

For connection *C* the boundary conditions are similar to those for connection *A* except that neither V_3^5 nor V_4^1 are equal to zero, and

$$I_4^1 = -I_3^5 = I_o$$

$$V_o = [V_4^1 - V_3^5] = I_o Z_L.$$

It is convenient to let V_1^1 be the independent variable in the analysis. One should also observe that (A.1) or (A.3), when written for a single turn of the coupler (see Fig. 3), must contain the proper transmission line lengths l_a or l_b .

REFERENCES

- [1] C. W. Allen and H. L. Krauss, "A wideband rotating coupler for VHF use," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-24, pp. 267-269, May 1976.
- [2] C. L. Ruthroff, "Some broadband transformers," *Proc. IRE*, vol. 47, pp. 1337-1342, Aug. 1959.
- [3] O. Pitzalis and Thomas P. Couse, "Broadband transformer design for RF transistor power amplifiers," *U.S. Army Technical Report*, ECOM-2989.
- [4] —, "Practical design information for broadband transmission line transformers," *Proc. IEEE*, pp. 738-739, Apr. 1968.
- [5] H. L. Krauss and C. W. Allen, "Designing toroidal transformers to optimize wideband performance," *Electronics*, pp. 113-116, Aug. 16, 1973.
- [6] I. Otawara and R. Sato, "Characteristics of the dc cut-off transmission-type transformer," *Electrical Engineering in Japan*, vol. 87, No. 7, July 1967.
- [7] S. Yokogawa and R. Sato, "Transformer using bifilar helices and its application," *IEEE, Japan*, vol. 87, p. 758, 1962.
- [8] A. Talkin and J. Cuneo, "Wideband balun transformer," *Rev. Sci. Instrum.*, vol. 28, p. 808, Oct. 1957.
- [9] Y. Shimada, "Fundamental considerations of broadband line type transformers," *Rev. Elec. Commun. Lab.*, vol. 18, No. 1-2, Jan.-Feb. 1970.
- [10] Y. Shimada, "Input impedance analysis of 1:1 balun," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, May 1970.
- [11] P. C. Magnusson, *Transmission Lines and Wave Propagation*. Boston, MA: Allyn and Bacon, 1971.

The Measurement and Use of Registered Reflectance and Range Data in Scene Analysis

DAVID NITZAN, SENIOR MEMBER, IEEE, ALFRED E. BRAIN, AND RICHARD O. DUDA, MEMBER, IEEE

Abstract—The theory and use of a laser sensor designed to provide arrays of registered intensity and range data for three-dimensional scene analysis are described in this paper. Equations are derived to relate the mean and variance of both the intensity and the range data to the physical properties of the target and the parameters of the sensor. Experimental results that verify these equations are given. Adaptive data smoothing and normalization procedures are presented. These procedures reduce photon noise and correct the intensity data for inverse square-law effects. The resulting intensity data give the product of the diffuse surface reflectance and the cosine of the angle of incidence for each picture element in the scene. Simple procedures are given for finding outlines of occluding objects, obtaining normal views of obliquely sensed planar surfaces, and extracting the images of horizontal surface regions from a scene. These procedures demonstrate the advantages of having registered reflectance and range data for scene analysis.

I. INTRODUCTION

IN RECENT YEARS, considerable attention has been devoted to the problem of analyzing pictures of three-dimensional scenes [1]-[3]. Much of this work has been concentrated on the analysis of single black-and-white images and has been devoted to overcoming the difficult problems

caused by projection, occlusion, and illumination effects such as shading, shadows, and reflections. While many monocular depth clues can be exploited to produce a three-dimensional interpretation of a two-dimensional image, the use of direct-range measurements can simplify these problems significantly. In this paper we describe an experimental system that provides registered intensity and range data for scene analysis. We also present simple, fast procedures that can be used to extract information useful for higher level analysis that would be difficult to obtain indirectly from intensity data alone.

Range information is so obviously relevant to the three-dimensional shape of objects that previous investigations of its use in scene analysis have by and large ignored the intensity information. Thus Shirai and Suwa [4], Shirai [5], Nitzan [6], Agin and Binford [7], Nevatia and Binford [8], Röcker [9], Popplestone *et al.* [10], and Röcker and Kiessling [11] have basically been concerned with various ways to describe objects by fitting surfaces to range data. Recently, Garvey [12] has demonstrated one way that both range and intensity data can be exploited for object recognition. Registered data of the type described in this paper should greatly facilitate any procedure that uses range and intensity data together.

Since the first part of this paper is devoted to the description and analysis of a particular instrument, a brief consideration of alternative approaches is appropriate. Two basically different

Manuscript received May 3, 1976. This work was supported in part by the National Science Foundation under Grant ENG75-09327 and in part by the Advanced Research Projects Agency under Contract DAHCO4-72-C-0008.

The authors are with the Stanford Research Institute, Menlo Park, CA 94025.

techniques can be used to measure range–triangulation and time of flight.¹ Let us consider the triangulation methods first.

Triangulation can be subdivided into stereoscopic and projected-light methods. The basic principles of stereo ranging are well known [15]. The stereo pair of images can be obtained either from two cameras or from one camera in two positions. A significant advantage of this approach is that it is passive, relying solely on normal scene illumination. However, to compute the range to a point in one image, one must find the corresponding point in the other image. Most of the work in using stereo for scene analysis has been devoted to solving this difficult correspondence problem [16]–[21].

One way to dispose of the correspondence problem is to use specially controlled illumination. Typical systems project sheets of light on the scene by means of a transmitter (a slit projector [4], a laser beam diverged by a cylindrical lens [7], or a collimated light beam rotated quickly by a mirror [22]), and record the image resulting from each sheet of light by means of a receiver (a television camera [4], [7] or a photodiode via a rocking mirror [22]; the reverse configuration can also be used [11]). Any point in the image determines a ray in space, and the range to the point hit by the ray is determined by computing the intersection of that ray with the sheet of light. Thus given the geometry of the transmitter and the receiver, the range is obtained by a simple trigonometric formula.

All triangulation range finders must make a tradeoff between accuracy and disparities between the two views. For accuracy, a large separation between the two cameras or between the transmitter and the receiver is desirable [23]. However, as the separation is increased, the disparities between the two views increase. In addition to complicating the correspondence problem, disparities lead to missing data because many points in the scene will be seen by one camera (or the receiver) but not by the other (or the transmitter).

A time-of-flight range finder eliminates this problem by keeping the transmitted and received beams coaxial. Thus an optical radar or lidar typically consists of a transmitter that illuminates the target with a collimated beam and a receiver that collects the portion of the reflected light that is essentially coaxial with the transmitted beam. The range is determined from the time needed for the light to travel from the transmitter to the target and back. A range image is obtained by using a scanning system to sweep the beam over the scene.

The time of flight can be determined either directly by using a pulsed laser and measuring the elapsed time, or indirectly by using a CW modulated beam and measuring the phase shift. The former method, which is the conventional approach for long-distance ranging, is used in a system designed at the Jet Propulsion Laboratory for a range finder to be used by a Mars explorer [24]. For our close-range application, we chose the latter method, primarily because it seemed technically easier to measure phase shift than elapsed time. In addition, a CW system provides a simple means for obtaining intensity as well as range data.

A detailed description of our sensor is given in Section II. In Section III we show how the raw data can be smoothed and normalized to obtain a direct measurement of the product of

the target reflectance and the cosine of the incident angle. Thus the intensity data provide quantitative information about the target surface without the effects of illumination that are so prominent in television imagery. Finally, in Section IV we show how straightforward procedures can be used to process these data to find outlines of occluding objects, to obtain normal views of obliquely viewed surfaces, and to segment a scene into regions corresponding to horizontal and non-horizontal surfaces.

II. THE REFLECTANCE/RANGE SENSOR

In this section we describe the reflectance/range sensor and its technical characteristics. Equations are derived relating the mean and variance of both the intensity and the range data to the physical properties of the target and the physical parameters of the system. These equations serve three purposes: they reveal the basic limits of the accuracy of such an instrument, exhibit the design tradeoffs, and provide the basis for calibrating the instrument so that actual surface reflectance can be obtained from the intensity data.

A. General Description

As shown in Fig. 1, the instrument has three major functional components—a transmitter, a scanning mirror, and a receiver. The laser beam emitted by the transmitter is amplitude-modulated with a sinusoidal waveform. This modulated beam is then deflected by the scanner to cover the field of interest.

When the beam strikes an object, a portion of the light is reflected back along a direction essentially coaxial with the incident beam, captured by the scanner, and deflected into the receiver. Here, an interference filter rejects ambient light and passes the return signal to a photomultiplier tube. The photomultiplier converts the received light into a sinusoidal current at the modulation frequency. The amplitude of this current is proportional to the reflectance of the object; when the range reading is used to correct for the inverse square law, television-like intensity data are obtained. The phase of this current, relative to the transmitted signal, is proportional to the time it takes the light to travel from the instrument to the object and back, and thus provides the range value.

One of the major problems is that the dynamic range of the signal from the photomultiplier can exceed 100 dB. The very weak return from a dark distant object yields a noisy photon-limited signal. The very strong return from a bright close object can overload amplifiers, introduce range-measurement errors, and can even damage the photomultiplier tube. Because of this wide dynamic range, the signal obtained from an amplitude detector is passed through a logarithmic amplifier before it is digitized to provide the intensity data. The received signal is also sent to a phase detector, which measures its phase relative to the transmitted signal. The output of the phase detector is smoothed by a low-pass filter and digitized to provide the range data. Further smoothing and correction of the intensity and range data are done later by software, as described in Section III.

B. Signal and Noise Relations

The accuracy of the instrument depends on the power of the signal reaching the photomultiplier. The received power is a function of the transmitted power, the reflectance of the target surface, the position and orientation of the target surface relative to the instrument, and the optics of the receiver. There are several sources of noise, but the dominant one is

¹ Other techniques, such as focus ranging [13] and grid coding [14], can also provide useful range information; however, the accuracy of focus ranging is limited and grid coding does not readily yield range values.

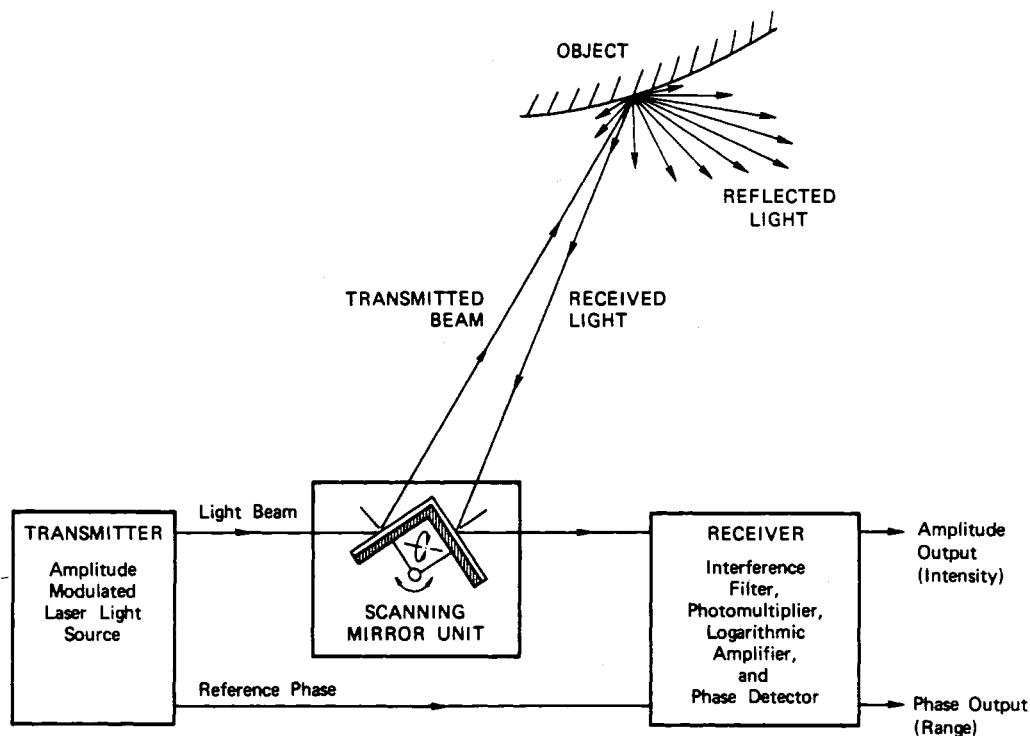


Fig. 1. Simplified block diagram of reflectance/range sensor.

photon noise resulting from the low intensity of the received light. The following equations, derived by applying general principles of laser communication theory [25], [26], relate the system parameters to the signal-to-noise (SNR) ratio.

1) *Surface Reflectance*: Fig. 2 shows a narrow collimated laser beam striking a surface at an angle of incidence θ_i with respect to the surface normal. At this surface, the time-average transmitted radiant flux \bar{F}_T (in watts) is partly absorbed and partly reflected. The reflected light contains two components—a mirror-like specular component that follows Fresnel's equations and a mat-surface diffuse component that is described approximately by Lambert's cosine law [27], [28]. It is usually thought that the specular component is significant for shiny surfaces and even for mat surfaces when the angle of incidence is large. However, because of the coaxial geometry of the instrument, specular reflections are received only at or very near normal incidence. Thus, except for these rather rare cases, only the diffuse component need be considered.²

Let \bar{I} be the time-average radiant intensity (in watts per steradian) of the light reflected at an arbitrary angle θ . Then \bar{I}_d , the diffuse component of \bar{I} , is approximately given by Lambert's cosine law:

$$\bar{I}_d = \frac{\bar{F}_T}{\pi} \rho_d \cos \theta \quad (1)$$

where ρ_d is the diffuse reflectance, the ratio of the total diffuse

² The subsequent analysis, though only approximate, is in good agreement with our experimental observations of common opaque solid objects measured at ranges sufficiently close so that atmospheric scattering is negligible. Highly specular surfaces usually appear dark rather than bright because they reflect back so little light, but on occasion they have led to anomalous results. As expected, with a truly mirror-like surface the instrument measures the range to the image of the object reflected by the mirror, not the range to the mirror. If the mirror surface itself reflects back an appreciable amount of light, an intermediate, average range value is obtained.

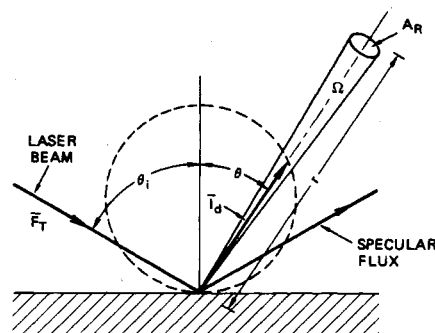


Fig. 2. Incident laser beam and reflected light components.

radiant flux reflected to the incident radiant flux \bar{F}_T . The value of ρ_d varies between zero for an ideally black surface and one for an ideally white surface. In general, ρ_d is wavelength dependent. However, since we illuminate the target with a monochromatic laser, we henceforth assume that the value used for ρ_d is the value at the wavelength of the laser.

2) *Average Received Signal*: The receiver captures only a small portion of the reflected radiant flux, and only a fraction α of that passes through the interference filter to the photomultiplier tube. Let A_R be the capture area of the receiver and let r be the range from the illuminated spot to the receiver, so that the solid angle Ω subtended by the receiver is A_R/r^2 . Then, the time-average radiant flux \bar{F}_p reaching the photomultiplier tube is given by

$$\bar{F}_p = \alpha \bar{I}_d \Omega = \left(\frac{\alpha A_R \bar{F}_T}{\pi} \right) \frac{\rho_d \cos \theta}{r^2} \quad (2)$$

Note that $\theta = \theta_i$, because the transmitted and received beams of our instrument are essentially coaxial.

Since the output current of the photomultiplier is proportional to \bar{F}_p , this equation shows that the amplitude of the

output signal is proportional to $\rho_d \cos \theta / r^2$. Range correction of the amplitude to eliminate the effect of the inverse square law is discussed in Section III-A.

3) *Intensity Noise*: When $\rho_d \cos \theta / r^2$ is small, as happens with dark distant objects viewed obliquely, the received power is very small and photon noise is significant. Each photon carries an energy E , given by

$$E = \frac{hc}{\lambda} \quad (3)$$

where h is Planck's constant, c is the speed of light, and λ is the wavelength of the illumination. Thus the average number of photons \bar{n} received in a time interval T is given by

$$\bar{n} = \frac{\bar{F}_p T}{E} = \frac{\lambda T}{hc} \bar{F}_p. \quad (4)$$

If the photocathode has a quantum efficiency η , the average number of emitted photoelectrons \bar{n}_{pe} is given by

$$\bar{n}_{pe} = \eta \bar{n}. \quad (5)$$

Under the assumption that photoemission is a Poisson process, the standard deviation σ_{pe} for n_{pe} is just $\sqrt{\bar{n}_{pe}}$, and thus the SNR ($\text{SNR} = \bar{n}_{pe} / \sigma_{pe}$) is also $\sqrt{\bar{n}_{pe}}$. Combining (2), (4), and (5), we obtain

$$\text{SNR} = \frac{\bar{n}_{pe}}{\sigma_{pe}} = \sqrt{\bar{n}_{pe}} = \left(\frac{\alpha \eta \lambda A_R \bar{F}_T T}{\pi h c} \frac{\rho_d \cos \theta}{r^2} \right)^{1/2}. \quad (6)$$

There are, of course, other sources of noise, such as laser noise, ambient noise, thermionic or dark-current noise, secondary-emission noise, and noise in the subsequent amplifier. However, photon noise dominates these other kinds of noise in our instrument.

4) *Range Noise*: Because of the amplitude modulation, the instantaneous transmitted flux $F_T(t)$ is given by

$$F_T(t) = \bar{F}_T (1 + m \cos \omega_m t) \quad (7)$$

where m is the modulation factor ($0 \leq m \leq 1$) and $\omega_m / 2\pi$ is the modulation frequency. The instantaneous flux $F_p(t)$ reaching the photomultiplier tube has a similar waveform, but is shifted in phase by an amount proportional to the total optical path length from the transmitter to the target and back to the photomultiplier. The phase detector measures this phase shift, and its output is a linear function of the range to the target. Photon noise also limits the accuracy of this range measurement. An exact analysis of the effect of photon noise on the measured phase value is difficult. However, an approximate analysis, given in Appendix A, indicates that the standard deviation σ_r of the range can be estimated by

$$\sigma_r \approx \frac{c}{\sqrt{2} m \omega_m \text{SNR}} = \frac{1}{2\sqrt{2} \pi} \frac{\lambda_m}{m \text{SNR}} \quad (8)$$

where the SNR is given by (6), and λ_m is the wavelength of the modulation:

$$\lambda_m = \frac{2\pi c}{\omega_m}. \quad (9)$$

Thus a high SNR and a large percentage modulation lead to small range errors, as one would expect. Furthermore, unless the modulation wavelength is small, σ_r can be appreciable even with high SNR's. Combining (6) and (8), we obtain

$$\sigma_r \approx \frac{\lambda_m}{m} \left(\frac{hc}{8\pi\alpha\eta\lambda A_R \bar{F}_T T} \frac{r^2}{\rho_d \cos \theta} \right)^{1/2} \quad (10)$$

This equation shows how the range error depends on system parameters and provides guidance for system design.

C. System Parameters

We constructed the instrument to obtain registered intensity and range data for an experimental investigation of machine perception. Thus many of the design decisions were made with this laboratory use in mind. In particular, the instrument was designed to sense objects that can be seen in a 128×128 image at a range of 1–5 m with a range resolution of about 1 cm.

Our present instrument uses a 9-MHz modulation frequency ($\lambda_m = 33.3$ m). Since the optical path from the scanner to the receiver and back is twice the range, a displacement of 1 cm corresponds to a phase shift of 0.0006 wavelength, or 0.2° . While it is possible to measure phase shift to better than 0.01° [29], the wide dynamic range of the received signal makes this a difficult task. As we observed earlier, the output of the photomultiplier is proportional to $\rho_d \cos \theta / r^2$. The reflectance ρ_d can vary from about 0.02 to almost 1, the range r from 1 to 5 m, and the practical angle of incidence θ from 0° to about 87° . Thus $\rho_d \cos \theta / r^2$ has a dynamic range of 25 000:1 or 88 dB. Furthermore, at normal incidence, specular reflection increases the output considerably above the levels predicted by Lambert's cosine law, and a dynamic range of 100 dB or more is encountered in practice.

The other major problem stems from photon noise. Safety and cost call for the use of a low-power laser. In theory, one can exchange power for integration time, since (6) and (10) show that only the total energy $\bar{F}_T T$ matters. However, excessively long integration times limit applications and introduce problems with drift. Since the present investigation did not require high-speed response, the tradeoff was made in the direction of low power at the cost of speed. Because the number of samples taken for each picture element (pixel) depends on the amplitude of the received signal (see Section III), the required time depends on the contents of the scene. However, typical times are 500 ms per pixel, or more than 2 h for a full 128×128 image. Clearly, most practical applications will require some combination of greater transmitted power, larger receiver area, and shorter modulation wavelength.

Fig. 3 shows the major elements in the complete system. The plane-polarized output of a 15-mW HeNe laser ($\lambda = 632.8$ nm) is amplitude-modulated by passing it through an ammonium dihydrogen phosphate (ADP) modulating crystal and a Glan-laser analyzer. Both a 9-MHz modulation voltage and a low-frequency bias voltage are applied across the ADP crystal to rotate the plane of polarization. The analyzer splits the beam into an on-axis beam and an off-axis beam. The on-axis beam provides the modulated output for the scanner. The average power of this transmitted beam is about 6 mW. The off-axis beam provides the reference signal for the phase detector. It also provides an automatic-gain-control (AGC) signal that corrects for thermal drift and maintains a constant output from the transmitter.

The on-axis beam is deflected by the scanner and illuminates the target. The received beam, which is a small fraction of the reflected light, is sent to the photomultiplier tube through a 632.8-nm interference filter having a 2-nm passband. A lens behind the interference filter images the illuminated target spot on a pinhole in a diaphragm in front of the photomultiplier tube. This pinhole limits the maximum off-axis angle to

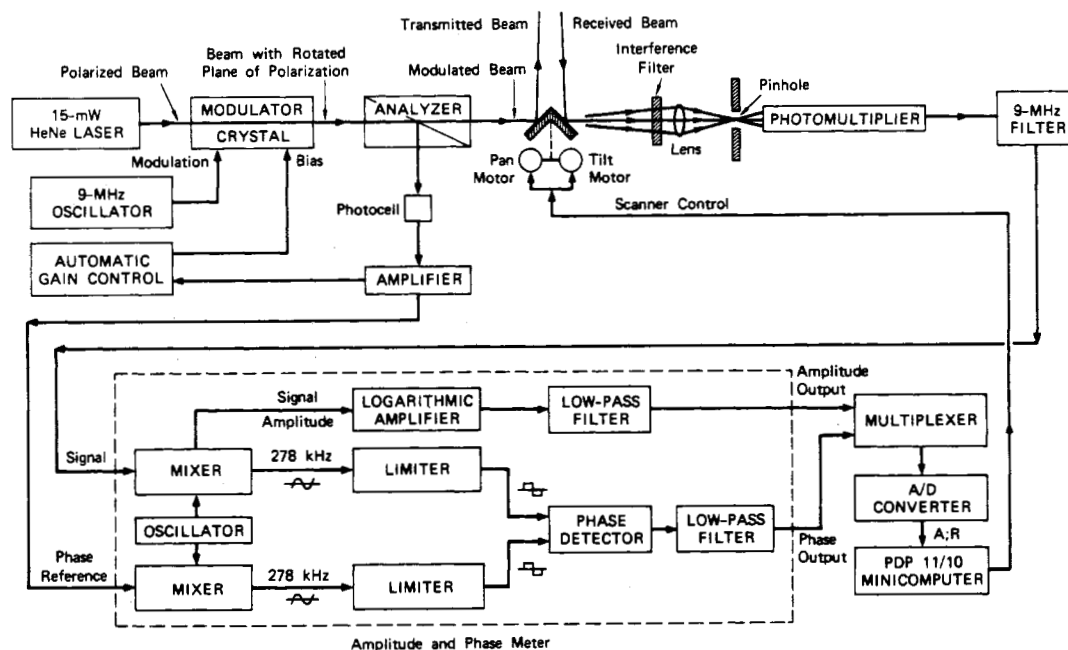


Fig. 3. Components of the reflectance/range sensor.

TABLE I
INSTRUMENT PARAMETERS AND COMPUTED SIGNAL PARAMETERS

Instrument Parameters		
$\bar{F}_T = 6 \times 10^{-3} \text{ W}$	$A_R = 1.5 \times 10^{-4} \text{ m}^2$	
$\lambda = 0.6328 \times 10^{-6} \text{ m}$	$\alpha = 0.25$	
$\lambda_m = 33.3 \text{ m}$	$\eta = 0.1$	
$m = 1.0$	$T = 0.01 \text{ s}$	
Signal Parameters		
Parameters	Low-Signal Strength	Intermediate-Signal Strength
ρ_d	0.02	0.3
θ (degrees)	87	45
r (m)	5	3
\bar{F}_p (W)	3×10^{-12}	1.7×10^{-9}
\bar{n}_{pe} (photoelectrons)	10^4	5×10^6
SNR	100 (40 dB)	2300 (67 dB)
σ_r (cm)	4	0.2

about 0.5° , and the combination of the interference filter and the pinhole effectively eliminates noise due to ambient illumination.³

A premium-grade gallium arsenide photomultiplier tube was selected because of its relatively high quantum efficiency and low dark current. The output of the photomultiplier passes through a 9-MHz filter to the amplitude and phase meter, which is a combination of a Hewlett-Packard 8407A network analyzer and an 8412A phase-magnitude display. Both the amplitude and the phase output signals are smoothed by single-pole low-pass filters. These filters have by far the smallest bandwidth in the signal path and effectively establish

³ Because the transmitted and received beams are not exactly parallel, at very close ranges the image of the target spot actually misses the pinhole. This serves to protect the photomultiplier tube from accidental damage caused by the large return obtained from very close targets.

the hardware observation interval T . The amplitude and phase outputs pass through a multiplexer and an analog-to-digital converter to a PDP-11/10 minicomputer. The minicomputer controls the scanner and averages these data further (see Section III) and records them on a cassette tape for subsequent scene analysis on another computer (see Section IV).

The equations derived in Section II-B can now be used to estimate the signal strengths, noise levels, and theoretical accuracy of the instrument. Numerical values are given in Table I for a low-signal case and an intermediate-signal case. These results show the difficulty of obtaining accurate range measurements, since even with a 40-dB SNR the estimated standard deviation for the range is 4 cm. At intermediate-signal levels, these calculations suggest considerably better than 1-cm resolution, although it should be remembered that with very strong signals photon noise is no longer the dominant limitation on resolution. In the next section, experimental results are given that are in general agreement with these computed results.

III. DATA SMOOTHING AND CORRECTION

In this section we describe how the raw amplitude and phase data are smoothed and corrected to produce compensated intensity and range data. The amplitude data are range-corrected to remove inverse square-law effects. To reduce noise, both the intensity and the range data are averaged in software. This smoothing is done adaptively, with longer time being spent in the dark areas where the SNR is lower. In addition, reference point readings are used to compensate for any long-term drift in the instrument.

A. Reflectance Data

Let A denote the output of the analog-to-digital converter when the amplitude output is sampled. From Fig. 3 we see that A is a linear function of the logarithm of the output of the photomultiplier tube. Since the photomultiplier output is proportional to the radiant flux input, and since (2) shows that the radiant flux input is proportional to $\rho_d \cos \theta / r^2$, it

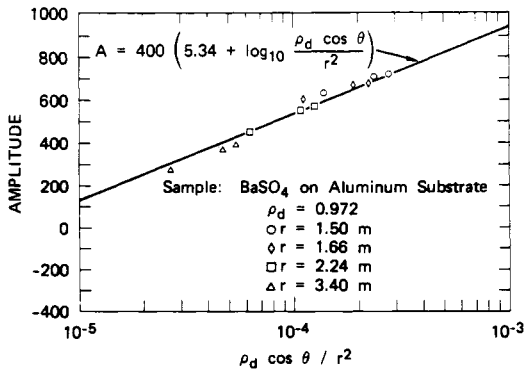


Fig. 4. Measured amplitude A from a standard BaSO_4 sample for different values of r and θ .

follows that

$$A = C_1 \left(\log_{10} \frac{\rho_d \cos \theta}{r^2} + C_2 \right) \quad (11)$$

where C_1 and C_2 are constants, independent of ρ_d , θ , and r . For our instrument, the values $C_1 = 400$ and $C_2 = 5.34$ were obtained from experimental data. Fig. 4 shows the agreement between (11) and experimental data obtained with a standard white-mat sample made by spraying 20 thin layers of barium sulfate on an aluminum substrate. (The reflectance of BaSO_4 at the 632.8-nm wavelength of a HeNe laser is 0.97 [28].) The amplitude values shown were measured for four r values, each at three values of θ . The right, middle, and left points for each r value correspond to θ values of 0° , 30° , and 60° , respectively.

From (11), it follows that

$$\rho_d \cos \theta = r^2 10^{(A/C_1 - C_2)}. \quad (12)$$

Once the range r is measured, this equation allows us to compensate for the inverse square law and thus to obtain $\rho_d \cos \theta$. This range correction is routinely done, and the quantity $\rho_d \cos \theta$ provides the intensity data for further scene analysis.

Because $\rho_d \cos \theta$ has a dynamic range of more than 200:1, for display purposes we found it more convenient to work with the logarithm of $\rho_d \cos \theta$. This commonly used technique compensates for the fact that the distribution of reflectance values over a scene is usually strongly skewed toward the darker values, and it also seems to correspond to the logarithmic sensitivity of the eye [30]. For convenience, we define the apparent reflectance ρ_d^* by the following linear function of the logarithm of $\rho_d \cos \theta$,

$$\rho_d^* = C_1 (\log_{10} \rho_d \cos \theta + C_2) \quad (13)$$

so that ρ_d^* can be computed easily from the digitized output A and the range r by

$$\rho_d^* = A + 2C_1 \log_{10} r. \quad (14)$$

This quantity, scaled to accommodate the dynamic range of the display, is what is shown in the intensity images in this paper.

The values of $\rho_d \cos \theta$ obtained from (12) yield television-like intensity data. However, because the illumination is provided by the sensor and is coaxial with the line of sight, there are no shadows, no shading effects, fewer highlights, fewer reflections, and no reduction in dynamic range due to secondary illumination. Highlights do arise when the angle of

incidence is very small. When this happens, Lambert's cosine law usually breaks down, and (12) yields values greater than one. However, such highlights appear less frequently than the highlights obtained from the extended sources of illumination typically used for television.

We note in passing that other techniques can be used to obtain more direct measurements of reflectance. For example, by locally fitting a plane to the range data, one could estimate the direction of the surface normal, determine $\cos \theta$, and thus obtain ρ_d . Alternatively, if several wavelengths were used, one could form ratios of the values of $\rho_d(\lambda) \cos \theta / r^2$. Since the common factor $\cos \theta / r^2$ cancels, the result is a spectral signature of the target that is independent of r and θ [31]. This approach does not require the use of range data and thus eliminates any errors in estimating r and θ .

B. Range Data

1) *Noise Statistics*: Let R denote the output of the analog-to-digital converter when the phase output is sampled (see Fig. 3). Since the phase shift is a linear function of range, in the absence of noise,

$$r = C_3 R + C_4. \quad (15)$$

As with the intensity data, the coefficients C_3 and C_4 were obtained by fitting pairs of measured (R, r) values [32].

To investigate the effects of noise on range measurement, we repeatedly measured A and R for 27 cases, each having three values of reflectance, incidence angle, and range. The sample means and sample standard deviations were computed for each case. The standard deviation for the amplitude signal proved to be negligible for the scene analysis procedures described in Section IV. However, the standard deviations for the range were quite significant, and varied with A as shown in Fig. 5.

It follows from (10) that σ_r should be inversely proportional to the square root of $\rho_d \cos \theta / r^2$, which is related to the amplitude A by (12). Combining (10) and (12), we obtain the theoretical relation

$$\sigma_r(A) = \sigma_0 e^{-A/A_0} \quad (16)$$

where

$$\sigma_0 = \frac{\lambda_m}{m} \left(\frac{hc}{8\pi\alpha\eta\lambda A_R \bar{F}_T T} \right)^{1/2} 10^{C_2/2} \quad (17)$$

and

$$A_0 = \frac{2C_1}{\ln 10}. \quad (18)$$

As Fig. 5 shows, this function fits the data quite well with the values $\sigma_0 = 4.3$ cm and $A_0 = 333$ analog-to-digital-converter units.

2) *Adaptive Smoothing*: The standard deviation $\sigma_r(A)$ indicates the expected error that will be encountered in sampling the range output once. By averaging repeated readings, one can reduce this expected error. The low-pass filter following the phase detector (see Fig. 3) effectively averages readings over a 20-ms interval (twice its time constant; see Appendix B), but this is a fixed amount of smoothing, independent of the noise level. If the amplitude A is large, this much smoothing might not be needed; if A is small, it may be inadequate.

Since A can be measured quite accurately from just one sample, we are able to use (16) to determine the expected single-sample range error and thus to determine how many additional samples are needed to reduce the expected error

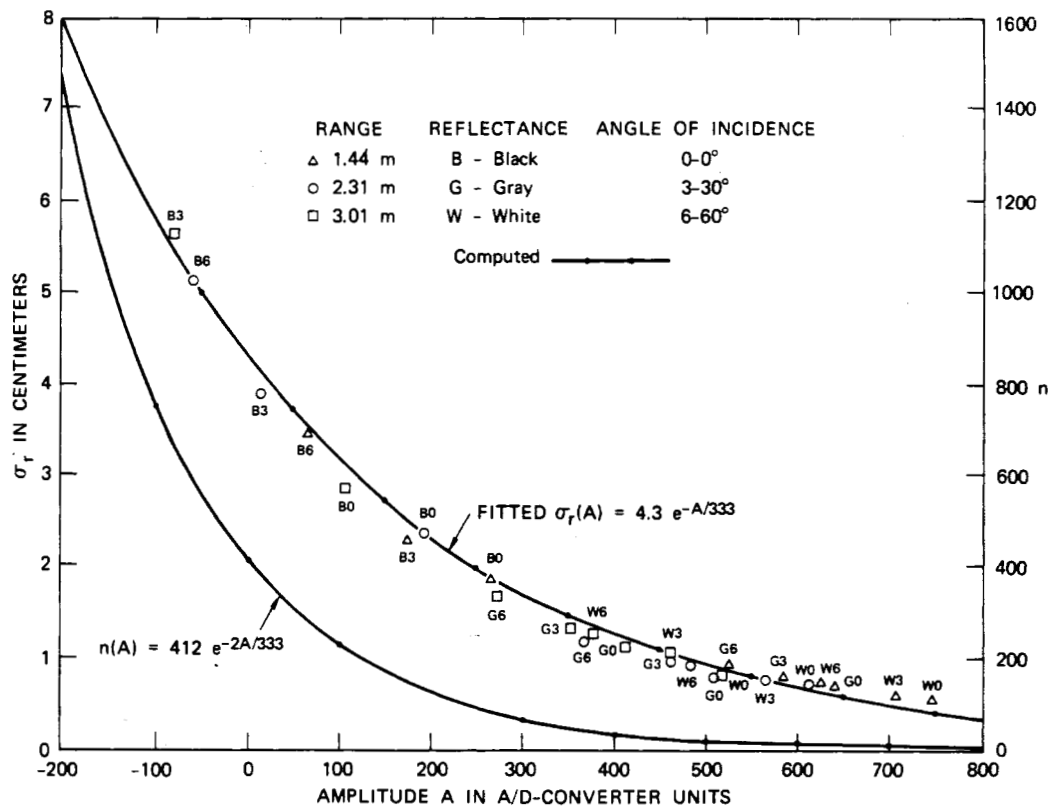


Fig. 5. Experimental and fitted standard deviation of range and computed number of samples as a function of amplitude.

to an acceptable value. The 1-ms sampling time interval ΔT of the analog-to-digital converter is much shorter than the 10-ms time constant T_f of the low-pass filter. Thus successive samples are highly correlated. It is shown in Appendix B that, under this condition, the standard deviation $\bar{\sigma}_r$ of the average of n samples is given approximately by

$$\bar{\sigma}_r = \frac{\sigma_r(A)}{\sqrt{n_{\text{eff}}}} \quad (19)$$

where

$$n_{\text{eff}} = \frac{\beta}{2 \left(1 - \frac{1 - e^{-\beta}}{\beta} \right)} \quad (20)$$

and where

$$\beta = \frac{n\Delta T}{T_f} \quad (21)$$

Note that if $n\Delta T \ll T_f$, then $n_{\text{eff}} \approx 1$, while if $n\Delta T \gg T_f$, then $n_{\text{eff}} \approx n\Delta T/2T_f$. If a desired $\bar{\sigma}_r$ is specified and if A is measured, then one can use (16) to find $\sigma_r(A)$, (19) to find n_{eff} , and (20) and (21) to find the required number of samples. We have adopted a simpler slightly conservative approach by approximating n_{eff} by the asymptotic value $n\Delta T/2T_f$; then, equations (16) and (19) yield

$$n(A) = \frac{2T_f}{\Delta T} \frac{\sigma_r^2(A)}{\bar{\sigma}_r^2} = \frac{2T_f}{\Delta T} \frac{\sigma_0^2}{\bar{\sigma}_r^2} e^{-2A/A_0} \quad (22)$$

The lower curve shown in Fig. 5 corresponds to $\bar{\sigma}_r = 0.95$ cm. This function was used to determine adaptively the number of

range samples to take and average to obtain the final range and amplitude reading.

3) *Compensation for Drift*: The actual time T_p required to obtain a range reading for a single pixel includes not only the sampling time $n\Delta T$, but also the time T_m to move the scanning mirror to the spot and the time (about five time constants) needed for any transient in the low-pass filter to die out. Thus

$$T_p = n\Delta T + T_m + 5T_f \quad (23)$$

With $\Delta T = 1$ ms, $T_m = 3$ ms, and $T_f = 10$ ms, the dominant term is $n\Delta T$, particularly in dark areas where n becomes very large. Unfortunately, over a 10-min interval, new errors begin to be introduced by drift. To correct for these errors, when the scanning of a 128-pixel column is completed, the scanner automatically reads the range to a convenient reference point. The reference-point readings are fitted by a polynomial to produce a range-drift function $\hat{r}(t)$. The range readings $r(t)$ for other pixels are then corrected by multiplying them by $r_0/\hat{r}(t)$, where r_0 is the true range to the reference point. Although the effects of drift are relatively small, such correction is necessary to maintain range accuracy on the order of 1 cm.

C. Experimental Results

Using the smoothing and correction procedures described in the preceding sections, we obtained intensity and range data for a number of indoor scenes. Under the control of a PDP-11/10 minicomputer, the scanner sampled a 128×128 raster (plus 128 reference-point readings). The average values of A and R were computed on the minicomputer. All subsequent processing, including the drift correction of the range data and the range correction of the amplitude data, was done on a

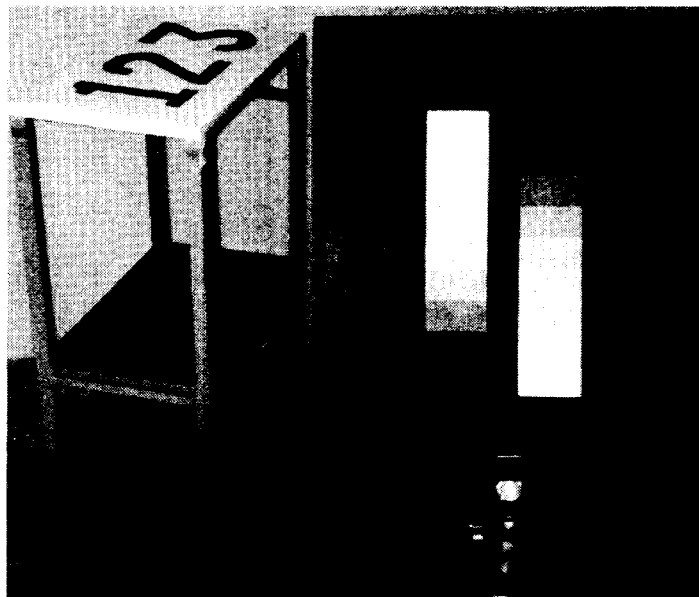


Fig. 6. Photograph of a test scene.

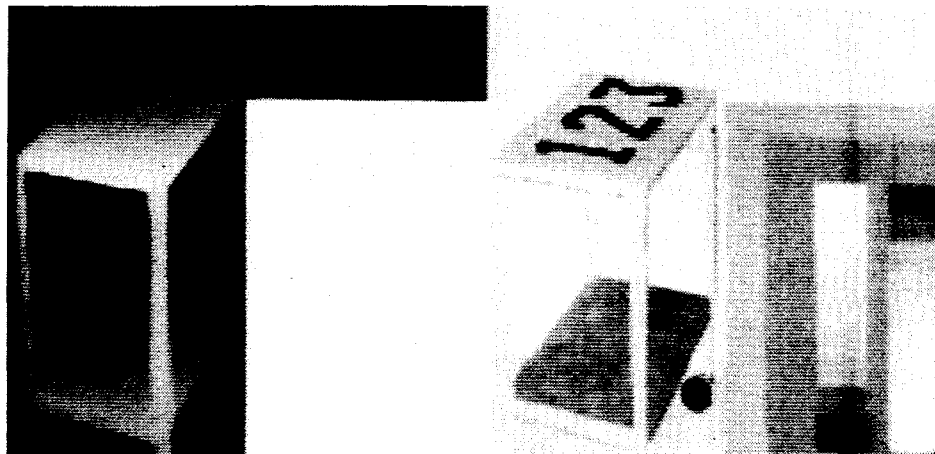


Fig. 7. Range and intensity data for the test scene.

PDP-10 computer; a Ramtek GX-100 display unit was used for pictorial output.

Two typical results are shown in Figs. 6-9. The photograph in Fig. 6 shows a cart and a linear gray-level chart viewed from the position of the sensor. The range and apparent reflectance (r and ρ_d^*) data for this scene are shown in Fig. 7. The range data are shown in a pictorial representation in which closer points appear to be brighter. Note that, except for a slight amount of cross talk on the top of the cart, the range image is insensitive to reflectance variations. In comparing the photograph and the intensity data, note that the latter is completely free of shadows. The noise in the intensity data is quite low and justifies quantization to about 200 levels.

Fig. 8 is a photograph of an office scene that contains some specular surfaces. Note that the shiny table legs appear to be dark in the intensity data shown in Fig. 9 because they reflect very little light back to the instrument. Highlights are picked up from the chair legs, as expected, but note the absence of the reflections present in Fig. 8. As with the previous scene, little amplitude information appears in the range data. Except for a very few anomalous pixels (such as at the center of the wall outlet), the accuracy of the range measurements is in

agreement with the specified maximum standard deviation of 1 cm.

IV. SCENE ANALYSIS OPERATIONS

Various types of information can be extracted from registered range and reflectance data for the purposes of object recognition and scene description. In this section we present some basic operations that are applicable to a variety of scene analysis problems. Specifically, in Section *A* we use the range data to extract the outlines of occluding objects, in Section *B* we show how the Cartesian coordinates for each pixel can be obtained, and in Section *C* we use these coordinates to obtain normal views of obliquely sensed planar surfaces and to extract the images of horizontal surfaces.

A. Jump Boundaries

A jump boundary is, by definition, the border between the images of two surfaces, one of which occludes the other; it is characterized by range discontinuity. Extraction of jump boundaries is helpful (and often sufficient) for finding the outlines of objects and holes in a scene. Detection of these outlines may be applied, for example, to the automated in-



Fig. 8. Photograph of an office scene.



Fig. 9. Range and intensity data for the office scene.

spection of manufactured parts, the avoidance of collision with other objects or falling into holes, and the recognition of known objects in a scene. A procedure for finding jump-boundary points is described briefly below.

The array of measured range values is scanned, first along each successive row and then along each successive column. In each scan, the difference Δr between the range values of adjacent pixels is computed. A jump boundary is established where, for given laser-spot size and scanning increment, the magnitudes and signs of neighboring Δr values indicate large changes in both r and Δr . Both conditions must be satisfied to exclude pixels corresponding to a highly oblique surface, which are characterized by a large change in r but not in Δr , and neighboring pixels corresponding to two intersecting surfaces, which are characterized by a large change in Δr but not in r .

We applied this procedure to the range data of the scenes in Figs. 7 and 9. The resulting jump-boundary pixels are shown in Figs. 10 and 11 superimposed on the intensity data of Figs. 7 and 9, respectively. Note how the occluding edges that were found are precisely registered with the intensity data. A few jump-boundary pixels are missing, primarily because the corresponding Δr values are smaller than the selected threshold value (7.5 cm). Lowering the threshold would solve this problem, but would also introduce spurious jump-boundary pixels.

B. Coordinate Transformation

Consider the Cartesian coordinate system (x, y, z) shown in Fig. 12. The position of the scanning mirror unit (see Fig. 1) is Point O_r , located at (x_0, y_0, z_0) . The direction of the laser beam striking the center of the scene defines the principal ray

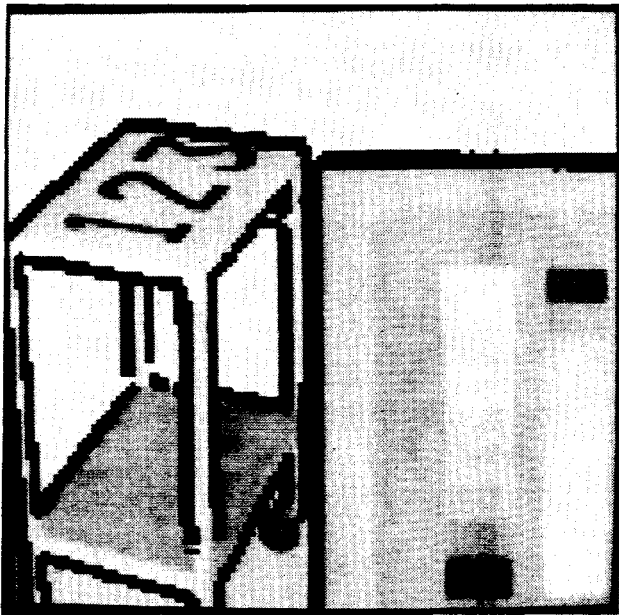


Fig. 10. Jump boundaries for the test scene.



Fig. 11. Jump boundaries for the office scene.

of the instrument. Axis y_r is defined along the principal ray so that $y_r = 0$ at Point O_r . A plane normal to y_r at a distance f_r from Point O_r is, by definition, the image plane. Image axes x_p and z_p are defined on the image plane so that $x_p = 0$ and $z_p = 0$ at the point of intersection between the principal ray and the image plane, and the directions of x_p and z_p coincide with the directions of panning and tilting of the scanning laser beam, respectively. Applying incremental steps Δx_p and Δz_p in these directions, we obtain an array of pixels indexed by J and I , respectively.

Consider in Fig. 12 a general point P located at (x, y, z) . The ray $O_r P$ pierces the image plane at Point P_p . Given the (J, I) or (x_p, z_p) values of Point P_p and the direction of the principal ray, we can compute the direction of ray $O_r P$; furthermore, having measured the range value $r = O_r P$, we can compute the coordinates x, y , and z .

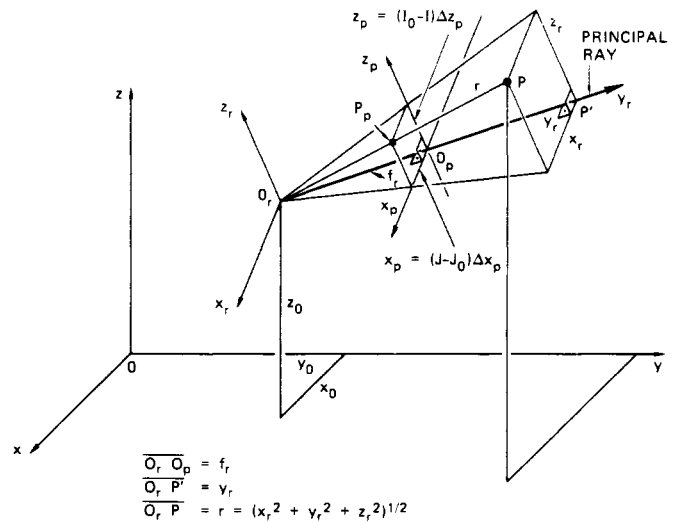


Fig. 12. Cartesian and image coordinates of a point in space.

The values of J, I , and r are conveniently transformed into the values of x, y , and z by multiplication of translation, rotation, and projection transformation matrices, using homogeneous coordinates [2]. The elements of these matrices are functions of eight parameters—three for translation, three for rotation, and two for projection. To determine these parameters, we use a calibration procedure, developed by Sobel [33] and improved by Agin [32], which fits the values of the eight parameters to four sets of measured (x, y, z, J, I) values by iteration. Given the parameter values, we can also easily obtain the inverse transformation from x, y , and z to J, I , and r .

C. Utilization of Cartesian Coordinates

From the (x, y, z) values of each (J, I) pixel, we can now extract three-dimensional information about the scene. Extraction of two types of information is illustrated next.

1) *Normal Views*: Given the intensity data of a planar surface that is not parallel to the image plane, we can eliminate the projective distortion in these data to obtain a normal view of this surface. This process may be applicable to recognition of skewed patterns, such as characters and labels on boxes that are positioned and oriented randomly.

Consider the special but important task of obtaining the top view of a horizontal surface of a given height H . We first find the set S of pixels whose z values fall within $H \pm \Delta z$, where Δz is a tolerance value depending on the expected error in the range data and on the geometry of the surface relative to the instrument. We then uniformly sample the xy plane and find the (J, I) values corresponding to each point (x, y, H) . Since these (J, I) values are not integers, the intensity displayed at (x, y) is obtained by interpolating among the four pixels surrounding (J, I) , with an intensity of zero (black) used for pixels not in S .

As an example, the normal view of the top of the cart in the scene of Fig. 7 is shown in Fig. 13. As expected, skewness has been eliminated: the top of the cart is rectangular, and the number "123" appears normal. The long line is the top view of a bright wall strip of height $H \pm \Delta z$, and the short line is the top view of a similar strip on the gray-level chart. If desired, these nonhorizontal surfaces could be eliminated by the methods described in the next section.

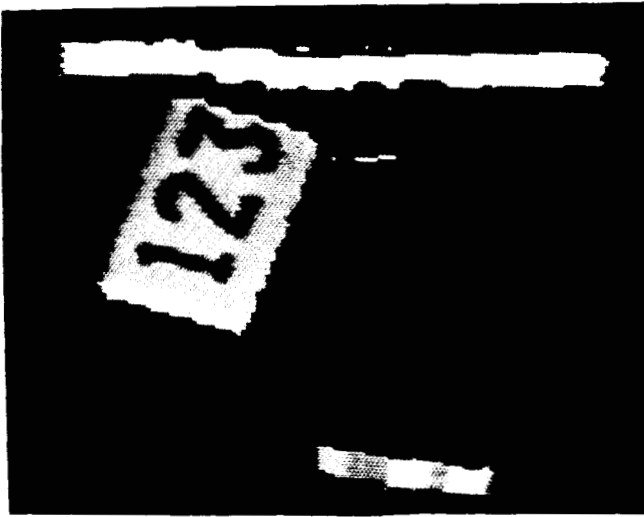


Fig. 13. Normal view of top of cart.

This method of obtaining normal views is not restricted to horizontal surfaces, but is applicable to any planar surface. Horizontal surfaces are particularly easy to find automatically, since the z values for a scene will be clustered at the corresponding height values. Vertical surfaces given rise to points clustered along straight lines in the xy plane, which can be found by the Hough transformation [2]. In principle, the Hough transformation can be used to find planar surfaces having arbitrary orientations. Regardless of how it is found, once a plane of interest is determined, a normal view of it can be obtained as easily as for a horizontal surface.

2) *Horizontal Surfaces*: Suppose that we wish to find the images of a number of horizontal surfaces whose heights are known *a priori*. Such information is useful for analysis of indoor scenes in which horizontal surfaces in general, and the floor in particular, are common.

Our procedure for finding the image of a horizontal surface of height H begins by finding the set S of pixels whose z values fall within $H \pm \Delta z$. The connected components of S typically include not only true horizontal surfaces, such as the tops of tables or boxes, but also stray portions of other objects, such as slices out of the wall. By fitting a plane to the (x, y, z) values for the pixels in each connected component, one can reject many of these nonhorizontal surfaces. However, it frequently happens that a component is a mixture of horizontal and nonhorizontal surfaces. For example, when a table touches a wall, a component of S will contain both the table top and a strip of the wall, and it is not meaningful to fit a single plane to such data.

We use the following simple heuristic procedure to divide S into homogeneous subsets, regions corresponding to one and only one plane. First, we shrink S to erase narrow necks and thin strips [34]. Next, we expand these results to form a set S' that regains many of the pixels around the border that were lost by shrinking. Because the necks and thin strips might be valid horizontal surfaces, we obtain them separately from the set difference $S - S'$. Finally, planes are fit to each of the connected components of S' and $S - S'$, and only those regions whose normal vectors are approximately vertical are retained as horizontal surfaces.

We applied this algorithm to extract the images of three horizontal surfaces—the floor, the bottom shelf, and the top shelf of the cart—in the test scene of Fig. 7. The results are

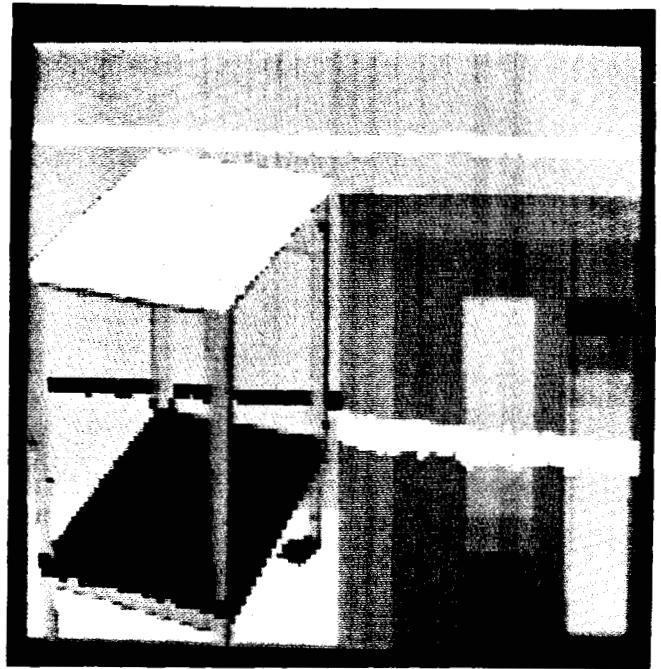


Fig. 14. Candidate horizontal surfaces for the test scene.

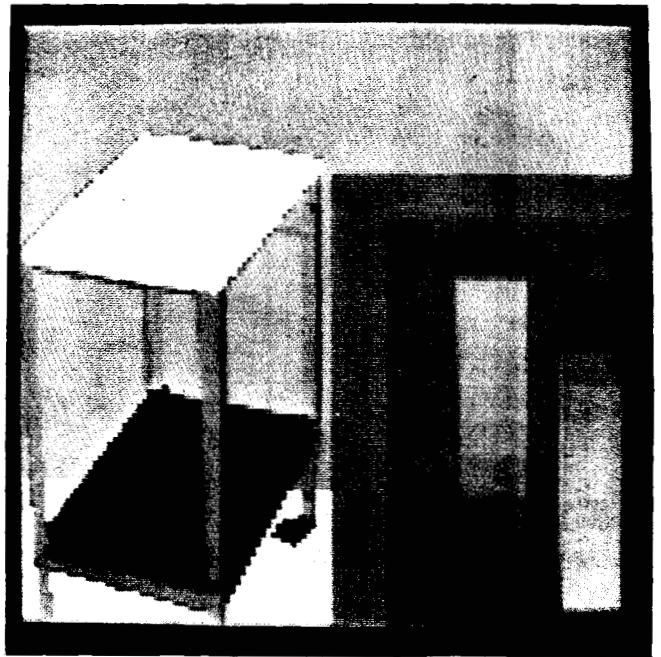


Fig. 15. Horizontal surfaces in the test scene.

shown in Figs. 14 and 15 superimposed on the intensity data. Fig. 14 shows the results of step 1 and thus includes strips of spurious pixels on the wall and on the chair. Black areas in this figure correspond to pixels with the height of the bottom shelf, and white areas correspond to pixels with either the height of the floor or the height of the top shelf. The final results are shown in Fig. 15. Another example is shown in Figs. 16 and 17. Here, the extracted images of the floor, the chair seat, and the desk top in the office scene of Fig. 9 are superimposed on the intensity data.

These results show how a straightforward and reliable procedure using range data can partition a scene into regions corresponding to horizontal and nonhorizontal surfaces. Our

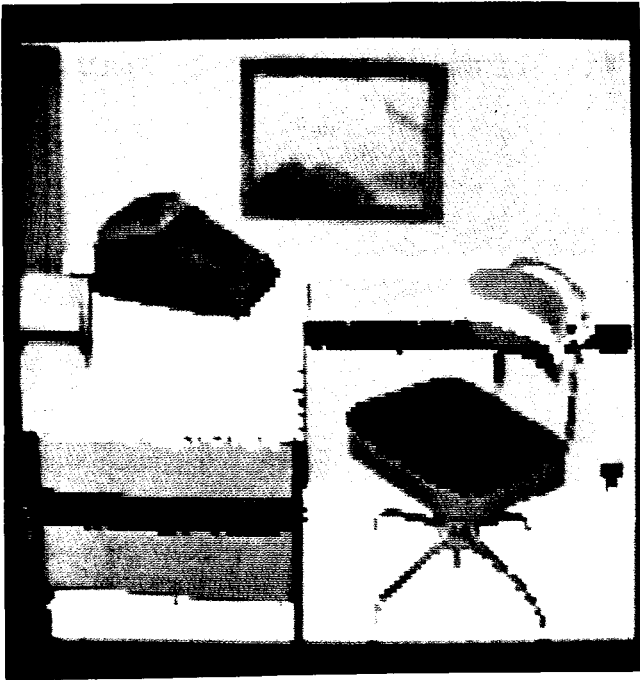


Fig. 16. Candidate horizontal surfaces for the office scene.

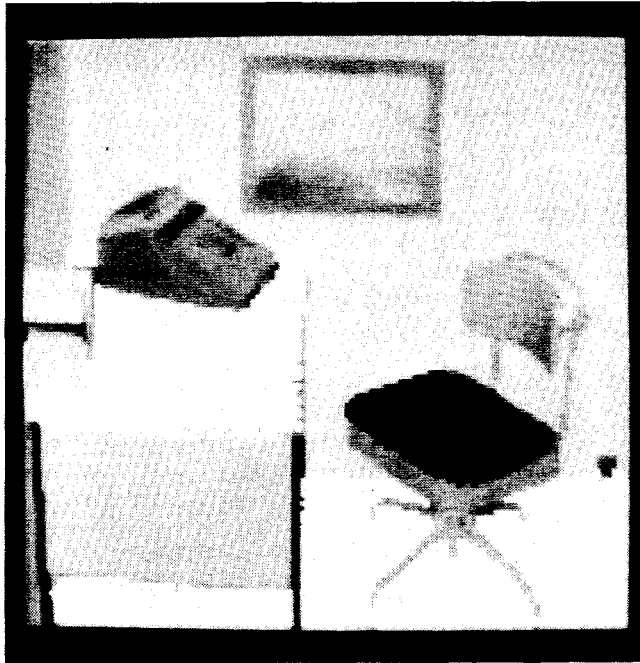


Fig. 17. Horizontal surfaces in the office scene.

earlier comments on automatically finding surfaces having other orientations hold here as well, and this can be viewed as the basis of a general segmentation procedure based on registered reflectance and range data.

V. CONCLUSIONS

The reflectance/range sensor described in this paper has several unique characteristics of importance to scene analysis. The arrays of intensity and range data that it provides are in precise register. The intensity data are free from illumination effects, such as shading and shadows, and contain relatively few highlights and reflections. In addition, since areas of low reflectance are not subject to extraneous illumination reflected

from brighter parts of the scene, the useful dynamic range is considerably wider than that obtained with television systems.

Except for the occasional appearance of highlights at points of normal incidence, the range-normalized intensity data are proportional to the product of the diffuse surface reflectance and the cosine of the incident angle. Using several wavelengths, one could obtain a spectral signature of the target surface that is independent of range and orientation. This kind of quantitative information provides new opportunities for the rapid discrimination and identification of objects.

The present system was designed for exploratory studies and is much too slow for most practical applications. However, by increasing the power, increasing the receiver area, using a higher modulation frequency, and/or adopting sequential analysis strategies that do not require a complete array, it should be possible to meet the time requirements of many applications. The equations developed in this paper show explicitly how these factors are related to the SNR of the intensity and the standard deviation of the range, and provide a theoretical basis for alternative designs.

Because the intensity and range images are always in register, occluding edges in the intensity image are easily identified as discontinuities in the range data. With a bit more difficulty, it should also be possible to find concave and convex edges from intersection boundaries in the range data [6]. We have presented a simple and effective procedure for extracting horizontal surfaces and have shown how normal views of planar surfaces can be obtained. These procedures can be extended to apply to arbitrary planar surfaces and thus provide information valuable for scene analysis, information that would be quite difficult to obtain from intensity data alone.

APPENDIX A ANALYSIS OF RANGE ERROR

In this appendix we derive an approximate expression for σ_r , the standard deviation for the range measurement.

Let r be the range from the exit pupil of the scanning unit to the object, and let r_0 be the optical path length from the exit pupil to the photomultiplier tube, so that the total length of the light path is $2r + r_0$. If the instantaneous transmitted flux $F_T(t)$ given by (7) is defined at the exit pupil, then the instantaneous flux $F_p(t)$ reaching the photomultiplier tube is given by

$$F_p(t) = \bar{F}_p [1 + m \cos(\omega_m t - \varphi)] \quad (A1)$$

where the time-average radiant flux \bar{F}_p is given by (2) and the phase shift φ is given by

$$\varphi = \frac{\omega_m}{c} (2r + r_0) = \frac{2\pi}{\lambda_m} (2r + r_0). \quad (A2)$$

The phase shift measured by the phase meter (see Fig. 3) differs from φ by a constant offset due to additional fixed parts of the signal path. However, without loss of generality, this offset can be absorbed in r_0 . By the same argument used to obtain (4) and (5), the instantaneous rate of emission of photoelectrons $a(t, \varphi)$ is given by

$$a(t, \varphi) = \bar{a} [1 + m \cos(\omega_m t - \varphi)] \quad (A3)$$

where the time-average rate \bar{a} is given by

$$\bar{a} = \frac{\bar{n}_{pe}}{T} = \frac{\eta \bar{F}_p}{E} = \frac{\alpha \eta \lambda A_R \bar{F}_T}{\pi h c} \frac{\rho_d \cos \theta}{r^2}. \quad (A4)$$

In statistical terms, what we have is a time-varying Poisson process with an unknown parameter φ . General formulas for the maximum likelihood estimate of φ have been derived by Bar-David [35]. The basic result can be stated as follows. Let $Q(\varphi)$ be the integral of the emission rate over the observation interval T ,

$$Q(\varphi) = \int_0^T a(t, \varphi) dt \quad (A5)$$

and let t_1, t_2, \dots, t_M be the successive times of emission of photoelectrons over that interval. Equation (29) of Bar-David [35] states that the maximum likelihood estimate $\hat{\varphi}$ for the unknown phase must satisfy

$$\sum_{j=1}^M \frac{1}{a(t_j, \varphi)} \frac{\partial a(t_j, \varphi)}{\partial \varphi} = \frac{dQ}{d\varphi}. \quad (A6)$$

Henceforth we will assume that the observation interval T is an integral multiple of the period $2\pi/\omega_m$ so that $Q(\varphi) = \bar{a}T$. Substituting this result and (A3) in (A6) yields

$$\sum_{j=1}^M \frac{\sin(\omega_m t_j - \hat{\varphi})}{1 + m \cos(\omega_m t_j - \hat{\varphi})} = 0. \quad (A7)$$

Following Herrmann [36], we consider the photoelectric current $i(t)$ given by

$$i(t) = e \sum_{j=1}^M \delta(t - t_j) \quad (A8)$$

where e is the charge of an electron and $\delta(\cdot)$ is the unit impulse function. Then (A7) may be written as

$$\int_0^T i(t) \frac{\sin(\omega_m t - \hat{\varphi})}{1 + m \cos(\omega_m t - \hat{\varphi})} dt = 0. \quad (A9)$$

It is difficult to solve this equation explicitly for $\hat{\varphi}$. However, an approximate solution can be obtained for the case of small percentage modulation. For that case, equation (A9) reduces to the classical matched-filter solution

$$\int_0^T i(t) \sin(\omega_m t - \hat{\varphi}) dt = 0. \quad (A10)$$

If we define the in-phase and quadrature components of $i(t)$ by

$$i_1 = \frac{1}{T} \int_0^T i(t) \cos \omega_m t dt \quad (A11)$$

and

$$i_2 = \frac{1}{T} \int_0^T i(t) \sin \omega_m t dt \quad (A12)$$

respectively, then it follows that the solution to (A10) is given by

$$\hat{\varphi} = \tan^{-1} \frac{i_2}{i_1}. \quad (A13)$$

This classic solution is not optimal in the more interesting case of large percentage modulation. The basic reason for this rather surprising result is that, if $i(t)$ is thought of as a sinusoi-

dal signal plus additive Gaussian noise, then the noise is not stationary. The variance of the noise is greater when the signal is strong than when the signal is weak, and (A9) takes this into account, whereas (A10) does not. Nevertheless, we shall assume the use of the matched-filter solution, since it is easier to analyze and it is easier to implement; it is essentially the method actually used in our phase demodulator, and Herrmann's analogous results for detection [36] suggest that the resulting loss in performance is not large.

Since (A2) shows that r is a linear function of φ , the corresponding estimate \hat{r} for the unknown range r is

$$\hat{r} = \frac{\lambda_m}{4\pi} \hat{\varphi} - \frac{r_0}{2}. \quad (A14)$$

To find the mean $E[\hat{r}]$ and variance $V[\hat{r}]$ of this estimate we must make use of the fact that $i(t)$ is a nonstationary Poisson process with mean $\mu(t) = e a(t)$ and autocovariance function $e^2 a(t_1) \delta(t_2 - t_1)$. From the central limit theorem and (A11) and (A12), it follows that i_1 and i_2 are essentially normally distributed. The mean and variance of i_1 are given by

$$\begin{aligned} \mu_1 = E[i_1] &= \frac{1}{T} \int_0^T E[i(t)] \cos \omega_m t dt \\ &= \frac{1}{T} \int_0^T e \bar{a} [1 + m \cos(\omega_m t - \varphi)] \cos \omega_m t dt \\ &= \frac{1}{2} e \bar{a} m \cos \varphi \end{aligned} \quad (A15)$$

and

$$\begin{aligned} V[i_1] &= E[(i_1 - \mu_1)^2] \\ &= E\left[\left\{\frac{1}{T} \int_0^T [i(t) - \mu(t)] \cos \omega_m t dt\right\}^2\right] \\ &= \frac{1}{T^2} \int_0^T \int_0^T E\{[i(t_1) - \mu(t_1)][i(t_2) - \mu(t_2)]\} \\ &\quad \cdot \cos \omega_m t_1 \cos \omega_m t_2 dt_1 dt_2 \\ &= \frac{1}{T^2} \int_0^T \int_0^T e^2 a(t_1) \delta(t_2 - t_1) \cos \omega_m t_1 \\ &\quad \cdot \cos \omega_m t_2 dt_1 dt_2 \\ &= \frac{e^2}{T^2} \int_0^T a(t_1) \cos^2 \omega_m t_1 dt_1 \\ &= \frac{e^2}{T^2} \int_0^T \bar{a} [1 + m \cos(\omega_m t_1 - \varphi)] \cos^2 \omega_m t_1 dt_1 \\ &= \frac{e^2 \bar{a}}{2T}. \end{aligned} \quad (A16)$$

Similar calculations show that

$$\mu_2 = E[i_2] = \frac{1}{2} e \bar{a} m \sin \varphi \quad (A17)$$

$$V[i_2] = \frac{e^2 \bar{a}}{2T} \quad (A18)$$

and

$$E[(i_1 - \mu_1)(i_2 - \mu_2)] = 0. \quad (\text{A19})$$

This latter equation shows that i_1 and i_2 are uncorrelated and hence statistically independent. Given that i_1 and i_2 are independent normal random variables, it is not difficult to obtain the exact distribution of $\hat{\varphi}$ from (A13). However, since the solution is only of interest when the standard deviations of i_1 and i_2 are small compared with their means, it suffices to linearize (A13) by using the Taylor's series expansion

$$\begin{aligned} \hat{\varphi} &= \tan^{-1} \frac{\mu_2 + (i_2 - \mu_2)}{\mu_1 + (i_1 - \mu_1)} \\ &\approx \tan^{-1} \frac{\mu_2}{\mu_1} + \frac{\mu_1(i_2 - \mu_2) - \mu_2(i_1 - \mu_1)}{\mu_1^2 + \mu_2^2} \\ &\approx \varphi + \frac{(i_2 - \mu_2) \cos \varphi - (i_1 - \mu_1) \sin \varphi}{\frac{1}{2} e \bar{a} m} \end{aligned} \quad (\text{A20})$$

so that

$$E[\hat{\varphi}] \approx \varphi \quad (\text{A21})$$

and, because of the independence of i_1 and i_2 ,

$$\begin{aligned} V[\hat{\varphi}] &\approx \frac{V[i_2] \cos^2 \varphi + V[i_1] \sin^2 \varphi}{(\frac{1}{2} e \bar{a} m)^2} \\ &\approx \frac{2}{m^2 \bar{a} T}. \end{aligned} \quad (\text{A22})$$

It then follows from (A14) that

$$E[\hat{r}] \approx \frac{\lambda_m}{4\pi} \varphi - \frac{r_0}{2} = r \quad (\text{A23})$$

and

$$V[\hat{r}] = \left(\frac{\lambda_m}{4\pi} \right)^2 V[\hat{\varphi}] \approx \frac{\lambda_m^2}{8\pi^2 m^2 \bar{a} T}. \quad (\text{A24})$$

But from (A4) and (6) we see that

$$\bar{a} T = \bar{n}_{pe} = (\text{SNR})^2. \quad (\text{A25})$$

Thus letting $\sigma_r^2 = V[\hat{r}]$, we obtain the desired final result:

$$\sigma_r \approx \frac{\lambda_m}{2\sqrt{2} \pi m \text{SNR}}. \quad (\text{A26})$$

APPENDIX B

VARIANCE OF SAMPLE MEAN OF CORRELATED DATA

In this appendix we determine the reduction in variance that can be obtained by averaging the correlated samples obtained from the output of a low-pass filter. The analysis is based on the following assumptions:

1) The filter input $x(t)$ is a zero-mean stationary white-noise process with a power density of N_0 W/(rad/s).

2) The filter is a single-pole linear filter with time constant T_f .

3) The filter output $y(t)$ is sampled at times t_1, t_2, \dots, t_n with a fixed time interval ΔT between samples.

Under Conditions 1 and 2, it is well known that the autocorrelation function $R_y(\tau)$ for the filter output is given by

$$R_y(\tau) = E[y(t)y(t+\tau)] = \frac{N_0}{2T_f} e^{-|\tau|/T_f}. \quad (\text{B1})$$

In particular, the variance of y , σ^2 , is given by

$$\sigma^2 = R_y(0) = \frac{N_0}{2T_f}. \quad (\text{B2})$$

Now, let $y_i = y(t_i)$ and let \bar{y} be the average of the n samples:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (\text{B3})$$

Since the mean of x is zero, the means of both y_i and \bar{y} are also zero. Thus the variance of \bar{y} , $\bar{\sigma}^2$, is given by

$$\begin{aligned} \bar{\sigma}^2 &= E[\bar{y}^2] - E^2[\bar{y}] \\ &= E\left[\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j\right] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n R_y(t_i - t_j) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{N_0}{2T_f} e^{-|i-j|\Delta T/T_f} = \frac{\sigma^2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \rho^{|i-j|} \end{aligned} \quad (\text{B4})$$

where

$$\rho = e^{-\Delta T/T_f}. \quad (\text{B5})$$

Now

$$\sum_{i=1}^n \sum_{j=1}^n \rho^{|i-j|} = 2 \sum_{i=1}^n \sum_{k=0}^{n-i} \rho^k - n \quad (\text{B6})$$

and, by using the well-known formula for the sum of a geometric series,

$$\sum_{i=0}^{m-1} \rho^i = \frac{1 - \rho^m}{1 - \rho} \quad (\text{B7})$$

we obtain

$$\begin{aligned} \bar{\sigma}^2 &= \frac{\sigma^2}{n^2} \left(2 \sum_{i=1}^n \frac{1 - \rho^{n-i+1}}{1 - \rho} - n \right) \\ &= \frac{\sigma^2}{n^2} \left(\frac{2n}{1 - \rho} - n - \frac{2\rho}{1 - \rho} \sum_{i=1}^n \rho^{n-i} \right) \\ &= \sigma^2 \left[\frac{1}{n} \frac{1 + \rho}{1 - \rho} - \frac{2\rho(1 - \rho^n)}{n^2(1 - \rho)^2} \right]. \end{aligned} \quad (\text{B8})$$

The behavior of $\bar{\sigma}^2$ can be clarified by introducing the average number N of samples during a time constant,

$$N = \frac{T_f}{\Delta T} \quad (\text{B9})$$

and the quantity

$$\beta = \frac{n}{N} = \frac{n\Delta T}{T_f} \quad (\text{B10})$$

which is the total observation time in units of time constants. From these definitions and (B5), we can write (B8) as

$$\frac{\bar{\sigma}^2}{\sigma^2} = \frac{1}{\beta N} \frac{1 + e^{-1/N}}{1 - e^{-1/N}} - \frac{2e^{-1/N}}{(\beta N)^2} \frac{1 - e^{-\beta}}{(1 - e^{-1/N})^2}. \quad (\text{B11})$$

This ratio decreases monotonically with N , asymptotically approaching the limiting value

$$\lim_{N \rightarrow \infty} \frac{\bar{\sigma}^2}{\sigma^2} = \frac{2}{\beta} \left(1 - \frac{1 - e^{-\beta}}{\beta} \right). \quad (\text{B12})$$

This limit is approached quite rapidly and, for practical purposes, is achieved with an N of four or five samples in a time constant. When the total observation time is not too large, sampling several times during a time constant produces a useful reduction in variance. If we define the effective number of samples n_{eff} by

$$\bar{\sigma}^2 = \frac{\sigma^2}{n_{\text{eff}}} \quad (\text{B13})$$

then (B12) yields

$$n_{\text{eff}} \approx \frac{\beta}{2} \frac{1}{1 - \frac{1 - e^{-\beta}}{\beta}} \quad (\text{B14})$$

However, for large β , n_{eff} approaches $\beta/2$. Thus when the total observation time extends by many time constants, little improvement is obtained over sampling once every two time constants.

ACKNOWLEDGMENT

We would like to express our appreciation to our colleagues at SRI, and most particularly to Dr. G. J. Agin, Ms. Phyllis Barrett, B. Meyer, Dr. C. A. Rosen, and Ms. Helen Wolf for their valuable contributions to this investigation.

REFERENCES

- [1] T. O. Binford and J. M. Tenenbaum, "Computer vision," *Computer*, vol. 6, pp. 19-24, May 1973.
- [2] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
- [3] *The Psychology of Computer Vision*, P. H. Winston, Ed. New York: McGraw-Hill, 1975.
- [4] Y. Shirai and M. Suwa, "Recognition of polyhedrons with a range finder," in *Proc. 2nd Int. Joint Conf. Artificial Intelligence* (London, England, Sept. 1971), pp. 80-87.
- [5] Y. Shirai, "A step toward context-sensitive recognition of irregular objects," *Comput. Graphics and Image Processing*, vol. 2, pp. 298-307, Dec. 1973.
- [6] D. Nitzan, "Scene analysis using range data," Artificial Intelligence Center, Stanford Research Institute, Menlo Park, CA, Tech. Note 69, Aug. 1972.
- [7] G. J. Agin and T. O. Binford, "Computer description of curved objects," in *Proc. 3rd Int. Joint Conf. Artificial Intelligence* (Stanford, CA, Aug. 1973), pp. 629-640.
- [8] R. Nevatia and T. O. Binford, "Structured descriptions of complex objects," in *Proc. 3rd Int. Joint Conf. Artificial Intelligence* (Stanford, CA, Aug. 1973), pp. 641-647.
- [9] F. Röcker, "Localization and classification of three-dimensional objects," in *Proc. 2nd Int. Joint Conf. Pattern Recognition* (Copenhagen, Denmark, 1974), pp. 527-528.
- [10] R. J. Popplestone et al., "Forming models of plane-and-cylinder faceted bodies from light stripes," in *Proc. 4th Int. Joint Conf. Artificial Intelligence* (Tbilisi, Georgia, USSR, Sept. 1975), pp. 664-668.
- [11] F. Röcker and A. Kiessling, "Methods for analyzing three-dimensional scenes," in *Proc. 4th Int. Joint Conf. Artificial Intelligence* (Tbilisi, Georgia, USSR, Sept. 1975), pp. 669-673.
- [12] T. D. Garvey, "An experiment with a system for locating objects in multisensory images," in *Proc. 3rd Int. Conf. Pattern Recognition* (Cornado, CA, Nov. 1976), pp. 567-575.
- [13] B. Horn, "Focusing," Project MAC Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Memo No. 160, May 1968.
- [14] P. M. Will and K. S. Pennington, "Grid coding: A preprocessing technique for robot and machine vision," in *Proc. 2nd Int. Joint Conf. Artificial Intelligence* (London, England, Sept. 1971), pp. 66-70.
- [15] B. Julesz, "Towards the automation of binocular depth perception," in *Proc. IFIP Congress 1962*. Amsterdam: North Holland, 1962, pp. 439-443.
- [16] D. N. Perkins, "Computer stereo vision: a combinatorial theory with implementation," Ph.D. dissertation, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, June 1970.
- [17] L. H. Quam, "Computer comparison of pictures," Stanford Artificial Intelligence Laboratory, Stanford Univ., Stanford, CA, Memo AIM-144, 1971.
- [18] M. D. Levine, D. A. O'Handley, and G. M. Yagi, "Computer determination of depth maps," *Comput. Graphics and Image Processing*, vol. 2, pp. 131-150, Oct. 1973.
- [19] D. A. O'Handley, "Scene analysis in support of a Mars rover," *Comput. Graphics and Image Processing*, vol. 2, pp. 281-297, Dec. 1973.
- [20] M. J. Hannah, "Computer matching of areas in stereo images," Stanford Artificial Intelligence Laboratory, Stanford Univ., Stanford, CA, Memo AIM-239, July 1974.
- [21] S. Ganapathy, "Reconstruction of scenes containing polyhedra from stereo pair of views," Stanford Artificial Intelligence Laboratory, Stanford Univ., Stanford, CA, Memo AIM-272, Dec. 1975.
- [22] P. E. Hart et al., "Artificial intelligence—research and applications," Annual Technical Report to ARPA, Stanford Research Institute, Menlo Park, CA, Contract DAHCO4-72-C-0008, Dec. 1972.
- [23] D. Nitzan, "Stereopsis error analysis," Artificial Intelligence Center, Stanford Research Institute, Menlo Park, CA, Tech. Note 71, Sept. 1972.
- [24] A. R. Johnston, "Infrared laser rangefinder," Jet Propulsion Laboratory, NASA New Technology, Pasadena, CA, Rep. NPO-13460, Aug. 1973.
- [25] W. K. Pratt, *Laser Communication Systems*. New York: Wiley, 1969.
- [26] M. Ross, *Laser Receivers*. New York: Wiley, 1966.
- [27] P. Moon, *The Scientific Basis of Illuminating Engineering*. New York: Dover Publications, 1961.
- [28] W. W. Wendlandt and H. G. Hecht, *Reflectance Spectroscopy*. New York: Interscience Publishers, 1966.
- [29] D. E. Maxwell, "A 5 to 50 MHz direct-reading phase meter with hundredth-degree precision," *IEEE Trans. Instrum. Meas.*, vol. IM-15, pp. 304-310, Dec. 1966.
- [30] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828-842, July 1972.
- [31] C. A. Rosen, "Combined ranging and color sensor," U.S. Patent 3 945 729, Mar. 23, 1976.
- [32] G. J. Agin, private communication.
- [33] I. Sobel, "Camera models and machine perception," Stanford Artificial Intelligence Laboratory, Stanford Univ., Stanford, CA, Memo AIM-121, May 1970.
- [34] A. Rosenfeld and J. L. Pfaltz, "Distance functions on digital pictures," *Pattern Recognition*, vol. 1, pp. 33-61, July 1968.
- [35] I. Bar-David, "Communication under the Poisson regime," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 31-37, Jan. 1969.
- [36] G. F. Herrmann, "Optimum versus suboptimum detection under the Poisson regime—Application to optical communications using biphasic and quadriphase modulated subcarriers," *IEEE Trans. Commun.*, vol. COM-21, pp. 800-809, July 1973.

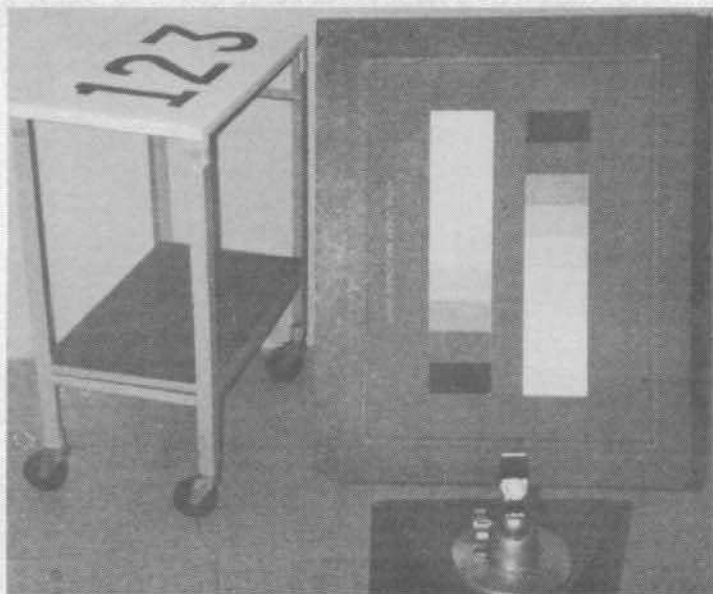


Fig. 6. Photograph of a test scene.

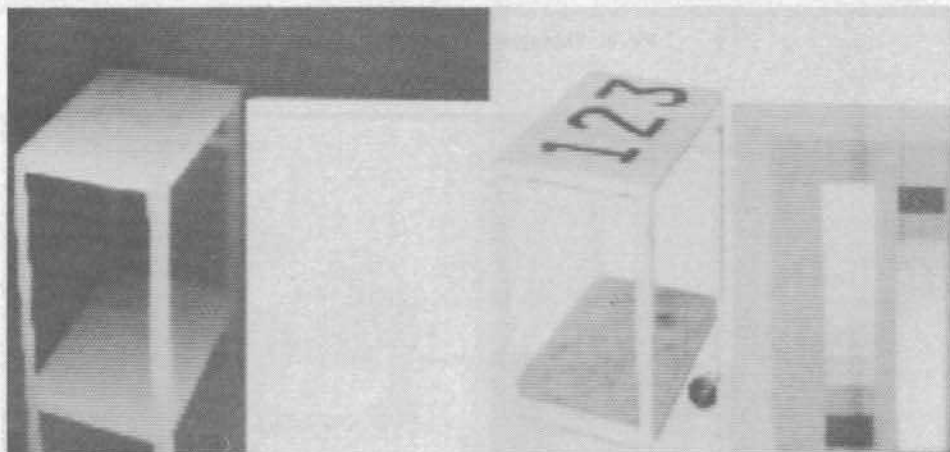




Fig. 8. Photograph of an office scene.

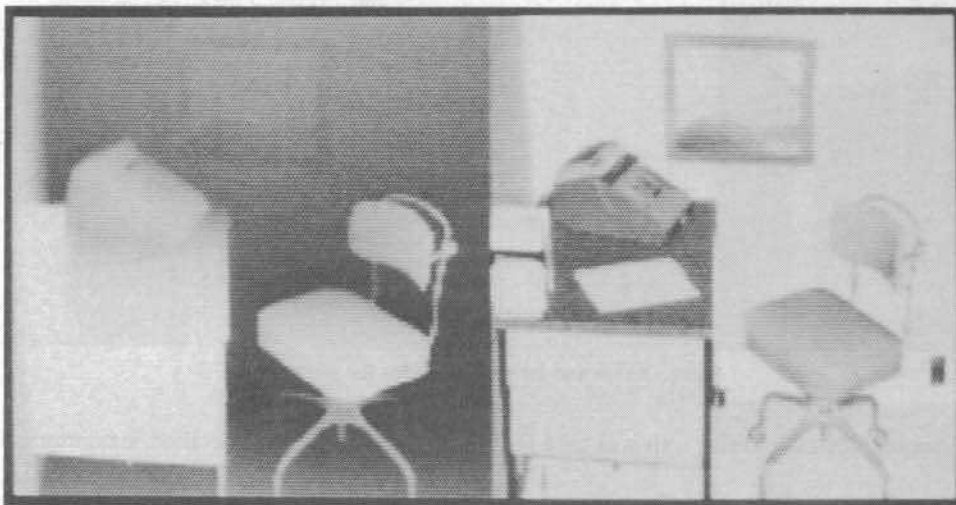


Fig. 9. Range and intensity data for the office scene.

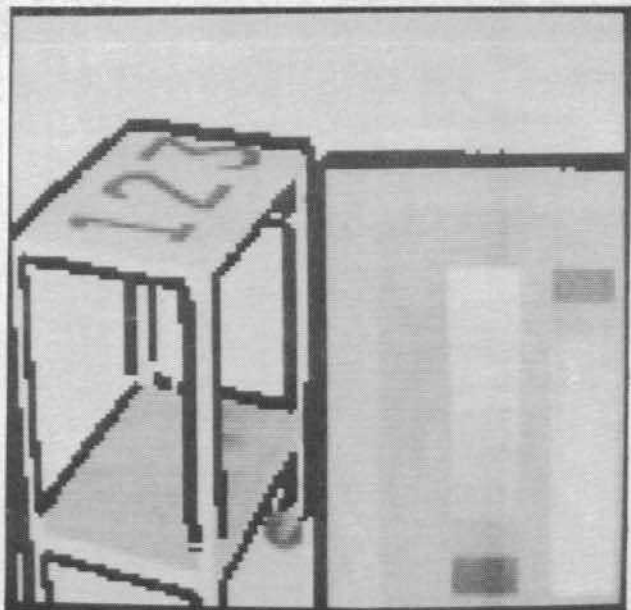


Fig. 10. Jump boundaries for the test scene.

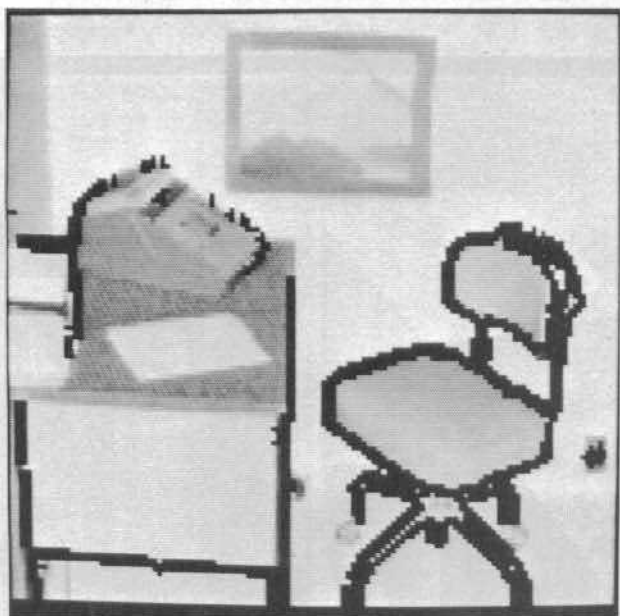


Fig. 11. Jump boundaries for the office scene.

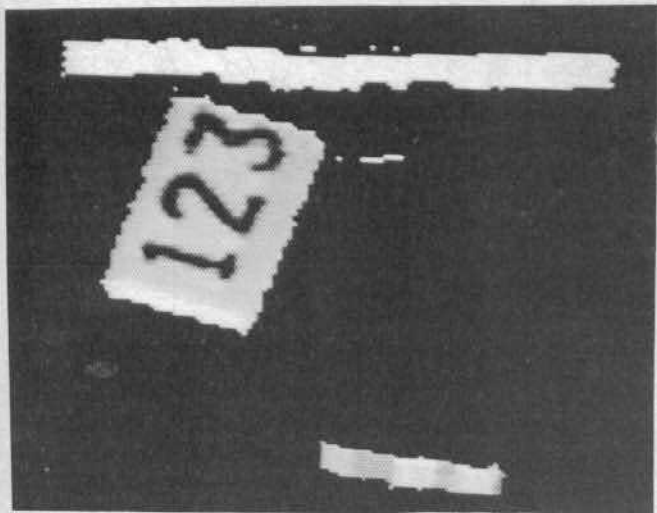


Fig. 13. Normal view of top of cart.

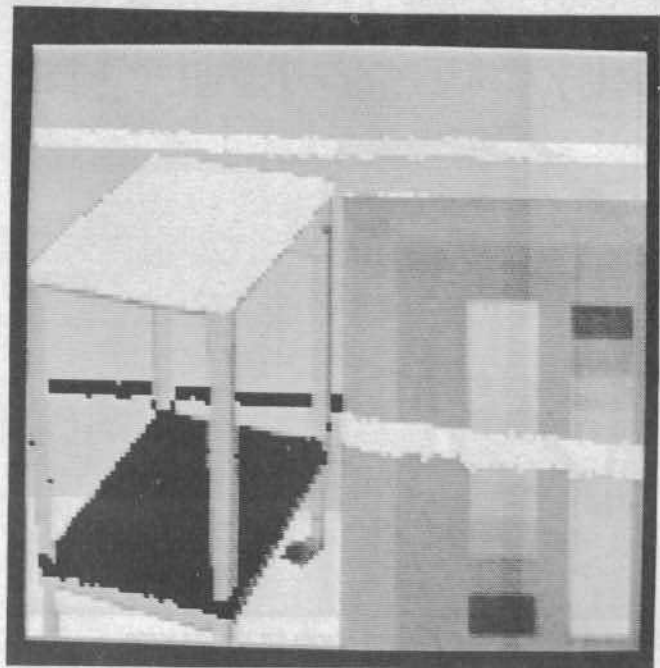
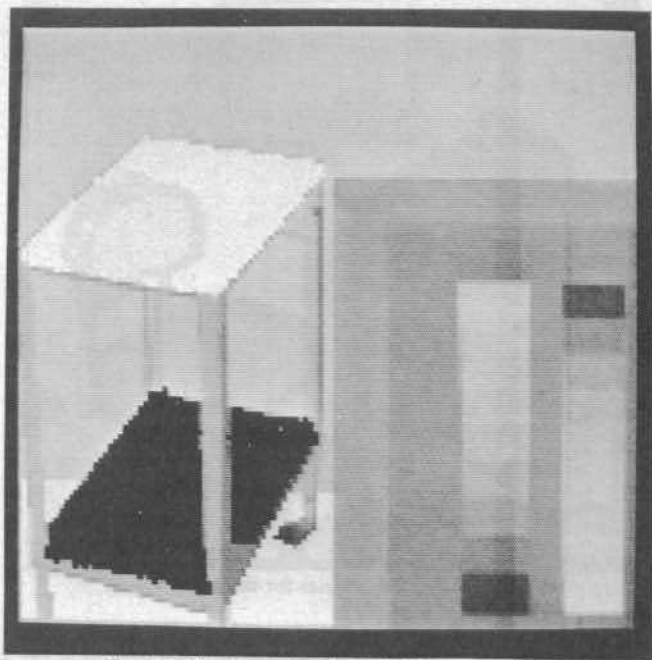


Fig. 14. Candidate horizontal surfaces for the test scene.



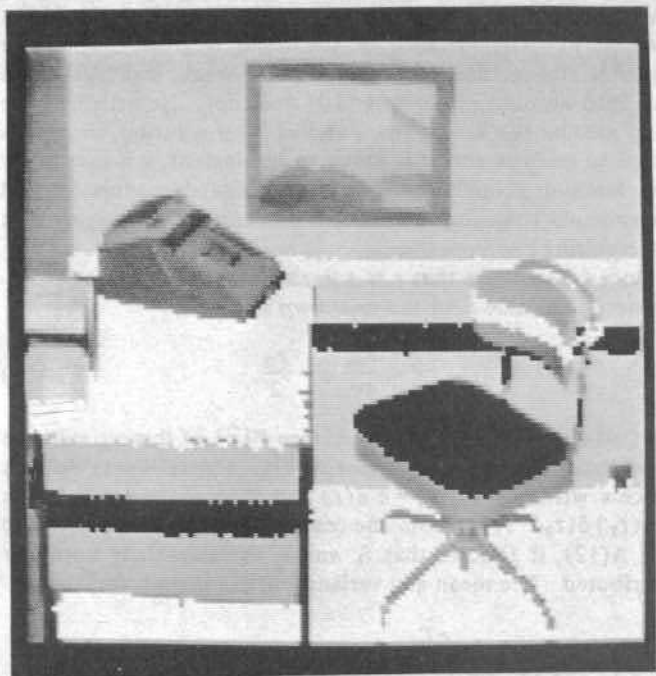


Fig. 16. Candidate horizontal surfaces for the office scene.

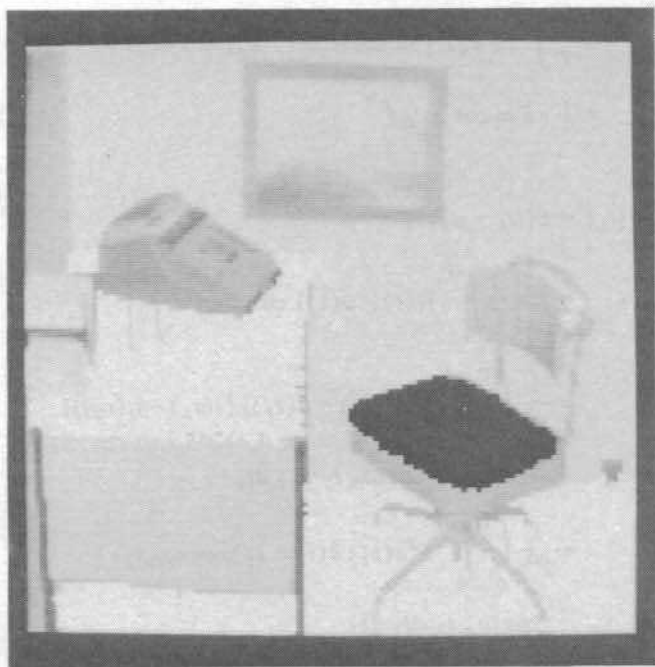


Fig. 17. Horizontal surfaces in the office scene.