INSTRUMENTATION & TECHNIQUES

Increased accuracy from a three-dimensional movement-tracking system

BRIAN McKEON

Massachusetts Institute of Technology, Cambridge, Massachusetts

This report describes the methods used to improve the performance of a three-dimensional movement-tracking system that is based on infrared cameras and light-emitting diodes. The technique has general application to similar systems with other types of cameras. The improvement was obtained by postprocessing to correct camera-image distortions. This procedure yielded a six-fold increase in accuracy. With a 900-mm-cube viewing volume, the error in locating a point was reduced to less than 3 mm.

In many studies of human movement, motion has been restricted to simple tasks to enable accurate measurement of some parameters of a task. The development of three-dimensional tracking systems allowed researchers to design experiments based on more natural movements, but these systems have disadvantages, and the accuracy of the data derived often restricts the user to a qualitative analysis.

One such system is the Selspot. Its operation is based on the use of two infrared cameras and light-emitting diodes (LEDs) fixed to different points on a subject's limb(s). The camera locations are known, and it is possible, using triangulation, to calculate the three-dimensional location of an LED based on the view seen by the two cameras. A major advantage of this system is that the three-dimensional data can be extracted directly from the camera data, in real time if necessary. A disadvantage, common to most camera-based systems, is that the calculation is prone to errors due to distortion of the view by the camera lens and electronics. Further errors arise due to misalignment of the cameras; for example, a camera may have a slight roll angle about its view axis.

Before attempting to improve the accuracy of the Selspot system, it was important to identify the sources of error and determine which errors could be reduced. The remaining error sources determined the limit of accuracy and indicated a practical value for which to aim. Some of the identified limiting factors were the following, specified in percent of the full range of camera output: (1) variation of LED intensity altered camera readings by up to 0.2%; (2) camera readings drifted by typically 0.03%, up to an extreme of 0.19% (19 LEDs, 3 days); and (3) reorientation of cameras after moving them was determined by the reproducibility of placement using bub-

ble levels permanently attached to the camera body and was 0.02%. These factors indicated that a practical limit of accuracy with the Selspot I system was probably around 0.2%. For a 1-m-cube viewing volume or 1-m-square view from each camera, an error of 0.2% translates to about a 2- to 3-mm position error.

Several sources of error could be reduced:

1. Distortion in the camera lenses and detector plates meant that the camera's *X* and *Y* readings did not reflect the true intersection point on the detector of a ray from a target. The Selcom technical manuals specified that the readings from the cameras had nonlinearity within 1% of full scale over an input range from 10% to 90% of full scale. The majority of error was assumed to be due to distortions in the camera lens and detector plate and, thus, would not change with time. Therefore, it was decided to correct these errors by postprocessing of the camera's *X* and *Y* data.

2. The orientation of optical axes that are normal to the detector plates of each camera are essential for calculation of three-dimensional object location. These axes are not necessarily parallel to the sides of the camera bodies. An assumption that is commonly used is that the detector plate is mounted so that its center *X*, *Y* reading (50%, 50%) is the intersection point of the optical axis with the detector plate. The results of a calibration procedure, described later, indicated that this would have been a bad assumption. For one of the cameras, the intersection point was actually over 12% from the nominal center point of the plate. If this offset was not accommodated, then errors of over 3% (30 mm for a 1-m cube) could result at edges of the viewing volume of the two cameras.

3. Focal length had to be determined for each camera. Calculations of three-dimensional target position are very sensitive to this parameter, and the calibration procedure described later indicated that, even for the two similar cameras of a Selspot system, calculated focal length differed by 10%. This difference did not have to be ex-

---

The author's mailing address is: Department of Psychology, E20-138, Massachusetts Institute of Technology, Cambridge, MA 02139.

plained by differences in the lenses; the distance of the detector plate from the lens was probably the important factor.

MEASUREMENT OF PERFORMANCE

The three-dimensional location of a target LED is calculated by projecting rays from the X, Y target locations on each detector plate through the focal point (fnode) of each camera and finding the three-dimensional point where these rays intersect. Misalignment of the two cameras and all of the other described sources of error prevent these rays from intersecting. Typically, a midpoint between the closest passing of the two rays is used as an estimate of the true intercept point. The minimum distance between the two rays can therefore be used as a measure of the performance of the system. This distance is known as the skew-ray error (SRE). The SRE was expressed in terms of camera output and, in the standard system, ranged up to 3%-4% which was not unexpected on the basis of the possible error sources listed for the standard system. Such an error translated to about 30-40 mm for a 1-m-cube viewing volume.

TECHNIQUES AND RESULTS

The system was improved in three areas: first, the distortions in the camera image were measured to allow correction of camera readings; second, a focal length figure was derived for each camera; and third, theoretical optical axes of the cameras were calculated for use with alignment of the cameras relative to each other. The details are described later.

Three tests were performed to measure the increase in accuracy after these changes. The cameras were aligned to give an approximate 900-mm-cube viewing volume. The first test involved moving a bar of 19 LEDs to different orientations and locations in the viewing volume and measuring SRE for the 19 LEDs at each position. The cases covered were vertical, horizontal, and 45° orientations at 1.6 m and 2 m from the camera bench and a vertical bar at 1.4 m from the bench. For 62 different LED locations, the worst-case SRE was 5.1 mm (mean = 2.6 mm, SD = 1.5 mm). The worst case of 5.1 mm was with a vertical array at the extreme point in the work space from both cameras. In this region, the SRE ranged from 2.9 to 5.1 mm.

The second test examined absolute inter-LED spacings for one of the above bar orientations and compared them to the known spacings. Only LEDs 4-16 were in range for this location, and the actual spacing between LEDs 4 and 16 was 505.5 mm. The computed prediction was 508.7 mm. The computed inter-LED spacing ranged from 41.1 to 43.0 mm for a known inter-LED spacing of 42 mm.

The third test examined the linearity of the system by moving a single LED along a straight-line path and ex-

amining the linearity of the calculated path based on samples from the movement. The path was not in a horizontal or vertical plane and was 670 mm in length. A regression line was fitted to 19 locations on the line, and the deviation of each point from that line was calculated. The root-mean-square error was 0.4 mm, and the maximum error was 0.7 mm.

DETAILED METHODS

The first step was to correct for camera distortion. Because this distortion is due in part to lens properties, changes in focus or aperture could alter the pattern of distortion. One solution was to use correction tables for different lens settings, but the mechanism of the infrared camera detector plate offered a simpler solution.

The detector plate of these cameras gave X and Y readings that were related to the centroid point of the light falling on the detector plate. Therefore, an unfocused image should provide the same information that a focused image provides. The camera focus length was therefore set at infinity and did not have to be altered for different experimental situations. In addition, the control unit for the LED targets was modified so as to allow adjustment of LED intensity for different distances. These two decisions meant that the camera lens settings did not have to be altered for different experiments, and therefore only one set of correction factors was needed.

The correction factors were derived by moving a column array of LEDs across a panel that had been set up so as to be normal to an axis from the camera. This axis was defined by use of a low-power laser that was temporarily attached to the camera body. The camera was leveled by use of a bubble level attached to the camera, and the LED panel was then moved so as to reflect the laser beam back into the laser.

The camera readings for each LED location were collected and a square grid was fitted to this data. The spacing of the lines on this grid, the distance to the panel, and the LED spacing on the panel were used to derive an equivalent focal length in the units of camera output. The deviations of camera readings from the square grid readings were used to generate a table that could be used to correct camera data.

An important decision was how far apart to place sample locations on the grid. Preliminary data indicated that the major component of distortion slowly changed over the full range of camera output, with a maximum of about six cycles of change over this range. To be able to properly interpolate between sampled points on such data, it was necessary to sample at least 18 points across the output range. This would have resulted in an 18×18 grid of input points. Our system was able to drive a column of 19 LEDs directly, and a 19×19 grid was chosen. An initial check of the assumptions was performed by use of a 38×38 grid. Every second location was used to create a 19×19 grid. These 19×19 data were then used to see

how well the intermediate positions of the 38×38 grid could be predicted. The errors were within 0.2%, validating the use of a 19×19 grid of samples.

The next step was to define the optical origin of the panel. This is the point at which a ray from the camera and normal to the panel intersected the panel. As the laser beam was normal to the panel, it was possible to use this beam, together with its known offset from the center of the camera, to define the origin. This was specified relative to a numbered LED location.

The final step was to put the cameras into their locations for the desired experiment and perform a global calibration. This was a quick process and had to be performed whenever the cameras were moved. A horizontal array of 19 LEDs, placed within view of both cameras, was used as a reference horizon. The two cameras were adjusted to be level with each other, and the array of LEDs was also placed at the same height. The data collected from the points of the array were corrected for distortion, and the corrected data enabled derivation of pitch-

and-roll angles of the two cameras. Finally, a single point on the array, together with some measurements from that point to each camera, was used to derive the included angles between the camera axes and a line drawn between the two cameras.

This last section of global calibration involved a compensation of some measurement errors by shifting other measurements from their true values. The result was a system that traded global accuracy for relative accuracy within the viewing volume. This was of little consequence, as all experimental measurements were made between points within the viewing volume.

SOFTWARE AVAILABILITY

Listings of the Fortran programs and subroutines used in the calibration procedure together with details of the calibration procedure can be obtained from the author.

(Manuscript accepted for publication January 10, 1986.)