from the Proceedings of the 1977 National Computer Conference
Vol. 46, pgs. 49-53.

A system for automatic acquisition of three-dimensional data

by HENRY FUCHS, JOE DURAN and BRIAN JOHNSON
The University of Texas at Dallas
Richardson, Texas

ABSTRACT

This paper presents the design of a three-dimensional data acquisition system based on multiple, single-dimensional optical sensors. The system can operate in any of three modes:

(1) the tracking of multiple, independent, point light sources
(2) the automatic digitization of opaque surfaces
(3) the real-time tracking of an unmarked moving object (e.g., tip of user's hand).

The design offers such advantages as a lensless sensing system, a minimum reliance on analog measurements, an ease of upgrading to higher precision measurements, an ease of portability, an adjustable field of view, and the ability to operate under normal ambient light conditions. A network of microprocessors is incorporated to minimize processing delays and thus increase data acquisition rates. In its initial application the system will digitize the cranio-facial surfaces of candidates for reconstructive surgery.

INTRODUCTION*

Recent developments in charged coupled device (CCD) technology have made generally available a variety of IC chips which contain a linear array of light-sensitive elements (Figure 1). Such an array can be treated in the system as a digital shift register in which successive clock pulses cause serial output of the entire array of values. (The sole complication, that the serial output is analog, is easily overcome with a single analog-to-digital converter.) Each of the individual values in the array is a function of the number of photons which have been absorbed by the associated cell in the array since the previous scanning of the shift register's contents. If the CCD chip and the analog-to-digital converter are considered as a single unit, then a sequence of clock pulses will obtain an array of *digital* values which define, as a function of distance along the array axis, the amount of light striking the CCD chip.

To build a one-dimensional sensor for 3-D input, we place in front of the CCD array a knife-edge in a plane parallel to the CCD array face with the line of the knife edge boundary perpendicular to the line of array elements (see Figure 2). If the environment contains a point light source which is brighter than the ambient light level, then the knife edge will cause a shadow to be cast somewhere along the line of light-sensitive elements.

MATHEMATICAL FORMULATIONS

With a number of such detector units placed around the room, the location of a light source can be determined by the taking of a single measurement, h_i, at each linear sensor, i, which can "see" the source. Each h_i measurement defines a plane in which the light must be located. When three such planes are defined, each containing the same light source, the location is uniquely determined (Figure 3). The use of homogeneous matrices considerably simplifies the mathematical formulations (c.f. Reference

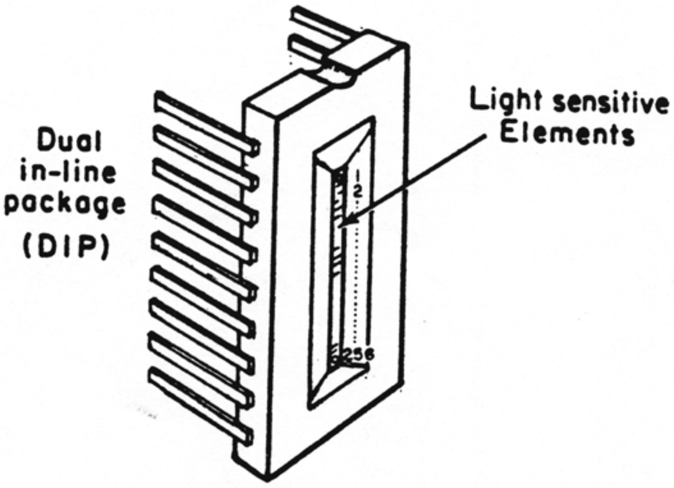

Figure 1—CCD linear sensing array chip

---

* Appendix A includes a short discussion of previous digitization devices and their effect on the design presented in this paper.

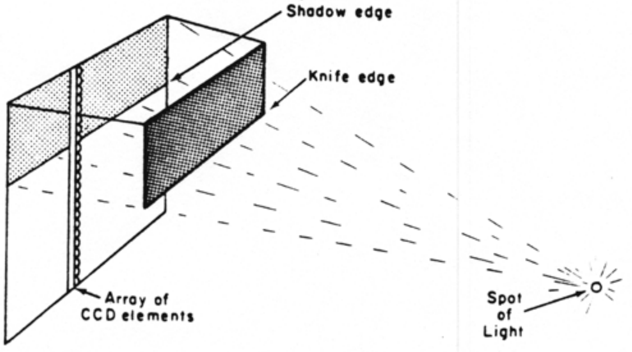

Figure 2—Basic sensor design

11). We show below that these simplifications allow the problem of computing the coordinates of a point light source to be stated as the problem of solving a system of planar equations of the form

$$a(h)X+b(h)Y+c(h)Z=d(h), \qquad (1)$$

where h is the position of the shadow edge (Figure 4). X, Y, Z are our system coordinates, and the coefficients a, b, c, d are linear expressions in h.

It is well-known (c.f. Reference 6) that an appropriate 4×4 matrix can be used to transform a 3-D point between 2 coordinate systems. For a single sensor, i, the transformation from room to sensor coordinates (see Figure 4) can be expressed by the 4×4 matrix, C^i, such that

$$(X\ \ Y\ \ Z\ \ 1)[C^i]=(x_i'\ \ y_i'\ \ z_i'\ \ w_i), \qquad (2)$$

Where $x_i=\frac{x_i'}{w_i}$, $y_i=\frac{y_i'}{w_i}$, and $z_i=\frac{z_i'}{w_i}$ are the actual sensor coordinate values. For the ith sensor, h_i, r_i, and the coordinate components, x_i and y_i, of the light source in the sensor coordinate system are related by the expression,

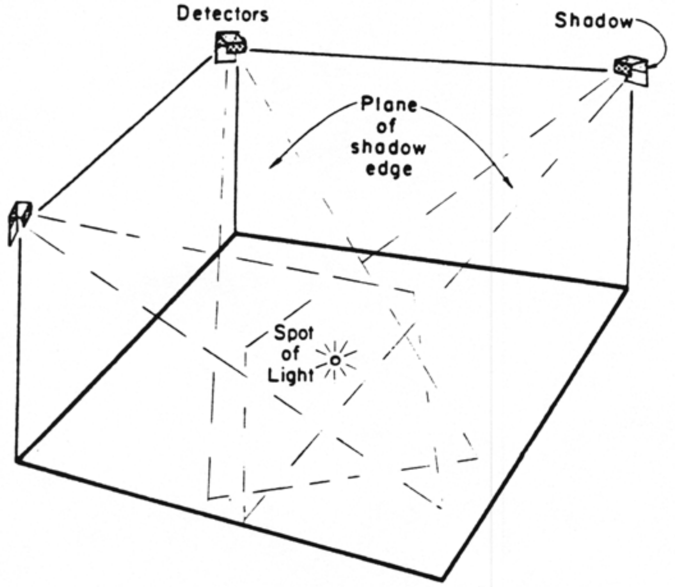

Figure 3—(3-D position from) 3 linear sensors

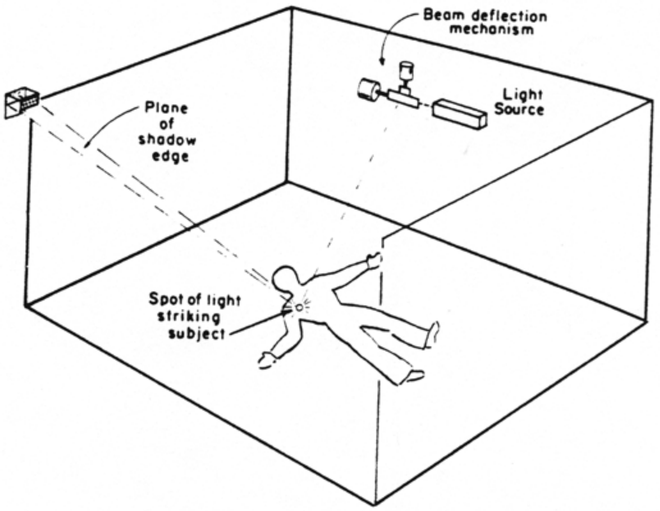

Figure 4—Sensor/room coordinate relationship

$$\frac{h_i}{-r_i}=\frac{y_i}{x_i}=\frac{C_{12}{}^iX+C_{22}{}^iY+C_{32}{}^iZ+C_{42}{}^i}{C_{11}{}^iX+C_{21}{}^iY+C_{31}{}^iZ+C_{41}{}^i}$$

Simplifying by collecting factors of X, Y, and Z yields

$$(h_iC_{11}{}^i+r_iC_{12}{}^i)X+(h_iC_{21}{}^i+r_iC_{22}{}^i)Y+(h_iC_{31}{}^i +r_iC_{32}{}^i)Z+h_iC_{41}{}^i+r_iC_{42}{}^i=0$$

One of the transformation matrix elements can be eliminated by dividing through by, say $r_iC_{42}{}^i$. This yields

$$(h_ip_{1i}+p_{2i})X+(h_ip_{2i}+p_{4i})Y+(h_ip_{5i}+p_{6i})Z+h_ip_{7i}+1=0, \qquad (3)$$

where $P_{1i}=\frac{C_{11}{}^i}{r_iC_{42}{}^i}$, $P_{2i}=\frac{C_{12}{}^i}{C_{42}{}^i}$, etc.

Eq. (3) is of the form of eq. (1) and is the equation of a plane in room coordinates. The light source lies in this plane. From three or more such planar equations, the unique location (X_1, Y_1, Z_1) of the light source in room coordinates can be calculated, either by standard matrix solution techniques or by least squares methods. If the p_{ji} are known for each sensor i, the h_i values measured for each of three or more sensors will, by eq. (3), uniquely determine the light source location.

Our sensors can be calibrated quite simply by determining the p_{ji}. Given the known position of seven light sources in room coordinates and the h_i value generated by each source at each sensor, we obtain a linear system of seven equations in the seven unknowns $p_{1i}, p_{2i}, \ldots, p_{7i}$ for each sensor. The seven points should be chosen so as not to give an ill-conditioned or singular matrix. The probability of such problems can be reduced by the use of more than seven calibration points, allowing a least squares solution.

Once the p_{ji} are calculated for each sensor, the location of an unknown point can be found by solving the linear system

$$\begin{pmatrix} a_i & b_i & c_i \\ a_j & b_j & c_j \\ a_k & b_k & c_k \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} d_i \\ d_j \\ d_k \end{pmatrix} \qquad (4)$$

(where $a_i = h_i p_{1i} + p_{2i}$, $b_i = h_i p_{3i} + p_{4i}$, etc.) for any three sensors i, j, k which see the light source. If more than three sensors see the source, error smoothing via a least squares solution of an overdetermined system is possible. If the values of $a(h_i)$, $b(h_i)$, $c(h_i)$, and $d(h_i)$ for each sensor are stored in advance for each possible h_i value (quite feasible, since h can take on 256 discrete values for our current sensor), then the coordinates of a point may be calculated with about 20 or so multiplications and divisions. The exact number depends upon such factors as the number of precalculations stored and the number of excess sensor equations which may be desired for error smoothing. This processing load is light enough so that the system can operate at sensor speeds with only a minicomputer serving as the central host.

With the use of a computer-deflected light beam, only one sensor needs to detect the light reflected from the surface being digitized, since the position of the light beam source in the environment will be previously determined and its orientation (deflection in two axes) will be under computer control. The plane and line measurements thus determine the 3-D location of the point on the target surface which is illuminated by the narrow beam of light (Figure 5). This might best be implemented by defining the line of the light beam in terms of the equation (in the form of eq. (3)) of two planes which intersect along the line. The sensor measurement thus defines the third necessary plane and the computational method described above may be used. This has the advantage of readily allowing a least squares error smoothing treatment if the use of more than one linear sensor is desired.

With some sacrifice in speed, it is possible to use the reflected light beam mode for interactive input applications. In this case the light beam deflection would be controlled by a real-time program which would acquire and track the object of interest (e.g., the tip of the user's hand) in a manner similar to a stage spotlight following a star performer.

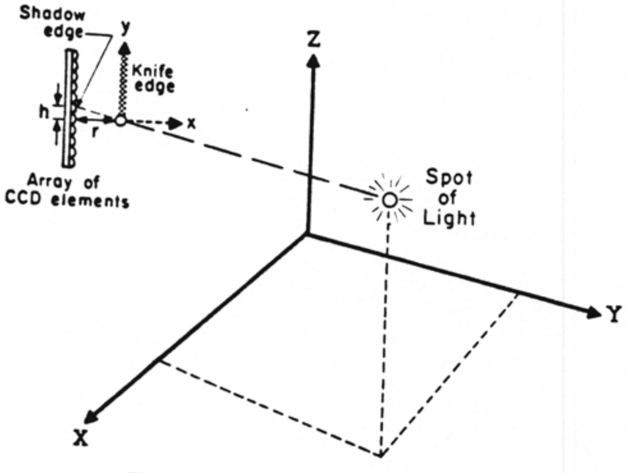

Figure 5—Automatic surface digitization mode

SYSTEM ARCHITECTURE

The current system layout has a microprocessor associated with each CCD array. The microprocessor is used to control the integration time of the CCD array (the amount of time it is allowed to sense), to collect the data from the array, and to compute the element position value (h_i) of the shadow boundary. Since current CCD arrays suffer from background noise level variations from cell to cell, individual cell calibration improves sensor performance. The microprocessor records these variations during an initial calibration phase and uses them to improve the shadow boundary determination.

The microprocessors are all connected to a host CPU, a small minicomputer with a minimal floating point capability (currently an HP 2100). The host computes the source point coordinates by solving the system of linear equations (4).

In the single point (or "wand") acquisition mode, the host stores the computed coordinates and waits for the next input point. It also carries out various control functions such as starting and stopping the process, changing the acquisition rate, etc. With multiple light source input, the host fires the sources at known times and keeps track of the association between coordinate computations and the corresponding sources. This mode allows tracking of not only the position, but also the rotational orientation of a moving object by the tracking of the position of three noncollinear light sources attached to the object.

In the light beam scanning mode, the host also must calculate coordinates for the scanning beam and deflect the beam to the desired orientation. The host polls the sensors, computes the coordinates of the reflected light spot, deflects the beam to the next desired position, then polls the sensors again, continuing in this manner until data acquisition is completed.

SENSITIVITY AND ACCURACY

Our prototype sensor uses a 256-element Fairchild model 110 CCD array. The measured effective sensitivity of the array is about 10^{-5} watts/cm^2. This indicates that a uniformly radiating point source of .1 watt can be sensed up to about 40 cm. (In practice we have been able to do slightly better, even without resorting to individual cell calibration.) We have sensed a 2 watt source at a distance of about 2 meters, in a room which was darkened but with still enough light for reasonable human vision.

The sensing distance for a projected spot depends both on the amount of energy projected onto the target surface and the scattering coefficient of this surface. Preliminary experiments indicate that the current sensor will enable digitization of human heads but not automatic scanning of

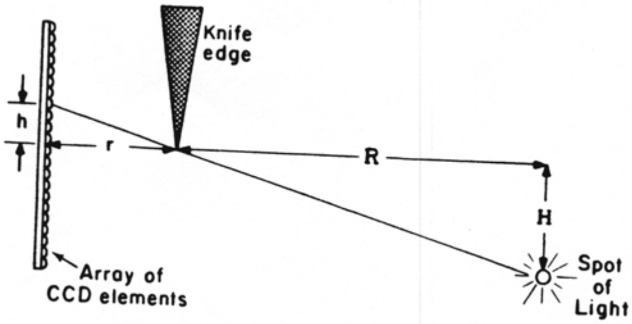

Figure 6—Sensor resolution distances

entire bodies, without resorting to high-powered light sources. We expect continual improvement in CCD sensitivity and can use other sensor technologies, such as arrays of individual photo-transistors or image intensifier devices, if desired.

The sensor's resolution is of course limited by the number of cells in the CCD array. Thus, the 256-element sensors yield a resolution of about .6 cm in a 1 cubic meter working volume. The smaller the working volume, the better the absolute resolution limit. The resolution depends upon the number of CCD elements and the relative distances of the CCD array, knife edge and source (see Figure 6). The absolute resolution in room coordinates is related to CCD element width, Δh, by the relation $\Delta H = \frac{R}{r}\Delta h$.

Accuracy increases can be realized by using sensors with more elements (sensors with over 1000 elements are currently available) or using more sensors to segment a desired working volume into several contiguous, smaller volumes.

CURRENT STATE OF IMPLEMENTATION

A prototype sensor system has recently been completed and interfaced to an HP 2100 computer. Figure 7 illustrates the output from the linear sensor array. A sharper distinction between the outputs of the shadowed and non-shadowed elements can be achieved by substituting a smaller light source.

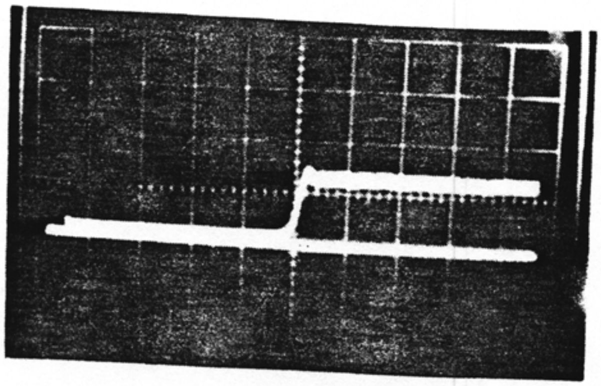

Figure 7—Output from sensor array

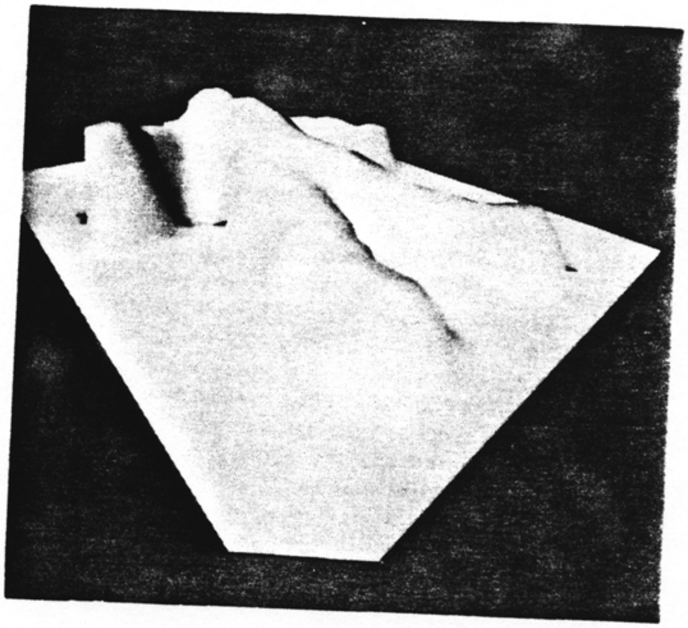

Figure 8—Reconstructed surface from laser scan

The data for Figure 8 was acquired by Fuchs,[3] by scanning a human model with a laser beam. The original Twinkle Box system detected the individual reflected spots of lights and calculated their three-dimensional positions.

CONCLUSION

We have presented a system which allows the use of very simple linear sensors to obtain 3-D information. These sensors involve no optical lensing systems or mechanical movement, and their use in the system leads to simple, tractable mathematical calculations. Although our current system was inspired by and uses linear CCD arrays, any sensor system capable of reporting values of a shadow edge position along one dimension can be used, with little or no changes in the total system. More expensive sensors may be used to increase sensitivity and/or resolution as needed.

It is hoped that this new system will make acquisition of 3-D information significantly faster, easier and less costly, thereby encouraging an expansion of 3-D man/machine interaction.

ACKNOWLEDGMENT

The authors are grateful for ideas contributed by Larry Evans of Evans and Sutherland Computer Corp., Salt Lake City, at the inception of this research project. Gordon McNorton assisted with the sensor/computer interfacing.

REFERENCES

1. Binford, T. O., "Visual Perception by Computer," *Proc. of the IEEE Conf. on Systems and Control*, Miami, Dec. 1971.

2. Burton, R. P. and I. E. Sutherland, "Twinkle Box—A Three-Dimensional Computer Input Device," *AFIPS Conference Proceedings*, Vol. 43, May 1974.
3. Fuchs, H., "The Automatic Sensing and Analysis of Three-Dimensional Surface Points from Visual Scenes," Ph.D. thesis, Univ. of Utah, Salt Lake City, August, 1975.
4. Gara, A. D., R. F. Majkowski, and T. T. Stapleton, "A Holographic System for Automatic Surface Mapping," General Motors Research Labs., Research Publication GMR-1342, Warren, Mich. March 1973.
5. Herron, R. E., "Biostereometric Measurement of Body Form," *Yearbook of Physical Anthropology*, Vol. 16, 1972, pp. 80-121.
6. Newman, W. M. and R. Sproull, *Principles of Interactive Computer Graphics*, McGraw-Hill, New York, 1973.
7. Roberts, L. G., "The Lincoln Wand," MIT Lincoln Laboratory Report, Lexington, Mass., June 1966.
8. Science Accessories Corp., "Graf/Pen Sonic Digitization," Science Accessories Corp., Southport, Conn., 1970.
9. Selective Electronic Company AB, "Selspot System," Molndal, Sweden.
10. Spreight, B. S., C. A. Miles, and K. Moledina, "Recording Carcass Shapes by a Moire Method," *Medical and Biological Engineering*, March 1974, pp. 221-226.
11. Sutherland, I. E., "Three-Dimensional Data Input by Tablet," *Proc. of the IEEE*, Vol. 63, No. 4, April 1974, pp. 453-461.

APPENDIX A—BACKGROUND

The widespread interest in acquiring 3-D information has stimulated the development of a variety of methods and systems. These systems can be loosely categorized into two general classes:

I. those systems which digitize the location of a special position-indicating device ("wand"), which is moved manually to specific points of interest, and
II. those systems which digitize (manually or automatically) an entire surface of interest.

Examples of type I systems include the early Lincoln Wand of Larry Roberts[7] and the current commercial product, the Graf-Pen.[8] Both systems digitize the position of a small ultra-sonic transmitter located at the tip of the wand. More recently, the Twinkle Box of Burton and Sutherland[2] and the Swedish commercial product, Selspot,[9] are both able to digitize a multiple number of points of light concurrently, using time-division multiplexing.

Type II systems, being usually more complex than type I, tend to be laboratory developments oriented toward a specific application: the digitization of automobile bodies using holographic techniques,[4] the measurement of the surfaces of animal carcasses by interference methods,[10] the scanning of small object clusters for automatic scene analysis using optical triangulations,[1] and the digitization of human body contours by classic stereometric procedures.[5]

This last application—the digitization of human body form—seems to be of greatest current interest. In this area, 3-D surface digitization has been used for such diverse studies as anthropomorphic dummy evaluation, spinal deformation measurements, and tumor detection and tumor growth monitoring. However, the presently implemented and proposed digitization methods are incompletely automated, and tend to be cumbersome and expensive, thus requiring much human intervention and effort. Additional processing steps, such as the transfer of images to photographic materials, often cause additional delays and expense.

Wand-like devices are also inconvenient for surface digitization in that the wands have to be moved by the user to every point to be measured. This can be a very time consuming task, sometimes made impossible if the surface to be digitized moves too rapidly. However, wand-like devices are particularly useful when real time tracking of only a few points is needed. For example, in limb mobility measurements, "wand tips" (small light sources) placed on a subject's leg allow leg movement to be tracked in real time.

The most direct predecessor of our system is the aforementioned Twinkle Box. We have attempted to retain most of its capabilities as well as overcome its major flaws. We consider the major assets of the Twinkle Box to be

(1) the digitization of multiple points,
(2) the use of arbitrarily positioned single-dimensional sensors instead of two-dimensional devices, such as TV cameras,
(3) a simple, unified approach to the problem of converting raw sensor measurements into a 3-D position in Cartesian coordinates.

The Twinkle Box's limitations, as reported by Burton and Sutherland, can be primarily attributed to the basic electro-mechanical sensor design. The sensors consist of spinning discs sandwiched between an optical lens and a photomultiplier tube. The system's four sensor units, each with a 22-inch diameter disc spinning at 3500 rpm, create a very noisy working environment, and the 2 h.p. motors needed to spin these discs severely overheat the sensors after about 30 minutes of operation. The disc fluctuations and non-synchronous rotations cause additional difficulties.

Our system eliminates the mechanical and optical problems by using an entirely new sensor design. The system also adds two new features (1) a focused light beam deflection system to enable automatic surface digitization and the tracking of unmarked objects; and (2) a network of microprocessors for signal pre-processing and other parallel computations. These features allow our system to be used both in the "wand" tracking mode of type I systems and also the automatic-surface-digitization mode of type II systems.