medioni, "Building 3-D human face models from two photographs", *Journal of VLSI Signal Processing Systems* February 2001 Volume 27 Issue 1-2, pp. 127-140, 2001.
6. Roger Mohr and Bill Triggs, "Projective Geometry for Image Analysis", a tutorial given at *ISPRS*, Vienna, July 1996.
7. M. Pollefeys, "3D Modeling from Images", tutorial notes organized in conjunction with *ECCV*, Dublin, Ireland, 26 June 2000.
8. F. Isgro, E. Trucco, "Projective Rectification without Epipolar Geometry", *Proceedings IEEE International Conference on Computer Vision and Pattern Recognition*, pp..94-99, Fort Collins, Colorado, USA June 1999.
9. Richard I. Hartley, "In Defense of the Eight-Point algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 19, No. 6, pp. 580-593, 1997.
10. S.Roy, J.Meunier and I.Cox, "Cylindrical Rectification to Minimize Epipolar Distortion", *Proc. IEEE conference on Computer Vision and Pattern Recognition*, pp. 393-399, 1997.
11. C. Loop and Z. Zhang. "Computing Rectifying Homographies for Stereo Vision", *IEEE Conf. Computer Vision and Pattern Recognition (CVPR'99)*, Vol. I, pp. 125-131, Colorado, June 1999.
12. F. Isgro, E. Trucco "On Projective Rectification", *Proceedings IEEE Conference on Image Processing and Analysis*, pp. 42-46, Manchester, England. July 1999.