Sensing the Position & Orientation of Hand-Held Objects: An Overview of Techniques

Krispin Leydon
12/3/01
Technical Report UL-CSIS-03-2
Interaction Design Centre
CSIS Department
University of Limerick

We live between two realms: our physical environment and cyberspace. Despite this dual citizenship, the absence of seamless couplings between these parallel existences leaves a great divide between the worlds of bits and atoms. At the present, we are torn between these parallel but disjoint spaces [14].

Bridging the gap between physical and virtual environments is a driving interest in the field of human computer interaction (HCI) with far-ranging implications for entertainment, communication and many other human activities. Interest in more seamless physical/virtual couplings is fueled by growing recognition of inherent limitations of the desktop computing paradigm on one hand, and renewed appreciation for the “rich” affordances provided by physical, tangible tools and media on the other. A sampling of the numerous phrases alluding to seamless physical/virtual couplings—“Hybrid Reality”, “Mixed Reality”, “Augmented Reality”, “Ubiquitous Computing”, “Pervasive Computing” — suggests a convergence of perceived need from multiple perspectives. The tasteful blending of physical and virtual worlds is a challenge of considerable contemporary relevance.

Why Track Position & Orientation of Hand-Held Objects?

One aspect of this challenge is real-time tracking the position and orientation of hand-held objects: six degree of freedom (6-DOF) tracking. This capability is important because of the central role that manual manipulation of objects occupies in human experience. The ability to hold and handle objects facilitates tool creation, artifact use, interpersonal communication, and the exploration of natural and artificial environments. Findings in brain science indicate that manual touch plays a primary role in human perception [10]. Observational studies in HCI emphasize the
value of manual object handling in educational activities [7,12]. According to research in behavioral science, an interdependence exists between manual manipulation of objects and the development of linguistic capabilities [9,16]. Since manual manipulation of objects shapes human experience in so many ways, tracking the movement of hand-held objects is an important aspect of integrating our virtual and the physical habitats.

A system capable of tracking the position and orientation of hand-held objects makes a number of potentially rich new forms of human/artifact interaction possible. Through simulated proprioception, objects would be able to sense and respond to the way they are handled – individually and as assemblies. Coffee mugs could sense consumption of the last drop, phones could sensing being held to one’s ear, balls could help to teach would-be jugglers, and physical chess pieces could dictate play on a virtual board [5]. Though these are “toy” scenarios, they hint at large, varied and potentially fertile interaction design spaces that have yet to be explored. The extent to which real-time tracking of hand-held objects can facilitate mediation of human activities remains unknown.

**Discussion of Position & Orientation Sensing Approaches**

In order for a general-purpose 6-DOF tracking system for hand-held objects to be viable, it must meet numerous requirements. An ideal solution must be accurate, inexpensive, safe and wireless. Any components of the system mounted in/on hand-held objects must have minimal power requirements. The active tracking volume must be void of physical obstructions and large enough to facilitate a normal range of human motion.

A number of tracking approaches are currently available or under development. Each approach addresses a number of the requirements stated above, however none at present offers a comprehensive solution. In this section, we present an overview of several interesting design tacks. For each approach we strive to answer the following questions: What is it? How does it work? Where is it in use? What are its chief advantages? What are its shortcomings?

[Note: For a quick summary of approaches and their respective strengths and weaknesses, see the table on page 12.]
Inertial + Magnetic Tracking

Inertial measurement units (IMUs) determine their own position $p$ and orientation $o$ in space through with respect to initial conditions: $p_0$ and $o_0$. Since IMUs don’t rely on an external coordinate system, inertial tracking is often described as “dead reckoning”.

IMUs typically sense the quantities: angular velocity, measured by gyroscopes or angular rate sensors, and acceleration, measured by accelerometers. Displacement and angular displacement are subsequently calculated through integration with respect to time.

Since IMUs are only capable of determining their position and orientation with respect to initial conditions, they are often combined with sensors that respond to an external frame of reference: electronic compasses (magnetometers) and global positioning system (GPS) receivers. Such augmentation enables conversion from relative to absolute coordinates and periodic re-calibration of the IMU.

Historically, inertial measurement techniques have been employed in biomedical motion analysis, airbag deployment and personal navigation systems. Recent advances in sensor fabrication techniques have yielded accelerometers and gyroscopes that are small, accurate, low power and low-cost. These new sensors make inertial measurement a candidate technology for real-time position and orientation tracking of hand-held objects.

The chief advantage of inertial measurement is that no external frame of reference is required. Inertial measurement places no limits on the active tracking volume or on the number of objects that can be tracked within this volume. Inertial measurement gives objects proprioception in a most literal sense.

Despite these advantages, IMUs are of limited use for tracking purposes for several reasons. First, since position and orientation are determined through single and double integration of original sensed quantities, even small sensor offsets and inaccuracies cause tracking to degenerate rapidly over time [5]. Second there is no certain way to distinguish contributions of gravity from other constant accelerations. (Typically, the assumption that gravity is the only constant acceleration is made.) Third, a 6-DOF IMU requires numerous
sensors—at least three accelerometers and three angular rate sensors. This requirement has negative implications with respect to IMU size, complexity, power consumption and cost.

Adding an electronic compass makes it possible to calibrate the IMU’s orientation periodically (and thus create some improvement in orientation tracking accuracy over time). This augmentation, however, requires additional sensors and introduces a complex problem: distinguishing earth’s magnetic field from field distortions common in indoor environments—distortions due to electric currents and the presence of ferromagnetic materials.

Optical Marker Tracking

Optical marker tracking reconstructs an object’s position and orientation from the position of “markers” mounted on the object’s surface. Marker position is determined by triangulation from three or more 2D images obtained from video cameras placed at the periphery of the active tracking volume. Markers are typically reflective surfaces or light emitting diodes (LEDs) that flashing at characteristic frequencies.

Optical marker tracking has proven effective in situations requiring a large active volume, high sampling rates and high accuracy [8]. Systems employing passive reflective markers have no on-object power requirements. Optical marker tracking is currently used in character animation and sports motion analysis.

Many factors limit the usefulness of marker based optical techniques for tracking hand-held objects. Most significant is the need for line-of-sight, since hand-held objects are by definition partially occluded by a person’s hand from all perspectives and totally occluded by a person’s body from certain perspectives. Additional camera angles can be employed to minimize the problems created by occlusion but introduce additional cost and computational complexity. Since marker position is determined by interpreting video data from multiple video streams, computational complexity is high enough to prohibit real-time tracking in most commercial systems. [Note: A sub-domain of optical tracking concerns the relative movement and deformation of facial features using a single camera. Real time tracking in such systems has proved possible][6]. In addition to cost, complexity and the need for line-of-sight, optical tracking systems have been plagued
by the problem of “crossover ambiguity”: markers being mistaken for one another.

**Acoustic Tracking**

Acoustic tracking makes use of ultrasonic sound waves to determine position and orientation. Relevant applications for acoustic tracking include automated spotlight control systems for theatre performance, and 3D mice and head-mounted displays (HMDs) for virtual reality applications [1,22].

Two approaches to acoustic tracking are time-of-flight tracking and phase-coherence tracking. Gregory Baratoff et al. provide the following description of these two techniques in “Tracking Devices” [4]:

Time-of-flight tracking works by measuring the amount of time that it takes for sound emitted by transmitters on the target to reach sensors located at fixed positions in the environment... By measuring when the sounds arrive at the various sensors, the system can determine the length of time it took for the sound to travel from the target to the sensors, and thereby calculate the distance from the target to each of the sensors. Since there will only be one point inside the volume delimited by the sensors that satisfies the equations for all three distances, the position of the target can be determined. [For] position, only one of the transmitters is needed. Orientation is determined by the differences in location indicated by these calculations for each of the three sensors...Time-of-flight trackers typically suffer from a low update rate, brought about by the low speed of sound in air... Another problem is that the speed of sound in air is affected by such environmental factors as temperature, barometric pressure, and humidity.

Phase coherence tracking works by measuring the difference in phase between sound waves emitted by a transmitter on the target and those emitted by a transmitter at some reference point. The phase of a sound represents the position on the sound wave... As long as the distance traveled by the target is less than one wavelength between updates, the system can update the position of the target. By using multiple transmitters, as with time-of-flight tracking, orientation can also be determined. Since they work by periodic updates of position, rather than by measuring absolute position at each time step, phase-coherence tracking devices are subject to error accumulation over time.

As with optical tracking, occlusion impedes effective tracking of hand-held objects. Acoustic reflections and the need for powered sound sources are additional limitations of acoustic tracking [23].
Ringdown Passive RFID Tag Tracking

Ringdown passive tag tracking, as envisioned by Kai-yuh Hsiao of the MIT Media Lab, makes use of an induced response from magnetically resonant ID tags in order to obtain information regarding tag position and orientation [11]. This approach to tracking rests on technical foundations laid by electronic article surveillance (EAS) systems.

Ringdown tracking works in the following way: First, an inductor coil within the tag “reader” generates a magnetic oscillatory pulse at a particular frequency. If an ID tag with that characteristic frequency is present, inductive coupling causes it to oscillate. Next, the reader stops oscillating and begins to “listen” for persisting oscillation from a resonant tag (a dying exponential with respect to time). The total power of this “ringdown” oscillation is a function of proximity to the tag reader and orientation with respect to magnetic field lines from the tag reader coils. By enabling the reader to emit then listen for magnetic oscillations of multiple frequencies, multiple tag tracking can be supported [11]. Through multiplexed use of mutually orthogonal Helmholtz coils (which generate roughly parallel field lines along three axes of an enclosed volume) and orthogonally oriented resonant tags, position and orientation in 3D space can be determined [11].

This manner of position/orientation tracking has not been pursued experimentally, due to Mr. Hsiao’s decision to continue with a related tracking modality better suited to robust real-time tracking of multiple objects. This alternate technique, Swept Frequency RFID Tag Tracking, is discussed in the following section. The disadvantages specific to the ringdown tracking method described above are:

- The ringdown approach requires waiting for a ringdown response at each frequency being tracked. As the number of tracking frequencies grows, overall system delay increases.
- Since the ringdown approach targets specific frequencies, frequency drift within reader or tags introduces problems necessitating recalibration.
- The reader system requires prior knowledge of the precise resonant frequencies of the tags being tracked [11].
Swept Frequency Passive RFID Tag Tracking

Swept frequency tracking, as pioneered by Hsiao, relies on the same underlying technology and geometry as ringdown tracking: passive resonant RFID tags and a reader which generates oscillating magnetic fields sequentially in a set of mutually orthogonal Helmholtz coils.

In contrast to ringdown tracking, swept frequency tracking does not “ping” certain frequencies, it repeatedly sweeps through a frequency range, identifying frequencies that correspond to tags in the active tracking region. Swept frequency tag tracking does not rely on power emitted by tags during ringdown to assess proximity. Instead, it makes use of the following phenomena:

When a tag enters [a] reader’s field and is exposed to a magnetically-coupled signal at its resonant frequency, it pulls current from the reader coil. This energy drawn from the field causes a detectible change in the perceived inductance of the coil, which manifests itself as a dip in the voltage or current being passed through the coil [11].

The dip’s magnitude is a function of tag position and orientation with respect to a reader coil. By obtaining dip magnitude for mutually orthogonal tags co-located in a handheld object through the use of mutually orthogonal helmholtz reader coils activated in sequence, the object’s 3D position and orientation can be determined.

Since the tag reader does not stop and listen for ringdown response, multiple tag tracking is possible without affecting tracking speed. Since the reader sweeps through a frequency range rather than checking specific individual frequencies, the system is more robust than ringdown tracking with respect to frequency drift in tags or reader.

The advantages of swept frequency tag tracking are numerous. Since tags are passive there are no on-object power requirements. Since tracking is accomplished via magnetic fields, line-of-sight is not required. The low cost of EAS tags and straightforward nature of the reader circuitry contribute to potentially low system cost. The ability to track multiple tags in real time makes simultaneous tracking of multiple handheld objects possible.

The chief disadvantages of swept frequency tag tracking (according to early experimental results) are low resolution
(± 10 degrees, ± 2 cm reported) and a relatively small tracking volume (on the order of .5 m$^3$) [13]. Through refinement of reader circuitry, coil geometry and addition of post-processing (special mapping and fitting functions, etc.) both limitations can be improved upon. The fundamental limits of this technique for tracking handheld objects remain unknown.

[NOTE: Ringdown passive tag tracking and swept-frequency passive tag tracking are special cases of AC magnetic tracking, convenient with respect to the power requirements and cost of on-object sensors. There are commercial wireless AC Magnetic tracking systems with powered sensors that offer better range and accuracy than the passive tag-based alternatives presented above – notably the “StarTrak” from Polhemus [15]. Currently, no commercially available system has wireless powered sensors small enough for tracking handheld objects.]

**Pulsed DC Magnetic Tracking**

A disadvantage common to AC magnetic tracking techniques is that accuracy is affected by the presence of metal objects in or near the active tracking volume. Alternating magnetic fields generate eddy currents in metal, and these eddy currents generate magnetic fields. As a result, the original field is distorted and tracking accuracy compromised.

An alternative to AC magnetic tracking that is more robust with respect to the presence of metal is pulsed DC magnetic tracking. Pulsed DC magnetic tracking makes use of DC magnetic fields generated for short periods of time rather than a continuously oscillating magnetic field. DC magnetic sensors take two readings per sample: the first is taken when no reference field is being generated, the second is taken during a pulse of external field generation. Through subtracting the first reading from the second reading, the effects of ambient magnetic influences (such as earth’s magnetic field) are removed. The second reading is taken after the rising edge of the reference field pulse. This gives eddy currents induced by the pulse’s rising edge the time to decay and approach steady state. In this way, field distortions from metal objects are minimized [1,2].

Pulsed DC magnetic tracking is relatively new technique. Its main applications have been for HMD tracking in metal-filled
environments (such as cockpits) and motion capture for virtual reality and animation [3].

This tracking technique does not require line-of-sight, has excellent range and accuracy, and is more robust with respect to metal distortion than AC magnetic tracking techniques. The principal disadvantages of pulsed DC magnetic tracking are an on-object power requirement and cost [20].

**Electric Field Sensing**

Another class of tracking approaches is electric field sensing. This class of modalities makes use of the measurable distortion that an object with electrical characteristics (such as capacitance) creates within an electric field oscillating at low frequency. Historically, electric field sensing has been known as “capacitive sensing”.

Electric field sensing exists naturally in certain species of fish, and has been implemented artificially as a musical interface called the “theremin”. Though electric field sensing has been accorded surprisingly little attention since the invention of the theremin (circa 1917) recent work at the MIT Media lab—primarily the work of Mr. Joshua Smith—has demonstrated the technique’s potential for sensing human motions in computer-mediated activity [18,19]. Demonstration applications include 3D mouse-like control via hand motion and a phone mock-up that can sense being held in a person’s hand and being held to a person’s head [19].

Smith presents three modes of field sensing—Transmitter Loading, Shunt, and Transmit—that are most conveniently discussed in reference to the lumped circuit model below [17]. This model consists of a low frequency ~50Hz transmitter, a receiver, and a field-altering object: a human hand.

![Lumped circuit model of electric field sensing parameters][1]
Transmitter Loading Mode: As the object being sensed moves towards and away from the transmitter, the value of $C_1$ changes measurably. Theremins make use of this mode. In transmitter loading mode, no receiver is required [18].

Transmit Mode: In this mode, the object is the transmitter, or is capacitively coupled to the transmitter via $C_1$. Object movement changes the effective distance between transmitter and receiver, and this creates measurable changes in $C_2$ [18].

Shunt Mode: In this mode, the object is not connected to transmitter or receiver. When no object is present, displacement current flows only through $C_0$. When an object enters the field, $C_0$, $C_1$ and $C_2$ values change. This results in measurable changes in the total displacement current flowing from transmitter to receiver [18].

By using multiple electrodes and multiplexing the transmitter, multiple measurements can be made. From these measurements and certain modeling assumptions, information concerning the position and orientation of an object in the field can be inferred.

Of the three modes of electric field sensing, shunt mode yields the most information. In transmit and transmitter loading modes, $n$ measurements can be made using $n$ transceivers, while in shunt mode it is possible to make $n(n-1)/2$ measurements with $n$ electrodes. Each unique measurement contributes unique information related to the geometry of the situation [19].

Electric field sensing has numerous advantages. The hardware required is simple and inexpensive. The technique is scalable; better resolution or a larger active tracking volume can be obtained by increasing the number of transceivers. This scalability is appealing because it makes it possible to “collect as much or as little information as [is] needed in a particular application” [19]. In this respect, electric field sensing is preferable to optical tracking. The active volume can be shielded from noise through the use of ground planes in electric field sensing. In this respect, electric field sensing is preferable to magnetic tracking [19].

A major disadvantage of electric field sensing is the mathematical complexity of inferring accurate information about an object’s position and orientation from indirect measurement of electrical properties. The couplings between
object transmitter and receiver are nonlinear. Since there are more unknown parameters than measurements, position and orientation must be found in terms of ambiguity classes and probability distributions. Another disadvantage of electric field sensing (as it has been explored thus far) is that it presumes the type of object being sensed. Thus, tracking different types of objects requires the use of different modeling assumption sets. Most of the work that has been done in electric field sensing concerns modeling the human hand and body. According to Smith, real time position tracking appears to be tractable, while the tractability of real time orientation tracking remains unknown [18]. The mathematical complexity of electric field sensing is daunting, however the modality has great “potential”.

Conclusions

Tracking the position and orientation of hand-held objects is an important challenge in the goal to create seamless couplings between our virtual and physical tools and media and environments. Though numerous approaches to this challenge have been investigated, none has emerged as clearly superior. Whether a robust, general-purpose solution will emerge through refinement, new innovation or clever combination remains to be seen.
## Approaches to Position/Orientation Sensing for Hand-Held Objects (6DOF, Wireless)

<table>
<thead>
<tr>
<th>Approach</th>
<th>Examples</th>
<th>Sensed Phenomena</th>
<th>Current Applications</th>
<th>Accuracy</th>
<th>Cost</th>
<th>Volume Constraints</th>
<th>Update Rate</th>
<th>Size of Handheld Component</th>
<th>Batteries Required in Handheld Component?</th>
<th>Tracking of Multiple Objects in Real Time Simultaneously</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inertial Measurement + Magnetometer</td>
<td>Masters Thesis Project, Ar Yosef Benbasat, MIT</td>
<td>Acceleration, Angular Acceleration, Earth’s Magnetic Field</td>
<td>Biomedical motion analysis, dead-reckoning navigation to support GPS</td>
<td>Variable; Displacement deteriorates rapidly over time</td>
<td>XXX$ per hand-held unit</td>
<td>Can accommodate any tracking volume.</td>
<td>Typically XX-XXXHz</td>
<td>Medium</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Optical Marker Tracking</td>
<td>InMotion Systems “CODA”</td>
<td>Light</td>
<td>Animation of human movement</td>
<td>Variable; depends on tracking volume. Xmm/degree resolution can be attained</td>
<td>XXX$ per hand-held unit</td>
<td>Large volumes possible. Extra volume surrounding active tracking area required to accommodate camera field of</td>
<td>Typically not real-time</td>
<td>?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Acoustic Tracking</td>
<td>Acoustic Positioning Research “Solospot”</td>
<td>Sound</td>
<td>Automated spotlight control, 6DOF Mouse Interface</td>
<td>Variable; depends on tracking volume. Xmm/degree resolution can be attained</td>
<td>XXX$ per hand-held unit</td>
<td>Large volumes possible. (Latency increases with volume)</td>
<td>XXHz</td>
<td>?</td>
<td>No</td>
<td>?</td>
</tr>
<tr>
<td>Electric Field Sensing</td>
<td>PhD Thesis Project, Joshua Reynolds Smith, MIT</td>
<td>Capacitance</td>
<td>Musical Interface (Theramin), 3D Mouse Interface (Experimental)</td>
<td>Variable; depends on tracking volume. Xmm/degree resolution can be attained</td>
<td>?</td>
<td>Small tracking volume; Xmm^3</td>
<td>XXHz</td>
<td>?</td>
<td>No</td>
<td>?</td>
</tr>
<tr>
<td>Swept-Frequency RFID Tag Tracking</td>
<td>PhD Thesis Project, Kai-Yuh Hsiao, MIT</td>
<td>Energy captured inductively by passive resonant tags</td>
<td>Musical Interface (experimental)</td>
<td>Variable; depends on tracking volume. Xmm/degree resolution can be attained</td>
<td>XXX$ for multiple hand-held units</td>
<td>Small tracking volume; Xmm^3</td>
<td>30Hz</td>
<td>Small</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Ringdown RFID Tag Tracking</td>
<td>PhD Thesis Project, Kai-Yuh Hsiao, MIT</td>
<td>Energy captured inductively by passive resonant tags</td>
<td>Shoplifting &amp; Library Tracking Systems</td>
<td>?</td>
<td>XXX$ for multiple hand-held units</td>
<td>Small tracking volume; Xmm^3</td>
<td>?</td>
<td>Small</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Pulsed DC Magnetic Tracking</td>
<td>Ascension Technologies “MotionBar Wireless”</td>
<td>Magnetic Field Strength</td>
<td>Animation of human movement</td>
<td>Variable; depends on tracking volume. Xmm/degree resolution can be attained</td>
<td>XXX$ for multiple hand-held units</td>
<td>Medium tracking volume; Xmm^3</td>
<td>Not a hand-held system</td>
<td>Medium</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

[1] Limit of the “example” commercial system
Useful Links

Domain Overview:

- [http://www.cis.upenn.edu/~hms/pelachaud/workshop_face/subsubsection3_7_1_4.html](http://www.cis.upenn.edu/~hms/pelachaud/workshop_face/subsubsection3_7_1_4.html)
- [http://www.media.mit.edu/resenv/papers.html](http://www.media.mit.edu/resenv/papers.html)
- [http://www.isdale.com/jerry/VR/MotionCapture_Links.html](http://www.isdale.com/jerry/VR/MotionCapture_Links.html)
- [http://www.wave-report.com/tutorials/MoTrak.htm](http://www.wave-report.com/tutorials/MoTrak.htm)

Inertial + Magnetic Tracking:

- [http://www.isense.com/company/whatismotion.htm](http://www.isense.com/company/whatismotion.htm)
- [http://user.cs.tu-berlin.de/~remuss/links.html](http://user.cs.tu-berlin.de/~remuss/links.html)
- [http://www.sem.samsung.co.kr/product/eappintroview.jsp?gcode=E1&code=JD0&jsessionid=604621004525737039](http://www.sem.samsung.co.kr/product/eappintroview.jsp?gcode=E1&code=JD0&jsessionid=604621004525737039)
- [http://www.sphere.net.au/](http://www.sphere.net.au/)
- [http://www.tri-m.com/products/precisionnav.products.html](http://www.tri-m.com/products/precisionnav.products.html)
- [http://www.inertial.co.uk/](http://www.inertial.co.uk/)

Optical Marker Tracking:

- [http://www.metamotion.com/motion-capture/optical-motion-capture-1.htm](http://www.metamotion.com/motion-capture/optical-motion-capture-1.htm)
- [http://www.dai.ed.ac.uk/CVonline/LOCAL_COPIES/RINGER1/mocap_overview.htm](http://www.dai.ed.ac.uk/CVonline/LOCAL_COPIES/RINGER1/mocap_overview.htm)
- [http://www.css.tayloru.edu/instrmat/graphics/hypgraph/animation/motion_capture/motion_optical.htm](http://www.css.tayloru.edu/instrmat/graphics/hypgraph/animation/motion_capture/motion_optical.htm)

Acoustic Tracking:


AC Magnetic Tracking (Ringdown & Swept-Frequency included)

- [http://www.polhemus.com/stardstech.htm](http://www.polhemus.com/stardstech.htm)

Pulsed DC Magnetic Tracking:

- [http://www.ascension-tech.com/](http://www.ascension-tech.com/)
Electric Field Sensing:

- http://www.media.mit.edu/~jrs/home.html

Works Cited


