Heinz Hügli & Gilbert Maître

3D by Structured Light: Implementation and Evaluation of a Vision System for Small Parts

Abstract

This paper presents principle, implementation and evaluation of a simple three-dimensional (3D) measurement system by structured light, used for computer vision of small parts. The system principle is given first: triangulation based both on structured light and the image from a tv camera. Then, the implementation part describes the various components of the system which are the hardware setup, the image processing software, the geometrical transform and the system calibration procedure. Finally, the performance of the 3D measurement system is presented. By giving examples of measured objects showing typical behavior of the system and by presenting the results of the overall error measurement.

Introduction

Computer vision needs input from various complementary sources. Among them, active ranging methods for measuring three-dimensional (3D) objects have been widely studied [1]. Two broad categories are the time-of-flight methods and the more classical methods based on triangulation. The time-of-flight methods have the advantage to be free of shadows but they are not well suited for short distances, that is for small objects. A very large number of active 3D measuring methods have been proposed, tested and some performant devices are available

on the market. However, few such systems are available which are capable of measuring whole objects and simple and inexpensive enough for small vision systems.

The 3D measurement system described in this paper has been developped as an alternative [3] range data input for a laboratory vision system. It uses basic triangulation, standard components available on the market, it is simple and inexpensive.

Next comes the formalisation of triangulation with structured light, then, a description of system implementation, and finally, a presentation of system performance.

3D by structured light

Perspective transform

We first formalize the perspective transform performed by the camera, leading from the 3 dimensional object space into the 2 dimensional image space. We refer to figure 1 and use homogeneous coordinate systems which represent points of

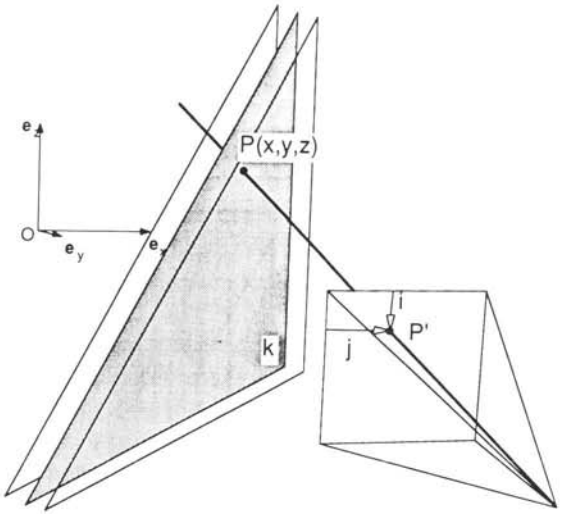

Fig. 1 Triangulation between camera and illumination plane

the n-dimensional space by vectors of dimension n+1. A point $\mathbf{x}=[\ x\ y\ z\ ]$ of the object space is represented by the corresponding homogeneous vector $\mathbf{X} = [\ X\ Y\ Z\ W\ ]$ where following relations hold

$$x = X/W,\ y = Y/W,\ z = Z/W\ ;\ W \neq 0 \qquad (1)$$

Similarly, a point $\mathbf{i}=[i\ j\ ]$ in the image space is represented by its corresponding homogeneous vector $\mathbf{I} = [\ I\ J\ H\ ]$ where following relations hold

$$i = I/H,\ j = J/H\ ;\ H \neq 0 \qquad (2)$$

Notice that the transform into an homogeneous space has arbitrary scaling: in above equations, W and H can be chosen arbitrarily.

Now the perspective transform from object to image space is expressed by [2]:

$$[\ I\ J\ H\ ] = [\ X\ Y\ Z\ W\ ] \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \\ c_{41} & c_{42} & c_{43} \end{bmatrix} \quad \text{or also} \quad \mathbf{I} = \mathbf{X} \cdot \underline{\mathbf{C}} \qquad (3)$$

Thus, expressed in homogeneous coordinates, the perspective transform is linear.

3D measurement is concerned with the inverse transform, leading from **I** to **X**. Obviously, the inverse perspective transform has infinite solutions: there are infinite object points with same projection **i**. In fact these points form the viewing line through **i**. In order to obtain a defined point **x** for each observation **i**, we constrain the points **x** to a plane and name constrained inverse perspective transform the way to compute it.

Intersection by a plane

A family of K planes in the homogeneous object space is given by the plane equations

$$A_k X + B_k Y + C_k Z + D_k W = 0; \qquad k = 1 \ldots K \qquad (4)$$

and the homogeneous vectors $\mathbf{N_k} = [A_k\ B_k\ C_k\ D_k]$ are the normals to the planes.

Constrained inverse perspective transform

Combining eq (3) and (4), we substitute Y in (3) in order to get the linear form:

$$[\ I\ J\ H\ ] = [\ X\ Z\ W\ ]\ \underline{\mathbf{T}}_k \qquad (5)$$

where $\underline{\mathbf{T}}_k$ is a 3x3 square matrix. Consequently the transform $\underline{\mathbf{T}}_k$ has an inverse $\underline{\mathbf{U}}_k = \underline{\mathbf{T}}_k^{-1}$ as long as $|\underline{\mathbf{T}}_k| \neq 0$, which means that the constrained inverse perspective transform exists and is simply given by:

$$[\ X\ Z\ W\ ] = \underline{\mathbf{U}}_k\ [\ I\ J\ H\ ] \qquad (6)$$

H can be chosen arbitrarily, i.e. H=1. Finally it results that the constrained inverse perspective transform consists of the successive application of equations (6), (4) and (1).

Calibration

Clearly the plane normal [A_k B_k C_k D_k] and the transform matrix $\underline{U_k}$ depend on the geometry of the measurement setup. These transform parameters are estimated in the calibration phase. The coefficients of the plane can be estimated from the measured [x y z] values of a number of points in the plane. Three non-colinear points are required for the plane calibration.

Concerning the $\underline{U_k}$ calibration, it can be shown [4] that developing X-xW=0 and Z-zW=0 leads to

$$\begin{cases} i \cdot u_{11} + j \cdot u_{12} + u_{13} - x \cdot i \cdot u_{31} - x \cdot j \cdot u_{32} - x \cdot u_{33} = 0 \\ i \cdot u_{21} + j \cdot u_{22} + u_{23} - z \cdot i \cdot u_{31} - z \cdot j \cdot u_{32} - z \cdot u_{33} = 0 \end{cases} \quad (7)$$

Therefore, the above system consists of two linear independant equations of the 9 coefficients u_{11}, u_{12}, u_{13}, u_{21}, ..., u_{33}. Because of the nature of homogeneous transforms, $\underline{U_k}$ can be scaled arbitrarily. Setting for instance $u_{33}=0$, 8 coefficients remain. Because we measure [x z] and [i j] for any given calibration point, we require at least 4 points for the calibration of $\underline{U_k}$. With 4 points, $\underline{U_k}$ is computed directly whereas least square estimation is used for more calibration measurements. Moreover, no 3-point-colinearity is allowed.

Illumination displacement versus object displacement

In order to scan an object completely with the illumination plane, we can either displace the illumination or displace the object. In the case of a displacement of the illumination, we have a family of different configurations, each characterized by its plane normal $\mathbf{N_k}$ and its transform matrix $\underline{U_k}$. Therefore a calibration is required for each configuration and also the parameters for each configuration must be stored.

In the case of object displacement, a single calibration is required and the constrained inverse perspective transform can accommodate the object displacement very simply. To do so, let us consider, together with the original coordinate reference R, a second coordinate reference R' bound to the object. All measurements are performed in the reference system R and then transformed into the reference system R' by considering the variable displacement $\mathbf{x_o}$:

$$\mathbf{x'} = \mathbf{x} - \mathbf{x_o} \quad (8)$$

In conclusion, only a single calibration is required for the case of object displacement ; the transform is the same for all single measurements, except for the subsequent translation of equation (8).

System design

Setup

Figure 2 shows the measurement setup. Its main parts are the illumination, the camera and the image analysis system. The illumination consists of a 5 mW He-Ne laser equipped with a cylindrical lens V-SLM-015 by Newport, which transforms the laser beam from a spot cross-section into a stripe cross-section; it is this stripe, when it propagates, that creates the illumination plane. The camera is a standard

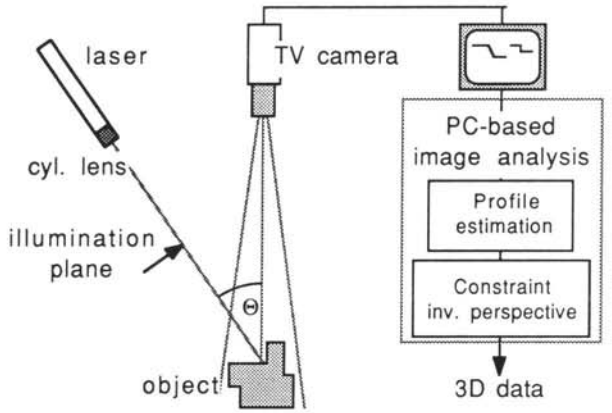

Fig. 2 3D measurement setup

black and white CCD camera K210 by Siemens, furnished with a 18-108 / 2.5 zoom lens. The image analysis system consists of a PC-AT equipped with a PCvision acquisition board by Imaging Technology showing a 512x512 pixel image. The complete hardware comes for less than $ 6000.-

The geometry of the measurement setup is chosen for a volume of interest given by a cube of 1 dm^3. The object is on the horizontal reference plane z=0; the camera is mounted vertically above it at a height of z=75 cm, with its viewing axis oriented vertically down. The plane of illumination falls from above on the horizontal reference plane with an incidence angle Θ (fig. 2). The choice of Θ is a

compromise between a higher measurement resolution in z (large values of Θ) and a higher attenuation of shadowing effects (small values of Θ). A value of Θ = 36° was chosen. Furthermore, the reference system R is chosen such that the equation of the illumination plane becomes $y = z\ tg\Theta$, i.e. $\mathbf{N} = [0\ 1\ -tg\Theta\ 1]$.

Measurement procedure

In a single measurement, the illumination plane intersects the object along a line which produces a profile line in the image plane of the camera (figure 3). The profile line usually consists of several segments. Each pixel [i j] belonging to the profile can now be transformed into the object space by the constrained inverse perspective transform followed by the translation of equation (8). The measurement of a complete object is performed profile by profile, repeating the above described single measurement for successively increasing values of the displacement

$$\mathbf{x_o} = k \cdot \Delta\mathbf{x_o} \quad ; k = 1 \ldots K \qquad (9)$$

Fig. 3 Light profile in the image

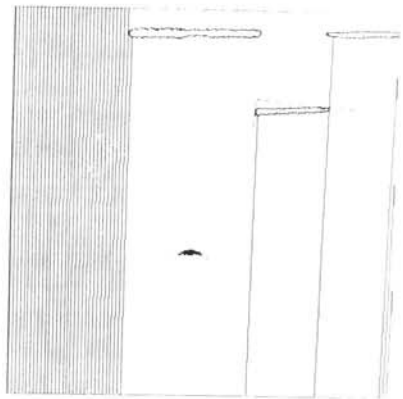

Fig. 4 Algorithm for profile estimation

Profile estimation

Ideally considered as a line of negligible thickness, the profile appears in the grey-level image of the camera as a collection of white thick segments. Profile estimation is the process which transforms such an image in a collection of line segments. The process includes three steps, namely image thresholding, segment extraction and segment thinning.

Image thresholding is performed interactively in the common way and once for all measurements.

The next steps use a propriety of the profile, namely the fact that in principle each beam of light leaving the light source has at maximum one intersection with the object. Obviously, this is also true in the transformed image space. Considering an ordered set of homogeneously spaced beams of light leaving the light source and belonging to the illumination plane, its transform in the image space is a set of N straight lines $a_n i + b_n j + c_n = 0$; n=1...N. This set of light beams can be used as guides for segment extraction and segment thinning.

Rather than testing the whole image area for the presence of a profile segment, the segment extraction algorithm proceeds in a more efficient way consisting of two phases (figure 4) . The first phase is aimed at finding a segment. It proceeds by exploring the image, using the set of light beams as guides, until a segment is found. Once a segment is found, the algorithm enters phase two, which performs contour following in order to completely extract the segment. Moreover, during contour following, the algorithm flags all light beams touching the segment, removing them from the list of beams to be explored.

Finally, segment thinning is performed by keeping as segment elements, the one dimensional centers of mass, computed on lines, parallel to the light beams.

Calibration

The calibration of the illumination plane is simplified by the particular choice of the setup geometry as previously described. The only parameter is $tg\Theta$ and its measurement is obvious. For the calibration of the constrained inverse perspective transform matrix $\underline{\mathbf{U_k}}$, we measure the $[i_m\ j_m]$ and corresponding $[x_m\ z_m]$ for m=1...8 points and compute the least square errore estimate of the matrix.

Experimental results

Performance

The typical performance of the 3D measurement system is illustrated by some examples. Figures 5, 6, 7, ..., 10 show three objects and their measured data. The measured data are the profiles, which are shown in form of a perspective view of the objects. In these examples, the object elementary translation $|\Delta\mathbf{x_0}|$ is 2 mm for object 1 and 1 mm for objects 2 and 3. The number of profiles is 43 (object 1), 108 (object 2) and 114 (object 3). Notice that often the profiles present

discontinuities. Each continuous part of it is a segment. It is represented in the figure by a polygonal line connecting its points.

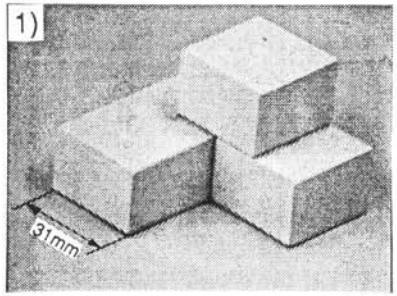

Fig. 5 Object 1

Fig. 6 3D data of object 1, illumination source view

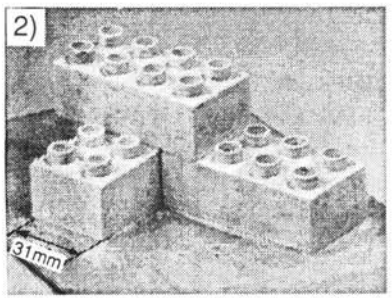

Fig. 7 Object 2

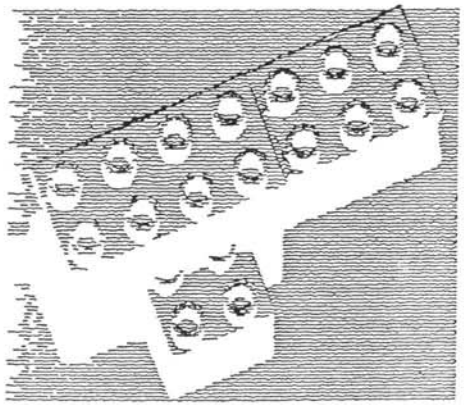

Fig. 8 3D data of object 2, camera view

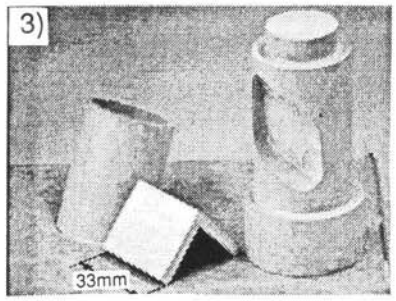

Fig. 9 Object 3

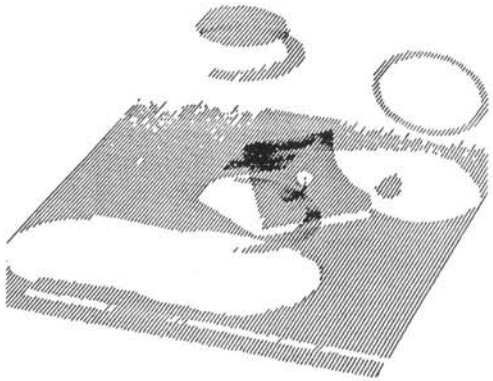

Fig. 10 3D data of object 3, arbitrary view

These examples also illustrate shadowing, which is a basic characteristic and limitation of the method. Shadowing designates the mechanism which makes certain parts of the object invisible. Basically, there are two types of shadows: shadows from the illumination and shadows from the camera. Only parts of the object which are visible both from the illumination source and from the camera are measured by the system. In figure 6, object 1 is viewed from the camera; in it, the white zones emphasize illumination shadows. Similarly, in figure 8, object 2 is viewed from the illumination source; the white zones emphasize camera shadows.

Overall measurement error

The system is subject to several errors such as calibration errors, device non-linearities, profile estimation errors, etc. In order to evaluate the overall system performance we compare a series of system measurements with a known reference and compile the differences as system errors.The error of a 3D measurement device is in general a vector in R^3. Because the present system is suited for surface measurements, we define here the error as the minimum distance between the measured point and the reference surface. Noting $(\bar{x}'\ \bar{y}'\ \bar{z}')$ the measured point and $ax' + by' + cz' + d = 0$ the reference surface, this distance is

$$e = \frac{|a\bar{x}' + b\bar{y}' + c\bar{z}' + d|}{\sqrt{a^2 + b^2 + c^2}}$$

Practically, the evaluation is performed with a special designed object that essentially consists of a large and precise square face 10 cm x 10 cm, whose plane equation is $x' + y' + \sqrt{2}\, z' = 0$. The corresponding measurement error is

$$e = (\bar{x}' + \bar{y}' + \sqrt{2}\,\bar{z}') / 2$$

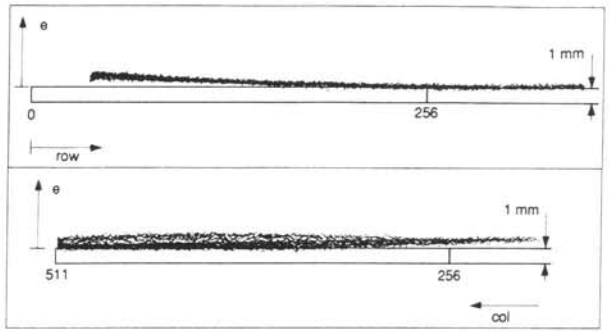

Fig. 11 Measurement error

Figure 11 is a plot of error e as a function of the image coordinates [i j]. The overall rms-error is σ_e = 0.39 mm. As a matter of comparison, the typical resolution due to quantizing in image space is 0.2 mm in this system and the residue errors of the $\underline{U_k}$ matrix calibration in the x-range are of ±0.38 mm.

Acknowledgement

This work was supported by the Swiss National Foundation for Scientific Research under project Nr. FN 2000-2256.

Conclusion

The 3D measurement system presented in this paper is a simple and inexpensive measurement system for laboratory use. It uses laser light stripe illumination, a tv-camera and a PC-based image analysis system. In the presented arrangement, both illumination and camera are fixed. The measurement proceeds profile by profile, with a successive displacement of the object at each profile

measurement. The useful measurement volume is 1 dm^3 and the number of profiles can vary from 1 to a few hundred. The rms-error is 0.39 mm.

References

[1] R. A. Jarvis, "A perspective on range finding techniques for computer vision", in IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-5, March 1983, pp. 122-139.

[2] D. H. Ballard and C.M. Brown, "Computer vision", Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

[3] H. Hügli and G. Maître, "Generation and use of color pseudo random sequences for coding structured light in active ranging", in SPIE Proc., vol. 1010, 1988, pp. .

[4] C.H. Chen and A. C. Kak, "Modeling and calibration of a structured light scanner for 3-D robot vision", in Proc. IEEE Int. Conf. on Robotics and Automation, 1987, pp. 807-815.