HAL
open science

Designing Disambiguation Techniques for Pointing in the Physical World

William Delamare, Céline Coutrix, Laurence Nigay

To cite this version:

William Delamare, Céline Coutrix, Laurence Nigay. Designing Disambiguation Techniques for Pointing in the Physical World. Proceedings of the 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2013), 2013, London, UK, pp.197-206, 10.1145/2480296.2480309 . hal-00953354

HAL Id: hal-00953354

https://inria.hal.science/hal-00953354v1

Submitted on 28 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Designing Disambiguation Techniques for Pointing in the Physical World

William Delamare, Céline Coutrix and Laurence Nigay
CNRS, Joseph Fourier University UJF-Grenoble 1
Grenoble Informatics Laboratory (LIG), UMR 5217, Grenoble, F-38041, France
{William.Delamare, Celine.Coutrix, Laurence.Nigay}@imag.fr

ABSTRACT
Several ways for selecting physical objects exist, including touching and pointing at them. Allowing the user to interact at a distance by pointing at physical objects can be challenging when the environment contains a large number of interactive physical objects, possibly occluded by other everyday items. Previous pointing techniques highlighted the need for disambiguation techniques. Addressing this challenge, this paper contributes a design space that organizes along groups and axes a set of options for designers to relevantly (1) describe, (2) classify, and (3) design disambiguation techniques. First, we have not found techniques in the literature yet that our design space could not describe. Second, all the techniques show a different path along the axes of our design space. Third, it allows defining of several new paths/solutions that have not yet been explored. We illustrate this generative power with the example of such a designed technique, Physical Pointing Roll (P2Roll).

Author Keywords
Physical Interaction; Interactive Physical Objects; Design Space; Selection Techniques.

ACM Classification Keywords
D.2.2. Software Engineering: Design Tools and Techniques - User interfaces. H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces - Theory and methods, User-centered design.

General Terms
Design; Human Factors.

INTRODUCTION
Ambient Intelligence involves a wide variety of smart objects. For instance, in a domestic context, there is a large range of physical objects that are digitally augmented: TV, light sources, coffee machine, washing machine, etc. These objects are not necessarily within reach of users in everyday situations. Thus, pointing techniques for distal selection of physical objects are important.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EICS'13, June 24–27, 2013, London, United Kingdom.
Copyright © 2013 ACM 978-1-4503-2138-9/13/06...$15.00.

However, selecting a physical object by pointing raises several problems that are related to the objects' sizes, their density in the physical environment and the occlusion of the targeted object by other items. For defining an efficient pointing technique, designers can draw their inspiration from studies of the Graphical User Interfaces (GUI) and Virtual Environments (VEs) communities, two communities that extensively studied the pointing task. However, the physical environment defines additional constraints: (1) exocentric techniques [16] that modify the user's point of view or techniques that modify the scene [7][9] are not possible, and (2) in the physical world, the system is not omniscient, whereas in virtual environments, the system creates the virtual world. As a consequence, designers cannot assume that the computer system has all the information about the physical surroundings [13].

Capturing these specificities induced by interaction in the physical environment, we present a design space for the disambiguation of the selection of physical objects by pointing at them. The design space adopts two points of view: the characteristics of the interaction techniques (*Interaction Group*) and the requirements for the disambiguation system to be developed (*Disambiguation System Group*). The identified axes in each group are useful for describing a significant range of existing techniques (descriptive power) as well as for helping designers to create new techniques for selecting physical objects (generative power).

In this paper, we first clarify the steps in a pointing task and highlight the disambiguation stage. After a review of related work, we present our design space. Its descriptive power is shown through concrete examples and its generative power through the design of a new disambiguation technique for the selection of a physical target.

POINTING TASK AND DISAMBIGUATION
When physical augmented objects are more than 1.1m away from the user, a distant interaction by pointing at physical objects is preferred [20, 24]. Nevertheless, pointing tasks in the physical world are difficult for two main reasons: accuracy and occlusion.

Previous studies have reported natural hand tremor and limited human precision as drawbacks for absolute pointing [8]. As a consequence, laser pointer interaction has been proven to be inaccurate, error prone and slow [14]. Jitters

appear when trying to maintain a steady position, and drifts of the beam are produced when pressing or releasing a trigger button. Thus, a volume selection (i.e., a cone) can make the selection of physical objects easier by facilitating the aiming action [24] (see Figure 1).

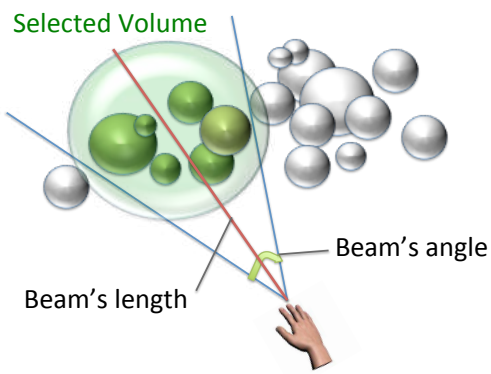


Figure 1: Volume selection: a rough aiming with a pointing gesture. A set of physical objects included in the volume selection (i.e. cone) are selected.

In addition to the accuracy problem, occlusion is another source of difficulty when multiple objects intersect the pointing direction [7].

For the two above cases, volume selection and occlusion, the pointing task requires a disambiguation process.

The finite state machine in Figure 2 models the implementation of a pointing task, including the disambiguation stage. In Figure 2, the initial state corresponds to the state where no physical object is selected. From this initial state, three paths are identified:

- *(t4)* describes a direct pointing, e.g., with an absolute laser pointing system [10, 15, 25, 26]. Direct pointing is possible and efficient for the selection of large physical objects (e.g., a TV) in a non-dense environment.
- *(t5.t6 = t5)* corresponds to an automatic disambiguation mechanism, i.e. performed by the system. The user does not perform any further action in order to reach the final state, that is, the state where a physical object is selected. This is the case for a disambiguation mechanism based on heuristics. Some heuristics rely only on characteristics at the moment of the trigger event (e.g., selecting an object within the selection cone) [17]. Some enhanced approaches define heuristics that are also based on characteristics before the trigger event, e.g., computing scores taking into account the object's size and distance from the selection cone as well as the velocity of the pointing gesture during the entire interaction [22].
- *(t1.t2*.t3)* describes an interactive disambiguation mechanism. Two strategies exist for interactively selecting an object among *n*:
  - *Navigation*: The user cycles through the previously selected set of targets (*t2**) until reaching and selecting the desired one (*t3*). For instance, after a pointing gesture in the physical world (*t1*), the names of the selected objects can be displayed on a handheld device and the user scrolls the list (*t2**) in order to select the targeted object (*t3*) [24].
  - *Designation*: The user performs another designation (*t3*) in the previously selected set of targets. For instance, the RFIG Lamps system [19] includes such a two-step technique. After a first pointing gesture (*t1*), the RFIG Lamps system projects tags onto the selected physical objects. Then, the user selects with a laser the targeted object by pointing at its corresponding tag (*t3*).

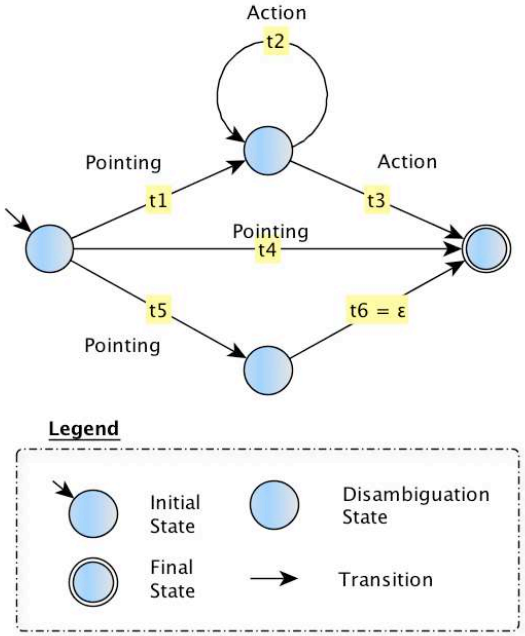


Figure 2: Starting by pointing in the physical world, three paths for the selection of a physical object. (t1t2*t3): pointing gesture followed by an interactive disambiguation process. (t4): single pointing gesture for selecting an object. (t5t6): pointing gesture enhanced by an heuristic method for automatic disambiguation.

Since directly pointing at a physical object (*t4*) reaches its limits when many physical objects can be selected in a dense environment, the selection techniques of a physical object imply a disambiguation stage. An automatic disambiguation stage based on heuristics (*t5.t6*) may lead to the selection of a wrong physical object [7, 17]. Therefore in this paper we focus on techniques that include an initial coarse physical pointing gesture and then an interactive disambiguation stage (*t1.t2*.t3*). We define a design space

for such two-step techniques for selection of a physical object by focusing on the disambiguation stage.

As presented in the following section, a literature review highlights that only five two-step techniques for the selection of a physical object have been developed. Our design space is therefore an analytical tool to empower the designers of interactive techniques in the physical environment, by organizing the design options and by highlighting gaps and under-researched areas.

RELATED WORK

There are two areas of research that are directly relevant to the present paper: (1) the existing two-step selection techniques of a physical object, and (2) the design spaces and taxonomies for the disambiguation stage.

Interactive Disambiguation Techniques for Physical Object Selection

As stated above, only five two-step techniques for the selection of a physical object have been designed. They all include an interactive disambiguation stage after a first pointing gesture. Bold names indicate how we will refer to the corresponding techniques in the remainder of the paper.

GesturePen [23] is a technique using a volume selection metaphor. It is based on an Infrared (IR) beam that is 1.5m long and with a conical angle of 30° (Figure 1). Users can select an IR tag by directing the beam towards a tag. This technique illustrates the problem of volume selection techniques: by easing the pointing task, GesturePen implies unwanted multiple selections. Indeed, GesturePen led to mistakenly selected tags neighboring the targeted object. Nevertheless the limited range of the technique was designed in order to minimize errors due to a too large beam span, but users felt uncomfortable when they had to walk close to the tag. To overcome these difficulties, the authors proposed a disambiguation technique: a dial on the pointing device in order to adjust the beam's length and/or the beam's angle.

Another technique with a similar disambiguation mechanism is **PICOntrol** [21]. With this technique, the user points at a physical object with a visible light projection that activates light sensors attached to physical objects. If several objects are located within the light projection, an envisioned solution is to reduce the projection size.

The Radio Frequency Identity and Geometry Lamps system (**RFIG Lamps)** is a one-handed device that includes a laser, a radio frequency identification (RFID) reader, buttons and a projector [19]. Physical objects are equipped with RFID tags. While the handheld RFID reader scans an area, the system automatically computes the locations of the tags and projects a stabilized interactive graphical presentation of the tags. A user can then more easily disambiguate the selection by pointing with a laser at the projected tags.

While the three above techniques keep the user's actions in the physical world during the disambiguation stage, the two following techniques include a disambiguation stage disconnected from the physical world.

A first solution is to display a list of selected objects on a handheld device [24]. After a coarse pointing gesture in the physical world, if several objects are located within the beam, a **List** of the selected objects' names are displayed on screen and the user selects the targeted object based on its name.

Freeze-Set-Go (**FSG**) proposes an Augmented Reality (AR) technique on a smartphone [11]. The user first places the device so that 2D markers attached to physical targets are visible on the screen through a real-time see-through video. Since several markers can appear on the screen, a touch gesture on the screen is needed for the disambiguation of the selection. In order to ease this disambiguation, FSG provides a custom button that allows users to freeze the video. The real-time see-through video is resumed when the button is pressed again.

With respect to the above five techniques, the authors highlight the need for efficient and usable disambiguation techniques, motivating the need for a design space. We now present the related existing design dimensions.

Design dimensions for Disambiguation Techniques involving 3D Targets in Virtual and Physical worlds

Design dimensions for disambiguation have been studied only for the case of virtual environments. Amongst these studies, the design space [9] for disambiguation mechanisms (selection by progressive refinement) introduces three dimensions:

- The *disambiguation criteria*, characterizing if the disambiguation is *spatial* (e.g., choosing an object based on its location in the scene), *by object attributes* (e.g., choosing red and/or large objects), or *out-of-context* (e.g., choosing objects from a menu).
- The *display* of the selected objects, that can be *in context* (e.g., by freezing the viewpoint as in **FSG** or by relocating targets in order to ease selection) or *out-of-context* (e.g., by using a menu as in **List**). This axis is related to the perceptual continuity criterion as described for Augmented Reality [5].
- The type, characterizing if the disambiguation is:
    - *Discrete*, e.g., if the disambiguation is done through several steps (e.g., *(t1.t2*.t3)* in Figure 2);
    - *Continuous*, e.g., the disambiguation is done through a continuous process implying one continuous action (e.g. *(t5.t6 = t5)* in Figure 2).

A similar axis is found in [7] for describing disambiguation techniques for the case of volumetric displays: the *concurrent/sequential* axis. A *concurrent* (or continuous) disambiguation mechanism means that both pointing and disambiguation tasks are performed at the same time. For instance, with the Depth Ray [7], users control the ray direction and a depth marker along the ray at the same time. Contrastingly, a *sequential* (or discrete) disambiguation mechanism needs a temporal separation between the pointing and the disambiguation tasks. For instance, with the Lock Ray [7], users control the ray direction before locking it so as to then control only the depth marker that appears on the ray.

These design spaces for virtual environments define relevant design dimensions that only partially contribute to our study, i.e. to *disambiguation* techniques for *physical* environments. Nevertheless they lay a foundation for characterizing the design and engineering of the interactive disambiguation stage in the selection of physical targets.

DESIGN SPACE

Our design space is made of 10 axes that are organized according to two groups, *Interaction* and *Disambiguation System*, as shown in Figure 4. The *Interaction Group* characterizes the interaction techniques for the disambiguation stage when selecting physical objects by pointing in the physical world. The *Disambiguation System Group* adopts an engineering point of view by identifying the implications of the *Interaction Group* for the disambiguation system. We first explain the axes of the *Interaction* and *Disambiguation System* groups before concluding on the links between them.

Interaction Group

From an interaction point of view, the designer can operate on two features of the disambiguation mechanism: the display of the targets (i.e. the display space) and the control from the user (i.e. the control space) [2].

Display Space

For the disambiguation phase, the user can (a) focus on the physical targets, or (b) switch her/his focus to a virtual representation (both ends of axis 1 in Figure 4). Between both ends, modifications of the display space can occur in a mixed manner, augmenting the physical target with an additional virtual representation.

Towards the virtual end, augmenting or replacing the physical targets can make the selection easier if selecting the *physical* target is too difficult due to its size, distance or density of objects. For instance, the disambiguation stage of the **RFIG Lamps** system proposes to project a disc onto physical targets for pointing refinement. Indeed the bigger size of the projections as compared to the original tags facilitates the pointing task.

The augmented or virtual representation can be spatially decoupled from the primary physical targets (e.g., the **List** technique, that displays a list of objects' names on a handheld device). However, such a solution divides the users' attention between the physical environment and the virtual representation and creates an additional cost for changing the focus of attention (perceptual discontinuity as described in [5]). To overcome this limitation, the augmented or virtual representation can maintain a spatial link with the physical target (e.g., the **RFIG Lamps** system and an augmented reality technique using the magic lens metaphor such as **FSG**).

For the designer, another way to avoid changes of attention focus between the physical and virtual worlds is to propose a solution towards the physical end of the axis, and allow the user to operate directly in the physical world.

Towards the physical end of the axis, designers will be able in the future to physically modify the display space. For instance, a system such as ZeroN allows actual movements of physical objects in 3D space [12], e.g., to rearrange the targets so that they are easier to point at. Alternatively, Jamming User Interfaces can increase the size of physical objects in order to facilitate the pointing task, while keeping the focus on the bigger primary physical target [6].

Control Space

Designers can help the disambiguation of preselected physical objects by modifying the control, i.e. the actions of the user and the interpretation of these actions by the system. The design approach is to modify the control space assuming that the display space (i.e., physical world) is not modifiable.

To do so, designers can decouple the control from the physical targets by modifying the reference frame (Figure 4, axis 2). Although initially introduced for describing physical pointing gesture [4], the Reference Frame of spatial input can be relevantly applied to disambiguation techniques. The Control Space axis characterizes how the absolute mapping of the pointing step has been transformed for the disambiguation step.

First, if the action of the user is directed to the physical objects, the reference frame is *absolute*. In this case, a disambiguation technique should be based on a Display Space modification (axis 1) in order to ease the pointing task.

Second, if the actions of the user are performed around a user's body part, the reference frame is *relative to the user's body*. The physical objects are abstracted and reorganized according to an origin defined by a body part. For instance the physical objects can be reorganized vertically from the hips to the head. In this example, the user selects a target by moving her/his arm along her/his trunk. The control space hence defines an input space that the user has to mentally map onto the physical world.

Third, if the actions of the user are performed according to the position of a device as a reference point, the reference frame is *relative to the device*. The physical objects are abstracted and reorganized with a neutral position of the device as origin. For instance the physical objects can be reorganized in a clock-manner disposition according to a neutral position of the device. In this example, the user selects a target by rotating the device. Again, the control space defines an input space that the user has to map onto the physical world.

The Control Space axis does not aim to describe in detail all the actions performed by the user: depending on the chosen disambiguation input (e.g., tangible, gesture, gaze, voice command, etc.), previous existing design spaces can further describe specific design characteristics that are not specifically related to disambiguation. Indeed it is an important attribute of our design space that it allows capitalizing on existing design spaces and that it facilitates interconnection between existing approaches.

Disambiguation System Group

We first refine the Extent of World Knowledge [13] of Figure 3. It was introduced by Milgram to characterize how much the system knows about the physical environment. This design dimension defines a continuum from a *world completely un-modeled* to a *world completely modeled*. Between these two extremes is a world partially modeled with *where* and/or *what* data, i.e. localization and identification data. A system knowing only *where* data about an object of interest could precisely place a graphical superimposition onto the physical object. Contrastingly, a system knowing only *what* data could precisely define the augmentation of the physical object, but without correct location, scale or orientation.

Figure 3: Extent of World Knowledge dimension from [13]**.**

We first define two other types of knowledge, dynamic and static, then refines the *what* and *where* subgroups in our design space.

Knowledge Computation: Static vs. Dynamic

As represented by axes 3 and 8 of Figure 4, we distinguish two kinds of knowledge: the one produced during a setup step and the one produced at runtime. For instance **FSG** needs a registration process of 2D markers. This is an imperative step for further video analysis in order for the system to know what the user is selecting with a tap on the video. Contrastingly, the **RFIG Lamps** system first receives dynamic spatial knowledge for the first pointing step using RFID tags and then the spatial locations of the tags for the disambiguation step using light sensors and encoded projections.

On the one hand, knowledge defined during a setup process will constrain the user to configure and maintain the system, e.g., naming or locating objects and additionally updating the information if the user moves, adds or removes physical objects. On the other hand, knowledge computed at runtime is more flexible but less robust. For example, the **RFIG Lamps** system may know the spatial coordinates of pointed objects by projecting grey codes on light sensors attached to objects. However, these sensors might be visually occluded from the projector's viewpoint, whereas the user can actually see the physical objects (the user's viewpoint being different from the projector's viewpoint).

The *What* and *Where* axes can have different kinds of *knowledge computation*, considering each axis one by one. However, a promising design option is the *dynamic* computation of knowledge, avoiding a setup process from the user. Thus, the *what* and *where* axes are globally qualified to be *static* if at least one of the axes is *statically* computed by the system. Consequently the *what* and *where* corresponding axes (3 and 8 of Figure 4) are qualified to be dynamic if all the information is dynamically computed at runtime.

"What" Subgroup

If the designer chooses to provide a *virtual representation* as a Display Space (axis 1 of Figure 4), then *What* information can enhance the cognitive continuity between the physical object and its representation [5]. The continuity at the cognitive level refers to the user's interpretation of the virtual representation. This interpretation is influenced by the *What* information.

The disambiguation system may know different types of information about the interactive augmented physical objects:

(1) *Identity* of the object (axis 4 of Figure 4) that would allow the disambiguation system to present a label "bedside lamp" instead of "object 1" to the user for example.

(2) *Functionalities* of the object (axis 5 of Figure 4) that would allow the disambiguation system to present another information to the user. For instance, if the user is aiming at two light sources but intends to select the one that can turn blue.

(3) *Appearance information* (axis 6 of Figure 4) such as the aspect of the physical object including its form, its color, its texture. Thus, the disambiguation system could reduce the cognitive discontinuity by representing the target with its actual picture for example.

(4) *State information* about the physical object (axis 7 of Figure 4), including its current power consumption, its on/off state, its dimming level. For instance, if the user is aiming at two light sources – one turned on and another turned off – and intends to select the one that is

turned on: the states of the objects are therefore a key characteristic during this disambiguation step.

Each axis can have several levels of knowledge: for instance *Appearance* includes the object's form, its color (either red-green-blue values or grey levels…), etc. However, in order to avoid too fine-grained descriptions that might be difficult to compare, we only consider two discrete values, namely yes/no, along the four axes [*Identity*, *Functionalities*, *Appearance* and *State*] of Figure 4. This representation allows a clear and simple visualization of *What* knowledge systems have, and is sufficient to compare existing techniques.

"Where" Subgroup

As in Milgram's initial axis, the system may have no spatial information about the physical surroundings (i.e., value none in the *Where Subgroup* of Figure 4). If the system maintains spatial information about the physical surroundings (statically or dynamically) we refine the Extent of World Knowledge axis [13] by defining two new axes:

(1) *Relative vs. Absolute* (axis 9 of Figure 4): In a *relative* situation, the system knows how objects are localized compared with each other (e.g., a target is below another). For example, the system could use a simplified map of the physical world (static knowledge) that only contains left/right relationships between objects. The *relative* situation is well adapted to the *navigation* strategy, when the user has to incrementally navigate through preselected objects, from an object to its neighbor. In an *absolute* situation, the system has absolute spatial coordinates in a common reference frame. For example, the **RFIG *Lamps*** system determine their spatial coordinates (x,y) in the video. The absolute situation is well adapted to the *designation* strategy, when the user has to directly designate the desired physical target amongst the preselected physical objects.

(2) *Of the world vs. Of a subpart* (axis 10 of Figure 4): In a *World* situation, the system knows spatial information about all the interactive physical objects. For example, a disambiguation technique using an OptiTrack[1] system, composed of IR tracking video cameras, would allow the system to dynamically know where all the interactive objects are (since they are augmented with reflexive markers). In a *Subpart* situation, the system knows only spatial information about a subset of objects. For example, the **RFIG Lamps** system uses dynamically computed spatial coordinates of the preselected objects only.

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

Links Between the Interaction and Disambiguation System Groups

While describing the 10 axes of our design space, we highlighted the links between the *Interaction Group* and the *Disambiguation System Group*. Indeed, interaction design choices will have an impact on the system to be developed and in particular on the knowledge that the system needs to maintain about the physical objects. But the design space clearly separates the two sets of axes, one set focusing on *interaction* characteristics and one set dedicated to the corresponding *disambiguation system* requirements. Furthermore the *disambiguation system* requirements will in turn have an impact on the technologies to be used for developing the whole system.

Moreover, it is worth noticing that the physical world knowledge may be a constraint for the user. Naming objects, or updating spatial coordinates may be a burden for the users. On the contrary, providing an autonomous system (i.e. value *dynamic* along the axes 3 and 8 of Figure 4 *Knowledge Computation*) might imply a heavy infrastructure and/or might be expensive or simply not feasible for real-world deployment.

Figure 4 illustrates how to use the design space. In order to be able to compare techniques based on the design space, we assume that a proper and complete description of each technique is available. A disambiguation technique for the selection of a physical object defines a path in this multidimensional space. The 10 axes of the design space enable us to precisely define an existing technique by a path in the design space or to design a new one by exploring an unexplored area of the design space. The number of design possibilities defined by the 10 axes is huge (2673 possible paths). In light of our design space, techniques are said to be similar if they have the same characteristics (i.e., the same path) and different if they have at least one difference on one axis. We do not provide any measure yet since we are not able to quantify the importance of the axes. Evaluation of the impact of the axes on usability is future work. The following section describes the existing techniques as paths in the design space (i.e., the descriptive power of our design space), while the next section presents a new technique for which the design is based on the 10 axes (i.e., the generative power of our design space).

DESCRIPTIVE POWER OF THE DESIGN SPACE

Both **PICOntrol** and **GesturePen** propose the same disambiguation mechanism: adjusting the selection volume. Thus, the display space is in the *physical* world, with an *absolute* reference frame since the interaction is the same as the one used for the pointing step, but with a different volume of selection. **GesturePen** proposes to *dynamically* get knowledge of the identity of selected objects using custom tags with an IR transceiver. **PICOntrol** does not have any knowledge about the spatial physical environment. The system uses light sensors attached to

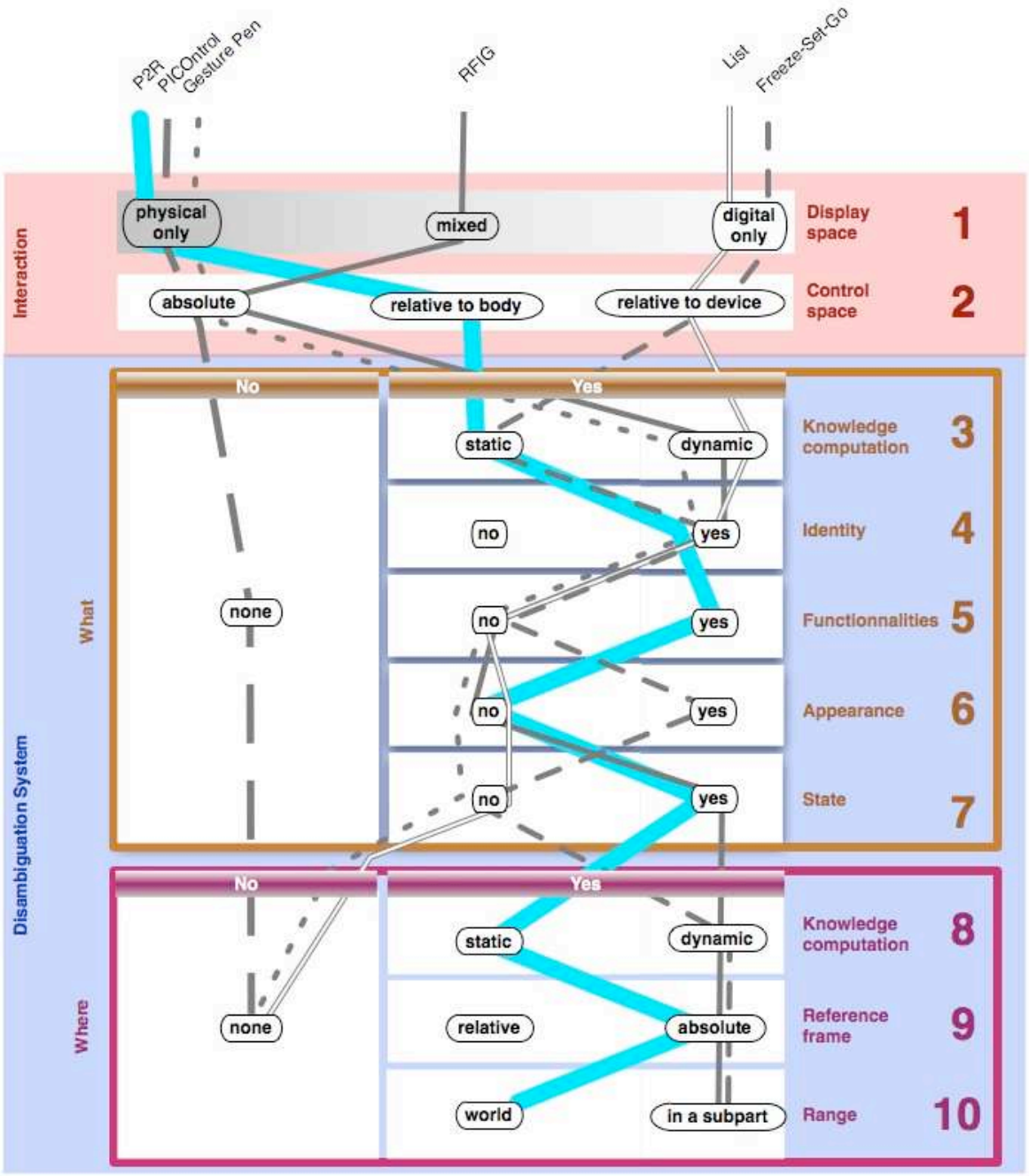


Figure 4: Design Space with 10 axes organized according to two groups: Interaction and Disambiguation System. Classification of existing disambiguation techniques (thin lines). A technique (P2Roll) designed by drawing a new line (under-researched area) in the design space (thick line).

objects in order to directly control the objects, and thus does not need any knowledge about them.

The **RFIG Lamps** system provides a *mixed* display space for the disambiguation step (digital objects projected onto physical ones). The reference frame is *absolute* since the user has to aim these projections with a laser. The system receives *dynamic* knowledge (RFID tags for *what* information and encoded projection for *where* information) that provides *identity* and *state* information (for the warehouse scenario) and *absolute* tags' locations in the current pointed area – *subpart* of the world.

FSG provides a *virtual* display space for the disambiguation step, i.e. after pressing the button that creates a frozen image of the video, breaking the link with

the physical world. The touching gesture for selecting the target is in a reference frame *relative to the device* since the user has to interact with the screen. In order to map the touch input and the content of the image, the system has *static* knowledge of 2D markers, using video-analysis techniques based on the *appearance* of the objects (markers, linked with objects' *identity*). This *What* knowledge about the physical world allows the disambiguation technique to *dynamically* obtain *absolute* coordinates of 2D markers in the video – *subpart* of the world.

The **List** of objects' names displayed on a handheld device defines a *virtual* representation as a display space. The user interacts on the screen of the handheld device, thus *relatively to the device*. The system receives *dynamic* knowledge of the *identity* of objects that are pointed at by the user with the IR beam. No more information is needed in order to disambiguate the selection.

GENERATIVE POWER OF THE DESIGN SPACE

Defining a new path in the design space allows designers to produce a new interaction technique. Based on the 10 axes, many design choices are possible. To illustrate this generative power of the design space, we now explain the design of a new technique, represented by a thick line in Figure 4.

We focus on the selection of light sources in the physical world. Indeed with Light Emitting Diode (LED) and Organic Light Emitting Diode (OLED) technologies, light sources will be in walls, ceiling, floors, furniture or fabric, so that the environment may have a very large number of interactive lights, possibly occluded by other everyday items. This context fits well with the topic addressed by the design space since it brings several problems that designers should address while designing selection techniques of light sources by pointing in the physical world. Moreover, this example could easily be transposed to other contexts since a classic way to augment physical objects for visual feedback is to attach LED to them [21, 23].

A design approach based on the generative power of our design space has led to the following disambiguation technique: a navigation strategy called Physical Pointing Roll (P2Roll). The user first performs a pointing gesture towards the light sources. This defines an initial coarse physical pointing phase. The user then has to precisely select the targeted light source from within the set of selected light sources. To do so, the user navigates through the preselected light sources using a roll gesture of the wrist, as illustrated in Figure 5.

First we made the design choice to keep the focus of the user in the physical world (value *Physical* on axis 1 of Figure 4) so that the user can directly observe LED lights, and more importantly the effect on the ambient lighting.

Concerning the control space, the wrist rolling input covers an unexplored area, with the reference frame having the value *Relative to Body* along the axis 2 of Figure 4. We had to choose an adequate body part as the origin of the reference frame. Since the arm is already used for the rough pointing gesture, we decided to apply the physiologic chain of the arm [1, 3]: Shoulder and elbow are used for the coarse physical pointing task, thus we chose a wrist input. Since the pronation/supination axis (roll) has been proven usable [18], our disambiguation technique is based on a *relative to the wrist* transformation.

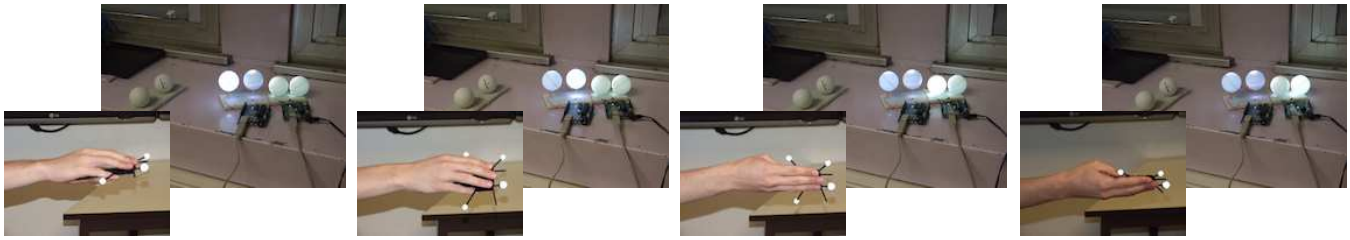
Having defined the P2Roll *Interaction Group*, we focus on the *Disambiguation System Group*. We make the distinction between two different systems: the prototype system of Figure 5 that we developed in order to experimentally evaluate the interaction technique and the final system that we will develop for real-world deployment if P2Roll is validated by the in-lab experiments.

Our current lab equipment allowed us to rapidly build a prototype in order to evaluate the P2Roll technique. We used the OptiTrack[2] with six IR cameras that track the instrumented user's hand as shown in Figure 5. Since our LED objects are very close to each other to simulate a very dense physical environment, the OptiTrack cannot be used for tracking LED objects. Thus, the developed system maintains a *static* (axes 3 and 8 of Figure 4) and *absolute* (axis 9 of Figure 4) knowledge of the entire physical *world* (i.e., the LED objects). Since LEDs are controlled by the system itself, the system keeps the *identity* (axis 4 of Figure 4), the *functionalities* (axis 5 of Figure 4), and the *state* on/off (axis 7 of Figure 4) of each LED object.

However, this lab configuration is not easily deployable in the real world: it has a heavy infrastructure for tracking purpose and a significant setup/maintenance process in a dynamic context in which objects can be moved. Thus, if the interaction technique is experimentally validated, we will design another system solution (*Disambiguation System Group*) for the engineering of the designed disambiguation technique.

As part of a collaborative project with a hotel chain, the final setting is a hotel room, designed with LED light sources in walls and furniture. For this particular real-world context, we can assume that the system can have a *static what* and *where* knowledge, the hotel rooms having the same fixed furniture and light sources. However, the key design issue related to the *Disambiguation System Group* is the tracking solution for the pointing and rolling gestures that compose P2Roll. Several design issues based on the axes of our design space are possible, we describe two technological design options: (1) RFID tags hidden in the physical environment for the pointing step that allow the system to *dynamically* receive objects' identities, and then

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



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

Figure 5: 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.

obtain *static* knowledge (e.g., a list of objects ordered from right to left and defined during the setup) for *relative* spatial information amongst the set of preselected objects. (2) Finding a way to communicate with LEDs by switching them on and off at high frequencies not perceived by users but only by the system with a high-frequency video camera. This could lead to *dynamic relative* spatial knowledge of preselected objects (i.e. *subpart* of the world).

CONCLUSION

This paper has introduced and described a design space for pointing in the real world focusing on disambiguation. This novel conception of the design space is intended to support design and research in the field of physical interaction. In contrast to a technology-centered design approach, the design space has introduced a new way of thinking of interaction design for the selection of physical objects in terms of interaction characteristics and their implications for the system.

A first contribution of the paper has been to demonstrate the descriptive power of the design space by precisely classifying existing techniques. This is ongoing work since new techniques are continuously defined. Till now this was successful because the existing techniques show different paths in the taxonomy and we have not found techniques in the literature yet that our design space left out. Moreover the design space capitalizes from existing design spaces of different domains. In particular for characterizing interaction, the design space refines the two main approaches (based on the display and control spaces) that are studied for enhancing pointing tasks in GUI [2] ; and for characterizing the system, it extends existing taxonomies dedicated to Augmented Reality [13]. This contributes to demonstrating the soundness of the underlying concepts of our design space.

A second contribution of the paper has been to demonstrate the generative power of the proposed design space: it allows defining of several new paths/solutions in the taxonomy that have not yet been explored. Although we presented here only one example of such a designed technique, namely the Physical Pointing Roll (P2Roll) technique, this is another form of validation of our design space. The P2Roll technique corresponds to an under-researched area of the design space.

As ongoing work, we are focusing on the design and the evaluation of other techniques in order to measure the benefits of the values along the axes of the design space. In particular we are currently designing techniques for selecting LED light sources, which will be deployed in a hotel room as part of a collaborative project involving a hotel chain. This will allow further testing of the design space.

ACKNOWLEDGMENTS

This work has been supported by the DELight project (French government's FUI -Single Inter-Ministry Fund-program, certified by the cluster Minalogic). The project is dedicated to the study of a new lighting system for Solid State Lighting (SSL) applications and is led by Schneider Electric.

REFERENCES

1. Balakrishnan, R. and MacKenzie, I.S. Performance differences in the fingers, wrist, and forearm in computer input control. In *Proc. CHI 1997*, ACM Press (1997), 303–310.

2. Balakrishnan, R. “Beating” Fitts’ law: virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies 61*, 6 (2004), 857–874.

3. Card, S.K., Mackinlay, J.D., and Robertson, G.G. A morphological analysis of the design space of input devices. *ACM Transactions on Information Systems 9*, 2 (1991), 99–122.

4. 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.

5. Dubois, E., Nigay, L., and Troccaz, J. Assessing continuity and compatibility in augmented reality systems. *Universal Access in the Information Society 1*, 4 (2002), 263–273.

6. Follmer, S., Leithinger, D., Olwal, A., Cheng, N., and Ishii, H. Jamming User Interfaces: Programmable Particle Stiffness and Sensing for Malleable and Shape-

Changing Devices. In *Proc. UIST 2012*, ACM Press (2012), 519–528.

7. Grossman, T. and Balakrishnan, R. The design and evaluation of selection techniques for 3D volumetric displays. In *Proc. UIST 2006*, ACM Press (2006), 3–12.

8. König, W., Gerken, J., Dierdorf, S., and Reiterer, H. Adaptive Pointing–Design and Evaluation of a Precision Enhancing Technique for Absolute Pointing Devices. In Proc. *INTERACT 2009*, 658-671.

9. Kopper, R., Bacim, F., and Bowman, D.A. Rapid and accurate 3D selection by progressive refinement. *In Proc. 3DUI 2011*, IEEE Computer Society (2011), 67–74.

10. 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. In *Proc. ICPS 2008*, ACM Press (2008), 37–46.

11. Lee, G.A., Yang, U., Kim, Y., et al. Freeze-Set-Go interaction method for handheld mobile augmented reality environments. In *Proc. VRST 2009*, ACM Press (2009), 143–146.

12. Lee, J., Post, R., and Ishii, H. ZeroN: Mid-Air Tangible Interaction Enabled by Computer Controlled Magnetic Levitation. In *Proc. UIST 2011*, ACM Press (2011), 327–336.

13. Milgram, P. and Kishino, F. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems E series D 77*, 12 (1994), 1321–1321.

14. Myers, B.A., Bhatnagar, R., Nichols, J., et al. Interacting at a distance: measuring the performance of laser pointers and other devices. In *Proc. CHI 2002*, ACM Press (2002), 33–40.

15. Patel, S. and Abowd, G. A 2-way laser-assisted selection scheme for handhelds in a physical environment. In Proc. *UbiComp 2003*, Springer (2003), 200–207.

16. Poupyrev, I. and Ichikawa, T. Manipulating objects in virtual worlds: categorization and empirical evaluation of interaction techniques. *Journal of Visual Languages & Computing 10*, 1 (1999), 19–35.

17. Rahman, A.S.M.M., Hossain, M.A., and Saddik, A.El. 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.

18. Rahman, M., Gustafson, S., Irani, P., and Subramanian, S. Tilt techniques: Investigating the Dexterity of Wrist-based Input. In *Proc. CHI 2009*, ACM Press (2009), 1943–1952.

19. Raskar, R., Beardsley, P., Van Baar, J., et al. RFIG Lamps: Interacting with a Self-DescribingWorld via Photosensing Wireless Tags and Projectors. *ACM Transactions on Graphics 23*, 3 (2004), 406–415.

20. 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*. Springer Berlin Heidelberg, 2007, 1-18.

21. Schmidt, D., Molyneaux, D., and Cao, X. PICOntrol: using a handheld projector for direct control of physical devices through visible light. In *Proc. UIST 2012*, ACM Press (2012), 379–388.

22. Steed, A. Towards a General Model for Selection in Virtual Environments. In *Proc. 3DUI 2006*, IEEE Symposium on (2006), 103–110.

23. Swindells, C., Inkpen, K.M., Dill, J.C., and Tory, M. That one there! Pointing to establish device identity. In *Proc. UIST 2002*, ACM Press (2002), 151–160.

24. Välkkynen, P., Niemelä, M., and Tuomisto, T. Evaluating touching and pointing with a mobile terminal for physical browsing. In *Proc. NordiCHI 2006*, ACM Press (2006), 28–37.

25. Wilson, A. and Pham, H. Pointing in intelligent environments with the worldcursor. In Proc. *INTERACT* 2003, 495-502.

26. Wilson, A. and Shafer, S. XWand: UI for Intelligent Spaces. In *Proc. CHI 2003*, ACM Press (2003), 545–552.