Pointing in the Physical World for Light Source Selection

William Delamare
Université Joseph Fourier, Laboratoire d'Informatique de Grenoble
41 rue des Mathématiques, BP 53
38041 Grenoble Cedex 9
William.Delamare@imag.fr

Céline Coutrix
CNRS, Laboratoire d'Informatique de Grenoble
41 rue des Mathématiques, BP 53
38041 Grenoble Cedex 9
Celine.Coutrix@imag.fr

Laurence Nigay
Université Joseph Fourier, Laboratoire d'Informatique de Grenoble
41 rue des Mathématiques, BP 53
38041 Grenoble Cedex 9
Laurence.Nigay@imag.fr

ABSTRACT

We focus on the selection of light sources in the physical world. Their selection is challenging for the user, since numerous Light-Emitting Diodes (LEDs) can be embedded into various materials as well as environments, thus creating high densities of interactive objects. In this paper, we describe an innovative technique for light sources selection based on the pointing paradigm, that is, allowing interaction at a distance. To address the limitations of the pointing paradigm (e.g., aiming at distant and/or small targets), we design a two-step pointing technique: a rough aiming with an arm pointing gesture and a disambiguation mechanism with a wrist rolling gesture. Feedbacks lean on the various capacities of LED lights. We expect that our technique is well suited for the selection task in dense environments, no matter how small and how distant the targeted light sources are. We also expect that the technique supports an efficient interaction based on proprioception and muscular memory properties for expert users, who may perform the two interaction steps by a single combined gesture for better performance.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces.

General Terms

Design, Human Factors.

Keywords

Physical World Interaction, Pointing Technique, Disambiguation Mechanism.

1. INTRODUCTION

In this position paper, we address the challenge of selection of distant light sources in the physical world. We focus on the selection of lights in dense environments. Indeed with LED (Light-Emitting Diode) and OLED (Organic Light-Emitting Diode) technologies, light sources will be in walls, ceilings, floors, furniture or fabric, so that the environment may have a large number of interactive lights, possibly occluded by other everyday items. Existing switches will be limited for controlling all the LED sources and their various parameters, i.e. color, source, brightness, color temperature, etc. An interesting property of LED lights is the entire range of visual feedback that can be defined as compared to other home appliances, which are often limited to on/off states only. Indeed LED lighting properties (e.g. brightness or color) define various possibilities for feedback while the user interacts with a LED source.

Allowing users to interact at a distance is important, as users can minimize physical effort. When interactive objects are more than 1.1m away from the user, a distant interaction is indeed preferred ([14], [16]). Since one remote controller per light is not an answer in a domestic context, an efficient selection mechanism is needed prior to any kind of interaction.

However, previous studies have reported natural hand tremor and limited human precision as drawbacks for absolute pointing [5]. As a consequence, laser pointer interaction has been proven to be inaccurate, error prone and slow [6]. Users usually do not know where they are pointing at when activating the pointer and they need one second for aiming at a target. Jitters appear when trying to keep a steady position, and drifts of the beam are produced when pressing or releasing a trigger button.

Moreover, in highly flexible domestic situations, users can freely move all the physical objects around. As a consequence, we cannot assume that the computer system has information about the location of the light sources. For this reason, existing pointing techniques, e.g. for large screen [5] or virtual environments [4] are difficult to transpose to the physical world. Techniques like target-aware techniques or spatial reconfiguration of potential targets to improve selection performance in dense environments are not directly applicable to our problem.

In this position paper, we propose a new interaction technique for light source selection in the physical world, which may be generalized to any interactive physical object. Its goal is to allow interaction with distant and/or small targets, partially or totally occluded due to a high density of interactive objects in their surroundings. Before explaining the designed interaction technique, we first review existing solutions for pointing in the real world.

2. EXISTING PHYSICAL POINTING TECHNIQUES

Several studies have explored the pointing paradigm in the physical world. To review the existing solutions for pointing at physical objects, we consider the distinction between the ray casting metaphor and the volume selection metaphor, a distinction well established in the Virtual Reality domain for pointing techniques in the virtual world.

2.1. Ray Casting Metaphor

Users positively perceive the pointing gesture for object selection if the object is out-of-reach but in line-of-sight, when compared to other object selection techniques: *touching, scanning* and *user-mediated*, i.e. when the user specifies an identifier of the object in a graphical user interface [14].

The pointing techniques in the physical world often use a handheld laser as a tool and objects equipped with sensors as targets. These imply typical problems such as cognitive effort and aiming difficulties that are addressed by existing techniques.

The XWand [18] allows physical world interaction by a pointing gesture for the selection and multimodal inputs for the control (i.e. speech and gesture). The user holds a pointing tool that has embedded orientation sensors. Two cameras with IR pass filters are installed in the environment to track the location of the pointing tool. Targets are not equipped with sensors. For the system to know the location of the targets, the physical world is modeled during a training phase: the user has to specify the targets to the system and their location by pointing at them. Since the system considers only known targets, this training phase should be repeated if targets or cameras are moved. The same constraint applies to the laser system proposed for handicapped persons to control their home devices [7]. Since the laser spot is video tracked, interactive objects are not known by the system until the user indicates them to the system. To do so, the user defines active areas on recorded images. If camera or objects are moved, the training process should be performed again.

The laser-based system called WorldCursor [17] allows an accurate aiming in indoor environments. It enhances the XWand system by removing direct camera-based tracking from the installation. It is composed of a platform fixed on the ceiling, and a handheld wand. Instead of directly pointing with a handheld laser, the user indicates a direction with the wand. Then, the platform fixed on the ceiling and composed of rotors and a laser, points at the desired location. By delegating the laser function from the user to the platform, and thus creating an indirection, the system can apply filters to avoid jittering and can thus help the user in precisely aiming at a target. However, the WorldCursor modifies the indoor environment with a platform attached to the ceiling. Moreover interactive physical objects should be in line of sight of the system, and not only of the user. Finally, like XWand, once a model of the physical environment is known by the system, the objects or the system itself cannot be moved without remodeling the environment. This is quite undesirable in a domestic context where users may want to have some freedom.

Another solution was proposed to decrease errors and improve precision, by providing a confirmation step and larger targets [8]. To avoid possible unintentional selections with the laser [6], a two-step button was added to the device. When pressed half way, the laser only is displayed, and when fully pressed, the selection message is sent. To increase the size of the targets, two alternatives are presented. The first is to add a dense set of custom light sensors around targets. The second is to add around the target a cone structure with reflective materials and a light sensor at the back. However, enlarging sensors on target side would be difficult for small objects and is invasive in an everyday environment. In addition, this system still needs an optical line of sight between the user and sensors, even if the user can see the object itself.

To short, ray casting metaphor involves an aiming accuracy problem. Existing solutions to overcome this problem have limitations: (1) a model of the physical world maintained by the system, thus preventing the domestic context from being dynamic, (2) invasive devices, such as large sensors or a platform attached to the ceiling. Techniques based on the volume selection metaphor, that we present in the following section, solves the aiming problem but involves other limitations.

2.2. Volume Selection Metaphor

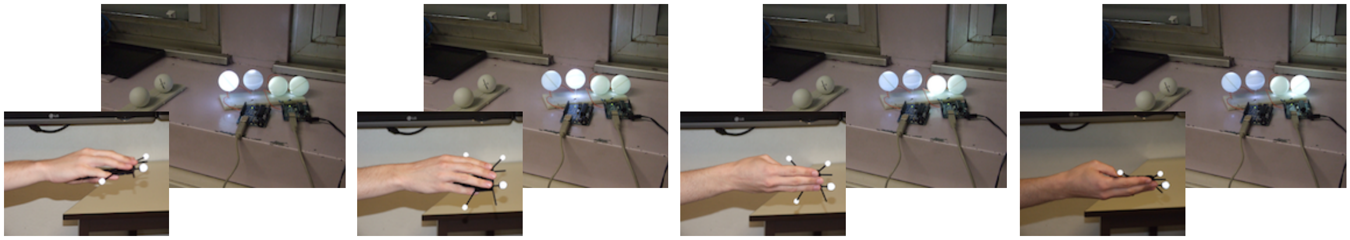
Volume selection is a good alternative to avoid accuracy problems. But while ray casting solutions report aiming problems, volume selection may encounter unwanted multiple selections. For example the GesturePen [15] is based on an Infrared beam, with a 1.5m range and a beam angle of 30°. This setting led to misidentifications during test, that is, users selecting tags neighboring the target. Moreover, the limited range was designed in order to avoid a too large beam span, but users felt uncomfortable when they had to walk close to the tag. This technique illustrates the problem of volume selection metaphor: unwanted selections and multiple selections that imply a disambiguation step for the selection of the target in a selected subset. First we present a technique that automatically performs this disambiguation step. We then consider two techniques for which the user selects the target in the set of selected objects.

The technique described in [10] is based on a smartphone's sensors for orientation and IR cameras for location tracking. Unfortunately, the technique requires a registration process for modeling objects of the physical environment. The technique can support different field-of-view angles and view distances of the selection volume, since this volume is computed by the system and not hardware-dependent. In order to select an object in the selected volume, the technique uses a heuristic mechanism for disambiguation, which assigns a weight to objects depending on the distance between the object and the centerline of the volume. However, users had difficulties for accurate selection with the disambiguation mechanism as the number of objects increased in the selection volume.

Other techniques let the user select the target object in the set of selected objects. Several techniques are based on a graphical user interface on a handheld device: the user selects the target object in a displayed list of selected objects [1] [16]. Those techniques force the user to switch her/his focus from the physical world to the handheld device.

To keep the user in the physical world during the disambiguation step, an RFID (Radio Frequency IDentification) – based system is described in [12]. The one-handed device includes lasers, an RFID reader, buttons and a projector. Targets are equipped with RFID tags. While the reader scans an area, the system supports a precise interaction by automatically modeling tags locations, projecting feedback in the physical world with a stabilized image, and then accurately selecting tags with a laser-pointer in a mouse-style manner, i.e. point-and-click. The technique was designed for environments with a high density of unnoticed tags and an accurate selection in the selected subset (i.e. the disambiguation second step) in the physical world. But the user has to be in line of sight of a specific tag if s/he wants to be precise. Moreover, the second step of the selection is a pointing gesture, with the identified limitations of jittering and aiming problems.

The volume selection metaphor solves the aiming problem, but brings an additional disambiguation step for the selection of the target in the volume. Displaying the list of selected objects on a handheld device forces the user to switch her/his focus from the physical world to the screen where the list is displayed. The technique we present now allows the user to perform the disambiguation step in the physical world but in contrast to other approaches like [12], our disambiguation step is not based on a pointing gesture.



(a) (b) (c) (d)

Figure 1: A pointing gesture turns on the lights of the current selected volume at medium brightness (a), and a rolling gesture (b, c, d) changes the current selected object (at maximum brightness) to the next one.

3. EFFICIENT DISAMBIGUATION MECHANISM

Our technique is based on two interaction steps: (1) a course-grained one where we aim at the area where the desired target is, and (2) a fine one, where we browse the current selection of objects by a roll gesture of the wrist (Figure 1). For feedback during this two-step interaction process, LED lighting offers several types of visual feedback that we will explore experimentally: for instance in Figure 1, the light sources in the selection volume are set to medium brightness and the currently selected light source is at maximum brightness.

With regards to the first step, i.e. the arm movement, after performing the gesture several times, pointing to a given direction can be done without looking at the target ([3], [13]) based on proprioception and muscular memory. Such a property is key in a domestic context where users may want to interact with their surroundings without loosing focus of their current tasks.

We now explain the design of the second step of our technique, i.e. the selection of a target in the pointed volume.

3.1. Rolling as an Input Gesture

In order to select targets that are too close to each other, previous work highlighted that pointing gestures are too difficult [6]. To overcome this difficulty, we consider rolling as an alternative gesture. Indeed, pronation and supination gestures for selection have been shown to be promising [2][11] but have limits in terms of the number of items: 16 items may be selected with visual feedback [11], and 5 for an eyes-free gesture [2].

We fix the range of the rolling gestures based on previous studies. Rahman et al. [11] used a 120˚ range for their rolling gestures while holding a phone, providing a quadratic discretization of the motor space and visual feedback for cursor and targets. Bailly et al. [2] found a maximum range up to more than 300˚, while holding a Wii remote controller. Contrary to [11], the technique is based on a linear discretization of the motor space and no feedback for the cursor is provided. While holding a laser, the comfortable range has been found to be 130˚ [9]. For the comfort of the user, 130˚ seems to be a good candidate for rolling with a free hand. However, as there are differences between rolling ranges identified in the literature depending on the handheld device, we plan to perform a study to experimentally determine the proper comfort range for our technique.

3.2. Overcoming Jitter

While performing the roll gesture, jitter can still occur since the user needs to maintain the arm in a stable position corresponding to the selected volume. Based on small movements of the arm, the selected volume may mistakenly be changed. For this reason, we assign objects with a fixed angular range centered on a fixed angle, in all the neighboring volumes where the object can be selected. In Figure 2, the user selected a volume with an arm position, and an object with a roll position. In Figure 3, the user mistakenly moved his arm position, causing the selected volume to change, but the angular range and center of the roll position remain unchanged, avoiding flickering of the selected object. An alternative solution consists of adding a confirmation step between the volume and the object selection, but possibly increasing the overall selection time.

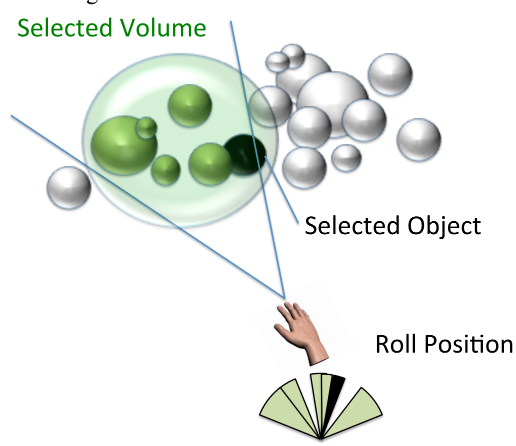


Figure 2: The user selected a volume (in green) with a pointing gesture, and a target (in black) with a rolling gesture.

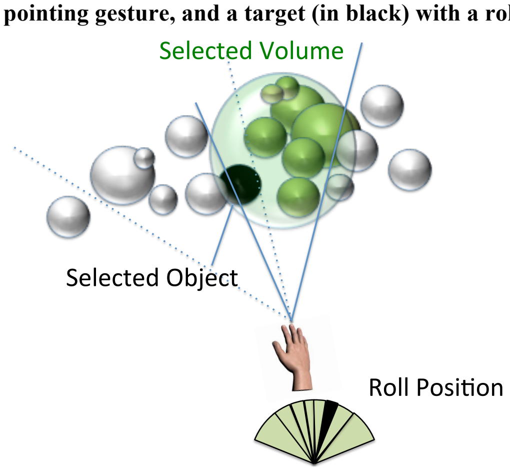


Figure 3: The user mistakenly moved his arm, but the angular range and center of the rolling position remains the same (in black) as in Figure 2.

We expect our solution to allow rapidity and stability in the selection process. In addition, we expect expert users to better remember a constant pointing and rolling position.

4. CONCLUSION

In this paper we presented the design of a technique for selecting distant light sources in the physical world. This technique aims at supporting efficient selection while keeping the users' focus in the physical world. Moreover the users can freely move their light sources since our technique is not based on a model of the physical environment.

We are currently developing the technique with the OptiTrack[1]: six infrared cameras track the instrumented hand (Figure 1), allowing six degrees of freedom. With the developed prototype, we will evaluate the technique itself, focusing on interaction and not on the tracking method for the arm and hand interaction.

We will first experimentally determine parameters of the technique: the hand rolling range, the several types of LED feedbacks and the size of the selected volume. We will determine if the later should be either fixed, dependant on the number of selected light sources or dependant on the distance to the user. Then we will evaluate the usability of our technique for overcoming the problems due to jitter. We will also compare the simultaneous aiming and rolling technique against the sequential aiming and rolling technique. For the case of the sequential aiming and rolling technique, the user will press a button after selecting a set of objects: s/he will then be able to select an object by performing rolling gesture in any arm position (i.e. wrist rotation without the need to maintain the pointing position). We finally plan to assess the benefits and limits of the technique regarding the density of objects, by considering low and high densities, as well as heterogeneous densities of objects. These experiments will allow fine-tuning of the technique, for a future comparison of the technique with existing ones, such as a list of selected objects displayed on a smartphone's screen [1] [16].

5. ACKNOWLEDGMENTS

This work has been supported by the French DELight collaborative project, dedicated to the study of a new lighting system for Solid State Lighting (SSL) applications.

6. REFERENCES

[1] Ailisto, H., Pohjanheimo, L., Välkkynen, P., Strömmer, E., Tuomisto, T. and Korhonen, I. 2006. Bridging the physical and virtual worlds by local connectivity-based physical selection. *Personal and Ubiquitous Computing*. 10, 6 (2006), 333-344.

[2] Bailly, G., Vo, D.-B., Lecolinet, E. and Guiard, Y. 2011. Gesture-aware remote controls: guidelines and interaction technique. *Proc. ICMI '11*, 263.

[3] Cockburn, a., Quinn, P., Gutwin, C., Ramos, G. and Looser, J. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *International Journal of Human-Computer Studies*. 69, 6 (2011), 401-414.

[4] Dang, N.-T. A Survey and Classification of 3D Pointing Techniques. *2007 IEEE International Conference on Research, Innovation and Vision for the Future* (2007), 71-80.

[5] König, W., Gerken, J., Dierdorf, S. and Reiterer, H. Adaptive Pointing–Design and Evaluation of a Precision Enhancing Technique for Absolute Pointing Devices. *Human-Computer Interaction–INTERACT 2009*. (2009), 658–671.

[6] 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. *Proc. CHI '02*, 33.

[7] de la O Chávez, F., Fernández de Vega, F., Olague, G. and Llano Montero, J. An independent and non-intrusive laser pointer environment control device system. *Proc. ICPS '08*, 37.

[8] Patel, S. and Abowd, G. A 2-way laser-assisted selection scheme for handhelds in a physical environment. *UbiComp 2003,* 200–207.

[9] Qin, Y., Shi, Y., Jiang, H. and Yu, C. Structured laser pointer: enabling wrist-rolling movements as a new interactive dimension. *Proc. AVI '10*, 163.

[10] Rahman, A.S.M.M., Hossain, M.A. and Saddik, A.E. Spatial-geometric approach to physical mobile interaction based on accelerometer and IR sensory data fusion. *ACM Transactions on Multimedia Computing, Communications, and Applications*. 6, 4 (2010), 1-23.

[11] Rahman, M., Gustafson, S., Irani, P. and Subramanian, S. Tilt techniques: investigating the dexterity of wrist-based input. *Proc. CHI '09*, 1943.

[12] Raskar, R., Beardsley, P., van Baar, J., Wang, Y., Dietz, P., Lee, J., Leigh, D. and Willwacher, T. RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. *ACM Transactions on Graphics*. 23, 3 (2004), 406.

[13] Reilly, D. Reaching the same point: Effects on consistency when pointing at objects in the physical environment without feedback. *International Journal of Human-Computer Studies*. 69, 1-2 (2011), 9-18.

[14] Rukzio, E., Broll, G., Leichtenstern, K. and Schmidt, A. Mobile interaction with the real world: An evaluation and comparison of physical mobile interaction techniques. *Ambient Intelligence*. (2007), 1–18.

[15] Swindells, C., Inkpen, K.M., Dill, J.C. and Tory, M. That one there! Pointing to establish device identity. *Proc. UIST '02*, 151.

[16] Välkkynen, P., Niemelä, M. and Tuomisto, T. Evaluating touching and pointing with a mobile terminal for physical browsing. *Proc. NordiCHI '06*, 28-37.

[17] Wilson, A. and Pham, H. 2003. Pointing in intelligent environments with the worldcursor. *INTERACT International Conference on Human-Computer Interaction* (2003).

[18] Wilson, A. and Shafer, S. 2003. XWand: UI for Intelligent Spaces. *Proc. CHI '03*, 545

[1] http://www.naturalpoint.com/optitrack/