

# 3-D Object Model Recovery From 2-D Images Using Structured Light

Philippe Lavoie, Dan Ionescu, and Emil M. Petriu, *Fellow, IEEE*

**Abstract**—Three-dimensional (3-D) object models are currently used in CAD/CAM, robotics, remote sensing, etc. The models (images) can be either directly acquired by using special devices such as range finders, CTR scanners, etc., or they can be recovered from a series of two-dimensional (2-D) images of the object. In this paper, the authors propose a method for determining a set of reference pixels in two simultaneous views of the same object, using two cameras, by projecting a pseudorandom encoded grid on the object. The grid nodes and their encoding values are extracted from 2-D images by applying first a smoothing and then a watershed algorithm. The pseudorandom information encoded in the grid nodes is used to match corresponding sets of points of the two 2-D images. The set of matched points are further used to calculate the disparity of each point of the object surface. Experimental examples illustrate the performance of this simple and elegant technique.

**Index Terms**—Computer vision, pseudorandom coding, stereo vision, structured light, three-dimensional (3-D) object modeling.

## I. INTRODUCTION

WITH THE ADVENT of new technologies, stereo vision has become an important field of research for both machine vision and graphic applications. One of the methods for three-dimensional (3-D) object reconstruction, known as computational stereo, is broadly defined as the recovery of 3-D characteristics of a scene from a series of images obtained from different points in the 3-D (Euclidean) space.

Barnard and Fischler [1] defined six steps necessary for stereo analysis: image acquisition, camera modeling, feature acquisition, image matching, depth reconstruction, and interpolation.

The 3-D content of a scene can be recreated using the “left” and “right” images of an object. In stereoscopy, the depth is estimated using triangulation [5]. The estimation process requires the knowledge of the global position and orientation of each cameras, the model of the cameras and the correspondence between all the feature points in both images. In our view, these steps can be further grouped into three more general stages which are: preprocessing, correspondence finding, and depth calculation.

Preprocessing involves finding the global position and orientation and the calibration of the cameras.

Establishing a correspondence between points from the two images is traditionally done using two different methods: area-based and feature-based matching. Both methods use the

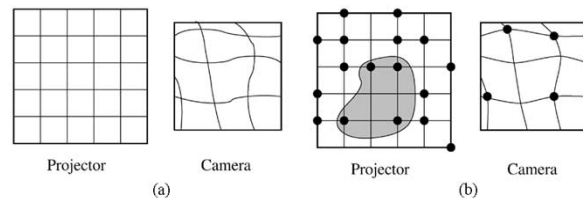


Fig. 1. Camera view of a grid. (a) The grid is not encoded, and the camera views the grid projected over a smooth object. (b) The grid is encoded with dots to form a PRBA, and the camera view of the PRBA is shown as a grey area over the grid.

epipolar constraint to limit their search space. The epipolar constraint implies that the points residing on an epipolar line of the left image will also be present on the corresponding epipolar line of the right image. This limits the search space to one dimension.

The image matching step is the most difficult one to solve. This step involves the determination of the correspondence between each point of the stereo pair.

In this paper, a new method for reconstructing 3-D images from the “left” and “right” two-dimensional (2-D) images of an object using a pseudorandom encoded structured light is introduced.

Based on a structured light grid, this new method allows for the introduction of a new procedure for the determination, of the so-called fundamental matrix as used in stereo vision and robotics, and a new technique for stereo fusion. Using a pseudorandom encoded mesh [3] of structured light, the procedure can determine with absolute precision a set of points on the object surface(s) on the “left” and “right” images.

The above proposed methods offer three distinctive advantages over a conventional stereo system.

- 1) It easily generates a list of matching points.
- 2) It adds structure to an object without textures.
- 3) It is less computational intensive.

In what follows, we will present the main algorithms of this procedure. In the Section II, we will present the pseudorandom grid methodology. In Section III, the information extraction process from the projected grid is outlined, while in Section IV, the matching of corresponding points from the “right” and “left” image is described. Eventually, in Section V, the depth calculating methodology, for each pixel on the object, is briefly exposed. Images obtained while recovering the 3-D image of a given object (a Greek goddess bust) illustrate the procedure. A conclusion section reviews the main steps of this paper.

Manuscript received July 29, 2003; revised December 6, 2003.

The authors are with the University of Ottawa, Ottawa, ON K1N 6N5 Canada.

Digital Object Identifier 10.1109/TIM.2004.823320

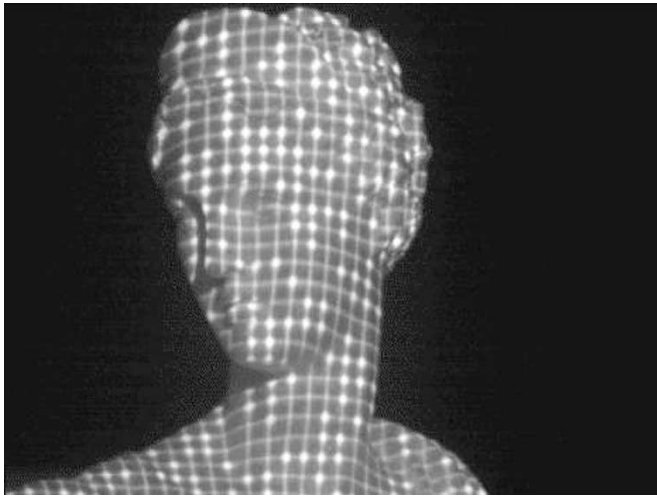


Fig. 2. Log of the image of the object with the projected PRBA.

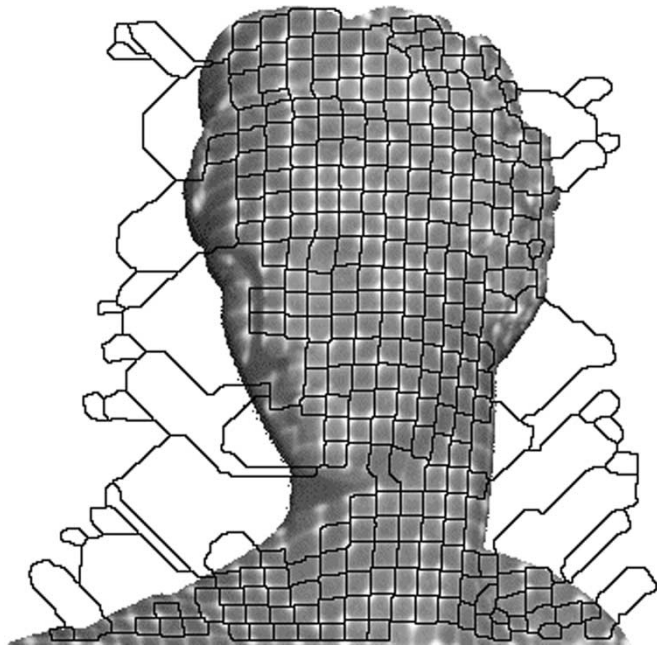


Fig. 3. Watershed result overlapping the log of the object.

## II. PLANAR PSEUDORANDOM ENCODED GRID

The techniques presented in this paper refer only to the 3-D object reconstruction using two 2-D views of the same object, obtained from two different cameras. By projecting a grid of pseudorandom encoded structured light on the object, the relative orientation of the two cameras is calculated from the correspondence of matched sets of points from the “right” and “left” image, and, eventually, one can obtain the depth of each pixel on the object surface.

When the grid is not encoded [as shown in Fig. 1(a)], the process becomes nearly impossible. When it is encoded with a pseudorandom binary array (PRBA) [as shown in Fig. 1(b)], a correspondence can easily be made since each small array of size  $k_1 \times k_2$  is unique in the PRBA; by simply looking at the

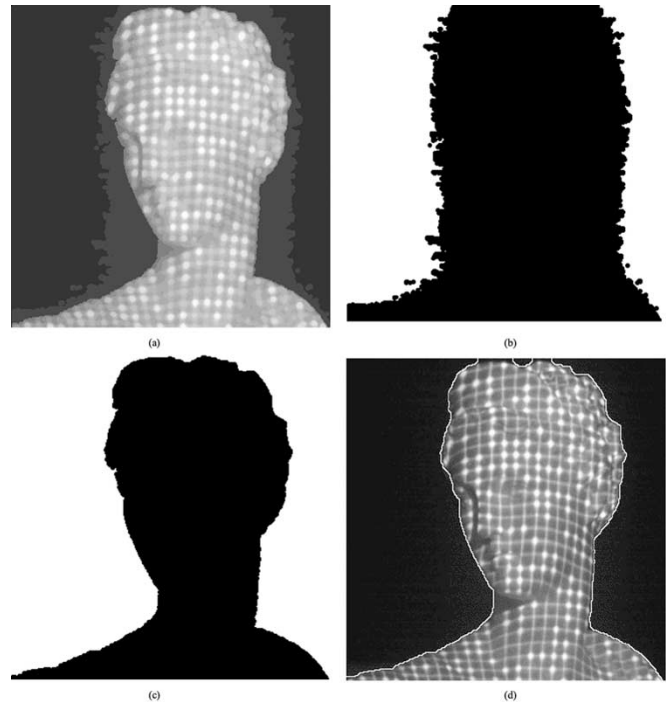


Fig. 4. Results of each background extraction process. (a) Opening the image with a disk of size 5. (b) Dynamic threshold of (a). (c) Dynamic threshold of (a) using (b) as a mask. (d) The smoothed result overlapping the log of the input image.

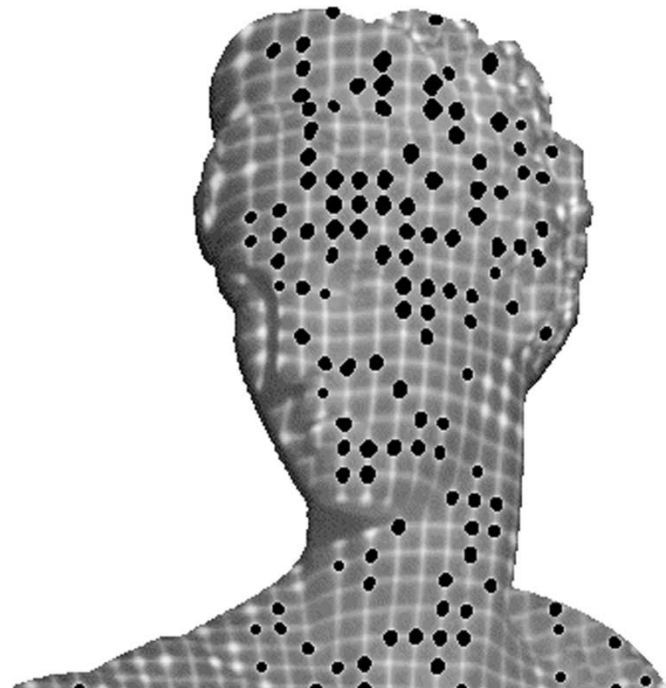


Fig. 5. Disks extracted superposed to the log of the input image.

pattern present in the small array, the exact position from the PRBA can be deduced. In the case shown in Fig. 1(b), a small array of  $3 \times 2$  uniquely defines a space of the PRBA.

A PRBA is defined by an  $n_1 \times n_2$  binary array encoded using a pseudorandom sequence such that a  $k_1$  by  $k_2$  window sliding

over the array is unique and fully identifies the window's absolute coordinate  $(i, j)$  within the array. The following relations hold for the PRBA [2]

$$2^2 - 1 = 2^{k_1 k_2} - 1 \quad (1)$$

$$n_1 = 2^{k_1} - 1 \quad (2)$$

$$n_2 = \frac{2^n - 1}{n_1} \quad (3)$$

where  $n_1$  and  $n_2$  must be relatively prime.

The process by which the "primitive polynomials modulo 2" method generates random bits from a primitive polynomial is described by Press and al. [4].

### III. EXTRACTING THE GRID INFORMATION

In this section, a description of how the PRBA projected onto the object is extracted is presented. In Fig. 2, the object with the projected PRBA is shown. The extraction process is done in two steps: extracting the grid and extracting the disks.

The extraction of the grid from the image (Fig. 2) is done in two steps: a smoothing and an extraction process.

The smoothing consists of applying a morphological closing operator with a disk structuring element of size 5. A greater disk size will tend to over-smooth the image and some of the grid information will be lost.

The information extraction process consists in detecting first the object, then the grid within the object borders, and eventually in the interpretation of the pseudorandom pattern.

The grid extraction process starts with the application of a watershed algorithm. The grid extraction process, as obtained by applying the watershed method, is shown in Fig. 3 overlapping the input image.

The background information is used to extract the disks from the grid. Extracting the background can be seen as detecting the foreground. The foreground corresponds to the object of interest. Detecting the object is done in three steps: filtering, thresholding, and smoothing.

Filtering is used to eliminate small bright areas in the background. It is done by applying a morphological opening operator with a disk structuring element of size 5.

Thresholding is done in two steps. The first step is performed dynamic threshold of the image. This creates a mask. In the second step, this mask is used to do another dynamic threshold on the image. The mask limits the region for which a dynamic threshold value is calculated.

Smoothing involves a morphological closing operator with a disk of size 11. This operation closes the small gaps in the foreground.

The object extraction steps are illustrated in Figs. 4 and 5.

Disk extraction is a two steps process: opening and thresholding.

The morphological opening operator is applied using a disk of size 5. This tends to diminish small bright regions (for example, the intersection of two lines) while preserving the larger ones (the disks).

A global threshold is applied using the background image as a mask. The mask defines the regions of interest in the images

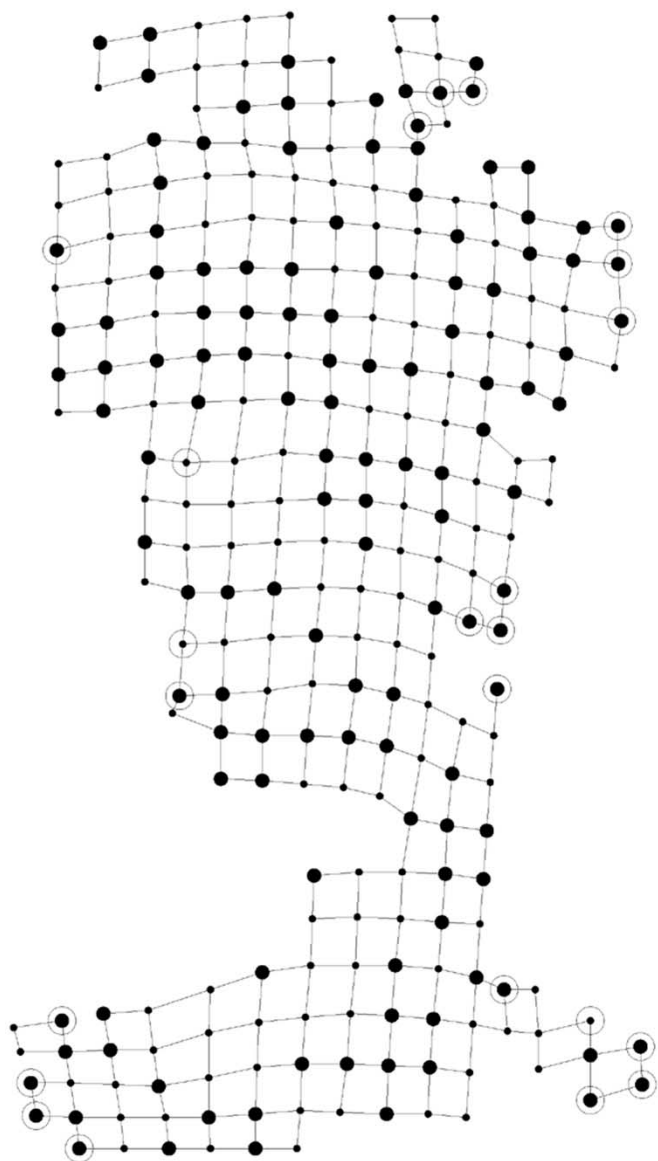


Fig. 6. Intersections found are shown as black points in the image.

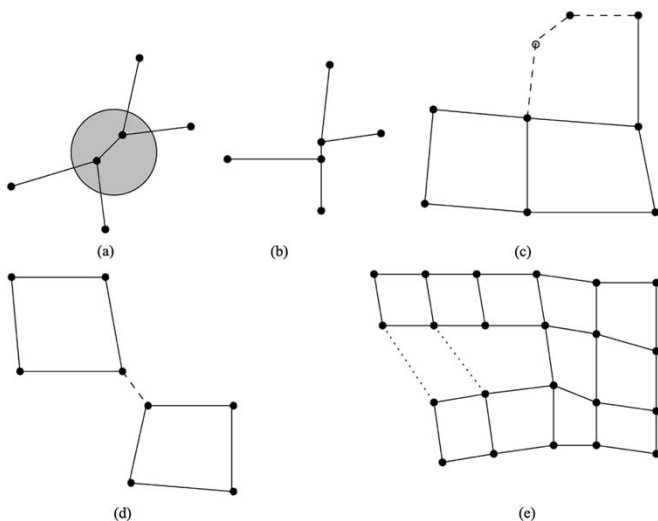


Fig. 7. Situations for which an intersection or a connection must be removed. (a) Two points in the same disk. (b) Two points close to each other. (c) Points not part of a square. (d) Connection not part of a square. (e) Connections not agreeing with the setup.

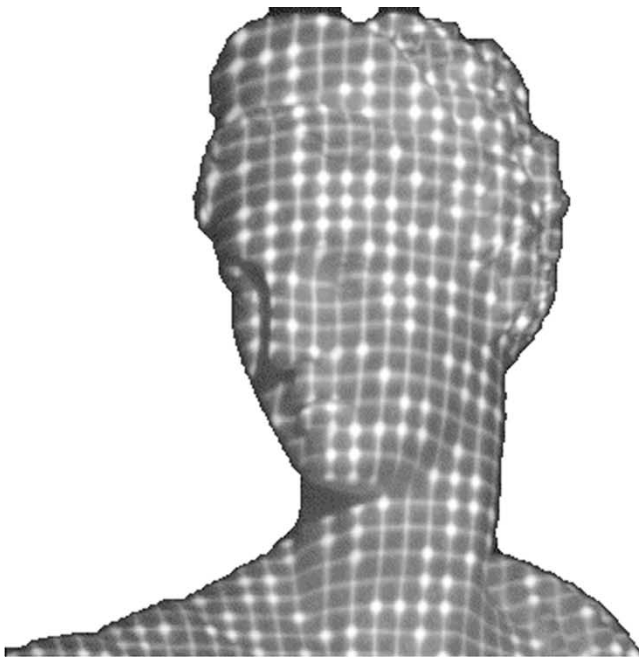


Fig. 8. Log of the input image with the label points obtained after the first step.

(the object). When limited to the masked region, the threshold operator will pickup only the brighter spots on the object.

This information cannot be used directly. More information on the precise location of the intersection points, their connectivity and their type is needed. To obtain this information, many steps are required.

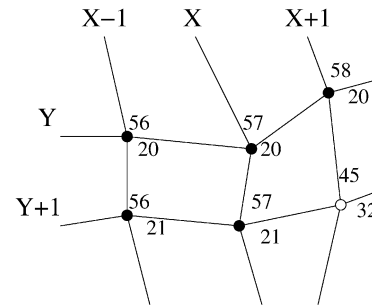
The intersections are found by using the morphological erosion operator. This operator removes all points in the grid which do not belong to a kernel. The kernels used are T-shaped kernels at  $90^\circ$  angles. The union of the erosion operator from all the kernels gives the location of all the intersection points. Fig. 6 shows all the intersections found by the kernels.

From the figures above, one can see that too many intersection points and too many connections are found. The following sequential steps remove or improve the location and/or the connectivity information:

- 1) the merging of the intersections:
  - if two or more intersections are part of the same disk [Fig. 7(a)];
  - if two intersection points are close to each other [Fig. 7(b)];
- 2) the removal of intersection points if they are not part of a square [Fig. 7(c)];
- 3) the removal of the connections if they are not part of the same square [Fig. 7(d)];
- 4) the removal of a connection if it does not agree with the grid setup. The grid setup defines how the grid points are related to each other. If a point is defined to be at  $(i, j)$ , then the point above it should be defined to be at  $(i-1, j)$ . If this is not the case, then the connection is removed. In Fig. 7(e), the two dotted connections are removed because they are not agreeing with the darker connections. The left connection is removed since it makes an upward connection between the point  $(i, j)$  and  $(i-1, j-1)$ , the right connection is eliminated for a similar reason.



Fig. 9. White point (45,32) is not compatible with the other labeled points.



#### IV. MATCHING THE “RIGHT” AND “LEFT” IMAGES

Matching the grid to the PRBA is done in three steps: 1) labeling  $k_1 \times k_2$  areas to their corresponding location in the PRBA, 2) finding the single intersection point which has the largest number of support and, finally, 3) spreading the information from that point.

The first step is to find all  $k_1 \times k_2$  areas in the grid. In the example in Fig. 2, the area is a  $2 \times 7$  region. The area is associated with a binary number of length  $k_1 k_2$ : the upper left corner is the lowest bit, the lower right corner is the highest bit. This number is compared to a lookup table for which a  $x$  and  $y$  location is associated in the PRBA. Each upper left corner of the  $k_1 \times k_2$  regions are labeled with their corresponding  $(x, y)$  location in the PRBA.

In this first step, a PRBA location for all the  $k_1 \times k_2$  regions in the grid is found. Due to errors in finding if a disk is present or not at a certain location on the grid, the result of the operation yields incompatible matches. An incompatible match means that two neighbors in the grid are not neighbor in the PRBA. The result of the first operation is shown in Fig. 8.

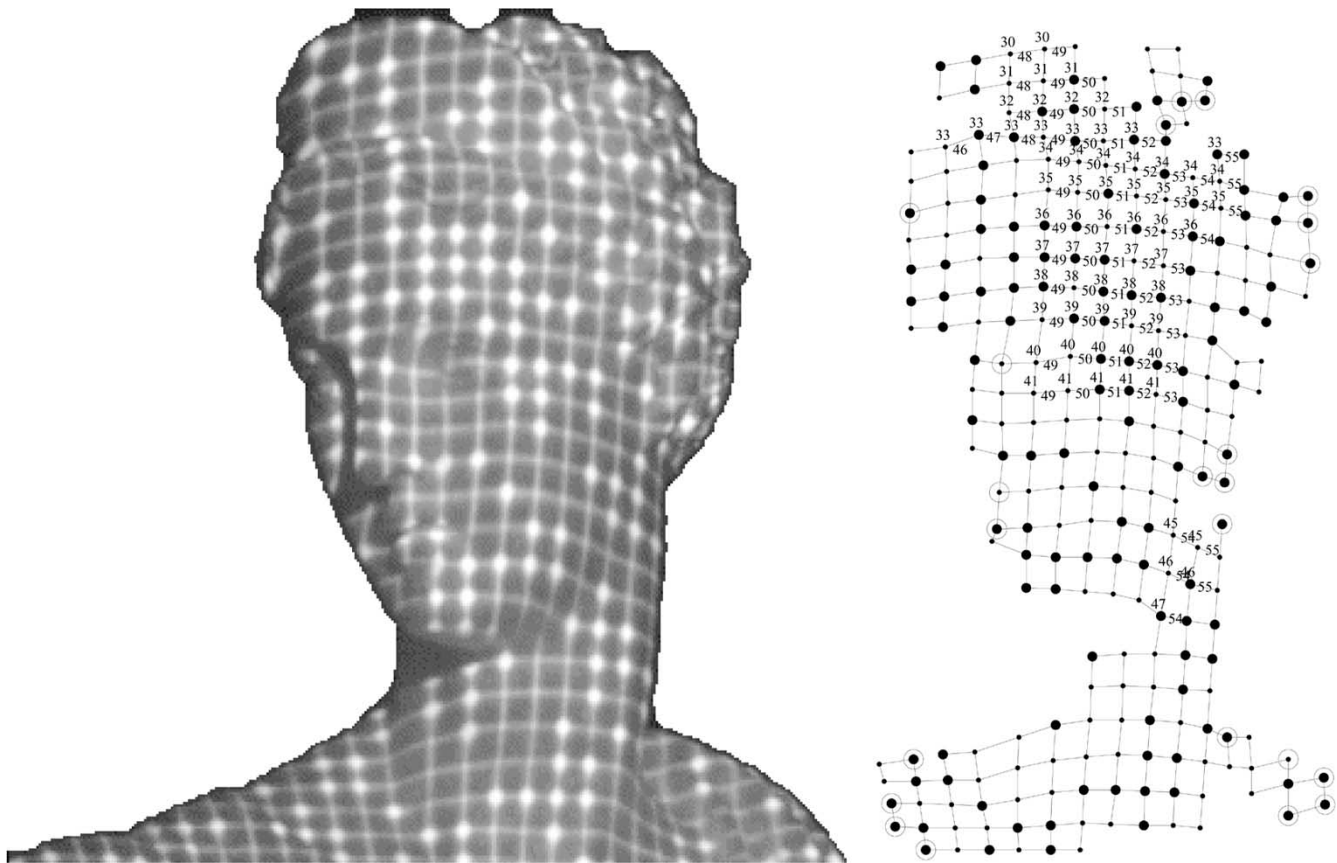


Fig. 10. Result of spreading the pseudorandom information from the point  $P$ .

The second step is to find an intersection point  $P$  which has the largest number of support. The number of support is the number of labeled points which are compatible with themselves. Being compatible with a point means that if a point is labeled  $(x, y)$  at a grid location  $(X, Y)$ , then the grid point at location  $(X + i, Y + j)$  must be labeled  $(x + i, y + j)$ .

In Fig. 9, all the black dots are compatible. However, the white point labeled  $(45, 32)$  at  $(X + 1, Y + 1)$  is not compatible with the other labeled points at  $(X, Y)$ ,  $(X + 1, Y)$ ,  $\dots$

The final step is to spread the information from  $P$ . This operation consists in setting, for all the points in the grid, a label for that point in accordance with the label at the point  $P$ . Sometimes, this operation will label two different points identically. If this happens, they are both removed from the set of labeled points. The result of this operation is shown in Fig. 10. In this figure, the only label points shown are the points which were originally compatible with  $P$ . This operation also spreads the type information (a 0 or a 1). The intersections where an error was originally made about the type are encircled.

In the Fig. 10, there are 264 intersections and 23 errors were originally made in the determination of the presence (or absence) of a disk. It should be noted that if there are too many errors in finding if a disk is present or not on an intersection, then it is possible that the grid points will be mislabeled.

## V. CALCULATING THE DEPTH OF EACH POINT

The depth of each point on the object surface is calculated using the disparity map calculated for the “right” and “left” images. When two images of an object with a PRBA projected onto it are taken, a disparity map can be easily obtained for the two images. The disparity of a point gives a scaled version of its 3-D location. Considering the two images of the same object as above, and extracting the information contained in the PRBA as shown in Fig. 11, and using the linear interpolation method as depicted in Fig. 12(a) one obtains the disparity value of the grid points if matching grid points exist in the other image. The disparity value is set to the absolute distance between the location of the point in the left image to its corresponding point in the right image.

A linear interpolation between disparity points of the grid is then be computed.

The linear interpolation method is shown in Fig. 12(a). The coarse disparity obtained from matching the two images is shown in Fig. 12(b). The coarse disparity gives a scaled 3-D perspective of the object. The lighter points corresponds to points which are closer to the left camera and the darker ones are farther away from the left camera. The points for which there is no 3-D information available are shown in black. For these points an interpolation method gives their disparity and, therefore, their 3-D location on the object image. The intensity level in this disparity map will eventually give the depth of each pixel.

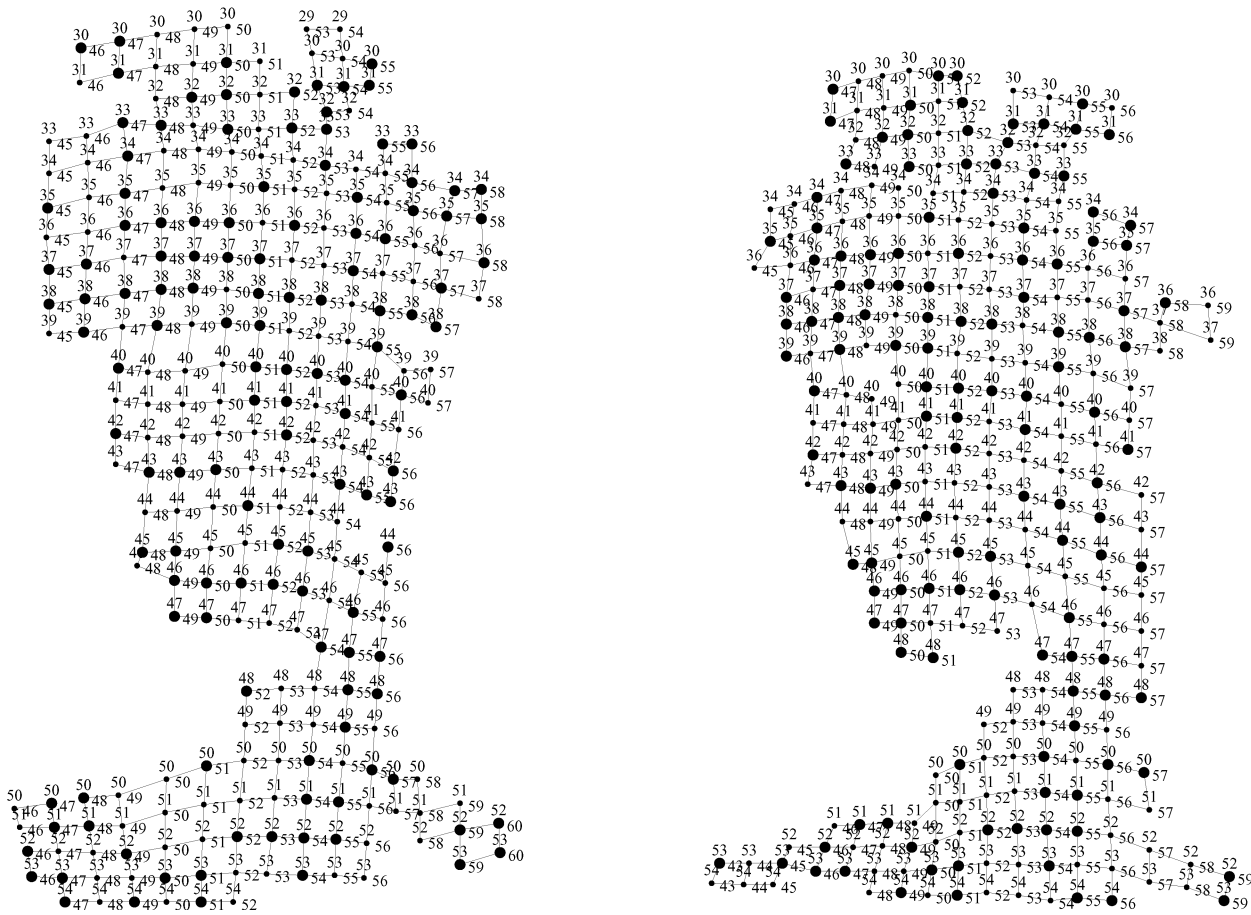


Fig. 11. Two PRBAs extracted from the images above.

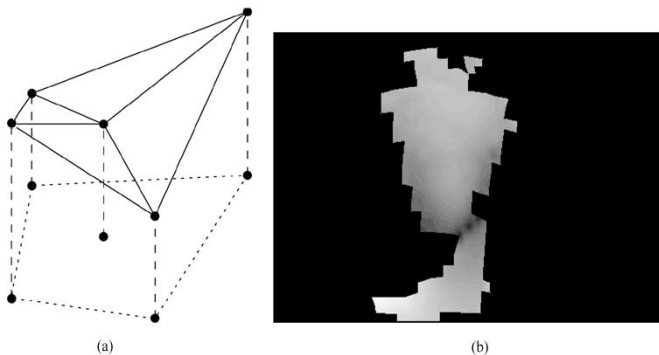


Fig. 12. Coarse disparity obtained from two PRBAs. (a) Linear interpolation for four connected grid points. (b) Coarse disparity obtained.

## VI. RESULTS

This paper focuses on the generation of a coarse disparity map. Results scaled in 3-D are shown to better illustrate the results obtained with the technique. The first set of results (shown in Fig. 13) is of the Greek goddess object which has 239 matching intersections.

The second test is done with a simple cup object. Fig. 14 shows the two input cup images and the scaled 3-D version of the object. This object has 460 matching intersection points between the left and right image.

The system doesn't recognize the handle of the cup since it requires a contiguous  $2 \times 5$  region to label a point properly. Errors

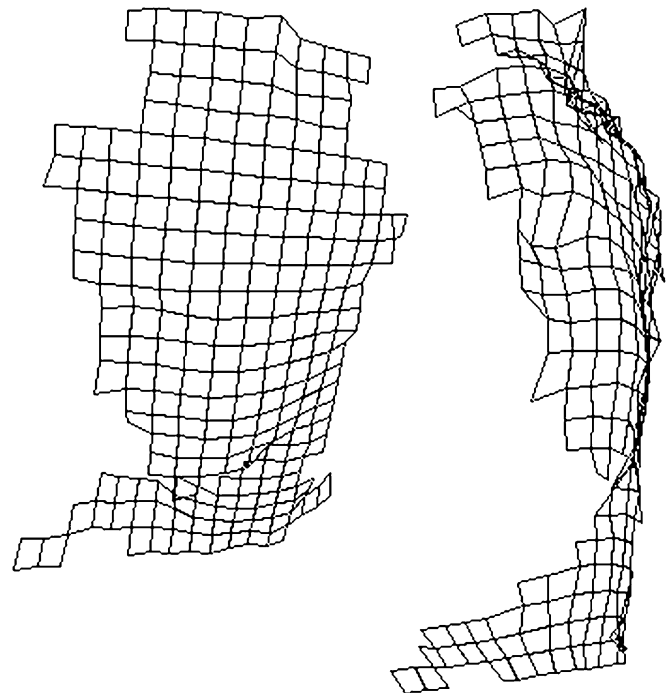


Fig. 13. Different views of the Greek goddess object in scaled 3-D.

in localizing the center of a disk are clearly seen in the scaled 3-D reconstruction of the cup object.

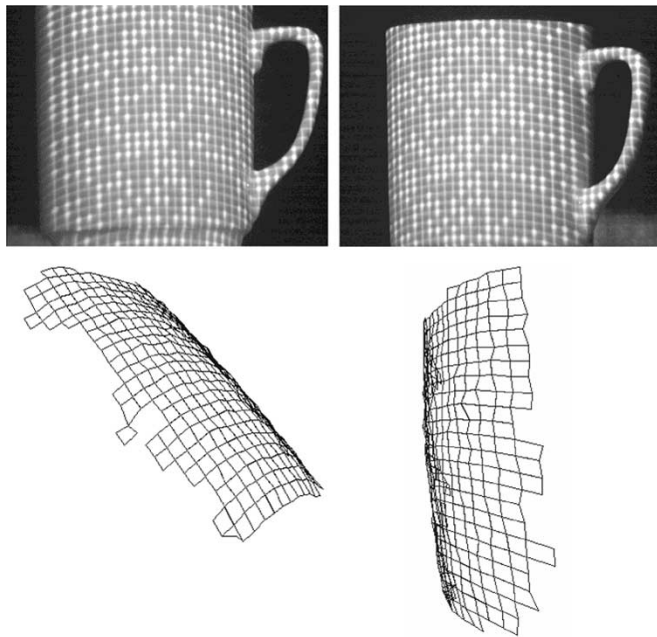


Fig. 14. Input cup images and different views of the cup object in scaled 3-D.

## VII. CONCLUSION

The problem with recovering the 3-D image of an object from two 2-D views of the same object is solved in this paper by using a structured light environment consisting of a pseudorandom encoded grid which is projected on the object. The random pattern encoded in the grid facilitates the matching of the similar points situated on the two 2-D images. The sets of matched points help in determining the disparity map and, eventually, a scaled value of the depth of each point on the object surface. The precision of the method is related to the precision of the PRBA projected on the object; the coarser the grid, the lower the accuracy. This precision can be considerably improved by further processing the two images. Among some of the algorithms which could be used to further process the images, one can mention the determination of the fundamental matrix, the dynamic programming, and autoregressive filtering techniques. The dynamic programming principle is used to perform the stereo fusion of all points contained in the two images. These post-processing aspects will form the subject of a further paper.

We can conclude stating that the method presented above is suitable for applications where the grid distance is well adapted to the variations of the object surface.

## REFERENCES

- [1] S. T. Barnard and M. A. Fischler, "Computational stereo," *ACM Comput. Surveys*, vol. 14, no. 4, pp. 553–572, Dec. 1982.
- [2] F. J. MacWilliams and N. J. A. Sloane, "Pseudo-random sequences and arrays," *Proc. IEEE*, vol. 64, pp. 1715–1729, Dec. 1976.
- [3] E. M. Petriu, D. Ionescu, D. C. Petriu, S. K. Yeung, P. Lavoie, and N. Trif, "Absolute position measurement applications of pseudo-random encoding," in *Proc. ETIMS IEEE Int. Workshop on Emergent Technologies for Instrumentation and Measurements*, Como, Italy, 1996, pp. 119–126.
- [4] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992, p. 298.
- [5] S. Roy, "Analyse D'images Stéréoscopique Basée Sur La Détermination Du Flux Optique," M.S. thesis, Univ. Montreal, Montreal, QC, Canada, 1992.

**Philippe Lavoie** received the B.A.Sc. and M.A.Sc. degrees from the Department of Electrical Engineering, University of Ottawa, Ottawa, ON, Canada, in 1993 and 1996, respectively.

His research interests are in the area of 3-D object reconstruction.

**Dan Ionescu** received the Ph.D. degree in control and computers from the Polytechnic Institute of Bucharest, Bucharest, Romania.

His research interests focus on artificial intelligence, image processing, and multimedia applications.



**Emil M. Petriu** (M'86–SM'88–F'01) received the Dipl.Eng. and Dr.Eng. degrees from the Polytechnic Institute of Timisoara, Romania, in 1969 and 1978, respectively.

He is a Professor at the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, where he has been since 1985. His research interests include test and measurement systems, interactive virtual environments, intelligent sensors, robot sensing and perception, neural networks, and fuzzy control. During his career, he has published more than 180 technical papers, authored two books, edited two books, and received two patents.

Dr. Petriu is a Fellow of the Canadian Academy of Engineering and Fellow of the Engineering Institute of Canada. He is currently serving as a Member of the AdCom, and Chair of TC-15 Virtual Systems and Co-Chair of TC-28 Instrumentation and Measurement for Robotics and Automation of the IEEE Instrumentation and Measurement Society. He is an Associate Editor of the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT and member of the editorial board of the IEEE INSTRUMENTATION AND MEASUREMENT MAGAZINE. He is corecipient of the 2003 IEEE Donald G. Fink Prize Paper Award.