34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A structured light-based laparoscope with real-time organs' surface reconstruction for minimally invasive surgery

Xavier Maurice[1], Chadi Albitar[2], Christophe Doignon[1] and Michel de Mathelin[1]

Abstract— **In this paper we present a new 3-D laparoscopic device based on structured light for minimally invasive surgery. Real-time reconstruction of internal organs' surfaces is very challenging as the numerous geometric and photometric variabilities and disturbances (bloody parts, specularities, smokes,...) often occur during the surgical operation, sometimes with manipulations by several assistants. We then conceived a structured light vision system to illuminate a coded pattern by means of an external video projector device or miniaturized diffractive optical elements and a laser source.**

Among the structured light techniques, the spatial neighbourhood scheme is the most relevant class of approaches to deal with moving and deformable surfaces, then to capture the depth map with only one shot. Each neighbourhood (a (3×3) window) is representing a codeword of length 9, and is unique in the whole pattern, even if there is a lack of information. To do so, a monochromatic subperfect map-based pattern is computed, driven by a desired minimal Hamming distance, H_{min}, between any couple of codewords. This provides patterns with high correction capabilities ($H_{min} > 1$). For practical considerations, each numerical codeword symbol is associated to a unique visual feature embedding the local orientation of the pattern, which is helpful for the neighbourhood retrieval during the decoding process. Together with the endoscopic device, in vivo real-time reconstructions (in mini-invasive surgical conditions) are presented to assess both the efficiency of the proposed pattern design, the decoding process and the 3-D laparoscope setup realized in the lab.

I. INTRODUCTION

The three-dimensional reconstruction of surfaces is a major problem for many applications of computer vision. It is particularly the case in medical applications for which the developments of robotic assisted Minimally Invasive Surgery (MIS) are remarkable. The great benefits of MIS are to reduce the incisions and trauma and the hospitalization time. The use of imagery in this field remains limited because of the nature, generally two-dimensional, of the information provided by most acquisition devices. Although stereo laparoscopes capture two two-dimensional video information, they involve human attention and brain activity to retrieve the depth mentally. This task is tiring, especially when the surgeon is focused on precise gesture. Moreover, stereovision principle fails with poor textured scene. In fact, a device providing the 3-D information of the close environment, automatically and in real time, would open large prospects. For example, within the framework of the MIS, this information could be registered with preoperative images in order to visually restitute the depth feeling to the surgeon or to achieve robotized surgical acts. Therefore, the aim of our work is to develop a device in order to estimate the 3-D moving surfaces of internal structures of the abdomen.

Many works ([5], [7], [17]) proposed to reconstruct the soft tissue structures of the abdomen using stereovision. Thus, they had to face to the well-known correspondences problem. To tackle this, structured light based methods were presented in [11]. The structured light technique replaces one of the cameras of the classical stereovision system by an active device which illuminates the scene with a known coded pattern [15]. Consequently, the correspondence problem turns into decoding the features in the captured image to determine their correspondences in the original known pattern. However, the long scanning time does not yield real-time 3-D reconstruction in [11].

Methods using the spatial neighbourhood strategy for coding the pattern elements adopt the projection of a unique pattern including all the code. Generally, this enables to deal with dynamic scenes but at lower resolution than other strategies and several errors may arise during the decoding stage. Then, the coding scheme must alleviate any ambiguity (uniqueness of codewords) and should also increase the robustness of the decoding algorithm, even in the presence of occlusions and shadows. To solve these problems, some authors used the theory of Perfect Maps (PM) to encode a unique pattern ([18], [14], [12], [1]). Briefly, if `M` is a matrix of dimensions $m \times n$ in which each element is taken from an alphabet of k symbols, and if each submatrix $u \times v$ appears exactly once, then `M` is a Perfect Map. To this end, the Hamming distance is used to compare two codewords by counting the number of indices with different symbols. Moreover, if `M` does not contain all the possible submatrix then `M` is an M-array.

The robustness of these patterns relies on their capability to decode visible parts of the observed pattern thanks to the properties of the M-arrays. The proposed patterns can be classified according to the size of the considered matrix, to the use of color and also according to the length of the considered alphabet. For the medical application of concern, a structured light projection based on the spatial neighborhood scheme using the M-array approach seems to be the most suited. Then the main difficulties lie on the pattern design and in the associated fast decoding algorithm.

PM are the 2-D extensions of De-Bruijn sequences which have been widely studied by many authors such as ([10], [13]). As it uses a local information, such approach reveals interesting capabilities for matching and correcting mislabeled features. If direct pattern constructions have been

---

[1] X. Maurice, C. Doignon and M. de Mathelin are with Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (UMR CNRS), Equipe Automatique Vision et Robotique, Université de Strasbourg, France.

[2] C. Albitar is with the Department of Electrical Engineering, Institut of Applied Sciences and Technology, Damas, Syria.

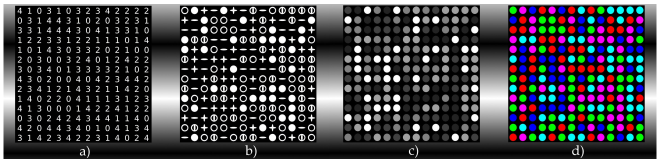


Fig. 1: A 5-ary, (14×14) *PSM* with a min Hamming distance of 5 between all the (3×3) neighbourhoods ($\equiv$ codewords), represented with a) digits, b) shapes, c) grey levels, d) colors.

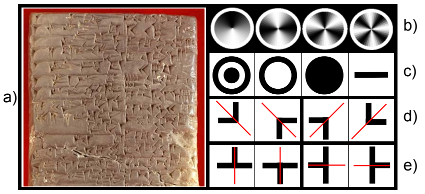


Fig. 2: a) The cuneiform writing system, created 5 000 years ago. b) Proposed shapes in [3]. c) Proposed shapes in [1]. d-e) Cuneiform shapes.

proposed by [4] and [6] two decades ago, they were often restricted to the binary case ($k = 2$). Furthermore they did not allow to choose an arbitrary pattern size, nor to impose Hamming distances greater than 1. Morano [12] and more recently Albitar in [1] and Claes in [3] proposed pattern generation algorithms, that we have improved in [9] with the knowledge of epipolar geometry (of the camera-projector images) and with fast decoding process [8]. These iterative algorithms build the k-ary pattern of a desired size by trial and error, checking the desired Hamming distance constraint at each step. Contrary to PM, not all the possible submatrices are used, therefore they were named *perfect submaps* (PSM, see Fig. 1 for coding examples). Finally, when the uniqueness is ensured only locally or in a particular direction, this kind of patterns is named *subperfect submap* (SPSM).

In comparison with previous articles about the coding ([9], [8]), in this article, we rather present the validation of a new monochromatic pattern based on the PSM and SPSM approaches for medical endoscopy with *in vivo* experiments and reconstructions. Grid-based patterns and epipolar-based patterns have been designed in such a way. The realizations have been made possible thanks to the technology of diffractive optics coupled with a laser (for grid-based pattern) or, alternatively, thanks to an external video-projector (for epipolar-based pattern).

II. PATTERN DESIGN

A. The visual features

As the foreseen application consists in acquiring images of internal structures of the abdomen, therefore, the choice of the light color is crucial because of the nature of the considered environment. In [16], the influence of the presence of tumors on the reflectance of body organs was studied at different wavelengths. Then, a monochromatic pattern seems suitable to avoid a color coding sensitive to interactions with injured or healthy tissues.

To associate a visual feature to a numerical symbol, many shapes have also been studied. Recent works focus on oriented features; In [3], the authors deal with a radial analysis of shapes in the frequency domain like those depicted in Fig. 2-b, but this approach, which seems very sensitive to spectral recovery, has not been further exploited. In [1] we proposed to use disk, ring and dash shapes. Such a choice was justified by the simplicity of shapes, well adapted to existing segmentation algorithms (see Fig. 2-c). Moreover, the dash direction was used to quickly find the neighbours. In both studies, the local orientation is brought by one or several symbols (see Fig. 3). Corners are also very relevant features as they are intersections of two edges and hence bring 2-D directional information. Since very efficient detectors are available, we recently seek for features based on corner composition. We have selected 8 cuneiform features, 4 with 2 arms (or segments) (Fig. 2-d) and the other 4 with 3 arms (Fig. 2-e). This way, the two dimensions may be fully extracted. The cuneiform shape has some advantages. Among them, we note:

1) the center is locally well-defined with the intersection of two orthogonal segments, which is more accurate than with circular-shaped features (see Fig. 4),
2) the directional information can be related to the epipolar geometry,

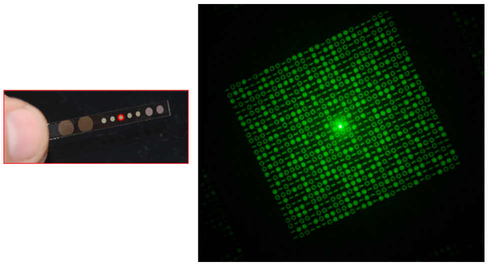


Fig. 3: Pattern realization and miniaturization by means of diffractive optical elements for grid-based pattern. Only three symbols ($k = 3$) are used and associated visual features are disks, ring and dash shapes [2] (drawn in fig. 2-c).

B. Epipolar-based pattern design

Prior to achieve the epipolar-based pattern, one may build a PSM-based pattern driven by the desired minimum Hamming distance. This pattern (Fig. 4) serves to estimate the Fundamental matrix with at least two viewpoints.

The search for corresponding features is the main time consuming step of the processing pipeline. Although this problem is easy to solve with planar surfaces, in the general case many difficulties may occur (occlusions, high curvatures, etc..), which makes the problem tricky. Therefore, if one can predict the features' position and orientation in the image of the camera whatever the scene is (as it is a projective invariant), it significantly improves the reliability and speeds up the features' matching/decoding process (see

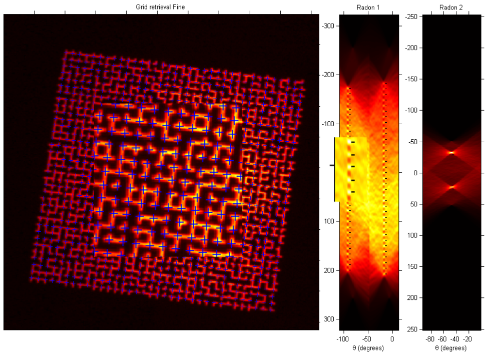
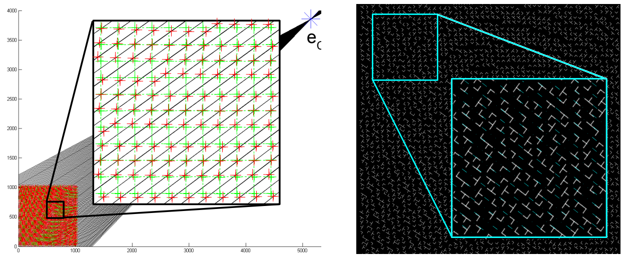


Fig. 4: A grid-based pattern (with a zoom in of central elements) and two successive Radon Transforms (the centers of cuneiform features are drawn with blue crosses).



(a) Relocalisation of the pattern elements (red) from a grid (green) onto the closest epipolar lines (black) in the image of the projector. $\mathbf{e}_c$: the epipole.

(b) A final pattern in the image of the projector after features alignment (translation, rotation) along chosen epipolar lines (Resolution : 50^2, $k = 8$, $H_{Min} = 6$).

Fig. 5: Design of a 50×50 pattern built upon the epipolar geometry (k is the number of symbols and H_{Min} is the minimum Hamming distance).

details in [9] and in [8]). To do so, we propose to embed the epipolar geometry in the pattern itself. Finally, once the coding is carried out and centers of cuneiform features are placed on the image of the projector, each cuneiform feature is rotated around its center in such a way that its horizontal arm(s) is superimposed with the associated epipolar line (see example in Fig. 5-b). This way, we enforce this(ese) arm(s) to project along the corresponding epipolar line in the camera image, whatever the scene is. This is very useful as it significantly simplifies the image segmentation as the features orientation is known *a priori* and hence it allows to discard many classification ambiguities.

III. SYSTEM SETUP: THE STRUCTURED LIGHT-BASED LAPAROSCOPES

A. *A 3-D endoscopic device for open surgery*

We have developed a miniaturized pattern illumination system. This version of the active imaging device requires at least two trocarts when applied to minimally invasive surgery (see Fig. 6). As diffractive optics are used to miniaturize the pattern (of size (29×27) in 1.3 mm of diameter), it is intended for monochromatic grid-based patterns (with green laser source herein - 532 nm, 200 mW).

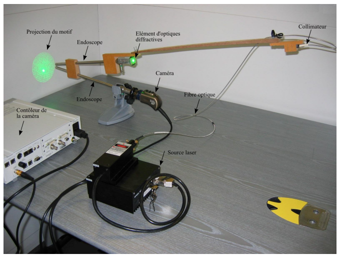


Fig. 6: Endoscopic device with coded structured light (Perfect Map) and laser illumination - 532 nm, 200 mW - for grid-based patterns (depicted in fig. 3).

B. *The 3-D laparoscope (mono-trocart)*

Like with stereo systems, the 3-D laparoscope is a two-channels Karl Storz 10 mm straight optics with right angle. A channel is used to connect the endoscopic full HD camera. In the second channel, a 5 mm straight laparoscope is connected through interface optics to an external (mini) video-projector for which the light source has been replaced by the powerful Karl Storz Xenon 300 white light (Fig. 7).

The 3-D laparoscope has been used in the IRCAD operating

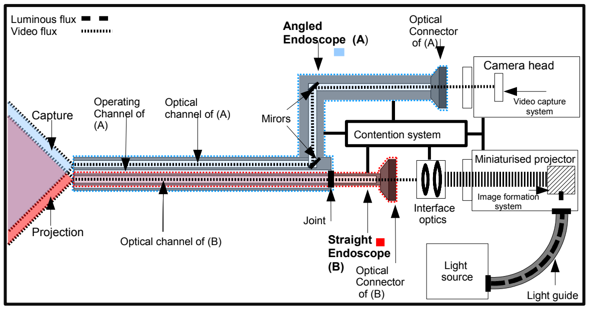


Fig. 7: Synopsis of the 3-D laparoscope.

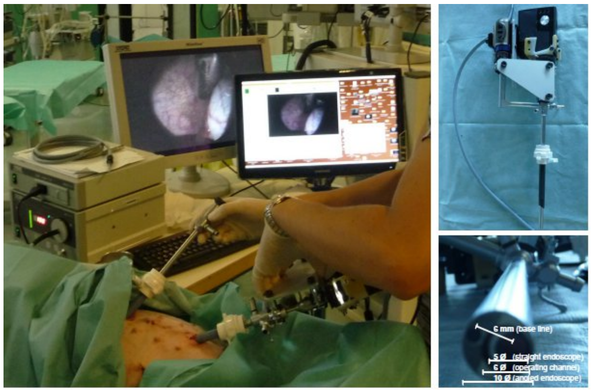


Fig. 8: (left) The 3-D laparoscope in the operating room at IRCAD France. (top right) The laparoscope with coded structured light (SubPerfect map) and a video projector (in black) for epipolar-based patterns. (bottom right) Details of the operating channels at the tip of the instrument: the base line is 6 mm for this 10 mm laparoscope.

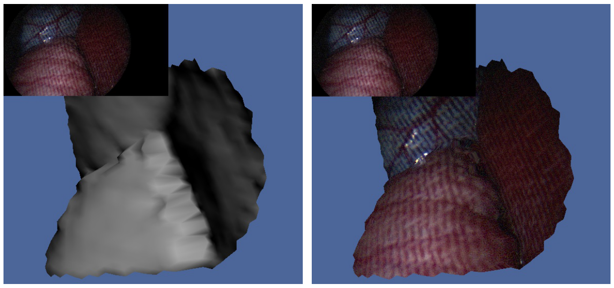


Fig. 9: 3-D reconstructions (original image in upper left). (left) : without texture mapping. (right) : with texture mapping. Other viewpoint can be selected thanks to the developed software with IHM. Many processings (image filtering and segmentation, distortions correction, rectification, features classifications, features epipolar alignments and decoding stages) are executed on GPUs with Cuda API (Nvidia GTX 470 graphic card).

room in Strasbourg hospital to perform experimental validations in minimally invasive situation (in situ calibration, insertion, dexterity) (see Fig. 8) and to get data set for invivo 3-D reconstructions (see Fig. 9).

IV. CONCLUSION

We present in this article the design and realization of a new endoscopic device by means of a robust structured light coding. This prototype has been entirely conceived and realized in the laboratory. Since the main objective of this work was to reconstruct in real time moving surfaces of internal organs for minimally invasive surgery, the pattern illumination has been carried out thanks to the diffractive optics technology for grid-based patterns and external mini video-projectors for epipolar-based patterns. Up to now, only the latter has been fully miniaturized, and this is the most promising technique to get dynamical patterns at low cost.

According to the theory of Perfect Map, monochromatic patterns have been computed to get the numerical symbols arrangement. Taking advantage of the known epipolar geometry, we have presented non grid-based patterns to speed up the decoding process and to simplify the feature segmentation and classification (see [9] and [8] for further details about the coding/decooding algorithms). Second, we model the pattern components with simple geometrical features instead of color coding. Each pattern component carries the directional information and this improves the detection of neighbours and speeds up the decoding process.

A structured light system was developed to validate the choice of the pattern with the associated algorithms. Finally, we integrated the structured light vision system in a single two-channels Karl Storz laparoscope of 10 mm of diameter. We have led a set of experiments in the operating room with this new 3-D laparoscope in order to capture internal structures of a pig abdomen. We found that our technique gave satisfactory results assessed by the 3-D reconstructions, with 50×50 patterns at a frame rate of 25 images/s with GPU processing of a full HD video stream.

As the structured lighting provides a relative depth map and 3-D structure of a moving object of interest, we plan to used it for improving visual 3-D tracking for medical diagnosis and visual servoing of robotized surgical instruments in minimally invasive and transluminal surgeries with endoscopic apparatus.

ACKNOWLEDGMENT

The authors would like to thank the Karl Storz company for their advices and material, the Institut de Recherche Contre les Cancers de l'Appareil Digestif (IRCAD) France, in Strasbourg Hospital, for surgical advices, operating room access and surgical instruments facilities, and the Alsace Region Council for its financial support.

REFERENCES

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

[2] C. Albitar, P. Graebling, and C. Doignon. Fast 3d vision with robust structured light coding. In *SPIE Medical Imaging 2009, Visualization and Image-Guided Procedures*, Orlando, USA, 2009.

[3] K. Claes. Structured light adapted to control a robot arm. *PhD Thesis*, pages 33–43, May 2008.

[4] T. Etzion. Constructions for perfect maps and pseudorandom arrays. *IEEE Int'l Trans. on Information Theory*, 34(5/1):1308–1316, 1988.

[5] M. Hu, G. Penney, P. Edwards, and D. Hawkes. 3d reconstruction of internal organ surfaces for minimal invasive surgery. In *Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 68–77, Brisbane, Australia, 2007.

[6] Glenn Hurlbert and Garth Isaak. On the de bruijn torus problem. *J. Combin. Theory Ser. A*, 64:50–62, 1993.

[7] W. Lau, N. Ramey, J. Corso, N. Thakor, and G. Hager. Stereo-based endoscopic tracking of cardiac surface deformation. In *Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 494–501, St Malo, France.

[8] X. Maurice, C. Doignon, and P. Graebling. A pattern framework driven by the hamming distance for structured light-based reconstruction with a single image. In *IEEE Intl conference on Computer Vision and Pattern Recognition*, Colorado Springs, USA, 2011.

[9] X. Maurice, P. Graebling, and C. Doignon. Real-time structured light coding with subperfect submaps. In *SPIE Photonics Europe*, Brussels, Belgium, 2010.

[10] Chris J. Mitchell. Aperiodic and semi-periodic perfect maps. *IEEE Trans. on Information Theory*, 41(1):88–95, 1995.

[11] H. Mitsuhiro, S. Naoki, and N. Yoshihiko. Laser-scan endoscope system for intraoperative geometry acquisition and surgical robot safety management. *Medical Image Analysis*, 10:509–519, 2006.

[12] R. 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.

[13] K. G. Patterson. Perfect maps. *IEEE Int'l Trans. on Information Theory*, 40(3):743–753, 1994.

[14] E. M. Petriu, Z. Sakr, H. Spoelder, and A. Moica. Object recognition using pseudo-random color encoded structured light. *In Proceedings of the 17th IEEE Instrumentation and Measurement technology Conference*, 3:1237–1241, 2000.

[15] J. Salvi, J. Pages, and J.Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, 2004.

[16] M. Sambongi, M. Igarashi, T. Obi, M. Yamaguchi, N. Oyama, M. Kobayashi, Y. Sano, S. Yoshida, and K. Gono. Analysis of spectral reflectance of mucous membrane for endoscopic diagnosis. In *IEEE Int'l Conf. on Engineering in Medecine and Biology*, pages 1026–1029, 2000.

[17] D. Stoyanov, A. Darzi, and G. Yang. Dense 3d depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery. In *Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*, pages 41–48, St Malo, France.

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