Robust Structured Light Coding for 3D Reconstruction

Chadi ALBITAR Pierre GRAEBLING Christophe DOIGNON
Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection
LSIIT, UMR ULP-CNRS 7005, Equipe de Recherche en Automatique, Vision et Robotique
Pôle API, Bd. Sébastien Brant, 67412 Illkirch
{albitar,graebling, doignon}@lsiit.u-strasbg.fr

Abstract

In this paper we present a new monochromatic pattern for a robust structured light coding based on the spatial neighborhood scheme using the M-array approach. The proposed pattern is robust as it allows a high error rate characterized by an average Hamming distance higher than 6. We tackle the design problem with the definition of a small set of symbols associated to simple geometrical primitives. One of these primitives embeds the local orientation of the pattern. This is helpful while performing the search for the relevant neighborhood during the decoding process. The aim of this work is to use this pattern for the real-time 3-D reconstruction of dynamic scenes, particularly in endoscopic surgery, with fast and reliable detection and decoding stages. Ongoing results are presented to assess both the capabilities of the proposed pattern and the reliable decoding algorithm with projections onto simple 3-D scenes and onto internal structures of a pig abdomen.

1. Introduction

3-D reconstruction of the environment surfaces is a major problem for many applications using computer vision. It is particularly the case of medical applications in which the use of imaging devices is increasing. Actually, the use of imagery in this field remains limited because of the nature, generally two-dimensional, of the information provided by the most of acquisition devices. Thus, a device providing, possibly in real time, the 3-D information of the environment would open new prospects. For example, within the framework of mini-invasive operations, this information could be registered with preoperative images in order to visually restitute the depth feeling to the surgeon. Also, it would be helpful to foresee and to achieve robotized surgical acts.

The objective of our work is to estimate the 3-D surfaces of internal structures of the abdomen in order to integrate this information in robotic medical systems. Among all the techniques usually used for surface reconstruction, we propose a solution based on a structured light projection. Here, we replace one of the cameras of the classical stereovision system by a device which illuminates the scene with a known pattern. Consequently, the well-known stereo correspondences problem turns into finding the correspondences between the image pixels and the projected pattern points. Therefore, the main problems of such systems are 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.

In this article, we focus on the design and the validation of a new robust pattern based on the "M-array" approach. It will be integrated in an endoscopic tool to reconstruct the 3D surfaces of dynamic scenes such as the abdomen structures. Also it can be used for robots vision-based control in surgery or in other applications.

2. Structured light codification

Structured light techniques can be classified, according to the strategy of encoding the projected pattern, into three categories [1]: time multiplexing, direct coding and spatial neighborhood. The methods based on the time multiplexing strategy are easy to be implemented and can achieve a good accuracy and a high resolution. The main drawback of these methods is their weakness to deal with dynamic scenes since multiple patterns must be projected successively. The methods adopting direct coding provide a good spatial resolution but their applicability is limited due to their sensitivity to noise and to light variations. Methods of the last category adopt the projection of a unique pattern including all the code. The code of each component of the pattern depends on its value and on those of its neighbors. Generally, this enables to deal with dynamic scenes but with lower resolution than the other strategies and some errors that may arise during the decoding stage. In order to minimize these errors, the coding and decoding algorithms must be more robust.

For our application, a technique related to the last category seems to be the most appropriate. The difficulties then lie

978-1-4244-1631-8/07/$25.00 ©2007 IEEE

in the choice of the pattern and in the associated decoding algorithm.

2.1. Spatial neighborhood strategy

The methods based on the spatial neighborhood technique tend to integrate the entire coding scheme in a unique pattern. The coding must alleviate any ambiguity while identifying the components (uniqueness of codewords) and must also increase the robustness of the decoding algorithm, even in the presence of occlusions. Some authors ([2]-[4]) proposed simple patterns with non-formal coding and without any mathematical background, which are consequently neither optimal nor robust. This can cause ambiguity because we may encounter identical regions within the pattern. Other strategies use a well-defined coding scheme based on a De Bruijn sequence ([5]-[7]). The drawback in this case is that occlusions or shadows may cause deletions and disorders among the pattern primitives so that it will be difficult to decode the observed pattern. To solve these problems, some authors adopted the theory of Perfect Maps (PM) to encode a unique pattern ([8]-[12]) taking advantage of its interesting mathematical properties. If M is a matrix of dimensions $n \times m$ in which each element is taken from an alphabet of k symbols, and if M has a window property $r \times v$ so that each different submatrix $r \times v$ appears exactly once, then M is called a Perfect Map. If M does not contain the submatrix filled with 0's, M 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-arrays. The proposed patterns in this category can be classified according to the size of the considered matrix, to the use of color or not and also according to the length of the considered alphabet.

2.2. Constraints of M-array based pattern

Since we are interested in non-controlled scenes as it is the case of the internal structures of the abdomen; we have to face the classical problems of shadows and occlusions. In such situations, a key factor is to elaborate a pattern with a significant Hamming distance between the codewords. Actually, the Hamming distance quantifies the differences between the codewords. If this distance is higher than 1, it quantifies the capability to correct the errors that can occur in the segmentation and decoding stages. It was shown in [9] 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 a significant number of symbols as it is the case in [9] and in [12] where, respectively, 7 and 8 symbols were used. Indeed, in order to simplify the decoding stage, it would be more efficient to handle fewer symbols. To represent the symbols, we can use color dots where each color is associated to a symbol. But we can also consider geometrical primitives and generate a monochromatic pattern. In practice, monochromatic patterns are more robust than colored ones. Firstly, it is much easier to deal with an image having only two intensity levels. Secondly, if the scene is very colored, the choice of colors to be used in the pattern can be very tricky. Moreover, the realization of compact projection devices with a monochromatic light is much simpler thanks to the technology innovations.
Since the code of a primitive depends on its neighborhood, it is necessary to detect this neighborhood easily and without imposing any constraints on the acquisition device. To solve this problem, the primitives can be linked as it was proposed in [11]. Unfortunately, this can cause new difficulties in the detection and segmentation stages.
So, it would be more convenient to consider an alphabet with minimum number of symbols in building an M-array while being able to detect the neighborhood correctly.

3. The design of the proposed pattern

The pattern we present is designed to solve the above mentioned problems taking into account the constraints imposed by the desired application. As we aim to integrate the pattern in an endoscopic tool, we searched for a solution that makes the design as compact as possible. That is why a solution based on the technology of diffractive optics with a monochromatic light seemed to be the most suitable. Diffractive Optical Elements (DOE) are used to shape an incident laser beam and implement a desired pattern. Generally, the pattern is generated using 40% of the energy of the laser beam. Moreover, if the pattern is central symmetric, the design of the DOE can be configured to use up to 80% of the energy of laser beam. So, we searched for a pattern based on the M-array theory and which respects the following properties:

1- Monochromatic light,
2- Central symmetry,
3- Uniqueness of the code of each window (3×3),
4- Hamming distance > 3.

Considering the choice of using a monochromatic light, the symbols cannot be coded with different colors, so we proposed an alphabet of three symbols associated to geometrical shapes as shown in Figure 1. We chose a circle and a disc to simplify the image processing during the segmentation stage. The third symbol is the stripe which will allow; thanks to the directional information it carries, to rotate the observation window correctly during the stage of neighborhood detection and without enforcing any positioning constraint on the acquisition device.

Figure 1: The considered primitives: disc, circle and stripe.

In our case, the length of the codeword associated to each primitive is 9 since we consider a 3×3 window centered on this primitive. The uniqueness constraint of each window corresponds to a Hamming distance 1. To increase this value, we imposed an additional constraint so that each window must remain unique even if the upper corners elements are missing. Since the matrix which we were looking for is central symmetric, each window will also remain unique even if the lower corners elements are missing. We implemented an algorithm to generate the matrix respecting the above mentioned constraints and we focused on maximizing the matrix dimensions in order to project the pattern on a surface of significant size. We obtained a matrix of dimensions 27×29, which means 783 primitives, verifying the first three desired criteria and using three symbols (see Figure 2).

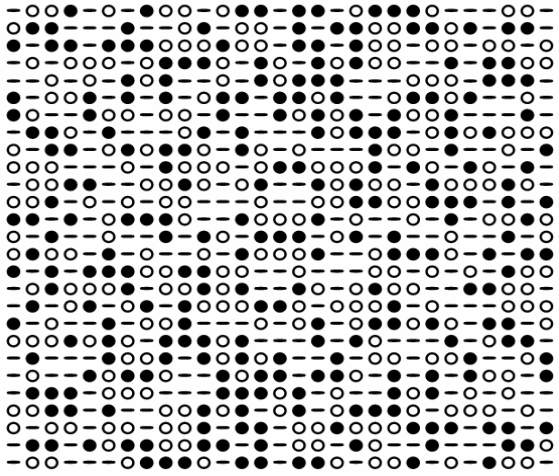


Figure 2: The proposed pattern (783 elements).

To verify the last constraint, we computed the Hamming distance between the codewords (Figure 3). Considering the distances between each codeword and all the other codewords, we obtained $H(x)$ which represent the percentage of the codewords that have a Hamming distance x between them : $H(1) = 0.05$, $H(2) = 0.71$, $H(3) = 3.45$, $H(4) = 10.40$, $H(5) = 20.24$, $H(6) = 27.07$, $H(7) = 23.55$, $H(8) = 11.83$ and $H(9) = 2.70$. Consequently, the average Hamming distance is $\overline{H} = 6.0084$. Note that 99,95% of our codewords have a Hamming distance greater than 1, while 95,79% of the codewords have a Hamming distance higher than 3. When we considered the distances between each primitive codeword and those of its neighbors only, the average Hamming distance that we obtain is $\overline{H} = 6.173$ while the best result obtained using other patterns is $\overline{H} = 3.15$.

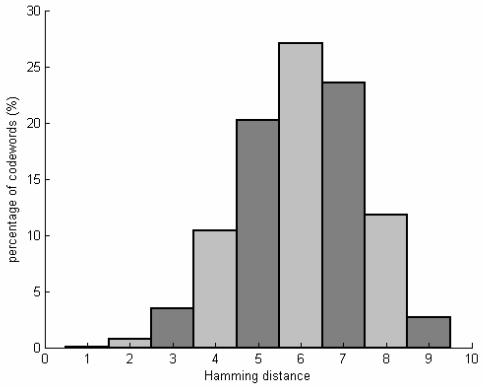
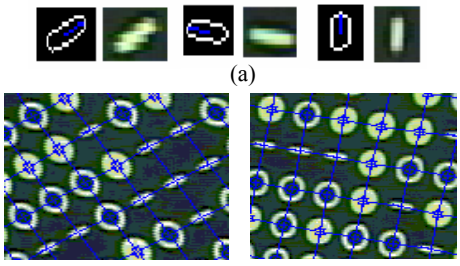


Figure 3: Hamming distance distribution.

So, the matrix that we designed verifies all the imposed criteria using only three symbols.

3.1. Pattern segmentation and decoding

The segmentation algorithm is based on an approach using the contours for detecting the symbols. First, the contours are detected then they 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. In practice, the center of each contour is estimated. The result of the contour detection for the symbol represented by a circle corresponds to two concentric circles. Consequently, to recognize this primitive, we have simply to search for two confused centers or, in practice, very close to each other. Concerning the remaining contours, the distances between the pixels of each contour and its center are computed in order to determine the closest pixel to the center and the farthest one. If the minimal distance is lower than an experimentally defined threshold, the primitive corresponds to a stripe; otherwise it corresponds to a disc. As for the stripe, the direction of the line passing by the center and the most distant pixel of the contour determines its local orientation (Figure 4-a). Then we can predict, thanks to this information, the directions to initiate the search for the neighbors of each symbol. Since the surfaces are locally smooth, their orientation varies slowly in a small neighborhood. That drastically simplifies the search of neighbors for each symbol while being reduced to a coarse estimated orientation (Figure 4-b).



(a) (b)

Figure 4: a) Stripe local orientation b) Searching the neighbors with respect to the local orientation.

Once the neighbors of each primitive are detected, its codeword is determined as illustrated in Figure 5. Therefore, an exhaustive search is carried out to decode the primitives (finding its location in the original matrix) taking advantage of the robustness of our pattern.

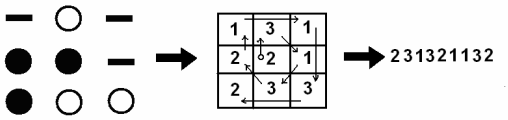
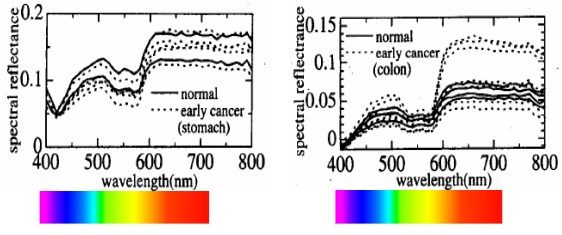


Figure 5: Determining the code of each window.

Then, we used the method of votes which determine the correct code of each primitive taking into account the decoding results of its neighbors. Finally, a projective reconstruction is performed to highlight the results on the observed images.

3.2. The choice of the pattern color

The foreseen application consists of acquiring images of internal structures of the abdomen. So the choice of the light color is very crucial because of the nature of the environment of the abdomen. Sambongi *et al.* [13] studied the influence of the cancer on the reflectance of the body organs at different wavelengths (Figure (6)).



(a) (b)

Figure 6: (a) The spectral reflectance of the stomach (b) The spectral reflectance of the colon (Sambongi *et al.*)

For both organs understudy (the stomach and the colon), the behavior of the reflectance stays the same passing by a maximum near the green wavelength and near the red wavelength. To assess these results, we undertook a study on light-organs interaction under the conditions of laparoscopic vision and we projected a matrix of color dots onto internal structures of a pig abdomen using a red laser source and a green one (Figure (7)).

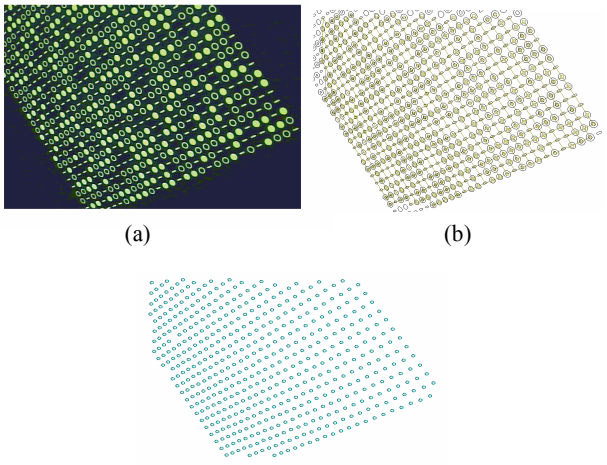
(a) (b)

Figure 7: The projection of a matrix of points onto the pig abdomen: (a) green points (b) red points.

We found that, contrary to the green color, the red color diffuses too much in the abdomen environment. That led us to choose a light source whose wavelength corresponds to the green color.

4. Experimental results

Firstly, experiments were undertaken to evaluate the choice of the pattern and to validate the proposed pattern and the detection and decoding algorithms in the laboratory environment. The system of validation is composed of a camera (JAI CV-S3200) which provides images of dimensions (568×760) pixels, and a video-projector (Sony VPL-CS6). We projected the pattern on a plan area (25×25 cm) (Figure 8-a), and we applied the detection and decoding algorithms previously described.



(a) (b)

(c)

Figure 8: a) The pattern projected on a plan b) Detection c) Decoding

Figure (8-b) shows the detected primitives, and in Figure (8-c) we present the primitives which were decoded correctly (correct attribution of row-column indices), which corresponds here to 99% of the detected primitives. We have also tested our pattern on other surfaces. For this purpose, we projected the pattern on a cylindrical surface (Figure 9-a).

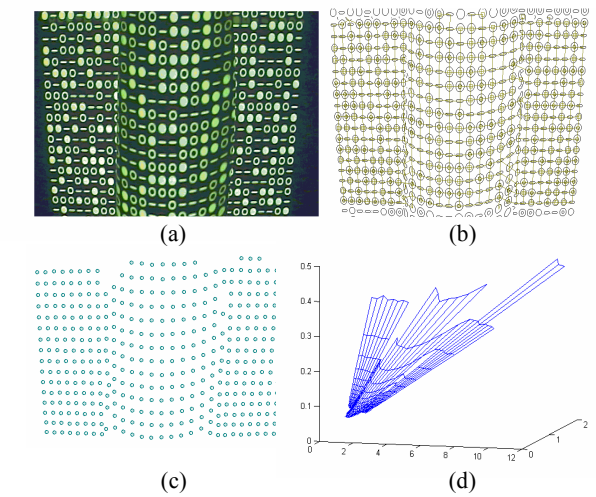


Figure 9: a) The pattern projected on a cylinder. b) Detection c) Decoding d) 3-D Reconstruction.

In spite of several deteriorations of the detection near strong discontinuities of the surface; the large majority of the detected primitives (95%) were decoded (Figure (9-b) and Figure (9-c)). The projective reconstruction of this surface is shown in Figure (9-d) and it attests the correct decoding. The whole procedure took 80ms on a P4 2.67GHz PC while it took 50ms when we worked with images of dimension (512×512). In table 1 below, we gather the results obtained for various conditions of projection in this stage.

Scene structure	detected primitives	decoded primitives	%
Plan1	756	754	99.7
Plan2	753	751	99.7
Cylinder	675	642	95.1

Table 1.

After having validated our pattern in the laboratory environment, we projected our pattern onto the internal structures of a pig abdomen and we acquired the images with an endoscopic camera (Figure (10-a)). In Figure (10-b), we present the result of testing our detection and decoding algorithms on these images. We can see that up to 75% of the primitives were detected (in blue) and up to 70% were decoded (in red). To evaluate the result, the projective construction was computed as it is shown in Figure (10-c) where the incorrect results are circled in red. The incorrect decoding is due to the errors that may occur in the stage of neighborhood detection and it is limited in the boundaries regions.

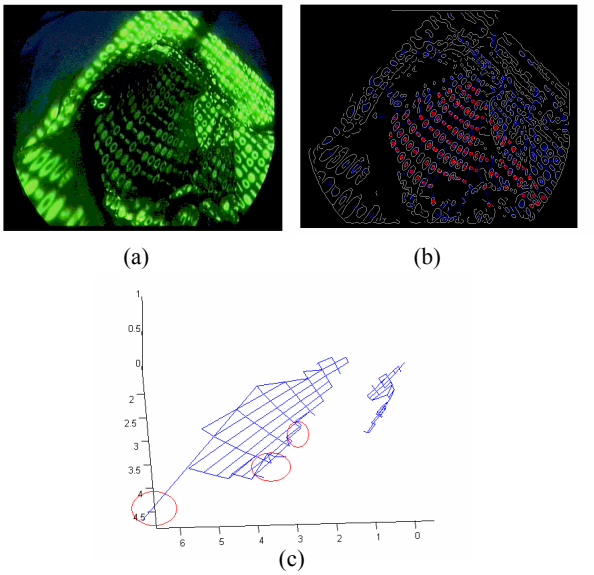


Figure 10: a) The pattern projected onto internal structures of pig abdomen b) Detection and decoding c) 3D reconstruction

As we aim to provide the surgeon with the 3-D information during the operation, we tested the pattern in the presence of an additional white light (Figure (11-a))

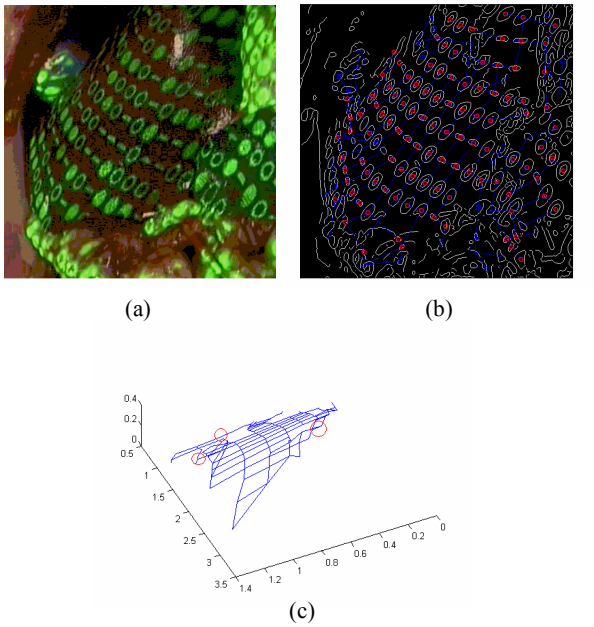


Figure 11: a) The projection onto internal structures of a pig abdomen with additional white light b) Detection and decoding c) The projective reconstruction

Here, we obtained good results during the detection stage which means that our algorithm works even in the presence of an additional white light. We had some errors in the stage of neighborhood detection which led to some errors in the decoding stage (Figure (11-b) and Figure (11-c)) but we can confirm that up to 85% of the detected primitives were decoded correctly.

5. Conclusion

We presented in this article a new robust pattern using monochromatic light. We adopted the coding strategy of spatial neighborhood based on the theory of M-array with three symbols. The final objective, which exceeds the framework of paper, is to reconstruct moving surfaces in real time and to integrate our pattern in an endoscopic tool using the diffractive optics technology. To this end, we adopted a method which models the pattern with simple geometrical primitives instead of color coding. The codewords were of length 9 with an average Hamming distance higher than 6. We considered three simple primitives; one of them carries the directional information. This simplifies the detection and speeds up the decoding. A traditional system of structured light was used to project our pattern on plane and cylindrical surfaces in the laboratory environment and the common projective reconstruction was employed to highlight the results. The first results show that more than 95% of the detected primitives are decoded correctly, even in the presence of partial occlusions. Then, we projected the pattern onto internal structures of a pig abdomen where we found that even in the presence of an additional white light, our algorithms gave a good results in decoding up to 80% of the detected primitives correctly.

The errors may be reduced by improving the neighborhood detection and the decoding algorithms taking advantage of the properties of our pattern. Then, a calibrated system must be used to evaluate the accuracy of the 3D metric reconstruction using the proposed pattern. The size of the area of interest, for the application of concern, is about (10x10) cm, and since the pattern is composed of (29x27) elements, the size of each projected primitive is typically of about (5x5) mm. Therefore, the dimensions of the primitives could be optimized since the scale of the primitives is a compromise between the number of primitives (783 in our case) and the minimal size allowing a good detection.

References

[1] J. Salvi, J. Pagès and J. Battle, "Pattern codification strategies in structured light systems", Pattern Recognition, Vol. 37, pp. 827-849, 2004.

[2] M. Ito and A. Ishii, "A three-level checkerboard pattern projection method for curved surface measurement", Pattern Recognition, 28 (1), pp. 27-40, 1995.

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

[4] N. Durdle, J. Thayyoor, and V. Raso, "An improved structured light technique for surface reconstruction of the human trunk", In IEEE Canadian Conference on Electrical and Computer Engineering, Vol. 2, pp. 874-877, 1998.

[5] J. Salvi, J. Battle and E. Mouaddib, "A robust coded pattern projection for dynamic 3d scene measurement", Int. Journal of Pattern Recognition Letters, Vol. 19, pp. 1055-1065, 1998.

[6] P. Lavoie, D. Ionescu and E. Petriu, "3D reconstruction using an uncalibrated stereo pair of encoded images", In Proceedings of the Int. Conf. on Image Processing, Vol. 2, pp. 859-862, 1996.

[7] L. Zhang, B. Curless and S. Seitz, "Rapid shape acquisition using color structured light and multipass dynamic programming", In Int. Symposium on 3D Data Processing Visualization and Transmission, Padua, Italy, June 2002.

[8] 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, Vol. 3, pp. 1237-1241, 2000.

[9] 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, Vol.20 (3), pp. 322 – 327, 1998.

[10] J. Pagès, Ch. Collewet, F. Chaumette and J. Salvi, "An approach to visual servoing based on coded light", In IEEE Int. Conference on Robotics and Automation, ICRA'06. 4118- 4123, 2006.

[11] S. Yee and P. Griffin, "Three-dimensional Imaging System", Optical Engineering 33(6), 2070–2075, 1994.

[12] A. Adan, F. Molina, and L. Morena, "Disordered patterns projection for 3D motion recovering", In Proceedings of the 2nd Int. Symp. on 3D Data Processing, Visualization, and Transmission, 3DPVT'04, pp. 262- 269, 2004.

[13] 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 Proceedings of the 22nd Annual EMBS International Conference, pp. 1026 – 1029, July 2000, Chicago.