

# PRESENCE

## TELEOPERATORS AND VIRTUAL ENVIRONMENTS

S  
P9145

This resource is also available  
on the WWW.  
Use MADCAT to launch.



VOLUME 10 / NUMBER 1 / FEBRUARY 2001

MIT PRESS

META 1007  
META V. THALES

**PRESENCE**  
**Teleoperators and**  
**Virtual Environments**

Volume 10, Number 1  
February 2001  
ISSN 1054-7460

**Editor-in-Chief**

Nathaniel I. Durlach  
Virtual Environment and  
Teleoperator Research Consortium  
Research Laboratory of Electronics  
MIT

**Managing Editor**

Rebecca Lee Garnett  
MIT

**Director of Corporate**  
**Contributions**

Thomas A. Furness III  
Human Interface Technology  
Laboratory  
University of Washington

**Editorial Address**

Editor-in-Chief, *PRESENCE*  
MIT  
77 Massachusetts Avenue  
Room 36-709  
Cambridge, MA 02139  
presence@mit.edu

Individuals wishing to submit  
manuscripts should follow the  
guidelines provided in this issue.

**Abstracting and Indexing**

*PRESENCE: Teleoperators and Virtual Environments* is included in *Computer Abstracts*, *Ergonomics Abstracts*, *Information Science Abstracts*, *International Aerospace Abstracts*, and *Multi-Index to Cyber-Space, Virtual, and Artificial Reality*. *CompuMath Citation Index*, *Current Contents/Engineering Computing and Technology*, *Research Alert*, and *SciSearch*.

**Sponsorship**

*PRESENCE* is indebted to the following organizations for their generous support:  
Air Force Office of Scientific Research  
Department of Biomedical Engineering, Boston University  
Human Interface Technology Laboratory at the University of Washington  
Office of Naval Research  
Research Laboratory of Electronics, MIT  
Visual Computing Department, Hewlett Packard Laboratories  
**Senior Editors**  
Woodrow Barfield  
Virginia Polytechnic Institute  
Gary Bishop  
University of North Carolina

Larry Hodges  
Georgia Institute of Technology  
John M. Hollerbach  
University of Utah  
Randy Pausch  
Carnegie Mellon University  
Thomas B. Sheridan  
MIT  
Mel Slater  
University College London  
Kay Stanney  
University of Central Florida  
Elizabeth Wenzel  
NASA Ames Research Center  
David Zeltzer  
Fraunhofer Center for Research in Computer Graphics  
Michael Zyda  
Naval Postgraduate School

**Editorial Board**

Bernard Adelstein  
NASA Ames Research Center  
Terry Allard  
NASA Ames Research Center  
Walter A. Aviles  
Teneo Computing, LLC  
Frank Biocca  
Michigan State University  
Jens Blauert  
Ruhr-University Bochum,  
Germany  
Grigore Burdea  
Rutgers University  
H. Steven Colburn  
Boston University  
Rudy Darken  
Naval Postgraduate School  
Paul Dizio  
Brandeis University  
Stephen Ellis  
NASA Ames Research Center  
Scott S. Fisher  
Telepresence Research  
Richard M. Held  
MIT  
Kenneth O. Johnson  
Johns Hopkins University  
Lynette A. Jones  
MIT  
James R. Lackner  
Brandeis University  
Jaron Lanier  
Columbia University  
Susan Ledeman  
Queen's University, Canada  
Jack M. Loomis  
University of California  
Michael Macedonia  
STRICOM  
Michael Moshell  
University of Central Florida  
Michael Naimark  
Interval Research Corporation  
Albert "Skip" Rizzo  
University of Southern California  
Warren Robinett  
University of North Carolina  
Jannick Rolland  
University of Central Florida  
Roy Ruddle  
University of Leeds

Karun Shimoga  
CMU Robotics Institute  
Barbara Shinn-Cunningham  
Boston University  
Gurminder Singh  
National University of Singapore  
Robert J. Stone  
VR Solutions/Virtual Presence  
Ltd. UK  
Martin Stytz  
Air Force Institute of  
Technology  
Susumu Tachi  
University of Tokyo  
James Templeman  
Naval Research Laboratory  
Geb Thomas  
University of Iowa  
Colin Ware  
University of New Hampshire  
Richard C. Waters  
Mitsubishi Electric Research  
Laboratory  
Suzanne Weghorst  
University of Washington  
Janet Weisenberger  
Ohio State University  
Robert B. Welch  
NASA Ames Research Center  
Thomas E. von Wiegand  
MIT

**Business Offices and**  
**Subscription Rates**

*PRESENCE: Teleoperators and Virtual Environments* is published bimonthly (February, April, June, August, October, December) by The MIT Press, Cambridge, MA 02142-1407. Subscriptions and address changes should be addressed to MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142-1407; (617) 253-2889; fax (617) 577-1545; e-mail journals-orders@mit.edu. An electronic, full-text version of *PRESENCE* is available from the MIT Press. Subscriptions are on a volume-year basis. Subscription rates:  
**Electronic only**—Individuals \$45.00, Students/retired \$43.00, Institutions \$342.00. Canadians add the 7% GST. **Print and Electronic**—Individuals \$80.00, Students/retired \$48.00, Institutions \$380.00. Outside the U.S. and Canada add \$30.00 for postage and handling. Current issues are \$15.00. Back issue rates: Individuals \$32.00, Institutions \$64.00. Outside the U.S. and Canada add \$5.00 per issue for postage and handling. Canadians add 7% GST. Claims may be e-mailed to: journals-claims@mit.edu. Claims for missing issues will be honored free of charge if made within three months after the publication date of the issue. Prices subject to change without notice. <http://mitpress.mit.edu/PRES>

**Postmaster**

Send address changes to *PRESENCE*:

*Teleoperators and Virtual Environments*, Five Cambridge Center, Cambridge, MA 02142-1407. Periodicals postage paid at Boston, MA, and at additional post offices.

**Advertising and Mailing List**  
**Rental**

Inquiries may be addressed to the Marketing Dept., MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142-1407. (617) 253-2866; fax (617) 258-5028; e-mail: journals-info@mit.edu.

**Copyright Information**

Permission to photocopy articles for internal or personal use, or the internal or personal use of specific client, is granted by the copyright owner for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the fee of \$8.00 per article-copy is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The fee code for users of the Transactional Reporting Service is 1054-7460/01 \$8.00. For those organizations that have been granted a photocopy license with CCC, a separate system of payment has been arranged. Address all other inquiries to Subsidiary Rights Manager, MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142-1407; e-mail: journals-rights@mit.edu.

*PRESENCE: Teleoperators and Virtual Environments* is available on microfilm from University Microfilms, Inc., 300 N. Zeeb Road, Ann Arbor, MI 48106.

© 2001 by the Massachusetts Institute of Technology.

# PRESENCE

TELEOPERATORS AND VIRTUAL ENVIRONMENTS

VOLUME 10, NUMBER 1, FEBRUARY 2001

Editorial Notes	iii
Guest Editors' Introduction: VRST'99 Special Issue	iv

<b>ARTICLES</b>		
S	High-Performance Wide-Area Optical Tracking: The HiBall Tracking System	1
	<i>Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'nardo Colucci</i>	
S	GNU/MAVERIK: A Microkernel for Large-Scale Virtual Environments	22
	<i>Roger Hubbard, Jon Cook, Martin Keates, Simon Gibson, Toby Howard, Alan Murta, Adrian West, and Steve Pettifer</i>	
S	Patterns of Network and User Activity in an Inhabited Television Event	35
	<i>Chris Greenhalgh, Steve Benford, and Mike Craven</i>	
S	Components for Distributed Virtual Environments	51
	<i>Manuel Oliveira, Jon Crowcroft, and Mel Slater</i>	
S	An Adaptive Multiresolution Method for Progressive Model Transmission	62
	<i>Danny To, Rynson W. H. Lau, and Mark Green</i>	
S	Testbed Evaluation of Virtual Environment Interaction Techniques	75
	<i>Doug A. Bowman, Donald B. Johnson, and Larry F. Hodges</i>	
S	An Introduction to 3-D User Interface Design	96
	<i>Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, Jr., and Ivan Poupyrev</i>	
	An Overview of the COVEN Platform	109
	<i>Emmanuel Frécon, Gareth Smith, Anthony Steed, Mårten Stenius, and Olov Ståhl</i>	

## FORUM

WHAT'S HAPPENING

128

UW-MADISON LIBRARY  
COLLEGE OF ENGINEERING  
MAY 2 - 2001  
UW-MADISON, WI 5370

**Greg Welch**  
welch@cs.unc.edu

**Gary Bishop**  
gb@cs.unc.edu

**Leandra Vicci**  
vicci@cs.unc.edu

**Stephen Brumback**  
brumback@cs.unc.edu

**Kurtis Keller**  
keller@cs.unc.edu  
Department of Computer Science  
University of North Carolina at  
Chapel Hill

**D'nardo Colucci**  
colucci@virtual-reality.com  
Alternate Realities Corporation

# High-Performance Wide-Area Optical Tracking

## The HiBall Tracking System

---

### Abstract

Since the early 1980s, the Tracker Project at the University of North Carolina at Chapel Hill has been working on wide-area head tracking for virtual and augmented environments. Our long-term goal has been to achieve the high performance required for accurate visual simulation throughout our entire laboratory, beyond into the hallways, and eventually even outdoors.

In this article, we present results and a complete description of our most recent electro-optical system, the HiBall Tracking System. In particular, we discuss motivation for the geometric configuration and describe the novel optical, mechanical, electronic, and algorithmic aspects that enable unprecedented speed, resolution, accuracy, robustness, and flexibility.

### I Introduction

Systems for head tracking for interactive computer graphics have been explored for more than thirty years (Sutherland, 1968). As illustrated in figure 1, the authors have been working on the problem for more than twenty years (Azuma, 1993, 1995; Azuma & Bishop, 1994a, 1994b; Azuma & Ward, 1991; Bishop, 1984; Gottschalk & Hughes, 1993; UNC Tracker Project, 2000; Wang, 1990; Wang et al., 1990; Ward, Azuma, Bennett, Gottschalk, & Fuchs, 1992; Welch, 1995, 1996; Welch & Bishop, 1997; Welch et al., 1999). From the beginning, our efforts have been targeted at wide-area applications in particular. This focus was originally motivated by applications for which we believed that actually walking around the environment would be superior to virtually "flying." For example, we wanted to interact with room-filling virtual molecular models, and to naturally explore life-sized virtual architectural models. Today, we believe that a wide-area system with high performance everywhere in our laboratory provides increased flexibility for all of our graphics, vision, and interaction research.

#### 1.1 Previous Work

In the early 1960s, Ivan Sutherland implemented both mechanical and ultrasonic (carrier phase) head-tracking systems as part of his pioneering work in virtual environments. He describes these systems in his seminal paper "A Head-Mounted Three Dimensional Display" (Sutherland, 1968). In the

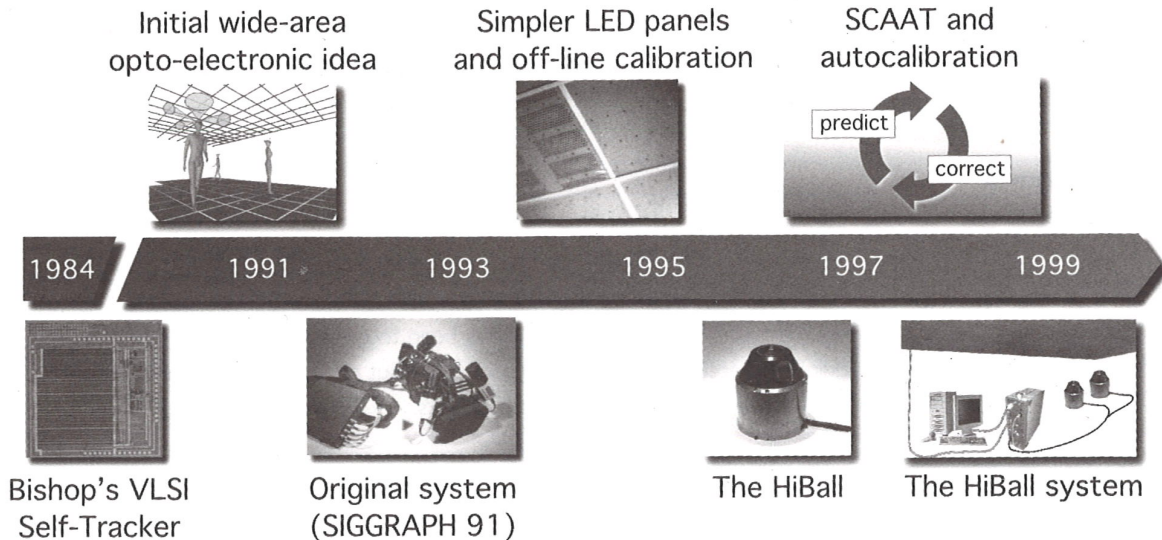


Figure 1.

ensuing years, commercial and research teams have explored mechanical, magnetic, acoustic, inertial, and optical technologies. Complete surveys include Bhatnagar (1993); Burdea & Coiffet (1994); Meyer, Applewhite, & Biocca (1992); and Mulder (1994a, 1994b, 1998). Commercial magnetic tracking systems for example (Ascension, 2000; Polhemus, 2000) have enjoyed popularity as a result of a small user-worn component and relative ease of use. Recently, inertial hybrid systems (Foxlin, Harrington, & Pfeifer, 1998; Intersense, 2000) have been gaining popularity for similar reasons, with the added benefit of reduced high-frequency noise and direct measurements of derivatives.

An early example of an optical system for tracking or motion capture is the Twinkle Box by Burton (Burton, 1973; Burton & Sutherland, 1974). This system measured the positions of user-worn flashing lights with optical sensors mounted in the environment behind rotating slotted disks. The *Selspot* system (Woltring, 1974) used fixed, camera-like, photodiode sensors and target-mounted infrared light-emitting diodes that could be tracked in a one-cubic-meter volume. Beyond the HiBall Tracking System, examples of current optical tracking and motion-capture systems include the *Flash-*

*Point* and *Pixsys* systems by Image Guided Technologies (IGT, 2000), the *laserBIRD* system by Ascension Technology (Ascension, 2000), and the *CODA Motion Capture System* by B & L Engineering (BL, 2000). These systems employ analog optical-sensor systems to achieve relatively high sample rates for a moderate number of targets. Digital cameras (two-dimensional, image-forming optical devices) are used in motion-capture systems such as the *HiRes 3D Motion Capture System* by the Motion Analysis Corporation (Kadaba & Stine, 2000; MAC, 2000) to track a relatively large number of targets, albeit at a relatively low rate because of the need for 2-D image processing.

## 1.2 Previous Work at UNC-Chapel Hill

As part of his 1984 dissertation on *Self-Tracker*, Bishop put forward the idea of outward-looking tracking systems based on user-mounted sensors that estimate user pose<sup>1</sup> by observing landmarks in the environment (Bishop, 1984). He described two kinds of

1. We use the word *pose* to indicate both position and orientation (six degrees of freedom).

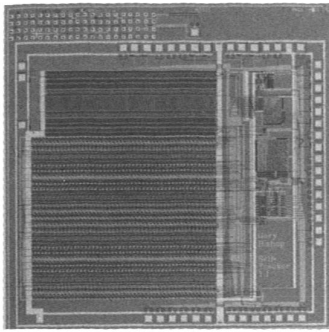


Figure 2.

landmarks: high signal-to-noise-ratio beacons such as light-emitting diodes (LEDs) and low signal-to-noise-ratio landmarks such as naturally occurring features. Bishop designed and demonstrated custom VLSI chips (figure 2) that combined image sensing and processing on a single chip (Bishop & Fuchs, 1984). The idea was to combine multiple instances of these chips into an outward-looking cluster that estimated cluster motion by observing natural features in the unmodified environment. Integrating the resulting motion to estimate pose is prone to accumulating error, so further development required a complementary system based on easily detectable landmarks (LEDs) at known locations. This LED-based system was the subject of a 1990 dissertation by Jih-Fang Wang (Wang, 1990).

In 1991, we demonstrated a working, scalable, electro-optical head-tracking system in the *Tomorrow's Realities* gallery at that year's ACM SIGGRAPH conference (Wang et al., 1990; Wang, Chi, & Fuchs, 1990; Ward et al., 1992). The system (figure 3) used four, head-worn, lateral-effect photodiodes that looked upward at a regular array of infrared LEDs installed in precisely machined ceiling panels. A user-worn backpack contained electronics that digitized and communicated the photo-coordinates of the sighted LEDs. Photogrammetric techniques were used to compute a user's head pose using the known LED positions and the corresponding measured photo-coordinates from each LEPD sensor (Azuma & Ward, 1991). The system was ground-breaking in that it was unaffected by ferromag-

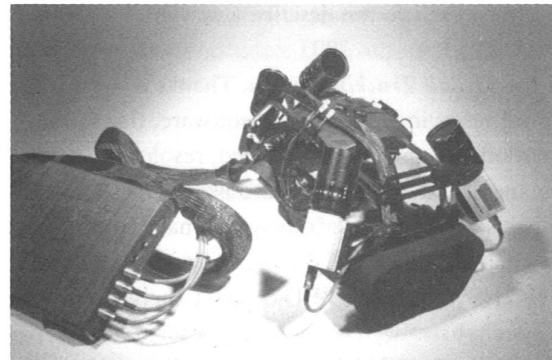
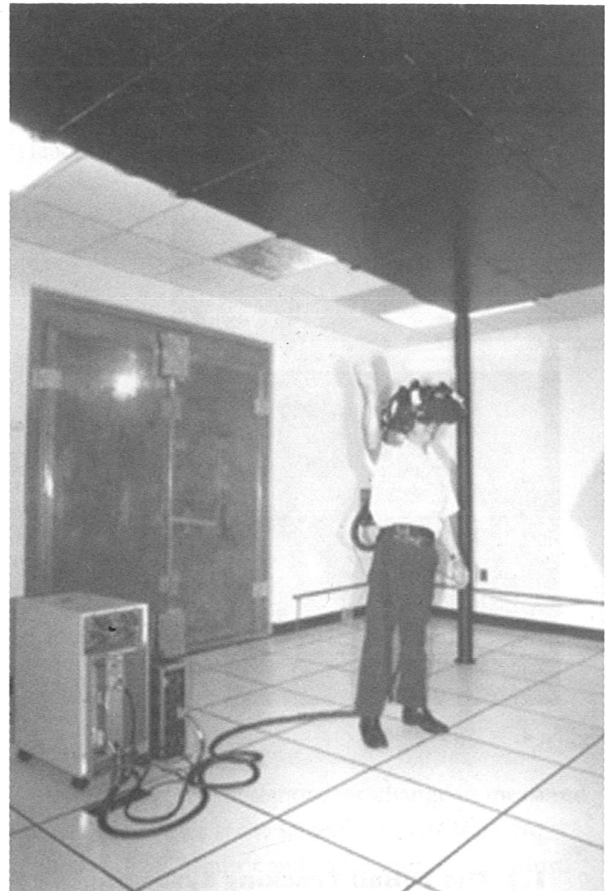


Figure 3.

netic and conductive materials in the environment, and the working volume of the system was determined solely by the number of ceiling panels. (See figure 3, top.)

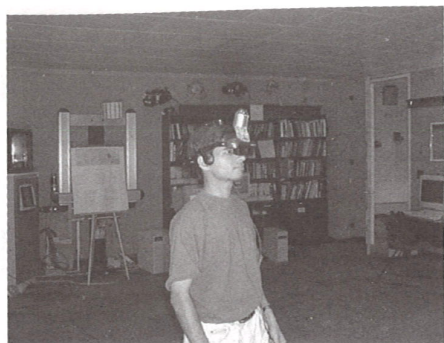


Figure 4.

### 1.3 The HiBall Tracking System

In this article, we describe a new and vastly improved version of the 1991 system. We call the new system the *HiBall Tracking System*. Thanks to significant improvements in hardware and software, this HiBall system offers unprecedented speed, resolution, accuracy, robustness, and flexibility. The bulky and heavy sensors and backpack of the previous system have been replaced by a small HiBall unit (figure 4, bottom). In addition, the precisely machined LED ceiling panels of the previous system have been replaced by looser-tolerance panels that are relatively inexpensive to make and simple to install (figure 4, top; figure 10). Finally, we are using an unusual Kalman-filter-based algorithm that generates very accurate pose estimates at a high rate with low latency, and that simultaneously self-calibrates the system.

As a result of these improvements, the HiBall Tracking System can generate more than 2,000 pose estimates per second, with less than 1 ms of latency, better

than 0.5 mm and 0.03 deg. of absolute error and noise, everywhere in a 4.5 m  $\times$  8.5 m room (with more than two meters of height variation). The area can be expanded by adding more panels, or by using checkerboard configurations that spread panels over a larger area. The weight of the user-worn HiBall is approximately 300 grams, making it lighter than one optical sensor in the 1991 system. Multiple HiBall units can be daisy-chained together for head or hand tracking, pose-aware input devices, or precise 3-D point digitization throughout the entire working volume.

## 2 Design Considerations

In all of the optical systems we have developed (see section 1.2), we have chosen what we call an *inside-looking-out* configuration, in which the optical sensors are on the (moving) user and the landmarks (for instance, the LEDs) are fixed in the laboratory. The corresponding *outside-looking-in* alternative would be to place the landmarks on the user and to fix the optical sensors in the laboratory. (One can think about similar outside-in and inside-out distinctions for acoustic and magnetic technologies.) The two configurations are depicted in figure 5.

There are some disadvantages to the inside-looking-out approach. For small or medium-sized working volumes, mounting the sensors on the user is more challenging than mounting them in the environment. It is difficult to make user-worn sensor packaging small, and communication from the moving sensors to the rest of the system is more complex. In contrast, there are fewer mechanical considerations when mounting sensors in the environment for an outside-looking-in configuration. Because landmarks can be relatively simple, small, and cheap, they can often be located in numerous places on the user, and communication from the user to the rest of the system can be relatively simple or even unnecessary. This is particularly attractive for full-body motion capture (BL, 2000; MAC, 2000).

However, there are some significant advantages to the inside-looking-out approach for head tracking. By operating with sensors on the user rather than in the

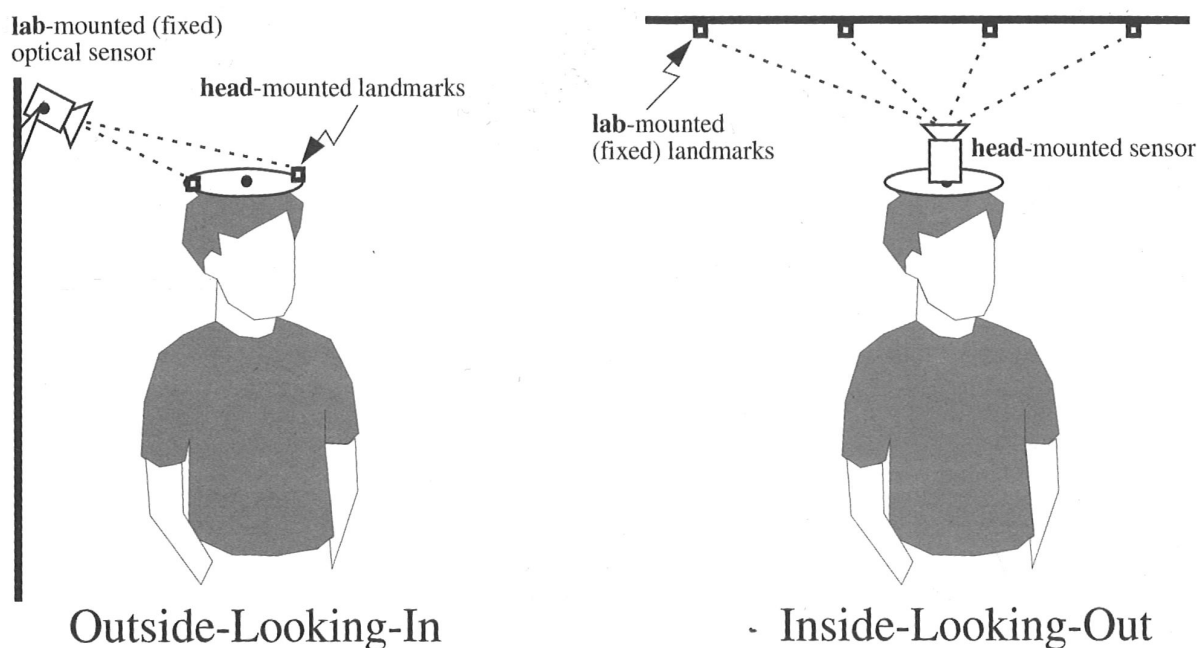


Figure 5.

environment, the system can be scaled indefinitely. The system can evolve from using dense active landmarks to fewer, lower signal-to-noise ratio, passive, and some day natural features for a Self-Tracker that operates entirely without explicit landmark infrastructure (Bishop, 1984; Bishop & Fuchs, 1984; Welch, 1995).

The inside-looking-out configuration is also motivated by a desire to maximize sensitivity to changes in user pose. In particular, a significant problem with an outside-looking-in configuration is that only position estimates can be made directly, and so orientation must be inferred from position estimates of multiple fixed landmarks. The result is that orientation sensitivity is a function of both the distance to the landmarks from the sensor and the baseline between the landmarks on the user. In particular, as the distance to the user increases or the baseline between the landmarks decreases, the sensitivity goes down. For sufficient orientation sensitivity, one would likely need a baseline that is considerably larger than the user's head. This would be undesirable from an ergonomic standpoint and could actually restrict the user's motion.

With respect to translation, the change in measured photo-coordinates is the same for an environment-mounted (fixed) sensor and user-mounted (moving) landmark as it is for a user-mounted sensor and an environment-mounted landmark. In other words, the translation and corresponding sensitivity are the same for either case.

### 3 System Overview

The HiBall Tracking System consists of three main components (figure 6). An outward-looking sensing unit we call the *HiBall* is fixed to each user to be tracked. The HiBall unit observes a subsystem of fixed-location infrared LEDs we call the *Ceiling*.<sup>2</sup> Communication and synchronization between the host computer and these subsystems is coordinated

2. At the present time, the LEDs are in fact entirely located in the ceiling of our laboratory (hence the subsystem name *Ceiling*), but LEDs could as well be located on walls or other fixed locations.



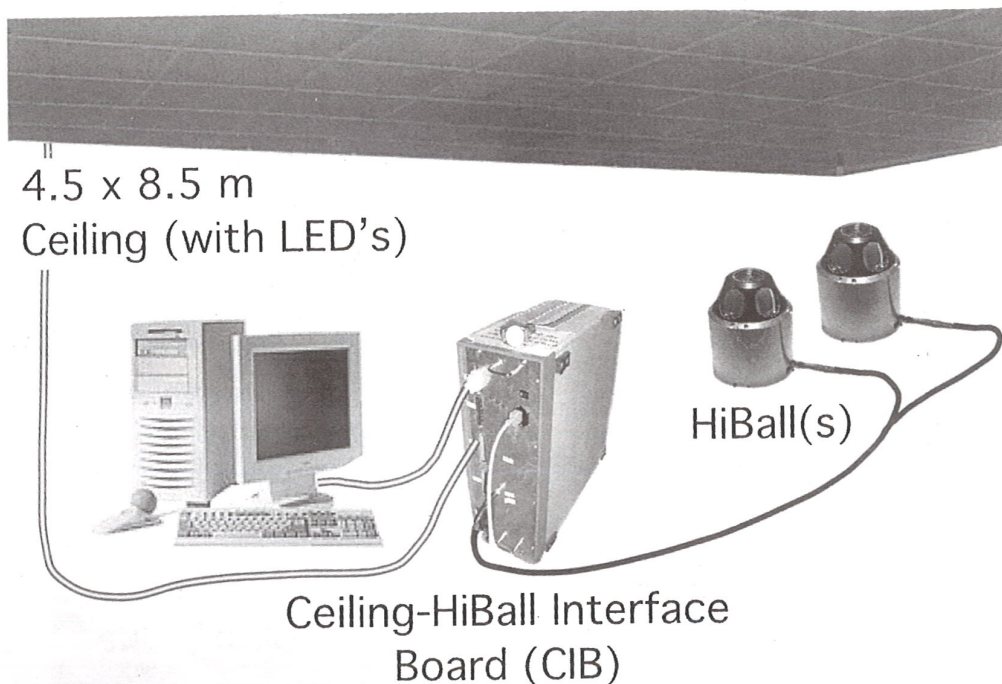


Figure 6.

by the *Ceiling-HiBall Interface Board (CIB)*. In section 4, we describe these components in more detail.

Each HiBall observes LEDs through multiple sensor-lens views that are distributed over a large solid angle. LEDs are sequentially flashed (one at a time) such that they are seen via a diverse set of views for each HiBall. Initial acquisition is performed using a brute-force search through LED space, but, once initial lock is made, the selection of LEDs to flash is tailored to the views of the active HiBall units. Pose estimates are maintained using a Kalman-filter-based prediction-correction approach known as *single-constraint-at-a-time (SCAAT)* tracking. This technique has been extended to provide self-calibration of the ceiling, concurrent with HiBall tracking. In section 5, we describe the methods we employ, including the initial acquisition process and the SCAAT approach to pose estimation, with the autocalibration extension.

## 4 System Components

### 4.1 The HiBall

The original electro-optical tracker (figure 3, bottom) used independently housed lateral-effect photodiode units (LEPDs) attached to a lightweight tubular framework. As it turns out, the mechanical framework would flex (distort) during use, contributing to estimation errors. In part to address this problem, the HiBall sensor unit was designed as a single, rigid, hollow ball having dodecahedral symmetry, with lenses in the upper six faces and LEPDs on the insides of the opposing six lower faces (figure 7). This immediately gives six primary “camera” views uniformly spaced by 57 deg. The views efficiently share the same internal air space and are rigid with respect to each other. In addition, light entering any lens sufficiently off-axis can be seen by a neighboring LEPD, giving rise to five secondary views through the top or central lens, and three secondary views

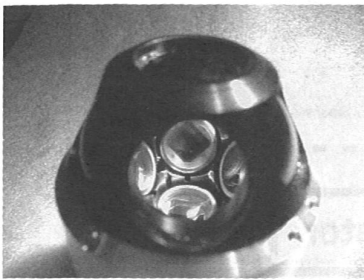
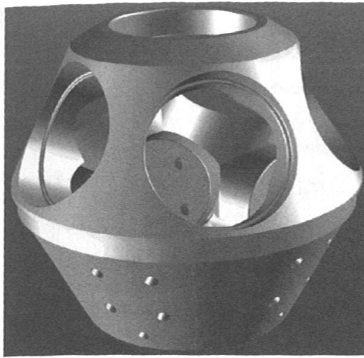


Figure 7.

through the five other lenses. Overall, this provides 26 fields of view that are used to sense widely separated groups of LEDs in the environment. Although the extra views complicate the initialization of the Kalman filter as described in section 5.5, they turn out to be of great benefit during steady-state tracking by effectively increasing the overall HiBall field of view without sacrificing optical-sensor resolution.

The lenses are simple plano-convex fixed-focus lenses. Infrared (IR) filtering is provided by fabricating the lenses themselves from RG-780 Schott glass filter material which is opaque to better than 0.001% for all visible wavelengths and transmissive to better than 99% for IR wavelengths longer than 830 nm. The longwave filtering limit is provided by the DLS-4 LEPD silicon photo-detector (UDT Sensors, Inc.) with peak responsivity at 950 nm but essentially blind above 1150 nm.

The LEPDs themselves are not imaging devices; rather, they detect the centroid of the luminous flux incident on the detector. The x-position of the centroid determines the ratio of two output currents, and the

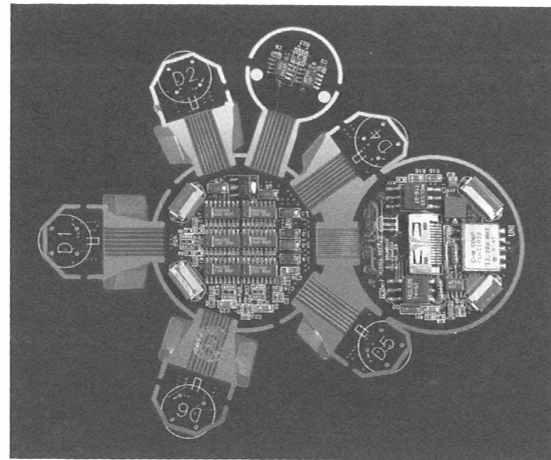


Figure 8.

y-position determines the ratio of two other output currents. The total output current of each pair are commensurate and are proportional to the total incident flux. Consequently, focus is not an issue, so the simple fixed-focus lenses work well over a range of LED distances from about half a meter to infinity. The LEPDs and associated electronic components are mounted on a custom rigid-flex printed circuitboard (figure 8). This arrangement makes efficient use of the internal HiBall volume while maintaining isolation between analog and digital circuitry, and increasing reliability by alleviating the need for intercomponent mechanical connectors.

Figure 9 shows the physical arrangement of the folded electronics in the HiBall. Each LEPD has four transimpedance amplifiers (shown together as one "Amp" in figure 9), the analog outputs of which are multiplexed with those of the other LEPDs, then sampled, held, and converted by four 16-bit Delta-Sigma analog-to-digital (A/D) converters. Multiple samples are integrated via an accumulator. The digitized LEPD data are organized into packets for communication back to the CIB. The packets also contain information to assist in error detection. The communication protocol is simple, and, while presently implemented by wire, the modulation scheme is amenable to a wireless implementation. The present wired implementation allows multiple HiBall units to be daisy-chained, so a single cable can support a user with multiple HiBall units.

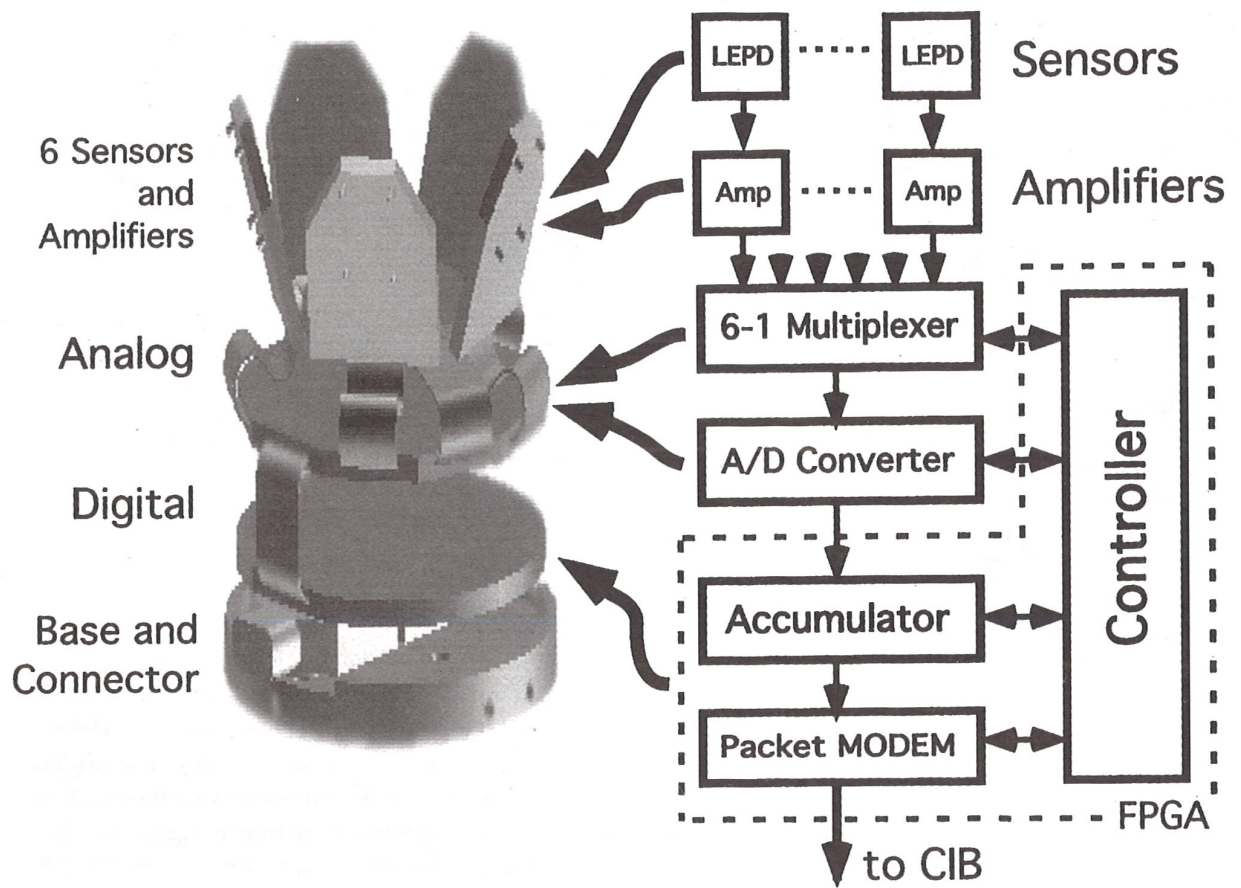


Figure 9.

#### 4.2 The Ceiling

As presently implemented, the infrared LEDs are packaged in 61 cm square panels to fit a standard false-ceiling grid (figure 10, top). Each panel uses five printed circuit boards: a main controller board and four identical transverse-mounted strips (bottom). Each strip is populated with eight LEDs for a total of 32 LEDs per panel. We mount the assembly on top of a metal panel such that the LEDs protrude through 32 corresponding holes. The design results in a ceiling with a rectangular LED pattern with periods of 7.6 cm and 15.2 cm. This spacing is used for the initial estimates of the LED positions in the lab; then, during normal operation, the SCAAT algorithm continually refines the LED position

estimates (section 5.4). The SCAAT autocalibration not only relaxes design and installation constraints, but provides greater precision in the face of initial and ongoing uncertainty in the ceiling structure.

We currently have enough panels to cover an area approximately 5.5 m by 8.5 m with a total of approximately 3,000 LEDs.<sup>3</sup> The panels are daisy-chained to each other, and panel-selection encoding is position (rather than device) dependent. Operational commands are presented to the first panel of the daisy chain. At each panel, if the panel-select code is zero, the

3. The area is actually L-shaped; a small storage room occupies one corner.

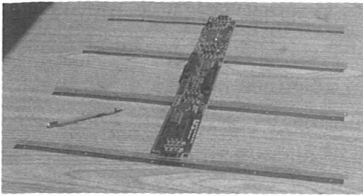


Figure 10.

controller decodes and executes the operation; otherwise, it decrements the panel-select code and passes it along to the next panel (controller). Upon decoding, a particular LED is selected and the LED is energized. The LED brightness (power) is selectable for automatic gain control as described in section 5.2.

We currently use Siemens SFH-487P GaAs LEDs, which provide both a wide-angle radiation pattern and high peak power, emitting at a center wavelength of 880 nm in the near IR. These devices can be pulsed up to 2.0 Amps for a maximum duration of 200  $\mu$ s with a 1:50 (on:off) duty cycle. Although the current ceiling architecture allows flashing of only one LED at a time, LEDs may be flashed in any sequence. As such, no single LED can be flashed too long or too frequently. We include both hardware and software protection to prevent this.

### 4.3 The Ceiling-HiBall Interface Board

The Ceiling-HiBall Interface Board (CIB) (figure 11) provides communication and synchronization between a host personal computer, the HiBall (section 4.1), and the ceiling (section 4.2). The CIB has four ceiling ports allowing interleaving of ceiling panels for up to four simultaneous LED flashes and/or higher

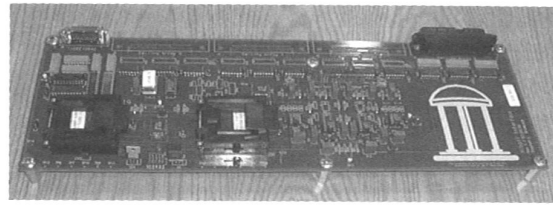


Figure 11.

ceiling bandwidth. (The ceiling bandwidth is inherently limited by LED power restrictions as described in section 4.2, but this can be increased by spatially multiplexing the ceiling panels.) The CIB has two tether interfaces that can communicate with up to four daisy-chained HiBall units. The full-duplex communication with the HiBall units uses a modulation scheme (BPSK) allowing future wireless operation. The interface from the CIB to the host PC is the stable IEEE1284C extended parallel port (EPP) standard.

The CIB comprises analog drive and receive components as well as digital logic components. The digital components implement store and forward in both directions and synchronize the timing of the LED “on” interval within the HiBall dark-light-dark intervals (section 5.2). The protocol supports full-duplex flow control. The data are arranged into packets that incorporate error detection.

## 5 Methods

### 5.1 Bench-Top (Offline) HiBall Calibration

After each HiBall is assembled, we perform an off-line calibration procedure to determine the correspondence between image-plane coordinates and rays in space. This involves more than just determining the view transform for each of the 26 views. Nonlinearities in the silicon sensor and distortions in the lens (such as spherical aberration) cause significant deviations from a simple pinhole camera model. We dealt with all of these issues through the use of a two-part camera model. The first part is a standard pinhole camera represented by a

$3 \times 4$  matrix. The second part is a table mapping real image-plane coordinates to ideal image-plane coordinates.

Both parts of the camera model are determined using a calibration procedure that relies on a goniometer (an angular positioning system) of our own design. This device consists of two servo motors mounted together such that one motor provides rotation about the vertical axis while the second motor provides rotation about an axis orthogonal to vertical. An important characteristic of the goniometer is that the rotational axes of the two motors intersect at a point at the center of the HiBall optical sphere; this point is defined as the origin of the HiBall. (It is this origin that provides the reference for the HiBall state during runtime as described in section 5.3.) The rotational positioning motors were rated to provide twenty arc-second precision; we further calibrated them to six arc seconds using a laboratory grade theodolite—an angle measuring system.

To determine the mapping between sensor image-plane coordinates and three-space rays, we use a single LED mounted at a fixed location in the laboratory such that it is centered in the view directly out of the top lens of the HiBall. This ray defines the  $z$  or up axis for the HiBall coordinate system. We sample other rays by rotating the goniometer motors under computer control. We sample each view with rays spaced about every six minutes of arc throughout the field of view. We repeat each measurement 100 times to reduce the effects of noise on the individual measurements and to estimate the standard deviation of the measurements.

Given the tables of approximately 2,500 measurements for each of the 26 views, we first determine a  $3 \times 4$  view matrix using standard linear least-squares techniques. Then, we determine the deviation of each measured point from that predicted by the ideal linear model. These deviations are resampled into a  $25 \times 25$  grid indexed by sensor-plane coordinates using a simple scan-conversion procedure and averaging. Given a measurement from a sensor at runtime (section 5.2), we convert it to an “ideal” measurement by subtracting a deviation bilinearly interpolated from the nearest four entries in the table.

## 5.2 Online HiBall Measurements

Upon receiving a command from the CIB (section 4.3), which is synchronized with a CIB command to the ceiling, the HiBall selects the specified LEPD and performs three measurements, one before the LED flashes, one during the LED flash, and one after the LED flash. Known as “dark-light-dark,” this technique is used to subtract out DC bias, low-frequency noise, and background light from the LED signal. We then convert the measured sensor coordinates to “ideal” coordinates using the calibration tables described in section 5.1.

In addition, during runtime we attempt to maximize the signal-to-noise ratio of the measurement with an automatic gain-control scheme. For each LED, we store a target signal strength factor. We compute the LED current and number of integrations (of successive accumulated A/D samples) by dividing this strength factor by the square of the distance to the LED, estimated from the current position estimate. After a reading, we look at the strength of the actual measurement. If it is larger than expected, we reduce the gain; if it is less than expected, we increase the gain. The increase and decrease are implemented as online averages with scaling such that the gain factor decreases rapidly (to avoid overflow) and increases slowly. Finally, we use the measured signal strength to estimate the noise on the signal using (Chi, 1995), and then use this as the measurement noise estimate for the Kalman filter (section 5.3).

## 5.3 Recursive Pose Estimation (SCAAT)

The online measurements (section 5.2) are used to estimate the pose of the HiBall during operation. The 1991 system collected a group of diverse measurements for a variety of LEDs and sensors, and then used a method of simultaneous nonlinear equations called *collinearity* (Azuma & Ward, 1991) to estimate the pose of the sensor fixture shown in figure 3 (bottom). There was one equation for each measurement, expressing the constraint that a ray from the front principal point of the sensor lens to the LED must be collinear with a ray from the rear principal point to the intersection with the

sensor. Each estimate made use of a group of measurements (typically twenty or more) that together overconstrained the solution.

This *multiple constraint* method had several drawbacks. First, it had a significantly lower estimate rate due to the need to collect multiple measurements per estimate. Second, the system of nonlinear equations did not account for the fact that the sensor fixture continued to move throughout the collection of the sequence of measurements. Instead, the method effectively assumes that the measurements were taken simultaneously. The violation of this simultaneity assumption could introduce significant error during even moderate motion. Finally, the method provided no means to identify or handle unusually noisy individual measurements. Thus, a single erroneous measurement could cause an estimate to jump away from an otherwise smooth track.

In contrast, the approach we use with the new HiBall system produces tracker reports as each new measurement is made, rather than waiting to form a complete collection of observations. Because single measurements underconstrain the mathematical solution, we refer to the approach as *single-constraint-at-a-time (SCAAT)* tracking (Welch, 1996; Welch & Bishop, 1997). The key is that the single measurements provide *some* information about the HiBall's state, and thus can be used to incrementally improve a previous estimate. We intentionally fuse each individual "insufficient" measurement immediately as it is obtained. With this approach, we are able to generate estimates more frequently, with less latency, and with improved accuracy, and we are able to estimate the LED positions online concurrently while tracking the HiBall (section 5.4).

We use a Kalman filter (Kalman, 1960) to fuse the measurements into an estimate of the HiBall *state*  $\bar{x}$  (the pose of the HiBall). We use the Kalman filter—a minimum-variance stochastic estimator—both because the sensor measurement noise and the typical user-motion dynamics can be modeled as normally distributed random processes, and because we want an efficient online method of estimation. A basic introduction to the Kalman filter can be found in chapter 1 of Maybeck (1979), and a more complete introductory discussion can be found in Sorenson (1970), which also contains

some interesting historical narrative. More-extensive references can be found in Brown and Hwang (1992), Gelb (1974), Jacobs (1993), Lewis (1986), Maybeck (1979), and Welch and Bishop (1995). Finally, we maintain a Kalman filter Web page (Welch & Bishop, 2000) with introductory, reference, and research material.

The Kalman filter has been used previously to address similar or related problems. See, for example, Azarbayejani and Pentland (1995), Azuma (1995), Emura and Tachi (1994), Fuchs (Foxlin) (1993), Mazuryk and Gervautz (1995), and Van Pabst and Krekel (1993). A relevant example of a Kalman filter used for sensor fusion in a wide-area tracking system is given in Foxlin et al. (1998), which describes a hybrid inertial-acoustic system that is commercially available today (Intersense, 2000).

The SCAAT approach is described in detail by Welch (1996), and Welch and Bishop (1997). Included there is discussion of the benefits of using the approach, as opposed to a multiple-constraint approach such as that by Azuma and Ward (1991). However, one key benefit warrants discussion here. There is a direct relationship between the *complexity* of the estimation algorithm, the corresponding *speed* (execution time per estimation cycle), and the *change* in HiBall pose between estimation cycles (figure 12). As the algorithmic complexity increases, the execution time increases, which allows for significant nonlinear HiBall motion between estimation cycles, which in turn implies the need for a more complex estimation algorithm.

The SCAAT approach, on the other hand, is an attempt to reverse this cycle. Because we intentionally use a single constraint per estimate, the algorithmic complexity is drastically reduced, which reduces the execution time, and hence the amount of motion between estimation cycles. Because the amount of motion is limited, we are able to use a simple dynamic (process) model in the Kalman filter, which further simplifies the computations. In short, the simplicity of the approach means that it can run very fast, which means it can produce estimates very rapidly, with low noise.

The Kalman filter requires both a model of the process dynamics and a model of the relationship between



Figure 12.

the process state and the available measurements. In part due to the simplicity of the SCAAT approach, we are able to use a simple position-velocity (PV) process model (Brown & Hwang, 1992). Consider the simple example state vector  $\bar{x}(t) = [x_p(t), x_v(t)]^T$ , where the first element  $x_p(t)$  is the pose (position or orientation) and the second element  $x_v(t)$  is the corresponding velocity; that is,  $x_v(t) = (d/dt) x_p(t)$ . We model the continuous change in the HiBall state with the simple differential equation

$$\frac{d}{dt} \bar{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_p(t) \\ x_v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \mu \end{bmatrix} u(t), \quad (1)$$

where  $u(t)$  is a normally distributed white (in the frequency spectrum) scalar noise process, and the scalar  $\mu$  represents the magnitude or spectral density of the noise. We use a similar model with a distinct noise process for each of the six pose elements. We determine the individual noise magnitudes using an offline simulation of the system and a nonlinear optimization strategy that seeks to minimize the variance between the estimated pose and a known motion path. (See section 6.2.2.). The differential equation (1) represents a continuous integrated random walk, or an integrated Wiener or Brownian-motion process. Specifically, we model each

component of the linear and angular HiBall velocities as a random walk, and then use these (assuming constant intermeasurement velocity) to estimate the HiBall pose at time  $t + \delta t$  as follows:

$$\bar{x}(t + \delta t) = \begin{bmatrix} 1 & \delta t \\ 0 & 1 \end{bmatrix} \bar{x}(t) \quad (2)$$

for each of the six pose elements. In addition to a relatively simple process model, the HiBall measurement model is relatively simple. For any ceiling LED (section 4.2) and HiBall view (section 4.1), the 2-D sensor measurement can be modeled as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} c_x/c_z \\ c_y/c_z \end{bmatrix} \quad (3)$$

where

$$\begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} = VR^T(\bar{l}_{xyz} - \bar{x}_{xyz}), \quad (4)$$

$V$  is the camera viewing matrix from section 5.1,  $\bar{l}_{xyz}$  is the position of the LED in the world,  $\bar{x}_{xyz}$  is the position of the HiBall in the world, and  $R$  is a rotation matrix corresponding to the orientation of the HiBall in the world. In practice, we maintain the orientation of the HiBall as a combination of a global (external to the state) quaternion and a set of incremental angles as described by Welch (1996) and Welch and Bishop (1997).

Because the measurement model (3) and (4) is nonlinear, we use an extended Kalman filter, making use of the Jacobian of the nonlinear HiBall measurement model to transform the covariance of the Kalman filter. Although this approach does not preserve the presumed Gaussian nature of the process, it has been used successfully in countless applications since the introduction of the (linear) Kalman filter. Based on observations of the statistics of the HiBall filter residuals, the approach also appears to work well for the HiBall. In fact, it is reasonable to expect that it would, as the speed of the SCAAT approach minimizes the distance (in state space) over which we use the Jacobian-based linear approximation. This is another example of the importance of the relationship shown in figure 12.

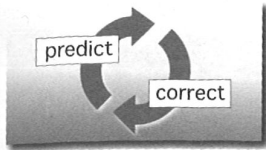


Figure 12a.

At each estimation cycle, the next of the 26 possible views is chosen randomly. Four points corresponding to the corners of the LEPD sensor associated with that view are projected into the world using the  $3 \times 4$  viewing matrix for that view, along with the current estimates of the HiBall pose. This projection, which is the inverse of the measurement relationship described above, results in four rays extending from the sensor into the world. The intersection of these rays and the approximate plane of the ceiling determines a 2-D bounding box on the ceiling, within which are the candidate LEDs for the current view. One of the candidate LEDs is then chosen in a least-recently-used fashion to ensure a diversity of constraints.

Once a particular view and LED have been chosen in this fashion, the CIB (section 4.3) is instructed to flash the LED and take a measurement as described in section 5.2. This single measurement is compared with a prediction obtained using equation (3), and the difference (or *residual*) is used to update the filter state and covariance matrices using the Kalman gain matrix. The Kalman gain is computed as a combination of the current filter covariance, the measurement noise variance (section 6.2.1), and the Jacobian of the measurement model. This recursive prediction-correction cycle continues in an ongoing fashion, a single constraint at a time.

A more detailed discussion of the HiBall Kalman filter and the SCAAT approach is beyond the scope of this paper. For additional information see Welch (1996) and Welch and Bishop (1997).

#### 5.4 Online LED Autocalibration

Along with the benefit of simplicity and speed, the SCAAT approach offers the additional capability of be-

ing able to estimate the 3-D positions of the LEDs in the world concurrently with the pose of the HiBall, online, in real time. This capability is a tremendous benefit in terms of the accuracy and noise characteristics of the estimates. Accurate LED position estimates are so important that, prior to the introduction of the SCAAT approach, a specialized offline approach was developed to address the problem (Gottschalk & Hughes, 1993).

The method we now use for autocalibration involves defining a distinct SCAAT Kalman filter for each LED. Specifically, for each LED, we maintain a state  $\bar{l}$  (estimate of the 3-D position) and a  $3 \times 3$  Kalman filter covariance. At the beginning of each estimation cycle, we form an augmented state vector  $\bar{x}$  using the appropriate LED state and the current HiBall state:  $\bar{x} = [\bar{x}^T, \bar{l}^T]^T$ . Similarly, we augment the Kalman filter error covariance matrix with that of the LED filter. We then follow the normal steps outlined in section 5.3, with the result being that the LED portion of the filter state and covariance is updated in accordance with the measurement residual. At the end of the cycle, we extract the LED portions of the state and covariance from the augmented filter, and save them externally. The effect is that, as the system is being used, it continually refines its estimates of the LED positions, thereby continually improving its estimates of the HiBall pose. Again, for additional information, see Welch (1996) and Welch and Bishop (1997).

#### 5.5 Initialization and Reacquisition

The recursive nature of the Kalman filter (section 5.3) requires that the filter be initialized with a known state and corresponding covariance before steady-state operation can begin. Such an initialization (or *acquisition*) must take place prior to any tracking session, but also upon the (rare) occasion when the filter diverges and “loses lock” as a result of blocked sensor views, for example.

The acquisition process is complicated by the fact that each LEPD sees a number of different widely separated views (section 4.1). Therefore, detecting an LED provides at best an ambiguous set of potential LED directions in HiBall coordinates. Moreover, before acquisition,



no assumptions can be made to limit the search space of visible LEDs. As such, a relatively slow brute-force algorithm is used to acquire lock.

We begin with an exhaustive LED scan of sufficiently fine granularity to ensure that the central primary field of view is not missed. For the present ceiling, we flash every thirteenth LED in sequence, and look for it with the central LEPD until we get a hit. Then, a sufficiently large patch of LEDs, centered on the hit, is sampled to ensure that several of the views of the central LEPD will be hit. The fields of view are disambiguated by using the initial hits to estimate the yaw of the HiBall (rotation about vertical); finally, more-selective measurements are used to refine the acquisition estimate sufficiently to switch into tracking mode.

## 6 Results

Three days after the individual pieces of hardware were shown to be functioning properly, we demonstrated a complete working system. After months of subsequent tuning and optimization, the system continues to perform both qualitatively and quantitatively as well—or, in some respects, better—than we had anticipated (section 6.1). The articulation of this success is not meant to be self-congratulatory, but to give credit to the extensive and careful modeling and simulation performed prior to assembly (section 6.2). In fact, the Kalman filter parameters found by the optimization procedure described in section 6.2.2 were, and continue to be, used directly in the working system. Likewise, much of the software written for the original simulations continues to be used in the working system.

### 6.1 Online Operation

The HiBall system is in daily use as a tool for education and research. For example, it was used by Martin Usuh et al. to perform virtual reality experiments comparing virtual “flying,” walking in place, and real walking (Usuh et al., 1999). (See figure 13.) The researchers used the HiBall system to demonstrate that, as a mode of locomotion, real walking is simpler, more straightfor-

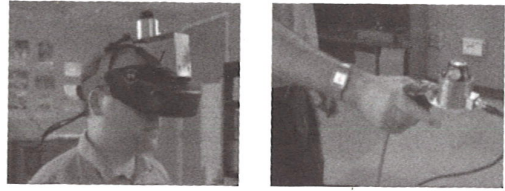


Figure 13.

ward, and more natural, than both virtual flying and walking in place. The unprecedented combination of large working volume and the high performance of the HiBall system led the researchers to claim that there was nowhere else that they could have meaningfully performed the experiments.

**6.1.1 Robustness.** As a result of a mechanical design tradeoff, each sensor field of view is less than six degrees. The focal length is set by the size of the sensor housing, which is set by the diameter of the sensors themselves. Energetics is also a factor, limiting how small the lenses can be while maintaining sufficient light-collecting area. As a result of these design tradeoffs, even a momentary small error in the HiBall pose estimate can cause the recursive estimates to diverge and the system to lose lock after only a few LED sightings. And yet the system is quite robust. In practice, users can jump around, crawl on the floor, lean over, even wave their hands in front of the sensors, and the system does not lose lock. During one session, we were using the HiBall as a 3-D digitization probe, a HiBall on the end of a pencil-shaped fiberglass wand (figure 14, left). We laid the probe down on a table at one point, and were amazed to later notice that it was still tracking, even though it was observing only three or four LEDs near the edge of the ceiling. We picked up the probe and continued using it, without it ever losing lock.

**6.1.2 Estimate Noise.** The simplest quantitative measurement of estimate noise is the standard deviation of the estimates when a HiBall is held stationary. With a tracker as sensitive as the HiBall, it is important to be certain that it really is stationary. The raised floor in our

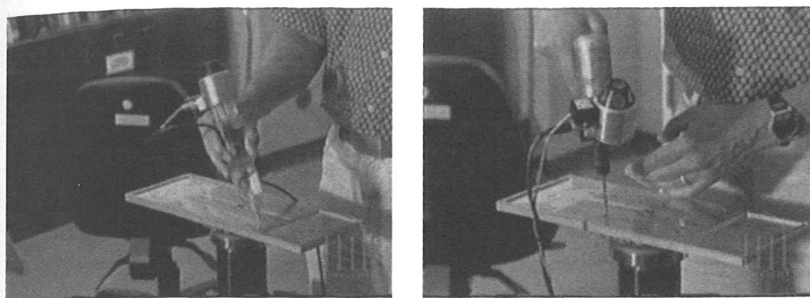


Figure 14.

laboratory allows motion, for example when a person walks by, that is larger than the expected error in the HiBall. We made careful measurements by resting the support for the HiBall on the concrete subfloor in our laboratory. The standard deviation of the HiBall estimates while stationary was approximately 0.2 mm and 0.03 deg. The distribution of the noise fit a normal distribution quite well.

To make measurements of the noise when the HiBall is in motion, we rely on the assumption that almost all of the signal resulting from normal human motion is at frequencies below 2 Hz. We use a high-pass filter (Welch, 1967) on the pose estimates, and assume the output is noise. The resulting statistics are comparable to those made with the HiBall stationary, except at poses for which there are very few LEDs visible in only one or two views. In these poses, near the edge of the ceiling, the geometry of the constraints results in amplification of errors. For nearly all of the working volume of the tracker, the standard deviation of the noise on measurements while the HiBall is still or moving is about 0.2 mm and 0.03 deg.

**6.1.3 Absolute Accuracy.** We have performed several experiments to measure the accuracy of the HiBall system; however, the most objective experiment took place in July of 1999. Boeing Phantom Works scientists David Himmel and David Princehouse (Associate Technical Fellows) visited our laboratory for two days to assess the accuracy of the HiBall system and its potential use in providing assembly workers with real-time feed-

back on the pose of handheld pneumatic drills during the aircraft manufacturing process. (The right image in figure 14 shows the HiBall attached to a pneumatic drill.)

The scientists designed some controlled experiments to assess the accuracy of the HiBall system. They brought with them an aluminum "coupon" (see figure 14 and figure 15) with 27 shallow holes drilled on 1.5-in. centers using a numerically controlled milling machine with a stated accuracy of 1/1000 in. The holes (except one) were not actually drilled through the coupon, but instead formed conical dimples with a fine point at the center. The center-most hole (hole 14) was actually drilled completely through to provide a mounting point. Using that hole, we attached the coupon to a military-grade tripod situated on the (false) floor of our laboratory, under the HiBall ceiling. As shown in the left image of figure 14, we mounted the HiBall on our standard probe, a rigid plastic, pencil-like object with a pointed steel tip. We used one of the coupon holes to perform our normal HiBall probe calibration procedure, which involves placing the tip of the probe in the hole, pivoting the probe about the point while collecting several seconds of pose data, and then estimating the transformation from the HiBall to the probe tip. (We have a standard application that assists us with this procedure.) Together with Himmel and Princehouse, we performed several experiments in which we placed the tip of the HiBall probe in each hole in succession, sampling the HiBall pose estimates only when we pressed the probe button. We performed several such sessions over the

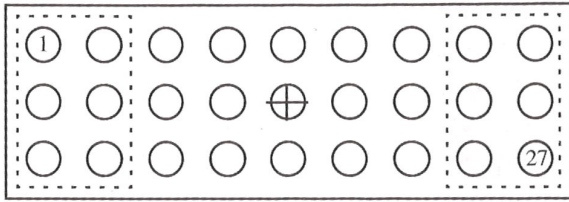


Figure 15.

course of one afternoon and the next morning. (We recalibrated the probe in the morning.)

For the data from each session, we used a least-squares optimization method to find an estimate of the full 6-D transformation (translation and rotation) that minimized the Euclidian distance from the probe data to a 2-D plane with 27 holes on 1.5-in. spacing. The resulting fit consistently corresponded to an average positioning error of 20/1000 in. (0.5 mm) at the metal tip of the HiBall probe, which is within the target Boeing specifications. The system might actually be more accurate than our experiments indicated. For one, the diameter of the (rounded) tip of the HiBall probe is 0.5 mm. In addition, at the time of the experiments, we unfortunately did not heed our own advice to position the experimental platform on the rigid concrete sub-floor. In any case, we are encouraged by the results, and are excited about the possibility that the HiBall system has uses beyond tracking for virtual reality.

## 6.2 Offline Simulation and Modeling

During the design of the HiBall system, we made substantial use of simulation, in some domains to a very detailed level. For example, Zemax (Focus Software, 1995) was used extensively in the design and optimization of the optical design, including the design of the filter glass lenses, and geometry of the optical-component layout. AutoCAD was used to design, specify, and fit check the HiBall body mechanicals, to visualize the physical design, and to transmit the design to our collaborators at the University of Utah for fabrication by the Alpha 1 System (Thomas, 1984; University of Utah Computer Science, 1999). A custom ray-tracing system

was built by Stefan Gottschalk (UNC) for the purpose of evaluating the optical behavior and energetics of the primary, secondary, and tertiary fields of view; the results were used by the noise model developed by Chi (1995) as described in section 6.2.1.

In addition, a complete simulator of the system was written in C++. This simulator, discussed further in section 6.2.2, was used to evaluate the speed, accuracy, and robustness of the system. In addition, it was used to “tune” the Kalman filter for realistic motion dynamics. This simulator continues to be used to evaluate mechanical, optical, and algorithmic alternatives.

### 6.2.1 HiBall Measurement Noise Model.

Signal-to-noise performance is a prime determiner of both accuracy and speed of the system, so an in-depth study (Chi, 1995) was performed to develop a detailed noise model accounting for properties of the LED, the LEPD (sensor), the optical system, the physical distance and pose, the electronics, and the dark-light-dark integrations described in section 5.2. The predominant noise source is shot noise, with Johnson noise in the sheet resistivity of the LEPD surfaces being the next most significant. Careful measurements made in the laboratory with the actual devices yielded results that were almost identical to those predicted by the sophisticated model in Chi (1995). A simplified version of this model is used in the real system with the automatic gain control (section 5.2) to predict the measurement noise for the Kalman filter (section 5.3).

### 6.2.2 Complete System Simulations.

To produce realistic data for developing and tuning our algorithms, we collected several motion paths (sequences of pose estimates) from our first-generation electro-optical tracker (figure 3) at its 70 Hz maximum report rate. These paths were recorded from both naive users visiting our monthly “demo days” and from experienced users in our labs. In the same fashion as we had done for Azuma and Bishop (1994a), we filtered the raw path data with a noncausal zero-phase-shift, low-pass filter to eliminate energy above 2 Hz. The output of the low-pass filtering was then resampled at whatever rate we wanted to run the simulated tracker, usually 1,000 Hz.

For the purposes of our simulations, we considered these resampled paths to be the “truth”—a perfect representation of a user’s motion. Tracking error was determined by comparing the true path to the estimated path produced by the tracker.

The simulator reads camera models describing the 26 views, the sensor noise parameters, the LED positions and their expected error, and the motion path described above. Before beginning the simulation, the LED positions are perturbed from their ideal positions by adding normally distributed error to each axis. Then, for each simulated cycle of operation, the “true” poses are updated using the input motion path. Next, a view is chosen and a visible LED within that view is selected, and the image-plane coordinates of the LED on the chosen sensor are computed using the camera model for the view and the LED as described in section 5.3. These sensor coordinates are then perturbed based on the sensor noise model (section 6.2.1) using the distance and angle to the LED. These noise-corrupted sensor readings are then fed to the SCAAT filter to produce an updated position estimate. The position estimate is compared to the true position to produce a scalar error metric that is described next.

The error metric we used combines the error in pose in a way that relates to the effects of tracker error on a head-worn display user. We define a set of points arrayed around the user in a fixed configuration. We compute two sets of coordinates for these points: the true position using the true pose and their estimated position using the estimated pose. The error metric is then the sum of the distances between the true and estimated positions of these points. By adjusting the distance of the points from the user, we can control the relative importance of the orientation and the position error in the combined error metric. If the distance is small, then the position error is weighted most heavily; if the distance is large, then the orientation error is weighted most heavily. Our two error metrics for the entire run are the square root of the sum of the squares of all the distances, and the peak distance.

**6.2.3 Tuning.** Determining the magnitudes of the SCAAT Kalman filter noise parameters (section 5.3)

is called *system identification* or *tuning*. We use Powell’s method (Press, Teukolsky, Vetterling, & Flannery, 1990) to minimize the error metric described above. Starting with a set of parameters, we run the simulator over a full motion run to determine the total error for the run. The optimizer makes a small adjustment to the parameters and the process is repeated. These runs required hours of computer time and some skill (and luck) in choosing the initial parameters and step sizes. Of course, it is important to choose motion paths that are representative of expected target motion. For example, a run in which the target is very still would result in very different tuning from a run in which the target moves very vigorously.

## 7 Future Work

### 7.1 Improving the HiBall

The current SCAAT filter form (section 5.3) and tuning values (section 6.2.3) are a compromise between the responsiveness desired for high dynamics, and the heavy filtering desired for smooth estimates during very slow or no motion. As such, we are investigating the use of a multimodal or multiple-model Kalman filter framework (Bar-Shalom & Li, 1993; Brown & Hwang, 1992). A multiple-model implementation of the HiBall should be able to automatically, continuously, and smoothly choose between one Kalman filter tuned for high dynamics and another tuned for little or no motion. We have this working in simulation, but not yet implemented in the real system.

As mentioned in section 4.3, the system was designed to support wireless communication between the HiBall and the CIB, without significant modification or added information overhead. Despite the fact that commercial head-worn displays are themselves tethered at this time, we are beginning work on a completely wireless HiBall and head-worn display system. We also intend to use the wireless HiBall with projector-based displays where the user is otherwise wearing only polarized glasses. Furthermore, the HiBall was designed with extra built-in digital input-output capabilities. We are considering possibilities for providing access to these signals

for (wireless) user-centered input devices and even body-centric limb tracking.

Finally, we note that a private startup company called 3rdTech (3rdTech, 2000) has negotiated a technology license with UNC for the existing HiBall Tracking System. 3rdTech is now marketing an updated system with simpler LED "strips" instead of ceiling panels.

## 7.2 Wide-Field-of-View HiBall

Beyond improving the existing system, we continue to head down a path of research and development that will lead to systems with reduced dependency on the laboratory infrastructure. For example, our current ceiling-panel design with 32 LEDs per panel provides far more dense coverage than we believe is necessary. The density of ceiling LEDs is a result of design based on the original sensor fixture shown in figure 3. Given a more sparse field of LEDs, we believe that we could achieve similar performance with a version of the HiBall that has a small number of wide-field-of-view optical sensor units. This would further reduce the packaging size of the user-worn sensor component.

## 7.3 To the Hallway and Beyond

By leveraging the knowledge gained from successful work in the laboratory, our long-term goal is to achieve similar performance with little or no explicit infrastructure: for example, throughout a building or even (some day) outdoors. Although high-performance 6-D tracking outdoors is a tremendous challenge that is unlikely to be solved any time soon, we believe that the eventual solution will involve a clever and careful combination of multiple complementary technologies. In particular, we are pursuing the hybrid approach initially presented by Welch (1995). We look forward to a day when high-performance 6-D tracking outdoors enables pose-aware devices for work such as Feiner's outdoor augmented reality (Feiner, MacIntyre, Höllerer, & Webster, 1997; Höllerer, Feiner, Terauchi, Rashid, & Hallaway, 1999), the "WorldBoard" initiative (Spohrer, 1999a, 1999b), and other wonderful applications.

## Acknowledgments

We acknowledge former Tracker Project members and contributors (alphabetically): Ronald Azuma, Henry Fuchs, Stefan Gottschalk, Pawan Kumar, John Thomas, Jih-Fang Wang, Mark Ward, Scott Williams, Mary Whitton, and Philip Winston. We thank Al Barr (California Institute of Technology) and John "Spike" Hughes (Brown University) for their contributions to the original offline LED calibration work that led to the simpler ceiling panels (figure 1 and figure 10). Finally, we want to acknowledge our many collaborators in the NSF Science and Technology Center for Computer Graphics and Scientific Visualization (below), and in particular our collaborators in mechanical design and fabrication at the University of Utah: Rich Riesenfeld, Sam Drake, and Russ Fish.

This work was supported in part by DARPA/ETO contract DABT 63-93-C-0048 "Enabling Technologies and Application Demonstrations for Synthetic Environments" (Principal Investigators Frederick P. Brooks, Jr. and Henry Fuchs (UNC)), and by the National Science Foundation Cooperative Agreement ASC-8920219 "Science and Technology Center for Computer Graphics and Scientific Visualization," Center Director Rich Riesenfeld (University of Utah) (Principal Investigators Al Barr (Caltech), Don Greenberg (Cornell University), Henry Fuchs (UNC), Rich Riesenfeld, and Andy van Dam (Brown University)).

## References

- 3rdTech. (2000, July 15). *3rdTech*<sup>™</sup> [Webpage]. Retrieved July 19, 2000, from the World Wide Web: <http://www.3rdtech.com/>.
- Ascension. (2000). *Ascension Technology Corporation* [Webpage]. Retrieved September 15, 2000 from the World Wide Web: <http://www.ascension-tech.com/> [2000, September 15].
- Azarbayejani, A., & Pentland, A. (1995). Recursive estimation of motion, structure, and focal length. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(6), 562-575.
- Azuma, R. T. (1993). Tracking requirements for augmented reality. *Communications of the ACM*, 36(July), 50-51.
- . (1995). *Predictive tracking for augmented reality*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill, Chapel Hill.
- Azuma, R. T., & Bishop, G. (1994a). A frequency-domain

- analysis of head-motion prediction. *Computer Graphics (SIGGRAPH 94 Conference Proceedings ed.)* (pp. 401–408). Los Angeles: ACM Press, Addison-Wesley.
- . (1994b). Improving static and dynamic registration in an optical see-through hMD. *Computer Graphics (SIGGRAPH 94 Conference Proceedings ed.)* (pp. 197–204). Orlando: ACM Press, Addison-Wesley.
- Azuma, R. T., & Ward, M. (1991). *Space-resection by collinearity: Mathematics behind the optical ceiling head-tracker* (Tech. Rep. 91-048). Chapel Hill: University of North Carolina at Chapel Hill.
- Bar-Shalom, Y., & Li, X.-R. (1993). *Estimation and Tracking: Principles, Techniques, and Software*. Boston: Artec House, Inc.
- Bhatnagar, D. K. (1993). *Position trackers for head mounted display systems: A survey* (Tech. Rep. TR93-010). Chapel Hill: University of North Carolina at Chapel Hill.
- Bishop, G. (1984). *The self-tracker: A smart optical sensor on silicon*. Unpublished doctoral Dissertation, University of North Carolina at Chapel Hill, Chapel Hill.
- Bishop, G., & Fuchs, H. (1984, January 23–25). *The self-tracker: A smart optical sensor on silicon*. Paper presented at the Advanced Research in VLSI conference, Massachusetts Institute of Technology.
- BL. (2000). *CODA mpx30 Motion Capture System [HTML]*. *B & L Engineering*. Retrieved April 27, 2000 from the World Wide Web: <http://www.bleng.com/animation/coda/codamain.htm>.
- Brown, R. G., & Hwang, P. Y. C. (1992). *Introduction to Random Signals and Applied Kalman Filtering* (2nd ed.). New York: John Wiley & Sons, Inc.
- Burdea, G., & Coiffet, P. (1994). *Virtual Reality Technology* (1st ed.). New York: John Wiley & Sons, Inc.
- Burton, R. P. (1973). *Real-time measurement of multiple three-dimensional positions*. Unpublished doctoral dissertation, University of Utah, Salt Lake City.
- Burton, R. P., & Sutherland, I. E. (1974). Twinkle Box: Three-dimensional computer-input devices. *AFIPS Conference Proceedings, 1974 National Computer Conference*, 43, 513–520.
- Chi, V. L. (1995). *Noise model and performance analysis of outward-looking optical trackers using lateral effect photo diodes* (Tech. Rep. TR95-012). Chapel Hill: University of North Carolina at Chapel Hill.
- Emura, S., & Tachi, S. (1994). *Sensor fusion based measurement of human head motion*. Paper presented at the 3rd IEEE International Workshop on Robot and Human Communication (RO-MAN 94 NAGOYA), Nagoya University, Nagoya, Japan.
- Feiner, S., MacIntyre, B., Höllerer, T., & Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring urban environments. *Personal Technologies*, 1(4), 208–217.
- Focus Software. (1995). *ZEMAX Optical Design Program User's Guide, Version 4.5*. Tucson, AZ: Focus Software.
- Foxlin, E., Harrington, M., & Pfeifer, G. (1998). Constellation™: A wide-range wireless motion-tracking system for augmented reality and virtual set applications. In M. F. Cohen (Ed.), *Computer Graphics (SIGGRAPH 98 Conference Proceedings ed.)* (pp. 371–378). Orlando: ACM Press, Addison-Wesley.
- Fuchs (Foxlin), E. (1993). *Inertial head-tracking (manual)*. Unpublished master's thesis, Massachusetts Institute of Technology.
- Gelb, A. (1974). *Applied Optimal Estimation*. Cambridge, MA: MIT Press.
- Gottschalk, S., & Hughes, J. F. (1993). Autocalibration for virtual environments tracking hardware. *Computer Graphics (SIGGRAPH 93 Conference Proceedings ed.)* (pp. 65–72). Anaheim, CA: ACM Press. ACM Press, Addison Wesley.
- Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., & Hallaway, D. (1999). Exploring MARS: Developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers & Graphics*, 23(6), 779–785.
- IGT. (2000). *Image Guided Technologies [HTML]*. Retrieved September 15, 2000 from the World Wide Web: <http://www.imageguided.com/>.
- Intersense. (2000). *Intersense IS-900 [HTML]*. Retrieved April 27, 2000 from the World Wide Web: <http://www.isense.com/>.
- Jacobs, O. L. R. (1993). *Introduction to Control Theory* (2nd ed.). New York: Oxford University Press.
- Kadaba, M. P., & Stine, R. (2000). *Real-Time Movement Analysis Techniques and Concepts for the New Millennium in Sports Medicine [HTML]*. Motion Analysis Corporation, Santa Rosa, CA. Retrieved September 15, 2000 from the World Wide Web: <http://www.motionanalysis.com/applications/movement/rtanalysis.html>.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transaction of the ASME—Journal of Basic Engineering*, 82(series D), 35–45.
- Lewis, F. L. (1986). *Optimal Estimation with an Introductory to Stochastic Control Theory*. New York: John Wiley & Sons, Inc.

- MAC. (2000). *HiRes 3D Motion Capture System* [HTML]. Motion Analysis Corporation. Retrieved September 15, 2000 from the World Wide Web: <http://www.motionanalysis.com/applications/movement/gait/3d.html>.
- Maybeck, P. S. (1979). *Stochastic models, estimation, and control* (Vol. 141). New York: Academic Press.
- Mazuryk, T., & Gervautz, M. (1995). Two-step prediction and image deflection for exact head tracking in virtual environments. *Proceedings of EUROGRAPHICS 95*, 14(3), 30–41.
- Meyer, K., Applewhite, H., & Biocca, F. (1992). A survey of position trackers. *Presence, a publication of the Center for Research in Journalism and Mass Communication*, 1(2), 173–200.
- Mulder, A. (1994a). *Human movement tracking technology* (Tech. Rep. TR 94-1). School of Kinesiology, Simon Fraser University.
- . (1994b, May 8, 1998). *Human Movement Tracking Technology: Resources* [HTML]. School of Kinesiology, Simon Fraser University. Retrieved September 15, 2000 from the World Wide Web: <http://www.cs.sfu.ca/people/ResearchStaff/amulder/personal/vmi/HMTT.add.html>.
- . (1998, May 8, 1998). *Human Movement Tracking Technology* [HTML]. School of Kinesiology, Simon Fraser University. Retrieved September 15, 2000 from the World Wide Web: <http://www.cs.sfu.ca/people/ResearchStaff/amulder/personal/vmi/HMTT.pub.html>.
- Polhemus. (2000). *Polhemus* [HTML]. Retrieved September 15, 2000 from the World Wide Web: <http://www.polhemus.com/home.htm>.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1990). *Numerical Recipes in C: The Art of Scientific Computing* (2nd ed.). Cambridge University Press.
- Sorenson, H. W. (1970). Least-squares estimation: From Gauss to Kalman. *IEEE Spectrum*, 7(July), 63–68.
- Spohrer, J. (1999a). Information in places. *IBM Systems Journal, Pervasive Computing*, 38(4).
- Spohrer, J. (1999b, June 16). *WorldBoard: What Comes After the WWW?* [HTML]. Learning Communities Group, ATG, Apple Computer, Inc. Retrieved December 24, 1999 from the World Wide Web: <http://worldboard.org/pub/spohrer/wbconcept/default.html>.
- Sutherland, I. E. (1968). A head-mounted three dimensional display. *Proceedings of the 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings* (vol. 33, part 1, pp. 757–764). Washington, D.C.: Thompson Books.
- Thomas, S. W. (1984, December). *The Alpha\_1 computer-aided geometric design system in the Unix environment*. Paper presented at the Computer Graphics and Unix Workshop.
- UNC Tracker Project. (2000, July 10). *Wide-Area Tracking: Navigation Technology for Head-Mounted Displays* [HTML]. Retrieved July 18, 2000 from the World Wide Web: <http://www.cs.unc.edu/~tracker>.
- University of Utah Computer Science. (1999). *Alpha 1 Publications* [HTML]. University of Utah, Department of Computer Science. Retrieved May 28, 1999 from the World Wide Web: [http://www.cs.utah.edu/projects/alpha1/a1\\_publications.html](http://www.cs.utah.edu/projects/alpha1/a1_publications.html).
- Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P., Jr. (1999). Walking > Walking-in-Place > Flying, in Virtual Environments. In A. Rockwood (Ed.), *Computer Graphics (SIGGRAPH 99 Conference Proceedings ed.)* (pp. 359–364). Los Angeles: ACM Press, Addison Wesley.
- Van Pabst, J. V. L., & Krekel, P. F. C. (1993, September 20–22). *Multi sensor data fusion of points, line segments and surface segments in 3D space*. Paper presented at the 7th International Conference on Image Analysis and Processing, Capotolo, Monopoli, Italy.
- Wang, J.-F. (1990). *A real-time optical 6D tracker for head-mounted display systems*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill, Chapel Hill.
- Wang, J.-F., Azuma, R. T., Bishop, G., Chi, V., Eyles, J., & Fuchs, H. (1990, April 16–20). *Tracking a head-mounted display in a room-sized environment with head-mounted cameras*. Paper presented at the SPIE 1990 Technical Symposium on Optical Engineering and Photonics in Aerospace Sensing, Orlando, FL.
- Wang, J.-F., Chi, V., & Fuchs, H. (1990, March 25–28). *A real-time optical 3D Tracker for head-mounted display systems*. Paper presented at the Symposium on Interactive 3D Graphics, Snowbird, UT.
- Ward, M., Azuma, R. T., Bennett, R., Gottschalk, S., & Fuchs, H. (1992, March 29–April 1). *A demonstrated optical tracker with scalable work area for head-mounted display systems*. Paper presented at the Symposium on Interactive 3D Graphics, Cambridge, MA.
- Welch, G. (1995). *Hybrid self-tracker: An inertial/optical hybrid three-dimensional tracking system* (Tech. Rep. TR95-048). Chapel Hill: University of North Carolina at Chapel Hill, Department of Computer Science.
- . (1996). *SCAAT: Incremental tracking with incomplete information*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill, Chapel Hill.

- Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter* (Tech. Rep. TR95-041). Chapel Hill: University of North Carolina at Chapel Hill, Department of Computer Science.
- . (1997). SCAAT: Incremental tracking with incomplete information. In T. Whitted (Ed.), *Computer Graphics (SIGGRAPH 97 Conference Proceedings ed.)* (pp. 333–344). Los Angeles: ACM Press, Addison-Wesley.
- . (2000, January 23, 2000). *The Kalman Filter* [HTML]. University of North Carolina at Chapel Hill. Retrieved April 29, 2000 from the World Wide Web: <http://www.cs.unc.edu/~welch/kalman/index.html>.
- Welch, G., Bishop, G., Vicci, L., Brumback, S., Keller, K., & Colucci, D. N. (1999). The HiBall tracker: High-performance wide-area tracking for virtual and augmented environments. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (pp. 1–11). University College London, London, United Kingdom (December 20–23): ACM SIGGRAPH, Addison-Wesley.
- Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio Electroacoustics*, *AU(15)*, 70–73.
- Woltring, H. J. (1974). New possibilities for human motion studies by real-time light spot position measurement. *Biometry*, *1*, 132–146.