SuperVision: Detection and Control of In-Sight and Out of Sight Objects In a Smart Home

Submitted in fulfillment of the requirements of BITS C421T Thesis

By
Sarthak Ghosh
2010A3TS211H

Under the guidance of
Dr. Gilles Bailly
Dr. Eric Lecolinet
Mr. Chetan Kumar

Department of Electrical And Electronics
Birla Institute of Technology and Science-Pilani
Hyderabad, India

7 May 2014

Certificate

This is to certify that the thesis entitled "The Control of Smart Homes and Offices With Pointing Gestures " and submitted by Sarthak Ghosh, ID No: 2010A3TS211H, in fulfillment of the requirement of BITS C241T Thesis embodies the work done by him under my supervision.

Signature of Supervisor	Signature of CoSupervisor
Name :	Name:
Designation :	Designation
Date :	Date :

Abstract

Today, there are a multitude of smart objects which come with their specific functions. As these become a part of everyday life, the concept of a ”smart home” becomes more and more realizable. A smart home is a home in which all, or most of the appliances are connected to a common network, so that they can communicate with each other and with the user [3].The user can communicate with the installed appliances in many ways. It can be through voice, through remote controls, using a smart phone, or a laptop, or through gestures[3]. Although various interaction methods have been proposed in this regard, rich and bidirectional interaction, is very often not possible and can get quite complex at times. For example, lets say all objects on a network have their unique identification addresses, and every time you want to choose one, you have to remember the assigned address. This is just to select the object. Further interaction can be even more difficult, depending on the range of functionality that the chosen device provides. Thus it is quite important to find a solution to intuitively detect and control objects from a distance effectively.

Here, we investigate the possible methods of selecting objects from a distance. We realize, that it is worthwhile to go beyond the current trend and try to control devices that are not directly in front of the user; devices that do not fall in the field of view. The possible ways of displaying an interaction interface are also explored, this involves the usage of a pico-projector. In this relation we also present an interaction technique called ”SuperVision” that lets the user

1. Select an Object by pointing at it
2. Look through a virtual hole in the wall to visualize objects in adjacent rooms
3. Interact with Objects using a projected Interface on the wall.

We also introduce a novel control device which houses a pocket projector, and a motorized slider for interaction. We implement the proposed solution and also validate it by carrying out an user study, to check if the results actually fall in line with the expectations.

Contents

List of Figures

Chapter 1

Introduction

1.1 Context

Since the advent of the electronic age, there has been a constant motivation for controlling electronic items remotely. The earliest example of a remote control was one which used radio waves to control a moving vehicle, and was developed by Tesla in 1898[7]. The first remote controlled domestic devices were probably radio transmitters and televisions. Today in an average home, we can find several devices that can be controlled using remote controls, including DVD/VCR, Air Conditioner, Cable Box, Television,Game Console, Electronic curtains etc. Also, instead of remotes, people now are more interested in much richer interactions with environmental objects using gestures, touch or even voice. With such a need in existence, there are bound to be more objects that are "smart" and capable of diverse bidirectional interaction as we slowly move towards a world of "smart homes" and "smart offices".

A smart home as defined by the UK department of Trade and Industry, is "A dwelling incorporating a communications network that connects the key electrical appliances and services, and allows them to be remotely controlled, monitored or accessed"[5]. Smart homes use 'home automation' technologies to provide home owners with 'intelligent' feedback and information by monitoring many aspects of a home. As is apparent there are two main directions of research related to the field of smart homes,

1. Bringing more home automation so that users do not have to participate
2. Making user participation easy and intuitive, so that users enjoy to participate in controlling objects

In this project we concentrate solely on the end-user oriented direction of research. Thus control of objects situated far away is extremely important to achieve.

1.2 Problem

The sudden growth in remote controlled devices has been a direct result of device miniaturization, and an effort to decouple the user interface from the device itself. This in fact provides access to a larger set of functions than on-device interfaces can provide. Despite all the appealing advantages, this has led to a new problem called, "too many remotes". In fact, according to the Consumer Electronics Association, an average American home has four remotes. It can be quite complex when one has to use 5-6 remotes in a particular sequence just to get a home theater system to work right. Because of this proliferation of remote controls, universal remote controls that manage multiple devices are becoming increasingly popular.

Universal remote controls can usually control upto 5 devices in their line of sight. They hold on to the traditional design of a remote control. But since they have to control multiple devices, the remotes become more complex to use . Moreover, just like the traditional remotes, these new ones also do not help to solve the problem of division of focus from the physical world to the hand-held remotes. Users have to divide and toggle their attention between the controller and the controlled, which reduces the effectiveness of interaction. Moreover there is also a limitation in terms of the display of interface for interaction. This leaves ample design space free for researchers to think of new innovative interaction methods with a multitude of objects. Thus, new techniques have surfaced, including pointing with a projector, or with a "wand", or with a hand-held camera and interacting through gestures.

The traditional remote control was based on Infrared Radiation, but today scientists are working on new and innovative methods like interacting through lasers or projected light. There is also a constant attempt towards controlling objects in a dynamic environment, where users do not have to be aware of the electronic environment, and new objects can be freely inserted and deleted from the environment. In the strive towards achieving an intuitive way of selecting an object,(for example one simply points at an object in sight to select it) quite a few related problems can be experienced. There can be the problem of too many objects close together, the problem of having very small objects, and also the problem of occlusion of the target object by another. Existing solutions [13],[21],[24],[16],[19],[20], all mostly concentrate on controlling objects in the line of sight of the user. But it is not clear how can objects in the same room as the user, and also in adjacent rooms, be controlled by the same system of interaction.

1.3 Objective

The objective of this study is to attack the problem of smart home control from a Human Computer Interaction perspective. We try to pinpoint the issues faced, in achieving a technological solution for pointing, selecting, and controlling objects. We also explore the issue of displaying information or an interaction interface once an object has been identified. We identify scenarios where it might be more useful to be able to control objects which are not directly in the line of sight. Motivated by such scenarios, we propose a "SuperVision" which is a new interaction technique using hand-held projectors. SuperVision lets the user control objects not in the direct surroundings and also lets users display information on and around objects.

1.4 Contributions

The principal contributions of this project are

- A special remote controller which is continuously tracked and has the capability of projecting information and interfaces on and around objects.
- Implementation of the "Super Man Vision " metaphor which allows users to drill a virtual hole through a wall, to visualize and control objects behind it.
- Implementation of several applications which allows smart interaction with objects, and also allows supervision over babies.

1.5 Outline

In this report, we first present a detailed review and comparative study of the literature that has been studied and . We try to identify the common issues which are faced in designing systems in the concerned field of "controlling objects". We then present a new interaction method in the form of "SuperVision" and explain the designing of the technique. The concept of using the two metaphors of "spotlight" and "Super Man Vision" are thoroughly explained thereafter. Implementation of the concept is underway and the system set up, the hardware used and details of the implementation exercises are mentioned in the "Implementation" section. Future work includes further implementation of the concept and development of its applications.

Chapter 2

Literature Review

In the next few subsections of this chapter, we present a summary of the literature that was found to be extremely relevant. To start with, we were chiefly interested in finding out how to control objects from a distance, using any kind of technology. This lead us to quite a few research papers that discussed and measured the performance of using laser pointers for pointing at objects and communicating with them. Myers et al.[11] compared the performance of lasers as compared to other devices like a mouse or a smart board, for pointing at objects. Many projects have made innovative use of small projectors to directly display information on or around objects. RFIG Lamps[16] makes use of both RFID tags and projectors to exchange information with objects. The Xwand project from Microsoft[24] on the other hand uses pointng using a "wand " like structure for selection of objects. This involves optical tracking of the wand using two cameras which flash infrared light. The GesturePen[21] also uses infrared light for object selection. The project called PICOntrol[19] goes a little out of the way to use only a Pico-projector to both display information and pass control signals.

We also looked into a few other papers which deal with the use of "Invisible" tags for augmenting objects[12]. We also researched about the existing technologies which concentrate on display of information using information on and around objects. ILight[10] and illumiroon are the two major works that we must mention in this regard.

To summarize we can say that the project SuperVision is based on work in the fields of Pointing interaction, Video Interaction, Smart Room control and Looking through Obstacles. In the following sections we present a comparative study of various interaction techniques which utilize pointing, and then we present the related works in the other fields, namely Video

Interaction, Transparent Obstacles and Smart Room Control.

2.1 Four Step Interaction: A Comparative Study

As understood from the literature, the whole objective of interaction can be broken down into the following subparts.

1. Pointing technique/technology
2. Selection method
3. Disambiguation
4. Control

2.1.1 Pointing

While referring to an object in sight, people often prefer to point at its direction, rather than thinking of its particular name. It is because of this intuitive nature of pointing, scientists have very often integrated pointing in their interaction technique design. For example, Patel et al.[13] they integrated a laser pointer in a mobile phone, for two way communication with objects with active tags. Pointing is achieved by using a wand like structure in XWand [24]. Xwand relies on Machine Learning algorithms and tracking of the "wand" to calculate the most probable item being pointed at, by the user.

Myers et al [11] showed that a pen like structure is ergonomically well suited for pointing at things. GesturePen[21] thus retains the pen-like structure of a laser pointer but moves away from the laser technology. GesturePen makes use of IR transceivers to send and receive encoded messages to and from tagged objects. But Ir transceivers often suffer from out of range and lack of directionality issues. On the other hand PICOontrol[19] and RFIGLamps[16] both use pico-projectors to establish the pointing gesture and also identify objects. This makes the pointing action more user friendly because the user now has a direct visual feedback, because of the illuminated projected area. "SuperVision" also utilizes an off the shelf pocket projector to establish the pointing direction. But it differs from RFIG Lamps and PICOntrol in the fact that objects in the room are not tagged. We are also not passing information encoded in the projected area. We use the projector only for visualizing interfaces and interacting with them through gestures.

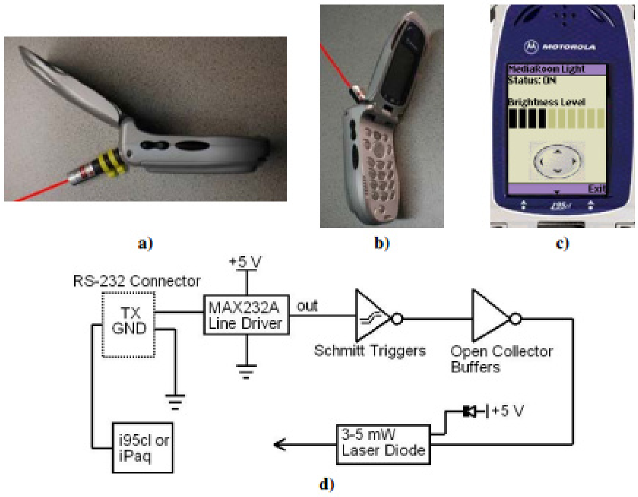


Figure 2.1: a-b) the customized handset with laser pointer.c) example control interface. d) Schematic of Laser Controller [13]

Figure 2.2: The Xwand pointing device.

2.1.2 Selection

The referred papers describe various technologies used to transmit signals and selecting an object for further control. Many of them make use of buttons placed on the customized pointing device to register a selection. For example the gesturePen uses a button on the "Pen" to transmit infrared radiation

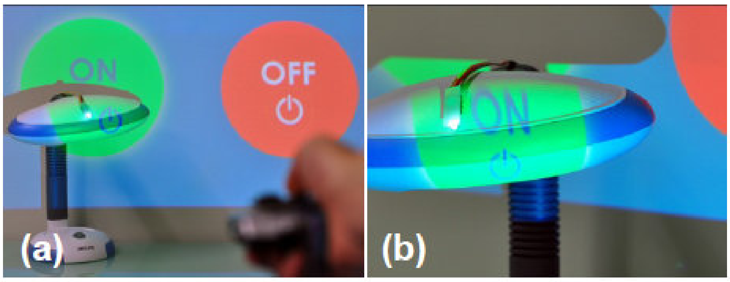


Figure 2.3: PICONtrol:A simple interaction example; projecting an the active region termed "ON" on the sensor and pressing the button on the projector switches on the lamp.

towards a tag on the target object. They make use of IR technology which is a directional technology. Unlike Radio frequency(RF) transmission, which is omni-directional. RF on the other hand is used in RFIG lamps, to select all the RFID tags in range, but only those are selected which are within the projected region. The PICOntrol does not involve the selection step because, any object on which an interface is projected gets implicitly selected. The selection step is actually over when the system knows which object was pointed at. To achieve this the gesturePen uses IR emitting tags which communicate back the receiver's IP address. The Laser Pointer also utilizes a similar approach but a different technology.The laser pointer itself emits the Pointer's IP address through the laser signal and the tag receiving this signal interprets that it is being pointed at. SuperVision employs a very

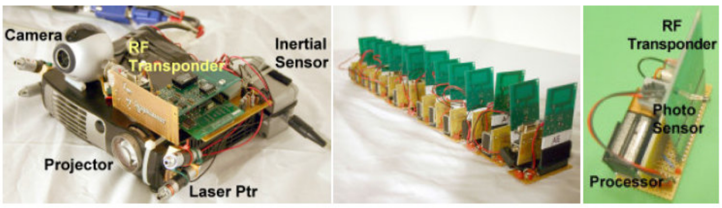


Figure 2.4: RFIG Lamps: a) The pointing device prototype b) The Active Tags.c) Close up of a Tag

simple method of registering a selection. Using a button of on the remote control, signals are passed to the host computer wirelessly , to tell the system what is the pointing direction. This direction is then mapped to a particular object in the room. There is a small balck cursor displayed at the center of projection always, to help users realize what they are pointing at actually.

2.1.3 Disambiguation

By disambiguation we mean telling the system, which particular object to select, when the system is confused among many closely spaced objects. For example, in fig 2.5, from the paper on disambiguation techniques[4] by Delamare et al. shows that for situations involving high object density, it might be more prudent to involve two steps in selecting the target object. The first step is a volume selection, and the second step is the disambiguation step, where the system chooses one particular object from amongst the selected volume of objects.

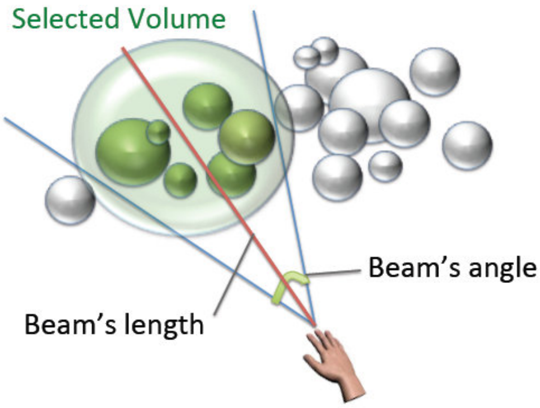


Figure 2.5: Volume selection for closely spaced objects

Thus the entire task of selecting a target object can be represented by a state diagram as shown in[4].

GesturePen which uses IR beam to select objects, works on the principle of volume selection.Authors proposed to bring in disambiguation by manually adjusting the beams length and/or angle. PICOntrol similarly has no apparent disambiguation step involved. But they propose to tackle this by reducing the projected area. In RFIG LAMPS[16], the user can easily disambiguate by pointing with a laser at the projected tags. On the other hand,[13] uses a very direct method, involving no disambiguation step.

2.1.4 Control

The control methods vary as well from one interaction technique to another.Xwand,[24] uses gestures like flicking or rolling of wrist to pass commands. This is done with the help of the gyroscope and accelerometers placed on the wand. PICOntrol[19] on the other hand depends entirely on the projected light.

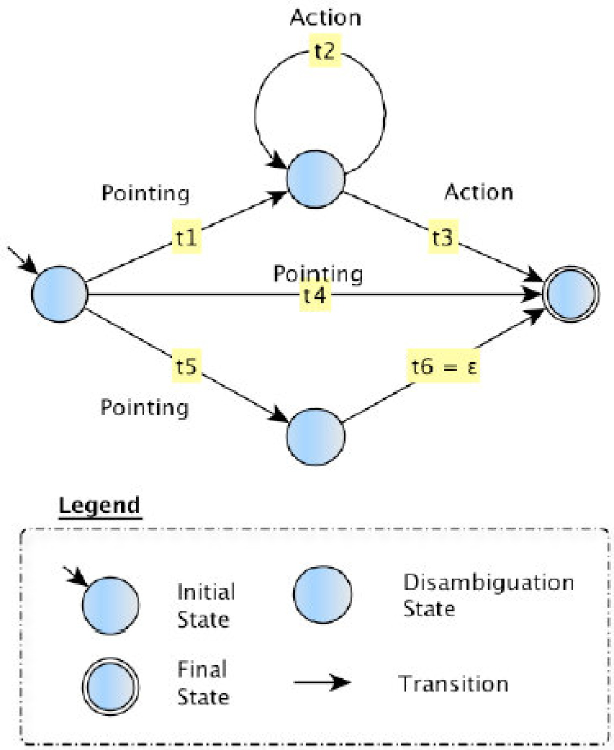


Figure 2.6: t1-t2-t3 : pointing+interactive disambiguation. t4: Direct selection of target in sight. t5-t6: Automatic disambiguation using a heuristic method.

They transmit gray codes through the projected pixels and the sensor on the object matches and understands the pattern of coded light falling on it. GesturePen[21] depends on a handheld device or a laptop to pass control information through the central network. In [13], all control interactions take place through the display screen of the mobile phone, into which the laser source is fixed. In SuperVision, all control takes place through the projected interface, which is displayed on and around the selected object. This makes sure that there is no division of focus from the controlled to the controller, and also there is a decoupling of the controller and the interface.

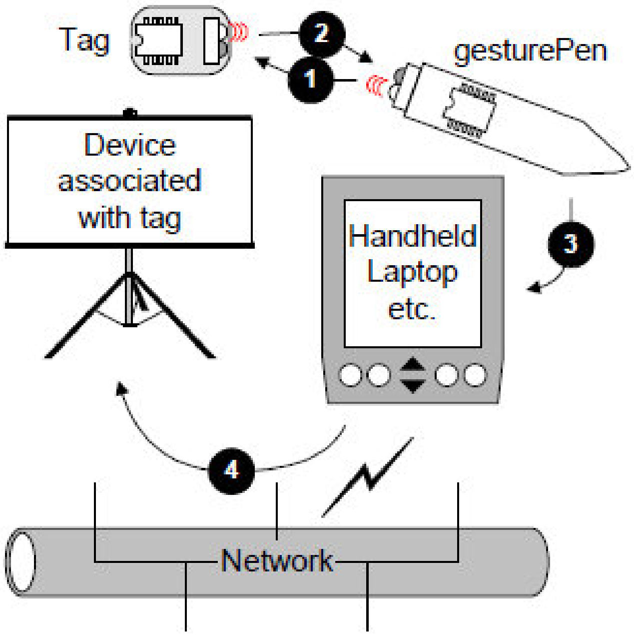
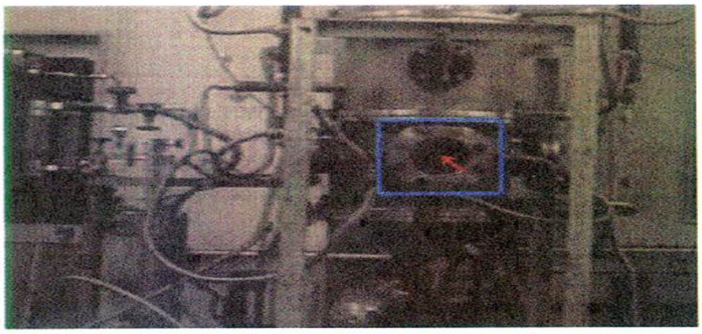


Figure 2.7: Communication channels between the gesturePen, the tags and the Handheld display.

2.2 Video Interaction

One of the core ideas of "SuperVision" is the use of live video to visualize objects which are out of sight, and also enabling the user to interact with the objects directly through the video. A direct relation of the concept can be seen with the work of Tani et. al[22]. In fact Tani et. al were the first ones to utilize graphical elements overlaid on a live video to interact with big machines in a plant. They fixed a video camera at a particular direction, to face a machine that had to be controlled from a computer in the control station. On the video image received from the camera, they defined certain areas(2D) to represent different parts of the machine. When such parts of the machine a re clicked, different graphical elements (like sliders) appear around the clicked area, which can be used to control the machine.

SuperVision employs a very similar interaction method of defining areas of the video image to represent different objects, but instead of the interaction taking place on a desktop, we project the entire video on the wall. Which lets the user make use of his spatial memory to directly point at the objects that he wants to control. Thus there is no more clicking required. Moreover, the video camera is capable of panning and tilting and the system



Tani Plate 4-1
Pointing to the window of the boiler causes it to be highlighted by the blue rectangle.

Figure 2.8: Tani's interaction technique: clicking on the circular part of the machine activates the overlaid graphics. Which can be used for further interaction

will adjust the camera's direction according to the pointing direction of the user.

An extension of Tani's concept can be seen in [2]. Boring et. al [2] introduced the method of manipulating virtual objects among different displays when there are multiple displays available, by making use of teh video camera present in hand held mobile phones. The interaction involves pointing at a an object on a display screen by looking through the phone camera. Then selecting the object with finger(touch) and "dragging" and "dropping" it on another display. This enables the user to emulate touch behaviour on screens which may not be touch sensitive actually, and also allows access to those displays which are out of reach of the user. The following figure from [2] describes the entire interaction process.

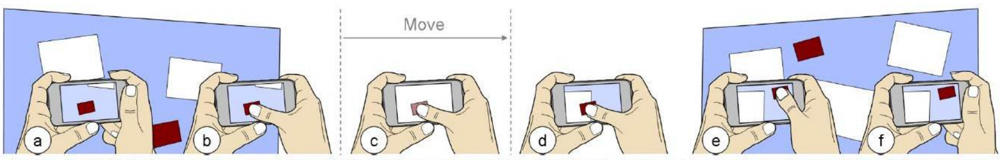


Figure 2.9: a-b)User aims at a display and touches the item of interest;c)When dragging the item from one display to another, a thumbnail of the item is shown on the mobile screen;d-f)The item is positioned accurately by pointing at the target screen and releasing the finger

While TouchProjector [2] applies the concept of video interaction, it is bounded in the world of manipulating virtual objects. SuperVision on the other hand extends the concept of physical pointing and controlling, to the world of real objects. This is achieved using hand-held pico projector, and a camera capable of panning and tilting.

An interesting application of video interaction is found in "Sketch and Run" [18]. "Sketch and Run" handles the specific problem of controlling a home robot, by defining its path on a handheld touch screen. To be precise, a camera is placed on the ceiling pointing downwards. the video feedback from this camera is fed to the hand held. Thus the user can see a top view of the room now. On this top view, the user draws a path using the finger. The system interprets the corresponding path in the real world coordinates, which is then followed by the robot.

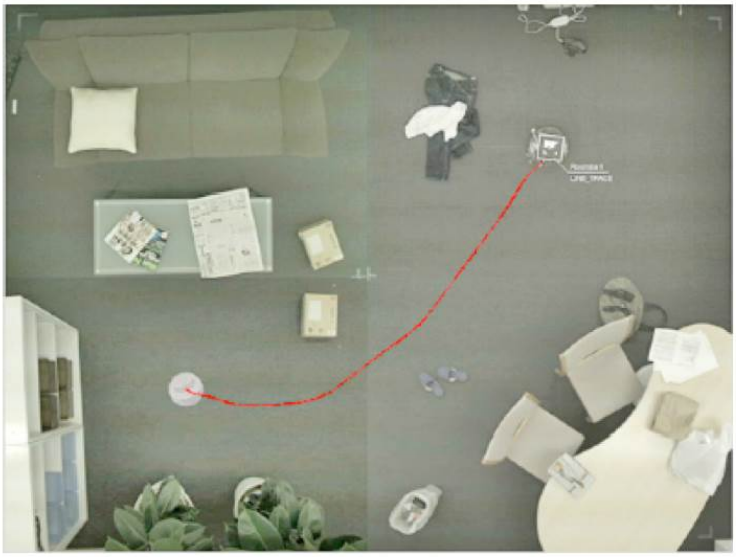


Figure 2.10: The top view of the room setup for the Sketch and Run Interaction technique

SuperVision takes a step further in making the interaction with a home robot (a vacuum cleaner robot) more intuitive. This is done by involving direct pointing on the ground for drawing the path which the robot has to follow. By using SuperVision, the user just has to point at the robot using the hand-held projector, and then the robot follows the virtual path drawn out by the projection center on the floor.

2.3 Transparent Obstacles

Since SuperVision attempts at visualizing the objects behind the walls of a room, the field of work dealing with "Transparent Obstacles" becomes extremely important. A great metaphor of transparent obstacles, specifically transparent walls is seen in "See through walls with WiFi"[1]. The authors Adib et. al present a new method of tracking movement behind a wall, using nothing but low power Wi-Fi signals. This technique can be used to map gestures performed in one room, to consequences in an adjacent room. For example, if a person performs a gesture in the bedroom, it can be tracked from the living room through wifi and a lamp can be switched on or off depending on the gesture. This concept is very close to the concept of SuperVision because we also attempt at controlling objects which are not in the same room as the user. But SuperVision is fundamentally different in the fact that we focus not on the tracking, but on the interaction more. Thus we allow the user to just point, and then visualize a distant object through the projector and control the desired object by just moving the projector. There are no specific set of gestures to be memorized.

Pindat et. al[14] also contributed towards visualizing through obstacles. Their work was focused towards how to visualize the insides of a complex 3D model through the model's outer shell. So they utilized the metaphor of "drilling" through the outer casing of a machine to locate and visualize the target part inside. Although this metaphor is similar to SuperVision's metaphor of "drilling" a hole in the wall and looking through it, the fundamental difference lies in the fact that SuperVision applies the virtual visualization technique to augment the real world.

2.4 Smart Home Control

Effective smart home control is one of the most discussed matters of research . The controlling of a smart home involves how the different objects interact with one another, and with the user. But the level of automation in the house that is comfortable for the user without taking away his/her freedom is also a matter of argument. Thus there are two directions of research in the field of smart home control, one that concentrates on making intelligent systems that can make decisions on their own, without the user having to interfere[3]. The other direction is one in which the user still retains the power to control all the appliances, but all the appliances are interconnected and are capable of rich interactions with the user. We concentrate on the second direction;

the user controls the objects.

Several of the existing methods that allow easy control of house appliances, rely on the use of smart phones or computers [6], [8], [9]. Often times this involves the use of lists or icons to choose from, and when there are many varieties of the same object(for example lamps), each has to have an unique identity (that has to be remembered by the user). Moreover, the use of a phone or a tablet, to interact with physical objects, divides the focus of the user from the controlled to the controller, which is undesired, because the user does not get an immediate feedback of the consequence of an action.

Russis et.al[17] proposed a wrist worn human-home interaction system which allowed the watch to interact with the Ambient Intelligence environment through messages. The figure below shows an example message delivered to the user through the wrist watch, by the intelligent system. Although the use of a wrist watch makes it extremely simple to carry around and interact, the small size of the display screen is indeed a constraint when it comes to the display of interfaces. SuperVision on the other hand utilizes the walls of a room as display surfaces. This allows for the display of interesting interfaces, (for example color wheel for setting color of a lamp) and makes interaction far more intuitive.

Figure 2.11: A message delivered to the user, demanding attention by activating the watch alarm

Chapter 3

Design

3.1 Scenarios

To design our system, it is quite important that we come up scenarios, in which it can be useful to control some objects from a distance. Since there has already been a lot of work about controlling "line of sight" devices, we specially focus on scenarios in which it becomes necessary to be able to control objects and devices which are not present in the current environment around the user, ie. objects which are kept in different rooms. Ofcourse we consider only a home or an office environment to fit in our scenarios.

3.1.1 Scenario 1

Bob is very busy working in his study room. But since he has been working for quite long, he sure does feel hungry as well. So he runs to the kitchen puts something to cook on the stove and returns back to his study room. As he slowly gets engrossed in his work again, the fact that there is something cooking in the kitchen quickly escapes from his mind. Then suddenly he smells something and now he remembers! He bolts across to the kitchen, only to find whatever he had hoped to eat, has been reduced to ash. If only he could switch of the stove from his study room!

3.1.2 Scenario 2

Ann is taking care of the baby who has suddenly woken up and crying a lot. There is nobody in the house. Suddenly it starts to rain outside. Ann remembers that she had put some clothes to dry, on the clothesline in the balcony. If she doesnot put the balcony shade on, everything will get wet and she has to get them rewashed. But at the same time se cant leave the baby alone. If only she could switch on the electronic shade from her current positionm inside the house!

Figure 3.1: Electronic balcony shade

3.1.3 Scenario 3

Olive has an important skype call scheduled this evening, but because of the traffic she is running late. She drives back home as quickly as possible, parks her car in the garage and rushes inside. When she is in the middle of her call, she realizes that she has forgotten to close the garage door. Now even if she could close the door from her room, she would still want to go, close the door and make sure that her car is still there. If only she could control her garage door and get direct realtime visual feedback from the garage as well!

3.2 Concept

The idea is to give users the super-human capabilities of having control over inanimate things, just by pointing at them. This involves all devices which are in the same room as the user, and also in other rooms. It is known that in a familiar environment, one can point at the approximate direction of an object, and one does not need to be able to visually perceive the object for this. We wish to capitalize on this spatial memory of users, when it comes to a familiar environment.

3.2.1 Objects in the line of sight

Objects which are in the same room as the user, may be selected just by pointing at them. Here we assume that the objects are static in the room. That is, the position of the devices inside the room cannot change, and if they change, their new locations have to be updated in the system as well. Once an object is selected, the related interface can be displayed using a hand-held projector. If there is no proper surface to project on, near the selected object, the interface can be "dragged " and dropped on a more suitable, plainer surface. Further interactions like selection of commands on the interface can be done using buttons and sliders on the interface.

3.2.2 Objects out of the line of sight

By saying, "Objects which are not in the line of sight" we basically mean objects which are not present in the same room as the user. The idea is to bring in a "Superman Vision" metaphor.

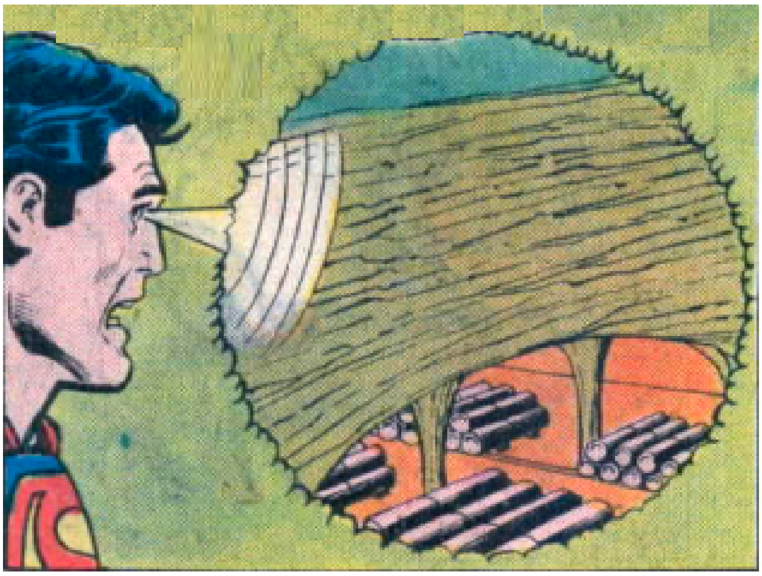


Figure 3.2: Superman's vision power to see through walls

As apparent from the Figure 3.2, superman possessed the power to see what is located behind a wall, through a virtual hole in the wall. It is to be noted that the whole wall does not become transparent to his eyes. On the other hand it is as if, he drills a hole through the wall with his eyes. But this hole is virtual. When our users get the power of this "Superman Vision" they can point at some part of the wall, and zoom in with a slider and see what is beyond the wall through only the projected area. So the projected area behaves like a virtual hole in the wall. Once they can see what is behind the wall, the rest of the interaction is meant to be similar to the case when they are interacting with objects in the same room.

3.2.3 Spotlight Metaphor

Another related idea is to use the metaphor of a spotlight in a virtual space. This metaphor has already been utilized in the project of Spotlight Navigation[15]. The basic concept is that we create a large virtual space, all of which is not visible at the same time. Imagine using a spotlight in a dark environment. So only that part of the environment becomes visible, which is illuminated by the cone of light falling from the spotlight. Extending this, a flashlight can be replaced by a projector, which reveals only some parts of a virtual environment depending on where you are pointing. Several applications of this metaphor can be found in [15], but how this can be utilized in our goal of interacting with objects, rests to be seen.

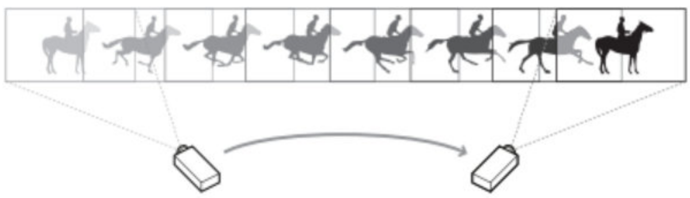


Figure 3.3: The Spotlight Metaphor

Figure 3.3 taken from the the paper on "Motion Beam"[23] rightly explains the spotlight metaphor. The set of 8 images of the horse, forms the virtua space, which is constant. But the user sees only the part where he points at, that is, the part which is illuminated by the "flashlight" like projector.

Chapter 4

Implementation

4.1 Assumption

In our implementation exercises, it becomes necessary that the system is aware of the location and orientation of the user or the pointing device. To fast-track our implementation process, we assume that some form of an user tracking system is already available for use. In fact five years from now this might not be a far fetched assumption. Currently we use a system named as DTrack2, supplied by a company calledm Advanced Realtime Tracking, (ART).

ART : ART tracking systems are infrared optical tracking systems. Tracking refers to measuring the position and orientation of objects in 3D space. To do this objects have to be equipped with rigid arrangement of markers. These markers are small balls arranged in a particular order and painted with retro-reflective paint, which reflects light back to the cameras, that flash the IR. Currently we use 8 infrared ART cameras and a "flystick" as a pointing device. A flystick is a special joystick, which comes equipped with markers, and hence can be tracked using the ART cameras.

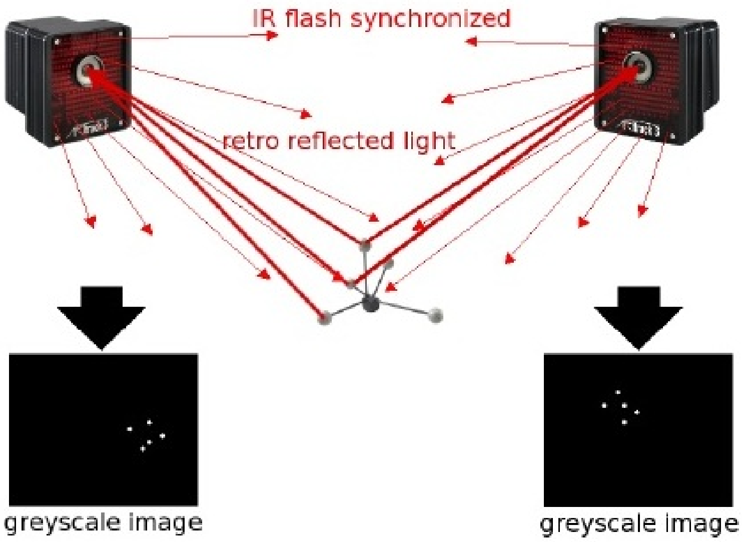


Figure 4.1: The ART tracking cameras , and the rigid marker reflecting back the IR flash.

4.2 Using Information From the ART controller

The ART controller can be set up to give out 6 degrees of freedom information regarding the flystick/body that it tracks. This is in the form of a rotation matrix and a location matrix. Position and orientation of a target are expressed by an affine transformation $(\vec{x}, S)$ that transforms a vector from the body coordinate system to the room coordinate system :

$$\vec{x}_{room} = R.\vec{x}_{body} + \vec{s} \tag{4.1}$$

This information can be transferred to a remote computer connected to the same network as the ART controller, using a "client-server" program running at the localhost. The code for such a program was written in C++, and using a few libraries of OpenGL, a 3D model of the room and the flystick was generated as shown in the following set of images.

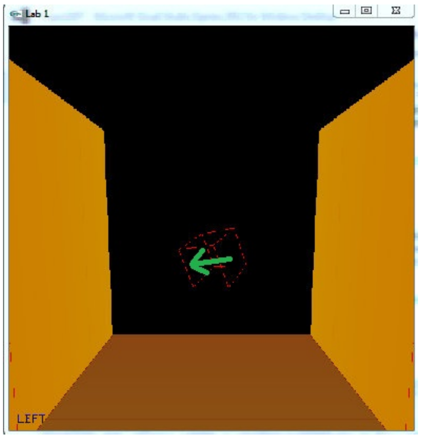


Figure 4.2: The wire-cube represents the handheld pointing device. the arrow shows the direction at which i am pointing. There is suitable text output at the bottom left corner, which shows whether I am pointing to the left wall or the right wall.

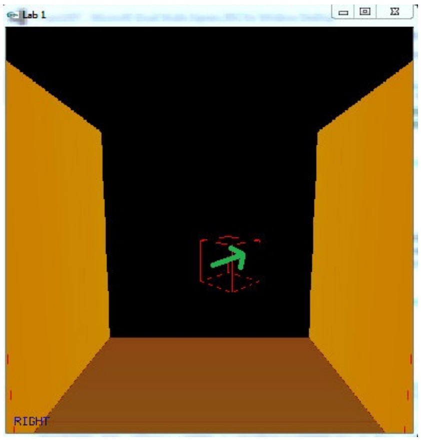


Figure 4.3: As before, the virtual cube copies the position and orientation of the pointing device in my hand. Here "Right" is displayed as we are pointing towards the right wall

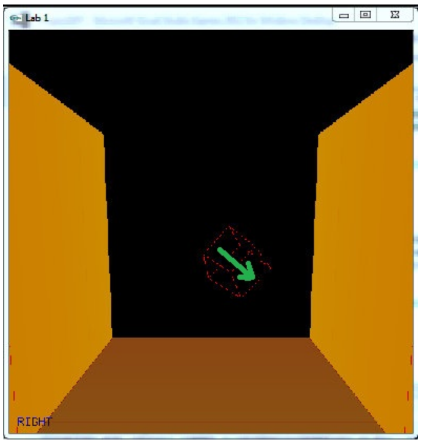
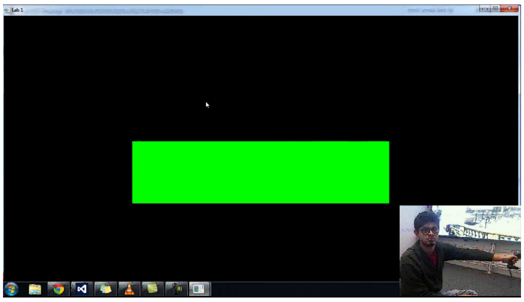
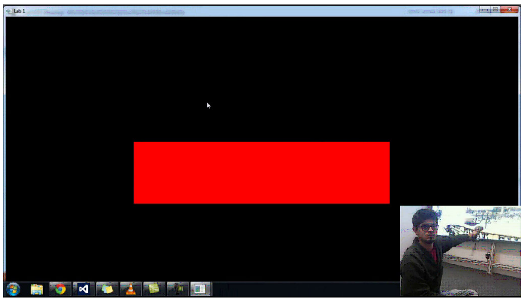


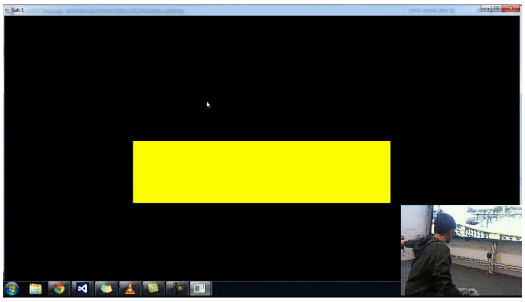
Figure 4.4: Using the same display technique as before, we are pointing right and down.



(a) Pointing in front of me towards a chair(unseen) kept in the room, makes the rectangle in the screen green



(b) Pointing towards the big screen, makes the rectangle red.



(c) Pointing behind towards a chair (unseen) kept in the room, turns the rectangle yellow.

Figure 4.5: Associating directions with colors. By using the rotation angle with the vertically up axis(here the y axis), we associated certain ranges of angles with some colors.

4.3 Interactions Using an Image

One of the ideas is to use a wide angled image of a room, to interact with the objects visible in the picture of the room. Such a picture can be displayed using a projector and interactions can be done, to control objects kept in other rooms. A small, mouse based application is built as a first stage of the prototype. This is done using c++ language in Qt environment.

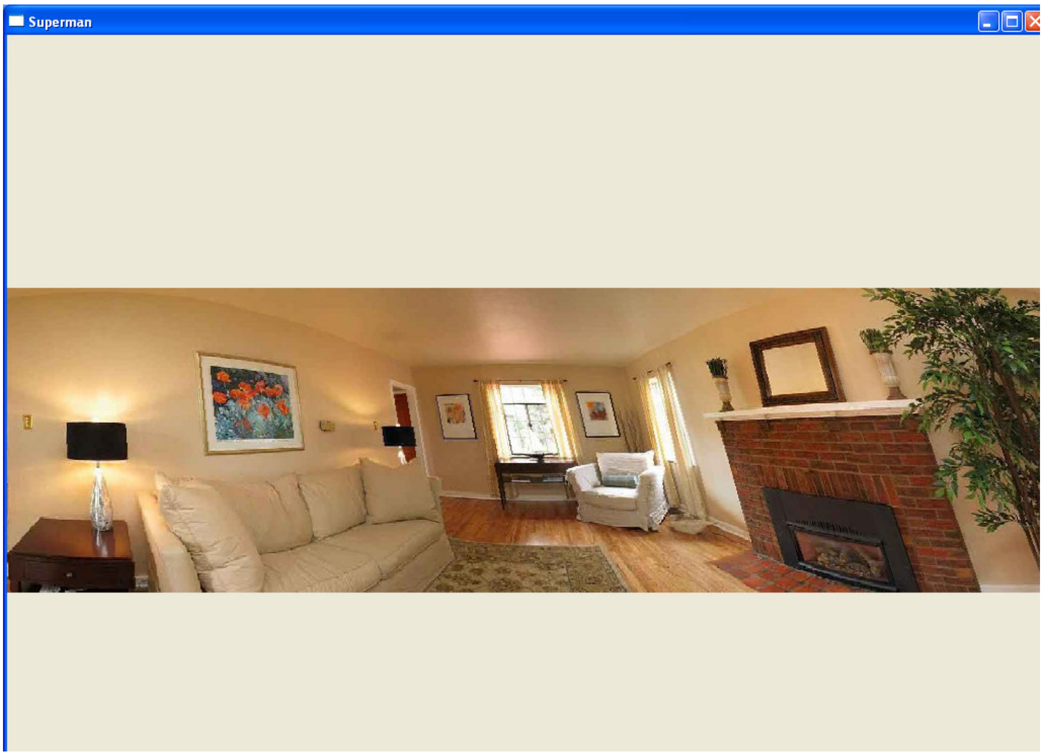


Figure 4.6: Displaying a wide angle photo of a living room

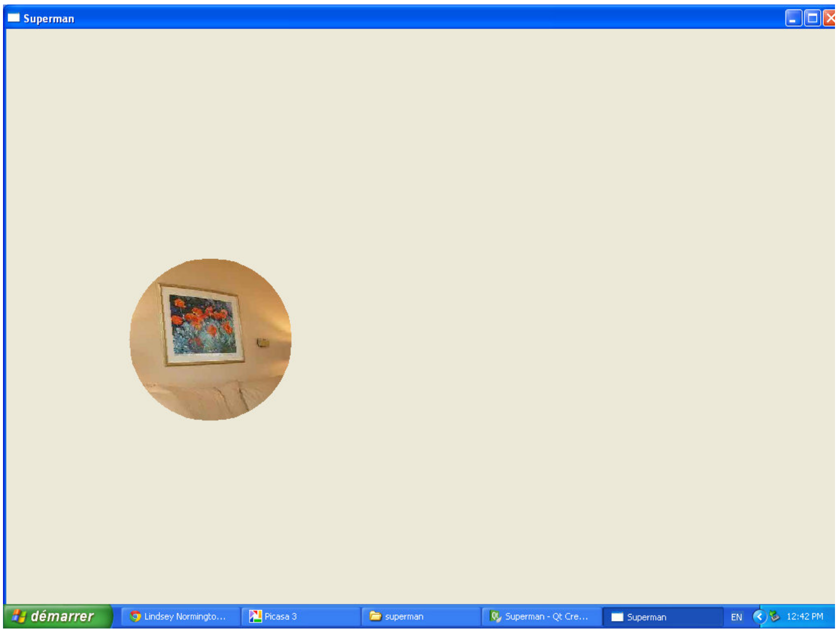


Figure 4.7: clicking the mouse hides the rest of the image and shows a circular "spotlight" through which the part of the room directly behind the circle can be seen. It becomes our selection circle

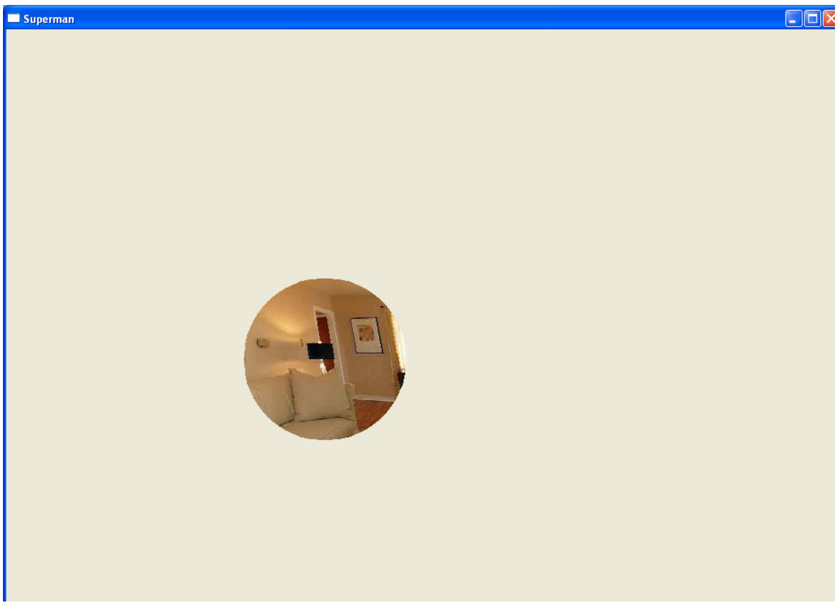


Figure 4.8: The mouse is moved while holding down the left click button to position the cursor over the small lamp

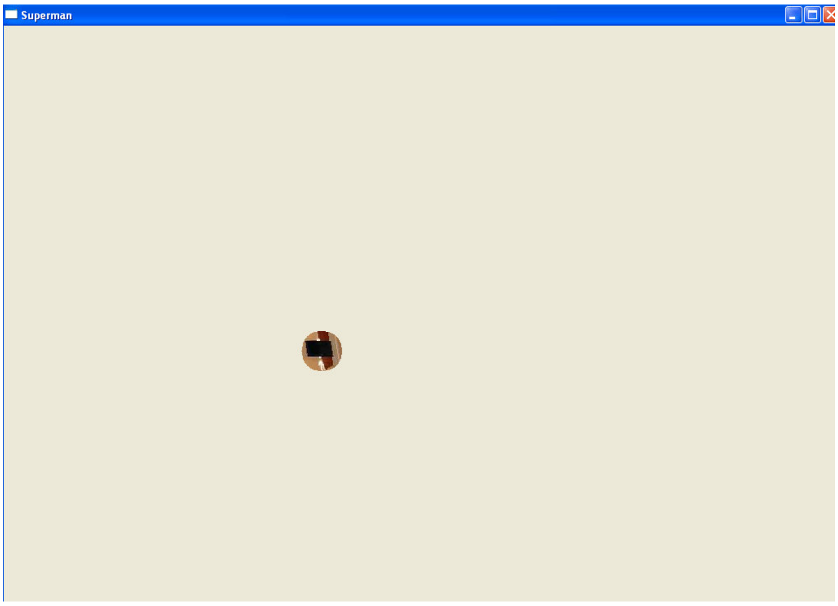


Figure 4.9: The selection circle size is reduced so that it just fits the lamp

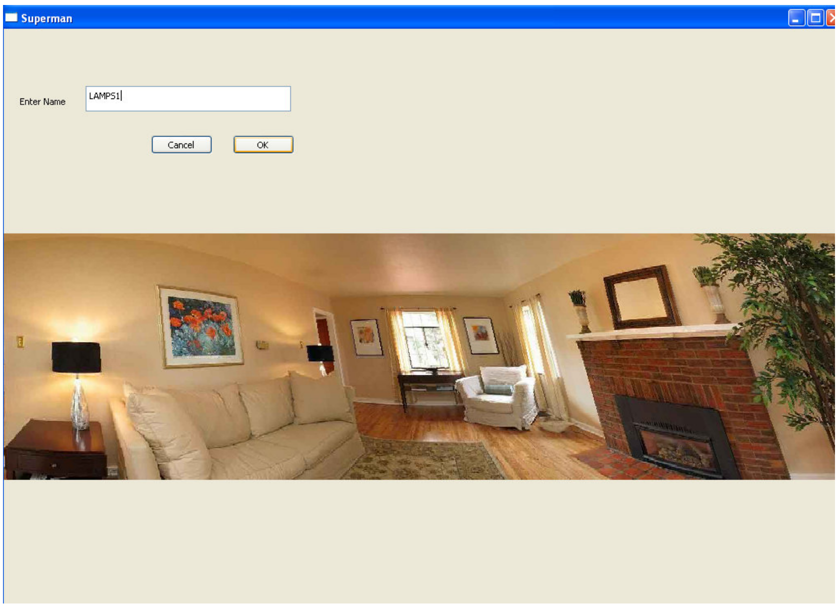


Figure 4.10: Releasing the mouse over the lamp, associates the current area inside the circle with the name you are about to enter. Here we enter LAMP1

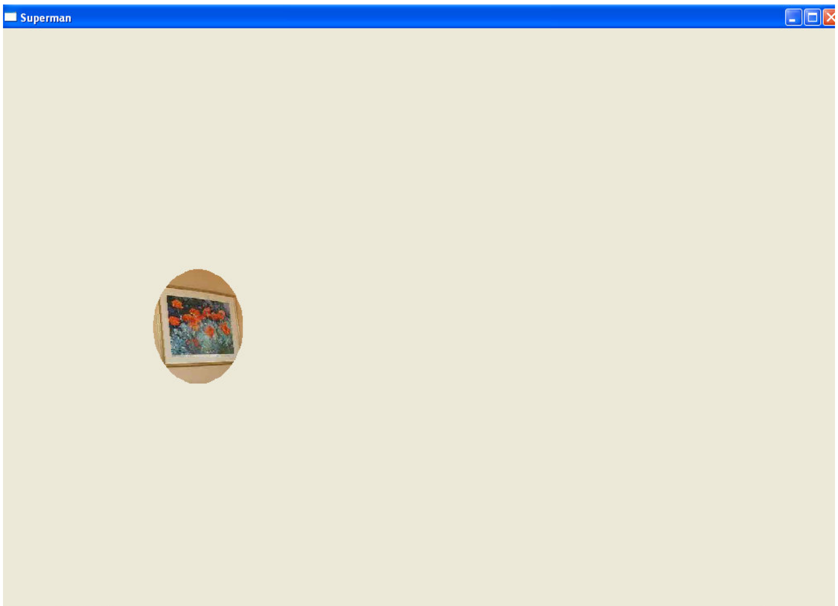


Figure 4.11: We click on teh picture once again and drag the circle to over the picture frame on the wall.

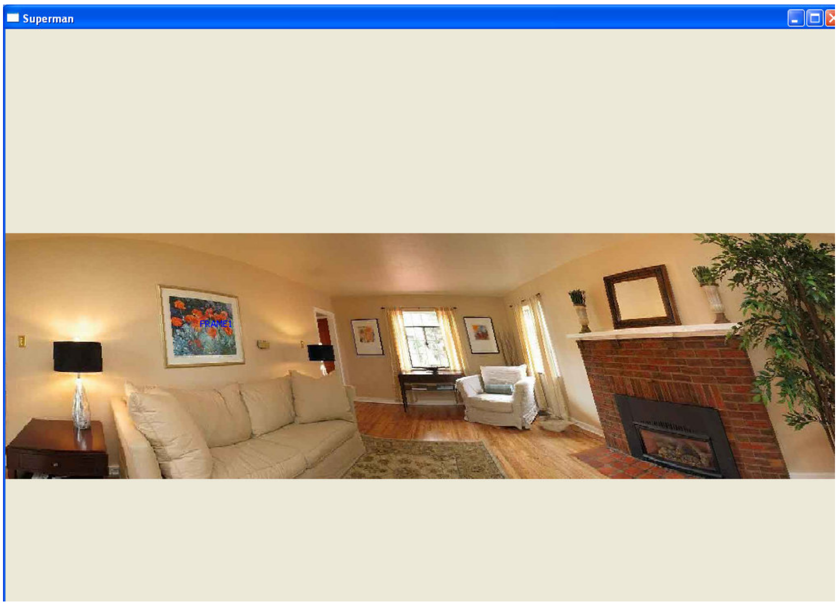


Figure 4.12: When we have entered the name of the last selection as "FRAME1" it shows the name on the frame after we press "OK"

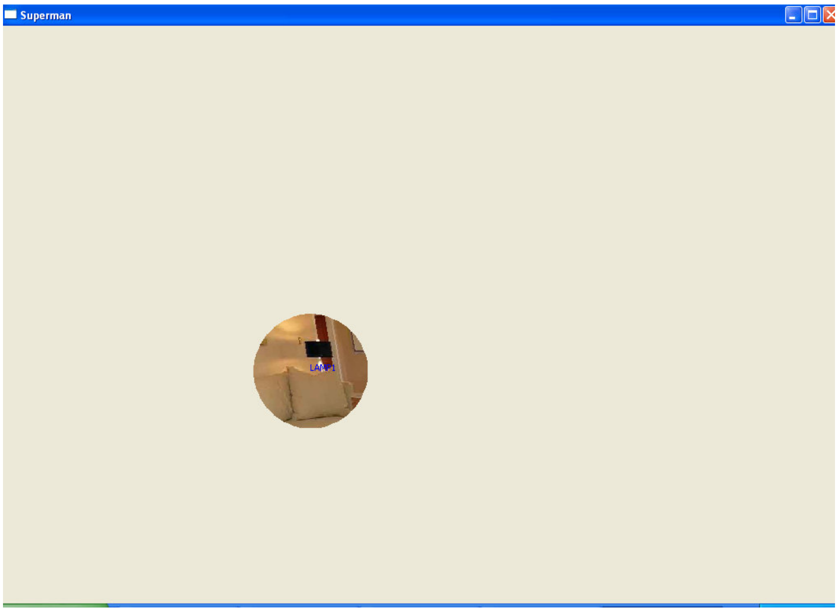


Figure 4.13: Now if we again click and drag teh selection circle to the lamp which we have already named, it shows that this device is called LAMP1

Figure 4.14: Interactions with an Image

4.4 The Superman Vision Metaphor

The idea is to let the user zoom past and through a wall, using a physical slider. It becomes important that it is made visually believable to the user, to increase the intuitiveness of the interaction. To begin with, two small applications are built in Qt, to express the superman vision metaphor. The first one, figure 4.15, is done by clipping successive image, with an ellipse, whose size is controlled by the slider on the top-right. The second one, is done by drawing a ragged region and turning it into a mask. The size of this mask is controlled by the slider and the mask is applied to the overlying image of the brickWall.

(a) Plaster Texture

(b) Bricks are shown

(c) A hole appears

(d) hole widens

(e) Kitchen is visible

(f) Final stage

Figure 4.15: From a-f, we show the stages in bringing out the metaphor of looking through the wall at the kitchen behind.

(a) Brick Wall

(b) A break in the wall

(c) The Hole Widens

(d) The kitchen is visible through the wall.

Figure 4.16: Another attempt at making apparent, the concept of seeing through the wall. This is done using masking techniques.

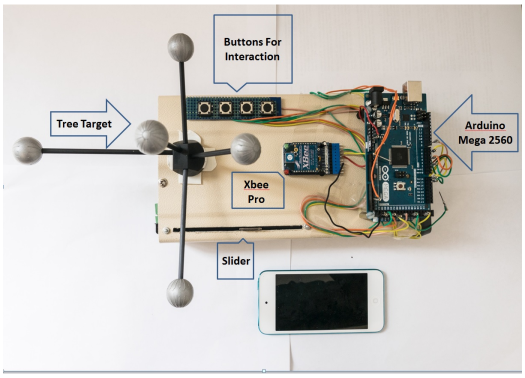
4.5 Hardware Description

We have built a special remote controller for SuperVision, which as a pocket projector in it. The hardware consists of and arduino mega 2560 micro-controller board,an XBee PRo module for radio communication, a motorized slider 10K potentiometer, and there are four normal buttons on the controller. Currently we are using just a single button to act as a select button, but other buttons are just for future use. The remote also has a 100 lumens off the shelf projector; Picopix 3610 from Philips. This Projector is Wifi enabled and the video information is passed from the computer to the remote through Wifi.

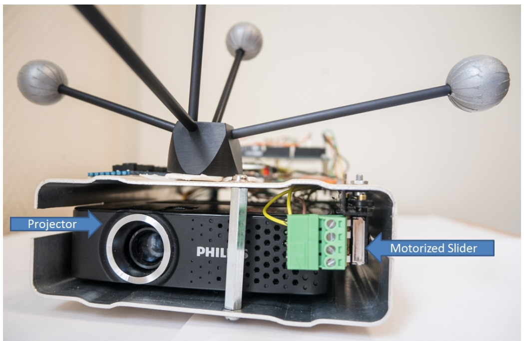
On the computer side, we have an Xbee Pro module connected to the computer through an FTDI cable. Thus the remote communicates the control signals with the computer through Xbee protocol, while the computer sends data to the projector through Wifi.

The camera being used for video feedback is an off the shelf Sony EVI-D90 pan, tilt and zoom camera. It has a wide angle capabilty of 55.8 degrees and a pan range of +-170 degrees. This camera is controlled through a serial interface by the VISCA protocol[?].

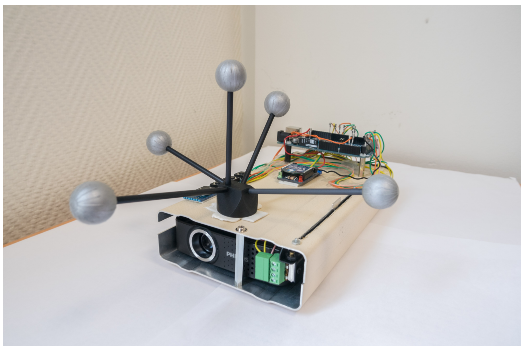
The figure4.17 below show both the prototype controller, with the projector, and the figure 4.18 shows the overall system design.



(a) Top View



(b) Front View



(c) Perspective View

Figure 4.17: The prototype of the hand-held controller. a)The pocket projector inside the controller, alongside the slider potentiometer. b) The different components of the controller c)Overall view

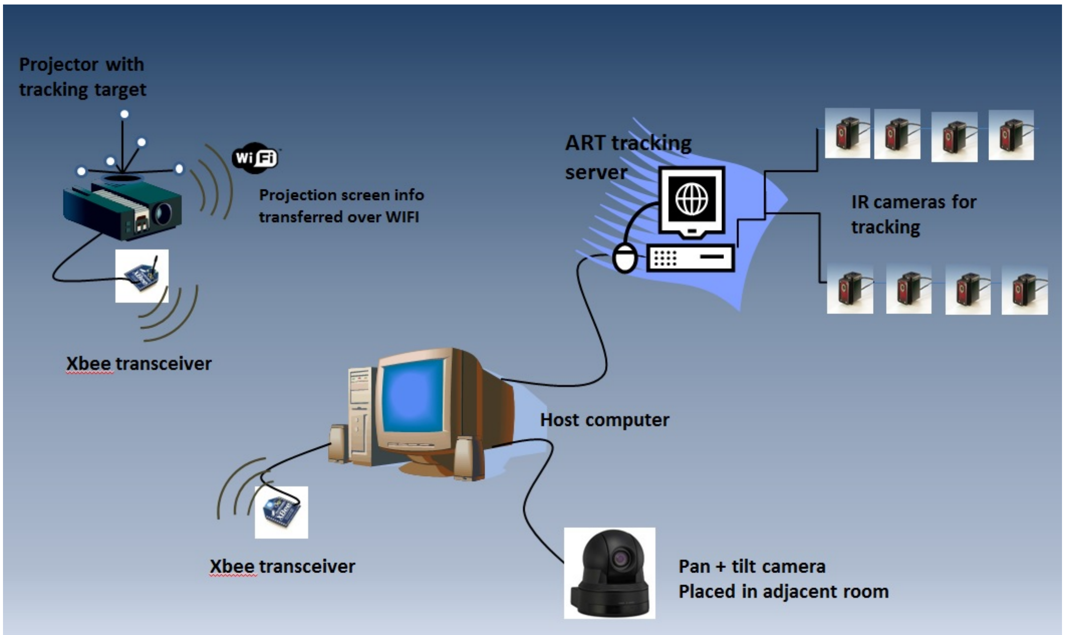
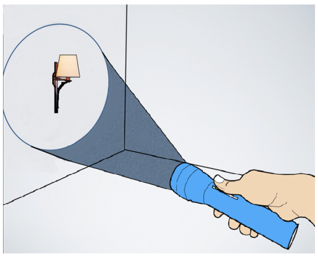


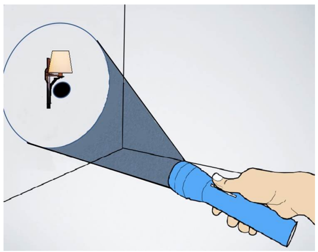
Figure 4.18: The System Design: The Host computer is connected with a Xbee module and a camera. It is connected to the ART tracking server through a common local area network. The controller has a Tree Target which is being tracked by the ART cameras. The controller also has an Xbee mmodule to communicate with the host computer, whether a button is pressed or not, and whether the slider is moved or not.

4.6 Example Interaction with Lamp

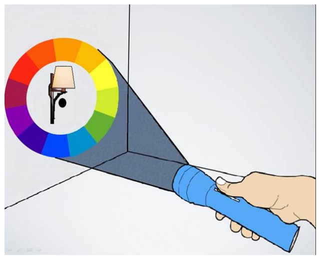
Here we give a step by step explanation of how an user can interact with a lamp located on a wall. Using the controller, the user first points at the lamp. A cursor can be seen at the center of the projection area when the user does that. By positioning this cursor on the lamp, the user presses the "select" button. Upon pressing the button, a color wheel appears around the lamp. It is to be noted that the positioning of the color wheel is fixed with respect to the real world coordinates, and not the virtual coordinates. So by moving the projector around the lamp, the user can see some or all or no part of the color wheel. Now the user clicks on the center of the wheel, and drags the cursor towards the outer circle, aiming at the desired color. The color wheel acts like a "crossing widget" and the color where the cursor crosses the wheel, gets selected. The lamp now glows with the same color.The Figure 4.19 explains all the steps pictorially.



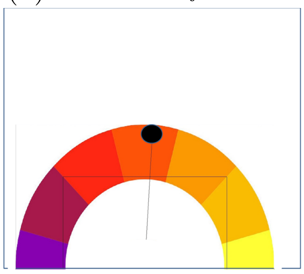
(a) Point at Lamp



(b) A cursor at Projection Center



(c) On Pressing Select



(d) Choosing the Color

Figure 4.19: Interaction with a Lamp on the wall: a)User points at the Lamp b) User sees the cursor in the projection Center c)User positions the cursor near the lamp and presses select. A color wheel appears around the lamp d) User clicks and drags the center of projection across the orange color to select it. The color wheel remains fixed in position in the real world coordinates, while the projected area moves.

Chapter 5

Future Work

Development of real world applications is the most important goal for the near future. This includes applications for controlling lamps, televisions and vacuum cleaner robots. Another important task left for the future is to introduce perspective corrections for the projected image so that the visualizations become more realistic. Implementing the camera control wirelessly, and also introducing semantic zoom depending on the user's position, are the other items on the agenda.

USER STUDY : Conducting an user study is highly important in any form of research in the field of Human-Computer Interaction. In our case also, once two or more applications are ready to be tested, we intend to conduct an user study. In this user study , the main focus will be to note down the time taken by each user to choose and control and object, both by pointing at it directly, and by using a more standard approach, or a hand-held device. A comparison of the results and observations can open up further scope of improvements in the prototype designs.

Chapter 6

Conclusion

From the related work,we saw that most of the literature stress on the fact that the interaction with an object is termed to be better if it is more intuitive and puts less cognitive load on the user. Thus we try to make our interaction more intuitive by including a pointing gesture in our first prototype. We also try to address the task of controlling objects that are not in the line of sight of the user. In the first prototype, we have developed a system to identify the direction at which an user is pointing using a tracking tool called "ART"; Advanced Realtime Tracking, the pointing direction can then be used to find out if any object falls in the same direction. This object can then be selected and information/interface can be projected on the identified object. We try to visualize objects situated in other rooms,(rooms other than the room where the user is present) by using wide angled (panoramic images) pan and tilt cameras which can be controlled to face the in the same direction as the user's pointing direction. We implement "spotlight navigation" to let the user see only that part of the image where he/she is pointing at. the object that falls inside the user's "spotlight" can be selected. Such a method integrates the concept of "spatial pointing memory" which makes the task of choosing an object interesting and effective.

In the future we will concentrate on building real world applications for the technique introduced here so that we can conduct a user study and evaluate it.

Acknowledgments

I would like to express my gratitude towards my advisor Dr. Gilles Bailly for giving me the wonderful opportunity to do my thesis at the Department of INFRES, Telecom ParisTech. His constant guidance and feedback has helped immensely in taking this project forward. I am also thankful to Dr. Eric Lecolinet, who has played a major role in formulating this project as well. Moreover I must also thank all the members of the Human Computer Interaction Department, especially Hind Gacem, and Thibaut Jacob, whose technical inputs have been extremely valuable.

My special thanks goes out to Mr. Chetan Kumar, who is my Co-supervisor from BITS Hyderabad Campus. I deeply appreciate all his help. I also wish to thank Dr. Y Yoganandam, and Dr. BVVSN Prabhakar Rao, of the Electrical Engineering Department of BITS-Pilani Hyderabad, for helping me in pursuing this off-campus Bachelor Thesis at Telecom ParisTech.

Sarthak Ghosh
May 2014
Birla Institute of Technology and Science-Pilani
Hyderabad Campus

References

[1] Fadel Adib and Dina Katabi. See through walls with wifi! *SIGCOMM Comput. Commun. Rev.*, 43(4):75–86, August 2013.

[2] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA, 2010. ACM.

[3] Scott Davidoff, Min Kyung Lee, Charles Yiu, John Zimmerman, and Anind K. Dey. Principles of smart home control. In *Proceedings of the 8th International Conference on Ubiquitous Computing*, UbiComp'06, pages 19–34, Berlin, Heidelberg, 2006. Springer-Verlag.

[4] William Delamare, Céline Coutrix, and Laurence Nigay. Designing Disambiguation Techniques for Pointing in the Physical World. In *Proceedings of the 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2013)*, pages 197–206, London, UK, Inconnu, 2013.

[5] http://smarthomeenergy.co.uk/what-smart home.

[6] http://www.control4.com/.

[7] http://www.fiercetelecom.com/special-reports/nikola-tesla-remote control.

[8] http://www.savantsystems.com.

[9] Sokratis Kartakis, Margherita Antona, and Constantine Stephanidis. Control smart homes easily with simple touch. In *Proceedings of the 2011 International ACM Workshop on Ubiquitous Meta User Interfaces*, Ubi-MUI '11, pages 1–6, New York, NY, USA, 2011. ACM.

[10] Sunjun Kim, Jaewoo Chung, Alice Oh, Chris Schmandt, and Ig-Jae Kim. ilight: Information flashlight on objects using handheld projector. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 3631–3636, New York, NY, USA, 2010. ACM.

[11] Brad A. Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, and A. Chris Long. Interacting at a distance: Measuring the performance of laser pointers and other devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 33–40, New York, NY, USA, 2002. ACM.

[12] Yusuke Nakazato, Masayuki Kanbara, and Naokazu Yokoya. Wearable augmented reality system using invisible visual markers and an ir camera. In *Proceedings of the Ninth IEEE International Symposium on Wearable Computers*, ISWC '05, pages 198–199, Washington, DC, USA, 2005. IEEE Computer Society.

[13] ShwetakN. Patel and GregoryD. Abowd. A 2-way laser-assisted selection scheme for handhelds in a physical environment. In AnindK. Dey, Albrecht Schmidt, and JosephF. McCarthy, editors, *UbiComp 2003: Ubiquitous Computing*, volume 2864 of *Lecture Notes in Computer Science*, pages 200–207. Springer Berlin Heidelberg, 2003.

[14] Cyprien Pindat, Emmanuel Pietriga, Olivier Chapuis, and Claude Puech. Drilling into Complex 3D Models with Gimlenses. In *VRST '13: 19th ACM Symposium on Virtual Reality Software and Technology*, pages 223–230, Singapore, Singapour, October 2013. ACM.

[15] S. Rapp, G. Michelitsch, M. Osen, J. Williams, M. Barbisch, R. Bohan, Z. Valsan, and M. Emele. Spotlight navigation: Interaction with a handheld projection device. In *International Conference on Pervasive Computing*, 2004.

[16] Ramesh Raskar, Paul Beardsley, Jeroen van Baar, Yao Wang, Paul Dietz, Johnny Lee, Darren Leigh, and Thomas Willwacher. Rfig lamps: Interacting with a self-describing world via photosensing wireless tags and projectors. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH '04, pages 406–415, New York, NY, USA, 2004. ACM.

[17] Luigi De Russis, Dario Bonino, and Fulvio Corno. The smart home controller on your wrist. In Friedemann Mattern, Silvia Santini, John F. Canny, Marc Langheinrich, and Jun Rekimoto, editors, *UbiComp (Adjunct Publication)*, pages 785–792. ACM, 2013.

[18] Daisuke Sakamoto, Koichiro Honda, Masahiko Inami, and Takeo Igarashi. Sketch and run: A stroke-based interface for home robots. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 197–200, New York, NY, USA, 2009. ACM.

[19] Dominik Schmidt, David Molyneaux, and Xiang Cao. Picontrol: Using a handheld projector for direct control of physical devices through visible light. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pages 379–388, New York, NY, USA, 2012. ACM.

[20] Thomas Seifried, Michael Haller, Stacey D. Scott, Florian Perteneder, Christian Rendl, Daisuke Sakamoto, and Masahiko Inami. Cristal: A collaborative home media and device controller based on a multi-touch display. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, pages 33–40, New York, NY, USA, 2009. ACM.

[21] Colin Swindells, Kori M. Inkpen, John C. Dill, and Melanie Tory. That one there! pointing to establish device identity. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology*, UIST '02, pages 151–160, New York, NY, USA, 2002. ACM.

[22] Masayuki Tani, Kimiya Yamaashi, Koichiro Tanikoshi, Masayasu Futakawa, and Shinya Tanifuji. Object-oriented video: Interaction with real-world objects through live video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pages 593–598, New York, NY, USA, 1992. ACM.

[23] Karl D.D. Willis, Ivan Poupyrev, and Takaaki Shiratori. Motionbeam: A metaphor for character interaction with handheld projectors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1031–1040, New York, NY, USA, 2011. ACM.

[24] Andrew Wilson and Steven Shafer. Xwand: Ui for intelligent spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 545–552, New York, NY, USA, 2003. ACM.