Journal of Zhejiang University SCIENCE
ISSN 1009-3095
http://www.zju.edu.cn/jzus
E-mail: jzus@zju.edu.cn

A simple rectification method for linear multi-baseline stereovision system*

DU Xin (杜 歆)†, LI Hong-dong (李宏东), GU Wei-kang (顾伟康)

(*Department of Information and Electronics, Zhejiang University, Hangzhou 310027, China*)

†E-mail: duxin@dial.zju.edu.cn

Received Apr. 3, 2003; revision accepted Sept. 5, 2003

Abstract: The linear multi-baseline stereo system introduced by the CMU-RI group has been proven to be a very effective and robust stereovision system. However, most traditional stereo rectification algorithms are all designed for binocular stereovision system, and so, cannot be applied to a linear multi-baseline system. This paper presents a simple and intuitional method that can simultaneously rectify all the cameras in a linear multi-baseline system. Instead of using the general 8-parameter homography transform, a two-step virtual rotation method is applied for rectification, which results in a more specific transform that has only 3 parameters, and more stability. Experimental results for real stereo images showed the presented method is efficient.

Key words: Rectification, Multi-baseline, Stereovision
Document code: A **CLC number:** TP391.41

INTRODUCTION

Stereo rectification is a crucial processing for any practical stereovision system; and is aimed transform images so that the conjugate epipolar lines are aligned horizontally and in parallel. Varied and numerous rectification algorithms have been given in literature (Ayache and Lustman, 1991; Fusiello *et al.*, 2000; Hartley, 1999; Robert *et al.*, 1995; Mulligan and Kaniilidis, 2000; Isgro and Trucco, 1999). Most of them involve applying a homography transformation for that purpose.

Multi-baseline stereo is a special type of stereovision system architecture, which was first introduced by the CMU-RI group (Kanade, 1995; Kanade *et al.*, 1996). It has the advantage of reducing mismatches during correspondences due to the simultaneous multiple baselines, thus being able to produce a more statistically accurate depth value. It is especially applicable to images with repetitive texture regions frequently emerging in robot navigation scenes.

In implementing the multi-baseline stereo matching method, it is very necessary that all of the N cameras be rectified. Unfortunately, because traditional rectification methods are devised mainly for binocular cases, they cannot be directly applied to the multi-baseline system. Even though such methods are employed on each pair of cameras that belong to the same multi-baseline system, simultaneous rectification of all the cameras will not necessarily be guaranteed. For example, in a 3-camera case, we first rectify the left camera with respect to the central one, and then rectify the central one with respect to the right one. Because these two steps are independent of each other, the latter step will often destroy the result of the former step.

* Project (No. 60105003) supported by the National Natural Science Foundation of China

In other words, the final result will hardly be a simultaneous rectification if such a pair-wise processing method is adopted.

This paper presents a new multi-baseline rectification algorithm which can rectify multiple cameras simultaneously. We consider the multi-baseline system as a whole rather than many separate image pairs. In Section 2, we briefly introduce the mathematical principles in the rectification, and then propose a new multi-baseline rectification method in Section 3. Experimental results and conclusions in Section 4 and Section 5 show the efficiency of the proposed algorithm.

PREVIOUS METHODS AND MATHEMATICAL PRINCIPLES

Traditional binocular rectification

We first consider a generally configured binocular case wherein the two cameras' axes are not necessarily coplanar.

In principle, binocular rectification is realized by applying such kinds of image transformations to each image, which can not only bring both epipoles to infinity, but also align the epipolar lines horizontally. Homography transform is widely adopted to realize rectification.

There are a number of rectification methods that have been devised for the binocular stereo system. Some assume that the cameras are already calibrated; others can handle the un-calibrated case. Hartley and Gupta (1993) presented an elegant theory and a practical rectification method based on the idea that to find two homographic transforms that can correctly align the geometric configurations, and also minimize the extra image distortion due to the transform. Robert *et al.*(1995), Loop and Zhang (1999), Isgro and Trucco (1999) also provided some other methods for un-calibrated cases from many different considerations. One disadvantage of such un-calibrated methods is that there seem to be no apparent geometrical explanations for the homographic transform, so it is hard to apply such methods to the multi-baseline configuration. Fusiello *et al.*(2000) provided a very simple rectification method and a very compact code im-plementation. Their method is based on the fact that both rectified cameras' axes are parallel and perpendicular to the baseline. Therefore an Euclidean coordinate frame was constructed according to the following rules: (1) The new X axis is parallel to the baseline; (2) The new Y axis is orthogonal to X; (3) The new Z axis is orthogonal to the X-Y plane. Then rectification was realized by transforming each camera from its original orientation to the one that coincided with the new X-Y-Z system, which assigned them a new and identical intrinsic matrix at the same time. This method provided a clear geometric explanation for binocular rectification. However, without modification, it still could not be applied to the multi-baseline case. The reasons are given below.

Multi-baseline rectification

Now we consider the rectification problem with a linearly configured multi-baseline system as shown in Fig.1, where there are many cameras arranged in line.

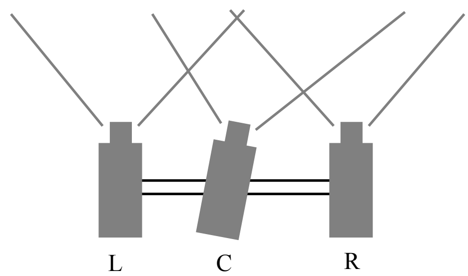

Fig.1 Linear multi-baseline system

Compared with the binocular case, there are much fewer papers in literature dealing with multi-baseline rectification problems. Williamson and Thorpe (1998) introduced a method for a trinocular case arranged in an L-shape rather than in line. He utilized the existing planar structures within images to obtain the appropriate homographic transforms that could rectify each pair of images (say, in a pair-wise manner). Mulligan and Kaniilidis (2000) also proposed a method for a trinocular stereo system of non-parallel configuration; and a pair-wise manner to rectify each pair of cameras. His method is especially useful in a 3D object reconstruction application scenario using a handheld camera, where the camera is revolving around the object, but is not particularly effective for the linear case, so he introduced an additional transform to warp the central camera image twice. Kang *et al*.(1994) introduced a similar method to rectify a linear 4-camera multi-baseline system, but they had to obtain six homographies computed for each pair of the four images.

The traditional binocular rectification methods cannot be directly applied to the multi-baseline case, because even if the binocular rectification method for each pair of cameras is employed, all the cameras cannot be guaranteed to be simultaneously rectified relative to a common frame.

In the next section we'll introduce a different idea to accomplish simultaneous multi-baseline rectification by a two-step virtual rotation method.

TWO-STEP VIRTUAL ROTATION METHOD FOR MULTI-BASELINE RECTIFICATION

From the above analysis of rectified cameras, it can be inferred that virtual rotations of cameras can be used for the purpose of rectification. The rotations should be based on their own optical center, around a suitable axis, and by a suitable angle.

Rotate ref-camera

Generally, the ref-camera's initial pose (i.e., its optical axis) is not necessarily perpendicular to the base line because of its inaccurate fabrication and assembly of a multi-baseline frame. For example Fig.2 illustrates a two-camera configuration, where O_l and O_r is the center of left and right camera respectively, and $\boldsymbol{l}$ and $\boldsymbol{r}$ are their axes. The vector $\boldsymbol{b}$ is the baseline vector. Without loss of generality we set the left camera as the ref-camera so that the ref-camera's initial axis is $\boldsymbol{l}$.

We intend to find a suitable rotation axis and a suitable rotation angle by which we can guarantee such perpendicularity and other important properties. At first, the angle is required to be as smaller as possible, because we know intuitionally that will

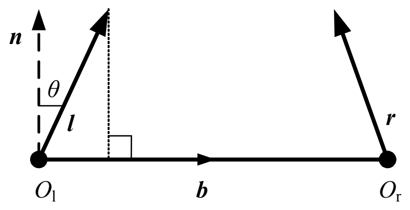

Fig.2 Binocular configurations

minimize extra rectification distortion in the process, as distortion minimization is an important criterion in rectification (Hartley, 1999; Mulligan and Kaniilidis, 2000). Secondly, we want to find one particular vector $\boldsymbol{n}$ perpendicular to the base line $\boldsymbol{b}$ that forms the smallest (minimum) angle (denoted by θ) between $\boldsymbol{n}$ and $\boldsymbol{l}$. In other words, we want to find the nearest (from $\boldsymbol{l}$) normal vector of $\boldsymbol{b}$. If such $\boldsymbol{n}$ is found, the suitable rotation axis is thus spotted.

Studying Fig.2, one can easily find that $\boldsymbol{n}$ can be calculated by an orthogonal projection of vector $\boldsymbol{l}$ onto vector $\boldsymbol{b}$, and the formula is $\boldsymbol{n}=\boldsymbol{l}-(\boldsymbol{l},\boldsymbol{b})\boldsymbol{b}/\|\boldsymbol{b}\|^2$. (The symbol $(,)$ is the inner product). The rotation angle θ is obtained by the inner product of $\boldsymbol{n}$ and $\boldsymbol{l}$, $\cos(\theta)=(\boldsymbol{n},\boldsymbol{l})/(\|\boldsymbol{n}\|\cdot\|\boldsymbol{l}\|)$, and the corresponding rotation axis (denoted by vector $\boldsymbol{q}$) can be described by the cross-product between $\boldsymbol{l}$ and $\boldsymbol{n}$:

$$\boldsymbol{q}=[q_1 \quad q_2 \quad q_3]=(\boldsymbol{l}\times\boldsymbol{n})/\|\boldsymbol{l}\times\boldsymbol{n}\|$$

The rotation matrix $\boldsymbol{R}^1$ on axis $\boldsymbol{q}$ by angle θ is given elsewhere in Zhang and Ma (2000).

After the above rotations, further rotation is still required in order to make the new X-axis be in the same direction as the baseline. It is easy to know that its rotation axis is just the new optical axis, and the rotation matrix (denoted by $\boldsymbol{R}_{\mathrm{img}}^1$) can be obtained in a similar method as above.

According to Hartley (1997), rotation of a camera with respect to its optical center will induce a special homography transform in the corresponding image space. Especially, when a camera (whose intrinsic parameter matrix is $\boldsymbol{K}$) is rotated by matrix $\boldsymbol{R}$, the corresponding image coordinate will be transformed from $\boldsymbol{m}$ to $\tilde{\boldsymbol{m}}$ as $\tilde{\boldsymbol{m}} = \boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^{-1}\boldsymbol{m}$. Therefore, rectification of the ref-camera image is given by $\tilde{\boldsymbol{m}}_1 = \boldsymbol{K}_1 \boldsymbol{R}_{\text{img}}^1 \boldsymbol{R}^1 \boldsymbol{K}_1^{-1} \boldsymbol{m}_1$.

Rotate other cameras and adjust their intrinsic matrices

After adjusting the ref-camera to the new pose $\boldsymbol{n}$, the second step is to rotate all the other cameras so that their optical axes are all parallel to $\boldsymbol{n}$. And then perform an image rotation on the new optical axis. Besides proper geometric configurations, having an identical intrinsic matrix is another necessary condition for a rectified system. So we have to adjust the other cameras' intrinsic parameters as well. This can be done by simple matrix manipulation as,

$$\tilde{\boldsymbol{m}}_i = \boldsymbol{K}_1 \boldsymbol{K}_i^{-1} \boldsymbol{K}_i \boldsymbol{R}_{\text{img}}^i \boldsymbol{R}^i \boldsymbol{K}_i^{-1} \boldsymbol{m}_i = \boldsymbol{K}_1 \boldsymbol{R}_{\text{img}}^i \boldsymbol{R}^i \boldsymbol{K}_i^{-1} \boldsymbol{m}_i$$

where $\boldsymbol{K}_1$ is the ref-camera's intrinsic matrix, and $\boldsymbol{K}_i$ the ith camera's intrinsic matrix.

Because every camera will undergo two steps of rotations during the rectifying process, we thus call the rectification method a two-step rotation method (In fact, in mathematics, these two-step rotations can be combined as a single 3D rotation, but for the sake of implementation, we would like to call it a two-step rotation).

EXPERIMENTAL RESULTS

Binocular rectification using our method

Although developed for multi-baseline system, our method is applicable to binocular case as well. For example, we may take the left camera as the ref-camera. First perform two rotations so that it might have a right attitude (pose), and then rotate the right camera to the right attitude and with the left-camera's intrinsic parameters. After that we are sure that they have been correctly rectified.

Figs.3a, 3b, 3a', 3b' show the original stereo images and the rectified images respectively, on which some horizontal lines are overlaid in order to demonstrate the result after rectification (A more professional way is to overlay epipolar lines. However, horizontal lines are feasible as well). It is interesting to notice that our results are different from what Fusiello *et al.*(2000) achieved, though the same image pair is used. That is because the considered transformations are different. Our method has avoided an arbitrary choice of the new Y-axis' direction.

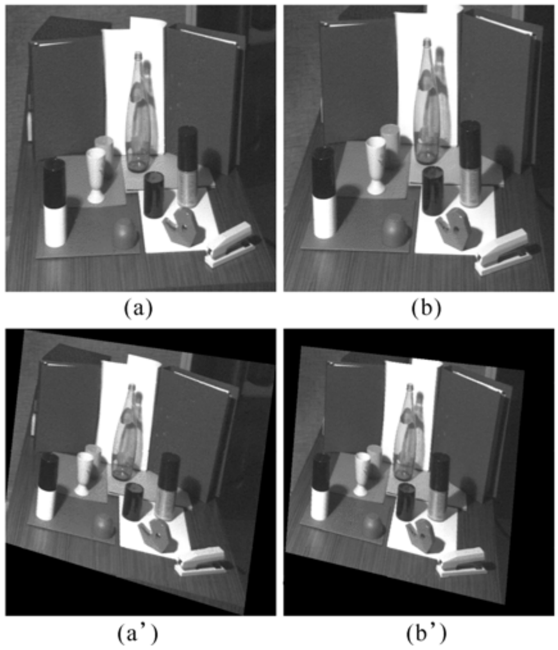

Fig.3 The original binocular images (a, b) and the rectified images(a', b')

Multi-baseline rectification using our method

We tested our method for a linear multi-baseline system with three cameras. Each was calibrated precisely, and the nonlinear lens distortions were removed before application of our rectification method. The three original images are shown in Fig.4a, and the rectified results are shown in Fig.4b, from which we can see that the epipolar lines of the three images are correctly aligned horizontally and matched; in other words, they have been rectified simultaneously. In addition, the new visions do not show great difference from their original counterparts.

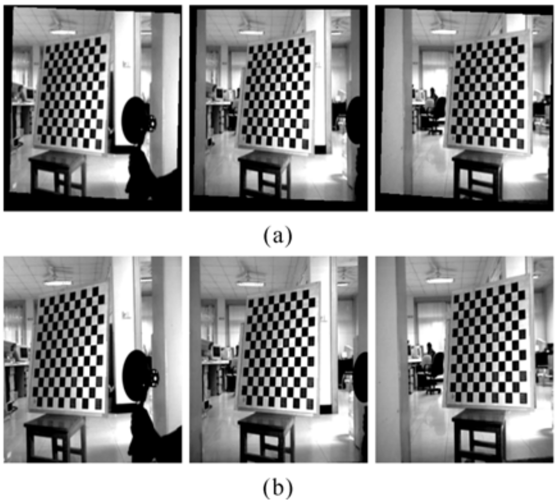

Fig.4 The original 3-camera images (a) and the rectified results (b)

CONCLUSION

A two-step rotational rectification method is proposed in this paper. It can be applied to the linear multi-baseline stereo scenario, where most traditional methods often fail. The proposed method is also applicable to a binocular stereo system, and has the advantage that distortion due to rectification process is minimized.

Another advantage of our method is that the geometrical transforms that each camera is undergoing during the process of rectification are precisely known, which is very useful for the following depth recovery calculations, otherwise one has to re-calibrate on the rectified images.

Our method will be ineffective in the case of a non-linear multi-baseline system, because pure rotation does not have the power of expressing the relative translations of cameras. If translation is small enough, some practical approximation measures could be taken for compensation. However, for more general cases, further research should be carried out.

ACKNOWLEDGEMENTS

The binocular stereo images used in experiments are from the INRIA-Syntim project.

References

Ayache, N., Lustman, F., 1991. Trinocular stereo vision for robotics. *IEEE T-PAMI*, **13**:73-855.

Fusiello, A., Trucco, E., Verri, A., 2000. A Compact Algorithm for Rectification of Stereo Pairs. Machine Vision and Application, USA.

Hartley, R., Gupta, R., 1993. Computing Matched Epipolar Projections. Proceedings of IEEE CVPR, p.549-555.

Hartley, R., 1997. Self-calibration of stationary cameras. *IJCV*, **22**(1):5-23.

Hartley, R., 1999. Theory and practice of projective rectification. *IJCV*, **35**(2):1-16.

Isgro, F., Trucco, E., 1999. Projective rectification without epipolar geometry. *IEEE-CVPR*, **1**:94-99.

Kang, S.B., Webb, J., Zitnick, C., Kanade, T., 1994. An Active Multibaseline Stereo System with Real-Time Image Acquisition. Tech. Report CMU-CS-94-167, Carnegie Mellon University.

Kanade, T., 1995. Development of A Video-Rate Stereo Machine. Proc. of International Robotics and Systems Conference (IROS'95), **3**:95-100.

Kanade, T., Yoshida, A., Oda, K., Kano, H., Tanaka, M., 1996. A Stereo Machine for Video-rate Dense Depth Mapping and Its New Applications. Proceedings of IEEE CVPR, p.196-202.

Loop, C., Zhang, Z.Y., 1999. Computing rectifying homographies for stereo vision. Proceedings of IEEE CVPR, **1**:125-131.

Mulligan, J., Kaniilidis, K., 2000. Trinocular Stereo for Non-parallel Configurations. International Conference on Pattern Recognition (ICPR), **1**:567-570.

Robert, L., Buffa, M., Hebert, M., 1995. Weakly Calibrated Stereo Perception for Rover Navigation. Proceedings of ICCV-95, p.46-51.

Williamson, T., Thorpe, C., 1998. A Specialized Multibaseline Stereo Technique for Obstacle Detection. Proceedings of IEEE CVPR, p.238-244.

Zhang, Z.Y., Ma, S.D., 2000. Computational Vision. Science Press, China (in Chinese).