

Engineering

Industrial & Management Engineering fields

Okayama University

Year 1991

High-speed processing for obtaining
three-dimensional distance image and its
application

Yutaka Tanaka
Okayama University

Hideki Tsukaoka
Okayama University

Hidetoshi Takeda
Okayama University

Kazuo Honda
Okayama University

Takaaki Sarai
Okayama University

This paper is posted at eScholarship@OUDIR : Okayama University Digital Information Repository.

<http://escholarship.lib.okayama-u.ac.jp/industrial-engineering/106>

HIGH-SPEED PROCESSING FOR OBTAINING THREE-DIMENSIONAL DISTANCE IMAGE AND ITS APPLICATION

Yutaka TANAKA, Hideki TSUKAOKA, Hidetoshi TAKEDA,
Kazuo HONDA and Takaaki SARAI

Department of Mechanical Engineering, Faculty of Engineering,
Okayama University, Naka 3-1-1, Tsushima, Okayama 700, JAPAN

ABSTRACT: A high-speed method of 3-D distance acquisition based on the triangulation principle is presented. This method uses conventional devices such as a CCD camera, a laser emitting semiconductor, and scanning mirrors; however, new circuits have been developed for detecting the position of spot-image on the CCD. This development enabled the high speed measurement and reduced the cost of the apparatus.

Experiments showed that the apparatus and the method gave the practical measuring accuracy and speed, and it was found that the system is useful for the method of image recognition. This method can easily display the stereoscopic image and cross-sectional figure of the object body. The method of real time processing has been also developed with the view to applying the device to the range finder for robots and blind men.

1 INTRODUCTION

There are many methods for obtaining the scene; however, there is quite a few [1,2] for obtaining a distance picture and recognizing object image. New types of intelligent robots have been developing in many institutions. The most important subject for promoting autonomous control of a robot and for awarding intelligence to him is to awarding the functions of recognizing the environments where he is situated. This means that he is granted appropriate eyes. Many studies have been undertaken on this standpoint; however, high-speed or real-time recognition device which can be used as practical visual organ has not yet been developed.

If the apparatus useful for recognizing 3-D image is developed, many applications are anticipated such as utilization as the sensors of 3-D profiling machine, obstruction detectors for automobiles and alternatives to guide dogs for blind people.

In the methods for acquiring 3-D information, there is the one which obtains the distance by detecting the parallax between the images of two TV cameras [3]. It is not easy in this method to find the corresponding pixels in the binocular images obtained for the same point on the object surface. Although it is possible to detect the corner points by some

softwares, it is not easy to detect the curved surface; speaking of the extreme case, it is impossible to decide the position of every pixels in the wall image of which whole the field of view is painted in quite the same color.

As a method to get rid of such difficulties, the method utilizing the light source with some patterns called textures has been considered [4]; however, it is difficult to use this as a practical one because long computation time is needed to differentiate one pattern from the others.

Instead of using two cameras, new methods [5] have been contrived in which one side camera is replaced by a light projector such as a semiconductor laser or a LED. In this case the other one can be replaced by a PSD (Position Sensing Detectors). The light beam is then projected onto the object and the brightest spot is detected by a CCD (Charge Coupled Device) or a PSD installed in the camera. Three dimensional distance image is obtainable by sweeping the light beam on all over the object surface. These methods are the ones rich in practicality although they are based on the principle of triangulation [6] which has been used from the ancient age.

The former methods which use a light source and a PSD or a CCD have some difficulties when the positions of the bright spot projected onto the body are detected; the PSD method has the limitation that the measuring distance is too short or is severely restricted because of the low sensitivity and the conventional CCD method has the disadvantage of low sampling rate. One point measurement requires 1/30 sec.(a frame period) in average. The 3-D scanning sensor developed in this paper uses a laser light source and a CCD camera; however, in consideration of the circumstances mentioned hitherto, we aim in the first place to develop a new method of signal processing for detecting the light spot.

Data processing and recognition in the existing CCD method use digital computers and give accurate information [7-9]. However, the throughput time from data sampling to the distance decision is too long to recognize the object in real time. In this paper, making use of the former CCD method, new techniques are presented in which efforts have been made for improving measuring speed, cost, convenience in use and applicability. The method of real time processing and the necessary circuits for computing 3-D coordinates analogously are described.

2 EXPERIMENTAL APPARATUS

2.1 Characteristics of the methods

The principle of triangulation has been used from the ancient age; therefore the most important objective, in the modern sense, for putting the methods into the practical level is at obtaining the distance image at a higher speed and a cost as cheap as possible. The distance image should then be comparable to the scene obtained by conventional cameras and 3-D distances for every pixels should be given. The characteristics of the present method are as follows:

(1) In this method the whole scene in the eye field is swept by laser light and is taken by a CCD camera. The position of the bright spot occurred when the laser beam is projected onto the object is detected by the circuits at a high speed. The output signal of the circuit is sent to a personal computer, where the distance image is obtained through the digital processing. As a laser source, both the spot and slit types can be used. Especially, the use of the spot source enables the random access of the projection direction and is useful in that the 3-D distance is acquired while confirming the spot position simultaneously with observing the scene with operator's eyes.

(2) With the intention of applying the method to the eyes of robots and blind men, the circuit has been produced which processes the calculation in real time and gives the analog output proportional to the distance.

Figure 1 shows the hardware configuration for acquiring the distance image on the basis of digital processing. The apparatus is composed of a laser light source, two galvanometers for scanning the light, a CCD camera, and two digital processing boards (modules for detecting the light spot and for scanning the laser beam).

2.2 Acquisition of 3-D position

Taking as an example the case when the spot light source (the wave length of 670 nm and the power of 4 mW) is used, the method of acquiring 3-D position is firstly explained in the same style as the references of [6] and [7]. The intersection between the light beam and the object surface is hereafter called the light spot. Fig.2 shows the space coordinates and the geometrical relationship among the laser source, the light spot, and the camera. In Fig.2, the origin of the coordinates is put on the focal point O of the camera lens. The Cartesian coordinate system $O[X, Y, Z]$ is considered for the arrangement of the camera, where the X -axis is taken as the right-hand-side direction, Y -axis as the direction vertical to the text, and Z -axis as the camera axis. Similarly, the system $O_L[X_L, Y_L, Z_L]$ for the light source is considered, whose origin O_L is situated at the point a distance (D_X, D_Y, D_Z) apart from the camera origin O . In general, it is regarded that the directions of $[X, Y, Z]$ do not coincide with those of $[X_L, Y_L, Z_L]$ because of mis-alignment of either system. The transformation matrix from the laser system to the camera system is then expressed by T_r . The horizontal angle of deviation of the light beam from Z_L -axis is designated as α and the vertical one as β . The horizontal and vertical distances

between the spot image and the reference point of the CCD-image plane are respectively designated as x and y , and the focal length of a lens as F . Then the coordinate (X_L, Y_L, Z_L) of the light spot on the object is expressed in the coordinate system $O[X, Y, Z]$ as follows:

$$[X \ Y \ Z] = [X_L \ Y_L \ Z_L] \cdot T_r + [D_X \ D_Y \ D_Z] \quad (1)$$

The relationships among the coordinate (X_L, Y_L, Z_L) of the light spot on the object surface and the deviation angles α and β of the laser beam are expressed as

$$X_L = -Z_L \tan \alpha, \quad Y_L = Z_L \tan \beta \quad (2)$$

and the relationships among the image coordinate (x, y) , the actual coordinate (X, Y, Z) , and the focus length F are

$$\frac{F}{Z} = \frac{x}{X} = \frac{y}{Y} \quad (3)$$

Owing to the mis-alignment of the axes of laser beam, it is assumed that the former coordinate system coincides with the latter one of $O_L[X_L, Y_L, Z_L]$ when the former system is rotated successively in the anti-clockwise directions: the angle θ_Z around the axis- Z , θ_Y around Y , and θ_X around X , then the transformation matrix T_r is expressed as

$$T_r = \begin{bmatrix} c_Y c_Z & c_Y s_Z & s_Y \\ s_X s_Y c_Z - c_X s_Z & s_X s_Y s_Z + c_X c_Z & -s_X c_Y \\ -c_X s_Y c_Z - s_X s_Z & -c_X s_Y s_Z + s_X c_Z & c_X c_Y \end{bmatrix} \quad (4)$$

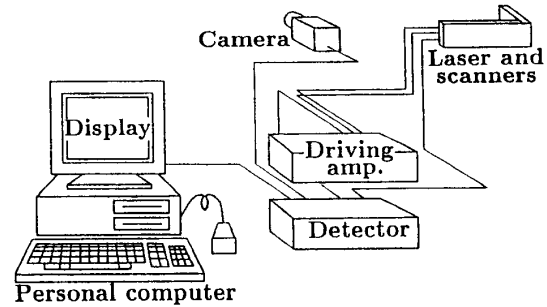


Fig.1 Schematic of the hardware system

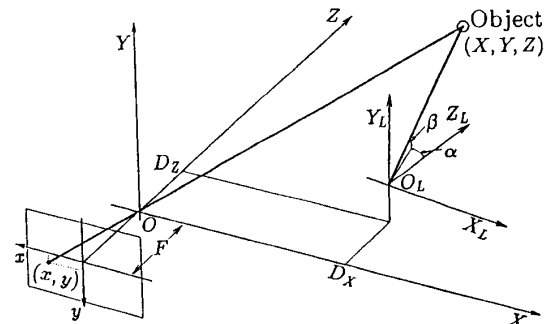


Fig.2 Coordinate system for detecting the position of the laser spot

From these equations the 3-D position (X, Y, Z) is summarized as

$$X = x B/E, \quad Y = y B/E, \quad Z = F B/E \quad (5)$$

$$B = D_X + D_Z \cdot T_A, \quad E = x + F \cdot T_A \quad (6)$$

In these equations, the following symbols are used.

$$T_A = \frac{T_\alpha \cdot c_Y c_Z - T_\beta (s_X s_Y c_Z - c_X s_Z) + c_X s_Y c_Z + s_X s_Z}{c_X c_Y - T_\alpha \cdot s_Y - T_\beta \cdot s_X c_Y} \quad (7)$$

$$\begin{aligned} c_X &= \cos \theta_X, c_Y = \cos \theta_Y, c_Z = \cos \theta_Z \\ s_X &= \sin \theta_X, s_Y = \sin \theta_Y, s_Z = \sin \theta_Z \\ T_\alpha &= \tan \alpha, T_\beta = \tan \beta \end{aligned} \quad (8)$$

It can be understood that, when there is a mis-alignment, the complicated calculation in the transformation matrix T_r is required for compensating the mis-alignment. If the alignment is accurate, the matrix T_r can be simplified into the unit matrix, therefore, T_A becomes equal to $\tan \alpha$. Even if the complete adjustment is difficult, the axis alignment should then be made in such a way that the term of T_β in Eq.(7) at least vanishes. This means that the angle θ_Z should be adjusted to be zero; this adjustment is greatly effective in reducing both the computation time and the measuring errors in absolute values. However, the use of spot light source diminishes such a difficulty in adjustment.

2.3 Standard video signal

For the detection of the light spot in the present method, no restriction is put on the kinds of CCD, storage tube, color, or monochromatic types if only the video signal is produced. The case with use of the CCD camera is hereafter described.

The video signal for a scene is generated pulsatively after the vertical (VSS) and horizontal synchronizing signals (HSS). This scanning time sequence is defined by the standard of NTSC (National Television System Committee in Japan). In this standard, the interlaced scanning method is employed. One frame of image plane is divided into two parts of odd and even fields, which contributes to the reduction in flicker of TV image. The gaps among the former scanning lines are traced by the succeeding lines, which fact is effective in increasing the apparent resolution. Fig.3 shows the scanning lines, the inner frame which corresponds to the visible range of TV-scope and Fig.4 the composite video signal output in which the synchronizing signal and video signal are combined.

The interlaced scanning employed in the standard video signal generates 525 scanning lines (480 visible). One field time is about a sixtieth sec. and one frame time is about a thirtieth sec. As can be understood from Fig.4, the composite synchronizing signal is very complicated, especially, there are many notches called equivalent pulses before and after the VSS. These pulses serve to stabilize the interlaced scanings. There are also invisible parts, which are called the blanking, horizontal front-porch, and back-porch periods. Fig.5 shows the waveform of the video signal for one horizontal scanning period in which a laser light spot is detected. In the signal, the HSS signal and the fly-back period are followed by the

signal of the scene containing the part of light spot. If this part is brighter, therefore, the signal is greater than that of the background image, the signal discrimination of the bright spot from the background is possible.

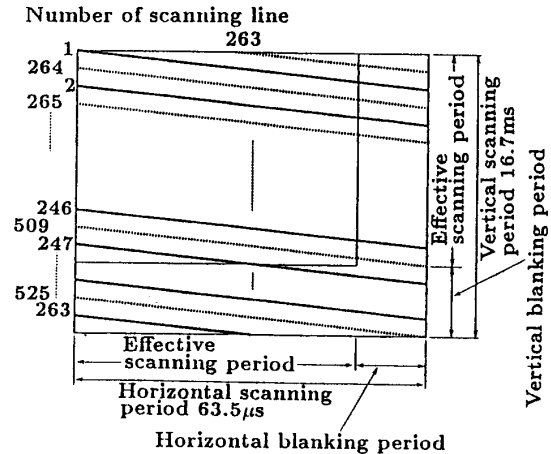


Fig.3 Scanning lines defined by NTSC

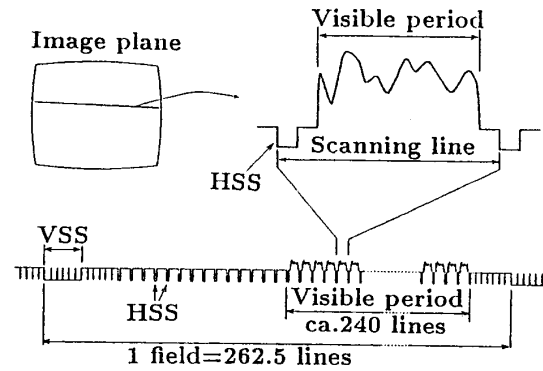


Fig.4 Composite video signal

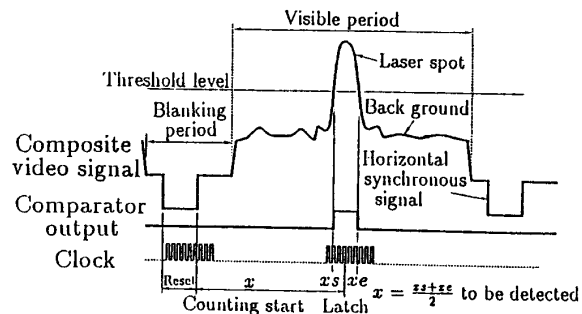


Fig.5 Detection of the brightest spot position

2.4 Signal processing

Figure 6 shows the block diagram of the circuit. This circuit is consisted of the board for detecting the brightest spot, the one of generating the signal for scanning the laser beam over the scene including the object, and a digital interfacing board to the personal computer where the calculation of determining the 3-D position and image processing are done.

Any of the spot and the slit types can be employed as a light source; however, the light beam in the spot type is swung in such a way that the beam image traverses the HSS lines in Fig.3 within one frame period. The CCD-camera is so arranged in such a way that the angle between the image stripe of the light beam and the HSS line becomes the right angles.

Figure 7 shows the sweeps of the spot light in the lateral(X or α) and longitudinal (Y or β) direction and the time elapse of the spot movement in the automatic scanning mode, then the spot movement is expressed in the relationship between lateral position and time. The sweep is made from top toward bottom in Y -direction and is returned again to the next starting point along the dashed line within the VSS period. The scanning mirrors are driven synchronously with the HSS and VSS by the wave generator according to the signals from the computer.

The light can be also randomly scanned with the use of a mouse. The scanning galvanometer with the aluminum-coated mirror of 10mm by 10 mm in size has the characteristics of the maximum swing angle of 40 deg. and the flat response frequency up to 300 Hz. The CCD sensor installed in the camera unit (CS3430B) has the pixels of 487(V) \times 764(H) and size of 6.6(V)mm \times 8.8(H)mm. It should be noted in the use of CCD camera that the electric charge in proportion to the light quantity received within the period of 1/30 sec. is discharged as the output of video signal in the next period. Consequently we must consider that there is inevitably the delay time of about 1/30 sec. in the video signal.

In the case of automatic scanning of spot light; however, the period while the spot image is projected onto the CCD module is very short compared to those for the background scene. If the stripe of the spot image is not detected, it is needed to increase the brightness of the spot or darken the background scene; otherwise to employ an optical filter. If only the spot can be detected by the CCD camera, there is no limit in detecting distance.

The signal flow of the circuit in Fig.6 are as follows: Firstly, the composite video signal of the CCD camera is fed to the Y/C (video/color signal) separator, where the VSS and HSS are discriminated and the signal of triggering the counter for measuring the brightest position is generated. Secondly, the video signal is passed through the comparator, where the spot signal is separated from the signals of the background scene under an appropriate threshold level.

The horizontal x -position of the light spot is detected by counting the number of pulses from the end of the HSS to the brightest image part of the signal using a counter. On the other hand, the y -position of this spot is given by counting the number of scanning lines from the VSS. The digital values of x - y position are sent to the computer through the digital

interfacing board. The clock frequency is 16 MHz, which is as much as 1024 times the frequency of HSS; therefore, giving 11-bit value for x position and maximum count value of 2030 within the horizontal scanning period. The employment of such a high clock frequency serves not only to increase the exactness of the spot detection but also to reduce a following fear. The CCD module, because it is consisted of a number of pixels, has unavoidably some defective pixels, thereby causing white and black noise spots. In a commercial handy camera, this defect is compensated with a help of a complementary ROM. However, the compensation is made in such a way as only to reduce apparent noises on the monitor image, and it can not be said that the compensation is a complete one. In addition the video signal may contain the noise owing to some outside circumstances. To avoid mis-counting caused by such noises, this counter is so devised as to exclude the spot image whose width is less than that of four clock pulses, and to detect the middle position of the spot.

The mirror scanning in the X - or α -direction can be made up to 2030 times but this has been limited to 500 times within the resolution of the CCD module (764 pixels in the horizontal direction). The maximum count of y position is of 10-bit number limited by the number of the video scanning lines. The circuit can also cope with the interlacing which sends out the signals of the even and odd fields alternately. As a result, this processing has enabled to detect the positions as much as the number of scanning lines within one frame period. Detection of 480-point (interlace) or 240-point (non-interlace) distances per one frame time has been achieved.

At the same time, the spot signal is synthesized again with the composite synchronizing signal and its output is observed on a TV-monitor for recognizing whether the threshold level is appropriate or not.

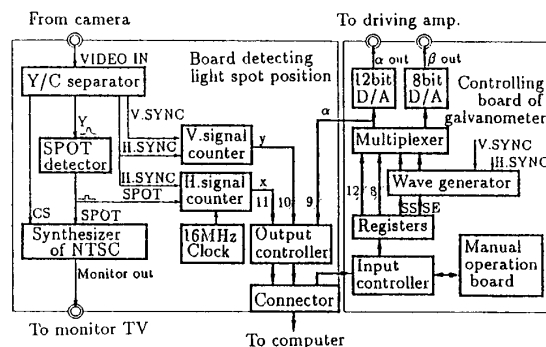


Fig.6 Block diagram of the circuits

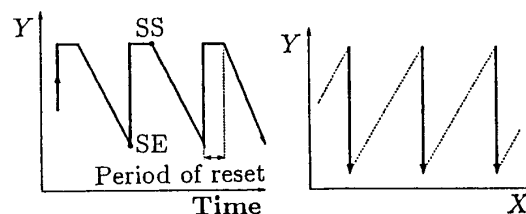


Fig.7 Auto-sweeps of a laser beam

The values x and y are sent to a personal computer (PC-9801DA), where the software written in C-language converts them into distance information on the basis of the triangulation principle with the swing angles α and β of a laser beam. When the sampling of 500 times is made at different angle of α , it takes about 17 sec. to get 720,000-data (3-D position information for 500×480 pixels) for one distance picture. Besides, it takes about 6 min. for calculating and recording the distance and for displaying the distance image. The most part of this time is spent to recording the distance data on a hard disk, because the accessible real memory space is restricted within 1 MByte in this type of microprocessor.

3 MEASUREMENT

3.1 Measuring accuracy

The view field and the resolution of the present method are decided by the size and pixel numbers of the CCD used, the focal length F , and the distance Z from the camera to the object. Now, let us estimate the view field and the resolution on the plane at $Z = 1$ m on the assumption that the effective CCD-size is $8 \text{ mm(H)} \times 6 \text{ mm(V)}$, the effective pixel numbers are $750(\text{H}) \times 480(\text{V})$, and the lens of $F = 12.5$ mm is used.

As can be easily understood from Fig.2, the view field is expressed by the CCD size \times the distance ratio Z/F . Therefore the use of the lens of $F = 12.5$ mm gives the view field of $600(\text{H})\text{mm} \times 480(\text{V})\text{mm}$. Since the resolution is given by the expression of CCD-size/pixel number $\times Z/F$, the resolution at $Z = 1$ m is estimated as about 1 mm.

The measurement of the distance to the wall situated 1 m ahead of the camera gave the result that the standard deviations of the x -values measured for the 10×10 points on the wall were 1.3 digits in average and that the positional dispersions of these deviations were little enough. Since the max-

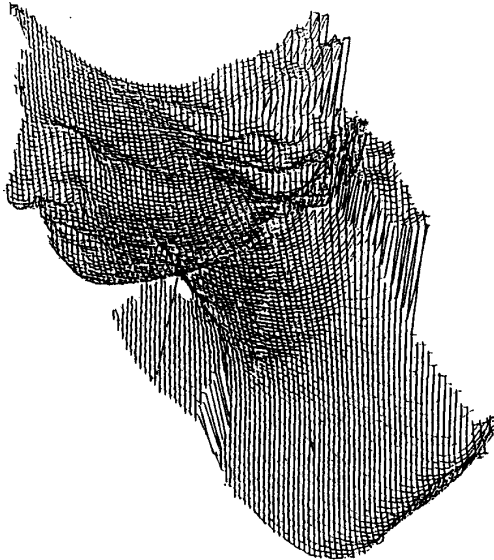


Fig.8 Distance net representation of the gypsum statue

imum count in x is 2030 and the maximum horizontal pixel number is 750, the standard deviation of 1.3 digits corresponds to that of 0.4 pixel. This deviation value corresponds also to the distance error of $\sigma_z = 0.4$ mm.

3.2 Example of application

The software system converts the 240000-set of information for x, y, α , into the data of 3-D position and displays

1. digital values of the projection angles (α and β) and of the position (x and y) of the spot image on the CCD,
2. distance image in color of 16-chromas and in monochromatic of poly-grey densities,
3. the distance at an arbitral pixel position,
4. cross-sectional curve on the straight line from an arbitral pixel to the other one,
5. distance net representation of the object,
6. extraction of the contour of the object and discrimination of the object.

As an example, Fig.8 shows the distance net representation of the well known gypsum statue and Fig.9 exhibits a printer-hardcopy of an image of a toy car displayed on the CRT display of the personal computer.

4 REAL TIME PROCESSING

As an application of the digital processing shown in the preceding sections, this section describes the real-time processing method of which the output signals of the circuit shown in Fig.6 are employed and are converted again into analog signals.

The digital values of distances x, y and projection angle α are converted into analog values in real-time; for instance, the analog distance z' of the object body measured from the camera center is analogously calculated from the following equation:

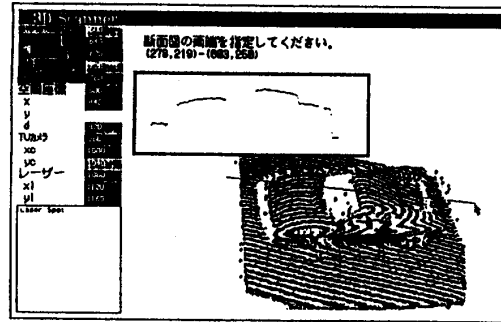


Fig.9 Distance image of a toy car

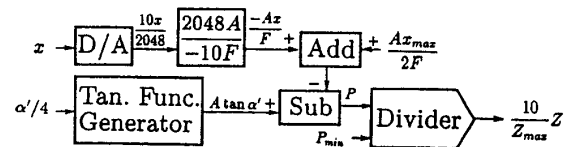


Fig.10 Block diagram of the circuits

$$z' = \frac{G}{x/F - x_{max}/(2F) + \tan \alpha'} \quad (9)$$

where, G is a proportional constant connecting above equation with the relationship (analog output $z' = 10Z/Z_{max}$), and the camera has been so arranged in such a way that the term corresponding to D_z in Eq.(6) becomes zero. In the measurement, a slit-type light emission semiconductor laser (wavelength of 870 nm, variable power up to 40 mW but usually it is used at 2 mW) and a monochromatic CCD camera were employed.

Figure 10 shows the block diagram of the main analog circuit. In this circuit the calculation necessary for obtaining the distance is carried out in the similar manner as digital ones. The stripe of light image obtained by one projection of the slit beam is then converted into a sequence of analog signal.

The symbol $10x/2048$ is the analog value which is D/A converted from x and the symbols A , P_{min} , X_{max} , and Z_{max} are the coefficients for suppressing the outputs within the dynamic range of 10 V.

The high-speed functional generating circuits such as a tangent generator and a divider, and an interface to the digital circuit have been newly developed. The accuracy and frequency characteristics of them were examined and whether they can cope with the high speed computation or not was evaluated. As total characteristics, it was confirmed that the computing board in which the above-mentioned circuits were combined gave the exact output in real-time, then the following application measurement was carried out.

Figure 11 shows an analog distance-record of the scene where a cone (15 cm in base diameter and 13 cm in height) is mounted on a box (17 cm in side length and 5 cm in height), thereby 15 distance lines under the different projection angles are drawn from the right-hand-side toward the left-hand-side direction. In the record, some ghost curves different from the actual scene are observed in the part on the left-hand-side contour of the cone. The parts like stages are ascribed to the shadow region where the laser stripes can not be detected by a CCD-camera. In the shadow region, the circuit holds the

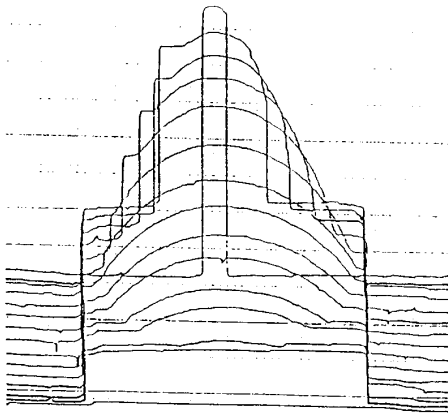


Fig.11 Real-time slice-record of a cone on a box

former value until it detects next signal, thereby causing a flat part and an abrupt fall in the recorded curve.

Excluding the small difference, the necessary information can be perceived even from this type of analog record. It is therefore concluded that this apparatus will be applicable as visual organs for robots and blind men.

5 CONCLUDING REMARKS

With the object to applying to the visual organs, high speed and low cost method for acquiring the distance image has been developed. Not only the apparatus for acquiring digitally the distance image but also the analog one which allows the real time display of the distance have been produced. In comparison to the conventional ones, the features of the present method are summarized as follows:

(1) High speed sampling in the distance measurement has been achieved by employing the counting method for detecting the position of the spot image, for which the detection is made at the rate as high as the frequency of the HSS. The characteristics of this method lies in the point that the light beam is scanned on the object surface synchronously with the camera synchronizing signal and the position of the spot image focused on the CCD module is detected by counting the time elapse from the end of synchronizing signal to its location.

(2) The signal of this spot image position is the digital one, and after they are transmitted to the personal computer the distance image corresponding to the whole pixels of the scene is obtained using the software developed at the same time. The time required for the measurement is as low as 17 sec. and that for calculating the distance, recording the data and displaying the scene is 6 min.

(3) The analog computing circuit has been devised in order to reduce the calculation time and apply it to recognize the object in real-time. This device will be applicable to the robot vision, automatic control of unmanned carriers, and the eyes for blind men.

REFERENCES

- [1] H. Tsuchiya and Y. Fukuda, *Image processing*, (1991-1), pp.174-186, Corona Pub.
- [2] M. Yachida, *Robot vision*, (1990-10), pp.149-206, Shokodo Pub.
- [3] M. Hanajima and T. Kurono, *Range imaging system and its application*, Tech. J. of Television Soc., IPA86(1986), pp.41-46.
- [4] M. Ejiri, Y. Ohta and K. Ikeuchi, *Machine Vision*, (1990), pp.30-50, Shokodo Pub.
- [5] M. Rioux, *3-D Camera based on Synchronized Scanning*, 13th Internat. Symp. on Indust. Robot, (1983), pp.17-48.
- [6] M. Ishii and T. Nagata, *Feature Extraction of 3 Dimensional Objects with a Laser Tracker*, Trans. SICE, 10-5(1974-10), pp.599-605.
- [7] T. Ninomiya, T. Hiroi and Y. Nakagawa, *Compact 3-D Sensor for Robots Using X-Shaped Slit Light*, Trans. SICE, 23-10(1987-10), pp.1009-1016.
- [8] O. Ozeki, T. Nakano and S. Yamamoto, *Real Time Range Measurement System Using Slit-Ray Projection*, Trans. Electro. Comm. Soc., J68-D, No.5(1985-5), pp.1141-1148.
- [9] T. Ito, D. DeMenthon and L.S. Davis, *Fast Range Scanning using an Optic RAM*, Proc. IEEE Int. Conf. Rob. Auto., 3(1990), pp.1503-1508.