TRACKING CAMERA CALIBRATION IN MULTI-CAMERA SEQUENCES THROUGH AUTOMATIC FEATURE DETECTION AND MATCHING

Federico Pedersini, Augusto Sarti and *Stefano Tubaro*
Dipartimento di Elettronica e Informazione, Politecnico di Milano,
Piazza L. Da Vinci, 32 – 20133 Milano – Italy
Tel: +39-2-2399.3647; Fax: +39-2-2399.3413
e-mail: pedersin/sarti/tubaro@elet.polimi.it

ABSTRACT

When using multi-camera acquisition systems, in order to extract accurate 3D information on the viewed scene, the geometrical, optical and electric characteristics of the camera system must be known with good accuracy. If the acquisition of the sequence of multi-views requires a certain time to be completed, mechanical shocks, vibrations or thermal effects on the cameras and their support can cause a drift of the initial camera parameters. In this paper we propose a technique that, during the acquisition session, is able to track the camera parameters and, whenever possible, to correct them according to the occurred modifications. This technique does not need any a-priori knowledge or test objects to be placed in the scene, but exploits luminance features already present in the scene, such as luminance corners and spots.

1 Introduction

Multi-camera acquisition systems are today often employed for 3D scene reconstruction in a variety of applications such as industrial quality control, content creation for virtual reality applications, and 3D measurements.

The geometrical, physical and electrical parameters that characterize a multi-camera system (position and orientation, focal length, pixel size, etc.) must be known with very high accuracy in order to be able to extract accurate 3D information from the scene views. The process that determines such parameters is called camera calibration and must be performed at the beginning of an acquisition session as it is based on the analysis of the views of a special test object called "calibration pattern" or "target set" which is preliminarily placed in the space that will be later occupied by the objects to be reconstructed.

The stability of the initially estimated camera parameters can become a critical problem when acquiring a long video sequence. The camera calibration is, in fact, very sensitive to mechanical shocks, vibrations and even thermal changes on both the cameras and the supports. This camera parameter drift can easily cause significant 3D reconstruction errors, as the 3D back-projection process is quite ill-conditioned with respect to the camera parameters. In order to overcome this problem, we could employ very heavy and rigid camera supports, but such frames are usually very expensive and quite cumbersome to handle. For these reasons a more desirable solution is to detect and track any modification of the acquisition system and, whenever possible, to correct the camera parameters "on the fly". This way the calibration will hold accurate throughout the sequence.

The original goal of this work was the detection of camera parameter changes through the analysis of scene features. In order to achieve this goal, we developed a method for detecting, matching and tracking point-like features that are naturally present in the viewed scene. The method does not need any test objects to be placed in the scene or any *a-priori* knowledge, but exploits luminance features that are already present in the scene (e.g. corners and spots), which can be localized in the image with high precision. After the localization process, which is performed with sub-pixel accuracy, a matching procedure is applied to the n sets (n being the number of cameras) of feature points, obtaining a set of n-tuples of homologous points. The matched n-tuples are then back-projected into the 3D scene space. Since any camera parameter change causes an unexpected error (larger than the predicted pre-calibration's accuracy) in the back-projection process, we can reveal and characterize any occurred incidental modification of the camera parameters through a proper analysis of the magnitude and the temporal behavior of the back-projection error. This approach we extended in quite a straightforward fashion in order to be able to correct the calibration parameters "on the fly", if possible. In its current state, the proposed technique performs a parameter correction in two possible ways, depending on the situation. If the camera system remains still during the acquisition session, then it exploits the "fixed points" of the scene for re-calibrating the parameters, otherwise it "chases" the stable points of the scene and uses them to perform self-calibration.

The proposed technique has been tested on real sequences acquired with a trinocular camera system, with both simulated and real variations of the camera parameters, providing very encouraging results. In all the experiments the algorithm was able to detect the modification of the camera parameters. Moreover, in the presence of the parameter changes that typically occur when the camera system is exposed to accidental shocks, change of focal length, etc., the algorithm was able to measure the parameter drift and re-calibrate the system.

2 Global Approach

The proposed tracking technique consists of the following steps:

1. **Accuracy evaluation** — the validity of the camera parameters is checked through the analysis of the back-projection accuracy. This requires an accurate preliminary extraction and matching of the image features;

2. **Accuracy analysis** — the temporal evolution of the back-projection accuracy is analyzed in order to reveal an increment of the back-projection error that could likely denote a change in the parameters of the acquisition system;

3. **Parameter correction** — when needed, if a sufficient number of accurate matched image features is available, the calibration parameters is corrected through either calibration or self-calibration.

2.1 Feature extraction

The accurate detection of image features (to be used as control points) is often required in applications of 3D reconstruction [4]. Spot detection is encountered when dealing with features that have been artificially added to the scene, and can be performed through template matching. The method that we developed for detecting features that are naturally present in the scene looks for vertices (crossings between 3D edges) through an improved version of [1, 2, 3].

Vertices are characterized by the fact that the Laplacian of their luminance profile is zero. Furthermore, the Baudet operator

$$DET = \det \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix} = I_{xx}I_{yy} - I_{xy}^2$$

has a relative maximum (in all directions) in the proximity of vertices and, when applied to a set of progressively more filtered versions of the image, the maxima can be shown to lie on a line. Such two constraints can be jointly used for determining a vertex with super-resolution accuracy. In order to do so, we can look for the zero-crossing of the Laplacian along the line of the maxima of the DET. The improvements with respect to the original algorithm [1, 2, 3] consist of a refinement of the localization of the maxima of the DET through the estimation of a quadric, followed by a linear regression of the line over four measurements and an accurate detection of the zero crossing of the Laplacian along that line. The achieved results show that such improvements allow us to reach a localization accuracy that is better than 0.2 pixel.

2.2 Feature matching

Once the features are correctly extracted, we apply an n-partite matching algorithm in order to find the stereo-corresponding n-tuples. The matching criterion is based on the similarity of the local luminance profiles (correlation-based matching) and, in part, on the epipolar geometry defined by the current calibration (the calibration is not considered reliable in this application).

2.3 Analysis of the back-projection quality

In the ideal case, the visual rays of homologous points are expected to intersect in the 3D feature that originated them. Calibration inaccuracy, modeling failures and superimposed noise, however, prevents this from happening. The "quality" of the intersection (distance between visual rays) can be monitored in order to detect unexpected changes in the calibration parameters. In order to do so, an "accuracy index" is computed from the back-projection error of all matched triplets of points. The statistical distribution of this index over the matched points and its temporal behavior is then analyzed, in order to reveal any anomalous increment of the accuracy index that could very likely denote a change in the system parameters.

2.4 Parameter correction

Assuming that the camera system is not subjected to a rigid motion with respect to the scene throughout the acquisition session, at the beginning of the sequence the most accurate and stable (fixed) back-projected points are detected and used as *control points*. These are the candidate points to be used as 3D targets for parameter correction, provided that their number is sufficient. When a parameter change is detected, the current set of matched n-tuples of image features is used for recovering the new camera parameters. Depending on the previous knowledge of the 3D position of the matched points, the algorithm adopts either a calibration or a self-calibration approach. More precisely, if the 3D position of some points had been measured at the beginning of the sequence when the system was still calibrated, then recalibration is performed through a standard procedure that uses the available 3D points as markers. If, on the contrary, no reliable information is available about the actual 3D position of the matched points, the calibration can only be corrected through a self-calibration procedure. Self-calibration allows to simultaneously determine the camera parameters and the 3D position of the fiducial points. This method, however, requires a larger number of matched points for accurate results, as the self-calibration problem is much more ill-conditioned than the calibration problem. We are currently working on a modified version of the method that is able to determine a *rigid* (rather than *fixed*) set of points and perform calibration with respect to a relative (rather than absolute) frame.

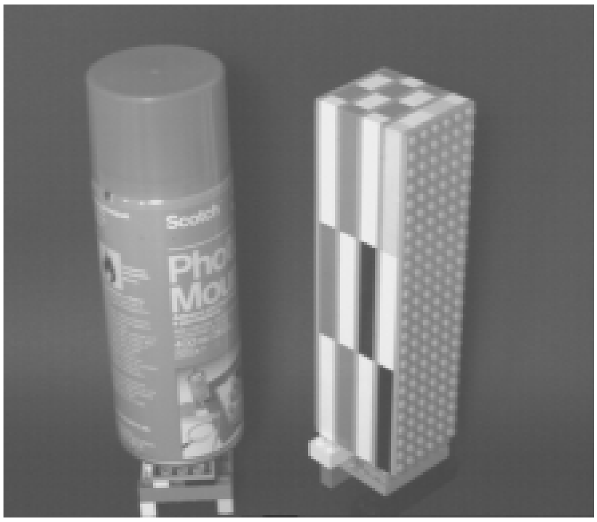
3 Experimental Results

In order to validate the proposed technique, we tested it on a trinocular system. Fig. 1 shows the two examples of scenes adopted for the experiment. Two types of tests have been performed on the acquired sequences: in the former case, the calibration parameters are artificially modified at a certain time instance in order to simulate a change in the acquisition setup. In the latter case, during the shooting of the sequence, the camera set-up is physically modified, by changing the relative pose of the cameras on their rigid frame and by slightly changing their focal length. Fig. 2 shows the result of the feature detection process, while Fig. 3 shows the resulting matched triplets.

The achieved results show that, in both cases, the system was able to immediately detect the changes in the camera parameters. This was revealed by a significant increase of the average accuracy index. It is worth noticing that the number of matched points was not significantly reduced by the parameter change, therefore we could use most of the matched points as fiducial points and correct the calibration parameters using such points. In order to test the accuracy of the corrected calibration parameters, a new feature matching was performed and the accuracy index was again evaluated and compared with the initial one. The results, collected in Table 1, confirm that the accuracy of the corrected parameters is comparable with that of the original calibration in both cases of re-calibration and self-calibration.

Method	av. dist. [mm]	std. dev. [mm]
std. calibration	0.0520	0.0240
recalibration	0.1016	0.0400
self-calibration	0.1020	0.0401

Table 1: Error in the estimation of the reconstructed calibration points.



(*a*)

(*b*)

Figure 1: Original images, taken with a trinocular acquisition system.

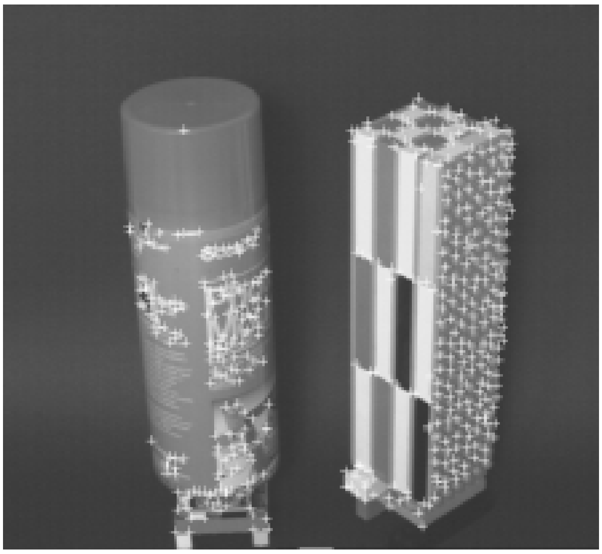
4 Conclusions

In this paper we proposed a technique for *tracking* the camera parameters through the analysis of luminance features that are naturally present in the scene. The method is based on subpixel feature localization, followed by feature matching. The accuracy of the back-projection of homologous features onto the 3D space is used as an index of quality for deciding whether to proceed with the correction of the parameters of the acquisition system. The proposed technique was proven effective through tests performed on real sequences acquired with a trinocular camera systems. Successful experiments were conducted with both simulated and real camera parameter drift, proving the method suitable for adaptive calibration.

Further research is being conducted in order to improve of the accurate feature detection strategy and add features of different nature. We are also focusing on self-calibration without any calibration target-set.

References

[1] L. Kitchen, A. Rosenfeld, "Gray-level corner detection", Pattern Recognition Letters, No. 1, 1982, pp. 95-102.



(a)

(b)

Figure 2: Detected image features.

[2] G. Giraudon, R. Deriche, "On corner and vertex detection", Proceedings Intl. Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, June 1991, pp. 650-655.

[3] K. Rohr, "Recognizing Corners by Fitting Parametric Models", Intl. Journal of Computer Vision, Vol. 9, No. 3, 1992, pp. 213-230.

[4] Z. Zhang, R. Deriche, O. Faugeras, Q.T. Luong, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry", INRIA, Report No. RR 2273, 1994.

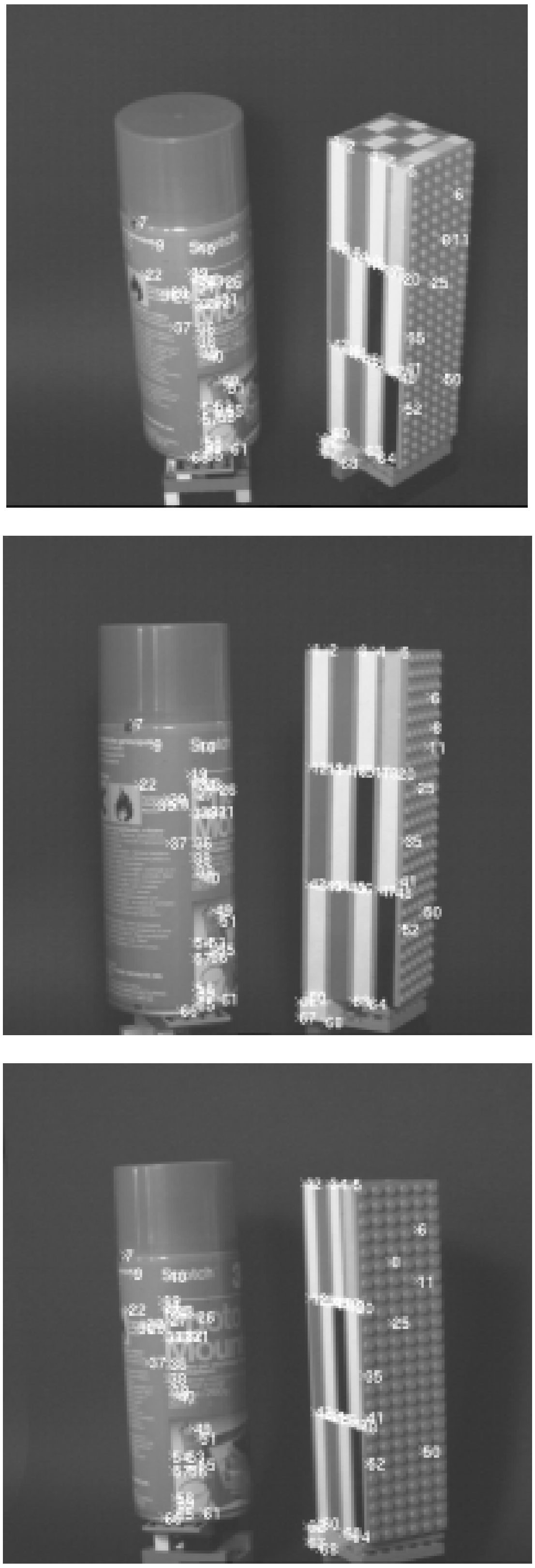


Figure 3: Set of matched triplets of image points.