

Computergraphic modeling and analysis: A portable system for tracking arm movements in three-dimensional space

HOWARD POIZNER

The Salk Institute for Biological Studies, San Diego, California

EDWARD WOOTEN

University of California, San Diego, California
and The Salk Institute for Biological Studies, San Diego, California

and

DOUGLAS SALOT

Arium Corporation, Los Angeles, California
and The Salk Institute for Biological Studies, San Diego, California

A system is presented that allows automated, three-dimensional tracking of hand and arm movements. The system incorporates commercially available optoelectronic cameras and provides portable and affordable, yet accurate, three-dimensional monitoring of multiple joints of the hands and arms. Special-purpose hardware components were developed, as was software for data acquisition, data processing, and graphic display. The hardware and software are described, along with such necessary procedures as system calibration and transformation of coordinate system frames of reference. Testing of the system revealed highly accurate three-dimensional spatial tracking. The three-dimensional numeric and graphic analyses of movement made possible by this system allow new studies into the nature of the neural control of movement.

That patterns of movement and spatial contouring alone are sufficient to convey many grammatical processes in American Sign Language has been demonstrated by presenting sign forms as patterns of moving points of light, following Johansson's (1975) classic procedures for isolating biological motion (Poizner, 1981, 1983, in press). Clearly, to understand the structure underlying this linguistic movement system and the biological constraints operating on it, it is necessary to monitor and analyze movement in three-dimensional space. Likewise, models of motor control increasingly rely on detailed analysis of kinematic patterns of movement displacement, velocity, and acceleration over time (Atkenson & Hollerbach, 1985; Kay, Munhall, Bateson, & Kelso, 1985), and thus require three-dimensional movement measurement and analysis.

Traditional methods of movement measurement have utilized video recording or high-speed cinematography. In these methods, the positions of multiple markers on the image from a camera are digitized frame by frame

by a human operator. This is a notoriously laborious and error-prone procedure (Woltring, 1984). Furthermore, the use of a single camera provides only two-dimensional information; the use of multiple cameras presents problems of camera synchronization.

There are commercially available systems that directly sense the positions of infrared-emitting diodes (LEDs) that can be attached to the body. One such system, the SELSPOT system of Selcom, has been fairly widely used for three-dimensional movement measurement (Antonsen, 1978, 1982; Atkenson & Hollerbach, 1985; Conati, 1977; Reece, 1982; Woltring, 1974, 1984). Although this system, with appropriate software, can provide highly accurate, automated three-dimensional data collection, it is quite expensive and is not portable, a drawback for many types of research. In this paper, we present the results of our development of a new system of automated movement measurement, a system that is affordable and portable and offers high spatial and temporal resolution for tracking up to eight moving body segments in three-dimensional space.

This work was supported in part by National Science Foundation Grant No. BNS83-09860 to The Salk Institute for Biological Studies, by National Institutes of Health Grant No. NS20204-02 to the University of Florida, and by National Institutes of Health Grants NS19096 and HD-13249 to the Salk Institute. We thank Nurit Weiss for her programming assistance and Richard Batch for his help throughout.

Please address reprint requests to: Howard Poizner, The Salk Institute, P.O. Box 85800, San Diego, CA 92138.

HARDWARE COMPONENTS

Figure 1 presents the main hardware components and shows the positioning of four LEDs on the major joints of a subject's arm. The main hardware components con-

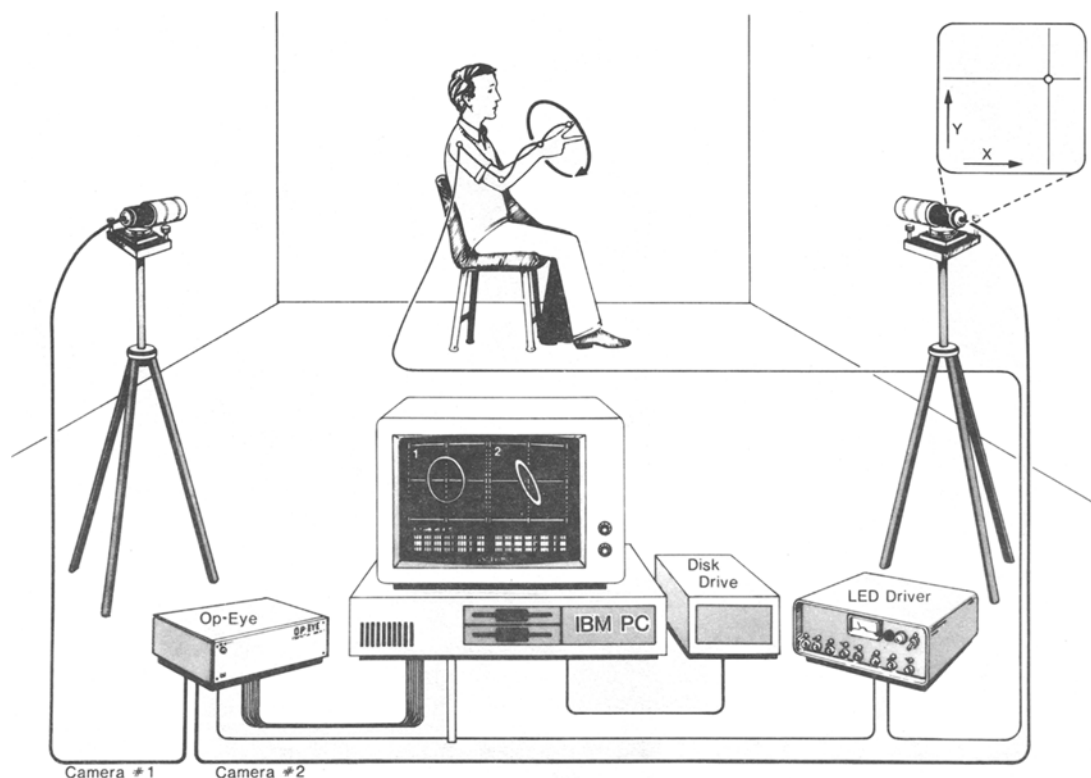


Figure 1. Main hardware components and the positioning of LEDs on a subject.

sist of Op-Eye cameras and amplifying electronics (United Detector Technology, Op-Eye 5 System)¹, together with an analog-to-digital (A/D) converter (Data Translation DT-2801); specially built LED drive electronics that pulse and time multiplex up to eight LEDs under computer control; and an IBM-PC microcomputer. The Op-Eye cameras are optoelectronic and directly sense the position of LEDs. An optical lens focuses light from an LED onto a detector plate (a lateral effect photodiode). The spot of light generates four output currents on the photodiode that are proportional to the position of the spot on the detector surface. The four output currents are then amplified by the Op-Eye amplifying electronics and digitized to 12-bit (1 part in 4096) resolution.

The Op-Eye system comes from the manufacturer with only the capability to continuously drive a single LED. However, several LEDs must be driven separately to measure the positions of multiple points on the hand and arm. To provide unique LED identification, we built an electronic module (LED driver) to rapidly pulse and sequence up to eight LEDs. Under control of the IBM-PC, the LED driver flashes one LED at a time in rapid sequence to provide good temporal resolution. Individual LED intensities may be controlled and monitored. High output LEDs were selected (Opto-Diode Corporation OD-100). These diodes have 180° divergence and radiate at 880 nm, closely matching the peak sensitivity of the detectors at 900 nm. An input/output interface card was built to allow the IBM to pass a sequence of control signals to the LED driver and to the A/D converter (see Figure 2).

Software

Data acquisition and display from the Op-Eye cameras was accomplished using a series of menu-driven modular routines. The programs allow for direct memory access transfers from the A/D converter, synchronization of the pulsing of the LEDs with the digitizing of the camera signals, and display of the x-y information provided by the Op-Eye cameras. The sampling frequency and the number of LEDs are selectable for any given measurement.

The IBM-PC synchronizes the digitizing of the camera outputs with the LED pulsing through a programmable timer (Intel 8253), which controls the sequence of events involved with turning on an LED and then sampling from the cameras. Figure 3 presents a timing diagram for the

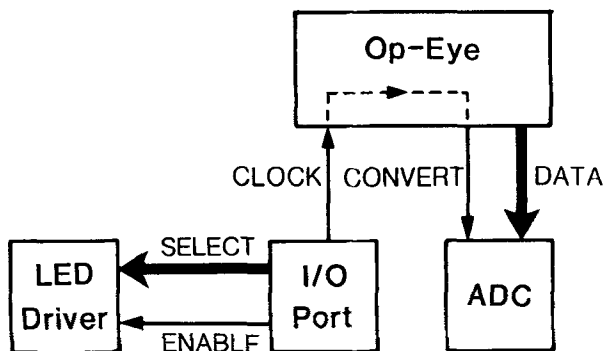


Figure 2. Control signals. During data acquisition, signals at the I/O port control the pulsing of LEDs and digitizing of Op-Eye data.

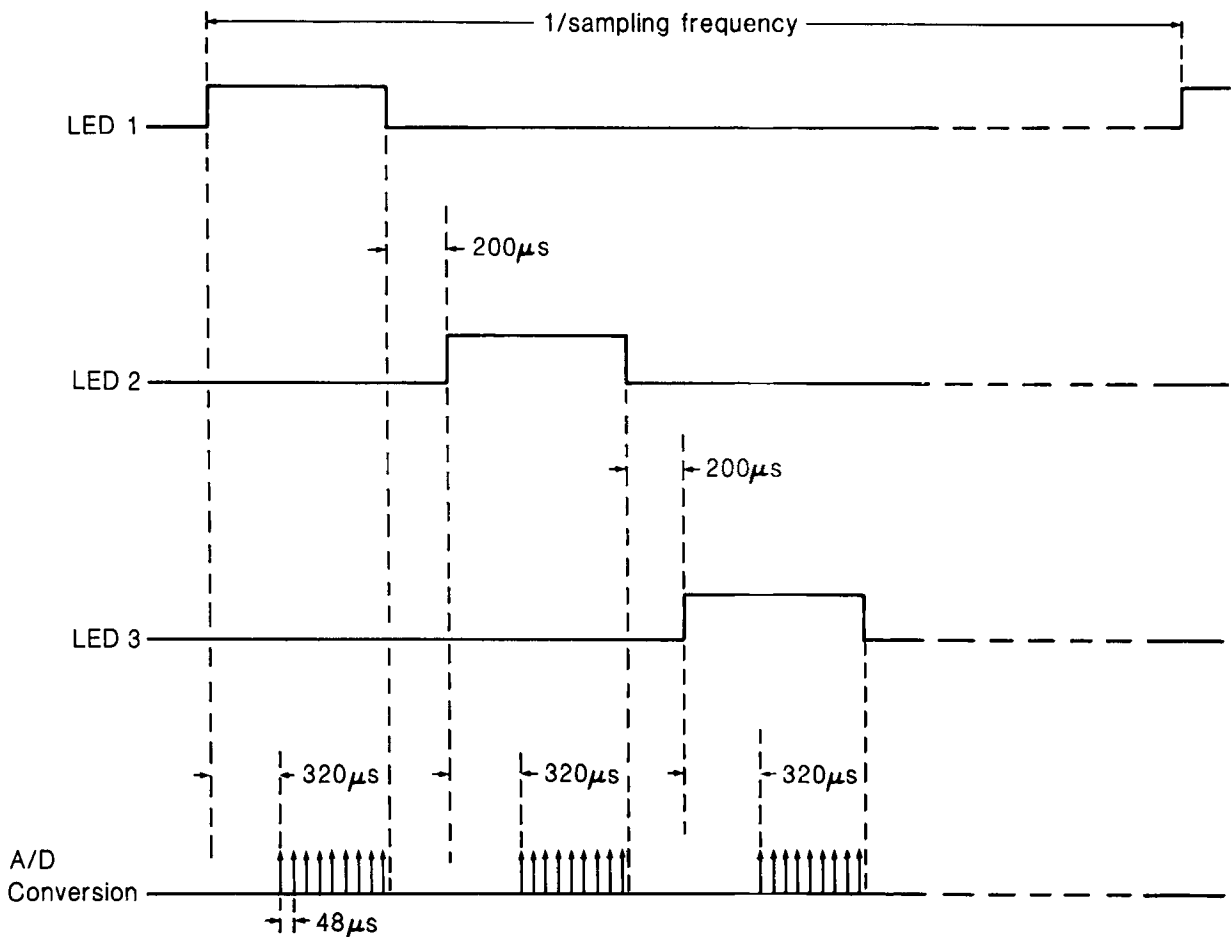


Figure 3. Timing for a two-camera, three-LED setup.

sequence of events in one sampling period. An initial clock interrupt signal is used to synchronize the control circuitry on the A/D converter board with the external clock. Data are collected by a scan of clocked A/D conversions. The first LED is turned on, the electronics are allowed to settle for $320\ \mu\text{sec}$, and then, in a two-camera set-up, eight A/D conversions are taken at intervals of $48\ \mu\text{sec}$. The first LED is then turned off, a delay of $200\ \mu\text{sec}$ follows to allow the detector to clear, and the sequence is repeated for all the LEDs. This entire sequence of events is then repeated at the specified sampling rate.

Data are acquired by the IBM-PC using its Direct Memory Access (DMA) channels in order to write rapidly to disk. The DMA buffer is dumped to either RAM disk for applications requiring extremely high sampling rates and large number of LEDs, or directly to the hard disk for our normal acquisition using four LEDs sampled at 100 Hz.

We tried to minimize the time each LED is turned on during a data collection cycle so as to pulse the LEDs with relatively high currents, thus improving the signal-to-noise ratio and increasing the viewing volume of the system. This, in conjunction with selection of very fast optical lenses (Computar 25 mm, $f/1.85$ C-mount lenses) and high-output LEDs, has allowed us to obtain accurate measure-

ments within a cubic meter viewing volume at camera distances of 3 m.

Calibration and Data Correction

The numerical outputs of the cameras constitute sampled image coordinates of the LEDs in the reference system of the camera. This coordinate reference system must be transformed to a laboratory system of reference. This transformation of coordinate systems is accomplished by first calibrating the system for optical and electronic distortion, and then reconstructing three-dimensional data points from corrected data of two cameras, with knowledge of camera positions and orientations. Imaging properties of the optical lenses and nonlinearities in the electronics and detectors introduce in the position measurements distortions that must be corrected for. A fixed grid of LEDs at known locations was used to calibrate each camera detector. A camera was first aligned perpendicularly to the center of a 15×15 grid of equally spaced LEDs on a square meter surface, at a distance of 3 m. Each point in the grid was sampled 50 times by the acquisition system, and the average measured position was calculated. By exhaustively placing LEDs at known locations, measured locations can be corrected using a table look-up, where the table contains the Op-Eye mea-

sured locations for the grid, the known spacing among grid positions, and the distance between the detector and the grid.² Quadratic interpolation procedures were used to read corrections from the tables.

Three-Dimensional Reconstruction

To reconstruct three-dimensional data points from the corrected data of two cameras, the orientations and positions of the cameras must be known. Manual measurements and a computer monitor routine were used to determine these parameters. To simplify reconstruction formulas, the cameras are carefully oriented prior to data acquisition such that their optical axes intersect at a point and their horizontal axes lie in the same plane. This orientation is achieved by using an orientation "tree" that has three LEDs defining a plane parallel to the floor. Tilt, rotation, and translation stages (Newport Corporation) were attached to the cameras to allow the operator to microadjust the cameras while viewing the corrected positions of the LEDs on the screen. In this manner, optimal positioning of the detectors was achieved. With the positions and orientations of the cameras known, the three-dimensional coordinates can be calculated from the two sets of corrected camera data by triangulation (Woltring & Marsolais, 1980).

Data Processing

The data are processed off-line, either on the IBM-PC, or, for large amounts of data, on a VAX 11/785. The raw data from each detector are converted to two-dimensional coordinates, and any occluded points are flagged. The data are corrected for lens and electronic distortion and are interpolated, if necessary, for occluded points. The data from each detector are then low-pass filtered with a modified Butterworth filter to remove high-frequency noise.³ The three-dimensional coordinates are then calculated. Velocities and accelerations are derived using a two-point central-difference differentiation algorithm (Kay et al., 1985).

Graphics

The movement trajectory of a given LED for a given detector can be displayed on the IBM. This is useful for an immediate check of data validity and for on-line monitoring of the movement prior to data acquisition. However, intensive graphic analyses are performed after three-dimensional coordinates have been calculated. The three-dimensional trajectories are reconstructed and are interactively manipulated on an Evans and Sutherland Picture System 330, which displays three-dimensional images with appropriate depth perspective and parallax on an extremely high-resolution cathode-ray tube. The Picture System 330 allows dynamic, interactive graphics, including real-time rotation, translation, scaling, and clipping. The digitized trajectories can be graphically edited, displayed, and analyzed, focusing either on three-dimensional spatial attributes or on temporal-dynamic attributes. (These computergraphic procedures will form the topic of a separate report.)

RESULTS

Stationary Targets

We tested the accuracy of the system by measuring the absolute and relative positions of target LEDs across the field of view of the detectors. For this purpose, target LEDs were positioned across a meter square surface, and the orientation "tree" was positioned at various positions within the cubic meter viewing volume. The orientation "tree" is a rigid aluminum-wand structure housing three noncolinear LEDs that define a plane. The distance between the back two LEDs on the tree is 53.3 cm, and the distance between each of the back LEDs and the front

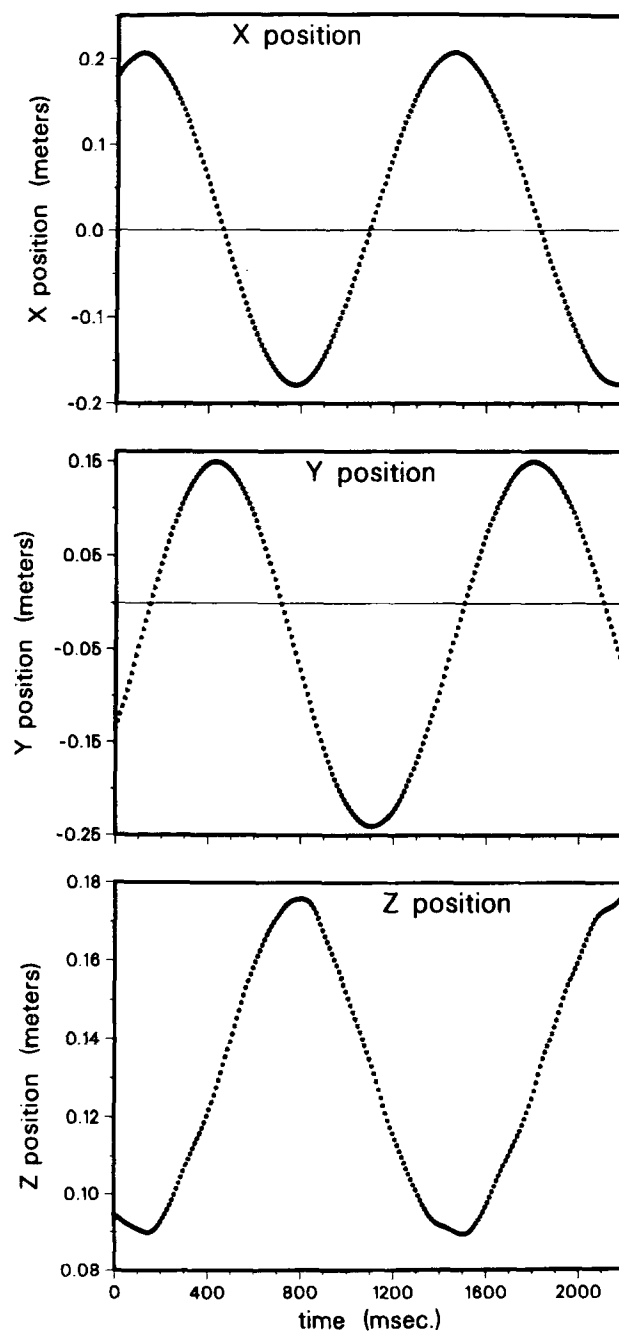


Figure 4. Measured displacements of a rotating wheel.

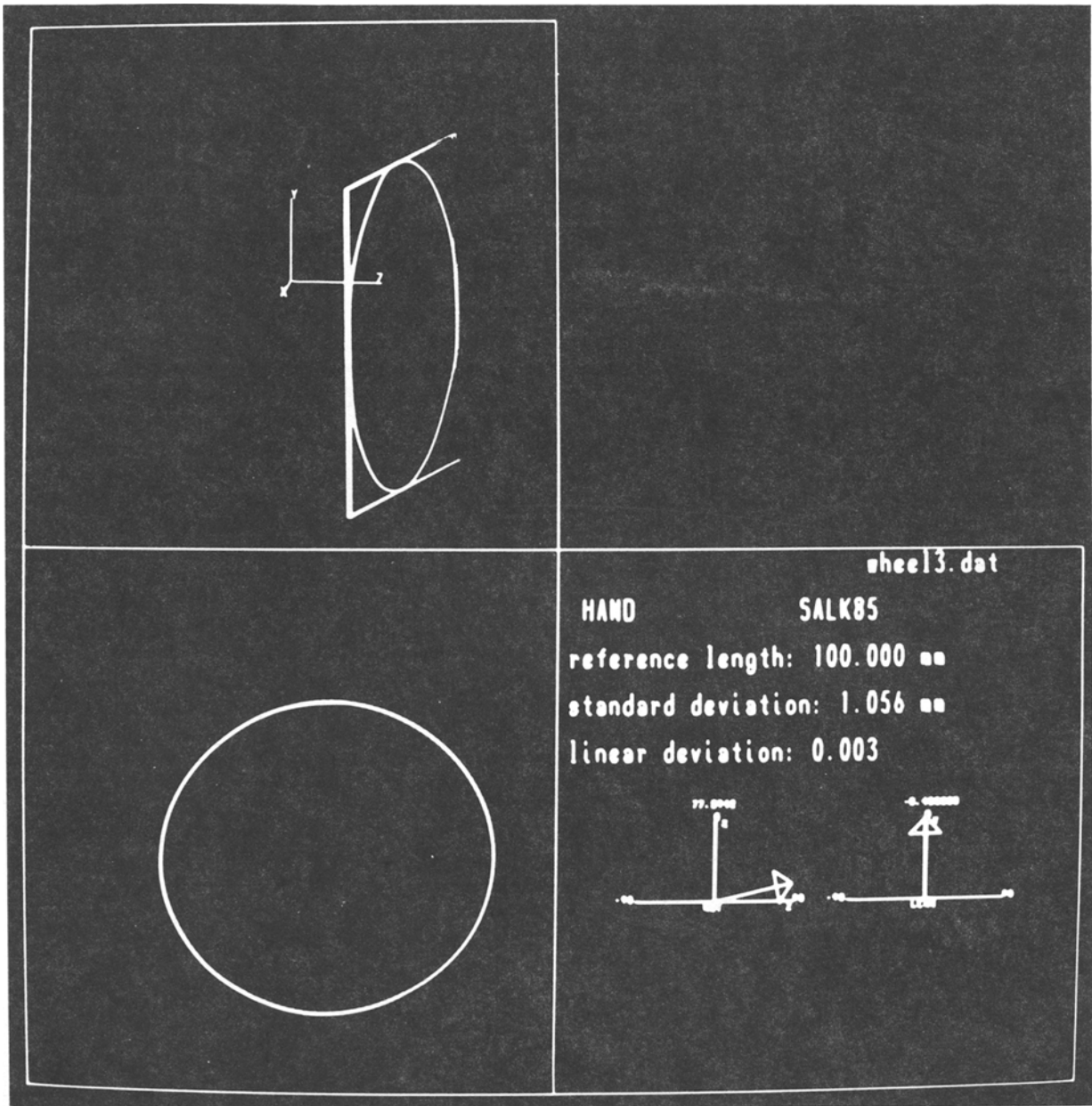


Figure 5. Top left: Three-dimensional reconstruction of the movement of an LED attached to a rotating wheel. The trajectory is embedded in a minimum boundary box drawn by the computer. The axes of the coordinate reference system are 100 mm on a leg. Bottom left: Trajectory of the wheel viewed perpendicular to the computed best-fitting plane for the data. Bottom right: Summary characteristics of the computed best-fitting plane.

LED is 31 cm. The orientation tree was mounted to a tripod and translated and rotated 10 times to position LEDs throughout the field of view of the detectors. The LEDs were sampled at 100 Hz for about 1 sec in each of the target positions. The standard deviations of individual x , y , and z coordinate values were calculated after three-dimensional reconstruction of the position of the stationary targets. These standard deviations were typically less than 1 mm.

The distances among the three LEDs on the orientation tree were computed from the three-dimensional coordinates and compared to the known distances. The computed distances differed from the actual distances by .2 cm to 4.6 cm, depending on the position in the field of view of the cameras. In general, the closer the target LEDs were to the center of the field of view, the more accurate the distance measurement.

Moving targets were measured next. We first mounted

a single LED to a rotating wheel, and then attached four LEDs to the major joints of the arms, while a person was producing American Sign Language.

A Rotating Wheel

A bicycle wheel was mounted to a honeycomb aluminum frame so that the wheel could freely spin. The frame itself was hinged to a wall and could be fixed at several rotational positions. An LED was mounted to the rim of the wheel, and the wheel was spun. Figure 4 presents the measured displacements along the horizontal, vertical, and transverse dimensions. Smooth sinusoidal motion was obtained, as we would expect. The three-dimensional trajectory was reconstructed, and the best-fitting plane to the movement was computed.

Figure 5 presents the trajectory embedded in the best-fitting plane, together with the trajectory viewed perpendicular to that plane. First, the diameter of the wheel trajectory was calculated from the measured three-dimensional data; it closely corresponded to the actual diameter. The diameter was computed for each of the 220 samples taken, by finding the distance between a given

sample point and the maximally distant point. The median computed wheel diameter (.392 m) differed from the actual diameter by only .6 mm; the standard deviation of the diameters was only .5 mm, and the maximum error was 1 cm. Second, Figure 5 shows that the measured motion was indeed quite circular. The major and minor axes of the trajectory projected onto the best-fitting plane were computed by algorithm, and their ratio was taken. The larger the ratio, the more linear the trajectory, whereas the closer the ratio to 1.0, the more circular the trajectory. The obtained ratio was 1.03, reflecting a highly circular trajectory shape, thus closely matching the actual movement. Finally, Figure 5 shows that measured movement of the wheel was highly planar, again matching the known trajectory. Indeed, the standard deviation of the points in the measured trajectory from the best-fitting plane was only 1.1 mm. The frame to which the wheel was attached was rotated with respect to the cameras, and the wheel was spun and its motion measured again. Similar results were obtained in this changed orientation. All of these data reflect the highly accurate three-dimensional monitoring of the system.

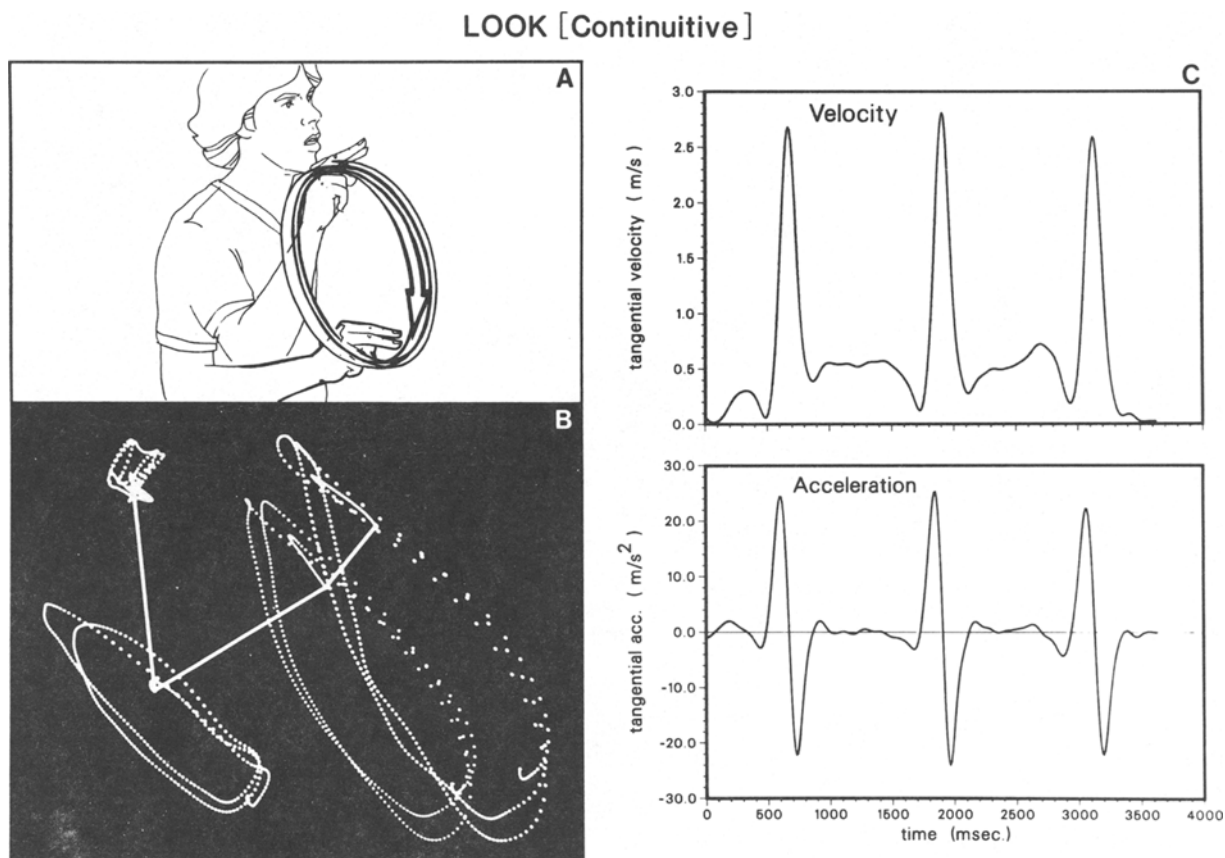


Figure 6. (A) Line drawing of the grammatically inflected sign, LOOK-AT[Continuative]. (B) Three-dimensional reconstruction of the trajectories of the hand, wrist, elbow, and shoulder. (C) Tangential velocity and acceleration of the hand.

American Sign Language

Our final measurement involved monitoring four positions on the arm while a signer produced a grammatically inflected sign from American Sign Language (ASL). Grammatical inflections in ASL are conveyed through proximal arm movements through space. We selected an inflection for temporal aspect, the Continuative, meaning "action for a long period of time." This inflection involves repeated elliptical movement of uneven rhythmic quality (Klima & Bellugi, 1979). LEDs were attached to the shoulder, elbow, wrist, and hand and were sampled at 100 Hz. Figure 6a presents a line drawing of the Continuative inflection (on the basic sign LOOK-AT); Figure 6b presents the three-dimensional reconstructed movement of the four points of the arm and shows the elliptical path shape of the sign. Figure 6c presents the tangential velocity and acceleration of the hand and clearly shows the rapid (2.7 m/sec) followed by slow (.5 m/sec) rhythmic quality of the movement across repetition cycles. The digitized and reconstructed sign trajectories presented in Figure 6 show capacities of the system for automated tracking of multiple joints of the arm in three-dimensional space. Only through such three-dimensional analysis can the precise nature of spatial and timing characteristics of movement be uncovered. We are finding, for example, that the temporal organization of sign language is very different from that of speech. This difference has important implications for our understanding of the underlying basis of left hemisphere specialization for language (Poizner, in press; Poizner, Klima, & Bellugi, in press).

SUMMARY

The Optical Position Acquisition System presented here uses affordable and portable hardware, yet provides accurate, automated, three-dimensional spatial monitoring of multiple joints of the hands and arms. We are currently using this system to help us better understand how the brain controls movement at various levels—linguistic, symbolic, and motoric. We are measuring and analyzing the breakdown of language and movement due to brain damage in deaf signers (see Poizner, Bellugi, & Irigui, 1984), a particularly revealing domain given that language and gesture are conveyed in one and the same modality in sign. The system is also being used to analyze the breakdown of movement in certain brain-damaged hearing subjects.⁴ The three-dimensional numeric and graphic analyses of movement made possible by the system are thus allowing new studies into the nature of the neural control of movement.

REFERENCES

- ANTONSSON, E. K. (1978). *The derivation and implementation of a dynamic three-dimensional linkage analysis technique*. Unpublished master's thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge.
- ANTONSSON, E. K. (1982). *A three-dimensional kinematic acquisition and intersegmental dynamic analysis system for human motion*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge.
- ATKINSON, C. G., & HOLLERBACH, J. M. (1985). Kinematic features of unrestrained arm movements. *Journal of Neuroscience*, 5, 2318-2330.
- CONATI, F. C. (1977). *Real time measurement of three-dimensional multiple rigid body motion*. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge.
- GINSBERG, C. M. (1983). *Human body motion as input to an animated graphical display*. Unpublished master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering, Cambridge.
- JOHANSSON, G. (1975). Visual motion perception. *Scientific American*, 232(6), 76-88.
- KAY, B. A., MUNHALL, K. G., BATESON, E. V., & KELSO, J. A. S. (1985). A note on processing kinematic data. Sampling, filtering and differentiation. *Haskins Laboratories Status Report (SR-81)*, pp. 291-303. New Haven, CT: Haskins Laboratory.
- KLIMA, E. S., & BELLUGI, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
- POIZNER, H. (1981). Visual and 'phonetic' coding of movement: Evidence from American Sign Language. *Science*, 212, 691-693.
- POIZNER, H. (1983). Perception of movement in American Sign Language: Effects of linguistic structures and linguistic experience. *Perception & Psychophysics*, 33, 215-231.
- POIZNER, H. (in press). Sign language processing. In J. Van Cleve (Ed.), *Gallaudet encyclopedia of deaf people and deafness*. New York: McGraw-Hill.
- POIZNER, H., BELLUGI, U., & IRAGUI, V. (1984). Apraxia and aphasia in a visual-gestural language. *American Journal of Physiology: Regulatory, Integrative & Comparative Physiology*, 246, R868-R883.
- POIZNER, H., KLIMA, E. S., & BELLUGI, U. (in press). *What the hands reveal about the brain*. Cambridge, MA: MIT Press/Bradford Books.
- REECE, D. A. (1982). *A Selspot-based data acquisition system for use in a clinical motion study laboratory*. Unpublished master's thesis, Case Western Reserve University, Department of Electrical Engineering and Applied Physics, Cleveland, OH.
- WOLTRING, H. J. (1974). New possibilities for human motion studies by real-time light spot position measurement. *Biotelemetry*, 1, 132-146.
- WOLTRING, H. J. (1984). On methodology in the study of human movement. In H. T. A. Whiting (Ed.), *Human motor actions—Bernstein reassessed* (pp. 35-73). North Holland: Elsevier.
- WOLTRING, H. J., & MARSOLAIS, E. B. (1980). Optoelectronic (Selspot) gait measurement in two- and three-dimensional space—A preliminary report. *Bulletin of Prosthetics Research*, 17, 46-52.

NOTES

1. Ginsberg (1983) has also proposed using the Op-Eye for three-dimensional movement measurement, but did not develop a working system.
2. We are currently developing a calibration routine that uses a computer-controlled dual-axis translation stage to position an LED to within 25 μ m resolution in x and y. This will allow an extremely precise, completely automated mapping of each detector.
3. We thank A. Dainis for providing the filter routine.
4. This work is being done in collaboration with Kenneth Heilman, Leslie Gonzalez-Rothi, and colleagues at the University of Florida.

(Manuscript received July 16, 1986;
revision accepted for publication September 18, 1986.)