Olwal, A. LightSense: Enabling Spatially Aware Handheld Interaction Devices. *Proceedings of ISMAR 2006 (IEEE and ACM International Symposium on Mixed and Augmented Reality)*, Santa Barbara, CA, Oct 22-25, 2006, pp 119-122.

# LightSense: Enabling Spatially Aware Handheld Interaction Devices

Alex Olwal
School of Computer Science and Communication, Royal Institute of Technology (KTH)
alx@csc.kth.se


**Figure 1. The LightSense system tracks the LED on commercial cell phones, enabling them to be used as spatially aware handheld devices. The outside-in approach tracks the light source and streams the data to the phone over Bluetooth. a) A wall-mounted map with embedded light sensors provides hotspot tracking. b) A table-top setup tracks the phone with a camera through a diffused glass surface. c) The spatially aware device augments a physical map with a detailed interactive road map of the area of interest.**

### Abstract

The vision of spatially aware handheld interaction devices has been hard to realize. The difficulties in solving the general tracking problem for small devices have been addressed by several research groups and examples of issues are performance, hardware availability and platform independency. We present LightSense, an approach that employs commercially available components to achieve robust tracking of cell phone LEDs, without any modifications to the device. Cell phones can thus be promoted to interaction and display devices in ubiquitous installations of systems such as the ones we present here. This could enable a new generation of spatially aware handheld interaction devices that would unobtrusively empower and assist us in our everyday tasks.

**CR Categories:** H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities; H.5.2. [User Interfaces]: Graphical user interfaces, Input devices and strategies; I.3.6 [Methodology and Techniques]: Interaction techniques.

**Keywords:** Augmented reality, mixed reality, ubiquitous, spatially aware, portable, mobile, handheld, LED, cell phone.

## 1 Introduction

In Weiser's vision of ubiquitous computing [18], computers assist us in the background, letting us focus on our real-world tasks, instead of steering our focus of attention to the human-computer interface. The interface to this invisible computing infrastructure is filled with Tabs, Pads and Boards. Tabs are tiny devices that provide access to situated digital content with basic user input and small displays for output. Pads are notebook-sized devices with more sophisticated user I/O and computational power. Boards are large interactive wall-size displays.

With embedded sensing in the environment we are presenting an approach that promotes today's commercial off-the-shelf cell phones into personal Tabs with the computational power that surpasses Weiser's Pads. Advances in screen resolution, storage capabilities, computational power and software development have made them an appealing candidate for a new generation of spatially aware mobile computing.

While it is tempting to fill the environment with smart sensors and digital content, we believe that one must respect the tangibility of traditional media. Thus, we have chosen a design that invisibly enhances current information sources, leaving them fully backwards compatible. Our supporting infrastructure is designed to remain in the background and be unobtrusive to traditional means of interaction.

Our LightSense system, shown in Figure 1, supports a variety of interaction techniques that has been previously described for spatially aware devices [2, 5, 16]. Our contributions are techniques for outside-in tracking of unmodified consumer cell phones and an exploration of associated interaction techniques and applications. Our motivation has not been to create a self-contained tracking system in a specially produced portable device, but to provide a readily available framework that can empower a huge number of existing devices. Our framework allows a user to wirelessly tap into an existing tracking system for access to tracking data and computing infrastructure at a location. Absolute or relative tracking data provides the means for intuitive interaction in the physical context with data in the system and on the device.

In Section 2, we discuss related work, followed by a presentation of our system and details on our sensing methods in Section 3. Section 4 describes some of the applications that we have implemented within the LightSense framework and finally we provide conclusions and future work in Section 5 and 6.

## 2 Related Work

The research on handheld devices as interactive displays in ubiquitous computing environments started out with the concepts of spatially aware displays. While the main technical problems with the devices (e.g., weight, computational power, etc.) have been addressed, issues in sensing and tracking have begun to receive attention in recent work.

### 2.1. Spatially aware handheld displays

The vision of handheld spatially aware displays dates back to the Chameleon [5] and NaviCam [14] systems. The Chameleon used a 4" video display tracked with a 6DOF electromagnetic sensor to show different views of imagery depending on its position and orientation. The NaviCam system extends this concept with a camera that tracks markers in the environment, allowing rendered graphics to be overlaid on the imagery, resulting in a video see-through augmented reality display. To address fatigue, Fitzmaurice [4] proposed the use of a mechanically tracked boom, which was later realized in the Boom Chameleon [15]. Yee explores pen-based interaction techniques on a lightweight PDA [16], but the proof-of-concept setups either require a backpack with an electromagnetic tracker plugged into an outlet, or a device attached with strings to a table. The focus was however to show a number of interesting interaction concepts, some of which have inspired this work.

### 2.2. Sensor-based techniques

Other approaches use different sensors attached to the device such that a richer set of data can be captured. Rekimoto proposes the use of a tilt sensor [13], Marsden and Tip mount a GyroMouse on the back of a PDA [10], whereas Hinckley et al. integrate proximity sensing, touch sensitivity and tilt-sensing accelerometers with a microcontroller on the device [9]. Reilly et al. use an RFID reader attached to a PDA and RFID tags placed on the back of physical maps [12]. On-board accelerometers are already available in certain cell phones (e.g., Samsung SCH-310), and GPS units provide absolute data on other devices. While relative data from inertial sensing is valuable for interaction, it is not sufficient to enable a spatially aware display, and the GPS is too coarse to be used in an interactive application as well as limited to outdoor use. However, devices with embedded hybrid sensing approaches indicate a promising development, which will lead to more engaging user interaction.

### 2.3. Inside-out camera-based tracking

Recent advances in processor performance and the abundance of integrated cameras on consumer cell phones has inspired researchers to achieve self-contained 6DOF tracking with the on-board camera — something that earlier required the use of a PDA.

Moehring et al. [11] present a custom marker tracking algorithm for Symbian phones, while Henrysson and Ollila [8] discuss a port of the ARToolkit to the same platform. PhoneGuide [6] explores object recognition as an alternative to computationally expensive marker tracking. TinyMotion [17] illustrate a robust optical flow implementation while Ballagas et al. [1] use optical flow and marker detection algorithms in their interaction.

Camera-based solutions provide cheap, robust real-time 3D tracking, but are not ideal in a ubiquitous computing environment. Markers clutter the physical space and can occlude important features in the real world. The camera-based tracking also forces the user to keep the phone at a distance from the object of interest, complicating interaction near or on surfaces. While supporting the Magic Lens metaphor [3], it might be somewhat cumbersome to use the device effectively for direct interaction with real-world objects.

## 3 THE LIGHTSENSE SYSTEM

Increasingly popular ultra bright LEDs are now built into many cell phones to help improve pictures taken in dimly lit environments as an alternative to using a real flash. The LED also allows the phone to be used as a small flashlight. We however make use of this light source on our devices (Sony Ericsson K750i and W810i) as an easily identifiable active marker, which we can track the position of with different light sensing methods in our outside-in tracking framework. We currently provide two methods for tracking the light source on the cell phone.

### 3.1. Continuous tracking for tables and walls

For interaction on surfaces where continuous tracking is required, we employ a system that consists of a laptop (Toshiba M200, Pentium M 1.80 GHz, 1.25 GB RAM), an attached firewire camera and a diffusing surface, such as a glass table, as shown in Figure 1b.

The camera is placed behind the surface such that it monitors the planar interaction space. The further behind the surface the camera is placed, the larger the interaction space it can monitor, as long as it is within the focal capabilities of the camera. The diffusion surface eliminates indirect and ambient light, while creating bright spots for focused light near the surface. As the distance between the focused light source and the surface is increased, the bright spot will grow in size (and diminish in brightness), which provides a useful depth cue. The light detection software was implemented in C++ using our modular filter framework that employs the open source computer vision library OpenCV (http://sourceforge.net/projects/opencvlibrary/). It consists of a set of filters that perform image processing operations on the input images (See Figure 2a). The first step is to filter out all image data below a certain intensity threshold, followed by dilation and erosion filters to eliminate small noisy regions. Subsequently, a fitting algorithm is applied to find the largest ellipse that can enclose all the remaining bright pixels.
The center of this ellipse corresponds to the position of the light source. For depth, we count the number of bright pixels in the image and normalize the value to fit in a ten-grade scale. This tracking approach provides us with continuous 2D tracking on the surface and discrete 1D sensing orthogonal to the surface. The device is manually calibrated and the parameters for each filter

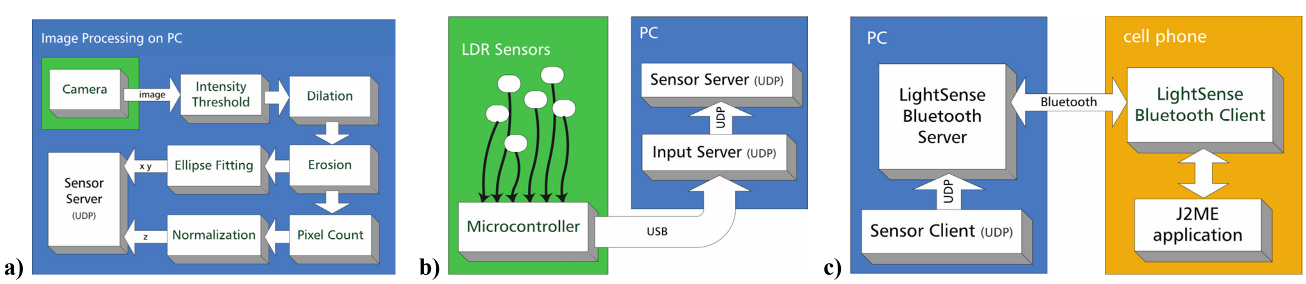


**Figure 2. LightSense system architecture. a) Continuous tracking for tables and walls. Image processing filters extract the centroid of the brightest area in the image. The distance to the surface is estimated using the size of the area, since the size of the blur spot will increase with distance (as long as the device is within range). The coordinates are streamed over UDP. b) Ubiquitous Sensing Tags. A microcontroller performs A/D conversions of data from the Light Dependent Resistors and transmits the values over USB to a PC. The position is streamed over UDP. c) A separate Java application handles the Bluetooth communication with the cell phone. It listens to sensor data on a local UDP port (could be from camera or LDR based data) and relays it to the handheld device, where it can be used in various applications. The J2ME applications on the device were developed using Mobile Processing (http://mobile.processing.org/).**

can be tweaked at runtime for optimal performance in the current conditions.

The data is made available in the Sensor Server and streamed over UDP to the LightSense Bluetooth Server, which handles all communication with the phone (See Figure 2a and c).

### 3.2. Ubiquitous Sensing Tags

In applications that do not benefit from continuous tracking we provide an approach based on hotspot-tracking. Reilly et al. [12] use RFID in their Marked-up maps, but this is a less ideal solution given the large size of the tags (20-50 mm), their discrete nature, inconsistencies when several tags overlap and the sensitivity to surrounding tags. We employ Light Dependent Resistors (LDRs) that are interfaced with a PIC microcontroller (Microchip PIC18F4550). These are placed in strategic positions under a diffused surface and as they are excited by strong light, the microcontroller reports the voltage level of the triggered sensors to the handheld device. The advantage is of course that no image processing is involved, which greatly reduces hardware requirements and overall system costs and complexity.

In our current setup, the microcontroller sends the data to a PC over USB, where the PC relays the data to the handheld device over Bluetooth (See Figure 2b and c). Adding a Bluetooth module to the microcontroller would however be straightforward and allow the sensing package to communicate directly with the device, providing embedded tracking at a low cost. We envision an environment where multiple cheap LDRs are bundled with a small microcontroller, a battery and a one-way radio communication module that is capable of reporting to a central microcontroller, which in turn relays the data over Bluetooth to the handheld.

Cheap tracking matrices could thus be built and embedded in the environment as part of a ubiquitous light tracking infrastructure.

## 4 Applications

We have begun to experiment with a number of applications in our LightSense framework, three of which we describe here.

### 4.1. A Multimodal Navigation Guide

We implemented an active map application that employs the PC and camera-based setup for the continuous tracking of the cell phone.

#### 4.1.1. Focus + context

A physical map of the Stockholm subway is augmented with digital content in a focus + context fashion [2]. The system identifies the closest station to the device and provides the user with a local map for the station's surroundings and the ability to zoom in and out by varying the distance of the phone to the physical map, similarly to the technique described by Yee [16] (See Figure 3).

#### 4.1.2. Audio and vibrotactile feedback for vision impaired users

We also see the potential in assisting people that typically cannot access physical information in public spaces, such as vision impaired users. The cell phone is used as a scanning device, where users can follow subway lines and receive vibrotactile feedback from the phone as it is slowly moved over the map. If users are about to loose contact with the subway line they are exploring, the device will start vibrating, indicating that the user should move back to get on the right track again. Additionally, as the phone is over a station, the station's name will be played back privately through the phone's earphones. Non-blind users can utilize the device as a personalized magnification glass, enlarging and rendering the station names such that they are easier to read. For color blind users, color names could be spelled out. Users with cognitive disabilities could have preprogrammed settings indicating a station as "home" and another as "park", for instance.

**Figure 3. Zooming in the navigation guide. As the user positions the device on top of a subway station on the map, a local road map becomes available and the station's name can be played back through the earphones. The user can interact with the map by lifting the device away from the surface to zoom out in the region, or by lowering it to zoom in.**

### 4.2. Peephole Interaction

This visualization tool for interaction with large datasets on small displays uses techniques introduced by Fitzmaurice [5] and later adapted to handheld devices by Yee [16]. The device acts as a viewport into a large image, for example and as the user moves the display around, different portions of the larger image will be seen (See Figure 4).

### 4.3. Ubiquitous Guides

We envision that many LightSense units could be placed in the environment, augmenting existing information that is currently communicated through traditional media, such as posters and signage. The wall-based LightSense system is a lightweight installation that could be suitable in many ubiquitous deployments. The front is embedded with LDR sensors in strategic places such that events are triggered when they are struck by the light from the phone LED. Figure 1a shows an application where detailed information (such as the number of inhabitants, size, year it was established, and a brief historic account) along with a picture is provided about each borough as the phone is moved above a printed satellite map of Stockholm.

## 5 Future Work

LightSense illustrates that commercial off-the-shelf cell phones can already be used as one of the key components in a ubiquitous computing environment. There are however a few limitations that we need to address.

### 5.1. Improved outside-in tracking

Tracking in the visible spectrum instead of, e.g., the IR spectrum leaves fewer options for filtering out undesired light and has more inconsistencies as the lighting changes over the day. This does not pose a significant problem for typical office environments, but could certainly do so in public spaces. Enclosing the camera or LDRs with opaque material and sealing with a diffused screen is


**Figure 4. Two superimposed images illustrate the implemented "Peephole" interaction technique. A small tracked display acts as a viewport into a larger virtual display and by moving the device the user is able to view different portions of the virtual display.**

required to ensure that unwanted light is kept out. Indirect scattering onto the diffused surface has little impact and a good design ensures robust results. The next step is to test the performance of our prototypes under more complex lighting conditions to assess the reliability. Additionally, more sophisticated computer vision analysis and tracking, such as employing the Kalman Filter, would improve the results in the camera-based version of our system.

We are currently limited to 2D tracking with discrete depth levels. The system could be extended with multiple cameras where the varying light intensity among the image sensors would give an indication of the current light direction and orientation of the light source. Orientation in the plane could be addressed through shape analysis, since our device has an elongated light source (consisting of two LEDs), but this becomes more complicated with increased distance to the surface.

### 5.2. Inside-out tracking

We are planning to explore egocentric light tracking with the phone's camera for tilt detection. The bright LED is clearly visible on a reflecting surface, such as glass or acrylic, and the reflection's position in the image gives an indication of the phone's orientation with respect to the surface. This is an interesting alternative approach to the printed marker pattern used by Hachet et al. [7] or to using a sensor-based tilt sensor [9, 13].

### 5.3. Tracking of multiple devices

Time division multiplexing of different devices would allow the LightSense system to uniquely identify multiple devices. This requires that the application can control the LED, which is currently not supported on all devices. With multiple targets, it becomes even more important that the tracking is robust and consistent.

### 5.4. Dynamic interaction surface

Our initial experiments with rear-projected imagery in the system confirm that there is no conflict between the rear-projected image and the LED tracking. Dynamically updated graphics could be especially useful in a collaborative setting, where each device acts as a personal display, with the large interaction surface visualizing public and shared data. A collaborative setting also presents interesting challenges such as display tiling and tangible interaction on the surface. Rear-projected imagery would also allow us to experiment in more depth with an interactive focus + context concept.

### 5.5. Power consumption

The use of a bright LED, Bluetooth connection, CPU-intensive graphics rendering and sporadic vibrotactile and audio feedback has so far not been an issue in terms of power consumption. In an informal experiment we experienced that the battery level decreased approximately 20% per hour. The phone was receiving tracking data over Bluetooth, rendering map updates and had the LED on, but did not utilize vibrotactile or auditory feedback during the test. Following up on these results with an in-depth study would be useful.

## 6 CONCLUSIONS

We have presented two solutions that employ commercially available components to enable spatial awareness in cell-phones, without any modifications to the device. There are several advantages to our approach. Most importantly, the tracking can be performed in public spaces, behind maps and bulletin boards or under tables and counters. This does not only hide the technology from users, but also protects the equipment and ensures robustness due to the controlled conditions. The platform-independent tracking allows any application on a Bluetooth-enabled phone with built-in LEDs to take advantage of the ubiquitous sensing. The device is at the same time freed from the often computationally intensive processing that self-tracked devices must address. In contrast to most other tracking approaches, our sensing works both on and above the surface of interest. This, for instance, allows the user to comfortably put down the device to avoid fatigue or to use it in a collaborative setting with multiple people. Since our tracking uses no resources on the device, it can coexist with other tracking techniques, such as marker-based inside-out camera tracking or optical flow algorithms [1, 7, 8, 11, 17], leveraging the best of multiple methods in a hybrid approach.

## ACKNOWLEDGEMENTS

We thank Sony Ericsson for providing the cell phones used in the project.

## REFERENCES

[1] Ballagas, R., Rohs, M., Sheridan, J., and Borchers, J. The Smart Phone: A Ubiquitous Input Device. *IEEE Pervasive Computing*, 5(1):70-77, 2006.

[2] Baudisch, P., Good, N., and Stewart, P. Focus plus context screens: combining display technology with visualization techniques. *UIST '01*, pp. 31-40, 2001.

[3] Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose, T. D. Toolglass and magic lenses: the see-through interface. *SIGGRAPH '93*, pp. 73-80, 1993.

[4] Fitzmaurice, G. W., Zhai, S., and M. Chignell. Virtual Reality for Palmtop Computers. *ACM TOIS* 1993, pp. 197-218, 1993.

[5] Fitzmaurice, G. W. 1993. Situated information spaces and spatially aware palmtop computers. *Commun. ACM,* 36, 7, pp. 39-49, 1993.

[6] Föckler, P., Zeidler, T., Brombach, B., Bruns, E., and Bimber, O. PhoneGuide: Museum Guidance Supported by On-Device Object Recognition on Mobile Phones. *MUM '05*, pp. 3-10, 2005.

[7] Hachet, M., Pouderoux, J., and Guitton, P. A camera-based interface for interaction with mobile handheld computers. *SI3D '05*, pp. 65-72, 2005.

[8] Henrysson, A. and Ollila, M. UMAR: Ubiquitous Mobile Augmented Reality. *MUM '04*, pp. 41-45, 2004.

[9] Hinckley, K., Pierce, J., Sinclair, M., and Horvitz, E. Sensing techniques for mobile interaction. *UIST '00*, pp. 91-100, 2000.

[10] Marsden, G. and Tip, N. Navigation control for mobile virtual environments. *MobileHCI '05*, pp. 279-282, 2005.

[11] Moehring, M., Lessig, C. and Bimber, O. Video See-Through AR on Consumer Cell Phones. *ISMAR '04*, pp. 252-253, 2004.

[12] Reilly, D., Rodgers, M., Argue, R., Nunes, M., and Inkpen, K. Marked-up maps: combining paper maps and electronic information resources. *Personal Ubiquitous Comput*., 10, 4, 215-226, 2006.

[13] Rekimoto, J. Tilting operations for small screen interfaces. *UIST '96*, pp. 167-168, 1996.

[14] Rekimoto, J. and Nagao, K. 1995. The world through the computer: computer augmented interaction with real world environments. *UIST '95*, pp. 29-36, 1995.

[15] Tsang, M., Fitzmaurice, G. W., Kurtenbach, G., Khan, A., and Buxton, B. Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. *UIST '02*, pp. 111-120, 2002.

[16] Yee, K. Peephole displays: pen interaction on spatially aware handheld computers. *CHI '03*, pp. 1-8, 2003.

[17] Wang, J. and Canny, J. TinyMotion: camera phone based interaction methods. *CHI '06 Extended Abstracts*, 339-344, 2006.

[18] Weiser, M. Some computer science issues in ubiquitous computing. *Commun. ACM,* 36, 7, 75-84, 1993.