Pergamon

Pattern Recognition, Vol. 31, No. 7, pp. 963–982, 1998

0031-3203/98 $19.00 + 0.00

PII: S0031-3203(97)00074-5

RECENT PROGRESS IN CODED STRUCTURED LIGHT AS A TECHNIQUE TO SOLVE THE CORRESPONDENCE PROBLEM: A SURVEY

J. BATLLE,*† E. MOUADDIB‡ and J. SALVI†

†Computer Vision and Robotics Group Universitat de Girona, Avda. Lluis Santalo, s/n. 17071 Girona, Spain

‡Laboratoire des Systèmes Automatiques, Université de Picardie Jules Verne, 7, Rue du Moulin Neuf. 80000 Amiens, France

(*Received* 5 *December* 1996)

Abstract—We present a survey of the most significant techniques, used in the last few years, concerning the coded structured light methods employed to get 3D information. In fact, depth perception is one of the most important subjects in computer vision. Stereovision is an attractive and widely used method, but, it is rather limited to make 3D surface maps, due to the correspondence problem. The correspondence problem can be improved using a method based on structured light concept, projecting a given pattern on the measuring surfaces. However, some relations between the projected pattern and the reflected one must be solved. This relationship can be directly found codifying the projected light, so that, each imaged region of the projected pattern carries the needed information to solve the correspondence problem. © 1998 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Pattern projection Correspondence problem Active stereo Depth perception Range data Computer vision.

1. INTRODUCTION

When 3D information of a given surface is needed,[1] we have to choose between a passive method and an active one.[2] The most widely known passive method is stereovision[3] which can be achieved by two different ways. In the first way, an optical sensor is moved to known relative positions in the scene. In the second way, two or more optical sensors are previously fixed in known positions.[4, 5] The surface to be measured is projected on the image plane of each sensor through each focal point. As described in this paper, 3D coordinates of the object point can be obtained by trigonometry[6–8] from the known projections of an object point and, furthermore, from the relationship between the optical sensors. But we have to know, with correctness, for each object point, its projections on the optical sensor image planes. In fact, in order to obtain the 3D coordinates of a given point from n given projections (one from each sensor), these projections have to be necessarily from the same object point. This problem is known as the correspondence problem.

The correspondence problem can be considerably alleviated by an active method.[9–13] One of the most widely used active methods is based on structured light projection.[14] Normally, only one camera is used to image the projection of a given pattern on the measuring surface. 3D information manifests itself in the apparent deformations of the imaged pattern from the projected one.[15] Analysing these deformations we can get information about the position, the orientation, and the texture of the surface on which the pattern has been projected. The 3D analysis can also be quite alleviated using Coded Structured Light. This technique allows us to know for each imaged point, its original point on the emitted projector plane. Then, the correspondence problem can be directly obtained, so it is not necessary to use geometrical constraints to get it. This technique is the main subject of the paper.

The survey is structured as follows: Firstly, a stereoscopic vision system made by two optical sensors is explained. In order to obtain the 3D coordinates of the measuring object from its projections, the mathematical equations and the geometrical constraints are analysed. Secondly, structured light as a method to solve the correspondence problem is presented, and we introduce the whole mathematical basis for a general system made by the relationship between a camera and a projector. Then, the purpose of the coded structured light is described, and a new classification is presented, analysing temporal dependence, emitted light dependence, and depth surface discontinuity dependence. Finally, several coded structured light techniques, presented in recent years, are shown and discussed.

2. STEREOVISION

A well-known stereoscopic vision system consists of one optical sensor (or camera) which can be moved, so

*Author to whom correspondence should be addressed. E-mail: jbatlle@ei.udg.es; qsalvi@ei.udg.es.

its relative positions are known at any time. A more common stereoscopic system is done by two or more sensors maintaining always the same known positions relative to each other. If more than one sensor is used, it will be supposed that they are all identical, and each one can be modelled as an ideal camera, often called a *pinhole* model.[16]

In practice, a camera does not run as a *pinhole* model as it is achieved by a set of lenses. As opposed to the *pinhole* model, distances between the lens centre and the image plane are not constants, and not equal to the focal length; of course, this system is rather difficult to be modelled. Even for measuring objects placed at large distances relative to the focal length, it can be assumed that the distance between the lens centre and the image plane is constant for any projection.

In the following discussions we will assume a general stereoscopic system made by the relationship between two cameras, as shown in Fig. 1. In such a system, we have defined a general coordinate system with origin $O = O_1$, and co-ordinates vectors (x, y, z) equal to (x_1, y_1, z_1), respectively. We suppose that optical sensors can be modelled as a *pinhole* model, with an optical axis which coincides with the z_1 axis, and a focal point F_1 placed at a distance f_1 from the origin along the same axis. Furthermore, we assume that the image plane π_1, with centre O_1, lies on the x_1y_1 image plane. Obviously, the second camera has the same properties as the first one, with a coordinate system (x_2, y_2, z_2) and origin O_2, where z_2 coincides with its optical axis, a focal distance f_2 along z_2, and an image plane π_2 which lies on the x_2y_2 image plane. This second coordinate system can be related to the first one using

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \mathbf{T}, \tag{1}$$

where $\mathbf{R}$ is a 3×3 orthonormal matrix, which describes the rotation, and $\mathbf{T}$ is a vector which describes the translation of the O_2 coordinate system with respect to O_1.

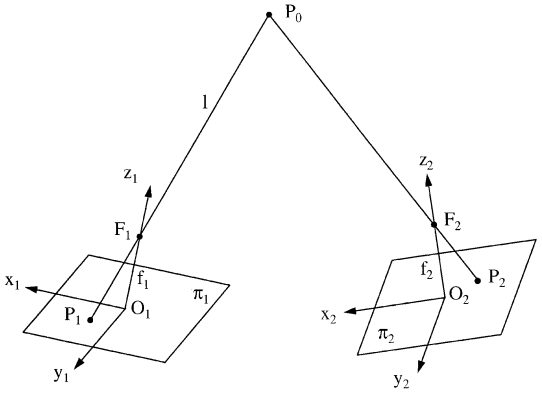

Fig. 1. A general stereoscopic system made by the relationship between two optical sensors.

Consider now an object point P_0 in the 3D space, which has a projection P_1 on the π_1 image plane, where the coordinates of P_0 with respect to O_1 are (x_{p0}, y_{p0}, z_{p0}), and the coordinates of P_1 with respect to O_1 are $(x_{p1}, y_{p1}, 0)$, then the projection P_1, which corresponds to the object point P_0 has to be projected through the focal point, so

$$P_1 = F_1 + \alpha(P_0 - F_1) \tag{2}$$

which can be expressed using a matrix relation as,

$$\begin{bmatrix} x_{p1} \\ y_{p1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ f_1 \end{bmatrix} + \alpha \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} - f_1 \end{bmatrix} \tag{3}$$

so, the two coordinate systems can be related as follows:

$$\frac{x_{p0}}{x_{p1}} = \frac{y_{p0}}{y_{p1}} = \frac{f_1 - z_{p0}}{f_1} \tag{4}$$

Consider now a second projection of P_0 on the image plane π_2, called P_2, with coordinates $(x_{p2}, y_{p2}, 0)$ with respect to O_2. If $(x'_{p0}, y'_{p0}, z'_{p0})$ coordinates express P_0 with respect to O_2, both co-ordinates systems can also be related as follows:

$$\frac{x'_{p0}}{x_{p2}} = \frac{y'_{p0}}{y_{p2}} = \frac{f_2 - z'_{p0}}{f_2}. \tag{5}$$

So, if we suppose that P_0 is an object point, the described formulas (1), (4) and (5) can be used to take out the (x_{p0}, y_{p0}, z_{p0}) or $(x'_{p0}, y'_{p0}, z'_{p0})$ positions from projections P_1 and P_2. Obviously, before the calculation of the P_0 co-ordinates, the most important matter is to ensure that P_1 and P_2 are projections of the same object point P_0.

We choose an image point P_1 on π_1, as the projection of an object point P_0 along the **l** line, defined from F_1 and P_0. As P_0 can be placed at any point on the **l** line, there is no unique position of P_2 on π_2, if we only want to apply geometrical constraints. However, we can affirm that the projection P_2 on π_2 has to lie on the segment made by the intersection between the plane defined by the points F_1, F_2 and P_1 and the image plane π_2. This constraint is known as the epipolar constraint.

If the projection point P_1 on π_1 is known, the epipolar constraint allows us to find the projection point P_2 on π_2 in only one direction. Of course, we cannot know whether the P_0 projection has been imaged on the π_2 image plane. It may be occluded by any other surface object of the scene, or it may be projected out of the range of the camera. The relationship between both projections is known as the correspondence problem, one of the most interesting topics in stereo.

3. STRUCTURED LIGHT

The mathematical solution of the correspondence problem can be simplified leaving out passive

methods such as stereo, and going to an active method based on the structured light concept. In this method the second stereo camera is replaced by a light source, which projects a known pattern of light on the measuring scene. A single camera images the illuminated scene. The required 3D information can be obtained by analysing the deformations of the imaged pattern with respect to the projected one. Of course, some correspondences between the projected pattern and the imaged one should be solved.

If a single light dot or a slit line is projected on the scene, then, there is no correspondence problem to be solved, but all the scene has to be scanned to obtain the 3D map. Shirai and Suwa in 1971, proposed a slit line projection to recognise polihedric objects.[17] In 1973, Agin and Binford generalised this idea to recognise curvilinear objects.[18] Two years later, Popplestone, Brown, Ambler and Crawford proposed a more general system which recognises either polihedrics or curvilinear objects.[19] In 1986, Yamamoto, Sato and Inokuchi[20] proposed a half plane illumination system instead of a slit line. In fact, binarizing an image of a scene illuminated by a half plane pattern corresponds to the boundary edge detection between the illuminated and the obscured area. Recently, some systems to obtain 3D maps of a scene have been presented. The most common systems are similar to the high-speed method presented by Ozeki, Nakano and Yamamoto,[21] and the video rate method presented by Yokoyama, Sato, Yoshigahara and Inokuchi.[22] There are also some authors, as Sato, Kitagawa and Fujita,[23] who project two slit lines with different orientation and position in the 3D coordinates system. In a similar way, Kemmotsu and Kanade[24, 25] chose to project three lines on the measuring scene.

In order to improve the accuracy of the system, an alternative way is to project a grid of dots or lines on the scene to cover the entire range of the camera. Asada *et al.* proposed to use a pattern made by a set of vertical, parallel and equidistant, stripe lines.[26–28] Wang *et al.*, have extended Asada's idea with the sequential projection of two orthogonal stripe patterns.[29–31] Furthermore, in order to obtain 3D surface properties, Hu *et al.*,[32,33] Stockman and Hu[34] and Shrikhande and Stockman[35] have proposed the widely known projection of a grid. Other authors also use the information given by an intensity image to obtain, with a better accuracy, the boundary edges of the scene.[36–40] Then, an easier correspondence problem has to be solved, we have to identify, for each point of the imaged pattern, the corresponding point of the projected pattern.

All these methods allow us to find out 3D information from geometric constraint propagation, and some of them are rather limited for measuring surfaces with depth discontinuities.

The correspondence problem can be directly solved codifying the projected pattern,[41] so each projected light point carries some information. When the point is imaged on π_1, this information can be used to determine its coordinates on π_2, from where it has been emitted.

In the following, the whole mathematical basis used to obtain 3D range information of a measuring scene is presented. The formulated equations have been adapted to the projection of an image and the capture of another, but, as can be seen, there is not much difference from the triangulation used in a stereoscopic system.

4. SURFACE MEASUREMENT

4.1. General principle

To understand the surface measurement principle we have considered it useful to explain the mathematical basis used to obtain 3D information from triangulation. For further details the reader may refer to references (6–8).

Consider a global coordinate system O, placed on the centre of the captured image, as shown in Fig. 2. The focal point of the captured image system is located at $F_1 = (0, 0, f_1)^t$. Then, given an object point $P_0 = (x_{p0}, y_{p0}, z_{p0})^t$, the corresponding point on the image plane, $P_1 = (x_{p1}, y_{p1}, 0)^t$, must be on the line which starts in P_0 and passes through F_1, so

$$P_1 = F_1 + \alpha(P_0 - F_1) \tag{6}$$

which can be expressed in a matrix form as

$$\begin{bmatrix} x_{p1} \\ y_{p1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ f_1 \end{bmatrix} + \alpha \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} - f_1 \end{bmatrix}. \tag{7}$$

Assume a second image plane, as shown in Fig. 2. Furthermore, assume that the origin of this second coordinate system has the same orientation as the first one, and moreover, it is also centred with respect to

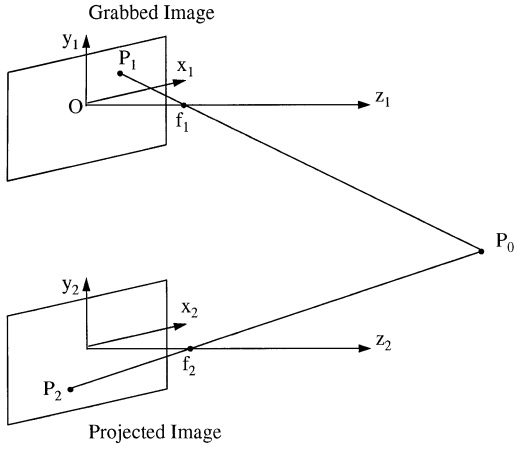

Fig. 2. Measurement system used in structured light

the image plane. So, with respect to O, it is located at

$$O_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}. \tag{8}$$

Suppose that the focal point, with respect to O_2, is located at the position,

$$F_2 = \begin{bmatrix} 0 \\ 0 \\ f_2 \end{bmatrix}. \tag{9}$$

With reference to the global coordinate system O, F_2 is located at

$$F_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 + f_2 \end{bmatrix}. \tag{10}$$

Then, if it can be supposed that the point imaged on the first image plane has been projected by the second imaging system, the image point $P_2 = (x_{p2}, y_{p2}, 0)^t$ with respect to the global coordinate system can be expressed as

$$P_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} + \begin{bmatrix} x_{p2} \\ y_{p2} \\ 0 \end{bmatrix} = \begin{bmatrix} x_2 + x_{p2} \\ y_2 + y_{p2} \\ z_2 \end{bmatrix}. \tag{11}$$

Since the same object point is considered, then,

$$P_2 = F_2 + \beta(P_0 - F_2) \tag{12}$$

$$\begin{bmatrix} x_2 + x_{p2} \\ y_2 + y_{p2} \\ z_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 + f_2 \end{bmatrix} + \beta \begin{bmatrix} x_{p0} - x_2 \\ y_{p0} - y_2 \\ z_{p0} - z_2 - f_2 \end{bmatrix} \tag{13}$$

So, summarising, from equations (7) and (13) we obtain

$$x_{p0} = \frac{f_1 - z_{p0}}{f_1} x_{p1}, \tag{14}$$

$$y_{p0} = \frac{f_1 - z_{p0}}{f_1} y_{p1}, \tag{15}$$

$$x_{p0} = x_2 + \frac{f_2 + z_2 - z_{p0}}{f_2} x_{p2}, \tag{16}$$

$$y_{p0} = y_2 + \frac{f_2 + z_2 - z_{p0}}{f_2} y_{p2}. \tag{17}$$

Then, equating the first and the third, and the second and fourth, gives

$$\frac{f_1 - z_{p0}}{f_1} x_{p1} = x_2 + \frac{f_2 + z_2 - z_{p0}}{f_2} x_{p2}, \tag{18}$$

$$\frac{f_1 - z_{p0}}{f_1} y_{p1} = y_2 + \frac{f_2 + z_2 - z_{p0}}{f_2} y_{p2}. \tag{19}$$

That, simplifying, gives

$$z_{p0} = \frac{f_1 f_2}{f_1 x_{p2} - f_2 x_{p1}} \left(x_2 + x_{p2} - x_{p1} + \frac{z_2 x_{p2}}{f_2} \right). \tag{20}$$

$$z_{p0} = \frac{f_1 f_2}{f_1 y_{p2} - f_2 y_{p1}} \left(y_2 + y_{p2} - y_{p1} + \frac{z_2 y_{p2}}{f_2} \right). \tag{21}$$

In summary, note that z_{p0}, and the following x_{p0} and y_{p0} can be calculated given their image point coordinates (x_{p1}, y_{p1}) on the captured image, and coordinates (x_{p2}, y_{p2}) on the projected image system. Of course, f_1, f_2, x_2, y_2, z_2 have to be previously determined in the calibration process.

Note that z_{p0} can be obtained only knowing either the coordinate x_{p2} or the coordinate y_{p2}. This deduction is very important when the projection light is coded to obtain directly the coordinates of P_2 which corresponds to the imaged point P_1. The coded projection technique will be further discussed.

If the orientation of both image planes is not the same, then a rotation matrix has to be included with the described translation vector, so a new matrix T, which describes the rotation and translation of the coordinate system O_2 with respect to the global coordinate system O, has to be defined. Equation (22) shows the representation with respect to O of a given projected point P_2 expressed with respect to O_2.

$$\bar{P}_2 = TP_2. \tag{22}$$

4.2. *Calibration*

In the calibration process of the system, the focal distances, and the translation and orientation of the coordinate imaging system with respect to the global coordinate system, have to be determined. An easy calibration process can be designed by locating several object points, and their projections on the image plane. The calibration process is described in homogeneous coordinates, where translation, rotation, scaling and perspective can be expressed in a single matrix. Then the perspective transformation imaging process can be described as

$$\begin{bmatrix} wx_{p1} \\ wy_{p1} \\ w \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \end{bmatrix} \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} \\ 1 \end{bmatrix}. \tag{23}$$

Arranging the variables and operating, it is obtained that,

$$A_{11}x_{p0} - A_{31}x_{p1}x_{p0} + A_{12}y_{p0} - A_{32}x_{p1}y_{p0} + A_{13}z_{p0} - A_{33}x_{p1}z_{p0} + A_{14} - A_{34}x_{p1} = 0, \tag{24}$$

$$A_{21}x_{p0} - A_{31}y_{p1}x_{p0} + A_{22}y_{p0} - A_{32}y_{p1}y_{p0} + A_{23}z_{p0} - A_{33}y_{p1}z_{p0} + A_{24} - A_{34}y_{p1} = 0. \tag{25}$$

As it is assumed by several authors, such as Faugeras and Toscani,[42] Bortolozzi[43] and Brassart,[44] the solution of the system can be largely simplified if it is supposed that the variable A_{34} is known and equal to 1. Then, the last equations can be rearranged, obtaining,

$$QA = B, \tag{26}$$

where

$$A = [A_{11}A_{12}A_{13}A_{14}A_{21}A_{22}A_{23}A_{24}A_{31}A_{32}A_{33}]^t \tag{27}$$

$$q_x = \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -x_{p1}x_{p0} \\ -x_{p1}y_{p0} \\ -x_{p1}z_{p0} \end{bmatrix}^t, \tag{28}$$

$$q_y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_{p0} \\ y_{p0} \\ z_{p0} \\ 1 \\ -y_{p1}x_{p0} \\ -y_{p1}y_{p0} \\ -y_{p1}z_{p0} \end{bmatrix}^t, \tag{29}$$

$$b_x = [x_{p1}] \tag{30}$$

$$b_y = [y_{p1}]. \tag{31}$$

Then, given a correspondence between the object point (x_{p0}, y_{p0}, z_{p0}), and the image point (x_{p1}, y_{p1}), two variables of the matrix A can be solved. So, six non-coplanar object points and their associated projections are needed, at least, to solve the 12-variable equation system:

$$\begin{bmatrix} q_{x1} \\ q_{y1} \\ q_{x2} \\ q_{y2} \\ \cdots \\ q_{x6} \\ q_{y6} \end{bmatrix} A = \begin{bmatrix} b_{x1} \\ b_{y1} \\ b_{x2} \\ b_{y2} \\ \cdots \\ b_{x6} \\ b_{y6} \end{bmatrix}. \tag{32}$$

Finally, with the aim to determine the transformation matrix of the projected image, the process has to be repeated.

In robot vision modelling applications, camera parameter calibration is a crucial problem. We have explained the basic ideas to calibrate a camera as a pinhole model, without lenses distortion.

Theoretically, only six non-coplanar points are needed to obtain the 12 unknown components of the transformation matrix. In practice, to obtain an accurate system, more than six points are needed. Some authors have considered that at least 60 points are necessary. Although, the number of points used depends on the accuracy of the vision system wanted, and on the resolution of the calibration system used. In fact, the resolution used to obtain the 3D object point components, the methodology used to get the projected points, and the position of them on the image plane, are some of the important subjects which influence a lot on the results obtained.

4.3. 3D points measurement

Given the transformation matrix, which relates an object point with its projection on the captured image and on the projector image, respectively. Given the image plane coordinates of both projections of the same object point, the coordinates of this object point can be determined.

Summarising, the equations obtained in the calibration process are

$$\begin{bmatrix} w_1x_{p1} \\ w_1y_{p1} \\ w_1 \end{bmatrix} = \begin{bmatrix} A_{111} & A_{112} & A_{113} & A_{114} \\ A_{121} & A_{122} & A_{123} & A_{124} \\ A_{131} & A_{132} & A_{133} & A_{134} \end{bmatrix} \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} \\ 1 \end{bmatrix}, \tag{33}$$

$$\begin{bmatrix} w_2x_{p2} \\ w_2y_{p2} \\ w_2 \end{bmatrix} = \begin{bmatrix} A_{211} & A_{212} & A_{213} & A_{214} \\ A_{221} & A_{222} & A_{223} & A_{224} \\ A_{231} & A_{232} & A_{233} & A_{234} \end{bmatrix} \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} \\ 1 \end{bmatrix}. \tag{34}$$

Operating, and rearranging the variables,

$$(A_{111} - A_{131}x_{p1})x_{p0} + (A_{112} - A_{132}x_{p1})y_{p0} + (A_{113} - A_{133}x_{p1})z_{p0} = A_{134}x_{p1} - A_{114}, \tag{35}$$

$$(A_{121} - A_{131}y_{p1})x_{p0} + (A_{122} - A_{132}y_{p1})y_{p0} + (A_{123} - A_{133}y_{p1})z_{p0} = A_{134}y_{p1} - A_{124}, \tag{36}$$

$$(A_{211} - A_{231}x_{p2})x_{p0} + (A_{212} - A_{232}x_{p2})y_{p0} + (A_{213} - A_{233}x_{p2})z_{p0} = A_{234}x_{p2} - A_{214}, \tag{37}$$

$$(A_{221} - A_{231}y_{p2})x_{p0} + (A_{222} - A_{232}y_{p2})y_{p0} + (A_{223} - A_{233}y_{p2})z_{p0} = A_{234}y_{p2} - A_{224}. \tag{38}$$

So, arranging the image point coordinates and the object point coordinates as a matrix, the relation can be expressed as,

$$PV = F, \tag{39}$$

where, the matrices are,

$$P = \begin{bmatrix} A_{111} - A_{131}x_{p1} & A_{112} - A_{132}x_{p1} & A_{113} - A_{133}x_{p1} \\ A_{121} - A_{131}y_{p1} & A_{122} - A_{132}y_{p1} & A_{123} - A_{133}y_{p1} \\ A_{211} - A_{231}x_{p2} & A_{212} - A_{232}x_{p2} & A_{213} - A_{233}x_{p2} \\ A_{221} - A_{231}y_{p2} & A_{222} - A_{232}y_{p2} & A_{223} - A_{233}y_{p2} \end{bmatrix}, \tag{40}$$

$$V = \begin{bmatrix} x_{p0} \\ y_{p0} \\ z_{p0} \end{bmatrix}, \tag{41}$$

$$F = \begin{bmatrix} A_{134}x_{p1} - A_{114} \\ A_{134}y_{p1} - A_{124} \\ A_{234}x_{p2} - A_{214} \\ A_{234}y_{p2} - A_{224} \end{bmatrix}. \tag{42}$$

Finally, the object coordinates can be obtained from,

$$V = (P^{t}P)^{-1}P^{t}F. \tag{43}$$

As the object coordinates depend on the corrected association of the captured image point (x_{p1}, y_{p1}) and the projected point (x_{p2}, y_{p2}), any mistake in the correspondence establishment leads to an error in the object point coordinates determination.

5. CODIFICATION OF THE PROJECTED PATTERN

Several different patterns have been proposed to realise the correspondence between the image plane and the projected image with a certain accuracy. They can be classified as: single scanned dot, slit line, grid and dot matrix. However, in order to solve the correspondence problem, many of these patterns have quite a few problems. For any kind of object, and surface, there is also the problem due to the lost projected points that do not have a projection on the image plane. This issue can be due to a low surface reflection, to a surface occlusion, or simply because dots are reflected out of the camera scope. All these troubles can be solved if the projected pattern is conveniently coded, so that, the projected light carries information about the point (x_{p2}, y_{p2}) from which it has been emitted.

As it can be deduced from equations (35)–(38) we can use four equations to calculate the three variables that determine the object point (x_{p0}, y_{p0}, z_{p0}). In fact, one of the four equations is linearly dependent on the other ones, so, with the aim to determine the three variables, only three of the four equations have to be used. As an image of the scene has to be captured to deduce scene depth information, the image points, coordinates (x_{p1}, y_{p1}) are known, as is the projection of the object point with coordinates (x_{p0}, y_{p0}, z_{p0}). Then, only one of the equations (37) and (38) has to be used, because with only one, there is enough information to deduce the 3D object points' coordinates. Therefore, only one of the two coordinates (x_{p2}, y_{p2}) of the projected point has to be known. If the x coordinate (x_{p2}) is known, then the equation (37) can be used. Otherwise, (y_{p2}) is the known coordinate and equation (38) will be computed.

This idea allows the projected pattern to be codified just along one component coordinate, so the captured light, at the point (x_{p1}, y_{p1}) on the image plane, carries information from row x_{p2}, or column y_{p2}, from which it has been emitted.

In the following, some classifications on coded structured light patterns are presented. The survey done allowed us to classify the different techniques which code the projected pattern into three classifications. Looking only at the temporal dependence, the pattern can be classified into:

Static: The pattern is limited to static scenes with motionless objects. This is always due to the necessity to project a set of different patterns to obtain, in a coded way, the label for each column (or row) of the pattern. Any movement inside the scene between two pattern projection frames always produces a correspondence error.

Dynamic: The pattern is not limited to static scenes. Then, if the scene objects can move, the column or row of the projector image has to be coded with a single pattern projection.

Looking only at the light projected, the pattern can be classified into:

Binary: Any of the (x_{p2}, y_{p2}) points of the pattern can only have one of two possible values that are codified with 0 and 1, respectively. This binary value normally represents opacity and transparency, the absence or presence of the projected light on the object.

Grey level: Each pattern point can have an associated grey value, which represents the transparency (or opacity) level of the point against the projected light. As the information is coded as a grey light level, normally two steps are necessary in order to find out 3D information. Firstly, we have to obtain an image of the scene illuminated with the same light intensity for each point (without coding). Secondly, we must obtain the reference light needed to cancel the surface reflection effect, which depends on the kind of surfaces where the light is reflected. This limitation means that the pattern has to be also classified as a static pattern.

Colour: Each pattern's point has to be associated with a hue value. In order to use the colour constancy property, the hue values used have to be quite different from each other. The main goal is to get an efficient and accurate segmentation. As the system projects colour on the scene, its use is limited to a neutral colour scene, as highly saturated colour objects can produce lost pattern regions in the segmentation step and posterior decodification.

Furthermore, another classification can be proposed, which classifies the pattern from its surface

depth discontinuities' dependence. Then, looking at this dependence, it can be classified as

Periodical: The codification is periodically repeated along the pattern. Normally, this technique is used to reduce the number of bits that codify the pattern, but produce a limitation of the depth discontinuity which can not be larger than the half of the period length.

Absolute: Each column or row of the projected pattern has a unique codification. So, it does not have any depth discontinuity dependence.

There follows a survey of the coded patterns presented in recent years, analysing for each one its advantages and disadvantages.

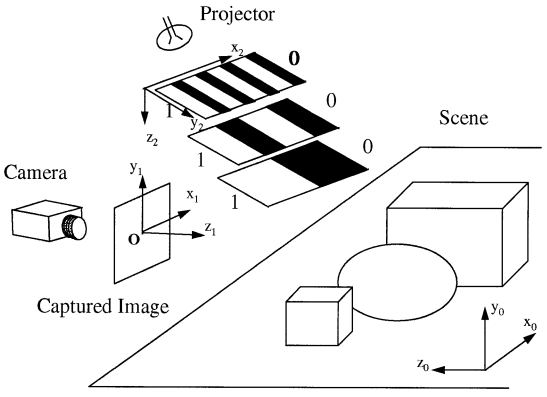

Fig. 3. The temporary codification proposed by Posdamer *et al.*[45] and Altschuler *et al.*[46, 47]

6. SURVEY: CODED STRUCTURED LIGHT TECHNIQUES

6.1. Posdamer–Altschuler

The system proposed by Posdamer and Altschuler[45] and Altschuler, Altschuler and Taboada,[46,47] shown in Fig. 3, is based on the utilisation of a pattern structured as a dot matrix of $n \times n$ binary light beams. Each n_i column of the pattern can be independently controlled, so it can be lighted or obscured. Then, several masks can be made to allow coding any pattern dot, in a temporal way, as a sequential projection of different patterns, as shown in Fig. 4. The number of patterns to be projected is determined by the number of columns to be coded.

Blocking light beams, a pattern image sequence can be projected which projects a sequence of 0 and 1 for each beam. Then, each position of the measuring surface is temporally and sequentially illuminated for the different values of the same dot beam, which produces a code that can distinguish one dot from its neighbours.

In fact the system works as follows. Firstly, an all dot illuminated pattern is projected on the scene. The camera images all the dots reflected from the surfaces of the scene, and keeps their positions in a datum structure. In the following steps, each coded pattern is projected and the system can ask for the light information in the kept positions. Each detected point, in the first step, on the measuring surface, at the last step, when all the patterns have been projected, has a code which allows us to determine from which beam column it has been projected.

The system proposed by Posdamer *et al.* and Altschuler *et al.* is limited to static scenes as it has to capture an image from each projected pattern. But, as proposed by Altschuler *et al.*,[46] the system can be improved to be used in dynamic scenes. In this case, all the patterns are projected with a different frequency wave and n cameras are used, one for each projected pattern. Each camera should have an optical filter, as each camera will image only one pattern frequency. Then, all the patterns can be projected at the same time. The cameras should be located as close as possible. Even so, we could also obtain some points

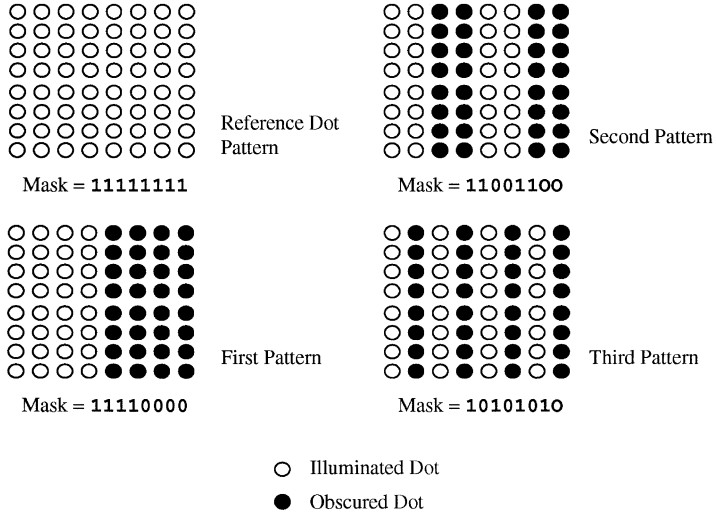

Fig. 4. Coding an $n \times n$ matrix of laser beam dots.

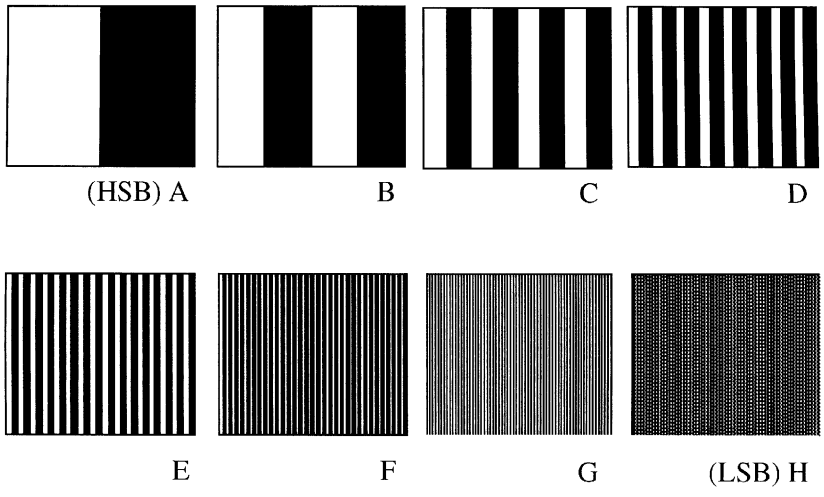

Fig. 5. 8 bits temporally binary-coded pattern projection.

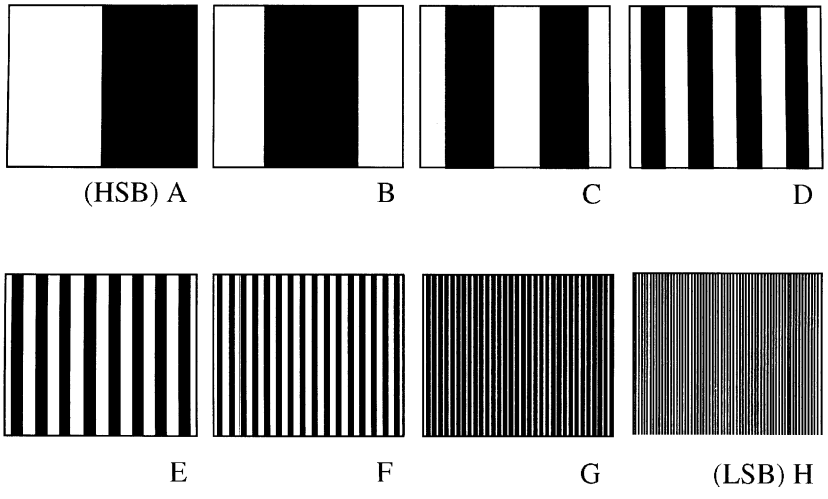

Fig. 6. 8 bits temporally Gray-coded pattern projection.

that cannot be decoded as they are not imaged from all the cameras.

Posdamer and Altschuler's idea has been widely studied. Basically, the codification and the speed measurement have been improved. For example, Inokuchi, Sato and Matsuda[48] in 1984 proposed changing the binary codification (Fig. 5) to a Gray codification (Fig. 6) which is more error robust.

Later, in 1986, Sato, Yamamoto and Inokuchi,[49] and in 1987, Sato and Inokuchi,[50] proposed to use a Liquid Crystal Device, already studied by Inokuchi, Morita and Sakurai[51] in 1972, which allows an increased number of columns to be projected with a high accuracy. The system also improves the coded speed, against a slide projector, so the LCD can be electronically controlled.

If an object has a high textural contrast or any high reflected surface regions, then, some pattern segmentation errors can be produced. Normally, this problem can be solved capturing the first image without pattern projection (or with the same light projection for each dot) to obtain a reference light intensity. This intensity gives a dynamic threshold for each pixel.

Sato *et al.*,[49] in 1986, and Sato and Inokuchi,[50] in 1987, proposed a complementary method of segmentation based on the projection for each coded pattern, of its positive and negative representation. Then, comparing the intensity image with both representations, a better segmentation of the projected pattern can be obtained.

The problem of a light projector is sometimes a result of heat irradiation onto the scene, and of its big size and weight. In 1995, Hattori and Sato[52] proposed to replace the light projector with a semiconductor laser, which gives a high power illumination with low heat irradiation. The proposed system, named Cubiscope, is shown in Fig. 7 and is composed

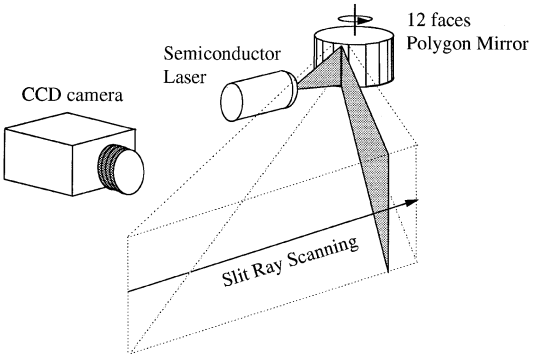

Fig. 7. The Cubiscope system.

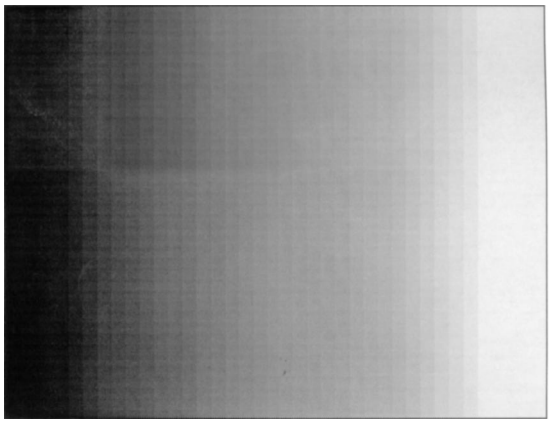

Fig. 8. The pattern proposed by Carrihill *et al.*[54]

by a CCD camera, a semiconductor laser and a scanned mirror, synchronously controlled. The main goal of this system is to obtain scene images based on the temporary codification proposed by Posdamer *et al.* and Altschuler *et al.*

In 1995, the same temporary codification described by Posdamer *et al.* and Altschuler *et al.* in 1982 is again proposed by Müller.[53]

6.2. Carrihill–Hummel

The system proposed by Carrihill and Hummel,[54] in 1985, is based on getting 3D information of any scene from two frame images. The first image is captured with a constant illumination projection on the scene. A linear wedge filter is used to project on the scene a pattern brightly illuminated on one side, and half brightly on the other. This pattern is shown in Fig. 8. This wedge filter is used to illuminate the scene for the second image capturing step.

For the first projection, it is supposed that the illumination intensity is constant along the x-axis and the y-axis of the projector. For the second projection, illumination has to be a constant along the y-axis. Then, for each obtained pixel, the intensity subtraction of the two images, can be calculated. This difference allows obtaining the column of the projector from which the dot has been emitted.

Using a difference ratio for each pixel makes easier to cancel possible surface reflections or highlights as an illumination ratio change is always produced by a change in the projected illumination.

Using electro-optics filters, the system can provide 3D information with a high speed, but, like systems which project more than one image, it is limited to static scenes. Obviously, any moving object in two consecutive frames would produce a correspondence error.

Carrihill *et al.* affirm that the system obtains an accurate measurement with pixels of 8 bits depth. The test has been limited to 12 bits depth, although Carrihill *et al.* affirm that there is a good agreement between the sensor perception and reality.

In order to improve the accuracy measurement without increasing the bit depth of the captured image, a new method based on a sawtooth pattern could be proposed. Two sawtooth patterns of different period length are shown in Fig. 9. Obviously, the maximum surface depth discontinuity to be measured is limited to the chosen period.

6.3. Boyer–Kak

In 1987, Boyer and Kak[55] proposed to illuminate the scene using a single pattern projection. The pattern is coded by vertical coloured slits. The correspondence problem, i.e. given an imaged slit, to obtain from which projector slit it has been emitted, can be solved as the slit codification used to make the pattern, is known. Obviously, as with any other system which projects colour, its use is limited to predominantly colour neutral surfaces, as highly saturated hues could produce slit identification errors.

Figure 10 shows only one slit pattern projection up to which the matching problem can be complicated between projected and imaged slits. Note that, slits projected out of the scope of the camera or reflected on occluded surfaces cannot be imaged. Also, the slits imaged by the camera can be obtained in a different order.

Boyer *et al.* proposed to codify the pattern using the three basic colour components: red, green and blue. The pattern will be made by a sequence of vertical slits coloured with any of the three basic components. The slit can also be white. Optionally, between each slit, a black gap can be placed, which can be used as a slit separator. Note that, white can be easily obtained and identified as the colour which has approximately the same red, green and blue intensity values.

If it is supposed that the pattern is made by n vertical slits, Boyer *et al.* proposed to divide this pattern into m subpatterns, each one made by k vertical slits, as $n = m*k$. In order to code the m patterns in a unique way, k is a value as large as necessary. As each subpattern is emitted without any beginning and ending code, multiple matchings can be obtained in reception. Actually, we do not know where a

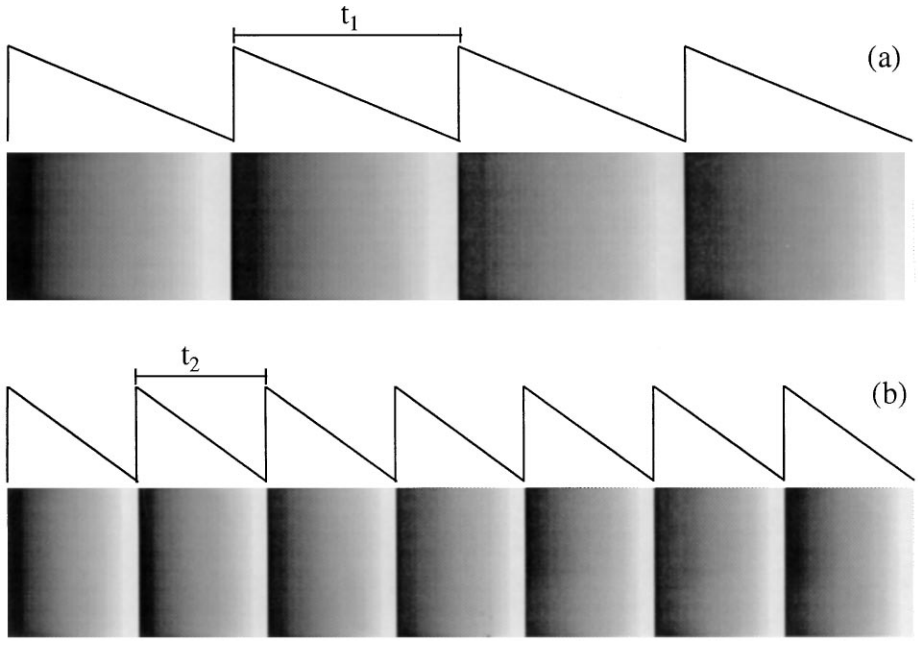

Fig. 9. The use of the period length to improve the resolution of the measurement without increasing the number of bits depth of the captured image. Obviously, period patterns are limited to measuring only surfaces with a depth discontinuity smaller than its period length.

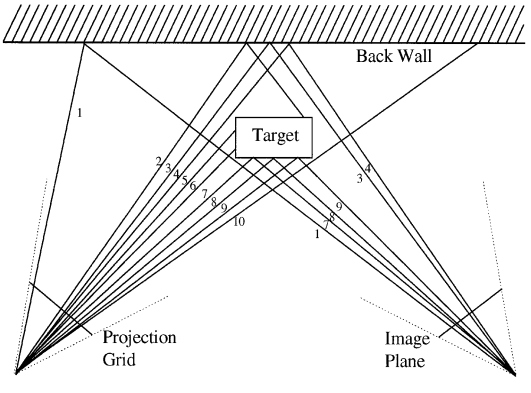

Fig. 10. An example of the correspondence problem that has to be solved between projected and imaged slits.

subpattern ends and the next begins, and moreover, the slits have been reordered as a function of the measured surface discontinuities. An example of the pattern deformation can be shown in Fig. 11, where a 72 slit pattern was emitted and a 50 slit one has been received.

For colour slit detection, Boyer *et al.* proposed hardware which detects red, green and blue peaks in the RGB signal given by the camera. So the specific hardware can obtain the index and position of each peak in real-time. In fact, the index can have a proper value for each different colour (e.g. the 1, 2 and 3 index values can be associated with the red, green and blue colours, respectively). So the position can be obtained from the camera synchronisation. Obviously, CAG feature allows the equalisation of the RGB signals.

In order to obtain the subpatterns from the index and position provided by the hardware, a software algorithm is used. In the first step all the possible matchings between emitted and received subpatterns are found. In the second one, a region growing algorithm is used to remove all the erroneous matchings. In the third one, for each subpattern not matched, a heuristic algorithm is used to find its better matching. When the third step is completed, a datum structure, which associates captured and projected subpatterns has been obtained. Finally, a triangulation process provides the 3D scene information we were looking for.

As we said before, the pattern proposed by Boyer *et al.*, is limited to measuring predominantly neutral colour surfaces. Even so, it is specially recommended in dynamic environments as only one projection is needed.

6.4. *Le Moigne–Waxman*

There are a lot of different patterns that can be made based on the grid pattern concept. Some examples can be shown in Fig. 12. When a grid pattern is chosen, the number of crossing points to be projected has to be chosen, but the line thickness has to be also chosen, as it depends directly on the smoothness of the imaged surface texture. A very thick line will give a low resolution, and, with a very thin one, we will obtain a lot of discontinuities, which will complicate the matching process. Obviously, the thickness of the line has to be chosen knowing the kind of scenes to be measured.

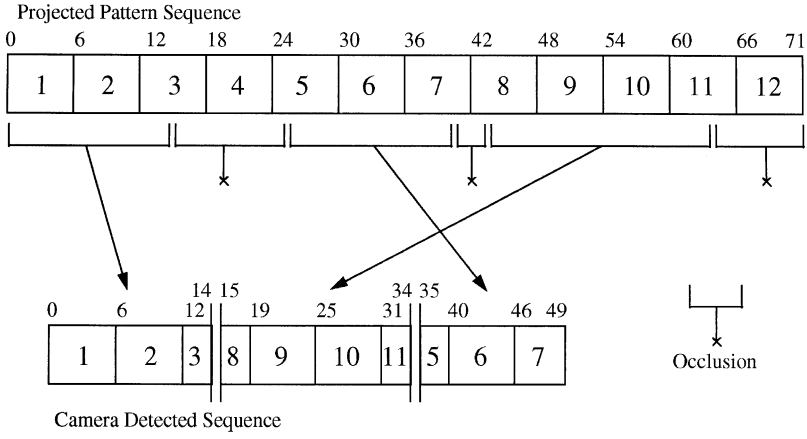

Fig. 11. A possible relation between a projected pattern and the captured one. A surface like the one shown in Fig. 10 has been measured.

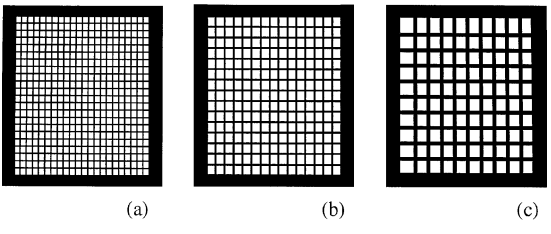

Fig. 12. Some not-coded grid patterns. The pattern resolution is given by the number of crossing points and its line thickness.

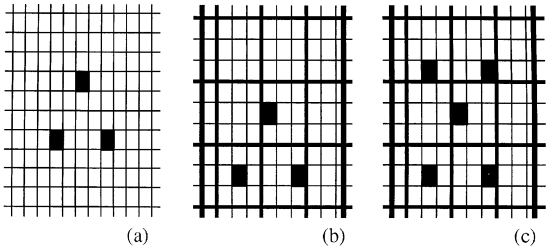

Fig. 13. Grid patterns partially coded by the position of some dots, which are used as landmarks to initiate the labelling process.

Note that the patterns shown in Fig. 12 are not coded. Le Moigne and Waxman[56–58] proposed to add dots on the grid which can be used as landmarks to initiate the decodification or labelling of the projected pattern. In Fig. 13 some grid patterns coded with dots are shown.

The configuration used for the system locates the projector at a fixed distance from the camera, along the y-axis. This means that vertical lines are imaged nearly without deformation, keeping its natural parallelism and continuity, so they will be easily detectable. A square neighbourhood operator is computed in each imaged pixel to reduce the albedo or highlight effect. The system uses the known location of the vertical lines as guides to search the dot and horizontal line intersections. Then the horizontal lines are explored to join their discontinuities. Note that, if lot of discontinuities are found, it is due to the line thickness chosen being too thin. Finally, for each landmark dot, two different edge labelling processes, using a greedy algorithm, are used. One process starts its labelling from the right side, and the other from the left side of the landmark. So, six different labellings are obtained from three landmark dot patterns. Then, a merging algorithm is used to label the whole pattern. Initially only two-labellings edges are considered. Then, they try to label the not labelled edges with a local interpolation.

Someone could think that the labelling algorithm is rather complicated. However, the horizontal lines can be highly broken due to depth discontinuities of the measuring surfaces, and they can also disappear partially. So, easily matching errors could be obtained using a simpler labelling process.

The pattern proposed by Le Moigne *et al.* is specially indicated to work in dynamic environments, as a single projection is needed to obtain 3D information of the scene. But it is true that the resolution given by the pattern is rather limited, basically for two reasons: firstly, vertical lines do not give any depth information as they are only used as searching guides. And finally, the matching process may be rather slow in density grids, so, dynamic high resolution is not permitted. But its utilisation as a vision sensor for mobile robot navigation in structured indoor environments is highly recommended. In such environments, the surfaces do not usually have high contrast textures, which can deform the pattern, and the scene is quite regular, and small differences do not give interesting information.

6.5. *Morita–Yakima–Sakata*

In 1988 Morita *et al.*[59] proposed a binary pattern of light dots as an M-array. The M-array has the already defined *coding window* property as any

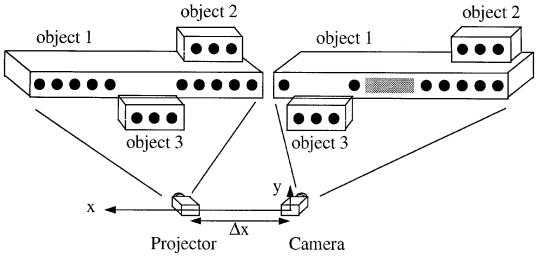

Fig. 14. The system configuration proposed by Morita *et al.*[59] and the dot disorders that can be obtained.

subpattern into a window exists only once and at one place within a period (for more information see the Vuylsteke–Oosterlinck method).

To simplify the correspondence problem, Morita *et al.* placed the camera and projector such that the Δ_y and Δ_z distances between them are zero, as shown in Fig. 14. Then, the dot identification process is quite simplified, as the dot position on the image plane can only move horizontally.

In fact, an M-array pattern is made by a set of circular bright, or dark, dots, represented by the binary number 0, or 1, respectively. From the M-array property, the indices of all the elements of the observed pattern can be uniquely determined relative to the projected pattern if no pattern disorders exists. However, for complex surfaces the following disorders can be obtained:

1. Missing dots: If two or more objects are measured, a dot projection can be reflected on an occluded surface, and, it cannot reach the image plane. An example can be seen in Fig. 14 where the points from #2 to #4 are projected on the back surface of the object #3.
2. Dot displacement: If the depth distance between two objects is different, the reflected dots will be imaged with a horizontal displacement between them. As can be seen in Fig. 14, the dots reflected on object 2 move right by one dot on the image plane.
3. Dot permutation: If an object is placed in front of another, the projected dots can swap positions. As shown in Fig. 14, the projected dots from #6 to #8 reach the image plane in the position of #2 to #4, as they have been occluded by object 3.

In order to obtain the image coordinates of each dot, this method has to project a whole illuminated dot matrix. Then, a binary (bright and dark) dot matrix is projected, so that each window, with a fixed size, determines the column index from which the imaged point has been projected.

The proposed system is quite simple, so that 3D information can be obtained without a lot of computing time. However, the system is limited to static scenes because they need two projections. As the system projects isolated light dots, not a lot of textural information can be obtained, and, if the dot size is reduced to get more resolution, then, high contrast textural surfaces could modify the dot shape, making its localisation difficult.

6.6. *Vuylsteke–Oosterlinck*

A model with a dynamic pattern, column coded, is proposed.[60] The basic structure of the pattern is like a regular chess-board with alternating bright and dark squares. Then, the pattern is modulated overlapping a bright or dark spot at every square vertex (Fig. 15), so that, any square of the regular chess-board pattern carries an additional information bit, which together with the neighbouring bits, will be used to code each column.

In order to identify each dot of the coded pattern, they are represented in binary. The dot gets the "1" value if it comes from a bright dot. Otherwise, it takes the "0" value, associated with dark. Additionally, the chess-board allows finding out naturally if the coded dot comes from a bright or dark square, represented by the symbols $-$ and $+$, respectively. Then, four different combinations, shown in Fig. 16, can be obtained.

Following the chess-board, the code assignment is based on two binary sequences, represented by c_k and b_k of 63 bits length. All the dots of the k column which corresponds to a $+$ square type are assigned the binary value c_k, while those of the $-$ type are assigned the value b_k. A 6 bits length code is needed to codify the 63 different columns of the pattern. Then any window of 2×3 squares allow coding any column index. In fact, any window with 6 squares size can be used, but obviously, a compact window is less affected by surface discontinuities than an elongated one. This is the main reason why a 2×3 window has been chosen instead of a 1×6.

Each 2×3 window has a triplet of c_k codes and another of b_k, easily differentiable. The sorting of these codes gives a 6 bits length number which codifies a column index of the pattern. An example for the

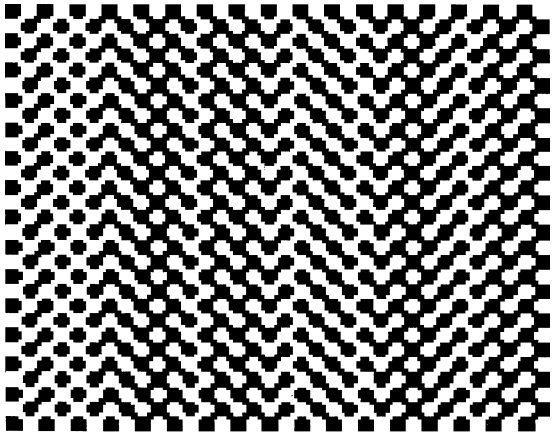

Fig. 15. Details of the proposed pattern.

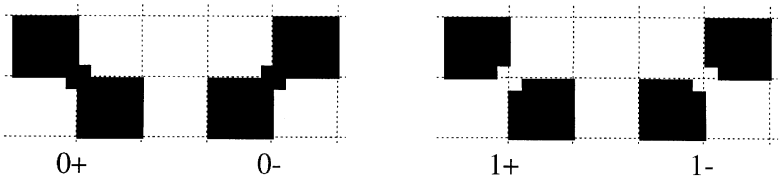

Fig. 16. The pattern is made by four different combinations. The top-left square is represented by the symbols $-$ and $+$, which represent a bright or dark square, respectively. The binary numbers 0 and 1 are used to code the dot at the square vertex, which can also be dark or bright, respectively.

$c_0\ b_1\ c_2\ b_3\ c_4\ b_5\ c_6\ b_7\ c_8\ b_9\ c_{10}\ b_{11}\ c_{12}\ b_{13}\ c_{14}\ c_{15}$
$b_0\ c_1\ b_2\ c_3\ b_4\ c_5\ b_6\ c_7\ b_8\ c_9\ b_{10}\ c_{11}\ b_{12}\ c_{13}\ b_{14}\ b_{15}$
$c_0\ b_1\ c_2\ b_3\ c_4\ b_5\ c_6\ b_7\ c_8\ b_9\ c_{10}\ b_{11}\ c_{12}\ b_{13}\ c_{14}\ c_{15}$
$b_0\ c_1\ b_2\ c_3\ b_4\ c_5\ b_6\ c_7\ b_8\ c_9\ b_{10}\ c_{11}\ b_{12}\ c_{13}\ b_{14}\ b_{15}$

Fig. 17. The code assignment determines that any 2×3 window can identify the column index. Each window is made by a triplet of c_i values and another triplet of b_i. In the example, for the two windows, the column is given by the same sequence order $((c_3, c_4, c_5), (b_3, b_4, b_5))$.

column $k = 3$ is shown in Fig. 17, where the code number which represents column 3 is given by the sequence: $((c_3, c_4, c_5), (b_3, b_4, b_5))$.

The code assignment cannot be chosen independently due the fact that the identification windows overlap. Then a kind of recursive code generation sequence, with an initial birth polynomy, has been used, which allows the obtaining of a desired length period repetition to code in a unique way all the columns of the pattern.

The pattern presented has basically two limitations. The first one is based on the difficulty to measure high textural surfaces, which produce partial lost regions of the pattern. And the second one is associated with the surface orientation. If these are not perpendicular to the optical axis of the projector, a deformation of the projected pattern is presented, which from a determined angle orientation, does not allow the identification of the pattern. However, the pattern is well recommended to measure dynamic surfaces where only one projection is permitted.

In 1995, Pajdla[61] has reimplemented the same pattern, explaining the calibration process. Pajdla has proposed, as an improvement, to use a hexagonal codification instead of a square one. In this case, the window's size is reduced and as a result, the number of not indexed columns due to depth discontinuities is decreased. However, the identification step is more complicated.

6.7. *Tajima–Iwakawa*

The system proposed by Tajima and Iwakawa[62] in 1990 is based on the vertical slit coding technique, where each slit is emitted with a different wavelength. The projected pattern is like a rainbow pattern, as the whole colour spectrum, from the violet up to the red,

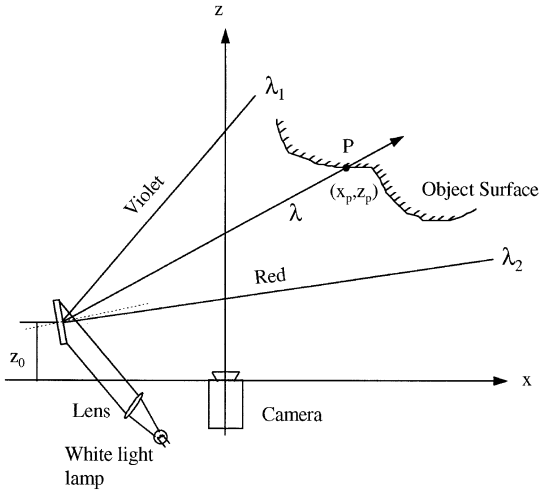

Fig. 18. The rainbow range pattern projected by Tajima *et al.*[62] made by a set of vertical slits which use the entire colour spectrum from violet to red.

is projected. This can be obtained diffracting white light as shown in Fig. 18.

Depth can be obtained using the triangulation principle, but first the slit angle (the x_{p2} coordinate) which has produced the colour pixel imaged on the image plane, has to be known. In fact, the slit emission angle can also be determined knowing the slit wavelength.

The objects illuminated by the rainbow pattern are imaged with a single monochromatic camera instead of a colour camera. Two different colour filters placed in front of the camera are used, and two images of the scene are captured. For the same point, the intensity relation between the two images does not depend on the illumination, nor on the colour object. Tajima and Iwakawa have proved that this intensity relation depends directly on the slit wavelength (λ).

As the system needs two frames for each measurement, it is limited to static scenes, but obviously the system can measure dynamic scenes using a colour camera.

In 1996, Geng,[63] seemingly without knowing the work already done by Tajima and Iwakawa proposed the same rainbow pattern in order to obtain 3D information from the measuring surface. However, the initial idea proposed by Tajima and Iwakawa have been improved by Geng using a CCD colour camera and using a Linear Variable Wavelength

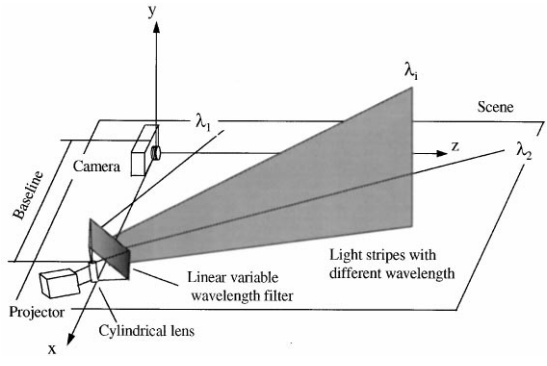

Fig. 19. The system proposed by Geng.[63]

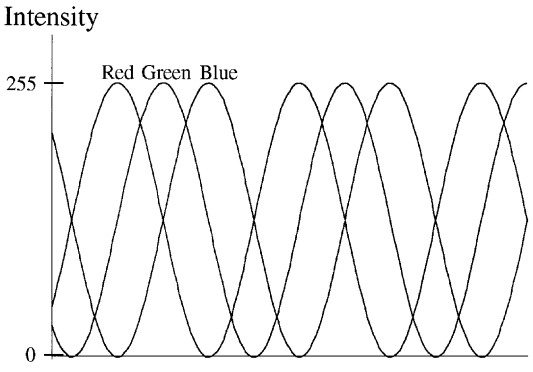

Fig. 20. The periodical pattern proposed by Wust and Capson,[65] made by the sinusoidal overlapping of the three primary colour components.

Filter (LVWF). Tajima's improved idea proposed by Geng can be seen in Fig. 19.

In this case, the projector emits a white light plane, which, using a cylindrical lens, generates a continuous fan of planes. From the natural design of the LVWF, a continuous colour spectrum is obtained, so no two light planes are projected on the scene with the same wavelength. Obviously, the system's resolution depends on the camera's features to distinguish among different colours, or different wavelengths.

Recently, Smutny and Pajdla[64] have reimplemented the system proposed by Tajima and Iwakawa. Regarding limitations of the rainbow pattern, they said that the surface can have any colour, but this has to be opaque. Obviously, no other wavelength which does not come from the projector, can be emitted on the scene. The scene has to be light controlled and the measuring objects cannot be fluorescent, nor phosphorescent.

6.8. Wust–Capson

In 1991, Wust and Capson[65] proposed the projection of a sinusoidal intensity pattern on the measuring surfaces. The proposed pattern is made by the overlapping of three sinusoids, of n periods along the x-axis. Each sinusoid is associated with each primary colour (red, green and blue). The green's sinusoid is shifted 90° with respect to the red, and the blue is shifted 90° with respect to the green, as shown in Fig. 20. The pattern is column coded, so all the rows are identical, resulting in a coloured vertical fringe pattern.

Instead of obtaining the column index by decoding the imaged pattern, Wust and Capson proposed to obtain the depth directly from the wave phase shifting. This technique is widely used in Moire methods to measure continuous surfaces.[66]

The pixel intensity obtained by the camera, as a function of its pixel position on the image plane, is given by the following equation:

$$I(x,y) = E(a(x,y) + b(x,y)\cos(\phi_r + \phi_z + \delta)), \quad (44)$$

where E is the power of the projected light, $a(x,y)$ the surface irradiance, $0 \le a(x,y) \le 1$, $b(x,y)$ the fringe visibility, $0 \le b(x,y) \le 1$, ϕ_r the carrier $(2\pi/p\, x\cos\theta)$, ϕ_z the displacement $(2\pi/p\, z\, \sin\theta)$, δ the introduced phase shift, p the period of projected fringes, θ the angle of projection, x the horizontal image coordinate, y the vertical image coordinate and z the depth (distance from the camera to the surface).

Then, to solve equation (44), Wust *et al.* proposed to obtain the intensity value $I(x,y)$ for different δ values, at the same angular distances. So, choosing the δ values for the three images I_1, I_2 and I_3 as $\pi/4$, $3\pi/4$ and $5\pi/4$, respectively, the following equation can be obtained:

$$\phi_z = \arctan((I_2 - I_3)/(I_1 - I_2)) - \phi_r. \quad (45)$$

So, if I_1, I_2 and I_3, are associated with the three primary colour components (red, green and blue),

$$I_r(x,y) = E_r(a(x,y) + b(x,y)\cos(\phi_r + \phi_z + \pi/4)), \quad (46)$$

$$I_g(x,y) = E_g(a(x,y) + b(x,y)\cos(\phi_r + \phi_z + 3\pi/4)), \quad (47)$$

$$I_b(x,y) = E_b(a(x,y) + b(x,y)\cos(\phi_r + \phi_z + 5\pi/4)). \quad (48)$$

If we are able to ensure, by calibration, that the system has a uniform spectral response, so, the intensity of the red, green and blue signals are comparable, then, it can be affirmed that

$$\phi_r + \phi_z = \arctan((I_r - I_g)/(I_g - I_b)). \quad (49)$$

Some limitations can be observed in the method. The scene has to be predominantly colour neutral, in spite of colour projection. As the pattern is made by periodical fringes, it is limited to measure surfaces without discontinuities larger than a fringe period. According to Wust and Capson, camera response depends on the frequency of the emitted fringes, and the histogram equalisation used to compensate for the nonlinear intensity response also produces some measuring errors. Improving these aspects the system can obtain a better resolution and robustness.

In 1993, Hung[67] proposed a grey level sinusoidal pattern. The period of the captured pattern depends on the depth of the surface where it is reflected. However, computation demands much time. Hung proposed, as Wust and Capson had also proposed, to triangulate from the column phase of the imaged point. For each pixel point, this phase can be approximately obtained from the light intensity. Obviously, this method suffers the same limitation as the method proposed by Wust *et al.*

6.9. *Griffin–Narasimhan–Yee*

Griffin *et al.*[68], in 1992, have carried out a mathematical study about which should be the largest size allowed for a coded matrix dot pattern. It is supposed that (1) A dot position is coded with information emitted by itself and the information of its four neighbours (North, South, East and West). (2) There cannot be two different dot positions with the same code. (3) The information is determined using a fixed basis, which determines the symbols used to code the matrix. (4) The biggest matrix is desired, i.e. the matrix which gives a better resolution.

If a basis equal to 3 is done, a possible dot codification is shown in Fig. 21.

Griffin *et al.* have proved that, given a basis b, the largest matrix (the biggest $n \times m$ matrix) can be obtained from its largest horizontal vector (Vhm), and its largest vertical vector (Vvm). Vhm is a vector made by the sequence of all the triplets of numbers that can be made without repetition using a b basis. Vvm is a vector made by the sequence of all the pairs of numbers that can be made without repetition. Then, the first row of the matrix is given directly by Vhm,

$$f_{0i} = \mathrm{Vhm}_i \tag{50}$$

and, the other matrix elements can be determined applying equation (51) where "i" is the row index and varies from 0 to the Vhm length, and "j" the column index and varies from 0 to the Vvm length.

$$f_{ij} = 1 + ((f_{i-1\,j} + \mathrm{Vvm}_j) \bmod b). \tag{51}$$

For example, if a basis equal to 3 is supposed, then its largest vectors are

$$\mathrm{Vhm} = (331321311231221211133232223333),$$

$$\mathrm{Vvm} = (3121132233).$$

1	3	2	1	1	2
2	2	2	3	1	3
1	3	3	1	1	2
1	1	2	3	2	2
3	2	2	1	3	1
2	1	1	3	2	3

Coded Order:
$(w_{ij}, w_{ij-1}\ w_{i-1j}\ w_{ij+1}\ w_{i+1j})$

Coded Word = 33212

Fig. 21. Dot codification example using its four neighbours and a basis equal to 3, i.e. only three different symbols can be used.

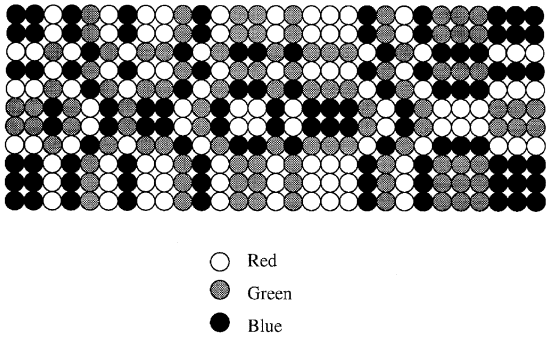

Fig. 22. A possible coded dot matrix, obtained by the method proposed by Griffin *et al.*[68] A basis equal to 3 has been supposed and to each symbol a coloured dot has been associated.

So, the obtained matrix is

331321311231221211133232223333
331321311231221211133232223333
112132122312332322213213333111
331321311231221211132132223333
112132122312332322213213333111
223213233123113133321321112222
223213233123113133321321112222
112132122312332322213213333111
331321311231221211133232223333
331321311231221211133232223333
331321311231221211133232223333

After the coded matrix is found out, a different projection can be associated for each value, i.e. for each number which belongs to the interval $\{1, b\}$. For example, a coloured dot pattern can be obtained if the red, green and blue colours are associated to the numbers 1, 2 and 3, respectively, obtaining a pattern like the one shown in Fig. 22.

The resolution of the pattern can be increased by simply increasing the basis value. Depending on the colour discriminating capability of the system employed to image the scene, almost any degree of resolution can be obtained.

In many applications, the scene is not made by colour neutral surfaces. A simple reason could be that the imaging system used is only able to capture monochromatic images. Then, monochromatic light has to be projected on the scene. In this case, the coloured association projected of each number can be changed for a geometric association. An example is shown in Fig. 23.

The method proposed by Griffin *et al.* is the unique method studied from the decodification of the pattern captured by the camera. For each image point (x_{p1}, y_{p1}), the projector position point (x_{p2}, y_{p2}) from which it has been emitted can be known. As shown in the mathematical section dedicated to surface measuring, it is not necessary to know both projector coordinates. Then, the pattern can be obviously simplified to obtain a single row coded or column coded pattern,

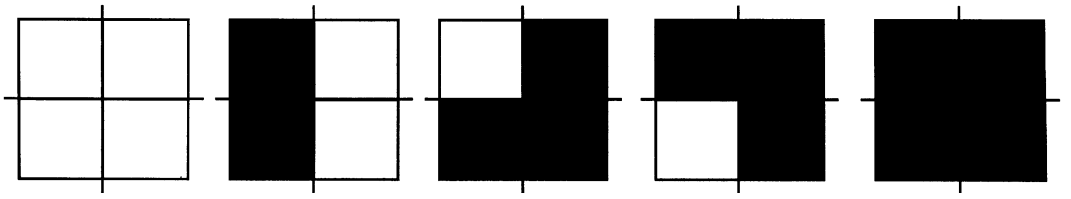

Fig. 23. A possible geometric association of a matrix dot coded using a basis equal to 5.

but we could possibly obtain a pattern which uses the windows proposed by Vuylsteke *et al.*,[60] or the coloured slits proposed by Boyer and Kak.[55]

6.10. Maruyama–Abe

The method described by Maruyama and Abe,[69] in 1993, is based on the projection of multiple vertical slits. Slits are coded from the position of some random cuts, as shown in Fig. 24.

All the random cuts are placed to obtain short lines, satisfying the condition that its length is in the interval $[L_0 - \Delta, L_0 + \Delta]$, where L_0 is a standard length and Δ is given by a random number.

Segment matching is performed based only on the correspondence of the short line end points along the epipolar lines. As the cuts are randomly distributed, more than one pattern line can be matched with a given image segment. An example of multiple matching is shown in Fig. 25. Then, the imaged segments are classified into three groups: (1) Segments with no competitors, as they have only one matching. (2) Segments that can be simply identified using adjacency relations between the segments and their neighbours. In this case, the matching cannot be done due to the noise on the image or due to a surface depth discontinuity, which has changed the line length. (3) Otherwise.

Maruyama *et al.* consider that the epipolar lines can always be horizontally computed, so that only the x-axis has to be examined to find out all the lines on the pattern which match with the imaged segment. This is not always true, as epipolar lines depend on the projector position and orientation, with respect to the camera axis, and Maruyama *et al.* do not impose any restriction. But, it is true that the image can always be transformed to get horizontal epipolar lines, as they said in their work.

In order to remove erroneous matching, the information of neighbour segments is taken up. Finally, with the aim to obtain the correspondence of all the segments from groups 2 and 3, a region growing algorithm matching from adjacency constraints is used.

To improve the system, we could think of using a pattern with an intelligent cut distribution which matches the same imaged segment, along its epipolar lines (in order to not obtain more than one matching). So, a bijection between captured image and projected pattern is determined. But, as Maruyama *et al.* said, if the adjacency constraints are used, the matching can be done using a region growing algorithm. However,

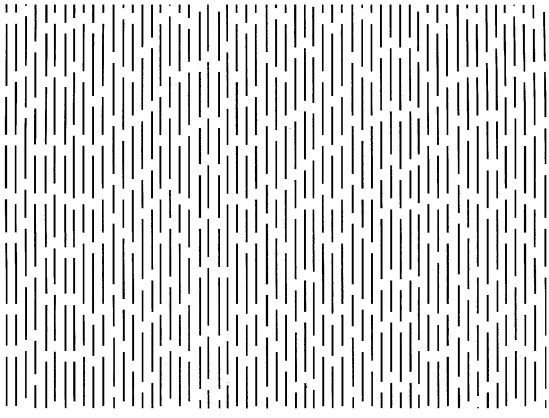

Fig. 24. The pattern proposed by Maruyama *et al.*[69] made by the projection of multiple slits with random cuts.

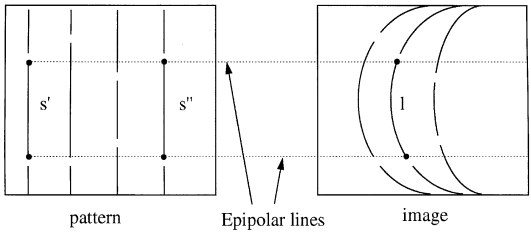

Fig. 25. A multiple matching example. In this case the end points of two short lines, s′ and s″ of the pattern, match the end points of a line projected on the image plane, along its epipolar lines.

surface depth discontinuity or noise problems in the image could provide segments which do not have a matching line. In this case, a region growing algorithm has to be used to match them.

The method is suitable for measuring 3D objects with relatively smooth surfaces. For objects with a lot of discontinuities and with highly textured surfaces, the method will have some difficulties. Obviously, line discontinuities produced in the captured segments complicate the matching process. However, the pattern is perfectly used to measure dynamic scenes, as only one projection is required.

7. AN OVERVIEW OF THE TECHNIQUES

In order to provide an easier overview of the techniques presented, we summarise them in the following

chronological list, relating also the authors who used or improved them lately.

1982 Posdamer–Altschuler. A temporal space-encoded projected beam system.[45–47]

Lately used: Inokuchi *et al.*[48] in 1984, Sato *et al.*[49, 50] in 1986 and 1987, Hattori *et al.*[52] and Muller[53] in 1995.

1985 Carrihill–Hummel. An intensity ratio pattern projection.[54]

1987 Boyer–Kak. A colour encoded stripped pattern.[55]

1988 Le Moigne–Waxman. A grid pattern with some landmark dots.[56–58]

1988 Morita–Yakima–Sakata. An M-array pattern projection system.[59]

1990 Vuylsteke–Oosterlinck. A single binary-encoded chess board pattern.[60]

Lately used: Pajdla[61] in 1995.

1990 Tajima–Iwakawa. A rainbow pattern[62].

Lately used: Geng[63] and Smutny *et al.*[64] in 1996.

1991 Wust–Capson. A colour sinusoidal fringe pattern[65].

Lately used: Hung[67] in 1993.

1992 Griffin–Narasimhan–Yee. A mathematical study about encoded patterns.[68]

1993 Maruyama–Abe. A multiple slits with random cuts pattern projection.[69]

8. CONCLUSIONS

The accuracy of surface measurement depends highly on the vision system used. It is widely known that a vision system based on stereoscopy is useful to measure surfaces with well-defined boundary edges. An algorithm to recognise singular points can be used to solve the problem of correspondence between points on both image planes. But the stereoscopy system becomes rather inefficient to measure continuous surfaces, where there are not many reference points. It has also several problems in textural surfaces or in surfaces with lots of discontinuities. In these environments the abundance of reference points can produce matching mistakes. In these cases, an active system based on a structured light concept will be useful.

In structured light, a correspondence between the captured pattern, imaged by the camera, and the pattern projected on the scene has to be determined. To solve the correspondence problem might demand too much computing time and sometimes can be unproductive, using only the geometrical constraints between the camera and the projector. In recent years, several coded structured light techniques have been proposed, so that, the emitted light carries information of the position with respect to the projector coordinate system from where it comes. This information allows us to solve the camera–projector correspondence directly.

This work surveys the different techniques which codify the pattern projected on the scene. The advantages and disadvantages have been discussed for each one of the following methods: (a) to measure dynamic scenes, where only one pattern can be projected to determine the correspondence, and static scenes, where the sequential projection of several patterns is allowed, (b) to measure scenes made by highly saturated colour objects, with basically only binary patterns projected (based on presence or absence of emitted light); or to measure scenes with a colour content predominantly neutral, where the emission of colour is permitted to make an easier pattern codification; specular or metallic surfaces should also be included, but they always lead to a specific system.[70–72] This reflection of a pattern region on another surface, already illuminated, produces an evident identification error of the image captured by the camera; (c) to measure scenes with a lot of discontinuities, of several different depths, where an absolute codification is only allowed. Then, the pattern projected is more complicated, as if we do not want to reduce the resolution, the number of bits has to be increased; or to measure scenes predominantly continuous, where very easy periodical patterns can be projected, which give on an easy identification and high speed solution to the correspondence problem.

Knowing whether there is, or there is not, scene illumination light control, is another important concept in structured light. If it is known that in the scene there is only the light projected by the pattern, or if there is any other source, its properties are known, then, the pattern can be quite simplified. Recently, some new light sources have been proposed, a dot matrix which emits a set of equal intensity dot lights,[73] and a 100% light efficiency intensity spot matrix,[74] which make the segmentation process easier.[75] There are also some projects based on microlaser matrix integration to project patterns with a high resolution,[76] and others which project a dot matrix on highly reflected and dynamic surfaces, as in the sea wave analysis.[77]

But a lot of work remains to be done to obtain a generic vision system which allows us to achieve an instantaneous measurement of an unknown scene, to obtain a 3D map from it, and then, to classify and recognise all the objects that form the scene.[78, 79] There is no doubt that electronic integration, and the improvement of computation power of new processors lead us in the correct direction, but human vision will never be surpassed.

9. SUMMARY

We present a survey of the most significant techniques, used in the last few years, concerning the coded structured light methods employed to get 3D information. In fact, depth perception is one of the most important subjects in computer vision. Stereovision is an attractive and widely used method, but, it is rather limited to make 3D surface maps, due to the

correspondence problem. The correspondence problem can be improved using a method based on structured light concept, projecting a given pattern on the measuring surfaces. However, some relations between the projected pattern and the reflected one must be solved. This relationship can be directly found codifying the projected light, so that, each imaged region of the projected pattern carries the needed information to solve the correspondence problem.

In order to obtain the 3D coordinates of the measuring object from its projections, the mathematical equations and the geometric constraints are analysed. The correspondence problem between two optical sensors or cameras will lead us to an active method. One of the most widely used active methods is based on structured light projection. Mathematical equations, which describe a structured light system, made by the relationship of an optical sensor and a light projector, are broadly defined. Then, a new classification of several studied methods which use coded structured light is presented, analysing temporal dependence, emitted light dependence, and depth surface discontinuity dependence. After that several coded structured light methods are described and compared, explaining the advantages and disadvantages of each method.

We do not need to mention the numerous advantages of obtaining accurate 3D information for many research subjects, such as: robotics, autonomous navigation, shape analysis, and so on. For example, to take out range information of the scene is almost the only way to provide inputs to the navigation algorithms. Industrial applications, as quality control and shape analysis, can also be improved with coded structured light techniques.

REFERENCES

1. P. Cavanagh, Reconstructing the third dimension: interactions between color, texture, motion, binocular disparity, and shape, *Comput. Vision Graphics Image Process.* **37,** 171–195 (1987).
2. R. Ahlers and J. Lu, Stereoscopic vision — an application oriented overview. *Int. Soc. Opt. Eng.* **1194,** 298–308 (1989).
3. S. T. Barnard and M. A. Fischler, Computational stereo, *Comput. Surveys* **14**(4), 553–572 (1982).
4. W. Hoff and N. Ahuja, Surfaces from stereo: integrating feature matching, disparity estimation, and contour detection, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(2), 121–136 (1989).
5. N. Ayache and F. Lustman, Trinocular stereo vision for robotics, *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(1), 73–84 (1991).
6. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision II*, Addison–Wesley, Reading, MA (1993).
7. E. Hall, J. B. K. Tio, C. A. McPherson and F. A. Sadjadi, Measuring curved surfaces for rob/ot vision, *Computer December*, 42–54 (1982).
8. O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA (1993).
9. R. Jarvis, A perspective on range finding techniques for computer vision, *IEEE Trans. Pattern Anal. Mach. Intell.* **5**(2), 122–139 (1983).
10. R. Jarvis, Range sensing for computer vision, Three-*Dimensional Object Recognition Systems*, A. K. Jain and P. J. Flynn, eds, Elsevier, Amsterdam, pp. 17–56 (1993).
11. J. P. Brady, N. Nandhakumar and J. K. Aggarwal, Recent progress in the recognition of objects from range data, *Proc. Int. Conf. on Pattern Recognition*, pp. 85–92 (1988).
12. F. Rocker and A. Kiessling, Methods for Analyzing three dimensional scenes, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 669–673 (1975).
13. P. J. Besl, *Active*, Optical Range Imaging Sensors, *Machine Vision Appl.* **1,** 127–152 (1988).
14. P. M. Will and K. S. Pennington, Grid coding: a preprocessing technique for robot and machine vision, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 66–70 (1971).
15. Z. Yang and Y. F. Wang, Error analysis of 3D shape construction from structured lighting, *Pattern Recognition* **29**(2), 189–206 (1996).
16. N. Alvertos, D. Brzakovic and R. C. Gonzalez, Camera geometries for image matching in 3-D machine vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(9), 897–915 (1989).
17. Y. Shirai and M. Suwa, Recognition of polyhedrons with a range finder, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 80–87 (1971).
18. G. J. Agin and T. O. Binford, Computer descriptions of curved objects, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 629–640 (1973).
19. R. J. Popplestone, C. M. Brown, A. P. Ambler and G. F. Crawford, Forming models of plane-and-cylinder faceted bodies from light stripes, *Proc. Int. Joint Conf. on Artificial Intelligence*, 664–668 (1975).
20. H. Yamamoto, K. Sato and S. Inokuchi, Range imaging system based on binary accumulation, *Proc. Int. Conf. on Pattern Recognition*, pp. 233–235 (1986).
21. O. Ozeki, T. Nakano and S. Yamamoto, Real-time range measurement device for three-dimensional object recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(4), 550–554 (1986).
22. A. Yokoyama, K. Sato, T. Yoshigahara and S. Inokuchi, Realtime range imaging using adjustment-free phot-VLSI — silicon range finder, *Proc. Int. Conf. on Intellligent Robots and Systems*, pp. 1751–1758 (1994).
23. Y. Sato, H. Kitagawa and H. Fujita, Shape measurement of curved objects using multiple slit-ray projections, *IEEE Trans. Pattern Anal. Mach. Intell.* **4**(6), 641–646 (1982).
24. K. Kemmotsu and T. Kanade, Sensor placement design for object pose determination with three light-stripe range finders, Research Report No. CMU-CS-94-152, School of Computer Science, Carnegie Mellon University (1992).
25. K. Kemmotsu and T. Kanade, Uncertainty in object pose determination with three light-stripe range measurements, *IEEE Trans. Robotics Automation* **11**(5), 741–747 (1995).
26. M. Asada and S. Tsuji, Utilization of a stripe pattern for dynamic scene analysis, *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 895–897 (1985).
27. M. Asada, H. Ichikawa and S. Tsuji, Determining of surface properties by projecting a stripe pattern, *Proc. Int. Conf. on Pattern Recognition*, pp. 1162–1164 (1986).
28. M. Asada, H. Ichikawa and S. Tsuji, Determining surface orientation by projecting a stripe pattern, *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(5), 749–754 (1988).
29. Y. F. Wang, A. Mitiche and J. K. Aggarwal, Computation of surface orientation and structure of objects using grid coding, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(1), 129–136 (1987).
30. Y. F. Wang and P. Liang, A new method for computing intrinsic surface properties, *Proc. Computer Vision and Pattern Recognition*, pp. 235–240 (1989).

31. Y. F. Wang, Characterizing three-dimensional surface structures from visual images, *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(1), 52–60 (1991).
32. G. Hu, A. K. Jain and G. Stockman, Shape form light stripe texture, *Proc. Computer Vision and Pattern Recognition*, pp. 412–414 (1986).
33. G. Hu and G. Stockman, 3-D surface solution using structured light and constraint propagation, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(4), 390–402 (1989).
34. G. Stockman and G. Hu, Sensing 3-D surface patches using a projected grid, *Comput. Vision Pattern Recognition*, 602–607 (1986).
35. N. Shrikhande and G. Stockman, Surface orientation from a projected grid, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(6), 650–655 (1989).
36. K. Ikeuchi and K. Sato, Determining reflectance properties of an object using range and brightness of images, *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(11), 1139–1153 (1991).
37. Y. F. Wang and D. I. Cheng, Three-dimensional shape construction and recognition by fusing intensity and structured light, *Pattern Recognition* **25**(12), 1411–1425 (1992).
38. G. Hu and G. Stockman, Representation and segmentation of a cluttered scene using fused edge and surface data, *Proc. Computer Vision and Pattern Recognition*, pp. 313–318 (1989).
39. B. C. Vemuri and J. K. Aggarwal, 3-D model construction from multiple views using range and intensity data, *Proc. Computer Vision and Pattern Recognition*, pp. 435–437 (1986).
40. J. Hoshino, T. Uemura and I. Masuda, Region-based reconstruction of an indoor scene using an integration of active and passive sensing techniques, *Proc. Int. Conf. on Computer Vision*, pp. 568–572 (1990).
41. O. D. Faugeras and G. Toscani, The calibration problem for stereo, *Proc. Computer Vision and Pattern Recognition*, 15–20 (1986).
42. F. Bortolozzi, Vision trinoculaire, une solution geometrique, Ph.D. Thesis, Universite de Technologie de Compiegne (1991) (in French).
43. E. Brassart, Localisation Absolue d'un Robot Mobile Autonome par des Balises Actives et un Systeme de Vision Monoculaire, Ph.D. Thesis, Universite de Technologie de Compiegne (1995) (in French).
44. E. Mouaddib, J. Salvi and J. Batlle, A survey: coded structured light, Technical Report No. LSA/CSL/96, Laboratoire des Systèmes Automatiques, Universite de Picardie, Jules Verne (1996).
45. J. L. Posdamer and M. D. Altschuler, Surface measurement by space-encoded projected beam systems, *Comput. Graphics Image Process.* **18,** 1–17 (1982).
46. M. D. Altschuler, B. R. Altschuler and J. Taboada, Laser electro-optic system for rapid three-dimensional (3-D) topographic mapping of surfaces, *Opt. Eng.* **20**(6), 953–961 (1981).
47. M. D. Altschuler, B. R. Altschuler, J. T. Dijak, L. A. Tamburino and B. Woolford, Robot vision by encoded light beams, *Takeo Kanade Three-Dimensional Machine Vision*, Kluwer Academic Publishers, Dordrecht (1987).
48. S. Inokuchi, K. Sato and F. Matsuda, Range-imaging for 3-D object recognition, *Proc. Int. Conf. on Pattern Recognition*, pp. 806–808 (1984).
49. K. Sato, H. Yamamoto and S. Inokuchi, Tuned range finder for high precision 3D Data, *Proc. Int. Conf. on Pattern Recognition*, pp. 1168–1171 (1986).
50. K. Sato and S. Inokuchi, Range-imaging system utilizing nematic liquid crystal mask, *Proc. Int. Conf. on Computer Vision*, pp. 657–661 (1987).
51. S. Inokuchi, Y. Morita and Y. Sakurai, Optical pattern processing utilizing nematic liquid crystals, *Appl. Opt.* **11**(10), 2223–2227 (1972).
52. K. Hattori and Y. Sato, Handy range finder for active robot vision, *Proc. Int. Conf. on Robotics and Automation*, pp. 1423–1428 (1995).
53. E. Müller, Fast three dimensional form measurement system, *Opt. Eng.* **34**(9), 2754–2756 (1995).
54. B. Carrihill and R. Hummel, Experiments with the intensity ratio depth sensor, *Comput. Vision Graphics Image Process.* **32,** 337–358 (1985).
55. K. L. Boyer and A. C. Kak, Color-encoded structured light for rapid active ranging, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(1), 14–28 (1987).
56. J. Le Moigne and A. M. Waxman, Projected light patterns for short range navigation of autonomous robots, *Proc. Int. Conf. on Pattern Recognition*, Vol. 1, pp. 203–206 (1984).
57. J. Le Moigne and A. M. Waxman, Multi-resolution grid patterns for building range maps, *Proc. Vision Conf.* **8**(22–39) (1985).
58. J. Le Moigne and A. M. Waxman, Structured light patterns for robot mobility, *IEEE J. Robotics and Automation* **4**(5), 541–548 (1988).
59. H. Morita, K. Yajima and S. Sakata, Reconstruction of surfaces of 3-D objects by M-array pattern projection method, *Proc. Int. Conf. on Computer Vision*, pp. 468–473 (1988).
60. P. Vuylsteke and A. Oosterlinck, Range image acquisition with a single binary-encoded light pattern, *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(2), 148–164 (1990).
61. T. Pajdla, BCRF—Binary-coded illumination range finder reimplementation, Technical Report No. KUL/ESAT/MI2/9502, ESAT Katholieke Universiteit Leuven (1995).
62. J. Tajima and M. Iwakawa, 3-D data acquisition by rainbow range finder, *Proc. Int. Conf. on Pattern Recognition*, pp. 309–313 (1990).
63. Z. J. Geng, Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems, *Opt. Eng.* **35**(2), 376–383 (1996).
64. V. Smutny and T. Pajdla, Rainbow range finder and its implementation at the CVL, Research Report No. K335 -1996, Computer Vision Laboratory, Cech Technical University (1996).
65. C. Wust and D. W. Capson, Surface profile measurement using color fringe projection, *Mach. Vision Appl.* **4,** 193–203 (1991).
66. G. T. Reid, Automatic fringe pattern analysis: A review, *Opt. Lasers Eng.* **7,** 37–68 (1986).
67. D. C. D. Hung, 3D scene modelling by sinusoid encoded illumination, *Image Vision Comput.* **11**(5), 251–256 (1993).
68. P. M. Griffin, L. S. Narasimhan and S. R. Yee, Generation of uniquely encoded light patterns for range data acquisition, *Pattern Recognition* **25**(6), 609–616 (1992).
69. M. Maruyama and S. Abe, Range sensing by projecting multiple slits with random cuts, *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(6), 647–651 (1993).
70. S. Hata, Shape detection of small specular surface using color stripe lighting, *Proc. Int. Conf. on Pattern Recognition*, pp. 554–557 (1992).
71. D. W. Capson and S. Eng, A tiered-color illumination approach for machine inspection of solder joints, *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(3), 387–393 (1988).
72. A. C. Sanderson, L. E. Weiss and S. K. Nayar, Structured highlight inspection of specular surfaces, *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(1), 44–55 (1988).
73. J. S. Darlin, P. Senthilkumaran, S. Bhattacharya, M. P. Kothiyal and R. S. Sirohi, Fabrication of an array illuminator using tandem Michelson interferometers, *Opt. Commun.* **123,** 1–4 (1996).
74. J. Glückstad, Adaptative array illumination and structured light generated by spatial zero-order self-phase modulation in a Kerr medium, *Opt. Commun.* **120,** 194–203 (1995).

75. R. M. Haralick and L. G. Shapiro, Survey: image segmentation techniques, *Comput. Vision Graphics Image Process.* **29,** 100–132 (1985).
76. H. R. Nahata and M. Murdocca, Decomposition of two-dimensional microlaser patterns, *Appl. Opt.* **35**(8), 1195–1204 (1996).
77. I. Grant, Y. Zhao, G. H. Smith and J. N. Stewart, Split-screen, single-camera, laser-matrix, stereogrammetry instrument for topographical water wave measurements, *Appl. Opt.* **34**(19), 3806–3809 (1995).
78. P. J. Besl and R. C. Jain, Invariant surface characteristics for 3D object recognition in range images, *Comput. Vision Graphics Image Process.* **33,** 33–80 (1986).
79. R. Hoffman and A. K. Jain, Segmentation and classification of range images, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(5), 608–620 (1987).

About the Author—JOAN BATLLE, graduated in Physics at the Autonomous University of Barcelona, received the Ph.D. degree in Computer Science at the Polytechnical University of Catalunya. At present, he is Director of the Computer and Robotics Group and head of the Department of Electronics, Automatics and Computer Engineering. His research activity is mainly focused on real-time vision and autonomous robots, and he is actually involved in some governmental projects about underwater robots and technology transference to industrial environments.

About the Author—EL MUSTAPHA MOUADDIB received the Ph.D. degree in robotics at the University of Picardie Jules Verne, Amiens, France, in 1991. Since 1992, he is an associate professor at the Institut Universitaire Professionnalisé, Electrical Engineering, where he teaches Robotics and Vision. His main interests are computer vision, camera calibration, mobile robotics and artificial visual perception. Since 1995, he is head of the Robotics Perception group at the Laboratory of Automatic Systems of the same university, where he is engaged in research projects on robot navigation with a conic sensor and in particular localisation and obstacle detection.

About the Author—JOAQUIM SALVI received the B.S. degree in Computer Science at the Polytechnical University of Catalunya in 1993. He joined the Computer Vision and Robotics Group at the University of Girona in September 1993, from where he received a M.S. degree in Computer Science in July 1996 and was engaged in the study of structured light applied on computer vision. At present, he is a lecturer of the Electronics, Automatics and Computer Engineering Department at the University of Girona, and he is working for his Ph.D. degree in the Computer Vision and Robotics Group at the same University of Girona. His current interests are in the field of computer vision, stereo vision, camera calibration, and he focuses on the study of structured light as a technique to recover 3D information from a lighted scene.