Automatic Camera Calibration and Rectification Methods

Authors: Hsien-Chou Liao and Hong-Jhih Wu

Qualifications: Ph.D. and M.S.

Affiliation: Department of Computer Science and Information Engineering, Chaoyang University of Technology

Address: 168 Jifong E. Rd., Wufong Township Taichung County, 41349, Taiwan

E-mail: hcliao@cyut.edu.tw, s9527614@cyut.edu.tw

Tel.: +886-4-23323000 ext 4211

Corresponding author: Hsien-Chou Liao

Executive Summary

The Global Positioning System (GPS) system is popularly used in an outdoor environment. A GPS tracking system usually displays the location of a moving object on a digital map. It is difficult to realize its actual situation without visual information. Therefore, cameras are incorporated with a GPS in order to provide a new GPS-based visual tracking service (GPS-VT). The tracking of an object with a GPS receiver can be performed not only on the digital map but also on the real-time camera image. GPS-VT was proposed in our previous study. It relies on the camera calibration for the transformation from a GPS coordinate to the coordinate on the camera image. Five point correspondences, i.e., a GPS coordinate and its corresponding image coordinate, must be established to generate the intrinsic and extrinsic parameters of the camera needed for coordinate transformation. There are two drawbacks in calibrating a camera. One is the difficulty in establishing five point correspondences manually for an uncalibrated camera. The other is the inconvenience of recalibrating a calibrated camera after adjusting its field-of-view (FOV). In this paper, an automatic calibration method and a rectification method are proposed to overcome separately the above-mentioned drawbacks. For an uncalibrated camera, the GPS moving vectors of the operator are matched with those moving vectors of objects in the camera image to locate the *operator. The operator simply walks to five different locations and the point correspondences can be established automatically. For the rectification of a calibrated camera, two sets of feature points are extracted separately from the images before and after the adjustment of the FOV. The correlations of feature points of two images are analyzed by using a normal cross correlation (NCC) method to obtain the offsets of the adjustment in x and y directions. The offsets are used to rectify the image coordinate transformed from the GPS coordinate. Two experiments were also designed for both methods separately. For an uncalibrated camera, the locating errors are estimated and compared for the five point correspondences that are established manually or automatically. The difference in locating errors of manual and automatic calibration only totaled 2.6 pixels. The results show that the accuracy of the automatic calibration method is almost the same as that of the manual calibration. For the rectification of a calibrated camera, the differences in locating errors are estimated for various adjustment offsets in x and y directions. The errors of the proposed method totaled less than five pixels for the adjustment up to 200 pixels, without recalibrating the camera. The tests show that the result of the rectification method is close to that of manual calibration. Therefore, the two proposed methods overcome the problems of using the GPS-VT in an actual environment.*

Keywords: camera calibration, visual tracking, global positioning system, motion detection, feature point matching.

1. Introduction

The Global Positioning System (GPS) is popularly used in the outdoor environment. The positioning error of a differential GPS (DGPS) receiver can be less than a few meters. Currently, the GPS is widely used for various kinds of services, such as car navigation, tracking of children and the elderly, and so on. In general, the GPS coordinates of an object are displayed on the digital map. However, determining the actual situation of the object simply based on the mark on the map is insufficient. On the other hand, cameras are widely deployed in the streets and community, or even in certain homes, especially in urban areas. Visual information is very important and helpful for realizing the actual situation of an object. Therefore, cameras were incorporated with GPS to provide a new GPS-based visual tracking (GPS-VT) service in the previous study[1]. When an object moves into the field-of-view (FOV) of a camera, it is not only displayed on the digital map, but also marked in the real-time camera image.

In order to provide a GPS-VT service, camera calibration is an important step. In general, calibration methods can be roughly classified into three categories: calibration pattern, self-calibration, and special techniques[2-4]. The method based on calibration patterns requires a calibration object with a known 3D geometry. The intrinsic parameters of a camera are estimated by observing the object. For example, M. Wilczkowiak et al. proposed a calibration method based on the parallelepiped in the camera image[5].

The self-calibration method does not need prior knowledge of the 3D to 2D correspondences. When the intrinsic parameters are initialized, the extrinsic parameters can be estimated from the multiple views acquired by moving the camera. For example, F. Lv et al. proposed a calibration method by using a walking human[6]. When there is no static structure available in the camera image, the method is useful for completing the camera calibration.

For a calibration method using a special technique, it depends on some specific features in the scene for camera calibration. For example, R. I. Hartley[7] as well as J. M. Frahm and R. Koch[8] proposed a method using three different images taken from the same point in space with different orientations of the camera. Z. Zhang proposed a calibration method from perspective views of daylight shadows in a scene that gives only minimal geometric information determined from the images[9].

In our previous study, the camera calibration was modified from the method proposed by R. Y. Tsai[10-12], which belongs to the first category of calibration pattern. The prototype of the GPS-VT service, called GODTA (GPS-based Object Detection and Tracking Approach), was implemented to demonstrate that GODTA is feasible in an actual environment. However, two drawbacks still exist in camera calibration when applied to the large number of cameras deployed in the public area. One is the trouble in establishing the five point correspondences needed for calibrating an uncalibrated camera. The other is the recalibration overhead when adjusting the FOV of a calibrated camera. Therefore, in this paper an automatic calibration method and a rectification method are proposed for overcoming these drawbacks.

The remainder of the paper is organized as follows. Section 2 presents the camera calibration principle and the prototype, i.e., GODTA, of GPS-VT service; Section 3 presents the calibration and rectification method; Section 4 presents the experimental study; and Section 5 presents the conclusion and suggestions for future work.

2. The principle and prototype of GODTA

Tsai's primitive calibration method requires five point correspondences. One correspondence consists of one three-dimensional world coordinate in conjunction with the corresponding two-dimensional image coordinate. Five point correspondences are needed to estimate the intrinsic and extrinsic parameters, including the focal length (f), translation (T) and rotation (R) matrices.

The coordinate acquired from the GPS receiver is in a DMS (degree, minute, second) format. It cannot be directly used as the world coordinate in Tsai's method. Therefore, the coordinate in the DMS format is transformed to the format of the WGS84 (World Geodetic System 1984) coordinate system. The WGS84 coordinate is transformed to a TM2 (2-degree Transverse Mercator) coordinate which is popularly used in digital mapping. Therefore, one point correspondence used in GODTA consists of one TM2 coordinate and the corresponding image coordinate. The matrices T and R are shown in Eq. (1). For a new TM2 coordinate, (x', y', z'), the corresponding image coordinate, (u', v'), can be computed by using Eq. (2).

$$T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix},\ R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

$$\begin{aligned} u' &= f\,\frac{r_{11}x' + r_{12}y' + r_{13}z' + t_x}{r_{31}x' + r_{32}y' + r_{33}z' + t_z} \\ v' &= f\,\frac{r_{21}x' + r_{22}y' + r_{23}z' + t_y}{r_{31}x' + r_{32}y' + r_{33}z' + t_z} \end{aligned} \tag{2}$$

The prototype of GODTA is shown in Fig. 1. The location of an object can be displayed on a digital map and the object can be marked simultaneously on the real-time camera image. The real-time image on the right-hand side is easier to realize the actual situation of the moving object than a red dot on the digital map. The inaccuracy of GPS coordinates means that the transformed image coordinate cannot match the object location precisely. Some moving object detection techniques can be integrated into GODTA to mark the object with greater precision.

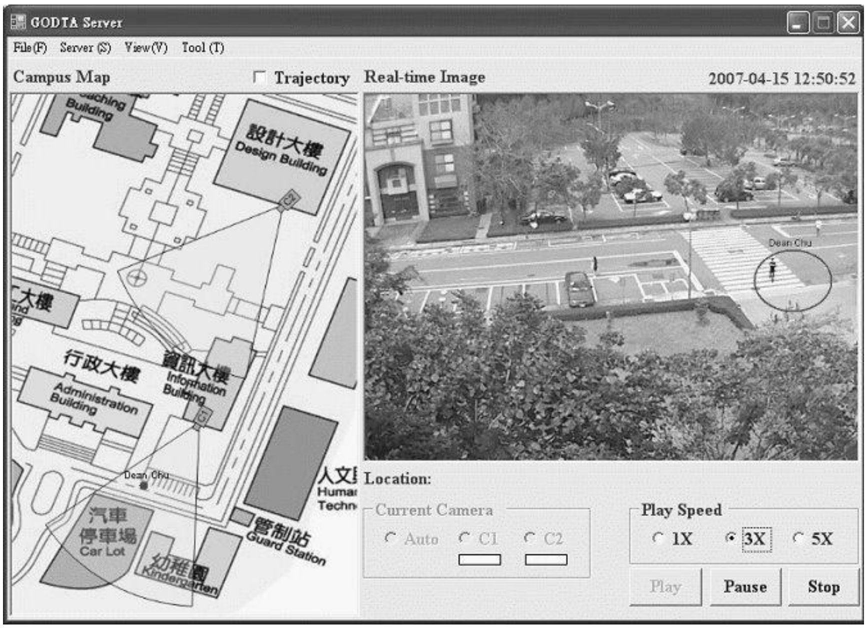


Figure 1: The screen shot of GODTA

3. Calibration and Rectification Methods

When the number of cameras is large, the camera calibration becomes a troublesome problem. It is necessary to increase the efficiency of the calibration procedure. There are two situations confronting camera calibration. One is to calibrate a new, i.e., uncalibrated, camera. The other is to re-calibrate a camera when its FOV is adjusted. Therefore, an automatic calibration method and a rectification method are designed for each of the above situations. They are presented in the following subsections.

3.1 The calibration method of a new camera

For the manual camera calibration in the previous study, two operators are needed to acquire the five point correspondences. The first operator carries the GPS receiver and walks into FOV of the camera. The second operator monitors the real-time camera image to record the image coordinate and the corresponding TM2 coordinate manually when the first operator is still. Obviously, the above procedure is time-consuming and should be automated.

In order to automate the above procedure, locating the operator is the first and primary task. The TM2 coordinates of the operator are matched with those moving objects in the camera image. However, the camera is unable to recognize the location or direction of the operator according to his/her TM2 coordinates before the camera calibration. Therefore, the vector angles (VAs) of successive moving vectors formed by the TM2 coordinates are matched with those formed by the image coordinates of moving objects. The principle is described by an example, as shown in Fig. 2. There are three curves of VAs in the figure. Two curves represent the VAs of TM2 and image coordinates of the operator. The third curve represents the VA of an unknown object. The maximum VA is 180 degrees. When an object moves in a straight line, the VA approaches zero degrees. Conversely, a large VA is found when the object turns around. By observing these three curves, the operator turns around twice and causes two peaks in the curves. The curves of VAs formed by TM2 coordinate and image coordinates have similar peaks, but cannot be found on the curve of the unknown object. According to these observations, the operator can be located by filtering out those moving objects not having a similar situation when the turnaround of TM2 coordinates is found.

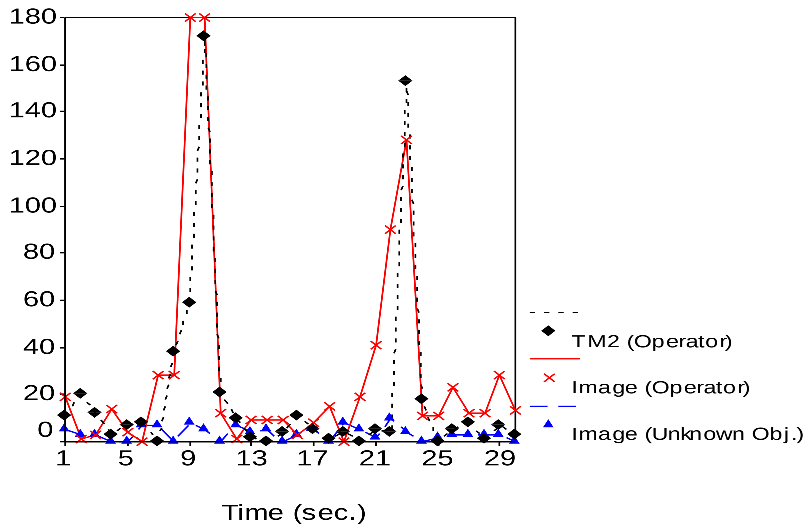


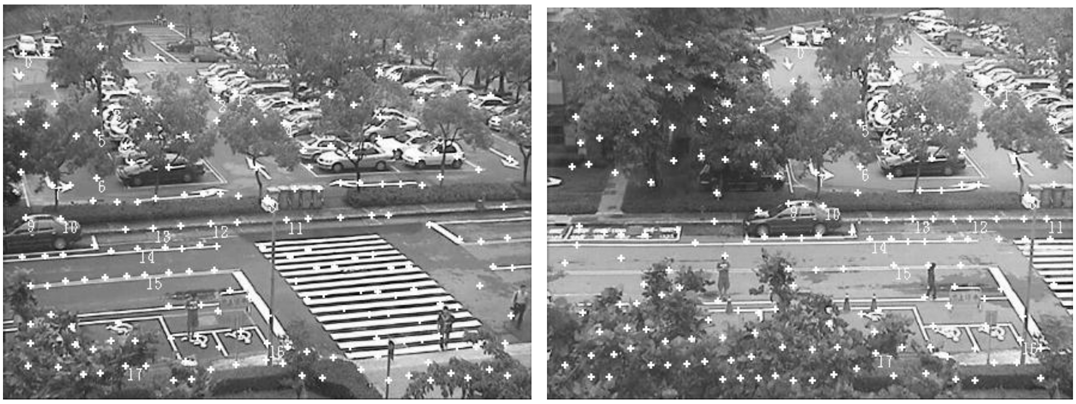
Figure 2: An example of VA curves

When the operator is located, s/he has to walk to five different locations and stop for a while. Five point correspondences can be recorded automatically and the calibration procedure is finished.

3.2 The rectification method of a calibrated camera

The adjustment of a calibrated camera is possible in order to obtain a better FOV. It invalidates the original five point correspondences. Therefore, a rectification method is designed in order to avoid the recalibration overhead for obtaining new point correspondences. The adjustment causes a displacement between the original image coordinates and latest one. Therefore, if the displacement of the image coordinate can be estimated, it can be used to rectify the original image coordinate. The rectification is based on the camera images captured before and after the adjustment. Two steps of the image mosaicing technique proposed by J. W. Hsieh are also used to estimate the displacement[13], the feature extraction and correspondence establishment. The original image (i.e., before the adjustment of FOV,) and the adjusted image are processed to extract the feature points. Two set of feature points are matched to establish their correspondence. The average displacement in the x-y-axis can be computed from all the correspondence. The displacement is used to revise the image coordinate transformed from the TM2 coordinate.

An example of feature extraction and correspondence establishment is shown in Fig. 3; the images in Figs. 3(a) and (b) are the original and adjusted image, respectively. The extracted feature points of two images are marked in white dots. Several correspondence points are found in common in both images. Every correspondence is labeled by the same number on both images, as shown in the figures. The average displacement of these correspondences is used to rectify the image coordinate. It is unnecessary to recalibrate the camera.



(a) (b)

Figure 3: An example of feature point extraction and correspondence establishment (a) original image (b) adjusted image

4. Experimental Studies

Two separate experiments are designed for the automatic calibration and rectification methods in order to evaluate their accuracy and efficiency.

4.1 Experiments with an uncalibrated camera

Experiments are designed to estimate the performance of the automatic calibration method, including the locating time of the operator, the locating successful rate, and the accuracy of the established five point correspondences. Three tools were implemented for the experiment. One is a client tool used by the operator for sending TM2 coordinates and acknowledging the calibration results. Another is a server tool for matching TM2 and image coordinates and establishing the five point correspondences. The third is a verification tool for evaluating the accuracy of the established five point correspondences.

The screen shots of the three tools are shown in Fig. 4. In Fig. 4(a), the real-time camera image is shown in the upper-left part. The image coordinates of moving objects and the received TM2 coordinates are listed in the lower-left part. When the operator is located by matching the TM2 and image coordinates, its image is shown in the upper-right part. The operator walks to five different locations, and the five point correspondences are established in turn, and shown on the right-hand side. In Figure 4(b), the client tool sends the TM2 coordinate to the server tool every second after clicking the “Start” button. When one point correspondence is generated by the server tool, it is transmitted to the client tool for the confirmation of the operator. The operator can restart the calibration procedure if the established point correspondence is not satisfactory. In Figure 4(c), the tool loads a set of testing images and the corresponding TM2 coordinate. One of the testing images is shown on the left-hand side. All the image coordinates transformed from the TM2 coordinates are listed on the right-hand side. A red circle, whose center is at the image coordinate transformed from the corresponding TM2 coordinate, is depicted on the testing image. The transformation is performed based on manual and automatic five point correspondences, separately, in order to evaluate the accuracy of the proposed method.

(a) (b) (c)

Figure 4: Screen shots of three tools (a) server tool (b) client tool (c) verification tool

By using the above tools, two results are used to evaluate the performance of the automatic calibration method. One is the elapsed time and successful rate of the server tool in locating the operator under various numbers of moving objects. The other is the locating error (LE) of the established five point correspondences. LE is the Euclidean distance between the actual coordinate and the center of the circle in Figure 4(c). After the experiments, the operator is located successfully 86 times among 105 tested. The successful location rate is 81.9 percent. The number of moving objects ranges from several to 50 persons. The server can still successfully locate the operator most of the time among 50 persons. The elapsed time on locating the operator is from 5 to 21 sec. The average time is 11.4 sec. It means the server tool can locate the operator within a short period.

For the second result of the LE, 30 testing images and the corresponding TM2 coordinates are used to compute the LE for the five point correspondences acquired manually or automatically. The average LE is 13.6 and 16.2 pixels for manual and automatic calibration, respectively. The difference of LE is only 2.6 pixels. It means that the accuracy of the proposed automatic calibration method is close to the manual calibration method.

4.2 The experiment with a calibrated camera

In this experiment, the FOV of a calibrated camera is adjusted in up, down, left, and right directions, separately. The adjustments are from 40 to 200 for every 40 pixels. The average LE of six testing locations is computed to evaluate the rectification method. The average LE is compared with that of manual calibration. The experimental results are depicted in Fig. 5. The y-axis represents LE difference, i.e., the difference of two average LEs based on rectification and manual calibration. The x-axis represents the values of left/right or up/down adjustment. Two curves show that the LE difference is less than three and five pixels for up/down and left/right adjustment, respectively. It means that the rectification method can provide enough accuracy without recalibrating the camera.

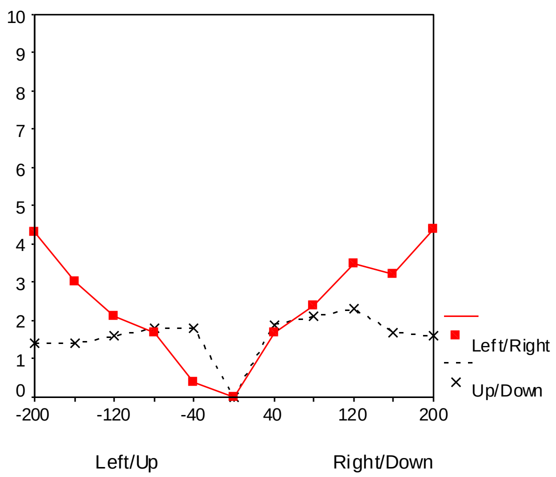


Figure 5: The experimental results of the adjustment of a calibrated camera

5. Conclusion and Future Work

In this study, an automatic calibration method and a rectification method are proposed separately for overcoming the problems in calibrating a new or calibrated camera. For the calibration of a new camera, a server tool can locate the operator among moving objects in the camera image according to the real-time GPS coordinates. Then the operator can move to five different locations to set the desired five point correspondences. A client tool is also designed for the operator to confirm every point of correspondence. The experimental results of the verification tool also show that the accuracy of the automatic calibration method is very close to that of manual calibration. That is, the proposed calibration method can overcome the inconvenience of manual calibration.

On the other hand, when a calibrated camera is adjusted manually, a rectification method is also designed based on the images before and after the adjustment. The feature points of two images are extracted and matched by using normalized cross correlation (NCC) to compute the displacement of two images. The displacement is then used to rectify the image coordinates transformed from the TM2 coordinates. The experimental results show that the LE difference is lowered within a few pixels under the adjustment up to 200 pixels. It means that the rectification method is useful to avoid the overhead of camera recalibration.

Two methods proposed in this paper are helpful in fulfilling the GPS-VT service in the practical environment. In our future work, GPS-VT service will be applied to track vehicles and will be operated via the active camera, i.e., pan/tilt/zoom (PTZ) camera, to provide a complete GPS-based visual tracking service.

Acknowledgement

This project was financially supported by the National Science Council under Grant No.: NSC 97-2221-E-324-043.

References

1. Liao, H. C., and Chu, P. T., *A Novel Visual Tracking Approach Incorporating Global Positioning System in a Ubiquitous Camera Environment*, Information Technology Journal, 8:4, 465-475, 2009.
2. Zhang, Z., *A Flexible New Technique for Camera Calibration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:11, 1330-1334, 2000.
3. Liao, H. C., and Cho, Y. C., *A New Calibration Method and its Application for the Cooperation of Wide-Angle and Pan-Tilt-Zoom Cameras*, Information Technology Journal, 7:8, 1096-1105, 2008.
4. Cao, X., and Foroosh H., *Camera Calibration Using Symmetric Objects*, IEEE Transactions on Image Processing, 15:11, 3614-3619, 2006.
5. Wilczkowiak, M., Sturm, P., and Boyer, E., *Using Geometric Constraints through Parallelepipeds for Calibration and 3D Modeling*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27:2, 194-207, 2005.
6. Lv, F., Zhao, T., and Nevatia, R., *Camera Calibration from Video of a Walking Human*,

IEEE Transactions on Pattern Analysis and Machine Intelligence, 28:9, 1513-1518, 2006.

7. Hartley, R. I., *Self-calibration from Multiple Views with a Rotating Camera*, Proceedings of the Third European Conference on Computer Vision, 1, 471-478, Stocklholm, Sweden, 1994.
8. Frahm, J. M., and Koch R., *Camera Calibration with Known Rotation*, Proceedings of the Ninth IEEE International Conference on Computer Vision, 2, 1418-1425, Beijing, China, 13-16 Oct. 2003.
9. Zhang, Z., *Camera Calibration with One-Dimensional Objects*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26:7, 892-899, 2004.
10. Tsai, R.Y., *A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses*, IEEE Journal of Robotics and Automation, 3:4, 323-344, 1987.
11. Lenz, R. K., and Tsai, R. Y., *Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 10:5, 713-720, 1988.
12. Shapiro, L. G., and Stockman, G. C., *Computer Vision*, Prentice-Hall, 444-456, 2001.
13. Hsieh, J. W., *Fast Stitching Algorithm for Moving Object Detection and Mosaic Construction*, Image and Vision Computing, 22, 291-306, 2004.