

# Calibration For Augmented Reality Experimental Testbeds

Valerie A. Summers\* & Kellogg S. Booth  
Dept. of Computer Science  
University of British Columbia

Tom Calvert  
Dept. of Computer Science  
TechBC

Evan Graham  
DSI Datotech  
Systems Inc.

Christine L. MacKenzie  
School of Kinesiology  
Simon Fraser University

## 1 Abstract

The Virtual Hand Lab (VHL) is an augmented reality environment for conducting experiments in human perception and motor performance that involve grasping, manipulation, and other 3D tasks that people perform with their hands. The hardware and software testbed supports both physical and virtual objects, and object behaviors that can be specified in advance by experimenters. A testbed for conducting experiments must provide visual stimuli that depend on the configuration of the experimental apparatus, on the specific tasks that are being studied, and on the individual characteristics of each subject. Calibration is an important concern and is the subject of this paper. A proper design leads to independent calibration steps that modularize the subsystems that require calibration and explicitly recognize and order the dependencies among them. We describe how the architecture for the VHL was designed to support independent apparatus-specific, experiment-specific, and subject-specific calibrations. The architecture offers benefits for any augmented reality environment by reducing re-calibration times and identifying appropriate modularization in the software that can result in a more robust and efficient implementation.

**CR Categories:** I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—virtual reality

**Keywords:** augmented reality, experimental systems, registration, calibration

## 2 Introduction

Augmented reality environments dynamically “augment” the user’s view of the real 3D world with computer-generated virtual objects. These virtual objects may overlay physical objects, creating hybrid objects, or the virtual objects may exist independently of the physical objects but still interact with them through behaviors that can differ from those of physical objects. Several domains are seen as effective applications of augmented reality: medicine and surgery [4, 15, 21], manufacturing and repair [8, 22], and complicated visualization domains that support hand-based interaction styles [6]. We built the Virtual Hand Laboratory (VHL), a desktop augmented reality environment, to learn more about how people interact with their

\*Contact author’s address: Department of Computer Science, University of British Columbia, 2236 Main Mall, Vancouver, British Columbia, Canada V6T 1Z4, summers@cs.ubc.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

1999 Symposium on Interactive 3D Graphics Atlanta GA USA  
Copyright ACM 1999 1-58113-082-1/99/04...\$5.00

environments. Our interest is in providing a testbed in which experiments can be conducted to examine human perception and motor performance when complex tasks are executed using the hands.

Kinesiologists and psychologists study how people behave in a variety of physical environments. The advantage of an augmented environment is that experiments can explore the limits of human perception and motor performance by simulating situations that would be difficult (even impossible) to create using the purely physical apparatus traditionally employed in studies of this type. In augmented reality (AR), accurate calibration assures visual registration between the virtual and physical objects that are combined to create “augmented” objects. It also ensures that purely virtual objects appear at the correct locations and orientations, and at the correct size.

It is clear that gross discrepancies in calibration cause human performance errors, but what about small discrepancies? To test the quantitative effects of calibration errors on the performance and perception of virtual environments users, it is necessary to systematically introduce errors into the calibration.

Our work is intended to address these issues by building a system in which experiments of this type can be conducted using an augmented environment whose calibration procedures naturally support an understanding of the errors inherent in them. It will benefit designers of augmented reality systems to be able to predict task error based on varying levels of calibration accuracy. A second benefit of our work is increased knowledge of basic human interaction within an augmented environment. Kinesiologists are using the VHL to perform extensive analyses on 3D kinematic data, and from that make inferences about human movement planning and motor control.

In the following sections we first discuss the requirements of calibration procedures for an augmented reality experimental testbed, provide a description of the experimental apparatus comprising the VHL, and then give a brief summary of the many coordinate systems that are used by the hardware and software. With this as background, the calibration procedure we have adopted are explained and analyzed in terms of the criteria we have established. The calibration procedure has three parts, calibration of the workspace, calibration of a subject’s point of view, and calibration of physical objects that will be employed in experiments. We conclude with an overall discussion of our system and of related work reported in the literature.

## 3 Requirements for Calibration

Calibration requirements for augmented environments are similar to those for robotics – in both cases computer hardware and software are used to coordinate objects in 3D where some of the objects are under computer control and some are not. In the case of augmented environments, some of the objects may be only virtual, but this does not change the fundamental requirement to accurately measure the current state of all of the objects and compute the next state for those that are under computer control. For this reason, we base our calibration requirements on those given by Hollerbach and Wampler [10], who say that in robotics “Ideally, the calibration

methods should be statistically robust, there should be a variety of approaches for different circumstances, and metrology equipment should be sufficiently accurate, convenient to use, and not too expensive”.

To these we add three more requirements that are specific to experimental testbeds in which experiments involving human subjects will be conducted:

**independent** - separate parts of the calibration should not rely on each other unless absolutely necessary;

**subject-specific** - the calibration should account for individual differences

**no cues** - no residual cues should be left that might be exploited by subjects in unanticipated ways that would introduce confounds into experiments.

The independence criterion derives from the need to measure the effects of calibration errors on performance by systematically introducing perturbations. If the calibrations are dependent on one another, then it is not possible to isolate the effects of a single perturbation. A benefit of independent calibration (for any augmented reality system, not just experimental testbeds) is that it decreases the total time spent on calibration, because changing one item in the environment does not produce a cascade of re-calibrations. Depending on the implementation, it may decrease errors by reducing error propagation.

Since the experiments examine differences between subjects, calibrations must reflect the individual being tested. Perturbations must be relative to the correct calibration for that subject.

The third experimental criterion, elimination of residual cues, arises from the goal of increasing internal validity for an experiment. Experiments have greater internal validity if the effects measured for dependent variables can be attributed to the independent variables under experimental control, and not to other sources. Some calibration procedures require fiducial marks which remain in the environment during its use. To increase validity, one must ensure that these marks cannot be used by subjects in ways unanticipated by the experimenter. Our approach is to eliminate fiducial marks in the environment wherever possible.

Our goal is to identify the requirements of an experimental system, merge them with traditional calibration requirements, and then provide procedures which satisfy both. A sub-goal is to identify methods for evaluating the accuracy of a calibration procedure and show how the evaluation procedure isolates the calibration from other calibrations.

## 4 Apparatus

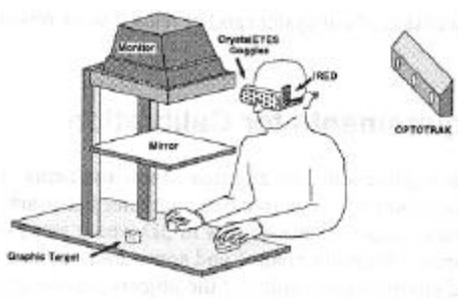


Figure 1: Physical configuration

The hardware supporting our augmented environment consists of an SGI Indigo 2 Extreme workstation a slaved secondary SGI mon-

itor (*monitor*) that is reflected through a mirror, a Northern Digital Optotrak 3D motion analysis system (*trucker*), StereoGraphics CrystalEyes stereo-graphic glasses (*stereo glasses*), and an ATI Force/Torque Sensor System force sensing device (*force sensor*). The SGI displays stereo images on its regular screen (which is seen by the experimenter) and also on the slaved monitor. The mirror places these images appropriately within the subject's view. The Optotrak senses the 3-D positions of infrared emitting diodes (*markers*) that are strobed under computer control. The markers are placed on all objects (including the subject) whose position or orientation is required during an experiment. Three or more markers rigidly mounted provide orientation information that can be computed from their relative positions. The stereo glasses, used in conjunction with position and orientation information for a subject's head obtained (from the Optotrak), enable the monitor to display head-coupled stereo images for the subject.

Figure 1 illustrates the physical set up. A metal rack on rollers supports the slaved monitor face down over the half-silvered mirror. The rack is placed so that the monitor and mirror are above a desktop, all parallel to each other, with approximately equal spacing between the face of the monitor on top, the mirror in the middle, and the desktop on the bottom. With this geometry, an image on the monitor is reflected through the mirror and a 2D virtual image appears just above and parallel to the desktop (the relative spacing of the monitor, mirror and desktop determine the exact distance).

A set up very similar to this was used by Knowlton [12] in his system for prototyping keyboard layouts for telephone operators. Subjects look through the half-silvered mirror and can reach underneath it, resulting in the virtual image reflected through the mirror appearing to be on top of the actual scene beneath the mirror. The subject's hands are visible beneath the mirror, which Knowlton describes as "eerie". He used this apparatus to change the virtual labels on the keys on a standard telephone operator's keyboard to conduct usability studies that compared various keyboard layouts and tested new functionality. His system used only a 2D virtual image because all of the labels for the keys are (approximately) in a common plane just above the keyboard. In our work, we want subjects to manipulate 3D objects within the entire workspace.

Using a mirror as a "beam-splitter" between the subject and the workspace allows computer augmentation of objects to supersede the visible attributes of the physical object. Turning a light (not shown) on or off under the mirror determines whether the display is see-through or purely virtual. If the light is turned off, none of the physical objects are visible, but the virtual ones are unaffected; with the light off, both physical and virtual objects are seen, just as in Knowlton's set up.

The workspace is the volume between the mirror and the desktop into which a subject can physically reach. To achieve the illusion of 3D images anywhere workspace, stereo glasses are used. Different left and right images are sequentially displayed on the monitor providing stereo separation. This is the typical field-sequential presentation common in desktop or fish tank VR, except that the subject is looking into the mirror and thus sees the images as if they monitor were embedded in the desktop, but without the physical inconvenience of the monitor protruding below the desk (which would limit foot room) and without the subject's hands obscuring the virtual objects. A flat-panel display laid on the desktop would achieve the first goal (foot room) but not the second, because images would be obscured any time a subject's hand (or any other object) blocked the line of sight to part of the image on the flat-panel display. The presence of a mirror provides some constraints on the workspace, but the ability to slide the mirror closer or farther away, and to move the entire apparatus left, right, closer or farther from the subject provides enough flexibility for our purposes.

The relative positions of the desk, rack with monitor and mirror, tracking equipment and subject vary according to the experiment.

The rack holding the monitor is built with supports on only one side, which cannot face the tracker, but the other side provides an unobstructed view of the workspace, allowing the markers to be sensed by the tracker. The monitor itself is awkward and heavy to move, so rearranging it in the rack requires more than one person.

The orientation of the workspace coordinate system is chosen relative to the side of the desk on which the subject will be seated rather than being associated with the monitor. Defining the coordinate system this way allows experimenters to place subjects on either side of the desktop, optimizing the use of the limited workspace without changing the orientation of the monitor in the rack, thus allowing for individual differences such as left and right handedness. The ability to slide the mirror forward and back makes it easy to adjust for differences in subject height and seating posture so that the virtual image is always visible to the subject through the mirror.

The SGI is programmed to update the display at 60 Hz in stereo mode. The glasses are shuttered at 120 Hz, giving each eye 60 images per second, a total of 120 images each second. Tracking is performed at 60 Hz by the Optotrak. Force sensing is conducted at 200 Hz or higher, but at the present time it is not used to control the system, only for data acquisition.

## 5 Coordinate Systems

There are a number of coordinate systems that are employed in the system. The software automatically converts between these as it displays the virtual scene based on data from the tracker. Some of the transformations are static, depending only on the geometry of the apparatus, while others are dynamic, depending on the current position and orientation of the subject's head or hands and of any physical objects that are in use.

From the experimenter's viewpoint, the most fundamental coordinate system is a *subject-specific* workspace coordinate system. This is a right-handed coordinate system in which X is to the subject's right, Y is pointing away from and in front of the subject, and Z is up. The coordinate system does not fluctuate based on a subject's movements while seated. It is aligned with the desktop, but it is oriented according to where the subject is sitting, which means that the relationship between the monitor's pixel-based coordinate system and the workspace coordinate system will be mirror imaged in some cases.

We define the workspace so that the XY-plane is aligned with the desktop (adjusting the height of the desk if necessary). Virtual objects entirely within the XY-plane thus have identical left and right stereo images because there is no binocular disparity in the XY plane. If the monitor, mirror and desktop were not parallel, the plane of the virtual image would be skewed relative to the mirror and not necessarily parallel to the desktop. This would not change the fundamental relationships, but it would make calibration more complicated because our procedure takes advantage of the geometry.

Many transformations are used during calibration of individual system components. Some of these calibrations can be "collapsed" to reduce the number of transformations needed during operation of the VHL. Table 1 shows the key transformations in the style of Tuceryan et al. [22]. This table is presented here as a high level overview. The descriptions will become more relevant in the sections describing the component calibrations. Figure 2 shows how these transformations interact.

## 6 Calibration Procedure

Three distinct types of calibrations are applied to the VHL. The workspace (volume of physical space which can be augmented), the point of view (for a head-tracked stereo display), and the physical

Trans.	Description	Nature	Calculation Process
W1	tracker to workspace	fixed	workspace calibration
H1	tracker to head	varying	tracker
H2	head to right eye	fixed	eye calibration
H3	head to left eye	fixed	eye calibration
P4	tracker to markers	varying	tracker
P21	markers to object	fixed	object calibration
P22	tracker to object	fixed	object calibration
P23	model transformation	varying	from P(4,21,22)

Table 1: Key transformations used in the VHL. Many more transformations are used during calibration (thus the skipped numbers).

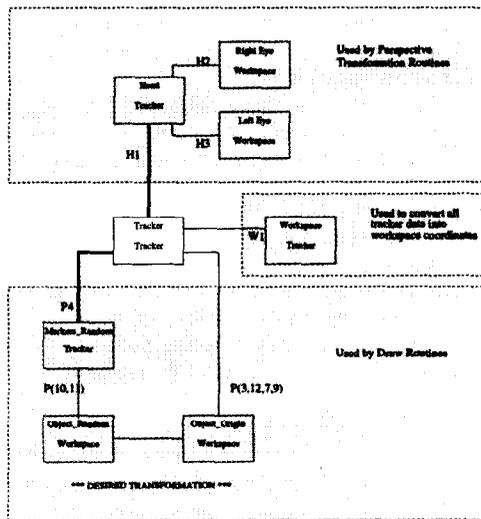


Figure 2: Key transformations in components of the VHL. Dark lines indicate transformations that vary during system use. Light lines indicate fixed or static transformations. The top label in each box indicates the physical item being measured. The bottom label indicates the coordinate system in which it is measured.

objects (to create augmented objects) must each be calibrated so that transformations between coordinate systems can be accomplished.

For each component, we describe the procedure an experimenter follows, how the technique satisfies the experimental calibration requirements, an evaluation of the technical accuracy of the technique, and a discussion of the effects a component calibration error has on the overall system. The tracking system is a commercial product, self-calibrated to 0.3 mm accuracy and is not evaluated here.

### 6.1 Workspace Calibration

The calibration procedure consists of aligning a set of virtual crosses with a set of markers according to a pre-defined one-to-one correspondence. The markers have identity because they are strobed, and the crosses are described relative to the workspace orientation chosen by the experimenter. For example, the first marker is placed in the "left front corner", where "left" and "front" are arbitrarily chosen by the experimenter.

For a given virtual cross, there is exactly one location in 3-space in which a marker can be placed so that the alignment of the virtual cross and the marker are not a function of head position. In all other positions, moving one's head will cause the virtual cross to "swim" with respect to the physical marker. We currently rely on human perception to determine proper alignment, but only to detect 2D misalignment, not differences in depth. Alignment in the third dimen-

sion is determined by the absence of swim, which is something that humans are very good at detecting. Because 2D alignment of the markers with the virtual crosses and the check for lack of swim do not vary with the head position, this calibration can be performed without stereo, i.e. independent of the point of view calibration.

Using standard techniques, a coordinate system is obtained from a set of points. We arbitrarily chose four points. If we used more markers, we could use the additional redundant data to reduce the variance further. The tradeoff would be increased time to perform the calibration.

The markers are removed after calibration and before the arrival of subjects, hence subjects cannot use them to gain any sort of information about the workspace. Unlike the experimenter, the subject is unaware of the coordinate system used. Specifically, the subjects cannot use the markers to facilitate alignment tasks, determine the center of the workspace, determine the boundaries of the workspace volume or anticipate target locations.

To evaluate the consistency of the calibration, the workspace was calibrated three times. For each calibration, data was collected four times without moving any markers, for a total of 12 readings. Each marker produces an (x,y,z) triplet, in tracker coordinates. Comparing readings within the same calibration allows us to evaluate the tracker error. The maximum range (maximum value minus minimum value for a given marker/axis combination) was 0.24 mm. We then evaluated the maximum range over all readings and calibrations. The maximum range was 1.49 mm. If we attribute 0.24 mm of that error to tracker error, then by subtraction, the calibration procedure contributes 1.25 mm to the total system error.

Marker Location	Range (mm)			Stdev		
	X	Y	Z	X	Y	Z
left front	0.62	0.45	0.85	0.29	0.19	0.35
left back	1.49	1.02	0.92	0.69	0.45	0.39
right front	1.37	0.51	0.34	0.58	0.24	0.10
center	0.12	1.41	1.34	0.04	0.68	0.52

Table 2: Range of values acquired during workspace calibration for each position.

Errors in workspace calibration affect the placement of virtual objects, but not the registration of augmented objects. A trial specification file specifies the time and place at which virtual objects are to appear. The three markers (left front, left back and right front) used to determine the workspace coordinate system define its orientation relative to the physical equipment, but not its scale. Errors in the placement of the center marker will move the origin of the coordinate system. All virtual objects will be affected in the same manner. In particular, virtual objects will retain the chosen relative distance and orientation despite any errors in workspace calibration. This is quite important, as many experiments are based on movements between specified locations.

Augmented objects draw virtual objects relative to their associated physical object. An error in the workspace calculation will cause the markers on the physical object to indicate an erroneous location in workspace coordinates. However, the virtual object will be drawn in that "erroneous" location. Hence, despite the error in the workspace calibration, the physical and virtual objects will still be aligned.

## 6.2 Point of View

Human performance is strongly influenced by visual input which, in a head-coupled, stereo, augmented environment depends on spatial and temporal alignment of the left- and right-eye virtual images with each other and with the real environment. Instead of tracking the eyes directly, we estimate the eye position relative to the glasses,

then track the glasses while the subject interacts with the environment.

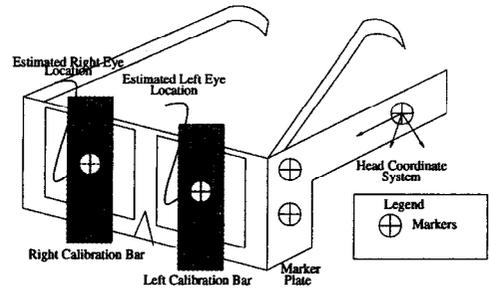


Figure 3: Point of View Calibration. Subjects align the calibration bars so that through each hole they see the same image. The relative position of the holes to the glasses is then calculated.

We start by outlining the calibration process.

1. The experimenter attaches light-weight rigid plastic plates to the sides of the stereo glasses in such a manner that they do not interfere with the flexible arms then places three markers on the plate facing the tracker in any non-collinear positions. These markers are used to track the head position and form a head coordinate system.

2. The subject dons the stereo glasses, which are secured against subsequent slippage with a strap of the type used for sports.

3. The subject closes his/her left eye. A bar with a small hole is provided. The bar is aligned over the subject's right eye so that a point in the center of the workspace is clearly visible. The bar attaches to the glasses with velcro.

4. The subject repeats this for the other eye.

5. The subject then looks through both holes simultaneously. The holes should align so that he/she sees the same fused image through both holes. It should appear as though there is only a single hole, not two separate or overlapping holes. If necessary, the subject rearranges the bars until this is the case.

6. The experimenter places markers directly over the holes.

7. Two seconds of data (120 frames) are captured, processed and analyzed. A summary of the analysis is produced. The experimenter will rerun the data collection if too many of the data frames are erroneous. This is typically solved by changing the subject's head position.

8. The bars and markers are removed from the subject's eyes. The markers which track head position remain.

Using the collected data, rigid body transformations between the position of the eye markers and the head coordinate system are calculated and used to estimate the eye positions after the calibration markers are removed from the eyes. Our estimates are actually 1-2 cm in front of the points we really want, but this results only in a slight magnification of the scene and avoids a number of pitfalls related to gaze direction.

A series of filters processes the data and determines the straight line distances between every pair of markers. An evaluation of the data quality is produced for the experimenters. During use of the VHL, distances are calculated between each of the markers on the head for every frame of data. These new distances are compared to the calibration metrics to identify frames where the data may be erroneous.

A major benefit of this approach is that there is no dependence between the right and left eye calibrations. Inter-ocular distance is correctly computed for each subject. It is not assumed that this distance is evenly divided by the nose piece. The distances from the nose to each eye will be similar for many people, but due to the asymmetry of the human body, not exact. For example, astigmatism and pre-

viously broken noses increase the discrepancy. The vertical placement of the point of view is computed. It need not be the center of the glasses. We do not assume the glasses sit levelly on the head. The position of ears on a subject's head will often cause the glasses to be skewed relative to the head.

The above calibration facilitates future experiments that will independently manipulate these parameters to determine to what extent they affect performance.

This procedure does not give the exact point of view. The markers used for estimating the eyes are placed on the glasses, not in the eyeballs. We also do not adjust the point of view based on distance to the point of interest. To evaluate the error, an optician's pupillometer [7] was used to measure interpupillary distance (IPD) for three subjects for each of seven focal distances: 35, 40, 50, 65, 100, 200 and  $\infty$  (cm). For the first subject, multiple readings with two experimenters were obtained. Over five readings, there was at most a 1 mm difference for a given focal distance. With this degree of accuracy, readings were taken only once for the other two subjects. As the focal length increased, the IPD monotonically increased, as expected. The difference in distance from the center of the nose piece to each eye was within 3 mm for each subject. Each subject was calibrated using the above procedure twice. The calibration IPDs all fell within range of the pupillometer readings as shown in table 3.

Technique	Subject 1		Subject 2		Subject 3	
	min	max	min	max	min	max
calibration	57	62	57	61	63	64
pupillometer	57	62	57	61	60	65

Table 3: An optician's pupillometer was used to record interpupillary distances over focal lengths ranging from 35 cm to infinity. The calibration technique produced values within the pupillometer range.

Not everyone has access to a pupillometer. The following test may be performed by anyone using an augmented environment. To quantify the effect of the point of view (POV) calibration, wire-frame blocks of various sizes were displayed at various locations using a randomized trial script. The width (Y-axis) of the virtual block was measured with a physical ruler. The head moved freely to obtain the best perspective. Accuracy was stressed over speed. Over 15 trials, the mean error was 0.47 mm with a maximum error of 1 mm. This experiment was repeated for the other dimensions. For height (Z) the mean error was 1.73 and the maximum error was 4 mm, and for depth (X) the mean error was 1.33 mm with a maximum error of 3 mm.

We speculate that the depth and height dimensions are more awkward to measure, and that the vertical dimension is more affected by screen curvature. An anonymous reviewer suggested that when the ruler is held horizontally, visual and/or tactile cues to its orientation are obtained from the desktop; in contrast, it is harder to see or feel if the ruler is vertical, and any tilt would contribute to measurement error in the vertical case.

These tests are independent of the workspace calibration. Errors in workspace calibration have only minor effects on the position or orientation of the virtual blocks, not on their size. The position and orientation of the targets did not have a significant effect on the errors. These tests are dependent on tracker calibration since this would affect any lag in head tracking.

### 6.3 Physical Objects

Physical objects provide tactile, kinesthetic and haptic feedback when they are manipulated [16], which is why we want to be able to make use of them in our experiments. By calibrating an object, we can track its motion. This information can be used to augment its visual appearance, and subsequently analyze its motion.

Marker Location	Range (mm)			Stdev		
	X	Y	Z	X	Y	Z
22	0.75	0.25	1.75	1.50	0.50	0.50
25	1.56	0.56	1.89	1.33	0.53	1.36
44	1.50	0.50	1.00	0.71	0.71	1.41
Total	1.33	0.47	1.73	1.29	0.52	1.16
Max Err.	3	1	4			

Table 4: Effects of eye calibration on perception of size. Bars of a virtual wire-frame block were measured.

Calibrating an object involves matching a computer model to the physical object. Both the model and the object are assumed to have distinguishable dimensions. This calibration technique is only dependent on the experimenter's ability to manipulate physical objects, not on an ability to manipulate physical and virtual objects together. This technique works for any non-deformable object.

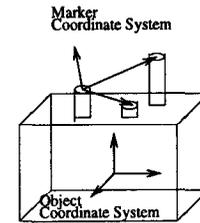


Figure 4: There is no assumed correspondence between object and marker coordinate systems.

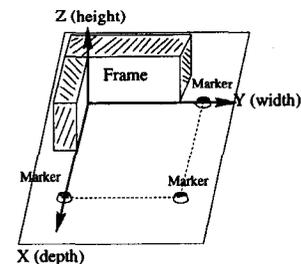


Figure 5: During calibration physical objects are placed in the corner of the calibration frame.

The procedure for this calibration is simple:

1. Attach three markers to the object in any non-collinear positions (See Figure 4). They may be placed on the side, top, placed flat on the object, raised using an adhesive or some combination thereof.
2. Place three markers in predefined locations of the calibration frame (Figure 5). The frame may be placed anywhere within view of the tracker. Its placement is completely unrelated to the location of the workspace except that it must be within range of the tracker.
3. Place the object in the corner of the calibration frame so that the XYZ orientations of the object and frame match.
4. Collect marker locations for both frame and object in tracker coordinates.

The position of the object markers relative to the frame markers plus knowledge of the size and shape of the physical object allow us to compute the internal coordinate system (CS) of the physical object relative to the object markers and thus we can overlay a corresponding virtual object if desired.

The markers are left on the object, but if care is taken in their placement, they have no specific orientation so they provide minimal cues to the subject. Specifically, they cannot be used to help

align the physical object with virtual objects because the markers are not aligned with any axis of either physical or virtual objects, nor do they indicate specific locations such as the center of the object.

Isolating object calibration from point-of-view calibration allows us to test the effect of calibration errors. For experiments involving perturbations of the virtual object with respect to the physical object, the object is not placed in the corner of the calibration frame, but rather is offset accordingly. Calibration in this case takes approximately 10 minutes, because the experimenters typically verify the selected rotation/translation by viewing it in the workspace. These types of experiments would not be possible with techniques that require someone to touch a sensor to specific points in the environment during calibration [2, 9].

Sources of error for this procedure are tracker error, incorrect measurement of the physical object and incorrect position of the object in the calibration frame. Tracker error is within 0.3 mm, physical measurements with a ruler are accurate within 1 mm, and position errors are negligible because both the objects and the frame are rigid and they fit tightly together. Recalibrating for new marker locations takes 30-60 seconds.

## 6.4 Transformation Between Coordinate Systems

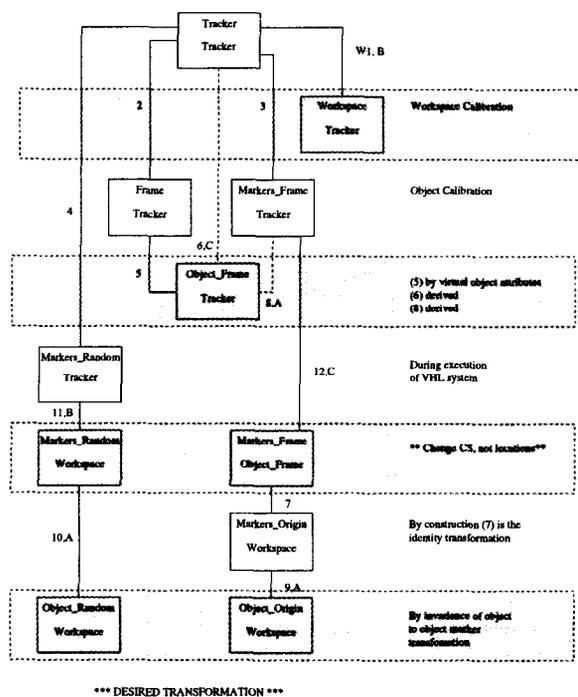


Figure 6: Transformations for object calibration. The top label in each box indicates the physical item being measured. The bottom label indicates the coordinate system in which it is measured. We wish to determine the transformation from the origin of the workspace to the object's new, random location.

In the following description, numbers identify the transformations in Figure 6. Occasionally the same transformation has been used in multiple places. In these cases, supplementary letter labels have been provided. Transformation (W1) is obtained from the workspace calibration. Transformations (2) and (3) are easily obtained from the sets of markers on the calibration frame and physical object while it is sitting in the frame.

Transformation (5) is based on the virtual object's encoded attributes of height, width and depth, and an adjustment which accounts for the height of the markers above the calibration frame.

With any such graph, there is always a transformation between any two nodes on the graph tree. Two identified by dashed lines will be used later. Transformation (6) with supplementary label (C) is derived from (2) and (5). Transformation (8) with supplementary label (A) is derived from (3) and (6). Transformation (8) is the transformation between the object CS and the marker CS; it is invariant no matter where the object is physically located.

Transformation 4 is computed for each data frame during system usage, not during the calibration phase.

Transformations can be used to change between different locations both measured in the same CS (as above), or to change the CS in which a point is measured. Transformations (11) and (12) do the latter by applying transformations (1) and (6) respectively.

Transformation (7) is the key transformation. We change both physical locations and the system in which it is measured. The CSs were defined so that the transformation from the workspace coordinate system to the coordinate system of the object while located at the origin of the workspace is the identity transformation, indicated by the supplementary label "I".

More specifically, the position of the markers while the object was sitting in the calibration frame, as measured in object\_frame coordinates, is exactly the position that the markers would be located, had the object been at the workspace origin, and the markers measured in workspace coordinates. (Note we are switching coordinate systems and locations at the same time). So, now we know where the markers would have been, had the object been centered in the workspace, as measured in workspace coordinates.

At this point, all of the off-line calculations have been done, and we need to make use of the real-time data to determine the "desired transformation" of the object from the origin position to the transformed position. It can be easily derived from the above transformations.

We collapse transformations to reduce overhead during use of the VHL. Key transformations are denoted as "Pn" and are used in Figure 2. We define transformation (P20) as the tracker to object origin in workspace coordinates, composed of transformations (3, 12, 7, and 9). (P21) is composed of (10,11) and the desired transformation is P23. (P4) is transformation (4).

## 7 Discussion and Related Work

We evaluate our complete set of calibration procedures based on Hollerbach and Wampler's criteria and then the additional experimental testbed criteria.

**robustness** – An experiment consists of a session for each subject in which a large number of trials are performed. Five full experiments and several pilots have been performed. Objects are calibrated for each experiment. One object was re-calibrated once during a session, presumably because a marker was moved. The workspace was calibrated for each experiment, but not between sessions except by one experimenter to ensure against the rack moving. Point of view calibrations were performed once for each subject.

**flexibility** – Depending on the experiment, the experimenter must place markers on the objects in different locations. For example, markers might be placed on the side, top, or some combination thereof. Subjects may be positioned anywhere around the workspace without affecting calibrations. The VHL physical setup can be easily and quickly rolled out of the way for other experiments and it is not necessary for the workspace to be precisely placed when a new VHL experiment takes place.

**accuracy** – Total static system accuracy is within 4 millimeters. Specifically: X (left/right)  $\leq 3$  mm, Y (away)  $\leq 1$  mm and Z (up)  $\leq 4$  mm. The tracker provides accuracy within 0.3 millimeters. Workspace error is  $\leq 1.3$  mm.

**convenience of use** – Once an experiment has been set-up, the kinesniologists who use the VHL successfully calibrate the entire sys-

tem independent of outside computer technical support. There have been two published papers [23, 24] and four more experiments are in progress. Calibrating the entire system takes between 10 and 30 minutes, depending on the experimenter.

**expense** – The VHL equipment (tracker, computers, rack) cost over \$100K. However the additional cost for the calibration materials is quite small. The frame used for object calibration is constructed of scrap wood and adhesives and the eye calibration accessories are created from readily available plastic and adhesives. Together these cost less than twenty dollars.

**independence** – All of the calibrations rely on the tracker being accurate, but this is hard to avoid. Changing subjects requires only point of view re-calibration. Moving either the tracker or the workspace requires only workspace re-calibration. Calibrating physical objects is independent of the other two calibrations.

**subject-specific calibration** – The workspace is defined according to the side of the desk where the subject is seated. The point of view calibration is individualized for each subject and accounts for the placement of the eyes relative to the stereo glasses for each subject, including such metrics as inter-ocular distance.

**no cues** – The markers used during calibration of the workspace and point of view are removed prior to experimentation, and thus provide no information which a subject could use during an experiment. The markers on the physical objects remain visible to the subjects and could potentially provide cues to the subjects. In practice, the fact that these markers are positioned without reference to the internal coordinate system of the physical object makes this risk minimal.

## 7.1 Augmented Environments for Experiments

Although there have been many experiments conducted using virtual environments, most of these have consisted of isolated experiments, rather than testbeds for conducting a variety of experiments [1, 5, 11, 17, 19], and much of this work has focused on virtual reality, rather than augmented reality. The closest work to ours are two virtual reality testbeds for subject testing, VRMAT and VEPAB.

Poupreva et al. [18] developed a testbed for studying manipulation techniques in a fully immersive environment which uses a non-see-through head mounted display. Their Virtual Reality Manipulation Assessment Testbed (VRMAT) focuses on basic direct manipulation tasks such as reaching, grasping, moving and orienting virtual objects. Besides the obvious difference of virtual versus augmented reality, the granularity of testing differs. Unlike our system, VRMAT does not focus on fine grained manipulation by fingers, but rather on whole arm motions.

The scope of Lampton et al.'s Virtual Environment Performance Assessment Battery (VEPAB) [14] is even broader. VEPAB is a set of tasks which are used to quantify human performance on vision, locomotion, tracking, object manipulation and reaction time tasks in virtual reality. Three manipulation tasks are considered, slide (grasp and move a control bar horizontally), dial (grasp and rotate to an indicated orientation), and bin (pull a ball out of one bin and push it into another). Two control devices were tested, joysticks and Spaceballs. The basic manipulation tasks are similar to ones we test in the VHL, but there are two major differences. The VHL is an augmented rather than a strictly virtual environment and we test as input natural hand movement using a prehension metaphor [15] rather than an indirect control device.

## 7.2 Calibration

Calibration in AR is hard and has been addressed by several people for several different types of equipment. Tuceryan et al. looked at monitor-based augmented reality [22], Bajur et al. and Koller et al. each merged video-camera images into synthetic scenes [3,

13], Azuma and Bishop studied see-through head mounted displays (HMDs) [2], and State et al. hybrid tracking systems [20].

Of the calibration requirements listed by Hollerbach and Wampler, accuracy is the hardest to satisfy in an augmented environment, and consequently the one to which most attention has been paid. According to State, "Current AR systems cannot convincingly meet this requirement. Typically a virtual object appears to swim about as the user moves..." [20]. We in fact take advantage of this effect and use it in the first calibration component to ensure that we have determined the plane of the virtual image by checking for a lack of visual swim.

Swimming can be the result of either spatial or temporal errors. This paper deals only with the first issue, spatial registration. We do not consider temporal registration, the problem of lag, but this must always be addressed. In our experiments the 60 Hz sampling rate and roughly 1.5 frame time lag have proved acceptable, although future experiments may require greater temporal accuracy. The limitations imposed by our 60 Hz update rate for the display suggest that the first problem to be tackled should be the spatial errors.

Some of the issues for calibration in augmented reality are obtaining a good geometric model of the physical objects, registration of locations (both physical and virtual) in the environment, calculating the correct point of view, and updating data in real time. System calibration is sometimes decomposed into several subsystem calibrations with predefined interactions. Prior calibration work has sometimes produced subsystems which satisfy the requirements for experimental systems. However, when these components are joined, the overall system does not. This is not a limitation in the other work, but merely reflects the simpler goal of overall system accuracy. By contrast, we have tried to satisfy a more comprehensive set of requirements, including those for experimental systems.

One common approach to calibration is to have the user visually align physical and virtual objects. This combines the point of view calibration with object calibration, which means that the system cannot be used for perception tests. For example, Ghazisaedy et al. calibrated a magnetic tracker with an ultrasound tracker by aligning both trackers with virtual blocks prior to use of the system [9]. This calibration satisfies the "no cues" requirement because there were no residual cues left after calibration. However, the calibration was not independent as it relied on the calibrator's perception of the 3D location of the virtual objects.

The alignment approach was also taken by Azuma and Bishop [2]. They calibrated a see-through HMD by registering one real object (the calibration cube) with one set of virtual objects (three bars representing a coordinate system) using a monocular view. To obtain the frame-to-tracker transformation, the user touches each corner of the cube. The eye-to-tracker transformation builds on this by using a bore-sight. They obtained  $\pm 4$  mm accuracy in X and Y directions, and  $\pm 5$  mm in Z (up/down).

Other systems isolate the point of view from other calibrations, but do so by using default values, which are not subject-specific. State et al. [20] implemented a hybrid approach by integrating landmark and magnetic tracking. Two-colored concentric landmarks are placed on non-moving physical objects. A mechanical arm is used to determine the center of the landmarks in world coordinates. From the camera images of the landmarks, head position is determined. The cameras are preset with an interocular distance of 64 mm and a 4 deg convergence angle.

The GRASP system is a platform for research in augmented reality applications. Its "most important feature...which differentiates it from an ordinary 3D graphics or virtual reality system, is its collection of calibration techniques" [22]. These techniques could, with fairly minor modification, be used for experimental systems. There are six sub-calibrations: image, camera, pointer, tracker, object and tracker markers. The majority of these sub-calibrations are independent with the exception of the object calibration. Two techniques are

provided: one of which is dependent on the pointer calibration and the other on the image calibration. The first technique gives an average distance error of 6.48 mm and the second 5.25 mm. In this system, the perspective transformation is dependent on the image and camera calibrations, not on the viewer's (or subject's) physical characteristics such as interocular distance, focal length or image center.

## 8 Conclusions

Augmented reality is a powerful tool which supports natural manipulation techniques as a means of improving the human-computer interface. Experimentation in this area will lead to even better interfaces if we can develop an understanding for how humans perform in augmented environments. One important area of experimentation is the effect of calibration on human performance.

It is possible to calibrate systems to support experimental requirements while still meeting the requirements of traditional calibration: robustness, flexibility, accuracy, convenience of use and expense. We have implemented this method, and it has been used by three kinesiology researchers in experimental research [23, 24].

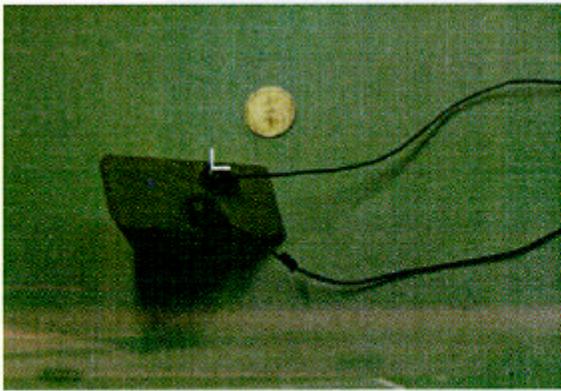
Similar methodology and calibration techniques are also useful for non-experimental systems utilizing different tracking devices because the independence criteria can prevent cascading recalibrations when components change and it can help to identify appropriate modularization in the software that often results in a more robust and a more efficient implementation.

## 9 Acknowledgments

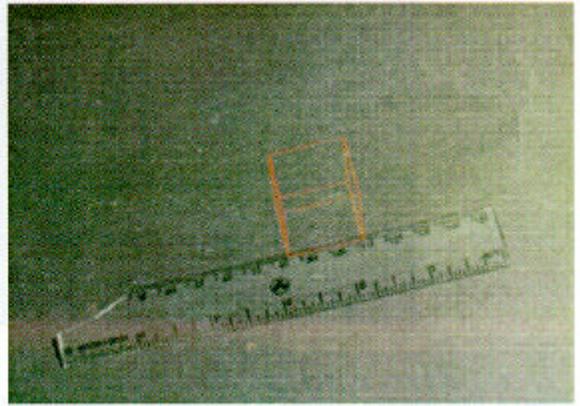
Feedback from Yanqing Wang and Alissa Fourkas (Simon Fraser University) and Dr. Satoru Kawai (Tezukayama Junior College, Japan), who are users of the VHL, is gratefully acknowledged. Figure 1 was provided by Wang and Kawai. Scott Ralph and Rob Walker provided useful comments on the paper, and Scott provided much appreciated last-minute help with photography. Comments from the anonymous reviewers were extremely helpful in revising the manuscript to clarify a number of issues. Financial support was provided by the Natural Sciences and Engineering Research Council of Canada through various grants and fellowships.

## References

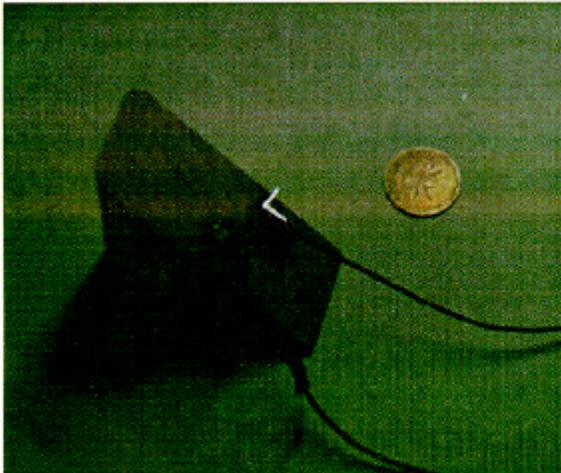
- [1] K.W. Arthur, K.S. Booth, and C. Ware. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*, 11(3):239–265, 1993.
- [2] Ronald Azuma and Gary Bishop. Improving static and dynamic registration in an optical see-Through HMD. In *Proceedings of SIGGRAPH '94*, pages 197–204, July 1994.
- [3] M. Bajura and U. Neumann. Dynamic registration correction in video-based augmented reality systems. *IEEE Computer Graphics and Applications*, pages 52–60, September 1995.
- [4] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, pages 203–210, July 1992.
- [5] Woodrow Barfield and Claudia Hendrix. Factors affecting presence and performance in virtual environments. In *Interactive Technology and the New Paradigm for Healthcare*, chapter 4, pages 21–28. IOS press and Ohmsha, 1995.
- [6] D. Drascic, Grodski J.J., Milgram P., Ruffo K., Wong P., and Zhai S. Argos: A display system for augmenting reality. In *Video Proceedings of INTERCHI '93: ACM Conference on Human Factors in Computing Systems*, page 521, April 1993.
- [7] Essilor International. *Digital C.R.P. (Pupillometer Reference Manual)*, 1997.
- [8] Steven Feiner, Blair MacIntyre, and Doree Seligmann. Knowledge-based augmented reality. *Communications of the ACM*, 36(7):52–62, July 1993.
- [9] M. Ghazisaedy, D. Adamczyk, D.J. Sandin, R.V. Kenyon, and T.A. DeFanti. Ultrasonic calibration of a magnetic tracker in a virtual reality space. In *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'95)*, pages 179–188, 1995.
- [10] John M. Hollerbach and Charles W. Wampler. The calibration index and taxonomy for robot kinematic calibration methods. *The international journal of robotics research*, 15(6):573–591, December 1996.
- [11] S. Honda, H. Tomioka, T. Kimura, T. Ohsawa, K. Okada, and Y. Matsusuhia. A virtual office environment based on a shared room realizing awareness space and transmitting awareness information. In *Proceedings of the ACM UIST '97 Symposium on User Interface Software and Technology*, pages 199–208. ACM Press, New York, 1997.
- [12] K.C. Knowlton. Computer displays optically superimposed input devices. *Bell System Tech. J.*, 56:367–383, March 1977.
- [13] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan. Real-time vision-based camera tracking for augmented reality applications. In *ACM symposium on Virtual Reality Software and Technology 1997*, pages 87–94, 1997.
- [14] Donald R. Lampton, Bruce W. Knerr, Stephen L. Goldberg, James P. Bliss, J. Michael Moshell, and Brian S. Blau. The virtual environment performance assessment battery (VEPAB): Development and evaluation. *Presence: Teleoperators and Virtual Environments*, 3(2):145–157, 1994.
- [15] C.L. MacKenzie, E.D. Graham, C.G.L. Cao, and A.J. Lomax. Virtual hand laboratory meets endoscopic surgery. In *Proceedings of Medicine Meets Virtual Reality (MMVR:7)*, pages 212–218, January 1999.
- [16] C.L. MacKenzie and T. Iberall. *The Grasping Hand*. Amsterdam: Elsevier Science, 1994.
- [17] R. Pausch, D. Proffitt, and G. Williams. Quantifying immersion in virtual reality. In *Computer Graphics (SIGGRAPH '97 Proceedings)*, pages 13–18, August 1997.
- [18] I. Pouprey, S. Weghorst, M. Billinghurst, and T. Ichikawa. A framework and testbed for studying manipulation techniques for immersive vr. In *ACM symposium on Virtual Reality Software and Technology 1997*, pages 21–28, 1997.
- [19] G. Robertson, M. Czerwinski, and M. van Dantzich. Immersion in desktop virtual reality. In *Proceedings of the ACM UIST '97 Symposium on User Interface Software and Technology*, pages 11–20. ACM Press, New York, 1997.
- [20] A. State, G. Hirota, D.T. Chen, B. Garrett, and M. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In *SIGGRAPH 96 Conference Proceedings*, pages 429–438, August 1996.
- [21] Andrei State, Mark A. Livingston, Gentaro Hirota, William F. Garrett, Mary C. Whitton, and Henry Fuchs. Technologies for augmented-reality systems: Realizing ultrasound-guided needle biopsies. In *SIGGRAPH 96 Conference Proceedings*, pages 439–446, August 1996.
- [22] M. Tuceryan, D.S. Greer, R.T. Whitaker, D.E. Breen, C. Crampton, E. Rose, and K.H. Ahlers. Calibration requirements and procedures for a monitor-based augmented reality system. *IEEE Transactions on Visualization and Computer Graphics*, 1(3):255–273, September 1995.
- [23] Y. Wang, C.L. MacKenzie, and V.A. Summers. Object manipulation in virtual environments: human bias, consistency and individual differences. In *Proceedings of ACM CHI'97 Conference on Human Factors in Computing Systems*, Extended Abstracts, pages 349–350, 1997.
- [24] Y. Wang, C.L. MacKenzie, V.A. Summers, and K.S. Booth. The structure of object transportation and orientation in human-computer interaction. In *Proceedings of ACM CHI'98 Conference on Human Factors in Computing Systems*, pages 312–319, 1998.



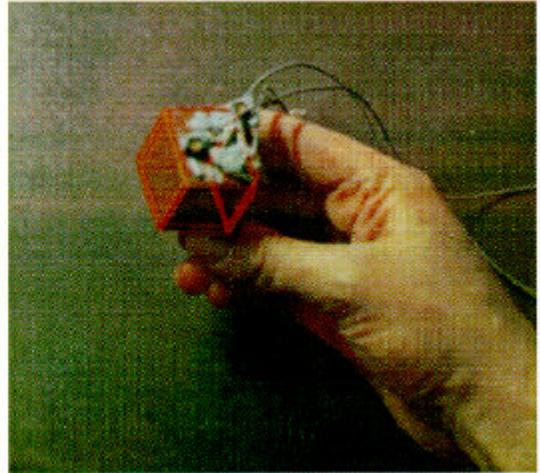
(a)



(d)



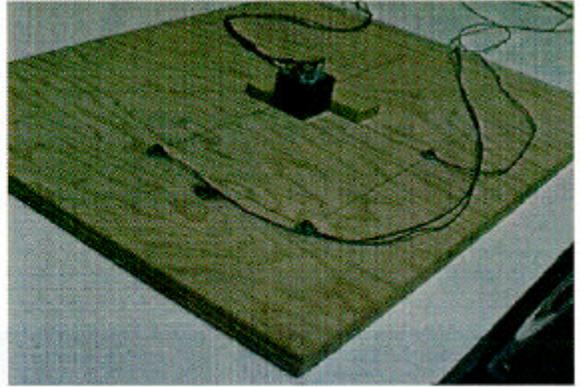
(b)



(e)



(c)



(f)

(a)(b) Workspace calibration. For demonstration purposes, 2 markers are mounted above the desktop. The “L” shaped virtual cross aligns with exactly 1 marker, even when seen from 2 different viewpoints. (c) Eye calibration. The subject is placing the second “calibration bar”. (d) A virtual cube is measured to test eye calibration. (e) Stereo image (2 viewpoints) of a a block augmented with a wire-frame. (f) Object calibration. Object sitting in calibration frame.