Head and Hand Tracking Devices in Virtual Reality

ENRICO GOBBETTI, RICCARDO SCATENI, AND GIANLUIGI ZANETTI

CRS4, VI Strada OVEST Z.I. Macchiareddu, C.P. 94, 09010 UTA (CA - Italy)

Telephone: +39 07027961

Fax +39 0702796216

e-mail: {gobbetti|riccardo|zag}@crs4.it

Head and Hand Tracking Devices in Virtual Reality

E. GOBBETTI, R. SCATENI, AND G. ZANETTI

CRS4, C.P. 94, I-09010 Uta (CA)

Abstract

This short introduction to head and hand tracking devices summarizes the characteristics of some of the technologies most relevant to medical application and presents the configuration chosen in three representative test beds.

Introduction

The ability to track human movement is an essential requirement for many virtual reality systems. More than thirty years are passed since the research activities in this field began (Sutherland, 1968). Nowadays, almost any virtual medical application that allows the user to interact actively with the scene needs to know where the user is looking, and, often, what she is doing with her hands. There are many different tracking scenarios and technologies. The selection among the several different technologies that can be used to implement a tracking device is obtained by weighting factors that range from technical requirements of specific applications to cost of specific systems.

In this short introduction to the subject, we will limit ourselves to the description of some of the techniques and systems that are specifically relevant to medical applications. For a general discussion see Meyer et al (1992), Azuma (1993), Bar-Shalom and Li (1993), Bhatnagar (1993), Kalawsky (1993), Burdea and Coiffet (1994), Sturman and Zeltzer (1994), Durlach et al (1995), and Youngblut et al (1996).

Trackers technologies

The evolution and the diffusion of virtual environments applications has spurred the development of many different kinds of tracking systems. Their characteristics are expressed in terms of the following parameters.

Accuracy: The measure of the error in position and orientation reported by the tracker.
Resolution: The smallest change in position and orientation that can be detected by the tracker.
Frequency: The rate at which position and orientation measurements are reported by the tracker to the host computer.
Latency: The delay between a change in position and/or orientation and the notification of the change to the computer.
Working volume: The volume within which the tracker can measure position and orientation with its specified accuracy and resolution.
An ideal tracker should have high accuracy, fine resolution and high frequency. Its latency should be as low as possible and it should work over wide areas, possibly scalable at will. It should also not need any specialized environment for operation and the wearable parts of the tracker should be small and light in weight, for the user comfort.
The purpose of a tracker is to report to a computer changes in the position of selected reference points. Depending on the tracking technology used, these changes may either be known in terms of relative motion or with respect to some external, fixed, reference system. For instance, gloves with piezo-resistive flex sensor are able to report on the relative motion of the fingers but not on the global movement of the hand. On the other hand, in medical applications, it is usually important to be able to describe the motion of the tracked part, e.g., the surgeon hand in a surgical simulator system, with respect to a fixed reference frame. From a technological point of view this may be obtained by either having a sensor that is attached to the tracked body and it is able to report its position by measuring the intensity of a location dependent field, or by having a detection system that it is able to measure, e.g., distances from some external reference points to some specific targets attached to the tracked body.
It is possible to summarize these considerations in the following classification categories, see Mulder (1994):

- **Inside-in systems** are defined as those which employ sensor(s) and source(s) that are both on the body (e.g., a glove with piezo-resistive flex sensors), thus allowing for an unlimited working volume. The sensors are generally small and are therefore especially suitable for tracking small body parts, but they are only able to provide relative motion information;

- **Inside-out systems** employ sensor(s) on the body that sense artificial external source(s) (e.g. a coil moving in a externally generated electromagnetic field), or natural external source(s) (e.g. a mechanical head tracker using a wall or ceiling as a reference or an accelerometer moving in the earth's gravitational field). Although these systems provide position information with respect to a fixed reference frame, their working volume and accuracy is generally limited due to use of the external source;
- **Outside-in systems** employ an external sensor that senses artificial source(s) or marker(s) on the body, e.g. an electro-optical system that tracks reflective markers or natural source on the body (e.g. a video camera based system that tracks the pupil and cornea). These systems generally suffer from occlusion, and a limited workspace.

The most common tracking technologies are: electromagnetic trackers; acoustic trackers; optical trackers; mechanical trackers; inertial trackers. For a selected number of systems we included references to the web documents containing further specifications.

Electromagnetic trackers

Electromagnetic trackers use orthogonal electromagnetic (EM) fields to determine the relative location of a receiver with respect to a known transmitter. The receiver is typically attached to the object to be tracked, while the transmitter is at a fixed reference position. Electromagnetic trackers systems (ETS) are widely used in VR because they are relatively inexpensive, and provide a fairly good resolution in a reasonably large workspace. Moreover, ETS do not suffer from optical occlusion. The main drawback of ETS is that the EM fields used could be seriously degraded and distorted by the presence of metallic structures and stray fields from other sources. Modern ETS, e.g., Flock-Of-Bird (Ascension ,2001), use pulsed DC magnetic fields to avoid the perturbation introduced by diamagnetic and paramagnetic materials (e.g., aluminum, stainless steel) but they cope with difficulty with ferromagnetic materials due to their high magnetic susceptibility. A single EM tracking system can pack up to 32 receivers

(Star*Trak, Polhemus, 2001) processing full-body movements with a frequency as high as 120 Hz, allowing to catch even human performers' fast-pace dancing.

Acoustic trackers

Acoustic trackers use ultrasonic waves to determine the position of the target object. They can be classified into two categories, Time of Flight (T.O.F.) trackers and phase coherent trackers.

T.O.F. trackers compute position and orientation of the target object by measuring the time taken by ultrasonic pulses to travel from a set of transmitters to a set of receivers. A typical system might consist of three transmitters and three receivers. The transmitters are mounted on the target object and the receivers are arranged at known fixed positions in the environment.

Phase coherent trackers track position and orientation by comparing the phases of emitted acoustic waves with the phase of a reference wave. Transmitters of acoustic waves are mounted on the target object and receivers are set up at fixed positions in the environment. The receivers periodically measure the phase difference between the waves emitted by the transmitters and a reference wave.

Acoustic trackers suffer from occlusion and are disturbed by large air movements in the working volume.

An example of acoustic tracker is the 6 degrees-of-freedom (DOF) Head Tracking System by Fakespace Labs Inc. (FakeSpace, 2001). It consists consists of a transmitter, helmet-mountable receiver and control box. The transmitter contains 3 speakers which emit ultrasonic signals and the receiver contains 3 miniature microphones to sense these signals. The trackers works at 50 Hz, in a cone-shaped working volume 150 cm long, 100 degrees wide.

Optical trackers

A wide variety of technologies is used for optical trackers. Most of them can be classified as beacon trackers, but there are some types, like the laser ranging trackers, which do not belong to this category.

Beacon trackers use a set of optical beacons which may be either active (emitters of light) or passive (reflectors of light). A set of sensors such as cameras or photodiodes is used to track them (InMotion Systems, 2001). This arrangement is extremely reliable and accurate and particularly suited for medical applications such as Computer Aided Surgery (CAS). In CAS the optical beacons, usually light-emitting diodes (LED), are mounted on rigid bodies such as surgical tools and/or bone fragments (FlashPoint 5000, Image Guided Technologies, 2001). The main drawback of optical tracking systems is that they require a free line of sight between the LED and the cameras and this could be too strong a limitation in an actual surgical setting.

Using low cost infrared LED's arrays as a reference frame, compact optical trackers can be used for very large areas coverage (hundreds of square meters) keeping a good level of accuracy: less than 1 mm in location; less than a prime in orientation (3rdTech, 2001).

Mechanical trackers

Mechanical trackers measure the position and orientation of a target object that is attached to the end of a movable mechanical arm. The arm is anchored at a fixed point of reference and it is made up of several sections that can rotate and move at the joints. The rotations and movements are measured by gears, potentiometers or optical encoders and are used to compute the position and orientation of the target object relative to the fixed point of reference. Due to the relatively small working volume they are mainly used as very accurate probing systems (FARO, 2001; FakeSpace, 2001).

When coupled with motors giving force feedback to the user's movements, they are basic components for virtual simulators (Sensable, 2001).

Inertial trackers

Inertial trackers use accelerometers and gyroscopes to compute changes in position and orientation from measurements of acceleration and velocity. These trackers tend to accumulate error over time and need to be updated by some external source from time to time. On the other hand they can be very light, small and cheap to be use in emerging markets like head mounted displays for videogames (Intersense, 2001). Apart from this, their main use is to complement EM trackers by providing a mechanism to compensate for EM field distorsions.

Tracking in medical applications

Medical applications have very specific tracking requirements: typically, accuracy and resolution have to be high, since very precise movements have to be followed; frequency and latency depend on the type of application (they are low in telesurgery, and higher in surgical simulation); the working volume is generally restricted to tens of centimeters.

Another important challenge is given by the medical environment: surgical applications tracking technology have to follow the same sterility standards as usual surgical instruments; the medical machinery should not interfere with the tracking systems and vice versa.

For all these reasons it is impossible to define the most appropriate technologies for a generic medical application; the choice has to be made on a case-by-case basis, after the analysis of specific application requirements.

In the following, we provide a short description of the tracking technology used in three recently developed applications.

Ultrasound-guided Breast Biopsy

This system, developed at the Dept.of Computer Science at University if North Carolina at Chapel Hill (State et al, 1996), merges rendered live ultrasound data and geometric elements, with stereo images of the patient acquired through head-mounted video cameras and presents these merged images to the physician in a head-mounted semi-transparent display (Augmented Reality: AR). This application is of great interest because the physician is able to see a visualization of the ultrasound data directly under the ultrasound probe, properly registered within the patient and with the biopsy needle. Although it is still experimental, it gave already useful clinical results allowing a physician, using this system, to successfully insert a needle into an artificial tumor within a training phantom of a human breast.

Accurate tracking is crucial for precise registration of real and synthetic imagery, especially in a medical application where surgical intervention is to be performed under AR guidance. In the system double tracking is required: the physician's head and the hand-held ultrasound probe.

The prerequisite for selecting the head tracking system was lightweight and minimal encumbrance for the physician, so a magnetic tracker appeared to be an ideal choice. Since the metallic structures in the lab interfered with EM tracking accuracy the developers combined magnetic tracking with vision-based landmark tracking for improved registration.

In addition tracking of the ultrasound probe had to be very precise for correct registration of ultrasound slices. Two registration tasks are to be performed: a slice-to-slice registration leading to reconstruct the volume of data, and an alignment of the visualized volume with the patient, seen trough the semi-transparent visor. The probe has usually a small working and tracking volume and is already tethered to the ultrasound machine. The solution of choice was, then, to track the probe with a 6-degree-of-freedom FARO mechanical tracker (Faro, 2001) even though it hinders probe motion to a certain extent.

The data reported by the authors, about the task precision are very demanding: the physician may be required to place a thin needle - for example, 0.7 mm diameter for cyst aspiration, 2.1 mm diameter for biopsy - into a 3 mm cyst. Since none of the commercially available trackers, had the accuracy and precision required for this medical application the system was set up combining the mechanical tracker with a magnetic tracker (Flock of Bird Ascension, 2001) corrected by a lookup table, and vision-based tracking to achieve improved registration of real imagery (patient, ultrasound probe, biopsy needle) and synthetic imagery (ultrasound slices, rendered visual and occlusion cues).

Catheter insertion

The insertion of a catheter into a vessel (artery or vein) is one of the most common procedures in clinical practice. Precise catheter insertion requires a perfect knowledge of the three-dimensional development of vessels and a high level of dexterity during vessel puncture, which is only attainable after considerable practice. If badly performed, catheter insertion can result in undue patient harm, and inadequate training in this technique is a leading cause of hospital-acquired infections (Zorcolo et al, 2000a).

The first part of the operation, needle insertion, is difficult to learn, and requires a combination of visual and tactile skills, to identify the needle insertion point and control needle position and orientation during penetration. It is thus important for a training system to provide co-registered visual and haptic responses. The system enhances touch perception with visual perception by using head-tracked

stereoscopic viewing and a custom made display system which provides hand-immersed interaction. The user looks into a mirror through shutter glasses and perceives the virtual image within the workspace of a PHANToM haptic device which controls needle insertion. By co-registering physical and virtual spaces beforehand and using stereoscopic projections that dynamically follow head movements, the virtual patient is made to appear at a fixed physical position on a virtual surgical table. This produces, without resorting to an head-mounted display, a combined haptic and stereoscopic view of a virtual volume in which users can manipulate the needle with their hands without obscuring the display.
As visual feedback and force feedback have different frequency requirements dictated by the human perceptual system, computation and tracking threads work in parallel and different tracking technologies are integrated in the same application.
Force feedback computation has to work at 1KHz, 6DOF hand position tracking requires about 100Hz, while stereo visual feedback guided by head tracking has to be provided at 10–20 frames per second (Zorcolo et al, 2000b). To reach these goals, the system combines a PHANToM haptic device (Sensable, 2001), for hand position sensing (6 degrees of freedom: x, y, z, yaw, pitch, roll) and force feedback (3 degrees of freedom: x, y, z) with an ultrasound Logitech tracker mounted on SGI CristalEyes (CrystalEyes, 2001)stereo goggles for hand tracking and stereo image presentation.

Mastoidectomy simulation

The most superficial and common surgery of the petrous bone is the mastoidectomy, which consists of removal of the air cavities just under the skin behind the ear itself.
A virtual reality system for simulating surgical procedures on the petrous bone is a human-computer interface tool that puts the user in the loop of a real-time simulation mimicking a realistic synthetic operating environment. The surgical simulator receives external input driving the positioning of surgical instruments. Such a system is being designed and developed in the framework of EU IST-1999-12175 IERAPSI project (John et al, 2001).
The project began with a detailed user task analysis of surgeons carrying out the procedures being targeted. Subjective analysis of video records, together with *in-*

situ observations highlighted a correlation between drilling behaviors and type and depth of bone (Stone, 2000).

The system requires to track both the dominant hand (controlling the burr) and non-dominant hand (controlling the sucker). The work volume for both hands is in the order of 5 cm. The range of motion for the dominant hand varies from 2-4 cm sweeps to static drilling while maintaining minimal surface pressure. The accuracy of the tracking must accommodate both rapid lateral strokes and *polishing* motion quality. The parameters for the non-dominant hand are less demanding. High quality force feedback is required for the dominant hand (sensing the tissue variations) while only collision detection is required for the non-dominant one.

The task analysis identified that the only commercially available system capable of replicating the qualities required for the dominant hand is Sensable Technologies' *PHANToM* with stylus encoder. For the non-dominant hand cheaper force feedback systems are available (e.g., Haptech/Immersion Corporation *PenCAT Pr*o, Immersion, 2001).

Careful consideration must be given to the head tracking and display-control technology used in temporal bone surgeons training system. The surgeons sees the patient through a surgical microscopes. The size of movements are not great: typically, the mimimum requirement are as follows:

- Translation left and right (relative to centre of surgeon's face): +/- 10cms (approx.),
- Translation up and down: +/- 5 to 10 cms
- Translation in and out: +/- 5cms
- Rotation of eyepiece unit (in pitch - away from and back to surgeon's face): +/- 30 to 45 degrees

Magnetic tracking devices are not well suited to the application because of the interferences of force feedback motors. Mechanical tracking systems would perfectly match these requirements. However, since the training application does not have stringent accuracy requirements ultrasound devices, that are much cheaper, or hybrid devices could be successfully integrated.

Conclusions

In this short introduction to head and hand tracking devices, we have summarized the characteristics of some of the technologies most relevant to medical application and presented the configuration chosen in three representative test beds.

This work was partially supported by the Sardinian regional authorities.

References

3rdTech. (2001) http://www.3rdtech.com.

Ascension. (2001) http://www.ascension-tech.com.

Azuma RT (1993) Tracking requirements for augmented reality. Communications of the ACM 36:50–51.

Bar-Shalom Y and Li XR (1993) Estimation and Tracking: Principles, Techniques, and Software. Boston: Artec House, Inc.

Bhatnagar DK (1993) Position trackers for head mounted display systems: A survey (Tech. Rep. TR93-010). Chapel Hill: University of North Carolina at Chapel Hill.

Burdea G and Coiffet P (1994) Virtual Reality Technology (1st ed.). New York: John Wiley & Sons, Inc.

CrystalEyes (2001) http://www.stereographics.com.

Durlach N. et al (1995). Position Tracking and Mapping. In N. Durlach & A. S. Mavor (Eds.), Virtual Reality: Scientific and Technological Challenges, (pp. 188-204). Washington, D.C.: National Academy Press.

FakeSpace (2001) http://www.fakespacelabs.com.

FARO (2001) Faro Arm. http://www.faro.com/products.

Image Guided Technologies. (2001). http://www.imageguided.com.

Immersion (2001) http://www.immersion.com.

InMotion Systems (2001) CODA mpx30. http://www.inmotionsystems.com/.

Intersense. (2001) http://www.isense.com.

John NW, Thacker N, Pokric M, Jackson A, Zanetti G, Gobbetti E, Giachetti A, Stone R, Campos J, Emmen A, Schwerdtner A, Neri E, Sellari Franceschini S and Rubio F (2001) An integrated simulator for surgery of the petrous bone In Westwood JD, ed, *Medicine Meets Virtual Reality 2001*, Amsterdam, The Netherlands. IOS.

Kalawsky RS (1993) The science of virtual reality and virtual environments. Wokingham, England: Addison-Wesley.

Meyer K, Applewhite H, and Biocca F (1992) A survey of position trackers. Presence, 1(2), 173–200.

Mulder A (1994). Human movement tracking technology (Tech. Rep. TR 94-1). School of Kinesiology, Simon Fraser University.

Polhemus. (2001). http://www.polhemus.com/home.htm.

Sensable (2001). http://www.sensable.com/haptics/products.html.

State A, Livingston MA, Garrett WF, Hirota G, Whitton MC, Pisano ED (1996) Technologies for Augmented Reality Systems: Realizing Ultrasound-Guided Needle Biopsies. Proceedings SIGGRAPH 96 (New Orleans, LA, 4-9, August, 1996) 439-446.

Stone RJ (2000) A Human-Centred Definition of Surgical Procedures, IERAPSI (IST-1999-12175), Work Package 2, Deliverable D2 (Part 1).

Sturman DJ and Zeltzer D (1994) A Survey of Glove-Based Input. IEEE Computer Graphics and Applications, 14(1):30-39.

Sutherland IE (1968). A head-mounted three dimensional display. Proceedings of the 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings 33(1):757–764. Washington, D.C.: Thompson Books.

Youngblut C, Johnson RE, Nash SH, Wienclaw RA, and Will CA (1996). Tracking Interfaces. In Review of Virtual Environment Interface Technology , IDA Paper P-3186 :47-77. Alexandria, VA: Institute for Defense Analyses.

Zorcolo A, Gobbetti E, Zanetti G, and Tuveri M (2000a) Catheter Insertion Simulation with co-registered Direct Volume Rendering and Haptic Feedback. In Westwood JD, Hoffman HM, Mogel GT, Robb RA, and Stredney D, eds, *Medicine Meets Virtual Reality 2000 -- Envisioning Healing: Interactive Technology and the Patient-Practioner Dialogue*, pp 96-98, Amsterdam, The Netherlands. IOS.

Zorcolo A, Gobbetti E, Zanetti G, and Tuveri M (2000b) A volumetric virtual environment for catheter insertion simulation. In van Liere R and Mulder J, eds, *EGVE2000*, Lecture Notes in Computer Science. Springer-Verlag Inc.