A 2-Way Laser-Assisted Selection Scheme for Handhelds in a Physical Environment

Shwetak N. Patel and Gregory D. Abowd

College of Computing & GVU Center
Georgia Institute of Technology
801 Atlantic Drive, Atlanta, GA 30332-0280, USA
{shwetak,abowd}@cc.gatech.edu

Abstract. We present a 2-way selection method to select objects in a physical environment with a novel feedback and transfer of control mechanism. A modulated laser pointer signal sent from a handheld device triggers a photosensitive tag placed in the environment. The tag responds via a standard wireless channel directly to the handheld with information regarding an object it represents. We describe a prototype implementation for a Motorola iDEN i95cl cell phone, discuss the interaction challenges and application possibilities for this physical world selection that extends a common handheld device. We also compare this solution to related attempts in the literature.

1 Introduction

It is becoming increasingly common to carry a handheld device (a PDA or a mobile phone), with nearly always-on network connectivity and significant computational capabilities. What is not so common is the ability for these devices to facilitate interactions with the physical world. Take, for example, the development of universal remote controls for use at home. Several commercial and research projects have considered the challenge of reducing control of a large variety of home devices down to one universal platform (see, for example the recent work on the Personal Universal Controller [8], or the Gesture Pendant [13]). But in an environment with many controllable devices, it is not clear the best way to select which device should be controlled. Pointing from a distance, naturally supported by a laser pointer, seems like a good solution to this problem. We are motivated, therefore, to explore direct and natural interactions with devices in the physical world mediated through a handheld device.

A number of researchers have explored ways to tag the physical world, using printed barcodes or 2-dimensional glyphs, RF ID, or active beacons, in order to connect the physical and electronic worlds. We are particularly interested in applications of lasers because they provide a simple means of visual feedback as well as at-a-distance interaction. Furthermore, unlike any other physical world selection technique in the literature, we see the advantages of creating 2-way communication between the object in the physical world and the laser-augmented handheld. The challenge we faced was to create a practical, 2-way selection technique using a laser mounted on a conventional handheld device that provides a two-way interaction. The

A.K. Dey et al. (Eds.): UbiComp 2003, LNCS 2864, pp. 200–207, 2003.

solution described in this paper demonstrates an augmented handheld device that communicates its identity via a modulating laser signal to active tags in the environment. These tags, logically linked to objects in the environment, can then communicate back to the handheld, establishing a two-way link. This solution allows for selection of physical objects in the environment that can then be further controlled or queried by the handheld device.

Integrating this 2-way laser pointer into a mobile phone can also enable many possibilities outside of the home. The active tags can be placed on road signs and billboards, using large tags connected to some wireless telephony service. When the billboard is spotted, the user selects it in order to receive additional information to the handset. While traveling by car, the natural visual feedback of the laser pointer might not work, so we developed a slower response vibration feedback. At an airport, these active sensor tags can be placed in the logos of the airline companies. As you walk through the airport, you can simply identify one of these tags and the gate and flight information is directly sent to the handset.

Overview

Having motivated the potential applications for this 2-way selection technique, we will next present a brief overview of existing techniques for selecting objects in the physical world. We describe the design of our laser-assisted selection technique as implemented on a commercial mobile phone using special-purpose active tags embedded in the environment. We then discuss some of the interaction challenges that influenced our design.

2 Related Work

Several researchers have explored laser pointer interaction recently, both for interaction at a distance with large displays [4, 7, 9] and also for the selection task studied in this paper [6, 12, 14]. Selection of physical objects in an environment has also been explored for various augmented reality tasks (e.g., the NaviCam system [11]). For the selection task, we identify three different approaches:

Static Tagging: A static label (e.g., a barcode or 2-dimensional glyph) is placed in the environment and read or scanned by some form of reader device [5, 10]. Barcode solutions are limited to a distance of about 1 meter [6], whereas camera-based solutions are limited only by the camera resolution and perception techniques used to decipher the glyphs. The scanning device must be connected to some service that converts the scanned information into a device identifier. This interaction is one-way and the same information is provided to every scanning device. This selection mechanism works well when it is suitable to require short-range interaction.

Camera-Tracked Laser Interaction: A popular laser pointer interaction scheme is to use a camera focused on a region of a wall or object where a laser spot may appear [1, 4, 7, 9, 14]. Simple computer vision techniques locate the red laser dot and follow it around the interaction region. Such a scheme is appealing for meetings or presentations, where you can interact with a display at a distance by simply pointing at it with an ordinary laser pointer [1, 4, 7, 9] and the XWand system demonstrates how it can be used with a collection of other sensors to support selection and interaction of devices through a special-purpose interaction device. An extension of XWand, called the World Cursor, removes the vision requirement by using the XWand to steer a remotely controlled laser pointer around a room [15]. The remotely controlled laser pointer has a model of where it is pointing in 3-space and has sufficient geometric information to know where its red laser dot is pointing. While these camera-based tracking solutions provide very flexible ways to point and select objects within the environment, they require a lot in terms of camera infrastructure and detailed geometric information and will not work well in large environments with much movement of objects.

Active Tagging: Our laser pointer system, Matthias Ringwald's Spontaneous Interaction, and MIT's FindIT Flashlight use active tags that respond to an incident laser [6, 12]. A modulated laser signal encodes information that is received by the tags and decoded. The tag is active in the sense that it can respond to the initiator of the interaction with any appropriate response. The response by the FindIT Flashlight is an indicator light to notify the user that the desired object has been found. For our system, the interaction with the handheld device is 2-way because the tag can use a number of wireless mechanisms to send a data response. Ringwald's Spontaneous Interaction makes a similar attempt by sending back web content from selected tags through 802.11b Wi-Fi. While active tags are an additional expense and may require separate power, they could be placed in a variety of locations or embedded in commercial appliances. No further knowledge of the environment is required, making them a more practical solution for the selection task compared with the camera-based solution discussed above. The HP CoolTown beacons are small hardware devices distributed throughout an environment whose function is to wirelessly broadcast device references (URLs) [3]. CoolTown beacons use IRDA to broadcast device references, which make it an attractive solution for IR ready handhelds. However, IRDA lacks the precise identification and natural visual feedback possible with lasers and is limited to only a few meters in range.

3 How Laser Selection Works

Our laser system is based on a PC-to-PC laser communication link first published by GKDesign [2]. We built prototypes for a Compaq IPAQ and a Motorola iDEN i95cl cell phone. Each handheld is assumed to have wireless data access (802.11b or Bluetooth for the IPAQ and the Nextel data service for the i95cl handset), providing an IP address for direct communication. The figures and description in this paper describe the cell phone prototype because the slim form factor and always on data service motivates a variety of indoor and outdoor applications.

3.1 The Instrumented Handheld

The laser is integrated into the handheld device by using a 3-5 mW diode laser module mounted on the antenna of the i95cl handset (see Figure 1). The RS-232 line from the handset is run through a MAX232A IC line driver and an open collector buffer to allow serial-controllable modulation of the laser beam. Using the J2ME environment on the handset, it is straightforward to encode outgoing messages, and we chose to encode the IP address of the handset to facilitate routing of return messages. Since the communication with tags is asynchronous, the message from the handset is padded with start and stop codes. When the laser button is pressed (using the lower button on the handset shown in Figure 1a), the message and its padding, or the message frame, is continuously transmitted. The user only has to keep the laser shinning on a sensor tag for the length of one complete frame. At a baud rate of 9600 bps, a message frame of 128 bits would take 13 ms to be detected. We found that speeds of 9600 bps or lower required only a simple parity check for error detection; more sophisticated error detection and correction schemes (e.g., CRC or Hamming codes) would be needed at higher transmission rates or for sensitive messages. A more sophisticated message-encoding scheme could be constructed by directly modulating the laser with the Motorola chipset. This would provide larger messages and faster transmission rates. The handset's 1400 mAh lithium ion battery pack powers the laser module. The circuit draws about 35 mA when the laser is activated. Moderate use of the laser does not significantly reduce the battery life of the handset.

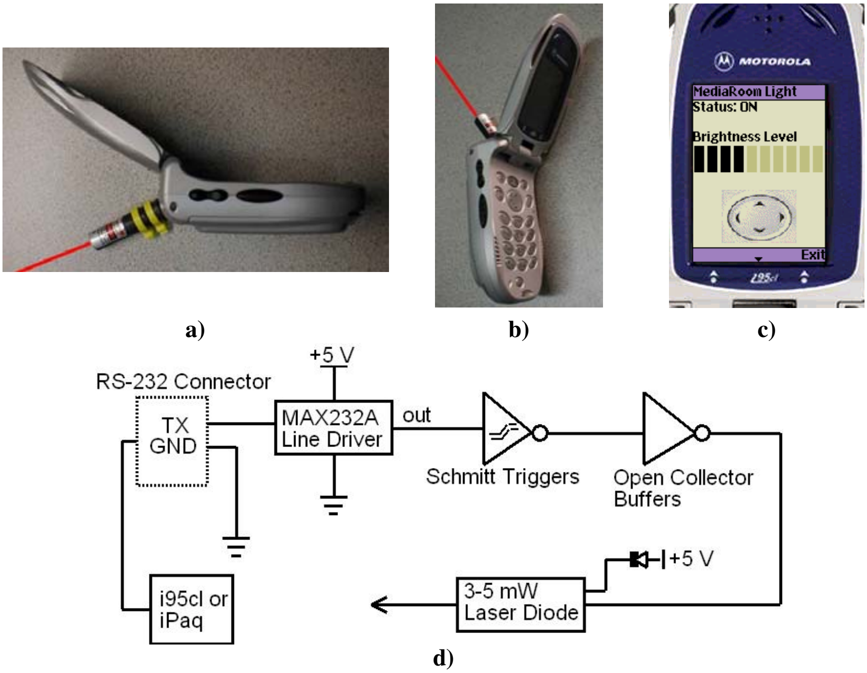

Fig. 1. a-b) The instrumented cell phone handset, a Motorola iDEN i95cl. The antenna is angled slightly to accommodate reading the screen while pointing the laser. c) A simulator screen for the phone showing a simple control screen for an X10-controlled light switch identified through laser interaction. d) A schematic for the laser controller.

3.2 The Active Tag

The sensor end of the system consists of a bed of phototransistors, a MAX232A IC line driver, and Schmitt triggers. The actual light sensing part of the tag is a bed of NPN IR phototransistors. The signal then runs through Schmitt triggers to square up the wave before it is fed into the MAX232A line driver. Out of the line driver comes the RS-232 signal, which is fed into a microcontroller (we used a PC in our prototype). The microcontroller produces the response back to the handset.

Since the message sent by the cell phone is an IP address, the sensor knows exactly where to send the feedback information. The route of the response can vary depending on the connectivity of the sensor tag. In one prototype, we directly connected the tag to the wired LAN. In another prototype, the tag uses X10 to transmit to a basestation (an X10 server) connected to the LAN. Routing to the handset is done using the Nextel data service (or 802.11b or Bluetooth in the case of the IPAQ).

We chose phototransistors as the sensor mechanism because they are less affected by ambient light. Phototransistors are available with a variety of spectral ranges. The ones used in our sensor tags have a range of 600 nm to 1000 nm. The response range is broad enough for a typical red laser (670 nm) and still able to ignore some of the lower wavelengths.

We found there is just enough red in ambient light, especially sunlight, making a red-pass filter on the sensor ineffective at times. For even more immunity to ambient light, a black epoxy 800 nm to 1000 nm phototransistor could be used, requiring an IR laser. Since IR lasers are invisible to the naked eye, we lose the natural visual feedback that the red laser provides. Coupling a small red laser with the IR laser produces both a higher-speed transmitter and a visual feedback loop. However, IR lasers present a greater eye safety concern than red lasers.

The reason for a bed of phototransistors is to allow for a larger target. Our prototype sensor tag is enclosed in a cone structure (see Figure 2a). The inside of the structure is padded with a reflective material. When the light beam hits anywhere inside the structure the light rays are reflected in many directions. The front part of the cone is covered with a defocusing material; when the beam hits the front of the structure, the light beam is refracted in many directions inside of the cone. The bed of phototransistors resides at the back of the cone structure, increasing the likelihood of a ray hitting at least one phototransistor and reducing the need for operator precision in aiming the laser. Another advantage of using a cone-like or cylindrical structure is that very precise optics can be added so that only one phototransistor is needed. In this case a beam coming in at any point can be refracted straight to the sensor. This approach is the opposite of the defocusing used in the FindIt flashlight [6]. By defocusing at the target instead of the source, we maintain the desirable features of a laser pointer, namely the long range and natural visual feedback.

The cone-like tag is useful for long-range selection. For short to intermediate ranges, a dense bed of phototransistors is suitable and can be embedded within the objects being selected. For example, figure 2b shows a prototype of a tagged light switch.

The size of the sensor constructed depends on the placement in the environment and desired precision for the laser. More accurately, inspired by the data reported by Myers *et al.*, we calculated that the average deviation resulting from laser "wiggle" is .0025 times the distance from the target [7]. Deviation here means the amount the laser dot wiggles while a user tries to hold it steady for 3 seconds. This means a tag diameter should be at least .005 times the expected distance to accommodate easy selection. Our prototype sensor tags are 5 inches in diameter, which is large enough to interact from within a very large open space (approx. 83 feet).

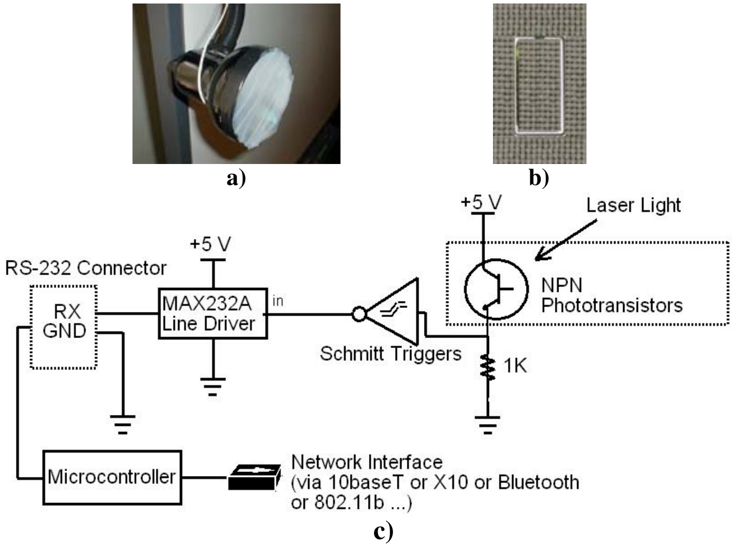

Fig. 2. a) A cone-shaped light-sensitive tag. b) A layer of phototransistors replacing a wall plate for a light switch shows a different form factor for the active tag. c) The schematic for the active tag.

4 Analyzing Handset-Tag Interaction

The interaction between handset and tag is relatively simple, with some interesting interaction challenges. The important human factors are acquisition time to locate a tag with the laser pointer, how long the user can comfortably keep the beam steady on the sensor tag to ensure a hit, and what feedback tells the user that the target is hit. Myers *et al.* showed that it takes approximately 1 second to move a laser beam to a target position [7]. This is perfectly reasonable for our scheme, since a likely alternative to pointing at a target for selection would be to look up the target device using a list on the handset, which would likely take longer than 1 second. The same study found that a laser mounted on a PDA is more stable than just holding a plain laser pointer, because the PDA provides a more effective grip because of its large size. In our experience, both the i95cl and iPAQ offer very stable control. The i95cl fits well in the palm of the hand and is the better of the two because of its intermediate size and ergonomic shape. The laser activation button resides on the left side of the phone, allowing thumb operation when held in the left hand or finger operation when held in the right hand. This allows for stable control as the laser is activated. We also angled the handset antenna to provide even more comfort and stability, allowing laser pointing and easy screen reading.

Since the dwell time for a tag to "read" the modulated laser signal is pretty low (13 ms), there are two ways a tag may be accidentally triggered. First, we know that a user cannot effectively predict where a laser pointer will hit when initially activated [7, 9], so it may accidentally hit the wrong target initially. Second, tags may be placed close to each other, as might happen in a cluster of home entertainment devices, causing inadvertent selection. One solution to this problem is to introduce a two-stage scan and select process using a two-position switch on the handset. When the button is depressed half way, the laser is turned on without any messages being sent and natural visual feedback is used to aim the laser over the appropriate target. Upon proper targeting, the handset button is depressed fully, sending out the modulated message frames. Another solution would be to deliberately slow down the active tags by requiring receipt of multiple message frames before being activated.

Another potential challenge is feedback, or knowing when a tag has made a reading. One solution, suggested by the FindIt flashlight, is to provide a LED on the tag itself that can be illuminated when the microcontroller detects a valid read [6]. Another solution, which we implemented, is to signal the handset over the air when a read is detected. The handset can respond with visual or vibration feedback. In our various prototypes, the latency for this handset feedback ranged from 700-1000 ms. The faster feedback occurred with the tag connected to the network via 802.11b wireless and the cell phone receiving feedback via its cellular network. The slower feedback was a result of using an X10 connection between the tag and the network. For many indoor applications, this latency is probably too great to be useful. However, in outdoor applications, where ambient sunlight negatively impacts the natural visible feedback of the laser, the vibration scheme is more useful.

5 Conclusions

We presented a laser-assisted 2-way selection technique for identifying and interacting with objects in the physical world. This technique uses active tags that can detect modulated signals from a handheld-mounted laser pointer and respond via some wireless route back to the handheld device. We demonstrated this prototype based on a popular mobile phone handset, the Motorola iDEN i95cl. The selection technique was originally designed for use in a home-based universal remote control application, but the use of a mobile phone with constant network connectivity opens up possibilities for this technique that would work outdoors as well as indoor.

Acknowledgements

The authors thank Motorola, and in particular Joe Dvorak, for the donation of the iDEN handsets for this research. We thank Matthias Ringwald who worked on an earlier version of laser-tracking for the remote control application. This work is sponsored in part by National Science Foundation (ITR grant 0121661) and the industrial sponsors of the Aware Home Research Initiative at Georgia Tech.

References

1. Cooperstock, J., Fels, S., Buxton, W. and Smith, K.: Reactive Environments: Throwing Away Your Keyboard and Mouse. *Communications of the ACM*, 40(7), pp. 65-73, September 1997.
2. GKDesign Engineering: RS-232 Laser Transceiver. *Electronics Australia*, pp. 56-61, October 1997.
3. HP Labs: The CoolTown Project. http://www.cooltown.com/research.
4. Kirstein, C and Müller, H.: Interaction with a Projection Screen using a Camera-tracked Laser Pointer. *Proceedings of the International Conference on Multimedia Modeling.* IEEE Computer Society Press, pp. 191-192, 1998.
5. Ljungstrand, P., Redstrm, J. and Holmquist, L.E.: WebStickers: Using Physical Tokens to Access, Manage and Share Bookmarks to the Web. *In Proceeding of Designing Augmented Reality Environments* (DARE 2000), pp. 23-31, ACM Press, 2000
6. Ma , H. and Paradiso, Joseph A.: The FindIT Flashlight: Responsive Tagging Based on Optically Triggered Microprocessor Wakeup. *Proceedings of Ubicomp 2002*, Springer-Verlag Lecture Notes in Computer Science, Volume 2498, pp. 160-167, 2002.
7. Myers, B.A., Bhatnagar, R., Nichols, J. Peck, C.H., Kong, D., Miller, R. and Long, A.C.: Interacting at a Distance: Measuring the Performance of Laser Pointers and Other Devices. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2002*). ACM Press, pp. 33-40, 2001, Minneapolis, Minnesota.
8. Nichols, J., Myers, B.A., Higgins, M., Hughes, J., Harris, T.K., Rosenfeld, R. and Pignol, M.: Generating remote control interfaces for complex appliances. *Proceedings of the 15th annual ACM symposium on User Interface Software and Technology (UIST 2002)*, ACM Press, pp. 161-170, Paris, France.
9. Olsen, D.R., Jr., Nielsen, T.: Laser pointer interaction. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2001*). ACM Press, pp. 17-22, March 2001, Seattle, Washington.
10. Rekimoto J. and Ayatsuka Y.: CyberCode: Designing Augmented Reality Environments with Visual Tags*, Proceedings of DARE 2000*, 2000.
11. Rekimoto, J and Katashi, N.: The World through the Computer: Computer Augmented Interaction with Real World Environments, *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '95),* ACM Press, pp.29-36, Pittsburgh, PA.
12. Ringwald, M.: Spontaneous Interaction with Everyday Devices Using a PDA. *Presented at the Supporting Spontaneous Interaction in Ubiquitous Computing Settings Workshop* (UBICOMP 2002). http://www.inf.ethz.ch/~mringwal/publ/ringwald-interaction.pdf. Göteborg, Sweden, September 2002.
13. Starner, T., Auxier, J., Ashbrook, D. and Gandy. M.: The Gesture Pendant A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. *The IEEE International Symposium on Wearable Computing (ISWC 2000)*, IEEE Computer Society Press, pp. 87-94, Atlanta, GA.
14. Wilson, A. and Shafer, S.: XWand: UI for intelligent spaces. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2003*). ACM Press, pp. 545-552, 2003, Ft. Lauderdale, Florida.
15. Wilson, A.: Pointing in Intelligent Environments with the WorldCursor. *Proceedings of Interact 2003, the Ninth IFIP International Conference on Human-Computer Interaction.* Zurich, Switzerland. To appear, September 2003.