

Annexe 5

Paper

“Laser-based interaction in a smart home environment”

Mai 2006

Michael Wissen and Guillaume Bernhard

Laser-based interaction in a smart home environment

ABSTRACT

Laser-based interaction devices are suitable only for a limited range of applications, as this kind of interaction requires well adapted user interfaces. However, concerning smart home environments they provide an easy and capable way to address and to interact with any equipment. This paper characterizes a reliable implementation, allowing the laser-based selection of devices in a room as well as the execution of predefined commands.

Categories and Subject Descriptors

H.5.2 [Information Systems]: User Interfaces – *Input devices and strategies*.

General Terms

Human Factors.

Keywords

Laser pointer, recognition, algorithm, smart home, laser-based interaction

1. INTRODUCTION

Several laser-based interaction techniques have been developed in recent years. The intended purpose of this kind of interaction is to control virtual objects on large projection screens and to position the mouse pointer of an operating system respectively.

Nevertheless the suggested solutions are not widely used within enterprises. A study presented in the Pebbles project [4] shows the poor performance of laser pointers when targets are too small. In fact, there are further reasons making a widespread application of a laser-based interaction difficult:

- Rear projection systems are comparatively expensive, the use of high quality cameras is necessary due to weak lightning conditions.
- Despite of low-cost web cameras being sufficient in case of front projection, their application is limited due to the fact that the low frame rate of these cameras doesn't allow fast movements of the laser point. Furthermore,

time and effort to install and calibrate the equipment must also be taken into account.

- For an exact positioning of the laser point, e.g. to control the mouse pointer, some practice by the user is demanded. Carrying out single mouse clicks or double clicks have been identified as problematic, because of the hardness to keep the position of the laser point while pressing a button on the laser pen. In addition, to allow mouse clicks, special pen devices or the recognition of mouse gestures are needed.

Unlike various alternative interaction techniques concerning business situations, the use of a classical laser pen within smart home environments is of particular interest to the simplified interaction with home appliances. Current approaches focus the usage of PDAs or Tablet PCs to control a huge quantity of household appliances. Apart from the fact, that these devices should be permanently reachable by the user, locating and controlling the desired home appliances can only be done within the user interface in several steps.

This paper presents an algorithm for the detection of a laser point in a room with the objective to interact with any equipment or software. The selection of an object can cause the presentation of possible commands related to this object, e.g. on a television screen. As an alternative, mouse gestures on the selected objects can also lead to an execution of the predefined command. We first present the functionality of the laser point detection and discuss its ability to meet the requirements concerning the reliability of the recognition.

2. RELATED WORK

Various attempts have been undertaken in order to create smart home environments by means of direct interaction. Kirstein and Mueller [2] presented a system for direct laser-based interaction with a video projection. Analogously, Wissen [6]. developed a using a laser pen for rear projection.

A further step towards direct interaction with arbitrary objects has been presented by Fails and Olsen [1][5]. Hereby, portions of common everyday objects can be excelled in order to serve as interactive objects. A continuous video surveillance enables arbitrary objects that have not undergone any manipulation to be sensitive to touching or pointing at.

Another way of providing smart environments involving interactive everyday objects has been suggested by Maes [3]. In this case, however, the objects are manipulated and contain invisible technology such as RFID tags or have been enabled to pro-actively show interactive behavior by blinking.

3. LASER RECOGNITION

The fundamental principle of the application is to detect a laser point in a room, its aim to interact with equipment or software. The detection of the laser point can be somewhat problematic: actually, there is a lot of interference that may lead to a wrong detection. Among the small interferences, we can name the brightness variation in a room and the different background colors that a laser point can meet. A black background is a problem for the detection, because it reflects very poorly the color's intensities.

Two other interferences are more problematic: first of all, reflective surfaces such as a glass bottle or door handle in aluminum. These surfaces can present very bright regions, which may interfere with the detection. The last interferences that we will name are the light rays, particularly sun beams. A stray ray of sunlight is a source of luminosity that can occlude a laser point in a room.

The camera that we use for our recognition can record an image every 33,3ms (30 fps), with a YUV 4:2:2 video output format (16 bits/pixel avg.) and a resolution of 640*480 pixels. We equip the camera with a red color filter (650nm) and use a standard laser pointer.

3.1 Algorithm

The laser point detection algorithm is divided into 5 steps that are illustrated in Figure 1:

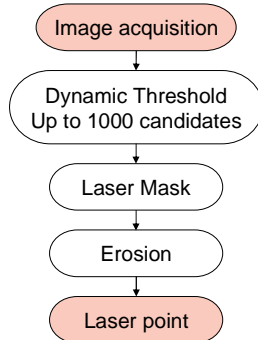


Figure 1. Laser point detection steps.

3.1.1 Image acquisition

The red, green and blue intensity of each pixel of the image recorded by the camera are measured on a discrete level. These intensity measurements are ranged from 0 to 255.

3.1.2 Candidates

An image recorded by the camera contains 307.200 pixels. With the aim to achieve a fast and efficient segmentation, the first step is to apply a filter to select the laser point's best candidates. Because a fixed threshold doesn't make sense in a room where the lighting is changing all day long; a practical solution is then to determine a dynamic threshold, which will select the brighter pixels in the image. This algorithm starts from the principle that a laser point, in the focus of the camera, is formed by two, maybe four pixels, which are brighter than their direct background. The idea is to select at most the N brighter pixels of the image as candidates to be the laser point (for example N=1000). The

selection is made in comparison to the discrete red-histogram of the image. We can actually distinguish two cases:

- Case 1: There are no isolated bright pixels in the histogram. Solution: we take the first 1000 brighter pixels as candidates (see Figure 2).
- Case 2: There are two, maybe three isolated pixels. Solution: we can be sure one of them is the laser point (see Figure 3).



Figure 2. Red-level histogram of the entranced image, without isolated bright pixel.

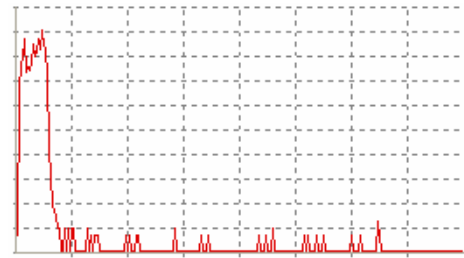


Figure 3. Red-level histogram of the entranced image, with isolated bright pixels.

This idea may be a little bit too restrictive, as written in the preliminary notes: a laser point on a dark surface might not belong to the N brighter pixels of the image. Another idea is then to divide the image into small squares (of dimensions 11*11px for example). In each square, if there is a potential laser point, there should be one or two isolated bright pixels in the red histogram for this square. If there are some, then we select the brighter pixel of this square as a candidate. This process makes it possible to recognize a potential laser point on a black surface.

3.1.3 Filters

Once the candidates are selected, it is important to check if they have all the traits of a laser point or not. For that, the idea is to apply a laser mask to each candidate. A study showed that the red level of the circles around a laser point decreases when we move away from the centre (see Figure 4).



Figure 4. Crude representation of a laser point.

That's why we check which of the candidates respond to these properties.

3.1.4 Erosion

After the two first steps, there may still be more than one candidate. For example, the corner of a door handle in aluminium, which reflects the light of the room, may be included in these candidates. A wrong candidate and the pixels around it have the property not to have a significant red-intensity difference. That's why the last step of this algorithm will be to remove these wrong candidates in making a focus on the region around each candidate. In these regions, a binarization is made with a new threshold. In the current implementation, we choose to define this threshold as $s = s_{\text{dynamic}} - 10$, where the value s_{dynamic} is the threshold defined in 3.2.2. Another idea is to take the red-level value of the candidate and to set the new threshold just under this value: *value-10* for example. The result is then a binary image, in which the white pixels represent the candidate and all the pixels around it, which don't have a significant difference of red level. Because a good laser pixel is quite isolated, two erosions will then make all the white pixels of the region around it disappear, whereas the region around a wrong candidate will still have a few white pixels. Thanks to these two erosions, we can detect which of the candidates is the laser point.

3.1.5 Laser point

The last step of the algorithm is to return the coordinates of the laser point pixel in the focus of the camera.

3.2 Anti jitter

The implementation gives the possibility to remotely command the cursor of a computer with the laser pointer (click, double click). This can be very practical when a screen is projected onto a wall during a presentation. But the natural shaking of the human hand leads to jitter of the mouse on the screen; which can be quite uncomfortable for the user. In [5], Wissen proposed a method with an adapting function between the position of the laser point in the focus of the camera and the position of the cursor on the screen. In this implementation the idea is used again (see Figure 5).

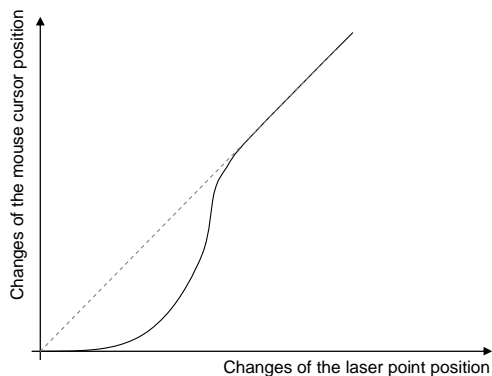


Figure 5. Function for the adaptation from laser point's position to cursor position, to avoid jitter.

The graphic shows how the position of the cursor should be adapted on the position of the laser point from one frame to another. A small movement of the laser point doesn't have an effect on the cursor (up to five pixels in the focus of the camera in our implementation). A larger movement of 5 to 12 pixels from one frame to the next will cause the cursor to move with a dampened effect, as shown in Figure 5. The movement of the cursor and the laser point become identical when the laser point moves more than twelve pixels from one frame to the next.

3.3 Mouse clicks

To create a complete functioning of the cursor, the user must have the possibility to execute a mouse click or double-click just with the laser pointer. A function has then been implemented to fulfil these possibilities. Let's take a look at the left click function: when a user wants to click on a zone of the screen, he has to make a small and fast movement of the wrist from right to left then left to right, as the Figure 6 describes it:

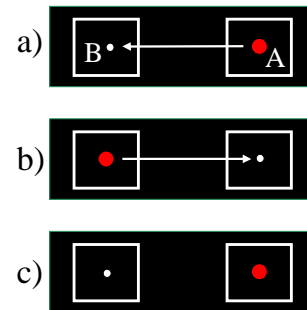


Figure 6. Description of the movement for the click.

In this example, the user wants to click on "A". To achieve this, he has to start the movement from point A and to move quickly to the left, until he reached a square centred on "B". "A" and "B" are 50 pixels apart in the focus of the camera. Then, to finish the click movement, the laser point has to return to the square centred on A. The square's dimensions are 20*20 pixels in our example. The movement is quite easy to execute and to repeat.

The principle of the double click is similar: the steps of the movement are successively left to right and right to left. It is clearly imaginable to design a function for the right click in the same way of the two functions described in this paragraph.

4. ROOM INTERACTION

4.1 Device zones

The main objective of the laser-based interaction technique is to control smart home appliances. Therefore it is necessary to define device zones and to assign various commands to each zone visible to the camera. As an example Figure 7 shows a definition of four device zones: TV, coffee machine and also one area for each roller shade. By selecting one of these areas using the laser pen, either possible commands regarding the selected device are shown on a presentation screen, e.g. a TV, or a command can be executed by a laser pen gesture.



Figure 7. Definition of device zones on the TV, the coffee machine and two roller shades.

4.2 Multiple cameras

The use of a single camera is insufficient in order to control all devices located in a room. In addition, not all sides of an object are visible to the same camera, so that multiple cameras are needed to allow the selection of an object from all accessible sides.

The number of cameras being used with a single computer is limited to the performance of the computers hardware. To allow the use of any quantity of cameras, multiple computers can be installed, each acting as a client and sending the laser point information to the server.

Actually, using multiple cameras, a laser point may be recognized by more than one video capturing device. The actual implementation holds a priority list, covering all active cameras.

5. CONCLUSIONS

Laser-based interaction devices have the potential to become a universal remote control for home appliances. The selection of objects by the user is quite simple while the execution of

predefined commands requires some practice. The next step is to analyze the usability of various pen gestures with the objective of improve their identification. In this respect, a set of up to five or six gestures would be sufficient to provide more operations than activating or deactivating a device.

As mentioned before, the recognition of a laser point on a dark and rough surface is difficult. Classifying areas of different brightness can improve the detection of a laser point by using an adequate threshold for this area.

6. REFERENCES

- [1] Fails, J., Olsen Jr., D. Light Widgets: Interacting in Everyday Spaces. Proceedings of IUI '02, San Francisco, 2002, 63-69.
- [2] Kirstein, C. and Mueller, H. Interaction with a projection screen using a camera tracked laser pointer. *Multimedia Modeling Conference.*, (1998), 191-192.
- [3] Maes, P. Attentive Objects: Enriching People's natural Interaction with Everyday Objects. *ACM interactions magazine*, July / August 2005, 45-48.
- [4] Myers, B., Bhatnagar, R., Nichols, J., Peck, C., Kong, D., Miller, R., and Long, C. Interacting At a Distance: Measuring the Performance of Laser Pointers and Other Devices. *In Proc CHI'02*, (2002), 33-40.
- [5] Olsen Jr., D., Nielsen, T. Laser Pointer Interaction. In: *Conference on Human Factors in Computing Systems*, Seattle, 2001, 17-22.
- [6] Wissen, M. Implementation of a Laser-based Interaction Technique for Projection Screens. *In: ERCIM NEWS*, Number 46 (2002), 31-32