The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Calibration of vision systems based on pseudo-random patterns

Chadi ALBITAR, Christophe DOIGNON and Pierre GRAEBLING

Abstract— **Solving visual features' correspondence and dealing with missing data are two factors of limitations for points registration techniques. To tackle this problem, we conceived a pattern, primarily designed for structured lighting vision systems, which may also be used for camera calibration purposes. The pattern design previously presented in [1] provides a huge of benefits. Among them, we firstly present a new calibration technique of a structured lighting system and secondly an automatic distortion compensation based on a printed pattern. These two well-known issues are very useful in 3D vision-based metrology with range data, for instance for model-based visual robot control, especially when the model is incrementally built with a real-time 3D reconstruction of moving surfaces. Perhaps, one of the most significant profit with a high Hamming distance pattern is the ability to reliably decode its projected individual elements even if several of items are missing, as it greatly extends the range of measurements volume.**

A technique which solves the distortion parameters by means of a robust M-estimator algorithm is presented. It uses a printed pattern and it allows the distortion be corrected with a single view and without the computation of other (intrinsic/extrinsic) parameters, even in presence of occlusions. Experimental results, in one hand by means of a printed pattern for the distortion compensation of a rigid endoscope and on the other hand by means of a projected pattern for the calibration of the structured lighting system, show very good performance for the 3-D reconstruction.

I. INTRODUCTION

For three decades, the field of vision-based robotics has been widely grown, and more and more complex 3D scenes are within robot reach due to deeper understandings of the scene perception, the increase of computer capabilities and control theory. For its wide field of applications, surfaces reconstruction is one of a very important topic in computer vision. This problem is related to structure from motion, stereovision, pose determination and so on, which results in applications including object modeling, mobile robot navigation and localization, environments building. Calibration is then a key step when accurate measurements of the 3D shape of an object or its 3D localization are required. Among the 3D shape measurement techniques, structured-light-based techniques are very efficient due to their fast speed and noncontact nature. Compared to a classical stereo vision system, structured lighting differs in that it solves the fundamentally problem of matching by replacing one camera by a light pattern projector. When the whole vision system is projectively identified, correspondences between views may be easily found for stereo vision or with an exhaustive search in the pattern array for structured light vision systems. In the latter case, the pattern design is a crucial step for quickly solving features' correspondence with neighborhood techniques, a class of techniques used for capturing the 3-D shape with a single shot. Therefore it can be applied for dynamic scenes. The matching problem is then closely related to the choice of the pattern and to the associated coding strategy since finding the correspondences depends on the ability of the decoding stage to locate the pattern elements in the image [12].

This work is supported by the Scientific Council of Strasbourg University

Ch. Albitar, Ch. Doignon and P. Graebling are with the Control, Vision and Robotics Team (LSIIT - UMR CNRS 7005), University of Strasbourg, Boulevard Brant, 67412 Illkirch, France, name@lsiit.u-strasbg.fr

A. Motivations and related works

We are currently being developing an endoscopic range finding system using a low-power laser beam illuminating the pattern. This system is designed so as to be used for robotic assistance in minimally invasive surgery. Except for robotic tasks concentrated on area close to the image centre, the distortion effect of such wide-angle camera lens is very significant in many situations, that is estimation errors could be very large for ulcer size determination or for surgical needle localisation for instance.

As the 3D reconstruction accuracy strongly depends on that of the calibration of the vision system, we seek a procedure which compensates for non-linear projections and outliers detection. Usually, the calibration of a vision system with structured lighting system requires two major stages. The first one is the camera(s) calibration and the second one is corresponding both to the identification of the position/orientation of the projector wrt to a common reference frame but also for its internal parameters. In the sequel, we provide a brief state-of-the-art mainly for structured lighting systems calibration but we also referred to some recent works about distortion camera evaluation since there is a vast literature about camera calibration and many techniques developed in the last two decades can be applied (see also [11] for a review).

Nakamura *et al.* [5] have designed an active scanning system with structured lighting for the reconstruction of 3-D intraoperative local geometry of pointed organs. With a 2-D galvano scanner and two cameras, a real-time registration of the area of interest was performed in order to alleviate the surgeon to mentally estimating the depth. A robot vision system that automatically positions a surgical cleaning instrument with a stationary camera mounted onto a long pipe laparoscope has been carried by Krupa *et al.* [6]. To this end, a laser pointer has been designed to emit 4 markers on the organ's surface. The distance estimation between the pointed organ surface and the tip of the instrument

was based on the projective invariance of the cross-ratio with a set of collinear active markers sticked on the instrument shaft and aligned with laser markers. With these two related works in the field of image-guided robotized surgery and active vision, the tracking behaviour and the positioning accuracy were disturbed by the high radial distortion effect with endoscopic devices. In both case, the distortion was corrected with one of the early techniques which simultaneously estimate the camera placement [11]. Hence, the video information related to the distortion may be disturbed by both perspective distortion, noise and the focal length value.

The distortion can be seen as a measure of the deviation of the imaged pattern elements with respect to those projected according to the pinhole model. The design of an extra geometric transformation which incorporates non-linear effects should lead to shrink the distance which separates the real projection from a true perspective. Then, a suitable distortion correction could be expressed as a prior image warping with a faithfully color interpolation.

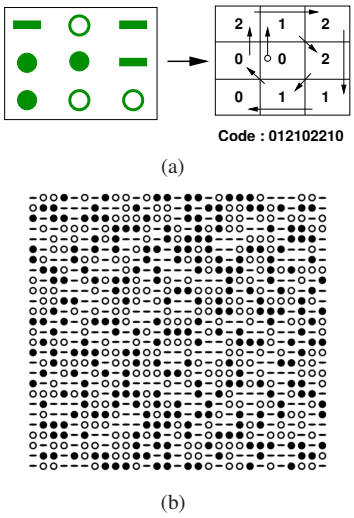


Fig. 1. Pattern of 783 elements based on central-symmetric perfect maps. (a) Definition of the codeword of each (3×3) submatrix. (b) The whole pattern of size (27×29) built with only three distinct geometrical features.

II. The coded structured lighting

Among the 3-D shape measurement techniques, structured light-based techniques are very efficient due to their non-contact nature. Compared to stereovision, structured light differs in that it solves the matching by replacing one camera by a light pattern projector. Structured light techniques can be classified, according to the strategy of encoding the pattern, into three categories [12]: time multiplexing, direct coding and spatial neighborhood. Time multiplexing can achieve a good accuracy and a high resolution, however it is unsuitable with dynamic scenes since multiple patterns must be projected successively. Direct coding techniques provide a good spatial resolution but their applicability is limited due to their sensitivity to noise and to light variations.

A. The neighborhood-based techniques

Techniques based on the spatial neighborhood strategy tend to integrate the entire coding scheme in a unique pattern. Therefore it can be applied for dynamic scenes. The coding must alleviate any ambiguity while identifying the components and must also increase the robustness of the decoding algorithm, even in the presence of occlusions. The code of each pattern component depends on its value and on those of its neighbors. To tackle the decoding of a partially visible pattern, some authors adopted the theory of Perfect Maps (PM) to encode a unique pattern ([9], [14], [7], [8]). If A is a $(m\times n)$ pattern matrix in which each element is taken from an alphabet of k symbols, and if A has a window property so that each $(p\times q)$ submatrix a_{pq} appears exactly once, then A is named a Perfect Map. If A does not contain the submatrix filled with 0's, it is called an M-array. The robustness of the methods adopting the projection of such patterns is due to their capability to decode the visible parts of the observed pattern thanks to the properties of the M-array's. Since we are interested in surfaces reconstructions of non-structured scenes as it is the case in medical endoscopy, we have to face with the classical problems of shadows and occlusions. In such situations, a key factor is to design an encoded pattern with a significant Hamming distance. The Hamming distance h quantifies differences between codewords. If this distance is higher than 1, some errors may be corrected. It was shown in [7] that we can obtain better results when the Hamming distance is higher than 3. However, most of referenced methods used such a distance but with more symbols as it is the case in [7] with $k=8$ symbols. To represent the symbols, some authors use gray levels or colors [8], [3]. For the application of concern, and after the study of the light-organs interaction, we chose to associate a geometrical feature to a symbol to generate a monochromatic pattern.

B. Design of the proposed pattern

The feasibility of building a pattern with a small h value depends on m, n, k and $p\times q$. To reduce the search space of each codeword location during the decoding stage, it is suitable to use the smallest value of k, as there are k^{pq} distinct codewords, and that $(m-[p/2])(n-[q/2])$ words are necessary to build A. Therefore, a lower bound of k can be defined as

$$k \geq (m-[p/2])(n-[q/2])^{(1/pq)} \ . \qquad (1)$$

In [7], the authors proposed an algorithm to generate M-arrays given m, n, k, h, and $p\times q$. The algorithm generates a M-array with fixed properties using a bruteforce approach. For example, to generate a M-array using three symbols with a window property of 3×3, we start by choosing a 3×3 subarray randomly. Then a (1×3)-column of random elements through the alphabet is added to the right of this initial subarray followed by adding random rows of (3×1) beneath the subarray maintaining the integrity of the window property and the Hamming distance between the resulted windows. The horizontal and vertical processes are

repeated while the starting coordinates is being incremented by one, until the entire array is filled. For our application of interest, we have used a similar algorithm where we imposed additional constraints. Besides the uniqueness of each (3×3) window, the pattern matrix must have a central symmetry for an efficient realization with diffractive optical elements. To increase the Hamming distance, the acceptance test has been extended to codewords with missing upper corners. This test led to generate a pattern with an average Hamming distance of $\bar{h}=6.173$ [1]. The pattern codewords are all of length 9 $(p=q=3)$ and only $k=3$ simple geometric features (disc= 0, circle= 1 and stripe= 2) have considered; the latest one carries a directional information, thus it simplifies the neighbors detection and speeds up the decoding. A codeword C_{ij} is defined from A by the shape value at (i,j) and the shape value of its 8 adjacent neighbors (see Fig. 1-a).

III. Calibration

A. Calibration of the structured lighting vision system

1) A global approach: The calibration of the overall active vision system could be tackled with the projector thought as a camera device, but with a constant view. This the case in [4] and [10]. Considering the perspective projection model, $\mathsf{P}^c=\mathsf{K}^c\ [\mathsf{I}\mid\mathbf{0}]$ such that $m^c\equiv\mathsf{P}^c\,\mathbf{M}^c$ and the upper triangular (3×3) matrix $\mathsf{K}^c=(\mathbf{k}_1^c,\mathbf{k}_2^c,[0\ 0\ 1]^\mathsf{T})^\mathsf{T}$ of camera parameters. This modeling leads to seek a solution for

$$\underbrace{\begin{bmatrix}(\mathbf{k}_1^c)^\mathsf{T}-\begin{bmatrix}0&0&u_{ij}^c\end{bmatrix}\\(\mathbf{k}_2^c)^\mathsf{T}-\begin{bmatrix}0&0&v_{ij}^c\end{bmatrix}\\(\delta_j^L)^\mathsf{T}\mathsf{K}^L\mathsf{R}_{LC}\end{bmatrix}}_{\mathsf{C}}\tilde{\mathbf{M}}_{ij}=-\begin{bmatrix}\mathbf{0}^\mathsf{T}\\\mathbf{0}^\mathsf{T}\\(\delta_j^L)^\mathsf{T}\mathsf{K}^L\end{bmatrix}\mathbf{t}_{LC}\ . \tag{2}$$

where $\tilde{\mathbf{M}}_{ij}$ are the 3D (inhomogeneous) coordinates of the projected element in position (i,j) in the matrix pattern, $m_{ij}^c=(u_{ij}^c,v_{ij}^c)$ its image coordinates, K^L a projector intrinsic parameters matrix (see Fig. 2-a). C is a rank-3 (3×3) real matrix. With a distortion-free laser model, the third row of C is computed with points M_{ij} of identical index j. Such points lie on a common plane π^j defined by the j^{th} pattern row δ_j^L, K^L and the laser centre. Then, δ_j^L and M_{ij} are related with $(\delta_j^L)^T\mathsf{K}^L(\mathsf{R}_{LC}\tilde{\mathbf{M}}_{ij}+\mathbf{t}_{LC})=0$ in the camera frame. Once the camera calibration is done, one needs to estimate only the epipole and the homography of a plane at infinite distance, H_∞, not the transformation $(\mathsf{R}_{LC},\mathbf{t}_{LC})$ between the projector frame (R_L) and the camera frame (R_C), since in (2) $\mathsf{H}_\infty\ \mathsf{K}^c=\mathsf{K}^L\ \mathsf{R}_{LC}$ and the epipole $\mathbf{e}\equiv\mathsf{K}^L\ \mathbf{t}_{LC}$. In practice, this can be achieved with two sets of printed patterns on two non-parallel world planes on which laser beams are simultaneously projected. Note that both printed and projected patterns may be pseudo-random array. The major problem with such approach is its sensitivity to non-linear effects coming from the projector lens. That's why we present in the next paragraph a different solution.

2) The proposed technique: We propose to calibrate each pattern element individually. In an earlier work, Zhou [16] has considered the calibration of each stripe line. In comparison to this work, we are dealing with each element and its own codeword. Furthermore, our technique takes into account the distortion of the projector lens and missing data which may occur in several images. To do so, a calibration plane is used both for camera calibration purposes but also to collect the projections of all pattern elements. Since the camera-to-plane transformation can be retrieved, the three coordinates of any point on this plane can be estimated provided it is visible. This the case for the intersections of each pattern element beam with the calibration plane. These intersections are all lying on the same line and are corresponding to an irregular sampling of the line, it doesn't matter whether the pattern beam element has been previously distorted or not through the projector lens. Therefore, we use this interesting property to conceive an algorithm.

For each attitude of the calibration plane, two images are acquired. The first one, with the projector switch off, is used to compute the plane parameters embedded in a 4-vector $\pi=(\mathbf{r}_3^\mathsf{T},-d_\perp)$ where $\mathbf{r}_3$ is the third column of the rotation matrix R_{co} between the camera frame and the plane frame ($x-y$ plane) and $d_\perp=\mathbf{t}_{co}^\mathsf{T}\mathbf{r}_3$ is the signed orthogonal distance from the calibration plane to the camera origin. $\mathbf{t}_{co}$ is the position vector between the frame origins. The second image is used to capture each imaged pattern projection $\mathbf{m}^\pi=(u^\pi,v^\pi,w^\pi)^\mathsf{T}$ when the projector is turn on. As the projection is issued from a 3-D point $\mathbf{M}^\pi=(X^\pi,Y^\pi,Z^\pi,1)^\mathsf{T}$ on the plane π, it must verified $\pi^\mathsf{T}\mathbf{M}^\pi=0$. Expressing the perspective projection with $\mathbf{m}^\pi\times\mathsf{P}^c\ \mathbf{M}^\pi=0$, it comes $Z^\pi=\frac{d_\perp}{\mathbf{r}_3^\mathsf{T}(\mathsf{K}^c)^{-1}\mathbf{m}^\pi}$. Each pattern line L_{ij} is then defined by a set p intersection points $\mathbf{M}_{ij}^{\pi(k)}$, the superscript $k=1,\cdots,p$ is referred to the k^{th} image, hence the p transformations $(\mathsf{R}_{co},\mathbf{t}_{co})$, then the p attitudes of the calibration plane. Finally, with any two distinct points $\mathbf{M}_{ij}^{\pi(k)}$ and $\mathbf{M}_{ij}^{\pi(l)}$ on L_{ij}, one may built the Plückerian matrix L_{ij}, a (4×4) rank-2 real matrix as (subscripts i and j are dropped for clarity)

$$\mathsf{L}=\mathbf{M}^{\pi(k)}\mathbf{M}^{\pi(l)\mathsf{T}}-\mathbf{M}^{\pi(l)}(\mathbf{M}^{\pi(k)\mathsf{T}}=\begin{bmatrix}[\mathbf{w}]_\times&-\mathbf{v}\\\mathbf{v}^\mathsf{T}&0\end{bmatrix}\ . \tag{3}$$

With more points, L may be evaluated with the Least Mean Squares technique, that is with $p\geq2$ geometrical primitives, one per image, emanating from the same pattern element. One can notice that the method can handle missing data for several attitudes, when for instance some points are out of the camera field-of view. The only condition is to capture at least two geometrical primitives with the same codeword in the set of p images, then it allows to consider a large range of position and orientation values of the calibration planes. The set of dual Plückerian matrices $\mathsf{L}_{ij}^\star$ (obtained directly from L by permuting the two vectors $\mathbf{v}$ and $\mathbf{w}$) are representing the calibration of the projector with respect to the camera frame. To conclude, since any 3-D point belongs to this line if $\mathsf{L}^\star\mathbf{M}=0$, the back-(perspective) projection rule is used for computing the 3D coordinates

$$\mathbf{M}=Z\begin{bmatrix}(\mathsf{K}^c)^{-1}\\\mathbf{0}^\mathsf{T}\end{bmatrix}\mathbf{m}\ +\ [0\ 0\ 0\ 1]^\mathsf{T}\ . \tag{4}$$

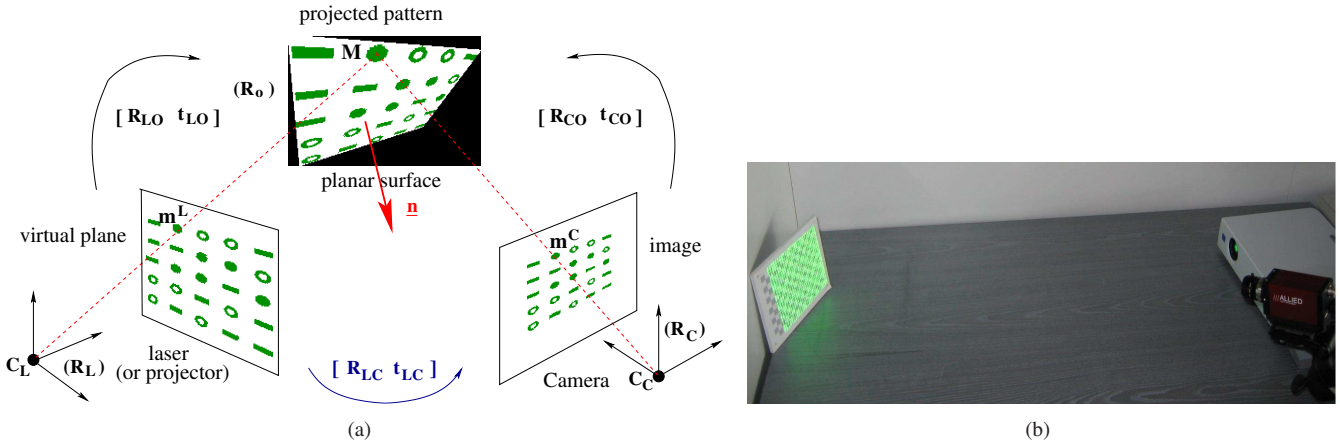


Fig. 2. (a) Geometrical description of the vision system with a structured light coded (5×5) upper-left sub-matrix of the original (flipped) pattern in Fig 1-b. (b) The experimental setup with a digital camera and a video-projector. The camera captures two images; the first one with the projector off is used to acquire the calibration plane texture (corners of the checkerboard). The second one, when the projector is turn on, grabs the images of the in field of view projections of the structured light pattern elements.

B. About the distortion model

In this part, we describe the distortion model we retained for high-distortion vision sensors. Although the following model has been thought for endoscopic devices and is not new, it may also be applied to any radially distorted fisheye optics mounted on more popular vision systems with a wide-angle lens. In this latter case, non-linear effects rather results of imperfections in the design or alignement of group of lenses due to low cost materials and quick implementation.

The camera model is based on the perspective projection affected by some distortion effects (radial pincushion, radial barrel, tangential or thin prism distortions). Brown [2] has presented fourty years ago a model for lens distortion. Although Brown's formulation can account for all the above-mentionned distortion effects, we shall focus on a radial distortion model. Let $\mathbf{m}_u = (x_u, y_u, 1)^\mathsf{T}$ and $\mathbf{m}_d = (x_d, y_d, 1)^\mathsf{T}$ be respectively undistorted and distorted image points. The distortion model we consider herein may be expressed as

$$\mathbf{m}_u \;=\; \mathbf{m}_{dc} \;+\; D(r_d)\;(\mathbf{m}_d - \mathbf{m}_{dc}) \qquad (5)$$

with $r_d^2 = (x_d - x_{dc})^2 + \frac{1}{a_r^2}\,(y_d - y_{dc})^2$ and $D(r_d) = 1 + \kappa_1\, r_d^2 + \kappa_2\; r_d^4 + \cdots + \kappa_n\; r_d^{2n}$. $\mathbf{m}_{dc} = (x_{dc}, y_{dc}, 1)^\mathsf{T}$ is the distortion centre and a_r is the aspect ratio. As a_r only depends on properties of the imaging array and digitizing electronics, we assume it is known as it is part of the set of instrinsic parameters which can be estimated with techniques based on the undistorted projection [15].

C. The robust distortion correction

In this section, we present a solution for the distortion's correction problem. The distortion parameters κ_i of the polynomial $D(r_d)$ (see section III-B) can be recovered linearly from the images of interest points that are collinear in space. Like in [13], let us assume that the distortion centre coincides with the image centre [1] and that the image coordinate system is centered in this point. Then consider the images of three collinear points $\mathbf{m}_u^{(i)} = (x_u^{(i)}, y_u^{(i)})^\mathsf{T}$. If these points are collinear once the distortion is corrected, we must have $\mathbf{m}_u^{(0)}\mathbf{m}_u^{(1)} \times \mathbf{m}_u^{(0)}\mathbf{m}_u^{(2)} = 0$. With the distortion model (5), this leads to

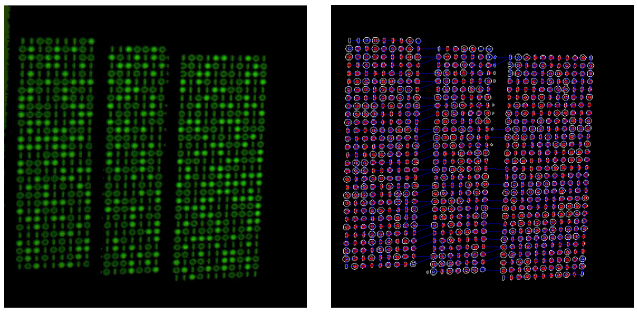
$$\begin{vmatrix} x_d^{(0)} & x_d^{(1)} & x_d^{(2)} \\ y_d^{(0)} & y_d^{(1)} & y_d^{(2)} \\ f(r_d^{(0)}) & f(r_d^{(1)}) & f(r_d^{(2)}) \end{vmatrix} \;=\; 0\;, \qquad (6)$$

with $\lambda_1 = -\kappa_1$, $\lambda_2 = -(\kappa_2 - \kappa_1^2)$, $\lambda_3 = -(\kappa_3 - 2\kappa_1\kappa_2)$, ... and $f(r_d) = 1/D(r_d) = 1 \;+\; \sum_{l=1}^{l=n} \lambda_l\; r_d^{2l}$. Contraints can be accumulated from all possible triplets of points that are projections of collinear points in the calibration pattern. We thus obtain a linear equation of the form $\mathsf{A}\;\Lambda = \mathbf{b}$ where $\Lambda = (\lambda_1, \lambda_2, \lambda_3 \cdots)^\mathsf{T}$ is the vector of unknowns. Minimizing the algebraic distance by the accumulation of the residuals $\sum_i \xi_i^2 = \|\mathsf{A}\;\Lambda - \mathbf{b}\|^2$ with the standard least squares through the SVD of matrix A yields the parameters that maximize the collinearity of the rectified points. However, to handle properly the estimation in presence of noise and corrupted data, a robust estimation has to be performed. The Iteratively Reweighted Least Squares algorithm (IRLS) aims at solving the following system $\mathsf{WA}\;\Lambda = \mathsf{W}\;\mathbf{b}$ where $\mathsf{W} = \mathrm{diag}(w_1, w_2, \cdots, w_n)$ is a diagonal matrix where w_i reflects the confidence of each feature. To do so, a sigmoid function is useful for favouring the confidence of farthest interest points from those close to the image center. Considering that the noise acts statistically with the same strengthness all over the image, it makes sense to pick up the distortion information rather close to the image borders. Then, the weight can be expressed as $w_i(r_d) = \frac{1}{1+e^{-a|r_d^{(i)} - c|}}$ and one has to follow the iterative minimization (at step k)

$$\min \sum_i w(\xi_i^{(k-1)})\;\xi_i^2\;, \qquad (7)$$

and the weight $w(\xi_i^{(k-1)})$ must be recomputed at every step.

[1] It seems advisable not to recover the distortion center if the lens has a very small distortion value.



(a) (b)

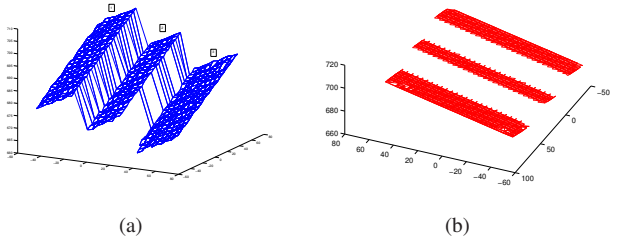
Fig. 3. (a) Projection on the three parallel planes. (b) Detection and decoding: the red color is corresponding to well-decoded primitives.

IV. Experimental results

A. Validation of the calibration

The experimental setup consists of a digital camera AVT Pike F100C and a video-projector Sony VPL-CS6 (Figure 2-b). The image size is (512×512) pixels. The sizes of the printed pattern are about (15×18) cm. 14 different positions of the pattern plane have been considered for the experiments. For each attitude, two images have been captured, the first one when the projector is switch off, the second one when it is turn on. With the set of captured images of the *printed* pattern, both the camera intrinsic parameters and the 14 euclidean transformations from the camera-to-plane euclidean transformations are estimated. With the images of the *projected* pattern, each pattern line (the laser beam) is evaluated as described in section III-A.2. In order to assess the accuracy of the calibration of the overall vision system (results of the camera parameters are not reported here), we have projected the pattern onto three parallel planes (Figure 3-a). The distance between each pair of planes has been measured previously at several positions with a mechanical measurement device (values are reported in the first column of TABLE I). The three planes have been placed inside a workspace corresponding to a depth range of 600-800 mm. Then, each image has been segmented so as to locate primitive centres. Then, each primitive has been classified according to its shape and the neighborhood search provides the associated codeword (see [1], for further details). In Figure 3-b, the primitives which are correctly decoded are drawn with a red color, the others are drawn in blue.

The set of calibration data corresponding to all pattern elements are used to reconstruct the three planes, independently, (Figure 4-a) that is the computation of the triple $\{\pi_1, \pi_2, \pi_3\}$, expressed in the camera frame (see section III-A.2). To do so, a robust estimation has been carried out. Several $k = 14$ viewpoints have been acquired so as to collect a set of vectors $\{\pi_{1k}, \pi_{2k}, \pi_{3k}\}$, allowing to compute the relative distance and orientation between planes and the average of interdistance differences. These data are gathered in the TABLE I. One can see that the distance error is always less than 0.9 mm and the angle error is always less than 1° (0.5° in average) in the considered workspace.



(a) (b)

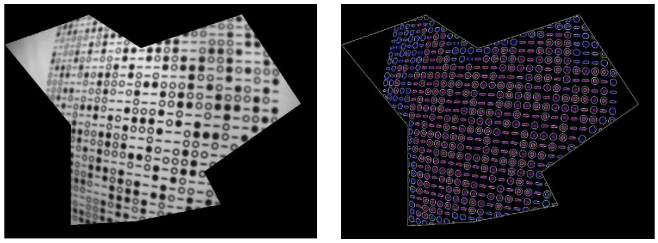
Fig. 4. (a) The planes reconstruction by means of the 3D coordinates estimates of the pattern projections. (b) A new view of the three planes computed from the 3D reconstruction and after the robust estimation of vectors $\pi_{1,2,3}$.

Surfaces	True distance (mm)	Mean distance (mm)	Distance errors (mm)	Angle (°)
1-2	16.50	16.27	0.48	0.15
1-2		16.00	0.84	0.53
1-2		16.75	0.75	0.87
1-2		15.98	0.88	0.77
2-3	16.70	16.71	0.21	0.11
2-3		16.58	0.52	0.09
2-3		16.98	0.75	0.74
2-3		16.78	0.45	0.91

TABLE I

Reconstruction errors for the three parallel planes.

B. Evaluation of the radial distortion correction



(a) (b)

Fig. 5. (a) A part of the imaged pattern. (b) Successfully decoded primitives are drawn in red color. Blue ones correspond to the only segmented (but not decoded) primitives.

Many experiments have been conducted with an endoscopic Karl Storz© vision system with a straight laparoscope of 35 mm of length and 10 mm of diameter. The segmentation approach use the contours for detecting the symbols. Then the contours are classified in order to label each contour by one of the three pattern primitives. The classification of the primitives (circle, disc or stripe) is based on first-order statistics for the circle and on second-order for the two other primitives (see [1] for more details). Once the neighbors of each pattern features are detected, its codeword is determined as illustrated on Fig. 1-a. Therefore, an exhaustive search is carried out to find the primitive location in the original matrix pattern taking advantage of the pattern robustness. On the Fig. 5-b, the segmented primitives (in blue) and the decoded ones (in red) have been drawn. One may observe that not all the pattern is visible (Fig. 5-a). Except close to

the boundaries of the selected polygonal image region, all the primitives have been successfully decoded.

The computation of the pattern primitives displacement has been performed. On Fig. 6, the red square are corresponding to corrected positions whereas green triangles are the initial primitives' positions. The distortion centre has been evaluated and its position is drawn on this figure (pink star). All the endoscopic images has been corrected, and a set of three sample images and corrected ones have been reported on Fig. 7, thanks to a backward warping and gray-level interpolation.

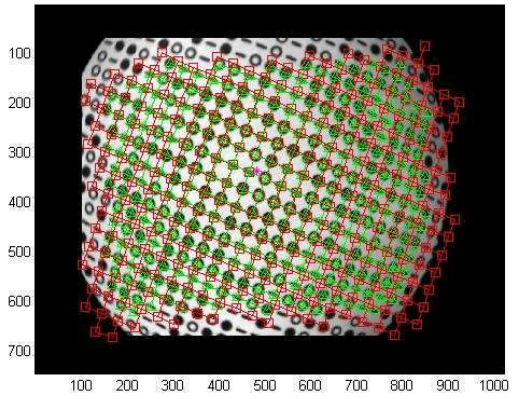
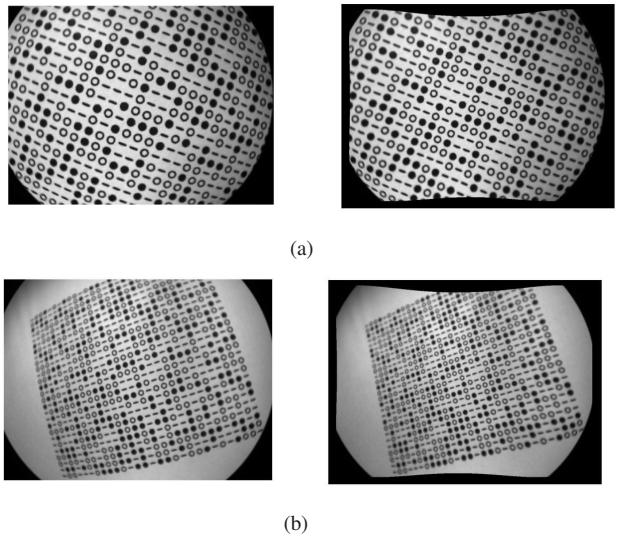


Fig. 6. Correction of the centres' coordinates of each decoded pattern primitives (red square). The estimated distortion centre (drawn with the pink star) is computed thanks to a sampling grid of (0.5 pixels step) dimensions (10×10) around the image centre.



(a)

(b)

Fig. 7. Two images acquired with a straight laparoscope at distinct distances and orientations from the calibration pattern. Left: distorted images. Right: corrected images.

V. Conclusion

We have presented a new calibration method for structured light vision system and a distortion compensation technique both using a robust pattern based on perfect maps. We have followed the strategy of spatial neighborhood for coding the structured light with only three symbols to design the overall pattern. One of them carries the local orientation so as to reduce the search of the neighbors during the decoding stage. The distortion compensation as well as the calibration of each pattern element assess the practicability of the proposed pattern for calibration purposes since the pattern (printed or projected) and its associated decoding algorithm allow to automatically and quickly solve the correspondences. Since the Hamming distance of any pair of codewords is greater or equal to one, the pattern exhibits a high degree of robustness. This is needed even for the distortion compensation as the pattern element are captured in disorded configurations and some of them may be out-of-field of view or blurred. Considering the accuracy, the proposed calibrations are not superior to early techniques, excepted in presence of a projector with strong distorsions. They are rather two convenient methods for in situ calibration with automatic matching. The experimental results we obtained with a high distorted straight laparoscope show very good performances in the distortion compensation and 3-D reconstruction accuracy.

References

[1] C. Albitar, P. Graebling, and C. Doignon. Robust structured light coding for 3d reconstruction. In *IEEE Int'l Conf. on Computer Vision*, Rio de Janeiro, Brazil, October, 14-20 2007.

[2] D.C. Brown. Decentering distortion of lenses. *Photogrammetry Engineering*, 32:444–462, 1971.

[3] C. Chen, Y. Hung, C. Chiang, and J. Wu. Range data acquisition using color structured lighting and stereo vision. *Image and Vision Computing*, 15:445–456, 1997.

[4] C. Doignon and D. Knittel. A structured light vision system for out-of-plane vibration frequencies location of a moving web. *IAPR Int'l Journal of Machine Vision and Applications*, 16(5):289–297, 2005.

[5] M. Hayashibe and Y. Nakamura. Laser-pointing endoscope system for intra-operative 3d geometric registration. In *IEEE Int'l Conf. on Robotics and Automation*, Seoul, Korea, May 2001.

[6] A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, L. Soler, J. Leroy, and J. Marescaux. Autonomous retrieval and 3d positioning of surgical instruments in robotized laparoscopic surgery. *IEEE Trans. on Robotics and Automation*, 19(5):842–853, October 2003.

[7] A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissanov. Structured light using pseudo-random codes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(3):322–327, 1998.

[8] J. Pagès, Ch. Collewet, F. Chaumette, and J. Salvi. An approach to visual servoing based on coded light. In *IEEE Int'l Conf. on Robotics and Automation*, pages 4118–4123, 2006.

[9] E.M. Petriu, T. Bieseman, N. Trif, W. McMath, and S. Yeung. Visual object recognition using pseudo-random grid encoding. In *IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 617–624, 1992.

[10] M. Ribo and M. Brandner. State of the art on vision-based structured light systems for 3d measurements. *International Workshop on Robotic and Sensor Environments*, pages 2–6, 2005.

[11] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

[12] J. Salvi, J. Pagès, and J. Batlle. Pattern codifications strategies in structured light systems. *Pattern Recognition*, 37:827–849, 2004.

[13] J.P. Tardif, P. Sturm, and S. Roy. Self-calibration of a general radially symmetric distortion model. In *European Conf. on Computer Vision*, pages 186–199 – LNCS 3954 Part IV, Graz, Austria, May 2006.

[14] S. Yee and P. Griffin. Three-dimensional imaging system. *Optical Engineering*, 33(6):2070–2075, 1994.

[15] Z. Zhang. A flexible new technique for camera calibration. In *ICCV*, pages 666–673, Kerkyra, Greece, September 1999.

[16] F. Zhou and G. Zhang. Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations. *Journal of Image and Vision Computing*, 23:59–67, 2004.