

Laser-Video Scanner Calibration without the use of a frame store

Bernard Tiddeman¹⁺²⁺³, Neil Duffy², Graham Rabey³ and Jamie Lokier³

¹Department of Surgery, University of Newcastle, Newcastle upon Tyne.

²Department of Computing and Electrical Engineering, Heriot-Watt University, Riccarton, Edinburgh.

³INTERFACE Research Unit, Stoke Mandeville Hospital, Aylesbury.

Laser-Video Scanner Calibration

Abstract

We present a new calibration method for a structured light surface scanner. The new method uses a model-based calibration, accounting for laser stripe and translation stage alignment and the internal and external camera parameters. The inclusion of non-linear radial distortion in the camera model gives an improvement in the calibration accuracy over a linear model. The method has been designed for laser-video scanner systems that use hardware-based measurement of the illuminated line position rather than a frame store. Standard camera-based calibration techniques cannot be used for such systems, however the new method is also applicable to systems that use software-based stripe segmentation.

1.0 Introduction

Structured light scanners using laser stripe light for triangulation-based range sensing are now commonly used for machine vision applications [1, 2, 3]. This paper is concerned with the calibration of a dual laser plus video scanning system that uses hardware-based measurement of the illuminated line position. Hardware-based stripe segmentation [4] (Figure 1) eliminates the storage and processing overheads imposed by capturing an image of each scan line using a large frame store. With some mechanical changes this type of scanner is capable of capturing changing surface profiles in real time over extended periods [5]. The disadvantage of not having a frame store is that standard model-based camera calibration methods involving the capture of a single image frame cannot be used. This has provided the motivation for a new calibration method.

The Interface scanner (Figure 2) has been developed as part of the Interface research project in collaboration with Heriot-Watt University. The system is installed within an analytic morphograph [6], a recording machine that produces standardised images of the face and head for research, diagnosis and treatment planning. The morphograph contains a cephalostat and marked axes that embody a fixed Cartesian reference frame. A requirement of the laser calibration system is that it meshes with the above geometric standard in addition to correctly rescaling the scanned data.

Calibration methods for laser-video scanners that use a frame store are most frequently based on a camera model with pinhole perspective [7, 8] with the possible inclusion of radial, tangential or other non-linear distortion terms [9, 10]. These methods can achieve very accurate camera calibration but they do not calibrate the laser stripe or the translation stage movement, which are frequently left to alignment by sight. Theodoracatos and Calkins [11] include laser stripe and camera calibration in a single model, but do not include translation stage alignment. This is because they only require calibration to the camera's internal reference frame and not to an externally defined reference frame.

Calibration methods such as ours that do not require a frame store are more varied and have been less widely adopted by the imaging community. The simplest such method is to measure all of the extrinsic parameters of the system, and rely on the manufacturer's data for the intrinsic parameters. This is inaccurate and time consuming and thus unsatisfactory.

A more promising approach is to scan a suitable 3D target and derive the necessary information from the known geometry and the imaged geometry. Manthey, Knapp and Lee [12] attempt to fit values to a model by scanning two geometrical objects - a sphere and a plane. They use a simple linear model that does not include camera perspective distortion or radial distortion. Trucco and Fisher [13] use a look-up table of the translations between the known and expected locations of points in a scanned calibration jig. This ensures that all systematic errors are accounted for but requires an expensive and accurately machined calibration jig and interpolation of the table's values.

2.0 Laser-Video scanner model.

The new method presented here is based on Tsai's camera model [9, 10], extended to include the full laser scanning system and modified for implementation on a system without a frame store

2.1 The Moving Camera Model

The extrinsic camera parameters represent the position and orientation of the camera in space. The initial position and orientation of the camera can be expressed as a 3 component rotation and a 3 component translation. The location of the camera at frame k is given by including an additional translation dependent on the translation step $\mathbf{v}$. Hence the transformation from co-ordinates in the external reference frame $\mathbf{x}$ to those in the camera reference frame $\mathbf{x}_c$ is given by

$$1) \quad \mathbf{x}_c = \mathbf{R} \cdot \mathbf{x} + \mathbf{T} + \mathbf{v}k$$

where $\mathbf{R}$ is the 3 x 3 rotation matrix, $\mathbf{T}$ is the 3 component translation vector and $\mathbf{v}$ is the 3 component velocity vector. The rotation matrix $\mathbf{R}$ is the result of rotating around each of the three axis (x, y, z) by the three angles (θ, ϕ, ψ) respectively. The result is the matrix

$$\mathbf{R} = \mathbf{R}_x \cdot \mathbf{R}_y \cdot \mathbf{R}_z$$

$$2) \quad \begin{bmatrix} \cos \phi \cos \psi & \cos \phi \sin \psi & -\sin \phi \\ -\sin \psi \cos \theta + \cos \psi \sin \phi \sin \theta & \cos \psi \cos \theta + \sin \psi \sin \phi \sin \theta & \cos \phi \sin \theta \\ \sin \psi \sin \theta + \cos \psi \sin \phi \cos \theta & -\cos \psi \sin \theta + \sin \psi \sin \phi \cos \theta & \cos \phi \cos \theta \end{bmatrix}$$

The intrinsic camera parameters represent the effective focal length, internal scaling and radial distortion coefficient of the camera. The pinhole perspective model will transform the 3D camera co-ordinate points $\mathbf{x}_c = (x_c, y_c, z_c)$ into undistorted sensor co-ordinates (X_u, Y_u) by

$$3) \quad X_u = f \frac{x_c}{z_c}$$

$$4) \quad Y_u = f \frac{y_c}{z_c}$$

where f is the effective focal length of the camera. Radial distortion terms are usually included only up to first order distortion, given by

$$5) \quad X_d = \frac{X_u}{(1 + \kappa_1 r_d^2)}$$

$$6) \quad Y_d = \frac{Y_u}{(1 + \kappa_1 r_d^2)}$$

where r_d is the distorted radius from the image centre given by

$$7) \quad r_d^2 = X_d^2 + Y_d^2$$

The measured pixel values (i, j) are then found by translation of the origin to the image centre (C_i, C_j) and by scaling the i axis to compensate for any differential scaling along each axis, i.e.

$$8) \quad i = X_d s_i + C_i$$

$$9) \quad j = Y_d + C_j$$

where s_i is the sensor's i scale uncertainty. The scaling in the j direction is calibrated by the effective focal length. The sensor element size does not effect the calibration as it is masked by the effective focal length f and the i scale uncertainty s_i . These variables will also include any scaling due to the hardware acquisition and internal processing on the PC.

2.2 Moving Laser Plane Model

We model the laser plane as an infinite plane of light with surface normal

$$10) \quad \mathbf{n} = (\sin \alpha, \cos \alpha \sin \beta, \cos \alpha \cos \beta)$$

which travels a distance $\mathbf{v}$ between frames, the same as that of the camera. This is equivalent to a stationary camera and laser plane, with the scanned object moving on a translation platform a distance $-\mathbf{v}$ between frames. This is represented by

$$11) \quad \mathbf{n} \cdot \mathbf{x} + n_0 = (\mathbf{n} \cdot \mathbf{v})k$$

where n_0 is a constant, dependent on the position of the plane at frame $k = 0$.

2.3 Complete System Model

The moving camera model and the moving laser plane model described above represent a complete description of the laser scanner system. They are combined by noting that all the points imaged by the camera lie in the laser plane (Figure 3). The model involves 17 individual parameters shown in Table 1.

3.0 Linear Calibration Method

If the radial distortion term is ignored, the equations representing the laser scanning system are given by

$$12) \quad i = s_i f \frac{r_1 x + r_2 y + r_3 z + T_x + v_x k}{r_7 x + r_8 y + r_9 z + T_z + v_z k} + C_i$$

$$13) \quad j = f \frac{r_4 x + r_5 y + r_6 z + T_y + v_y k}{r_7 x + r_8 y + r_9 z + T_z + v_z k} + C_j$$

$$14) \quad k = \frac{n_x x + n_y y + n_z z}{n_x v_x + n_y v_y + n_z v_z} + n_0$$

By absorbing the image centre (C_i, C_j) into the division expressions, and grouping variables, these can be written in matrix form as,

$$15) \quad \mathbf{M} \cdot \mathbf{x} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} i' \\ j' \\ k' \\ l' \end{bmatrix} \text{ and } \begin{bmatrix} i \\ j \\ k \end{bmatrix} = \begin{bmatrix} i'/l' \\ j'/l' \\ k' \end{bmatrix}$$

Because the depth map's vertical component k is not affected by perspective the system is non-linear with respect to the co-ordinate variables. The model is linear with respect to the model parameters m_{ij} and so linear methods can be used. The coefficients of the matrix $\mathbf{M}$ can be found using at least five non-coplanar points and solving the resulting set of 15 linear equations. We over-define our equations using 110 calibration points and solve them using singular value decomposition [14] to give a least squares fit.

To find the reference frame co-ordinate, (x, y, z) from a measured point (i, j, k) the inverse transform is applied. This is given by,

$$16) \ x = \frac{(1 - n_{43}k)(n_{11}i + n_{12}j + n_{14})}{(n_{41}i + n_{42}j + n_{44})} + n_{13}k$$

$$17) \ y = \frac{(1 - n_{43}k)(n_{21}i + n_{22}j + n_{24})}{(n_{41}i + n_{42}j + n_{44})} + n_{23}k$$

$$18) \ z = \frac{(1 - n_{43}k)(n_{31}i + n_{32}j + n_{34})}{(n_{41}i + n_{42}j + n_{44})} + n_{33}k$$

where $\mathbf{N} = [n_{ij}] = \mathbf{M}^{-1}$. Here the system non-linearity is even more clearly demonstrated.

4.0 Non-linear Calibration Method

Including radial distortion in the model equations produces a set of non-linear simultaneous equations. These can be solved iteratively using a standard technique such as the Levenberg-Marquardt algorithm [14]. This tries to minimise an error function, χ^2 by searching the solution space with a combination of the inverse Hessian and the steepest descent methods. The algorithm below finds the three component error χ given an uncalibrated point and its expected reference frame co-ordinate.

ERROR FUNCTION

INPUTS

Model parameters: $\mathbf{n}, \mathbf{v}, s_i, f, \kappa_1, \mathbf{R}, \mathbf{T}, C_i, C_j$
 Uncalibrated co-ordinate: (i, j, k)
 Reference frame co-ordinate: $\mathbf{x} = (x, y, z)$

VARIABLES

Scalar $I, J, K, X_u, Y_u, R_d, R_u, X_d, Y_d$
 Vector $\mathbf{x}_c, \chi$

BEGIN

Calculate the frame number K

$$K = \frac{\mathbf{n} \cdot \mathbf{x}}{\mathbf{n} \cdot \mathbf{v}} + n_0$$

Calculate the camera co-ordinates $\mathbf{x}_c$

$$\mathbf{x}_c = \mathbf{R} \cdot \mathbf{x} + \mathbf{T} + \mathbf{v}K$$

Calculate the undistorted sensor co-ordinates

$$X_u = f \frac{x_c}{y_c}, \quad Y_u = f \frac{y_c}{y_c}$$

Calculate the undistorted and distorted radius

$$R_u = \sqrt{X_u^2 + Y_u^2}, \quad R_d = \text{Solve_cubic}(R_u = R_d(1 + \kappa_1 R_d^2))$$

Calculate the distorted sensor co-ordinates

$$X_d = \frac{R_d}{R_u} X_u, \quad Y_d = \frac{R_d}{R_u} Y_u$$

Calculate the image co-ordinates

$$I = X_d s_i + C_i, \quad J = Y_d + C_j$$

Calculate the error vector

$$\chi = \begin{bmatrix} I \\ J \\ K \end{bmatrix} - \begin{bmatrix} i \\ j \\ k \end{bmatrix}$$

OUTPUT
Error vector χ

END

Here R_u and R_d are the undistorted and distorted radii from the image centre (C_i, C_j) , (X_u, Y_u) and (X_d, Y_d) are the undistorted and distorted image co-ordinates, (I, J, K) is the calculated depth map co-ordinate and (i, j, k) is the corresponding measured co-ordinate. The cubic polynomial is solved using Cardan's method, selecting the smallest positive real root.

The Levenberg-Marquardt method will only converge to a local minimum and so a reasonable initial guess is required. In order to make such a guess we use certain assumptions about the scanner as follows.

- 1) The radial distortion term is small and so can be ignored in a first guess.
- 2) The image centre is close to the centre of radial distortion.

We use the linear model parameters along with the approximate image centre to establish a first guess.

As with the linear method, the inverse transform can be defined using a matrix transform, provided the radial distortion term is removed first, using

$$19) X_u = X_d (1 + \kappa_1 r_d^2)$$

$$20) Y_u = Y_d (1 + \kappa_1 r_d^2)$$

$$\text{where } r_d^2 = X_d^2 + Y_d^2, X_d = (i - C_i)/s \text{ and } Y_d = (j - C_j).$$

The forward transform matrix is then,

$$21) \mathbf{M} = \begin{bmatrix} r_1 + v_x n_x & r_2 + v_x n_y & r_3 + v_x n_z & T_x + v_x n_0 \\ r_4 + v_y n_x & r_5 + v_y n_y & r_6 + v_y n_z & T_y + v_y n_0 \\ n_x & n_y & n_z & n_0 \\ r_7 + v_z n_x & r_8 + v_z n_y & r_9 + v_z n_z & T_z + v_z n_0 \end{bmatrix}$$

$$\text{where } \mathbf{n}' = (\sin \alpha, \cos \alpha \sin \beta, \cos \alpha \cos \beta) \text{ and } \mathbf{n} = \frac{\mathbf{n}'}{\mathbf{n}' \cdot \mathbf{v}}.$$

The linear inverse transform is then applied to the co-ordinates $(X_u/f, Y_u/f, k)$, using the matrix $\mathbf{N} = \mathbf{M}^{-1}$.

5.0 Calibration Data Collection

To fit parameters to the models described above we use points extracted from two scans of a calibration target (Figure 4). The target consists of 55 circular white dots on a black background constructed by writing a raw postscript file and printing on a phototypesetter to an RMS accuracy of 0.09mm. By adjusting

the threshold on the video signal intensity, used to segment the laser stripe, only the white dots of the target are detected. The calibration range maps therefore have the appearance of a number of approximately circular discs floating in space. For robust calibration, the calibration points range across as much of the scanned volume as possible. The first target is placed in the morphograph's cephalostat, with all points lying in the plane $x=5\text{mm}$, and the second is placed on the back of the morphograph's diffuser in the plane $x=169\text{mm}$. The use of 110 calibration points reduces the effects of random errors in the detected centroids (i.e. due to system vibration and quantisation errors) reducing the need for highly accurate detection of any single point. The dots are detected automatically, by thresholding followed by a nearest neighbour connect routine. The centroids are found by averaging the 3D values over each segmented dot.

6.0 Results

The following results demonstrate that reasonable accuracy can be achieved with a one step calibration technique. Results are given for the inverse transform (transforming measured depth-map co-ordinates into real world points) in Table 2 and for the forward transform (transforming the expected real world location of a point into the measured co-ordinate system) in Table 3. The errors in the inverse transform are affected by the distance of the point from the camera (approximately 1m,) and the relatively coarse quantisation (approx. 1mm^3 per voxel.) These results could be improved by using sub-pixel laser stripe segmentation, a higher resolution CCD camera or smaller translation stage steps. The errors in the forward transform are not affected by the dimensions of the scanner or the voxel size and so give an indication of the true accuracy. This method predicts the depth-map co-ordinate to within ± 1 voxel for the vast majority of points and so may be regarded as close to the theoretical limit.

To measure the error away from the calibration points we used a multi-plane calibration test jig (Figure 5,) machined on a milling machine accurate to 0.01mm. The perpendicular distance of each calibrated point $p = (x, y, z)$ from the plane $x = ay+bz+c$ was calculated for almost every point in each of the 6 planes. The planes span approximately 60% of the depthmap area, about 50, 000 points. The distribution of errors for the planes (Figure 6) obeys a Gaussian model so we conclude that quantisation errors and vibrations cause the large maximum error on some of the planes. The results for the non-linear calibration method are shown in tables 4 and 5, giving the results for the inverse and forward transforms respectively.

7.0 Conclusions

We have demonstrated a new one-stage calibration technique for calibrating a laser-video surface scanner. Although the technique is applicable to scanners with a frame store, it has been designed to solve the calibration problem for a scanner without one. The position and orientation of the camera and the laser stripe, the translation stage movement and the internal camera parameters are all included in a single model.

The results show that the method gives an accuracy approaching the limit of the scanner. For greater accuracy the method could be extended to include more non-linear distortions such as tangential or barrel distortion.

Acknowledgements

We are indebted to the Jimmy Savile Stoke Mandeville Hospital Trust for the funding of this work and to Dr Roger Bodley and other colleagues in the Radiology Department at Stoke Mandeville Hospital Aylesbury, Buckinghamshire for their support of the Interface Research Unit.

References

1. Besl, P.J.: 'Active optical range imaging sensors,' *Machine Vision Appl.*, 1988, **1**, (2), pp.127-152.
2. Strand, T.S.: 'Optical three-dimensional sensing for machine vision,' *Opt. Eng.*, 1985, **24**, (1), pp. 33-40.
3. Jarvis, R.A.: 'A perspective on range finding techniques for computer vision,' *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1983, **5**, pp.122-139.
4. Duffy, N.D. and Yau, J.F.S.: 'Facial image reconstruction and manipulation from measurements obtained using a structured lighting technique,' *Pattern Recognition Letters*, 1988, **7**, pp. 239-243.
5. Duffy, N.D., Drummond, G.D., McGowan, S. and Dessesard P.: 'A real time optical scanning system for measurement of chest volume changes during anaesthesia,' *SPIE Biostereometric Technology and Applications*, 1990, **1380**, pp. 46-52.
6. Rabey, G.P.: 'Current principles of morphanalysis and their implications in oral surgical practice,' *British Journal of Oral Surgery*, 1977-78, **15**, pp. 97-109.
7. Vernon, D.: 'Machine vision: automated visual inspection and robot vision,' (Prentice Hall, UK, 1991).
8. Arridge, S., Moss, J.P., Linney, A.D. and James, D.R.: 'Three dimensional digitisation of the face and skull,' *J. Maxillofacial Surgery*, 1985, **13**, pp. 136-142.
9. Tsai, R.Y.: 'An efficient and accurate camera calibration technique,' *Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition*, 1986, Miami Beach, Florida, pp. 364-374.
10. Tsai, R.Y.: 'A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses', *IEEE J. Robotics Autom.*, 1987, RA-3, (4), pp. 323-344.

11. Theodoracatos, V.E. and Calkins, D.E.: 'A 3-D vision system model for automatic object surface sensing', *International Journal of Computer Vision*, 1993, **11**, 1, pp.75-79.
12. Manthey, D.W., Knapp, K.N. II and Lee, D.: 'Calibration of a laser range-finding co-ordinate-measuring machine', *Optical Engineering* 33(10), October 1994, pp. 3372-3380.
13. Trucco, E. and Fisher, R.B., 'Acquisition of consistent range data using local calibration', *Proc. IEEE International Conference on Robotics and Automation*, 1994, San Diego California, pp. 3410-3415.
14. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: 'Numerical Recipes in C: The Art of Scientific Computing', (Cambridge University Press, 1992), 2nd Ed.

Tables

Table 1.

Variable	Description
f	Effective focal length
S_i	image i scale uncertainty
κ_1	radial distortion coefficient
$\mathbf{v}$	3 component translation stage velocity vector
(α, β)	2 laser plane angles
n_0	laser plane constant
$\mathbf{T}$	3 component camera initial translation vector
(θ, ϕ, ψ)	3 camera rotation angles
(C_i, C_j)	2 component image centre (centre of radial distortion)

Table 2.

	Linear Method (x, y, z) mm	Non-linear Method (x, y, z) mm
RMS error	(0.549, 0.245, 0.0897)	(0.306, 0.0595, 0.00814)
Max. error	1.38489	0.98138

.

Table 3.

	Linear Method (i, j, k) voxels	Non-linear Method (i, j, k) voxels
RMS error	(1.382, 0.377, 0.0909)	(1.318, 0.374, 0.0908)
Max. error	3.4289	2.4275

Table 4.

Plane	Mean error	R.M.S. error	Max. error
$x = 0$	-0.050	0.388	1.417
$x = 125-z$	-0.005	0.291	0.972
$x = 125-y$	-0.117	0.241	0.850
$x = (y+z)/2$	0.007	0.447	1.268
$x = z+125$	0.081	0.265	0.897
$x = y+125$	0.011	0.319	1.082

.

Table 5.

Plane	Mean error	R.M.S. error	Max. error
$x = 0$	0.091	0.612	2.501
$x = 125-z$	0.014	0.465	1.512
$x = 125-y$	0.237	0.489	1.700
$x = (y+z)/2$	0.044	0.791	2.337
$x = z+125$	0.127	0.426	1.402
$x = y+125$	0.015	0.535	1.780

Table Captions

Table 1. The 17 parameters included in the scanner model.

Table 2. The measured error for the inverse transform using the calibration points.

Table 3. The measured error for the forward transform using the calibration points.

Table 4. Errors on planar surfaces for non-linear inverse transform in mm.

Table 5. Errors on planar surfaces for non-linear forward transform in voxels.

Figure Captions

Figure 1. The Interface Laser Video Scanner Architecture.

Figure 2. The video processing hardware measures the time between the line sync pulse and the centre of the laser stripe pulse in real time, eliminating the need for a frame store.

Figure 3. A diagram of the Scanner Model.

Figure 4. The calibration target as captured by the laser scanner. The detected centroids are marked with crosses.

Figure 5. A diagram of the multi-plane calibration test object.

Figure 6. A histogram showing the distribution of errors (in mm with 0.01 mm bins) over the six planes of figure 5.

Figures

Figure 1.

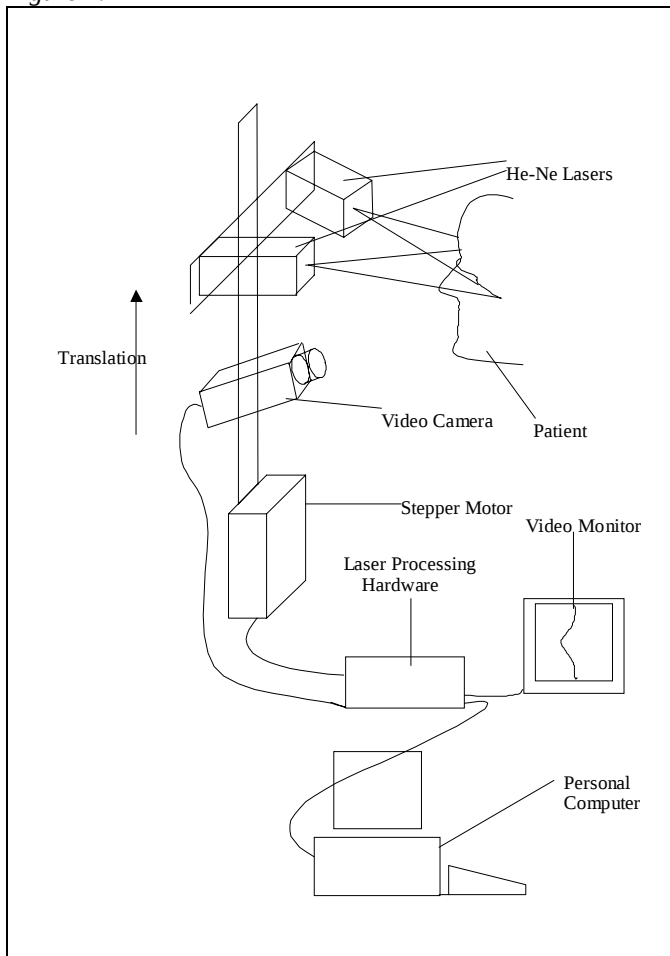


Figure 2.

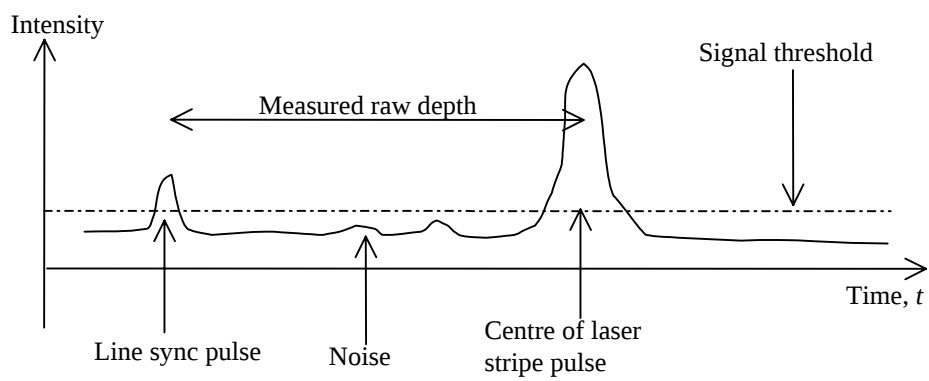


Figure 3.

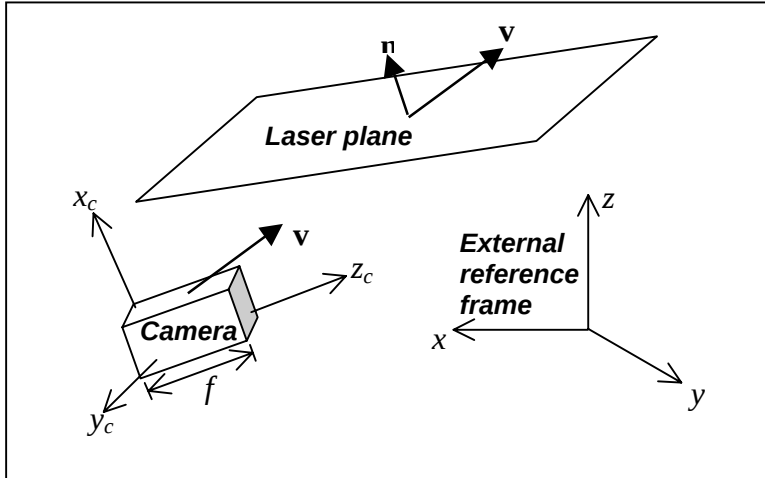


Figure 4.

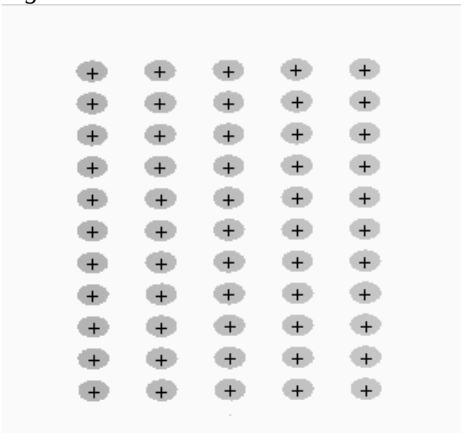


Figure 5.

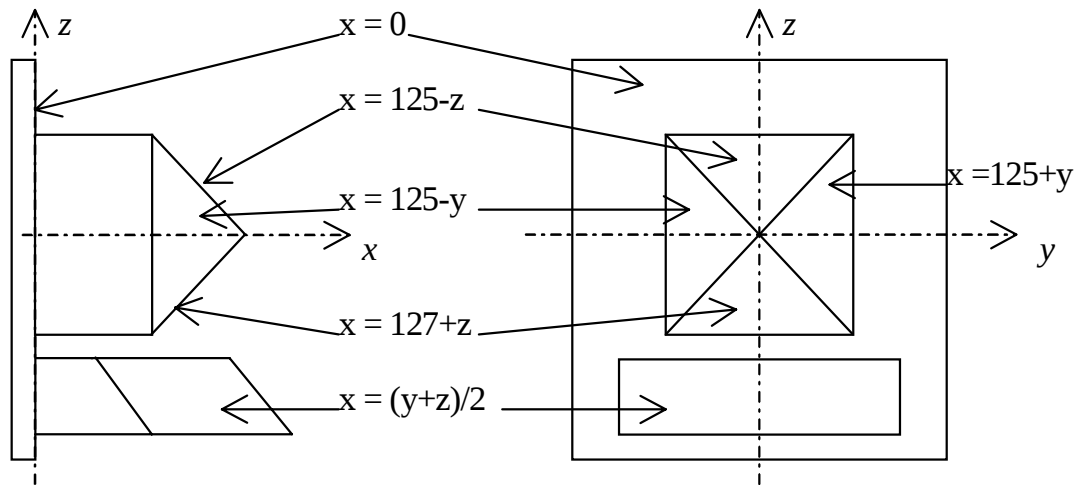


Figure 6.

