ISOMA-016.1

A Visual Inspection System Based on Fundamental Matrix and Structured Light

Nuno Martins and Jorge Dias
I.S.R. - Instituto de Sistemas e Robótica
Departamento de Engenharia Electrotécnica da Universidade de Coimbra - Portugal

ABSTRACT

This article describes a system for estimation of distances, areas and volumes of opaque and static objects by using stereo images and exploring the epipolar geometry based on the fundamental matrix and structured light. The article describes the design of a low cost, fast and compact system for visual inspection and three-dimensional reconstruction using images. To obtain the stereo images, the system simulates a binocular stereo system by using mirrors and one stereo camera. The correspondence between points in the stereo images, is easily obtained by the epipolar geometry generated by the fundamental matrix and the structured light. A fast and robust computation of the fundamental matrix is achieved by a geometric configuration equivalent to the visual fixation. The scene is illuminated with a liner radiation in the infrared band, allowing the definition of features generated by the interference of light with the objects in the scene. In the case of planes of light, the curves can be associated with the epipolar lines generated by the fundamental matrix to give a mechanism to detect the correspondent point without ambiguity. Future developments will use different light patterns (circles and dots) to obtain other type of measurements (e.g. curvatures).

INTRODUCTION

The appearing of the investigation area of the computer vision allowed the development of a new kind of machines that, exploring the visual information, are capable of accomplish tasks minimizing the human intervention. One of the goals of this domain is the deduction of the three-dimensional information of a scene, analyzing the images from that scene. The recover of three-dimensional information from images will allow its use in tasks of great importance for the industry because, being performed by machines, can improve metrology inspection and quality testing. The majority of efforts on solutions for three-dimensional imaging for metrology and inspection rely on structured light projection.

The structured light projection in a scene, for extraction information of its structure, was first proposed by Will and Pennington[7], in 1971. The technique is based in the projection of a pattern, composed by perpendicular lines equally spaced. Wang, Mitiche and Aggarwal[8], use an identical pattern to obtain the local orientation of the surface of an object, which rest in a base plane. The possibility proposed by Will and Pennington[7] give a place for various techniques, which differ uniquely in the type of the pattern. Posdamer and Altschuller[9] proposed a temporal modulation, using a controlled beam laser matrix. Each beam has a unique code, allowing its identification in the image without ambiguity. Le

Moigne and Waxman[10] use a labeling technique to identify the cross points of the pattern (this pattern has reference points). Hu and Stockman[11] obtain this labeling by a restraint propagation. The pattern used by Vuylsteke and Oosterlinck[12] is composed by white and black squares of two sizes. Boyer and Kak[13] proposed a pattern encoded by color.

This work describes another approach for the use of structured light, developed to create a low cost, fast and compact system for visual inspection and three-dimensional reconstruction, addresses the problem of the estimation of distances, areas and volumes of opaque and static objects by using stereo images (captured by only one camera), and exploring the epipolar geometry based on the fundamental matrix and structured light. In the first stage, the goal is the reconstruction of three-dimensional information of the scene points by image triangulation, using the stereo images.

In order to build a low cost system it is simulated a binocular stereo system by using only a camera and a set of mirrors for the acquisition of stereo images, showed in the figure 1. It gives the equivalent geometry of two cameras observing the same scene from different viewpoints and makes easy the reconstruction solution.

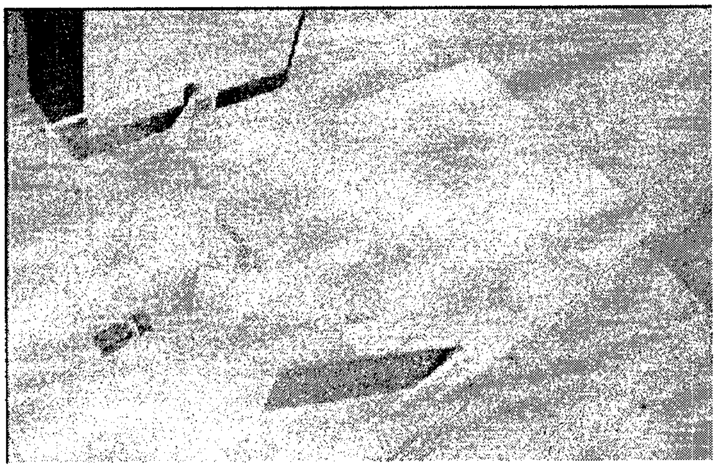


Figure 1. Image of the experimental system.

The geometry involved, called epipolar geometry, allows to establish a transformation that relates the projection on the images and can be described by a 3 x 3 matrix, called fundamental matrix F. This matrix helps to solve the correspondence problem.

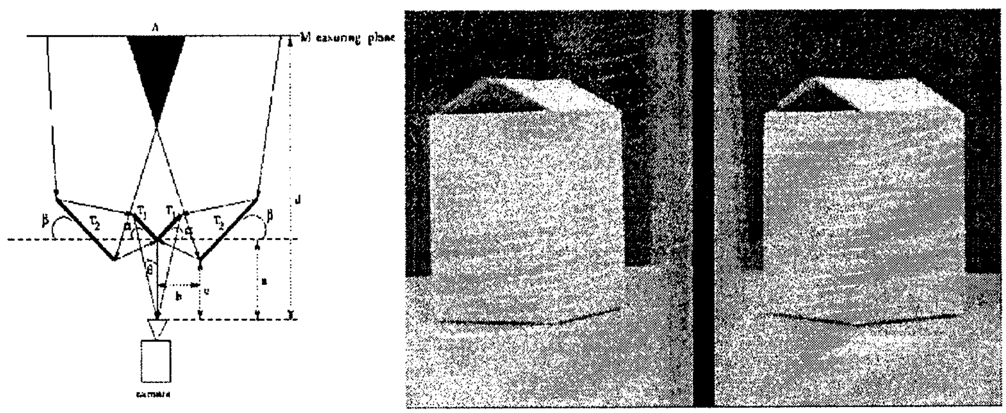
The scene is illuminated with a laser radiation in the infrared band, allowing the definition of features generated by the interference of light with the objects in the scene. This interference effect is registered in the stereo images by using an interference filter, that filters all radiation that not belongs to the infrared band. The obtained images are the curves of the interference of the light with opaque objects in the scene. In the case of planes of light, the curves can be associated with the epipolar lines generated by the fundamental matrix to give a mechanism to detect the correspondent point without ambiguity.

THE MIRROR ACQUISITION SYSTEM

To recover the three-dimensional scene from images, exploring the disparity between projected points, is necessary to grab, at least, two images of the same scene from different positions. A mirror system based on stereoscope allows the acquisition of the two images in only one snapshot, which is the liaison of two different views of the same scene. Each image is taken by a virtual camera, cloned from the real camera, i.e. with the same characteristics.

A process based on a mirror system of this kind has two great advantages: the cost and the versatility. The mirrors give us the possibility y to replicate expensive sensors, obtaining the same results. A system like this, also, allow us to create a compact system, easy to handle, where we can control its size and volume.

To achieved the final geometry for the mirrors acquisition system, we simulate the presence of two virtual cameras. Due the rectangular format of the camera sensor, the simulation could be done in a bi-dimensional plane (see e.g. case in the figure 2a). Knowing the size of the camera's sensor and using geometrical calculation, we complete the 3D study of the system. This symmetrical configuration looks like the one revealed by INRIA[1], being different only in the possibility of $a \neq b$. The motive to this little change is due to the attempt to decrease the space occupied by the camera and the set of mirrors. The information considered important to the choice of the geometry we want to set **in** the construction of the mirror system can be seen in the figure 2a. This figure is the final configuration of the mirrors.



(a) **(b)**

Figure 2. (a) Lateral cut of a symmetrical configuration; (b) Compact stereo image, resulting from mirror acquisition system.

EPIPOLAR GEOMETRY

Two viewpoints of a single scene or object, obtained either from stereo or a moving camera, are related by the epipolar geometry, can be described by a 3×3 singular matrix. This article uses a constraint for a fast and robust computation of the this matrix and it is achieved by a geometric configuration equivalent to the visual fixation (keeping the same feature on the scene, in the center of the two stereo images) and similar to the approach used on active vision systems. The resulting geometry simplifies the equations, and implies only the knowing of six correspondent points in the stereo images in order to compute the this matrix.

The projective model

The camera model consider in this article, which is the most widely used, is the pinhole model that we refer as the perspective projection. The co-ordinates of the generic 3D point $P(X, Y, Z)$ in a world co-ordinate system and its image co-ordinates $p(x, y)$ are related, in homogeneous co-ordinates, by

$$k\,\overline{p} = C \begin{bmatrix} R & t \end{bmatrix} \overline{P} \tag{1}$$

where *k is* an arbitrary scale factor. The rigid transformation between the two image planes is given by a rotation and a translation (see figure 3). From these camera's positions, the generic three-dimensional point, *P,* projects onto the image planes at p_l and *p*,. The line formed by C_l and C_r intersects I_l and *I*,, originating the epipoles e_l and e,.

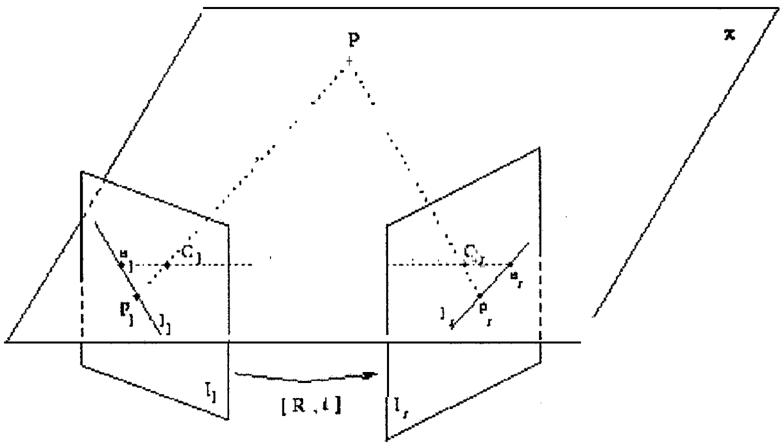


Figure 3. Epipolar Geometry. Two images of a point P are taken whose focal point is at positions C_l and C_r. The plane through $C_l\ PC_r$ is the epipolar plane for P. This plane intersects the images along the two epipolar lines l_l and l_r.

Every epipolar lines of I_l runs through e_l, as well as epipolar lines of I_r holds e,. This property has important implications when attempting to match points in the two images. Suppose that *P* has been observed at position *p,, in* the left image. Thus, the position of p_r *is* constrained to lie somewhere on the epipolar line 1,. Consequently, given a mechanism for calculating 1,, corresponding to a point p_l, the search space for the matching point p_r can dramatically reduced. The epipolar constraint

$$p_r^T M p_l = 0 \qquad (2)$$

provides just such a mechanism, where M is the singular 3 x 3 matrix named fundamental matrix[2][4].

Visual fixation constraint

Assuming that the co-ordinates system of the virtual cameras and the world do not have relative motion, the calculus of the fundamental matrix become simplified. We also assumed that $Y_l = Y_r = -Y$ and that the planes $X_l Z_l$, $X_r Z_r$ and XZ are in the same plane, or in an other way the system has no cyclotorsion. To derive *F* for such situation, let us consider geometry of the system for the general case, presented in figure 4. The left and right cameras have different vergence ϕ_l and ϕ_r, respectively, and their optical axes pass through P_f, the fixation point of the cameras. This geometry, as it can be verified in the figure 4, can be regarded as two cameras related by a translation *t* along the baseline *B,* and a rotation *R* about Y axis, obtaining, thus, *F.* Those are given by

$$t=\begin{bmatrix} B\sin(\phi_r) \\ 0 \\ B\cos(\phi_l) \end{bmatrix} \qquad R=\begin{bmatrix} -\cos(\phi_l+\text{fib,}) & 0 & -\sin(\phi_l+\phi_r) \\ 0 & 1 & 0 \\ \sin(\phi_l+@) & 0 & \cos(\phi_l+\phi_r) \end{bmatrix} \qquad F=\begin{bmatrix} 0 & f_{12} & f_{13} \\ f_{21} & 0 & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

This new form of F is called common elevation fundamental matrix[3]. For the chosen geometric configuration of the mirrors, ϕ_l =@. =@.

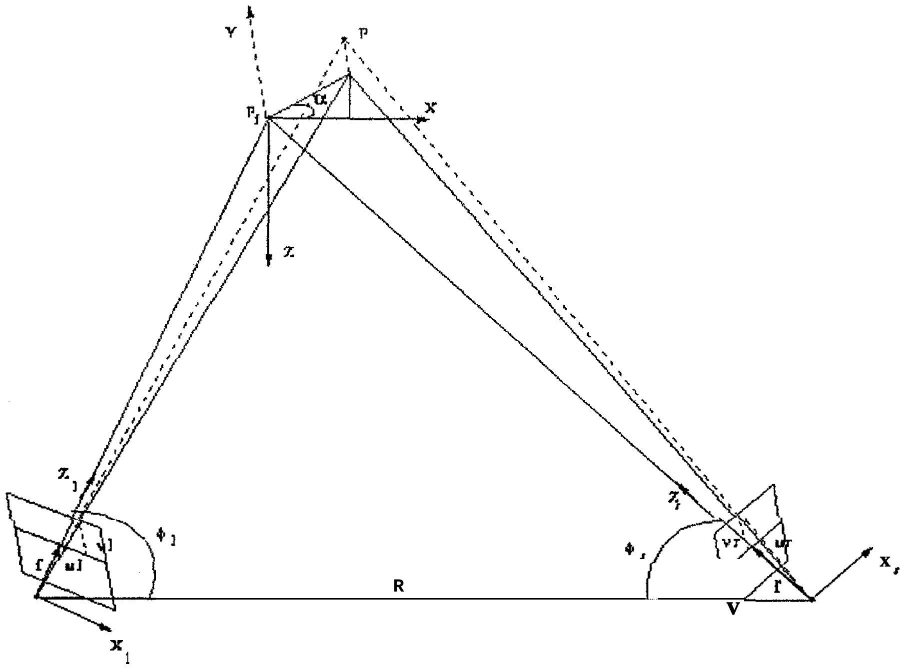


Figure 4. Representation of the co-ordinate system of the two virtual cameras and the world. The optical axes of each camera pass in P_f .

Estimating the fundamental matrix

The problem consider in the sequel is the estimation of F from a sufficiently set of point correspondences: $(\overline{p_{l\,i}}, \overline{p_{r\,i}})\ i = 1,\dots,n$, where $n \geq 6$.

Rewriting the equation (2) as a linear and homogeneous equation with the nine parameters of F, its aspect would be

$$mf = 0 \tag{3}$$

where

$m = [\ x_r y_l\ x_r\ \ x_l\ \ y_r\ \ y_r\ \ x_l\ \ y_l\ \ 1\ ]$ and $f = [f_{12}\ f_{13}\ f_{21}\ f_{23}\ f_{31}\ f_{32}\ \ 1]^T$.

If we have $n >> 6$ matched points, using (3), we could built a set of linear equations $f = [m_1\ m_2\ \ .\ .\ .\ \ m_n]^T$. This improve the precision on the computation of f, by the least-squares estimation method, breed on singular value decomposition[5]. Joining, afterwards, the rank constraint to fundamental matrix[4], it allow us to have a better estimation of the epipolar geometry. The result of the estimation of the fundamental matrix for the image in figure 2b is presented in figure 5 with the correspondent epipolar lines for each point used to compute the fundamental matrix.

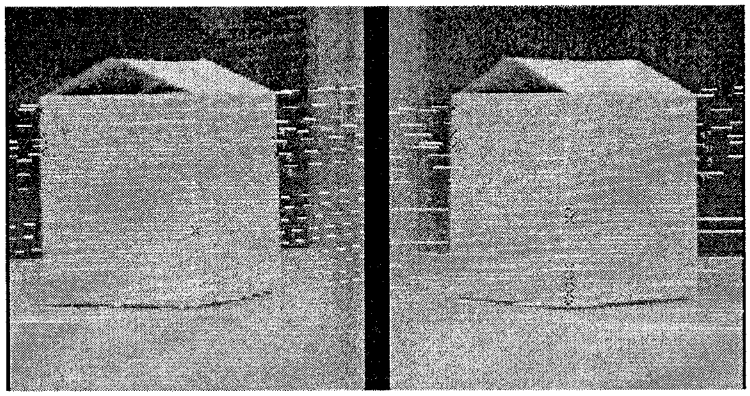


Figure 5. Intimation of the fundamental matrix and representation of the epipolar lines.

LIGHT CONSTRAINT AND 3D RECONSTRUCTION

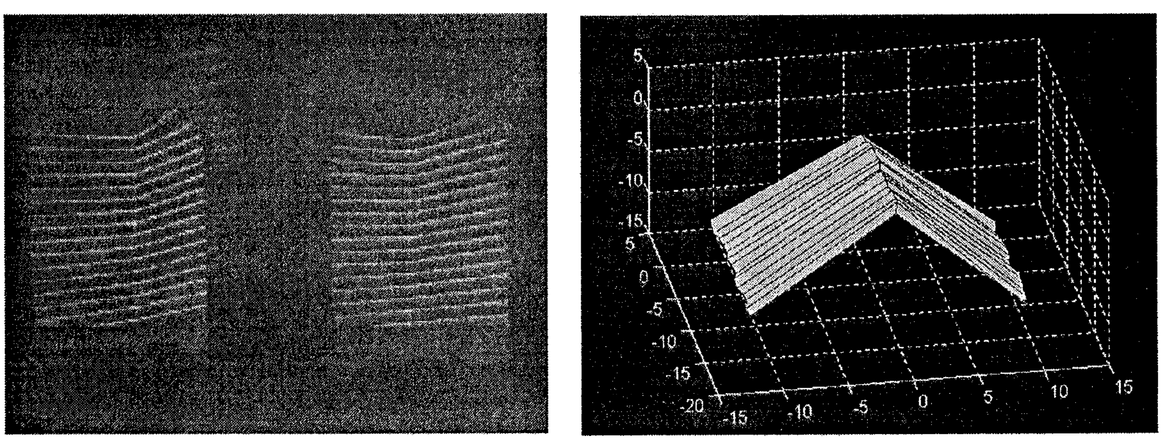
As you can see in figure 1, in the back of the minor mirrors, there are a strutted light sensor. This sensor has the capacit y of emitting various patterns of light, like dots, lines and circles. For a more inexpensive approach of a system like this, we can substitute the structured light sensor by a normal source of light, which send a particular pattern, design to achieve some especially purpose.

In this phase, we only use planes of light and smooth objects. The reasons are that the projection a light plane through a empty measuring plane is a line, which is an easy feature to identify the equation. With the smooth objects the stripe change very little, so we can tag and identify, from the initial point of the stripe to the final point, with no doubt.

The acquisition of a compact image of the illuminated scene, as you can see in figure 6a. As the radiation is in infrared band, we put on the camera an interference filter that cut all radiation that not belongs to this band. The integral pattern could be seen in both stereo images, making the correspondence between stripes automatically done.

Joining the epipolar knowledge with the light constraint, we obtain a univocal correspondence of two image points, from the intersection of the projected stripe and epipolar line. Both lines pass through those points. For a better reconstruction, we must obtain a group of this true correspondent image points much larger as possible.

The three-dimensional points reconstruction is based in the positions of its projections in



(b)

Figure 6. (a) Compact image of the laser stripes projected on the object; (b) Result of the reconstruction algorithm applyed to the object represented in figure 2b, and obtained from the information given by the figures 5 and 6a.

the images. Since the univocal correspondence between some image projections are establish, the next step is to recover its 3D positions using triangulation. Lets consider the geometry of the figure 4. Due the positioning of the mirrors we consider that both of the virtual cameras have the same vergence ($\phi_l = \phi_r = \phi$). The baseline, B, is the distance between the optical centers of the virtual cameras. The cameras have the same characteristic, because they origin is the same real camera and the mirror properties don't change those characteristics.

In our case, from the projective model equations for each camera, where

$$\overline{P_l} = [R_l,\ t]\overline{P} \qquad \overline{P_r} = [R_r,\ t]\overline{P} \tag{4}$$

and the rotations and translation matrix given by

$$R_l = \begin{bmatrix} \sin(\phi) & 0 & \cos(\phi) \\ 0 & -1 & 0 \\ \cos(\phi) & 0 & -\sin(\phi) \end{bmatrix} \qquad R_r = \begin{bmatrix} \sin(\phi) & 0 & -\cos(\phi) \\ 0 & -1 & 0 \\ -\cos(\phi) & 0 & -\sin(\phi) \end{bmatrix} \qquad t = \begin{bmatrix} 0 \\ 0 \\ \frac{B}{2\cos(\phi)} \end{bmatrix}$$

Due the geometric configuration of the system, we know that $\overline{P_l}(2) = \overline{P_r}(2)$ (note that $P(i)$ is the line i of the matrix P). For a general 3D point, as can be verified in the figure 4, $\tan(\alpha)$ = Z/X (in [6] it can be seen how to obtain tan(a)). Using this knowledge in the equation (4), we obtain

$$Z = \frac{c_2 \tan(\alpha)}{1 - c_1 \tan(\alpha)} \qquad X = \frac{c_1 c_2 \tan(\alpha)}{1 - c_1 \tan(\alpha)} + \frac{B}{2}$$

with

$$c_1 = \frac{y_l - y_r}{y_l + y_r}\tan(\phi) \qquad c_2 = \frac{B(y_r - y_l)}{2(y_l + y_r)\cos^2(\phi)} \qquad \tan(\alpha) = \frac{x_r y_l - x_l y_r}{x_r y_l + x_l y_r}\tan(\phi)$$

The Y co-ordinate is obtained using

$$\frac{f_y \overline{P}(2)\left[\overline{P_l}(3) - \overline{P_r}(3)\right]}{y_r - y_l} = \frac{f_x\left[\overline{P_r}(1)\overline{P_l}(3) - \overline{P_l}(1)\overline{P_r}(3)\right]}{x_r - x_l}$$

and result as

$$Y = k_a \frac{y_r - y_l}{x_r - x_l}\left\langle \left\{ \frac{c_2 \tan(\alpha)\left[c_1 + \tan(\alpha)\right]}{1 - c_1 \tan(\alpha)} + c_2 \right\} \sin(\phi) - \frac{B \tan(\alpha)}{2\cos(\phi)} \right\rangle$$

where k_a is the aspect rate of both stereo images (due the symmetric geometry).

The three-dimensional reconstruction is obtained by an interpolation of a group of 3D points, as you can see in figure 6b. The horizontal and vertical axes are the X and Z axes, respectively. The other axe is Y.

CONCLUSION

Any technique that leads to the acquisition of two images of a scene, taken from two different three-dimensional positions, can use the triangulation principle. If in one of the positions, the acquisition sensor is replaced by the projection of a pattern of light on the scene, the projected pattern can be faced as a projection image created in the scene. So, using the potential of the mirrors, the duplication of this sensor can be simulated. This perspective gives us the possibility to create a thought based on the duality of the measuring results, in which one uses two receivers and an emitter of light and vice-versa.

Future developments will use different light patterns (circles and dots) to obtain other type of measurements (e.g. curvatures). We expect to apply this system in applications in industry for quality test and metrology, because its characteristics of low cost.

REFERENCES

1. Mathieu, Hervé et Devernay, Frédéric, "S ystème de miroirs pour la stéréoscopie", Rapport technique, n. 172, Institut National de Recherche en Informatique et en Automatique, Juin, 1995.
2. Longuet-Higgins, H., "A computer algorithm for reconstructing a scene from two projections", *Nature*, 293:133-135, 1981.
3. Fuxing Li, Michael Brady and Charles Wiles, ``Fast Computation of the fundamental matrix for an active stereo vision system", Deportment of Engineering Science, University of Oxford.
4. Zhengyou Zhang, "Determining the epipolar geometry and its uncertainty: A review", Technical report n°2927, July, 1996.
5. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, "Numerical recipes in C: The art of scientific computing", Cambridge University Press, 1992.
6. Daphna Weinshall, "Qualitative depth and shape from stereo, in agreement with psychophysical evidence", Massachusetts Institute of Technology Artificial Intelligence Laboratory and Center for Biological Information Processing Whitaker College, December, 1987.
7. P. M. Will, K. S. Bennington, "Grid Coding: A preprocessing technique for robot and machine vision", *Artificial Intelligence*, *2*, pp. **319-329, 1971.**
8. **Y.** F. Wang, A. Mitiche, J. K. Aggarwal, ``Computation of surface orientation and structured of objects using grid coding", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 129-137, January, 1987.
9. J. L. Posdamer, M. D. Altschuller, ``Surface measurement by space-encoded project beam system", *Computer Graphics and Image Processing*, 18, pp. 1-17, 1982.
10. J. J. Le Moigne, A. M. Waxman, "Structured light patterns for robot mobility", *IEEE Journal of Robotics and Automation*, vol. 4, no. 5, pp. 541-548, October, 1988
11. G. Hu, G. Stockman, ``3D surface solution using structured light and constraint propagation", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 4, **pp.** 390-402, April, 1989.
12. P. Vuylsteke, A. Oosterlinck, "Range image acquisition with a single binary-encoded light pattern", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 148-164, February, 1990.
13. K. L. Boyer, A. C. Kak, ``Color-encoded structured light for rapid active ranging", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 14-28, January, 1987.