Radu-Daniel Vatavu. (2017). Smart-Pockets: Body-Deictic Gestures for Fast Access to Personal Data during Ambient Interactions. *International Journal of Human–Computer Studies 103*, 1-21. http://dx.doi.org/10.1016/j.ijhcs.2017.01.005

Smart-Pockets: Body-Deictic Gestures for Fast Access to Personal Data during Ambient Interactions

Radu-Daniel Vatavu

MintViz Lab | MANSiD Research Center
University Ştefan cel Mare of Suceava
720229 Suceava, Romania

Abstract

This work introduces Smart-Pockets, a new set of whole-body gesture recognition techniques that enables users to access their personal digital content efficiently for visualization on ambient displays. Smart-Pockets works by recognizing users' *body-deictic gestures* entailing access to their pockets, for which associations between specific pockets and personal digital content anchored to those pockets has been managed a priori. The "pocket metaphor" that we explore in this work enables links to digital content using physical personal containers (*i.e.*, pockets) placed at convenient locations on the user's body, containers that have been specifically devised over decades of fashion design to store and carry people's personal belongings comfortably and conveniently. Consequently, Smart-Pockets gestures are fast, require absolutely no precision to perform effectively, and are robustly recognized in user-independent scenarios with absolutely no training required from the user of the ambient display. Also, the Smart-Pockets technique is flexible and easily extensible to other physical containers, such as bags and hand-held objects, which we demonstrate in the form of Smart-Containers. We evaluate the accuracy of several techniques for recognizing Smart-Pockets access gestures, for which we report +99% accuracy for user-independent classification and explicit segmentation. We discuss users' kinematic performance with Smart-Pockets and Smart-Containers and show that the average pocket access time of 2.2 seconds is comparable to the average production time of touch gestures on smart mobile devices and is much smaller than the time required to produce other whole-body gestures. Beyond their practical implications for advancing knowledge in gesture-based interface design for ambient interactions, we believe that the contributions introduced by the Smart-Pockets concept will also foster new developments by pointing the community attention toward (i) more examination of the potential of a new class of whole-body gestures, *i.e.*, body-deictics, (ii) more attention toward how users access their personal digital content on public displays, an important preliminary step before actual interaction, and (iii) inspiring work in the community to examine new and

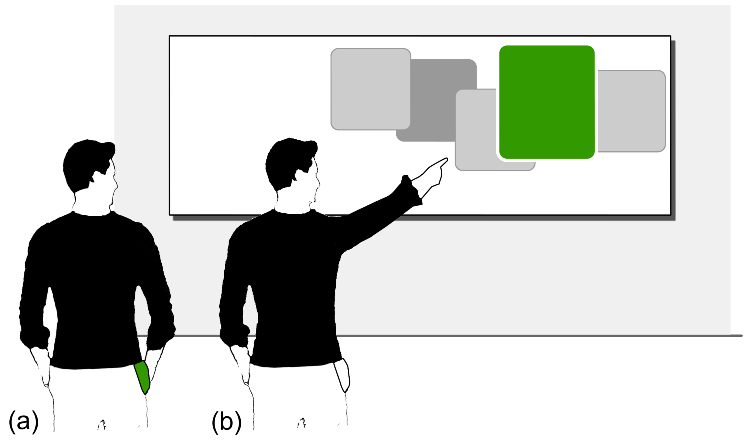

Figure 1: Our vision for Smart-Pockets: users reach into their pockets (a) to link to personal digital content (*e.g.*, calendar, email list, etc.) that they wish to access. When pointed toward the ambient display (b), the content is automatically transferred and visualized. Smart-Pockets act as placeholders for their users' personal digital content.

creative associations between users' physical personal objects and their digital content visualized on ambient displays.

Keywords: body-deictic gestures, pointing, whole-body gestures, ambient displays, pockets, containers, gesture interaction, gesture recognition

1. Introduction

Ambient displays have become almost pervasive in public settings, affording their audience great opportunities to visualize digital content in a large variety of forms, contexts, and application scenarios (Ardito et al., 2015; Börner et al., 2013). Interactive content presented on ambient displays enables users to control and share information by transitioning between various interaction zones and phases (Vogel and Balakrishnan, 2004; Michelis and Müller, 2011; Müller et al., 2010). However, such systems still need appropriate interaction techniques for users to control content in fluent and intuitive ways. So far, gesture-based interaction has been examined by the research community as one practical solution to address this need, with focus on making touch, pointing, and whole-body gestures effective and efficient for controlling content during ambient interactions (Castellucci et al., 2013; Müller et al., 2014; Jota et al., 2010; Nancel et al., 2015; Vogel and Balakrishnan, 2005; Dingler et al., 2015; Vatavu, 2012a; Walter et al., 2013).

However, in the context of public ambient displays, for which interactions are usually brief and to-the-point, getting to the content one needs (*e.g.*, personal files) represents an important, frequently occurring task. Depending on how easy it is to retrieve that content, the efficiency of such a preliminary operation may affect the efficiency and the user experience of the entire interaction process. In this work, we introduce Smart-Pockets, a gesture-based technique that implements the "pocket metaphor": users reach for their pockets in order to link to and access their personal digital content stored elsewhere (*e.g.*, in the cloud), which is transferred to and visualized on the ambient display. Figure 1 shows a visual illustration of the Smart-Pockets concept. Smart-Pockets can provide instant access to one's daily agenda that is relevant to a specific context, such as the user's location; retrieve documents that the user wishes to upload to the display and make public, just like posting or handing out a flyer that was kept in the pocket; or they can act as links to multimedia content that the user wishes to resume playing from the small-screen personal device to a large display. Smart-Pockets can also help users on the go, who cannot afford stopping and reading the information provided by the public display, with a fast way to retrieve and store that information for later scrutiny. Various pockets may be associated with different types of digital content, such as the email input box could be linked to the outer breast pocket, the calendar with the trousers rear pocket, a menu of additional options to the inner breast pocket, and so on. By implementing associations to frequently used digital content, the Smart-Pockets technique comes to address our dependency on carrying mobile technology, such as smartphones, tablets, etc., which made the fashion industry adjust to fit this need; see, for example, the SCOTTeVEST vest designs,[1] some with as many as 18 pockets, including the "cellphone pocket," the "tablet pocket," the "camera pocket," as well as bud buckets that come incorporated in the vest.

Smart-Pockets gestures represent a hybrid between deictic and body-referenced gestures, which are executed in an integrated fashion and posses unique meaning. Such hybrid deictic and body-deictic gestures have not been examined so far in the literature of gesture-based interaction. However, as we show in this article, Smart-Pockets gestures are fast to perform and are robustly recognized with no training required from the user, *i.e.*, user-independent recognition. Note that we are not interested in this article in user authentication aspects on public ambient displays, as these topics have been examined thoroughly elsewhere (Roalter et al., 2013; Wilson and Sarin, 2007; Patel et al., 2004; Aumi and Kratz, 2014; De Luca et al., 2009); equally, we are not interested in how people associate content and physical locations in space, as such experiments have been conducted many times in the mnemonics literature (Ängeslevä et al., 2003; Bellezza, 1996; Guerreiro et al., 2008; Strachan et al., 2007). Instead, our interest in this work is on providing a new interactive experience to the users of public ambient displays, once some form of authentication has been established, so that users would be able to access and visualize their personal digital content fast and intuitively by exploiting the pocket metaphor. In our vision, *Smart-Pockets act as placeholders or links to their users' personal digital content, facilitating fast access to that content, similarly to how conventional pockets facilitate access to one's personal belongings.*

The contributions of this article are as follows: (1) we introduce the concept of Smart-Pockets that employs body-deictic and deictic gestures to enable fast access to one's personal digital content using the "pocket metaphor"; (2) we present three techniques for detecting and recognizing Smart-Pockets gestures from data streams of whole-body gesture movement, for which we report +99% accuracy for a set of 20 distinct gestures under user-independent training and explicit segmentation; (3) we extend the Smart-Pockets concept to generic containers, such as clothing accessories and bags, for which we conduct a second evaluation consisting of 28 distinct Smart-Pockets and Smart-Containers

[1] SCOTTeVEST, Multi-pocket clothing for travelers, gadget lovers, photographers, and people on the go, `http://www.scottevest.com`

access actions; and (4) we evaluate the kinematics and users' perceived experience of Smart-Pockets and Smart-Containers and we discuss implications for practitioners and designers of gesture-based user interfaces for interactive public ambient displays.

2. Related work

We discuss in this section prior work on gestures for ambient interaction as well as techniques for interacting on and around the human body to position our Smart-Pockets technique in the context of relevant research in the community. We review related work regarding smart garments that recognize human activity in context and we point to connections between Smart-Pockets and tangible user interfaces. A history of pockets from the perspective of fashion design and pockets evolution over time in terms of functionality and style, including recently augmented pockets with sensing technology, completes this section.

2.1. Gesture interaction with ambient displays

Gesture input has been explored extensively by researchers and practitioners for designing interactions for ambient displays; see (Castellucci et al., 2013; Jota et al., 2010; Nancel et al., 2015; Vogel and Balakrishnan, 2005; Dingler et al., 2015; Haque et al., 2015; Jurmu et al., 2013; Müller et al., 2014; Vogel and Balakrishnan, 2005; Vatavu, 2012a; Walter et al., 2013). Previous work has considered gesture commands of many types, ranging from pointing and close-up touch gestures (Müller et al., 2014; Vogel and Balakrishnan, 2005) to elaborate whole-body movements to effect actions on remote displays (Walter et al., 2013; Vatavu, 2012a). Gestures are appealing for interacting with ambient displays because (i) gesture input makes supplementary devices superfluous, enabling users to interact instantly, without the need to carry and fetch other devices and (ii) gestures are flexible and adaptable to a wide range of interaction scenarios: from remote interactions implemented with pointing (Vogel and Balakrishnan, 2005) to interactions performed at medium distance (Müller et al., 2014; Walter et al., 2013; Vatavu, 2012a), and to close-up interactions supported by touch input (Vogel and Balakrishnan, 2004; Müller et al., 2014). Users' natural transition between various interaction zones and gesture input modalities has been remarked and examined in the literature before (Jurmu et al., 2013; Müller et al., 2014; Vogel and Balakrishnan, 2004). In this direction, Dingler et al. (2015) even proposed a technique to adjust the interaction modality based on users' proximity to the ambient display, and examined four interaction zones corresponding to touch, fine-grained, general, and coarse gestures.

Deictic gestures (or pointing) have been studied substantially in the literature as they deliver simple and effective means to perform selections: users simply point with their hands to the on-screen items they wish to operate (Bolt, 1980; Jota et al., 2010; Vogel and Balakrishnan, 2005; Haque et al., 2015; Nancel et al., 2015). For instance, Vogel and Balakrishnan (2005) introduced three techniques for pointing (*i.e.*, finger ray, relative pointing with clutching, and a hybrid technique combining relative and absolute pointing) and two techniques for confirming selection with finger gestures (AirTap, a "down and up" gesture of the index finger, and Thumb Trigger, an "in and out" gesture of the thumb). Jota et al. (2010) evaluated several variants of ray pointing for large displays (*i.e.*, laser, arrow, image-plane, and fixed-origin) and reported that techniques that rely on rotational control perform better for targeting tasks, while techniques with low visual parallax (*i.e.*, the distance between the center of the device and the point of view of the user) perform better for tracing tasks. In a study that compared ray casting with absolute position mapping, Castellucci et al. (2013) found that pointer movements were more accurate and smoother in terms of depth variability, rotation movement ratio, and rotation direction change. Avellino et al. (2015) examined the accuracy of pointing gestures in terms of distance and angle errors and reported that eye gaze alone is more accurate than hand gesture input, and that the relative position between the viewer and the on-screen video feed has little effect on accuracy. Working with very high resolution displays, Nancel et al. (2015) showed that traditional pointing techniques are inefficient for precision pointing on ultra-high resolution and proposed a tunable acceleration function and a framework for dual-precision input for such displays. Haque et al. (2015) showed that consumer-level electromyographic and inertial motion unit sensors, such as the Myo armband[2], are practical for free-hand pointing and clicking on a remote display.

Researchers have created design spaces for gestures and ambient interactions. For instance, Walter et al. (2014) introduced a gesture design space to implement item selection organized across five dimensions: selection gesture, confirmation gesture, input space, layout, and user representation. The selection gesture enables users to browse available options on-screen, *e.g.*, a swipe gesture of the hand effects a change in the current selection; see (Vatavu, 2012a) for an example. The selection is then committed with a confirmation gesture implemented as dwell, push, grasp, wave, a second point toward the selected item (Walter et al., 2014) or with expert-designed hand gestures (Vogel and Balakrishnan, 2005). The input space designates the specific volume around the user in which gestures are performed; see (Vatavu, 2012b) for body-referenced gestures proposed by users during a gesture elicitation study. The layout in which items are arranged on screen determines the type of gesture input that may be used to access those items. User representation can be as simple as a hand cursor or can be more elaborate and take the form of an avatar or of a body silhouette mirroring the actual image of the user (Vatavu, 2012a, 2015; Walter et al., 2013).

Designing gesture input for interacting with remote displays must consider important aspects that influence user experience, such as the type of feedback provided to users, the design of the gesture set, and ways to deal with fatigue that may install after repeated arm movement. Providing appropriate feedback during gesture interaction is important to let users know that their gestures have been properly detected and understood by the system. To this end, Vogel and Balakrishnan (2005) proposed visual and auditory cues to compensate the lack of kinesthetic feedback during free-hand gesture interaction, while other researchers explored electrical muscle stimulation and vibrotactile feedback (Pfeiffer et al., 2014; Schönauer et al., 2015). For instance, Schönauer et al. (2015) found that people are approximately 80% accurate at recognizing various patterns and amplitudes of vibrotactile feedback applied to the wrist of the hand when performing mid-air gestures. Interacting with hands in mid-air becomes fatiguing after some time, so researchers have also looked at techniques to reduce fatigue. For instance, Bailly et al. (2011) introduced the "Finger-Count" menu that enables users to select items on a remote screen by mapping the number of raised fingers to menu options. Liu et al. (2015) introduced

[2] `https://www.myo.com/`

“Gunslinger,” a mid-air gesture interaction technique for interacting with large displays that keeps the user’s arms in a relaxed position along the body. Vatavu and Zaiţi (2014) examined users’ preferences for low-effort free-hand gestures during a gesture elicitation study for lean-back gesture input for TV and suggested guidelines for designing low-effort gestures for generic remote displays.

Gesture commands are not self-revelatory, so careful design must be considered to inform users about the specific gesture types that the system was trained to recognize. In this direction, Walter et al. (2013) introduced several strategies to reveal gesture commands on public ambient displays by implementing concepts such as spatial division, temporal division, and integration. Vatavu (2012a, 2011) addressed the gesture discoverability problem by introducing the “nomadic gestures” concept. Nomadic gestures are gesture sets that reside on their users’ smart devices and are uploaded to the interactive display just before the interaction. The display performs a best match between users’ gesture descriptions and the available functions in the system, after which users can interact with the display with the gesture commands they already know and have practiced before.

Researchers also looked at ways to infer users’ intent to interact with public displays by sensing users’ body postures, actions, and positions relative to the display. For instance, Huber et al. (2015) examined feet positions as indicators of the intention to interact with the nearby display. Tanase et al. (2008) and Annett et al. (2011) instrumented interactive tabletops with proximity sensors to determine when potential users approach the tabletop, from which side they approach it, and how users move around the tabletop during actual interaction. Proxemic interactions (Ballendat et al., 2010; Greenberg et al., 2011) implement a concept and tools that formalize awareness of users and devices at the level of the environment by operating with distance, orientation, identity, movement, and location measurements for users and devices to enable intuitive interactions in smart environments. The proximity toolkit (Marquardt et al., 2011) enables system developers to access proxemic information from environmental motion detection sensors to design proxemic user interfaces, while the “gradual engagement design pattern” informs designs of transparent applications that unveil devices’ connectivity and information exchange capabilities, according to their proximity to each other (Marquardt et al., 2012).

2.2. Interactions on and around the human body

Prior work has looked at designing user interfaces *on* and *around* the human body. For instance, “Skinput” is a technology leveraged by detecting mechanical vibrations propagating through the body that enables users to employ the skin as an input surface (Harrison et al., 2010). “OmniTouch” is a wearable, vision-based system that allows any surface, including the arms and palms, to detect and react to multi-touch input (Harrison et al., 2011). Harrison et al. (2012) developed “Armura,” a system that enables real-time interaction with on-body projected content. Other work focused on user studies to inform on-body interface design and addressed aspects such as the locations on the body most suited for wearable displays (Harrison et al., 2009) or the implications of location and touch for on-body projected interfaces (Harrison and Faste, 2014). Besides touch gestures, participatory design studies revealed users’ preferences for body-referenced gestures to effect actions, such as to control the functions of the TV set (Vatavu, 2012b). Other researchers developed on-body interaction techniques in conjunction with smart devices. For instance, body mnemonics represent a design concept for portable devices that use body locations to execute actions (Ängeslevä et al., 2003). As an example, Strachan et al. (2007) developed "BodySpace," a system that enables users to control a music player by placing the device at different locations on the body, *e.g.*, the left shoulder is for browsing the playlist. Guerreiro et al. (2008) examined mnemonical body shortcuts in the form of gestures made with a mobile device pointed toward various body areas. Chen et al. (2012b) developed body-centric interactions with mobile devices, for which users position and orient the device to navigate and manipulate content in the space around and on the human body.

Our Smart-Pockets actions represent a hybrid of deictic and body-deictic gestures (McNeill, 1992; Kendon, 1994, 2004; Lausberg, 2013). Deictic gestures are the familiar "pointing" used to indicate objects in the concrete world, although they can also be used to indicate abstract concepts as well (McNeill, 1992) (p. 18). While deictics generally refer to pointing to an *external loci*, body-deictics are pointing gestures to a specific body part (Lausberg, 2013) (p. 176). Such hybrid gestures and the pocket metaphor have not been explored so far in the literature of gesture-based interaction and, therefore, we examine in this work both recognition and performance aspects for body-deictics in order to evaluate the practical validity of our Smart-Pockets technique.

2.3. Access to digital content using physical objects, body-referenced virtual tools, and smart garments

In the following, we connect our Smart-Pockets concept and recognition techniques with previous work from the Tangible User Interfaces (TUIs) literature (Ishii, 2008; Shaer and Hornecker, 2010), which enables us to discuss Smart-Pockets using the broad perspective of accessing and manipulating digital content with physical objects, *e.g.*, physical pockets and accessories in our case. We also discuss Smart-Pockets from the perspective of recent advances in prototyping smart garments and designing interactions for smart textiles (Cheng et al., 2013), including gesture-based approaches (Harms et al., 2009; Profita et al., 2013), which enables us to develop more connections between body-deictics and gesture-operated wearables.

Smart-Pockets is a concept that connects a physical location on the user's clothes or accessories with personal digital content stored in the cloud. Physical movement in the form of a body-deictic gesture to the physical location of the pocket enables instant access to digital content, assuming that an association was previously created between that content and the specific pocket. From this perspective, the Smart-Pockets concept connects to the paradigm of tangible user interfaces that builds on connections instantiated between physical objects (*e.g.*, phicons) and digital content that is controlled by direct manipulation of the physical objects; see Shaer and Hornecker (2010), Fishkin (2004), and Ishii (2008) for authoritative overviews on designing tangible interactions. While the community has designed TUIs for various applications and interactive contexts, such as learning (Markova, 2013; Markova et al., 2012), storytelling (Zhou et al., 2004), gaming (Vatavu et al., 2007), and even software design (Wu et al., 2011), joint exploration of tangible user interfaces and public displays has been practically neglected, with only few investigations performed so far, such as (Claes and Moere, 2015). In this context, we believe that Smart-Pockets may represent one valuable step in this direction to explore more intuitive connections between digital content and real-world physical containers in the context of interacting with public ambient displays.

Smart-Pockets also connect to recent research in smart garments, *i.e.*, electronics and textiles operating jointly in one common wearable system (Haladjian et al., 2016; Heller et al., 2014; Profita et al., 2013; Schneegass et al., 2015) that enable many opportunities for convenient ambient interactions (Cheng et al., 2013; Harms et al., 2009; Randell and Muller, 2000). For instance, the "Shopping Jacket" system of Randell and Muller (2000) alerts its owner of the availability of retail services in nearby shops by exchanging personal data, *e.g.*, the owner's shopping list, with servers installed in shops. This sort of information exchange enabled by connections between wearable and ambient systems that are performed smoothly and conveniently for the user (Weiser and Brown, 1995), next to the intuitiveness of mapping objects to pockets, is key for our concept of Smart-Pockets. The smart garments community has also looked into implementing gesture and activity recognition with embedded touch and motion sensors. For instance, Harms et al. (2009) were interested in activity-aware applications and proposed an approach for rapid prototyping of smart garments. The authors described the SMASH system, a smart long-sleeve shirt, which enables classification of users' body postures from motion data delivered by accelerometers embedded in the garment. The "FabriTouch" touch-sensitive fabric of Heller et al. (2014) was designed to be easily integrated in garments, such as to enable fabrication of touch-sensitive pockets. Recent technical advances in gesture sensing through smart textiles and wearables led Profita et al. (2013) to examine the implications of using gestures in public contexts to interface smart textiles at different on-body locations. Results revealed that system location (*e.g.*, the wrist or the waistline) and interactions at that location represent two distinct dimensions important for how such gestures are perceived by others in a public context. For instance, one interesting result observed by Profita et al. (2013) was the "gender effect": overall, on-body interactions with wearable interfaces appeared more acceptable by/on male participants, except when the wearable system was placed on the participants' waistline. Smart-Pockets connect to smart garments as they enhance the functionality of physical pockets and accessories by extending their capacity range to store digital objects. From this perspective, body-deictic gestures performed on pockets located on clothes and accessories are likely to be more acceptable to perform in public scenarios than interacting with other gesture types designed for other on-body or mobile interfaces (Profita et al., 2013; Rico and Brewster, 2009, 2010). Moreover, we estimate that, with careful interaction design, the subtlety of the gestures entailed by Smart-Pockets operation may even make them pass unnoticed by public observers (Anderson et al., 2015; Ashbrook, 2010).

Our Smart-Pockets concept also connects well to previous work about designing tools for virtual reality systems as well as with user interfaces employing body-centric visual representations (Ilmonen and Reunanen, 2005; Shoemaker et al., 2010). For instance, Ilmonen and Reunanen (2005) proposed the metaphor of a "virtual pocket user interface" for interacting in virtual environments, which presents the user with virtual replicas of real-world pockets for storing virtual tools. Shoemaker et al. (2010) proposed a suite of body-centric techniques for interacting with large wall displays, which they evaluated for a map browsing and editing application. Among those techniques, the authors proposed "body-based tools," which are virtual tools located on the user's body, such as around the waist, that are visualized on the display. The body-based data storage technique creates a direct connection between the location of the user's torso and a virtual container, from which personal files can be accessed and shown on the large display. The work of Shoemaker et al. (2010) showed that pointing to body parts represents a creative and useful way to access virtual content by referring to one's own body space in the context of proprioception. Our Smart-Pockets concept connects to Shoemaker et al. (2010) in how the metaphor of a physical pocket is employed for interacting with a computer system, yet it also adopts the new perspective of transforming real clothes and accessories (*e.g.*, pockets, bags, etc.) into actual placeholders for digital content, facilitating fast access to that content, similarly to how conventional pockets facilitate access to one's personal physical belongings.

2.4. About pockets: history, fashion, and augmenting pockets with technology

It is interesting, for the purpose of informing the design of Smart-Pockets, to take a look at the history of pockets in women's and men's clothing and fashion. Cambridge Dictionaries Online defines a pocket in relation to clothing as *"a small bag for carrying things in, made of cloth and sewn into the inside or onto the outside of a piece of clothing,"* but also more generally as *"a container, usually made of cloth, that is sewn into or onto a bag or attached to a set or door in a vehicle."*[3] The Victoria and Albert Museum of Art and Design provides a brief history on the subject[4], while a history of men's pockets can be found in (McKay and McKay, 2015). For example, we know from the literature of the 17th to the 19th century that most women had at least one pair of pockets, which were worn underneath petticoats. Men's pockets were sewn into the linings of their coats, waistcoats, and breeches. A variety of objects were kept in pockets, such as money, jewelry, everyday implements (*e.g.*, knife and scissors, or pincushions), objects of vanity for personal grooming, and even food and drinks. Although items that people carry in their pockets change over time, the wide range of personal items that have once been, are now, or will be relevant in the future for individuals to carry in their pockets is definitely impressive. This observation is key to our design of Smart-Pockets, as it informs the generic use and purpose of pockets that will need to extend their capacity range in the ubiquitous computing era to cover personal belongings that go beyond physical items. To reflect further the wide spectrum of items to store in pockets, we point readers to Robert D. Abrahams' *"A pocket history of Milton J. Wurtleburtle"* (Esquire, 1937, p. 64-65), a satire listing the contents of a man's pockets over the years, which shows that the contents of one's pockets tell a lot about that person. Such premises, resorting from the evolution of fashion design over time, history reports, and even from the fictional literature (Esquire, 1937), are informative for our concept of Smart-Pockets in the way that personal digital content, residing one's pockets, is reflective to a high degree of that person's belongings, interests and, more generally, life events. Moreover, when one's possessions involve content pertaining to the digital world, pockets need to be redesigned to cover the complex physical and digital lives of their owners.

The Victoria and Albert Museum's history of pockets informs us further that pockets could be purchased to match clothes, and they were often given as gifts. However, changes in fashion influenced the use of pockets and the type of objects that could be carried in pockets. For instance, the line of a dress with a high waistline could not be ruined by traditional pockets and, consequently, pockets were either re-designed or

[3] Pocket Meaning in the Cambridge English Dictionary, `http://dictionary.cambridge.org/dictionary/english/pocket`

[4] A history of pockets - Victoria and Albert Museum, `http://www.vam.ac.uk/content/articles/a/history-of-pockets/`

replaced by decorative bags. Modern pockets are tools with various uses and styles, and can be found in a variety of forms, such as watch-pockets, breast pockets, inner breast pockets, ticket pockets, coin pockets, cargo pockets, etc. Pocket styles vary greatly as well, from patch pockets to flap pockets, buttoned-flap patch pockets, jetted, jetted with zip, welt pockets, tailored pockets, etc. Professional fashion design details on many design techniques for pockets (Moyes, 1997). In this work, we inspire from fashion design to select our experimental conditions (*e.g.*, type and locations for modern pockets for our experimental design) to evaluate body-deictic gestures for accessing digital objects mediated by Smart-Pockets. We also use this knowledge to reflect on the future of physical pockets, having seen how tendencies in fashion design evolve so starkly over time, even for small things, such as pockets. This informs our explorations in this work to go beyond pockets, which we implement and evaluate in the form of Smart-Containers.

Pockets have also been considered by researchers in Human-Computer Interaction, but cases are extremely few. For instance, "PocketTouch" (Saponas et al., 2011) represents a capacitive sensor that, when connected to a mobile device and placed inside a pocket, enables multi-touch input on the device directly through the garment of the pocket. PocketTouch enables eyes-free touch gesture interaction with mobile devices without users having to remove those devices from their pockets. As we show later in the article, the PocketTouch technology represents a good option for implementing explicit segmentation of Smart-Pockets access actions. Related research also includes smart fabrics (Orth et al., 1998; Karrer et al., 2011; Heller et al., 2014). For instance, "Pinstripe" (Karrer et al., 2011) represents a textile user interface element made of parallel conductive lines sewn onto the fabric that enables user input by pinching and rolling the fabric. "FabriTouch" (Heller et al., 2014) is another example of touch-sensitive fabric that retains the flexible properties of the fabric. Researchers have also looked at pockets from a different perspective. For instance, Shimozuru et al. (2015) implemented a device for recognizing objects put into pockets by using a matrix of infrared sensors, and Jadidian and Katabi (2014) created systems for wireless charging mobile devices while still in the pockets of their users. However, there has been very little attention devoted overall to pockets, which were largely overlooked by the research community, despite their intuitive use as physical containers to store and carry personal objects.

2.5. Summary

The Smart-Pockets technique that we introduce in this work goes beyond prior work on ambient interactions in several ways. First, Smart-Pockets gestures represent a distinct class of gesture types that involve, in a short period of time, pointing to two different locations: to a specific area on the body and toward the display. From this perspective, Smart-Pockets access actions represent a hybrid between pointing to external loci and body-deictics, never examined before. Second, Smart-Pockets link physical objects (*e.g.*, a pocket or a generic container, such as a bag) to digital content and, consequently, they create the premises for exploring a specific form of tangible user interfaces for ambient interactions. The Smart-Pockets concept emerged from the increased interest manifested in the community for designing interaction techniques for ambient displays that rely on whole-body and deictic gestures (Müller et al., 2014; Vogel and Balakrishnan, 2005; Walter et al., 2013; Vatavu, 2012a), the recent focus on on-body interactions (Harrison et al., 2010; Harrison and Faste, 2014; Harrison et al., 2012), and our idea to associate personal digital content to physical pockets, similarly to how conventional pockets facilitate access to one's personal belongings.

3. Techniques for detecting Smart-Pockets access actions in whole-body data streams

In the following, we present the techniques that we use in this work to recognize Smart-Pockets access actions in continuous streams of whole-body gesture movement. First, we make the distinction between *segmentation* and *recognition* of Smart-Pockets actions. Segmentation refers to spotting a pocket action in a continuous stream of whole-body movement, while recognition is the process of identifying the specific pocket, hand, and action type (*i.e.*, take out content from the pocket or put content into the pocket) involved by the specific Smart-Pockets access action. Overall, we consider and evaluate techniques that fall into two distinct classes of approaches:

1. *Explicit segmentation* of Smart-Pockets access actions followed by action recognition. In this approach, users segment pocket access actions themselves during actual interaction with the system by mimicking a form of "click"-like input. Click-like events can be implemented with a secondary, worn sensor or with a hand-held input device, or by performing a specific hand posture that acts as a delimiter for Smart-Pockets gestures. For instance, through-fabric capacitive touch input (Saponas et al., 2011; Heller et al., 2014), mid-air devices (Baudisch et al., 2006; Wilson and Shafer, 2003) or specific postures that delimit meaningful motion (Haque et al., 2015; Malik et al., 2005; Vatavu et al., 2009) represent good candidates of techniques for click-like events to segment Smart-Pockets access actions in continuous whole-body movement.
2. *Implicit segmentation* of Smart-Pockets access actions followed by action recognition. In this approach, a detection algorithm analyzes the continuous stream of whole-body movement delivered by the sensor, in which it searches for Smart-Pockets gestures. In this work, we present and evaluate two techniques that follow this approach. The first technique relies on detection of specific events, such as the moment when the hand reaches a specific pocket; we call this technique *event-wise* detection. The second technique is a *brute-force* approach that considers all the possible subsequences of body movement from the continuous stream of whole-body data as potential candidates of Smart-Pockets access actions, which are filtered according to their similarity with Smart-Pockets samples from a training set.

We use the terms *explicit* and *implicit* in relation to how Smart-Pockets access actions are segmented. While explicit segmentation actively involves the users, which have to specifically delimit their commands, implicit segmentation automatically extracts Smart-Pockets access actions from the continuous stream of whole-body movement. We will come back later to discuss the pros and cons of explicit and implicit segmentation for Smart-Pockets access actions in the Discussion section of this article.

In the following, let $\mathcal{S}$ be a stream of whole-body movement data captured by some given motion tracking equipment, such as the Microsoft Kinect depth sensor[5] or a Vicon

[5] `http://kinectforwindows.org/`

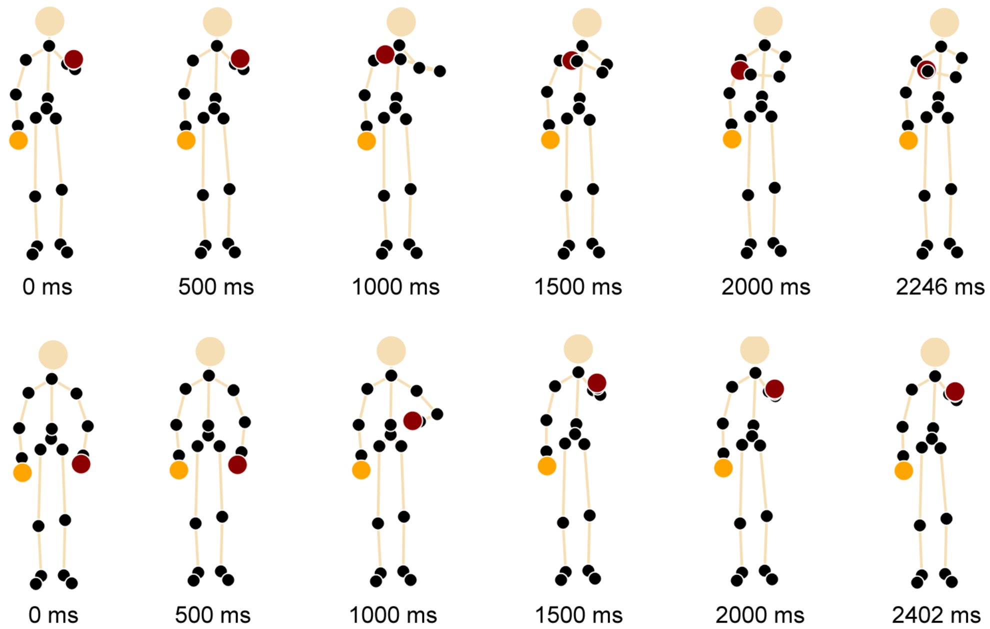

Figure 2: Two examples of Smart-Pockets actions (frontal view): putting content into the outer breast left pocket (top) and taking content out of the trousers front left pocket (bottom). NOTE: the dominant hand performing the Smart-Pockets gesture is shown in dark red; for clarity purposes, only six frames per gesture are shown in this figure.

motion tracker[6]. For convenience, we assume in the following that each data frame, *i.e.*, each body posture reported by the sensor at any timestamp t, consists of a fixed number n of 3-D joints tracked on the user's body; for example, n=20 joints are reported by the Kinect sensor (SDK versions up to 1.8), n=30 joints are reported by Kinect for Windows SDK v2, and the number of tracked joints can be customized to practically any value for marker-based motion capture systems, such as Vicon. Figure 2 illustrates two examples of Smart-Pockets access actions captured with the Kinect sensor. We denote by $P_t = \{p_{t,i} = (x_{t,i}, y_{t,i}, z_{t,i}) \,|\, i = 1..n\}$ the set of n joints representing the user's body posture at time t. With these notations, the data stream $\mathcal{S}$ can be expressed as the discrete set of postures $\mathcal{S} = \{P_t \mid t = 1, 2, ...\}$. Also, let $\mathcal{C}$ represent the whole-body gesture candidate that we wish to classify into one of predefined Smart-Pockets access actions.

3.1. Explicit segmentation

Explicit segmentation is implemented by employing a delimiter provided by the user in the form of a "click" event. The technique has been successfully implemented before (Ruiz and Li, 2011; Vatavu and Pentiuc, 2008) and various technology can be employed for this purpose (Saponas et al., 2011; Heller et al., 2014; Haque et al., 2015; Malik et al., 2005; Vatavu et al., 2009). Actually, how the delimiter is implemented is irrelevant for the

[6]`https://www.vicon.com/`

moment; all that matters is that clear "click"-like events are available for the system to segment Smart-Pockets actions. In the Discussion section, we elaborate on several techniques that can be used to generate "click"-like events in practice. In our experiments (see next in the paper), delimiters were already available from the gesture logs that specifically marked the timestamps for the moment when participants were prompted to execute a whole-body gesture and they initiated movement up to the moment when movement stopped (timestamps were marked by the experimenter during the gesture acquisition experiment with the press of a key). For an actual system deployed in real-world conditions, the "click" would implement the users' option to be in control of when their actions should be interpreted by the system. Let t_1 and t_2 be two consecutive timestamps at which such events are produced. The sequence of body postures captured between t_1 and t_2, *i.e.*, $C = \{P_{t_1}, ..., P_{t_2}\}$, represents the candidate input for our whole-body action recognizer, described later in this article.

3.2. Implicit segmentation

During implicit segmentation, the system automatically detects candidates for Smart-Pockets access actions in the continuous data stream of users' whole-body movement. In this work, we introduce and evaluate two techniques that implement implicit segmentation:

1. *Brute-force segmentation.* The brute-force technique considers all the subsequences of stream $\mathcal{S}$ as potential candidates for Smart-Pockets access actions. Each subsequence is fed into the recognizer and, according to the degree of similarity to known templates, that subsequence may be reported as a valid Smart-Pockets access action.
2. *Event-wise segmentation.* This technique searches for specific body posture events in stream $\mathcal{S}$, which filter the set of all possible subsequences considered for processing by the brute-force implementation. The specifics of our interaction scenario (*i.e.*, the hand either points in front of the body toward the display or the hand reaches for some pocket on the body) informs an efficient solution for this segmentation problem based on two *key events* that are easy to spot using simple arithmetics on z coordinates, as follows:
   (a) Pointed-Hand represents the event occurring when the user's hand points in front of the body, which makes the sensor report a z value for the hand different than that of the body center. This event is detectable as follows:

   $$\text{event Pointed-arm is True if } z_{\text{Hand}} - z_{\text{Body}} \geq \epsilon_1 \tag{1}$$

   where z_{Body} is the z location of the body measured by the sensor, *e.g.*, Microsoft Kinect reports the locations of the head, spine, and hip, which are all good approximations of the body's z coordinate; ϵ_1 represents a threshold, beyond which we can assume that the user's hand is sufficiently away from the body to be considered as pointing toward the ambient display. Our experiments (reported in the next sections of the article) informed the value ϵ_1=0.50 m.
   (b) Hand-in-Pocket represents the event occurring when the user's hand reaches a pocket, which makes the sensor report a z value for the hand similar in magnitude to the z value of the body. This event is detectable as follows:

   $$\text{event Hand-in-Pocket is True if } |z_{\text{Hand}} - z_{\text{Body}}| \leq \epsilon_2 \tag{2}$$

where ϵ_2 is our second threshold, below which we can assume that the hand is touching the pocket. Our experiments (reported in the next sections) informed a value of $\epsilon_2 = 0.20$ m for this threshold, which includes both Smart-Pockets and Smart-Containers actions.

The candidate $\mathcal{C}$ represents the set of body postures between two opposite key events. Furthermore, we can safely make another simplification assumption for our segmentation problem, according to which we only need to detect if the candidate $\mathcal{C}$ is a postfix of stream $\mathcal{S}$, instead of searching $\mathcal{C}$ at all possible locations in stream $\mathcal{S}$. For example, if the most recently-acquired body posture in $\mathcal{S}$ is of type POINTED-HAND, we go back in time in stream $\mathcal{S}$ to see if the opposite event, *i.e.*, HAND-IN-POCKET, has occurred recently. If true, the sequence of body postures between these two events constitutes the candidate $\mathcal{C}$ of a potential Smart-Pockets access action, to be further validated.

The validation of the candidate $\mathcal{C}$ as a specific Smart-Pockets action, the specific pocket (*e.g.*, whether it was the breast pocket or the trousers pocket that the user has touched), as well as rejection of non-pocket actions (*e.g.*, the hand may go into a resting state alongside the body, after having pointed toward the screen) are solved by treating $\mathcal{C}$ as a candidate gesture in a whole-body gesture recognition problem (Vatavu, 2012a), for which we dispose of a training set of examples of Smart-Pockets access actions. We describe this approach in the next section. Also, more sophisticated discrimination between Smart-Pockets access actions referring to digital objects stored inside pockets and actual reaching for the pocket to grab a physical item, such as the phone, need to be handled by users with explicit segmentation, *i.e.*, it is users that are in control of when their actions should be interpreted by the system. Alternatively, discrimination between the two actions could potentially be achieved automatically at the system level by using context information (Dourish, 2004), keeping and using records of a user's interaction history (Nakamura and Igarashi, 2008), or collecting information from other devices and/or sensors (Arase et al., 2010; Wiese et al., 2013), which are legitimate and interesting directions for future work on interpreting human gestures performed with reference to a physical or a digital object. The Discussion section contains a dedicated analysis of the pros and cons of explicit versus implicit segmentation of Smart-Pockets access actions, with explicit segmentation being our preferred implementation for Smart-Pockets. In the following, we strictly focus on discriminating Smart-Pockets access actions from each other as well as on detecting Smart-Pockets actions in continuous whole-body movement.

3.3. Classification of Smart-Pockets access actions

Let $\mathcal{T} = \{(\mathcal{T}_k, \phi_k) \mid k = 1..|\mathcal{T}|\}$ be a training set of previously recorded and annotated whole-body movement sequences representing valid Smart-Pockets access actions, for which sample $\mathcal{T}_k = \left\{T_i^k \mid i = 1..|\mathcal{T}_k|\right\}$ belongs to Smart-Pockets class ϕ_k; *e.g.*, ϕ_k may be "put content into the front trousers pocket." According to our definitions, each posture T_i^k of sample $\mathcal{T}_k$ consists of n points in 3-D, $T_i^k = \{t_{i,j}^k \in \mathbb{R}^3 \mid j = 1..n\}$, where index k identifies the training sample, i identifies a specific body posture in that sample, and j refers to a specific point of the i-th body posture of training sample $\mathcal{T}_k$.

Let μ be a metric defined over sequences of whole-body movement that computes a real positive value representing the dissimilarity between any two movement sequences. The whole-body gesture literature has implemented μ as the Euclidean distance or the

Dynamic Time Warping (DTW) function; see (Vatavu, 2012a; De Silva et al., 2014). The next sections of the paper report experiments that evaluate both these metrics. In this work, we adopt the Nearest-Neighbor classification approach (Webb, 2003) (p. 93), according to which a candidate is classified to the class of its closest template from the training set. Let $\mathcal{T}_{k^\star}$ be the template from the training set $\mathcal{T}$ that is closest to $\mathcal{C}$ with respect to μ, *i.e.*, $\mu(C, \mathcal{T}_{k^\star}) = \min_k\{\mu(C, \mathcal{T}_k)\}$. We reject sequence $\mathcal{C}$ if the measure's result is larger than a given threshold ϵ_3; our experiments informed a value of 0.125 for this threshold for normalized body gestures. Otherwise, $\mathcal{C}$ is validated as a Smart-Pockets access action and classified to the same type as template $\mathcal{T}_{k^\star}$. For completeness purposes, we provide below the formulas for the Euclidean (eq. 3) and DTW (eq. 4) measures of dissimilarity:

$$\mu_{\text{ED}}(\mathcal{C}, \mathcal{T}_k) = \sum_{i=1}^{|\mathcal{C}|} \sum_{j=1}^{n} \left\| c_{i,j} - t_{i,j}^k \right\| \tag{3}$$

where $\|a - b\|$ represents the Euclidean distance between points $a = (x_a, y_a, z_a)$ and $b = (x_b, y_b, z_b)$ in $\mathbb{R}^3$, *i.e.*, $\|a - b\| = \left((x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2\right)^{\frac{1}{2}}$. Note that the two sequences $\mathcal{C}$ and $\mathcal{T}_k$ must be equal in length (*i.e.*, they must present the same number of postures, $|\mathcal{C}| = |\mathcal{T}_k|$) in order for the Euclidean distance to match body postures one to one in chronological order. As this requirement is not met in practice, we re-sample whole-body sequences of movement to make them equal in length. The DTW measure is:

$$\mu_{\text{DTW}}(\mathcal{C}, \mathcal{T}_k) = \texttt{cost-matrix}[|\mathcal{C}|, |\mathcal{T}_k|] \tag{4}$$

where `cost-matrix` is a $|\mathcal{C}| \times |\mathcal{T}_k|$ matrix that iteratively computes the optimum alignment between subsequences of body postures from $\mathcal{C}$ to subsequences of $\mathcal{T}_k$, until the whole sequences $\mathcal{C}$ and $\mathcal{T}_k$ are aligned. The result of the DTW measure is found in the lower right corner of the matrix, *i.e.*, `cost-matrix`$[|\mathcal{C}|, |\mathcal{T}_k|]$; see (De Silva et al., 2014; Vatavu, 2012a).

4. Experiment #1: Smart-Pockets access actions

We conducted a data acquisition procedure to collect whole-body movements produced by users when they reach their hands into pockets at various locations on clothes (*e.g.*, front and rear trousers pockets, breast pockets, etc.) in order to understand how accurate our detection and classification techniques are and, consequently, to validate the feasibility of implementing Smart-Pockets for practical ambient display interactive scenarios.

4.1. Participants

Ten participants (5 were females) aged between 22 and 31 years old (M=23.3 years, SD=2.8 years) took part in the data collection experiment.

4.2. Apparatus

A Microsoft Kinect sensor for Windows (SDK $v1.8$) was employed to capture participants' whole-body movements as time sequences of body postures with each posture represented by 20 joints evenly distributed on the human body; see joints' locations in (Microsoft, 2013, p. 8). The Kinect sensor, running at a resolution of 640×480 pixels and frame speed of $\approx 25-30$ fps, was connected to a desktop PC (3.2 GHz Quad-Core) running our custom software application for the acquisition and storage of whole-body movements.

4.3. Design

The experiment was a within-subject design with 3 independent variables:

1. Pocket represents the specific location on the human body (*e.g.*, upper chest or trousers rear) at which pockets are sewn on clothes and from which/to which digital content can be taken out/put in. We designed Pocket as a nominal variable with 8 conditions: *outer breast pockets (left and right side)*, *inner breast pockets (left and right side)*, *front trousers pockets (left and right side)*, and *rear trousers pockets (left and right side)*; see Figure 3 for visual illustrations of the experimental conditions for the Pocket variable. These conditions cover common locations for pockets as per today's common styles of fashion.
2. Action represents the type of operation performed on a specific smart pocket. Action is a nominal variable with 2 conditions: *put in* and *take out* content into/from the smart pocket.
3. Hand represents the user's active hand employed to access the smart pocket, nominal, with 2 conditions: *left* and *right* hand.

Note that the ergonomics of comfortable hand access actions to specific pockets restricts the number of possible Hand $\times$ Pocket combinations for this experiment. For instance, except for the two outer breast pockets that can be reached comfortably with both hands, all the actions on all the other pockets can only be performed with one hand (*e.g.*, the left inner breast pocket is reachable with the right hand only), which results in a total number of 2 (outer breast pockets) $\times$ 2 (hands) + 6 (remaining pockets)$\times$ 1(hand) = 10 Pocket $\times$ Hand combinations. When we further multiply these combinations by two Action types, we get 20 distinct experimental trials to collect Smart-Pockets access actions from participants.

The analyses that we run to evaluate the performance of our detection and recognition techniques of Smart-Pockets access actions add three more independent variables to our experimental design, as follows:

4. Recognizer represents the specific classification technique employed to recognize body movements as Smart-Pockets access actions. Recognizer is a nominal variable with 4 conditions: Euclidean-All-Joints, Euclidean-Hands, DTW-All-Joints, and DTW-Hands. The Euclidean and DTW metric functions were defined in equations 3 and 4, while the "All-Joints" and "Hands" suffixes specify whether all the 20 joints or only the 2 hand joints reported by the motion sensor were employed to compute the metrics.

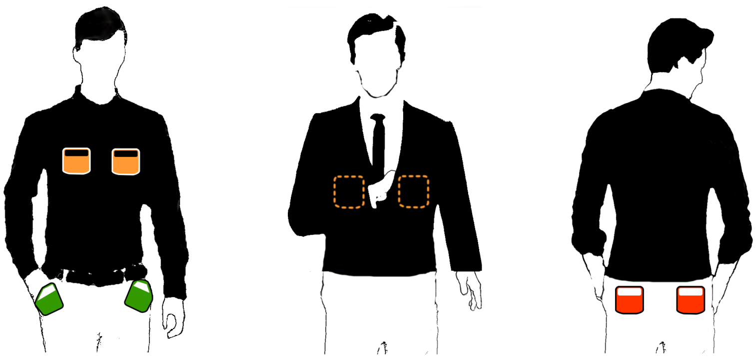

Figure 3: Smart-Pockets locations for the first experiment: *outer breast pockets* (left and right side), *front trousers pockets* (left and right side), *inner breast pockets* (left and right side), and *rear trousers pockets* (left and right side).

5. P represents the number of participants from which body movement data was collected to train the recognizers. P is an ordinal variable with 8 conditions: 1, 2, 3, 4, 5, 6, 7, and 8 training participants.
6. DETECTOR represents the method employed to segment users' pocket access actions in the continuous stream of whole-body movement delivered by the sensor. DETECTOR is a nominal variable with 3 conditions: *explicit*, *implicit brute-force*, and *implicit event-wise*. In the *explicit segmentation* condition, Smart-Pockets are delimited in advance (timestamps were marked by the experimenter during gesture acquisition by a press of a key in the software application implementing the experiment design), while in the two *implicit* conditions the system automatically detects Smart-Pockets actions using our two segmentation techniques described earlier in the article.

The three independent variables employed for the analysis of Smart-Pockets recognition performance generate a total number of 4 (RECOGNIZERS) $\times$ 8 (values for the number of training participants P) $\times$ 3 (DETECTORS) = 96 trials.

4.4. Procedure

Participants stood at a distance of about 4 meters from a large display, on which text instructions were presented by our custom software application regarding the specific POCKET, active HAND, and ACTION to be performed, according to our experimental design, *e.g.*, *"[Take out] content from your [trousers rear left pocket] with your [right hand] and point it toward the screen."* Text in square brackets indicates conditions for our three independent variables regarding Smart-Pockets access actions. Pointing was performed by participants by outstretching their arms toward the display. There was absolutely no constraint that we imposed on the hand pose to use (*e.g.*, index finger outstretched, index and thumb fingers pinching, etc.), which we allowed participants free to vary as they wished in order to encourage them to use their own interpretation of "holding" or "manipulating" digital objects and, by that, to make their experience feel as natural and intuitive as possible. Once participants confirmed they understood the action they were required to perform, their whole-body movement was recorded and the experiment continued with the next trial. In total, there were 20 trials of various pockets and pocket access actions (see the experiment design), which resulted in 20 Smart-Pockets gestures acquired from each participant. Participants were not required to wear any particular outfit other than the clothes they worn the day of the experiment. In case a specific pocket was not available, participants were asked to mimic the action required to access that pocket. By not particularly enforcing our participants to wear a specific outfit with pockets at fixed locations, we were able to collect a wide range of variation in body movement for pocket access actions determined by (i) slight differences in pocket locations for different clothes and (ii) mimicking the presence of a pocket, an action useful for actual practical scenarios where such a pocket is not available, yet users might want to employ nevertheless. By adopting this procedure, we will evaluate recognition algorithms that are invariant to variations in pocket locations on clothes and, even more, also invariant to the actual presence of pockets. The order of trials was randomized across participants. At the end of the data collection procedure, participants filled in a questionnaire in which they rated the *perceived comfortability* to reach each Pocket using ratings from a 5-point Likert scale ranging from 1 (very uncomfortable) to 5 (very comfortable). Participants also rated their actual *frequency of use* of each Pocket in their everyday lives with ratings from 1 (never) to 5 (very often).

5. Results #1: Recognition accuracy of Smart-Pockets actions

We present in this section experimental results on the recognition accuracy of Smart-Pockets access actions as well as results on body movement analysis from a dataset of 9 (participants) × 20 (Pocket × Hand × Action conditions) = 180 body movement records[7], comprising a total number of 12,174 body postures with 243,480 joints tracked in 3-D.

To understand the spatial variation in the articulation of Smart-Pockets access actions between our participants, we looked at their hands' locations reported by the sensor when hands were placed in pockets as well as when hands were pointing toward the display, according to our two event types from the event-wise implicit segmentation technique. For each Smart-Pocket action type and for each of the two conditions (hand in pocket, hand pointing), we computed the average Euclidean distance between the hands' locations for all participants. Results showed an average variation in hands' locations of 0.16 m (SD=0.08 m) when hands were in pockets and an average variation of 0.18 m (SD=0.08 m) when hands were outstretched pointing to the display. Figure 4 shows the average variation values computed for each Smart-Pockets action. We found a significant effect of Pocket on the amount of variation in hands' locations when hands were in pockets ($\chi^2(7) = 38.925$, $p < .001$) as well as when they were pointing ($\chi^2(7) = 29.574$, $p < .001$). We found no significant difference between the *left* and *right* Hand conditions (Wilcoxon signed-ranked tests showed $Z = -1.745$ and $Z = -0.732$, respectively, $p > .05$, *n.s.*). These results show that the type of Pocket influences users' precision and consistency when performing hand movements toward that pocket, which may impact the classification accuracy of Smart-Pockets actions. Possible causes for this

[7]The body movement data of one participant was unfortunately lost, but his responses to the questionnaire were not.

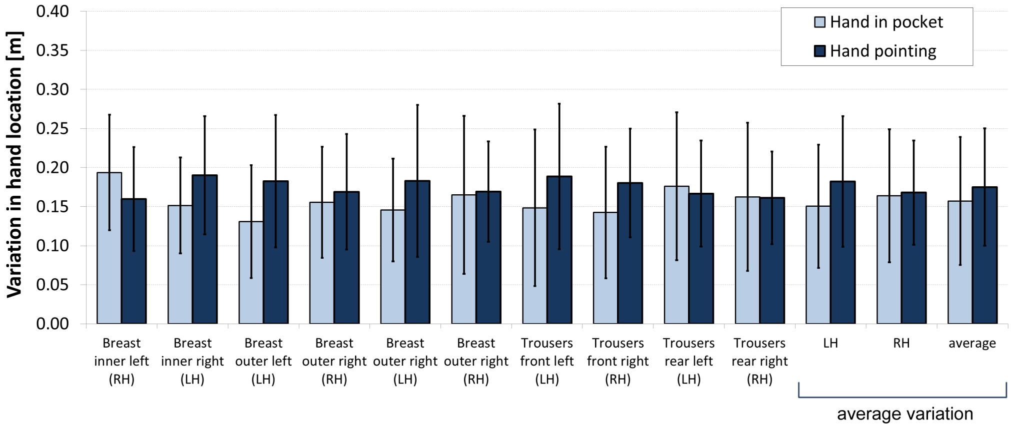

Figure 4: Variation in hand location when the hand was in pocket and when pointing, computed for each Pocket and Hand experimental conditions. Error bars show ±1 SD.

variation are represented by different locations of similar pockets on our participants' clothes as well as different precision of accessing those pockets, *i.e.*, actually placing the hand into the pocket or just pointing to the pocket. Our results also indicate that users show similar precision for their movements performed with either the left or the right hand. In the following, we compute and report recognition accuracy rates for both explicit and implicit segmentation conditions.

5.1. Explicit segmentation of Smart-Pockets access actions

We start our recognition analysis by measuring and reporting the recognition accuracy of Smart-Pockets access actions in the explicit segmentation condition in order to understand the *upper margin of recognition accuracy* expected for the Smart-Pockets technique. All the recognition results that we report in this work are computed from *user-independent classification* procedures, *i.e.*, different participants were used for training and testing.

Our recognition experiment followed the design of similar experiments conducted in the literature for user-independent gesture recognition; see (Vatavu et al., 2012, p. 275) and (Vatavu, 2013, p. 395) for examples. A number of P participants were randomly selected to deliver training data and one additional participant (which was different from the first P) was randomly selected for testing. The number of training participants P was varied according to the experimental design, from 1 to 8 participants. This selection procedure was repeated for 100 times for each value of the P independent variable. In total, we report recognition results from 8 (conditions for the number of training participants P) × 100 (repetitions) × 20 (Pocket × Hand × Action conditions) = 16,000 classification trials. All body movement samples were pre-processed before recognition: they were uniformly scaled into the $[-1, 1]^3$ unit cube, translated to origin, and re-sampled into a fixed number of 32 body postures (which corresponds roughly to 70 ms time duration between consecutive frames or a speed frequency of 15 fps). Scale and translation normalizations applied to the data make our recognizers scale and translation invariant, while body movement re-sampling is required by the Euclidean metric that needs the same number of body postures for the body movements it compares.

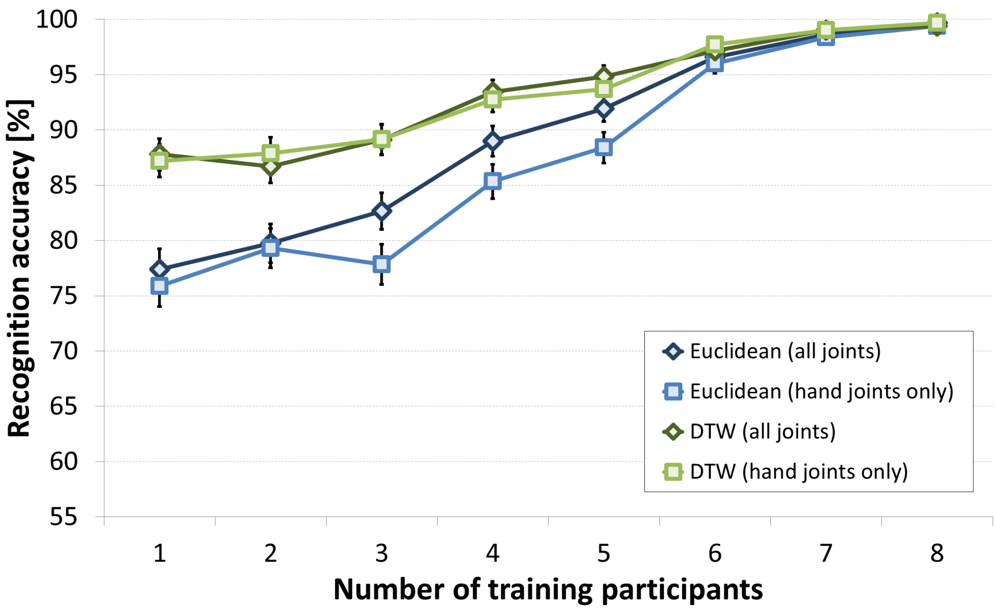

Figure 5: User-independent recognition accuracy of Smart-Pockets actions for all Recognizer × P conditions and the *explicit segmentation* assumption, *i.e.*, users segment their Smart-Pockets access actions themselves. Note how recognition accuracy increases with more training samples up to 99.4% accuracy for DTW with 8 training participants × 1 sample per each Smart-Pockets gesture type. Note: error bars show 95% CIs.

Figure 5 illustrates recognition results for all Recognizer and P conditions. A Cochran's Q test showed a significant effect of Recognizer on the recognition accuracy of Smart-Pockets actions ($\chi^2_{(3,N=16000)}$=1239.846, p<.001). Follow-up post-hoc McNemar's tests (Bonferroni corrected at $p = .05/4 = .0125$) revealed that DTW significantly outperformed the Euclidean recognizer with an average accuracy of 93.4% (SD=2.5%) versus 89.5% (SD=3.1%) in the all-joints condition and 93.4% (SD=2.5%) versus 87.6% (SD=3.3%) in the hand joints only condition, respectively, all p<.001. We also found that the Euclidean recognizer was more accurate when all the joints were considered for recognition than in the hand joints only condition (average recognition rate was 89.5% versus 87.6%, p<.001). There was no significant difference between the recognition accuracy reported by DTW in the two joints conditions (average 93.4% and 93.4%, respectively, $n.s.$ at p=.01).

We found an overall significant effect of the number of training participants P on recognition accuracy ($\chi^2_{(7,N=8000)}$=3625.701, p<.001), a result that we also obtained for each individual metric (all p<.001). For instance, DTW delivered 87.2% recognition accuracy when data from only one participant was used (with one training sample per pocket action), which increased to 93.5% with data from 4 participants (p<.001), and reached 99.4% when training data from 8 participants was available (p<.001).

The high recognition accuracy of Smart-Pockets access actions obtained for user-independent training with just few samples from a few people (*i.e.*, +99% accuracy with only 8 training samples) show that the Smart-Pockets technique is feasible for implementation in practical scenarios under the explicit segmentation assumption. Note that our results were obtained with only one training sample from each training participant and, consequently, we expect accuracy to improve further when more samples per participant are added to the training set leading to steeper curves in Figure 5. Also, these results recommend the DTW-Hands recognizer for the implicit segmentation scenario, described in the next section.

5.2. *Implicit segmentation of Smart-Pockets access actions*

We know from the previous section that the Smart-Pockets technique can be successfully implemented with an accuracy of +99% when users segment their pocket access actions themselves. In this section, we are interested to learn how an automatic segmentation procedure would compare to those results. Our hypothesis is that such a procedure will lead to segmentations of Smart-Pockets access actions that will sometimes be inaccurate (*e.g.*, because of candidates that present slightly different starting and ending timestamps when compared to the actual pocket access body movement) which, in turn, will cause an increase in recognition error. We report in this section the recognition accuracy of Smart-Pockets access actions when body movement is automatically segmented by the system and we also evaluate the detection accuracy of Smart-Pockets actions by computing and reporting the false detection rate of our Detectors. As in the previous section, all the recognition results that we report were computed from *user-independent classification* procedures.

The number of body postures captured for each Smart-Pockets access action varied between 48 and 106 with a mean of 67.6 postures (SD=12.3). At roughly 30 fps, these values correspond to pocket access times that fall between 1.56 and 3.53 seconds (M=2.22 s, SD=0.41 s). The difference in the z coordinates between the pointed arm and the center of the users' body (*i.e.*, the value $z_{\text{Hand}} - z_{\text{Body}}$; see eq. 1) varied between 0.42 m and 0.83 m (M=0.66 m, SD=0.06 m), while the difference between the z coordinates of the hand while in pocket and the body's z value (as per eq. 2) varied between 0.00 m and 0.20 m (M=0.04 m, SD=0.05 m). A Wilcoxon signed-rank test showed a significant difference between these two means ($Z_{(N=180)}$=11.635, p<.001) with a large Cohen effect size (r=.613). These results, as well as the large effect size, informed values 0.20 m and 0.50 m for thresholds ϵ_1 and ϵ_2, respectively (see eqs. 1 and 2 for the meaning of these thresholds); see also Figure 6 for the histogram distributions of the differences in z coordinates.

We now proceed to the evaluation of the performance of our Detectors of Smart-Pockets access actions. The design of the recognition experiment was as follows: P participants were randomly selected to deliver training samples and 1 additional participant (different from the first P) was randomly selected for testing. Samples from the training participants were normalized with respect to scale, translation, and number of body postures per pocket access action as in the previous experiment. Each sample from the testing participant (not normalized) was inserted at a random timestamp in a continuous stream of body movement $\mathcal{S}$ that was generated with body postures from the testing participant's dataset and hand joints generated at random locations. The length of the testing stream $\mathcal{S}$ was 900 body frames, which corresponds to 30 seconds of continuous movement data. As the average time duration of a Smart-Pockets access action

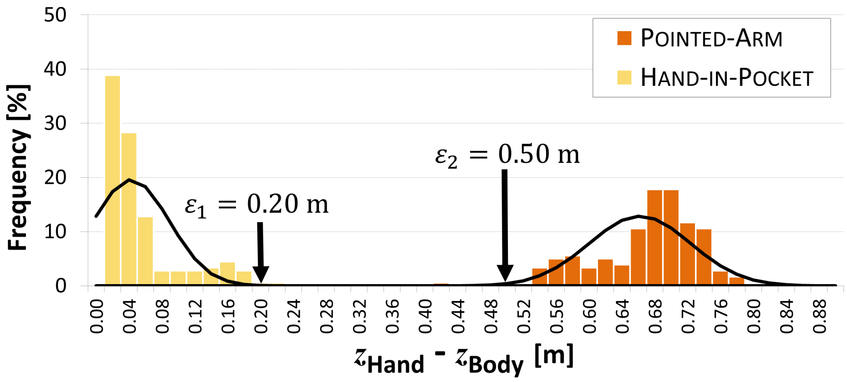

Figure 6: Histogram of the difference in z coordinates for the Pointed-Arm and Hand-in-Pocket events with normal curves overimposed. These results inform threshold values $\epsilon_1 = 0.20$ m and $\epsilon_2 = 0.50$ m, respectively.

was 2.2 seconds only, the "interesting" part of each body movement stream that our automatic procedures had to detect represented on average just 7% of the data submitted to processing. The testing stream $\mathcal{S}$ was fed into each Detector that reported both the type of the Smart-Pockets access action and its timestamp in the stream. We counted correct recognition and detection when the type of the reported pocket action matched the type of the testing sample and the difference between the timestamp reported by the detector and the true timestamp was less than 0.5 seconds. The procedure for selecting training and testing participants was repeated for 100 times for each P and the generation of the testing stream $\mathcal{S}$ was repeated for each pocket access action type. In total, we report recognition results from 8 (conditions for the number of training participants P) $\times$ 100 (repetitions) $\times$ 20 (Pocket $\times$ Hand $\times$ Action conditions) $\times$ 2 (Detector implementations for implicit segmentation) = 32,000 detection and classification trials.

Figure 7 illustrates recognition results for the DTW-Hands recognizer, which was the recognizer that delivered the best performance under the explicit segmentation assumption; see the previous section. A Cochran's Q test showed a significant effect of the number of training participants on recognition accuracy for both the *brute-force* and the *event-wise* technique ($\chi^2_{(7,N=7000)}$=2958.885 and $\chi^2_{(7,N=7000)}$=3843.380, respectively, all p<.001). Follow-up McNemar's tests (Bonferroni corrected at p=.05/2=.025) showed that the *brute-force* detector was significantly more accurate than the *event-wise* detector ($\chi^2_{(1,N=16000)}$=229.493, p<.001). The maximum accuracy attained by *brute-force* was 94.0% with 8 training participants, while *event-wise* delivered only 84.1% under the same training conditions. However, even *brute-force* was significantly less accurate than DTW-Hands running in the explicit segmentation condition ($\chi^2_{(1,N=16000)}$=3709.167, p<.001), with the maximum recognition performance being 94.0% versus 99.4% for P=8 participants.

To understand more about the performance of each Detector, we computed two additional accuracy measures: (1) the percentage of false detections and (2) the F-measure reporting the weighted average of our detectors' precision and recall performances; see Figure 8. On average, false detection rates were 5.13% (maximum 7.00%) for *event-wise*

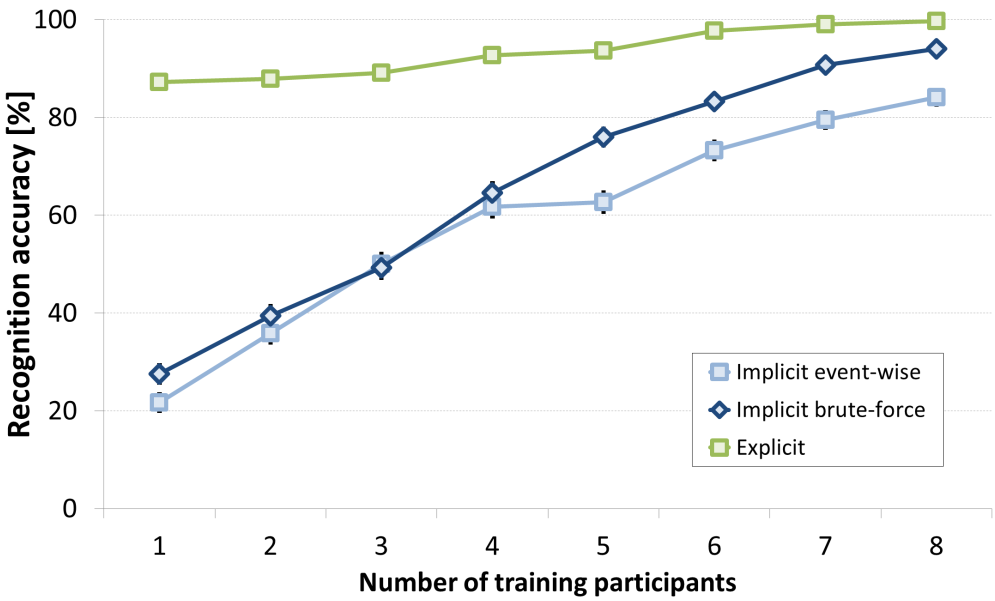

Figure 7: User-independent recognition accuracy of Smart-Pockets actions for the DTW-Hands recognizer and *implicit segmentation*, *i.e.*, actions are automatically segmented by the system. Note how recognition accuracy increases with more training samples up to 93.4%. Notes: error bars (too small to be visualized in this figure) show 95% CIs; the performance of the DTW-Hands recognizer in the explicit condition is also shown (in green) for comparison purposes, as it represents the upper margin of accuracy expected for Smart-Pockets.

and 4.63% (maximum 6.40%) for *brute-force*, which decreased to 2.10% and 1.35%, respectively, for the maximum number of 8 training participants. We found a significant effect of the number of training participants P on the false detections rate for both Detectors ($\chi^2_{(7,N=2000)}$=92.416, p<.001 for *event-wise* and $\chi^2_{(7,N=2000)}$=103.422, p<.001 for *brute-force*, respectively). A McNemar's test showed no significant difference in terms of the false positives rates delivered by the two Detectors ($\chi^2_{(1,N=16000)}$=4.914, *n.s.* at p=.01). The value of the F-measure increased for both detectors with more training participants, from 0.34 and 0.42 for P=1 to 0.90 and 0.96, respectively, for P=8 participants; see Figure 8. The F-measure evaluates the tradeoff between recall and precision in the [0..1] interval: the higher its value and the closer it gets to 1.0, the better the performance (Zhang and Zhang, 2009). Our results show that the F-measure increases almost linearly with the number of training participants up to 0.96 for both Detectors, confirming our previous results about the better performance of the *brute-force* technique over *event-wise*.

While the results of the automatic, implicit segmentation of Smart-Pockets access actions are inferior to those obtained in the explicit segmentation condition, recognition rates are still high, *i.e.*, 94% for 8 training participants × 1 training sample for each Smart-Pockets access gesture. It is likely that the recognition accuracy will improve further with more training data. However, there are several pro and con arguments for the implicit versus explicit segmentation that we present in the Discussion section of the

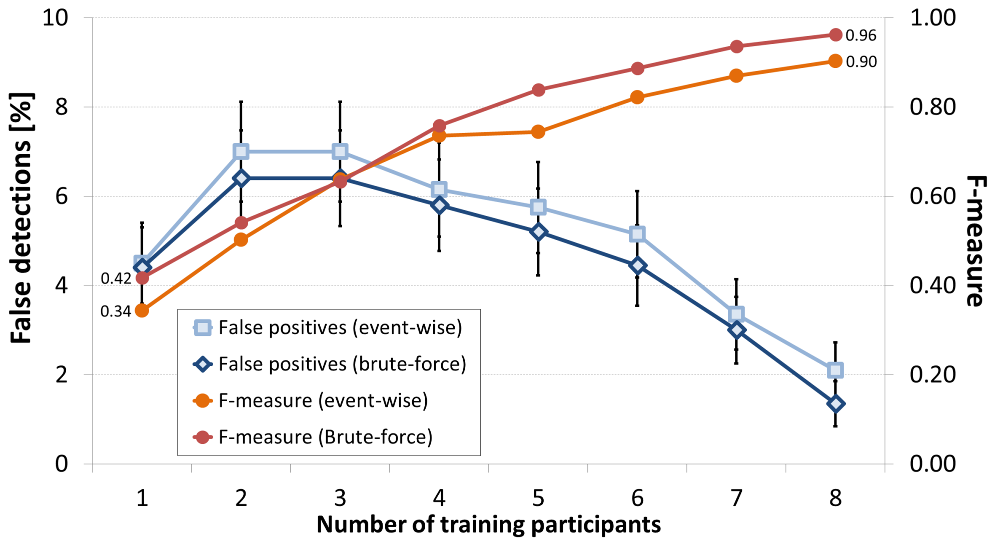

Figure 8: False detection rates and F-measure values obtained for the implicit DETECTORS of Smart-Pockets access actions. NOTES: error bars show 95% CIs; the false detections rates decrease when data from more training participants is available; higher F-measure values (closer to 1.0) indicate better performance.

paper, with explicit segmentation being our preferred choice for Smart-Pockets.

6. Experiment #2: Smart-Containers access actions

Motivated by the high recognition results obtained for the 20 types of Smart-Pockets access actions, we wanted to explore even more options for users to access personal content with body-deictic gestures. As pockets are available for briefcases, backpacks, and similar items, we wanted to explore those options as well. Consequently, we conducted a second data acquisition procedure to collect whole-body movements produced by the nine validated participants from the first study. This time, participants were asked to access content from various hand-held or worn accessories and containers, such as backpacks, briefcases, or cups. The same experimental design was used for this study as well, except that the independent variable POCKET was renamed to CONTAINER to better reflect the new actions. CONTAINER is a nominal variable with seven (7) conditions: *backpack*, *sleeve* (left and right), *briefcase* (held in the left or the right hand), and *coffee cup* (held in the left or the right hand); see Figure 9 for a visual illustration of these conditions. Except for the backpack that can be accessed easily with either the left or the right hand, all the other containers require one specific hand explicitly, *e.g.*, the right sleeve can only be accessed with the left hand, etc. Therefore, the maximum number of conditions for Smart-Containers generated by CONTAINER × HAND combinations is 8, which makes the total number of Smart-Pockets and Smart-Containers experimental conditions 20+8=28. Just like for the Smart-Pockets experiment, we did not enforce participants to adopt a specific pose or location in space for container objects, *e.g.*, no other express indication was given for the “coffee cup” other than “Please hold the coffee cup in your hand." By adopting this approach, we were able to collect a wide range of variations in whole-body movement during access to various containers.

Figure 9: Smart-Containers employed for the second study with 8 experimental conditions: *sleeves* (left and right arm), *backpack* (accessible with the right or left hand), *briefcase* (held in the left or the right hand), and *coffee mug* (held in the left or the right hand). Except for the backpack, all containers are accessible with one hand only because of holding and grasping constraints, *e.g.*, the right sleeve can only be reached with the left hand, etc.

7. Results #2: Recognition accuracy of Smart-Containers actions

In this section, we report results from a recognition experiment involving all the 28 Smart-Pockets and Smart-Containers access actions. Before the recognition experiment, we ran a similar analysis for evaluating the variation in the performance of Smart-Containers access actions, using the same approach as we used before for Smart-Pockets. We found an average variation in hands' locations of 0.20 m (SD=0.10 m) when hands were accessing containers and 0.19 m (SD=0.08 m) when hands were outstretched pointing to the display. Figure 10 illustrates the average variation for each CONTAINER and HAND experimental conditions. We found a significant effect of CONTAINER on the amount of variation in producing Smart-Containers access actions when hands were accessing containers ($\chi^2(7) = 18.560$, $p < .001$), but not when hands were pointing ($\chi^2(7) = 11.202$, $p > .05$, *n.s.*). A Wilcoxon signed-rank test revealed a significant effect of HAND ($Z = -4.560$, $p < .001$) when hands were reaching for the containers, but not when hands were pointing ($Z = -.791$, $p > .05$, *n.s.*). The average amount of variation when hands were accessing the containers (0.20 m) is larger in magnitude than the variation observed for hands in pockets (0.16 m), a result that can be explained by more degrees of freedom to hold and move containers when compared to pockets sewn on clothes. We focus in the following on the recognition accuracy rates for Smart-Pockets and Smart-Containers access actions.

As in the previous sections, all the recognition results that we report were computed from *user-independent classification* procedures, *i.e.*, different participants were used for training and testing the detectors and recognizers. Overall, we report the accuracy of recognizing access actions from 28 (Smart-Pockets and Smart-Containers) $\times$ 8 (conditions for the number of training participants P) $\times$ 100 (repetitions for each P) $\times$ 4 (RECOGNIZERS) = 89,600 classification trials for the explicit segmentation condition and

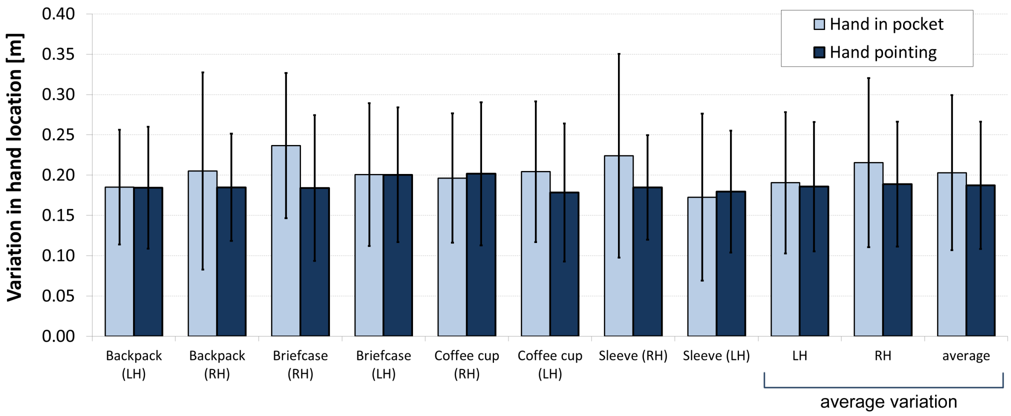

Figure 10: Variation in hand location for each CONTAINER and HAND experimental conditions. Error bars show ±1 SD.

22,400 classification trials for the implicit brute-force segmentation[8]. Overall, a total number of 112,000 attempts to classify Smart-Pockets and Smart-Containers actions was performed. Recognition results are shown in Figure 11.

A Cochran's Q test showed a significant effect of RECOGNIZER on recognition accuracy ($\chi^2_{(3,N=28800)}$=1545.129, p<.001). This time, the all-joints DTW recognizer was significantly more accurate than DTW hands-only by 1.1% as revealed by a follow-up McNemar's test ($\chi^2_{(1,N=28800)}$=70.534, p<.001), but the effect size was small (r=.035). Again, DTW outperformed the Euclidean recognizer for both all-joints and the hands-only conditions (p<.001, r=.110). The maximum recognition accuracy was attained by DTW with 8 training participants: 98.8% and 98.2% in the all-joints and hands-only conditions, respectively. The difference in accuracy was marginally significant as revealed by a follow-up McNemar's test ($\chi^2_{(1,N=3600)}$=6.682, p=.01, r=.031) (Bonferroni corrected at p=.05/3=.017). The best performance of DTW achieved for the 28 actions (98.8%, see Figure 11) was just 0.9% smaller than the best performance achieved for the 20 Smart-Pockets access actions (99.7%, see Figure 5). We found a significant effect of the number of training participants P on recognition accuracy ($\chi^2_{(7,N=14400)}$=12127.999, p<.001): accuracy increased from an average of 64.6% for P=2 training participants to 86.8% for P=4 participants, and to 98.0% for P=8 participants (average values computed for all four RECOGNIZERS). These results show that the Smart-Pockets technique can easily accommodate recognition of more access actions without considerable loss of accuracy. In the following, we consider the performance of Smart-Pockets in the implicit segmentation condition, using the *brute-force* approach and the DTW-HANDS recognizer for all the 28 Smart-Pockets and Smart-Containers access actions.

Figure 12 illustrates the recognition performance of DTW-HANDS under the implicit segmentation condition. Recognition accuracy starts as low as 15.3% for one training participant, increases to 56.6% for P=4 participants, and attains the maximum value of 93.3% for P=8 training participants. A Cochran's Q test revealed a significant effect

[8]Only DTW-Hands was employed for the implicit segmentation recognition experiment.

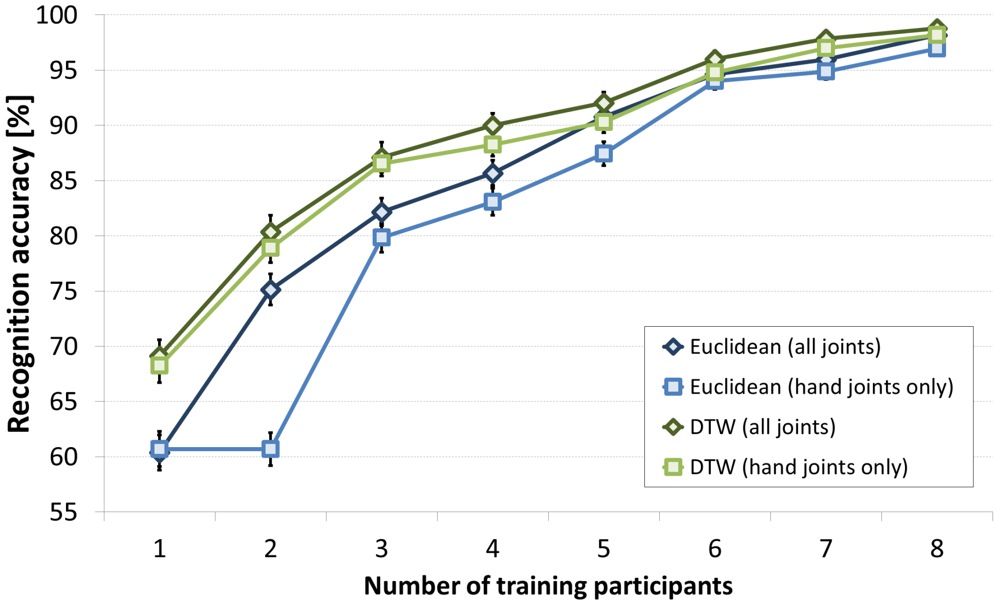

Figure 11: User-independent recognition accuracy of Smart-Pockets and Smart-Containers access actions for all Recognizer × P conditions under the *explicit segmentation* assumption, *i.e.*, users segment their actions themselves. Note how recognition accuracy increases with more training samples up to 98.8%; error bars show 95% CIs.

of the number of training participants on recognition accuracy ($\chi^2_{(7,N=3600)}$=8388.435, p<.001). A McNemar test showed that DTW-Hands was significantly more accurate in the explicit than in the implicit condition ($\chi^2_{(1,N=28800)}$=6960.120, p<.001). The smallest difference in performance between the two conditions occurred for P=8, with 98.2% accuracy in the explicit and 93.3% accuracy in the implicit condition. Cochran's Q test showed a significant effect of the number of training participants P on the false detections rate ($\chi^2_{(7,N=3600)}$=103.327, p<.001); see Figure 13. The average false detection rate was 3.02% and reached the minimum value 1.11% ($\text{CI}_{95\%} = [0.77\%, 1.45\%]$) for P=8 training participants. The F-measure increased from 0.26 (P=1) to 0.96 (P=8), confirming the high performance of the *brute-force* technique.

8. Discussion

We discuss in this section users' performance with Smart-Pockets and Smart-Containers access actions and we also point to alternative technology to implement our recognition techniques. We discuss pro and con arguments for explicit versus implicit segmentation of body-deictic gestures, and we conclude that explicit segmentation represents our preferred solution for implementing Smart-Pockets in practical ambient scenarios at this moment. We also point to interesting research directions regarding body-deictic gestures and to further developments envisioned for the Smart-Pockets concept.

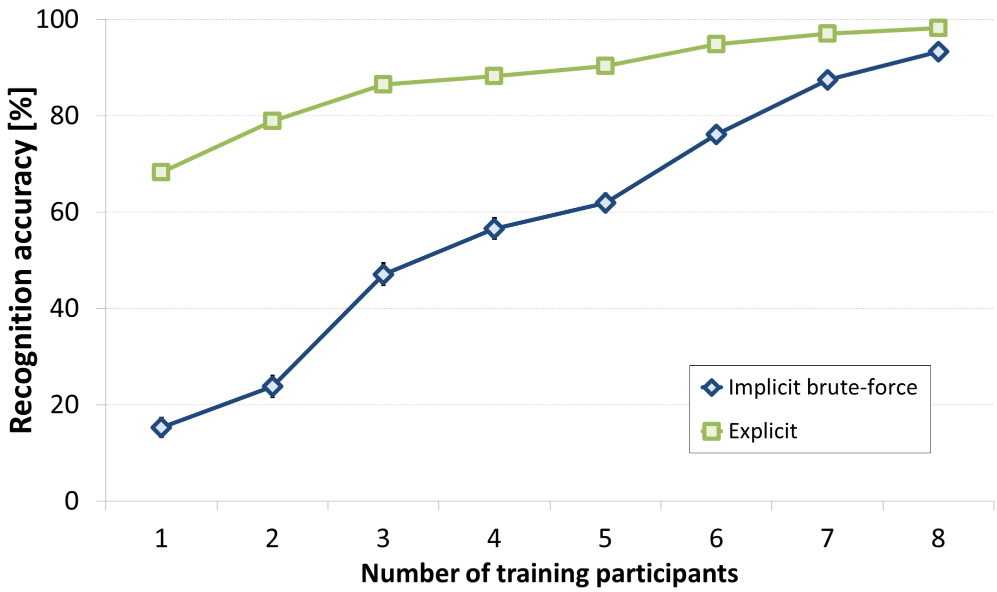

Figure 12: User-independent recognition accuracy of Smart-Pockets and Smart-Containers access actions for the DTW-Hands recognizer and *implicit segmentation*, *i.e.*, actions are automatically segmented by the system. Notes: error bars (too small to be visualized in this figure) show 95% CIs; the performance of the recognizer in the explicit condition is also shown (in green) for comparison purposes, as it represents the upper margin of accuracy expected for Smart-Pockets and Smart-Containers.

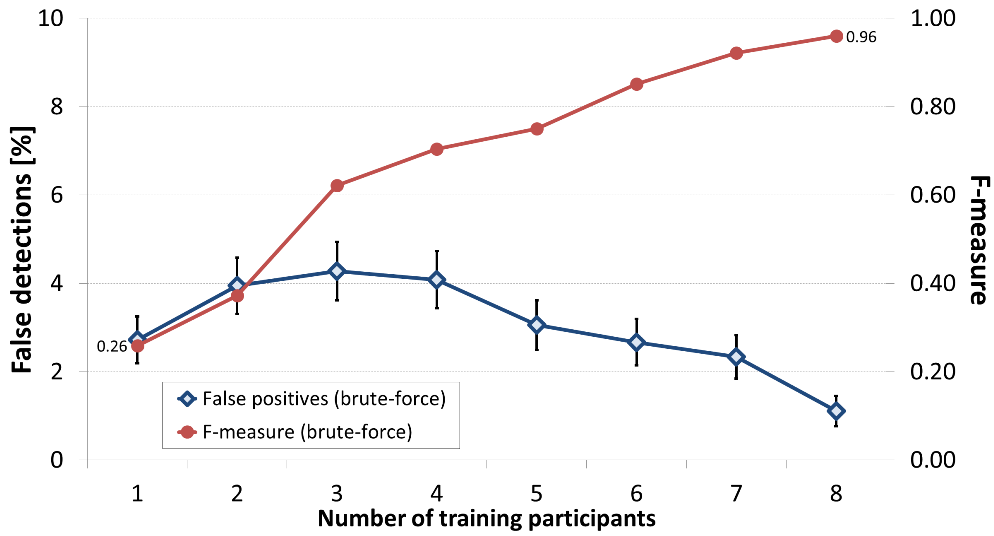

Figure 13: False detection rates and F-measure values for Smart-Pockets and Smart-Containers access actions for the *brute-force* implicit recognition technique. Note: error bars show 95% CIs; higher F-measure values (closer to 1.0) indicate better performance.

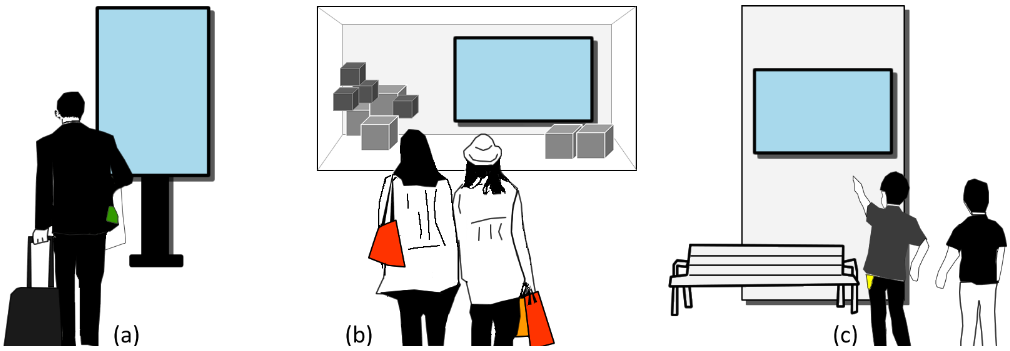

Figure 14: Illustrative examples for Smart-Pockets and Smart-Containers providing assistance to ambient users and consumers of ambient services during traveling (a), shopping (b), and consuming multimedia content in public spaces (c); see discussion in the text.

8.1. Usage scenarios for Smart-Pockets and Smart-Containers

We start the Discussion section with a consideration of a few prospective usage scenarios for Smart-Pockets and Smart-Containers to point readers to practical, relevant examples of everyday interactions with public ambient displays enabled by our Smart-Pockets concept. In the following, we show through a series of fictional stories how Smart-Pockets and Smart-Containers can find applications in a variety of interactive contexts (*e.g.*, collection of public information from ambient displays, multimedia rendering, or retrieval of personalized content from public information displays) and how they can be customized accordingly to suit various needs; see Figure 14:

(a) John is a frequent traveler, as a result of his job responsibilities. He travels mostly by train to connect between cities and takes transit buses in the city. As he has been doing his job for a long time, he has learned time schedules by heart for many of his train and bus connections. However, he also needs to travel to new cities quite often and, sometimes, it happens that even time schedules for his frequent routes may change unexpectedly in the last minute, because of delays or other events in public transportation. To still be on time for his day agenda, John needs to be able to make fast decisions, sometimes within minutes or even seconds, to change his travel connections or even the entire travel route to his destination. John keeps a link to his day agenda in the right pocket of his vest. For time critical situations, he uses this Smart-Pockets connection to take the link from his pocket and point it to an information display with train time schedules; see Figure 14a. The ambient display reads John's calendar and instantly highlights the fastest route and the first connection that John needs to take to arrive on time at his destination. John can save this information by storing it into his left pocket, and use it later to receive confirmation or more directions inside the train station by pointing it to other ambient displays, while he heads to his train platform. Displaying personal events from one's calendar on public displays for various applications has been explored by researchers with various technologies and interactive techniques (Vogel and Balakrishnan, 2004; Cao et al., 2007), while route planning using ambient displays to help travelers navigate in public transportation has started to receive attention recently (De Marchi et al., 2015). In this context, Smart-Pockets enables fast access to personal content, such as John's day agenda in our example, reducing the search time for retrieving that content to under 2.5 seconds (see next for performance measures for Smart-Pockets).

(b) Esther and Olivia are on a day out in the city. While they have fun window shopping, they sometimes see interesting items on display, and they enter the various shops. Other times, however, shops are unfortunately closed, yet the window displays advertise savings and deals that the two girls cannot miss; see Figure 14b. To save that information, Olivia points to the display and then to her purse: a link to the specific offers from that shop is now mapped by a Smart-Container connection directly to Olivia's purse. Olivia continues to link her Smart-Container purse to various pieces of shopping offers which, at the end of the day, she can browse, learn more about, and share with Esther. Assisting users accessing retail services in public places has been an active research direction in the ambient intelligence community (Garcia-Perate et al., 2013; Meschtscherjakov et al., 2009; van Doorn et al., 2008) and also for wearables and smart textiles; *e.g.*, the "Shopping Jacket" of Randell and Muller (2000) enables both subliminal and active shopping by submitting the owner's shopping list to nearby shops. Smart-Containers, such as purses and bags, create an intuitive mapping between purchased items and physical containers to carry those items, even when the sale is done entirely electronically and purchased items will be later dispatched by the shop to the buyer's address.

(c) Joseph and David, two 12-year-old boys, enjoy playing video games on their new game console. However, it just happens that today they need to attend to their chores, outside the house. Unfortunately, this means that they need to pause their video game, right in the middle of its exciting action. Joseph points his hand toward the TV and, from the menu that pops up on the screen, he selects "Save game state." As the console performs the action, Joseph puts his hand into his trousers' front pocket. As a result, a link to the current game state has been mapped by a Smart-Pockets connection directly to Joseph's pocket, making the game state easily accessible at any time, just like any other physical pocket game. The boys leave the house and head to the bus station. However, today is different from other days. While they wait at the station for their bus to arrive, Joseph takes out the game state from his pocket and points it to a nearby display; see Figure 14c. The game resumes, and the two boys are able to continue playing and, in just few minutes, they succeed advancing to the next level. As the bus arrives, Joseph carefully puts back the new saved game state into his pocket, ending the connection with the public display. Ambient games have been designed in the community for various purposes, such as to foster interaction, socialization, and learning in public spaces, but also to implement persuasion to change behaviors (Korozi et al., 2012; Kuramoto et al., 2013; Salvador and Romão, 2011), while a recent study showed that interruptions and pausing of games on public displays occur often (Feuchtner et al., 2016). In this context, Smart-Pockets represent an efficient way to access personal multimedia content for consumption in public spaces, enabling fast retrieval, play, and resume operations that treat digital content, *e.g.*, the game state in our example, just like any another personal possession stored in one's pockets.

8.2. Explicit versus implicit segmentation for Smart-Pockets and Smart-Containers

The results of our two recognition experiments showed that Smart-Pockets access actions can be recognized with 99.4% accuracy in the *explicit* segmentation condition and with 94.0% accuracy under *implicit* segmentation using the *brute-force* approach. Combined Smart-Pockets and Smart-Containers actions were recognized with 98.8% and 93.3% accuracy in the explicit and implicit segmentation conditions, respectively. The immediate advantage of explicit over implicit detection of Smart-Pockets and Smart-Containers access actions clearly emerges from these results.

However, more aspects need to be considered for this discussion. For instance, explicit segmentation has the benefit of creating a tight coupling between the user's interaction intentions and system responses. Users are familiar with click-like events from operating standard desktop PCs that expose WIMP (windows, icons, menus, pointer) interaction styles. Users are also familiar with tap input on touch-screen devices, where a tap simply acts as another implementation of a "click." Practically, almost all current styles of interaction with computers have been heavily based on the notion of a "click." Furthermore, the literature has shown that users move fluently between different interaction zones and phases when they interact with public ambient displays, while they shift from pointing to direct touch interaction (Vogel and Balakrishnan, 2004; Müller et al., 2014). In this context, reusing concepts, such as the "click," during the multiple interaction phases of the same user interface represents a good design decision. Lastly, explicit interaction gives the users the feeling of being in control as they tell the system when their actions should be interpreted. From this perspective, explicit segmentation creates the premises for a fluent user experience, for which command execution is clearly defined by start and stop events. Our recognition results already recommend the explicit segmentation approach over implicit techniques due to increased accuracy. The above arguments strengthen our conviction that explicit segmentation is the preferred choice for implementing Smart-Pockets at this moment.

There are many ways to implement click-like events for explicit segmentation, such as by using a secondary sensor, worn on the body or held in hand, or by performing a specific hand gesture that acts as a delimiter for Smart-Pockets gestures. For instance, through-fabric capacitive touch input (Saponas et al., 2011; Heller et al., 2014), mid-air devices (Baudisch et al., 2006; Wilson and Shafer, 2003) or specific hand postures that delimit meaningful motion (Haque et al., 2015; Malik et al., 2005; Vatavu et al., 2009) represent good candidates of techniques that deliver click-like events to segment Smart-Pockets gestures in continuous whole-body movement. Our preferred choice is the PocketTouch technique (Saponas et al., 2011) that places sensors inside the pocket to enable touch sensing through the garment. Also, consumer-level electromyographic sensors, such as the Myo armband, represent a good alternative for implementing explicit segmentation. Note that electromyographic and IMU sensors have been recently adopted by researchers for implementing techniques to point and control content on large displays (Haque et al., 2015).

However, implicit segmentation should not be discarded entirely. After all, implicit segmentation requires no instrumentation of clothes and no in-between devices to mediate interaction. Even though our recognition results show only 94.0% accuracy for Smart-Pockets and 93.3% accuracy for Smart-Pockets and Smart-Containers combined, recognition performance is likely to improve with more training samples per gesture type

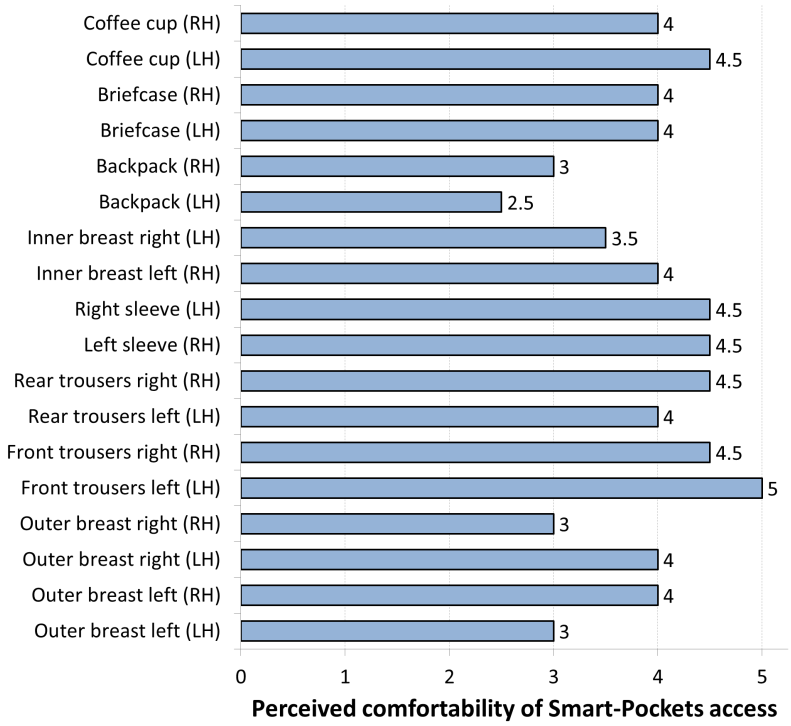

Figure 15: Participants' perceived comfortability of accessing Smart-Pockets and Smart-Containers. NOTE: rating values were 1–"uncomfortable", 2–"somewhat uncomfortable", 3–"medium", 4–"comfortable", and 5–"very comfortable".

and data from more training participants. Future work will look at practical ways to improve recognition accuracy for the implicit segmentation approach.

8.3. Users' perceived experience with Smart-Pockets and Smart-Containers

We asked participants to evaluate their perceived comfortability of accessing each POCKET and CONTAINER with ratings on a 5-point Likert scale ranging from 1 ("uncomfortable") to 5 ('very comfortable"). The median rating across all conditions was 4 ("comfortable"). Figure 15 illustrates ratings for each POCKET and CONTAINER. The least comfortable locations were the *backpack* when accessed with the left hand (2.5, between "somewhat uncomfortable" and "medium"), *backpack* accessed with the right hand, and the *outer breast right* and *left pockets* (3, "medium"). The pocket that was the most comfortably accessed was the *front trousers left pocket* with the left hand (median rating 5, "very comfortable").

We also asked participants to rate how frequent they used pockets in daily life on a 5-point Likert scale with ratings ranging from 1 ("never") to 5 ("very often"); see Figure 16 for results. The most used pockets and containers were *front trousers left* and *right* (median ratings 5, "very often"), followed by *briefcase* (4, "often"), and *rear trousers right* and *left pockets* (3.5, between "now and then" and "often"). The least used pockets were the *inner* and *outer breast pockets*. These results inform the design of Smart-Pockets gestures to specific locations on the body that are frequently employed by users. They also suggest ideas for designing specific associations; for instance, while frequently-used pockets may be associated to frequently-accessed digital content, a specific association between content and a rarely-used pocket might also be considered to reflect a particular aspect of that content.

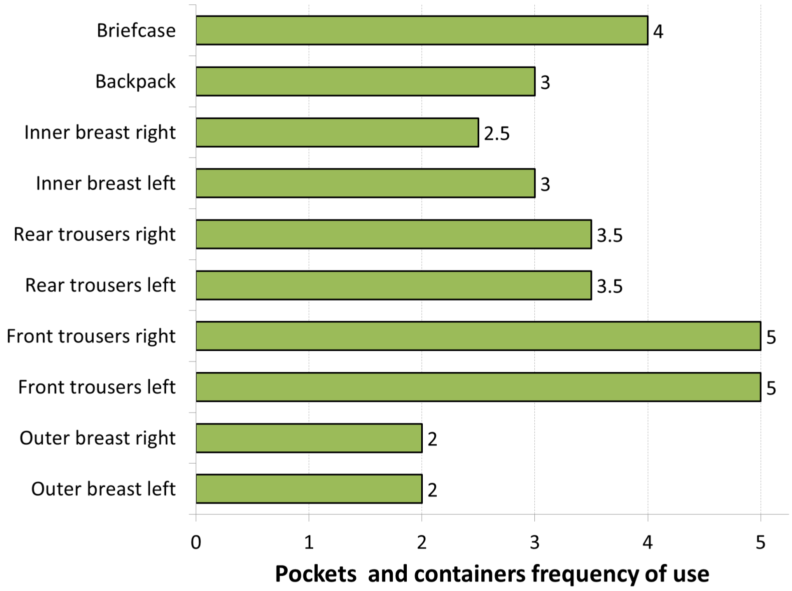

Figure 16: Participants' self-reported use of pockets and containers. NOTE: rating values were 1−"never", 2−"very rare", 3−"now and then", 4−"often", and 5−"very often."

8.4. Kinematics of Smart-Pockets and Smart-Containers access actions

We examine in this section our participants' performance with Smart-Pockets and Smart-Containers gestures, which we measure using a new dependent variable, ACCESS-TIME, expressed in seconds. Overall, the time that our participants needed to access Smart-Pockets and Smart-Containers varied between 1.56 and 3.53 seconds with a mean of 2.22 seconds (SD=0.41); see Figure 17 for both individual and average ACCESS-TIME values. A Friedman test revealed a significant effect of POCKET and CONTAINER on ACCESS-TIME ($\chi^2(14)$=49.378, p<.001). The fastest access occurred for the outer breast pockets, as follows: the *outer breast left pocket* was accessed the fastest with the right hand (2.01 seconds), followed closely by the *outer breast right pocket* accessed with the right hand (2.05 seconds), the *outer breast left pocket* with the left hand (2.07 seconds) and the *outer breast right pocket* with the left hand (2.09 seconds). In between, the *coffee cup* container was accessed the fastest among all Smart-Containers with both the right and left hands (2.02 and 2.03 seconds, respectively). The explanation of these findings probably lies with the fact that the outer breast pockets and the coffee cup are at the shortest distance from the hand pointed in front of the body. The slowest access was

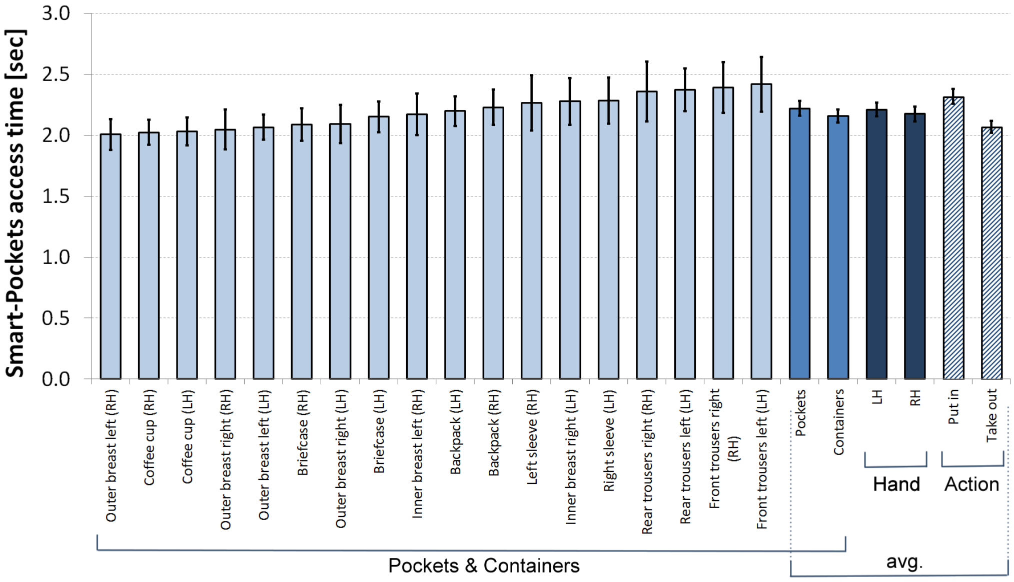

Figure 17: Access times for each Pocket and Container experimental conditions and average access times for each Hand and Action condition.

registered for the *front trousers left* and *right pockets* when access was performed on the same side of the body (2.39 and 2.42 seconds, respectively); see Figure 17. Overall, the difference between the fastest and the slowest Access-Time of all Pocket × Hand combinations was under a half of a second (0.41 seconds).

We found that access times were significantly larger when users put content into their pockets than when they took content out (2.32 versus 2.07 seconds, Z=−6.070, p<.001, r=−.369). We found no significant effect of Hand on Access-Time (2.21 versus 2.18 seconds, Z=−1.729, *n.s.* at p=.05). Also, we found no significant effect of the type of smart container (*i.e.*, Pocket or Container) on Access-Time (2.22 versus 2.16 seconds, Z=−.392, *n.s.* at p=.05); see Figure 17 for average access times for each condition of the Hand and Action independent variables.

These findings show that access to Smart-Pockets is fast (*i.e.*, 2.2 seconds on average) and that the location of the pocket on the body has a significant, yet small influence on the time needed to access that pocket. To understand these results better, we positioned them in a broader context by relating to similar results reported in the literature. For instance, Ashbrook et al. (2008) were interested in the effect of placement of on-body interfaces on the time required to access those interfaces. In their study, similar in size and participants' age range to ours (15 participants, mean age 24.87 years, see p. 221), they found that participants' average access times were 2.78 seconds for the wrist, 4.62 seconds for pocket, and 5.52 seconds for the hip (p. 221-222). We also wanted to learn how Smart-Pockets gestures compare with other whole-body gestures and, in general, to other types of gestures as well, such as touch and pen gestures performed on smart mobile devices. To this end, we computed the average production times of gestures from several public datasets (Hoffman et al., 2010; Chen et al., 2012a; Wobbrock et al., 2007;

Anthony and Wobbrock, 2010; Vatavu et al., 2011; Fothergill et al., 2012), as follows:

- The Wiimote motion gestures dataset of Hoffman et al. (2010) consisting in 25 distinct acceleration gestures performed by 17 participants for 20 times, with a total number of $25 \times 17 \times 20 = 8500$ gesture samples; the dataset is available for download from the Interactive Systems & User Experience Research Cluster of Excellence website[9].

- The 6DMG dataset of Chen et al. (2012a), composed of 20 acceleration gestures performed by 28 participants with 10 repetitions, with a total number of 5600 samples; available from the 6DMG website[10].

- The $1 stroke gesture dataset of Wobbrock et al. (2007) composed of 16 distinct unistroke gestures performed by 10 participants at 3 speeds (slow, medium, and fast) for 10 times using the stylus on a Pocket PC, with 4800 gesture samples in total; the dataset is available to download from the $1 Unistroke Recognizer website[11].

- The MMG multi-stroke gesture dataset of Anthony and Wobbrock (2010) composed of 16 multi-stroke gestures performed by 20 participants for 10 times with the stylus and the finger, with 3200 total samples; the dataset is available from the $N Recognizer website[12].

- The two Unistroke Gesture Difficulty datasets of Vatavu et al. (2011) consisting in 38 distinct unistroke gestures performed by 25 participants with a stylus on a Wacom interactive display, with 9440 samples in total; the two datasets are available at the Gesture Difficulty website[13].

- The Microsoft Research Cambridge-12 Kinect gesture dataset of Fothergill et al. (2012), consisting of 6244 whole-body gesture samples of 12 distinct gestures performed by 30 participants; the dataset is available to download from the MSRC-12 website[14].

We wanted to learn how Smart-Pockets access times compare to various gesture types performed in various conditions. In total, we compared Smart-Pockets access times with 14100 accelerated motion gestures, 17440 unistroke and multi-stroke finger and stylus gestures, and 6244 whole-body gestures. Figure 18 shows the summary of our findings. Smart-Pockets gestures (average 2.22 seconds) were comparable in terms of production time to unistroke and multi-stroke gestures performed on tablets (average 2.01 and 2.09 seconds) (Anthony and Wobbrock, 2010; Vatavu et al., 2011) and they were 60% faster than other whole-body gestures (Fothergill et al., 2012). Overall, these results show that Smart-Pockets in the form of body-deictic and deictic gestures are very efficient to perform, making them suitable for implementing fast ambient interactions.

---

[9] http://www.eecs.ucf.edu/isuelab/downloads.php
[10] http://www.ece.gatech.edu/6DMG/6DMG.html
[11] https://depts.washington.edu/aimgroup/proj/dollar/
[12] http://depts.washington.edu/aimgroup/proj/dollar/ndollar.html
[13] http://www.eed.usv.ro/~vatavu/index.php?menuItem=pengestures2011
[14] http://research.microsoft.com/en-us/um/cambridge/projects/msrc12/

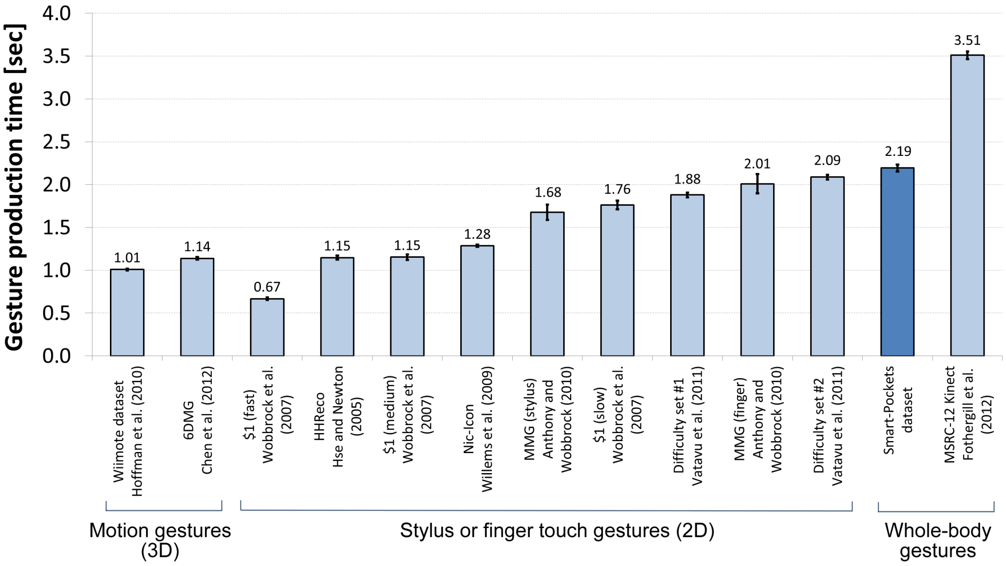

Figure 18: Average production times computed for touch, stroke, motion, and whole-body gestures from public datasets. Note how the Smart-Pockets average access time is very close to touch gesture production time (Anthony and Wobbrock, 2010; Vatavu et al., 2011) and smaller than the production time of other whole-body gestures (Fothergill et al., 2012). Error bars show 95% CIs.

8.5. Alternative technology and future work on Smart-Pockets and body-deictic gestures

In this work, we demonstrated the Smart-Pockets concept with an implementation using whole-body gestures captured with the Microsoft Kinect sensor. However, Smart-Pockets implementation is flexible and other technology can be considered to achieve the same effect. For instance, sensors worn at arm level report the movement characteristics of the hands in 3-D space: smart-watches, the Myo armband[15], the Leap Motion controller[16], and the Ring Zero sensor[17] are just a few examples. These devices incorporate accelerometers and gyroscopes that report many parameters of the hand movement in real-time and have been recently considered for interactions with ambient displays; see the "Gunslinger" prototype of Liu et al. (2015) and the "Myopoint" technique of Haque et al. (2015). While such sensing equipment is very practical to consider for implementing interactions with Smart-Pockets and Smart-Containers for outdoor scenarios, such as the ones discussed at the beginning of this section, more punctilious investigations of users' performance with body-deictic gestures can be achieved with more precise measurements delivered by state-of-the-art motion capture systems. For instance, Figure 19 shows an example of a participant's body-deictic gestures tracked by a Vicon Bonita system. Note that the mathematical formalism that we introduced in this work for classifying and detecting Smart-Pockets and Smart-Containers access actions in sequences of continuous whole-body gesture movement is generic, not tied to any particular motion

[15] `https://www.myo.com/`
[16] `https://www.leapmotion.com/`
[17] `http://ringzero.logbar.jp/`

Figure 19: Our mathematical formalism for classification and detection of Smart-Pockets and Smart-Containers actions access can be used with any motion tracking device and any types of body-deictic gestures. In this figure, a Vicon Motion Capture system tracks the user's dominant hand and a large display shows a collection of generic graphical items.

sensing equipment and readily applicable to work with any number of joints tracked on the human body as well as to any type of body-deictic gestures.

We list below a few interesting directions for future work regarding body-deictic gestures:

1. Examination of the design space of body-deictics by considering factors as such gesture type, performance and accuracy, but also social acceptability for gestures produced in public (Ahlström et al., 2014; Montero et al., 2010; Rico and Brewster, 2009, 2010) or for gestures performed on the body (Harrison and Faste, 2014; Profita et al., 2013). A taxonomy for body-deictic gestures will also be useful to guide further explorations of potential applications.
2. Precise measurements of user performance with body-deictic gestures in terms of pointing accuracy, production speed, perceived difficulty, and recall rates. Specific measures to evaluate the geometric and kinematic accuracy of body-deictics might be useful to capture and report users' performance, similar to existing measures available for other types of gestures (Vatavu, 2017; Vatavu et al., 2014, 2013), general pointing (Fitts, 1954; MacKenzie et al., 2001), or to the velocity profile measurements of the hand performing movement in the context of the Kinematic Theory of Rapid Human Movements (Leiva et al., 2017; Plamondon et al., 2014).
3. Further investigations of application opportunities, including the design and evaluation of interaction techniques based on body-deictic gestures, as well as integration with other gesture types, such as whole-body gesture input (Lou et al., 2016; Vatavu, 2012a; Walter et al., 2013, 2014) or touch input on mobile and wearable personal devices (Lee et al., 2011; Perrault et al., 2013; Vatavu et al., 2016) will lead to richer interactive experiences for users.

The Smart-Pockets concept builds on the "pocket metaphor": personal belongings are held in one's pockets, from which they are easily retrieved. However, although metaphors play an important part in user interface design to help users understand and experience new concepts in relation to their knowledge about the real world (Hamilton, 2000; Lakoff and Johnson, 1980; Sanford et al., 2014), we also acknowledge that metaphors may not always be appropriate as they might break the operation consistency across the digital and physical worlds, especially for tangible interactions (Bakker et al., 2012; Celentano and Dubois, 2012; Oppl and Stary, 2011; Svanaes and Verplank, 2000). Nevertheless, we believe that the simple operation of Smart-Pockets is likely to make our technique free of such metaphor consistency problems. In fact, Smart-Pockets represent an instance of both "body-inspired" and "body-as-a-surface" metaphors, as inventoried by Cho and Yang (2002) and Kim et al. (2004). Body-inspired metaphors employ parts of the body for interaction, such as touching the head, pointing to the ear, or even pointing to specific parts on clothes, such as pockets, to trigger a command. The body as interaction surface repurposes parts of the body as the physical support for interactions, such as the forearm can turn into an interactive surface. As body-inspired metaphors were found to produce the lowest error rates among four different types of body-based interfaces (Kim et al., 2004), Smart-Pockets may benefit of similar advantages. However, such a hypothesis needs to be examined by future studies. We list below several future work directions for Smart-Pockets and Smart-Containers:

1. Further evaluations of the naturalness and intuitiveness of the pocket metaphor for accessing digital contents are needed, with experimental conditions involving various object types and application contexts, *e.g.*, document browsing, editing, information retrieval, etc.
2. Specific interaction techniques need to be designed for Smart-Pockets and Smart-Containers to act as basis for evaluating the appropriateness of the pocket metaphor. The recognition techniques that we introduced in this work showed that Smart-Pockets and Smart-Containers access actions can be reliably recognized with high accuracy in both explicit and implicit interaction contexts. Our results create the basis for future integration of Smart-Pockets into actual interaction techniques for various application contexts.
3. Future explorations regarding users transitioning between implicit and explicit interaction (Vogel and Balakrishnan, 2004) will likely reveal new interesting findings on people's capacity to use Smart-Pockets and Smart-Containers in public scenarios. For instance, some frequently occurring Smart-Pockets access actions may need to be explicitly segmented to help the ambient system avoid disambiguation against pocket access for actual physical objects, such as the phone. Understanding the context of the interaction (Lopes et al., 2012; Dourish, 2004) can play a key part in achieving natural and fluent interactions with ambient displays mediated by Smart-Pockets access actions.
4. Integration of the Smart-Pockets concept in the recent context of advances in prototyping smart garments and designing gesture-based interaction techniques for smart textiles (Profita et al., 2013; Heller et al., 2014). Recent efforts, such as the "Interactex" visual programming environment of Haladjian et al. (2016) will likely foster new developments in the community toward moving smart garments technology from niche to mass production (Cheng et al., 2013). c recognition techniques for body-deictics in such tools will enable readily integration of Smart-Pockets into the next generation of natural user interfaces.

9. Conclusion

We introduced in this work Smart-Pockets, a new concept and set of recognition techniques that rely on the pocket metaphor and body gestures to enable users to access personal content: the users' pockets act as placeholders or links to their digital content, just how conventional pockets facilitate access to one's personal belongings. We implemented Smart-Pockets by introducing and evaluating three whole-body gesture detection and recognition approaches. Our experimental evaluations showed that Smart-Pockets gestures are fast and robustly recognized (99%) in user-independent conditions, while the concept is easily extensible to include other physical containers as well. We hope that the Smart-Pockets concept will inspire researchers and practitioners to explore the opportunities offered by this new type of body-deictic gestures to create novel ways to associate physical objects with digital content in the context of ambient interactions.

10. Acknowledgments

This work was supported from the project *"Interaction Techniques with Massive Data Clouds in Smart Environments"*, project no. 47BM/2016, financed by UEFISCDI, Romania. The work was carried out in the Machine Intelligence and Information Visualization Lab (MintViz) of the MANSiD Research Center. The infrastructure was provided by the University of Suceava and was partially supported from the project *"Integrated center for research, development and innovation in Advanced Materials, Nanotechnologies, and Distributed Systems for fabrication and control"*, No. 671/09.04.2015, Sectoral Operational Program for Increase of the Economic Competitiveness, co-funded from the European Regional Development Fund.

References

Ahlström, D., Hasan, K., Irani, P., 2014. Are you comfortable doing that?: Acceptance studies of around-device gestures in and for public settings. In: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services. MobileHCI '14. ACM, New York, NY, USA, pp. 193–202.
URL `http://doi.acm.org/10.1145/2628363.2628381`

Anderson, F., Grossman, T., Wigdor, D., Fitzmaurice, G., 2015. Supporting subtlety with deceptive devices and illusory interactions. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15. ACM, New York, NY, USA, pp. 1489–1498.
URL `http://doi.acm.org/10.1145/2702123.2702336`

Ängeslevä, J., Oakley, I., Hughes, S., O'Modhrain, S., 2003. Body mnemonics: Portable device interaction design concept. In: Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology. UIST '03.
URL `https://uist.acm.org/archive/adjunct/2003/pdf/posters/p3-angesleva.pdf`

Annett, M., Grossman, T., Wigdor, D., Fitzmaurice, G., 2011. Medusa: A proximity-aware multi-touch tabletop. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11. ACM, New York, NY, USA, pp. 337–346.
URL `http://doi.acm.org/10.1145/2047196.2047240`

Anthony, L., Wobbrock, J. O., 2010. A lightweight multistroke recognizer for user interface prototypes. In: Proceedings of Graphics Interface 2010. GI '10. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, pp. 245–252.
URL `http://dl.acm.org/citation.cfm?id=1839214.1839258`

Arase, Y., Ren, F., Xie, X., 2010. User activity understanding from mobile phone sensors. In: Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct. UbiComp '10 Adjunct. ACM, New York, NY, USA, pp. 391–392.
URL `http://doi.acm.org/10.1145/1864431.1864452`

Ardito, C., Buono, P., Costabile, M. F., Desolda, G., Feb. 2015. Interaction with large displays: A survey. ACM Comput. Surv. 47 (3), 46:1–46:38.
URL `http://doi.acm.org/10.1145/2682623`

Ashbrook, D. L., 2010. Enabling mobile microinteractions. Ph.D. thesis, Atlanta, GA, USA, aAI3414437.

Ashbrook, D. L., Clawson, J. R., Lyons, K., Starner, T. E., Patel, N., 2008. Quickdraw: The impact of mobility and on-body placement on device access time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08. ACM, New York, NY, USA, pp. 219–222.
URL http://doi.acm.org/10.1145/1357054.1357092

Aumi, M. T. I., Kratz, S., 2014. AirAuth: Towards attack-resilient biometric authentication using in-air gestures. In: CHI '14 Extended Abstracts on Human Factors in Computing Systems. CHI EA '14. ACM, New York, NY, USA, pp. 1585–1590.
URL http://doi.acm.org/10.1145/2559206.2581157

Avellino, I., Fleury, C., Beaudouin-Lafon, M., 2015. Accuracy of deictic gestures to support telepresence on wall-sized displays. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15. ACM, New York, NY, USA, pp. 2393–2396.
URL http://doi.acm.org/10.1145/2702123.2702448

Bailly, G., Walter, R., Müller, J., Ning, T., Lecolinet, E., 2011. Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. In: Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part II. INTERACT'11. Springer-Verlag, Berlin, Heidelberg, pp. 248–262.
URL http://dl.acm.org/citation.cfm?id=2042118.2042143

Bakker, S., Antle, A. N., Van Den Hoven, E., Apr. 2012. Embodied metaphors in tangible interaction design. Personal Ubiquitous Comput. 16 (4), 433–449.
URL http://dx.doi.org/10.1007/s00779-011-0410-4

Ballendat, T., Marquardt, N., Greenberg, S., 2010. Proxemic interaction: Designing for a proximity and orientation-aware environment. In: ACM International Conference on Interactive Tabletops and Surfaces. ITS '10. ACM, New York, NY, USA, pp. 121–130.
URL http://doi.acm.org/10.1145/1936652.1936676

Baudisch, P., Sinclair, M., Wilson, A., 2006. Soap: A pointing device that works in mid-air. In: Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology. UIST '06. ACM, New York, NY, USA, pp. 43–46.
URL http://doi.acm.org/10.1145/1166253.1166261

Bellezza, F., 1996. Mnemonic methods to enhance storage and retrieval. In: Bjork, E., Bjork, R. (Eds.), Memory: Handbook of perception and cognition. San Diego, CA, Academic Press, pp. 345–380.

Bolt, R. A., 1980. "Put-that-there": Voice and gesture at the graphics interface. In: Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '80. ACM, New York, NY, USA, pp. 262–270.
URL http://doi.acm.org/10.1145/800250.807503

Börner, D., Kalz, M., Specht, M., Jul. 2013. Closer to you: Reviewing the application, design, and evaluation of ambient displays. International Journal of Ambient Computing and Intelligence 5 (3), 16–31.
URL http://dx.doi.org/10.4018/ijaci.2013070102

Cao, X., Forlines, C., Balakrishnan, R., 2007. Multi-user interaction using handheld projectors. In: Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology. UIST '07. ACM, New York, NY, USA, pp. 43–52.
URL http://doi.acm.org/10.1145/1294211.1294220

Castellucci, S. J., Teather, R. J., Pavlovych, A., 2013. Novel metrics for 3D remote pointing. In: Proceedings of the 1st Symposium on Spatial User Interaction. SUI '13. ACM, New York, NY, USA, pp. 17–20.
URL http://doi.acm.org/10.1145/2491367.2491373

Celentano, A., Dubois, E., 2012. Metaphor modelling for tangible interfaces evaluation. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. AVI '12. ACM, New York, NY, USA, pp. 78–81.
URL http://doi.acm.org/10.1145/2254556.2254573

Chen, M., AlRegib, G., Juang, B.-H., 2012a. 6DMG: A new 6D motion gesture database. In: Proceedings of the 3rd Multimedia Systems Conference. MMSys '12. ACM, New York, NY, USA, pp. 83–88.
URL http://doi.acm.org/10.1145/2155555.2155569

Chen, X. A., Marquardt, N., Tang, A., Boring, S., Greenberg, S., 2012b. Extending a mobile device's interaction space through body-centric interaction. In: Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services. MobileHCI '12. ACM, New York, NY, USA, pp. 151–160.
URL http://doi.acm.org/10.1145/2371574.2371599

Cheng, J., Lukowicz, P., Henze, N., Schmidt, A., Amft, O., Salvatore, G. A., Troster, G., Jul. 2013. Smart textiles: From niche to mainstream. IEEE Pervasive Computing 12 (3), 81–84.
URL http://dx.doi.org/10.1109/MPRV.2013.55

Cho, C., Yang, H., 2002. Body-based interfaces. In: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces. ICMI '02. IEEE Computer Society, Washington, DC, USA, pp. 466–.
URL http://dx.doi.org/10.1109/ICMI.2002.1167040

Claes, S., Moere, A. V., 2015. The role of tangible interaction in exploring information on public visualization displays. In: Proceedings of the 4th International Symposium on Pervasive Displays. PerDis '15. ACM, New York, NY, USA, pp. 201–207.
URL http://doi.acm.org/10.1145/2757710.2757733

De Luca, A., von Zezschwitz, E., Hußmann, H., 2009. Vibrapass: Secure authentication based on shared lies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09. ACM, New York, NY, USA, pp. 913–916.
URL http://doi.acm.org/10.1145/1518701.1518840

De Marchi, M., Eriksson, J., Forbes, A. G., 2015. Transittrace: Route planning using ambient displays. In: Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems. SIGSPATIAL '15. ACM, New York, NY, USA, pp. 67:1–67:4.
URL http://doi.acm.org/10.1145/2820783.2820857

De Silva, S., Barlow, M., Easton, A., 2014. An evaluation of DTW approaches for whole-of-body gesture recognition. In: Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI. BCS-HCI '14. BCS, UK, pp. 11–21.
URL http://dx.doi.org/10.14236/ewic/hci2014.2

Dingler, T., Funk, M., Alt, F., 2015. Interaction proxemics: Combining physical spaces for seamless gesture interaction. In: Proceedings of the 4th International Symposium on Pervasive Displays. PerDis '15. ACM, New York, NY, USA, pp. 107–114.
URL http://doi.acm.org/10.1145/2757710.2757722

Dourish, P., Feb. 2004. What we talk about when we talk about context. Personal Ubiquitous Comput. 8 (1), 19–30.
URL http://dx.doi.org/10.1007/s00779-003-0253-8

Feuchtner, T., Walter, R., Müller, J., 2016. Interruption and pausing of public display games. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '16. ACM, New York, NY, USA, pp. 306–317.
URL http://doi.acm.org/10.1145/2935334.2935335

Fishkin, K. P., September 2004. A taxonomy for and analysis of tangible interfaces. Personal and Ubiquitous Computing 8 (5), 347–358.
URL http://dx.doi.org/10.1007/s00779-004-0297-4

Fitts, P. M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47 (6), 381–391.
URL http://dx.doi.org/10.1037/h0055392

Fothergill, S., Mentis, H., Kohli, P., Nowozin, S., 2012. Instructing people for training gestural interactive systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12. ACM, New York, NY, USA, pp. 1737–1746.
URL http://doi.acm.org/10.1145/2207676.2208303

Garcia-Perate, G., Dalton, N., Conroy-Dalton, R., Wilson, D., 2013. Ambient recommendations in the pop-up shop. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '13. ACM, New York, NY, USA, pp. 773–776.
URL http://doi.acm.org/10.1145/2493432.2494525

Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M., Jan. 2011. Proxemic interactions: The new ubicomp? interactions 18 (1), 42–50.
URL http://doi.acm.org/10.1145/1897239.1897250

Guerreiro, T., Gamboa, R., Jorge, J., 2008. Mnemonical body shortcuts: Improving mobile interaction. In: Proceedings of the 15th European Conference on Cognitive Ergonomics: The Ergonomics of Cool Interaction. ECCE '08. ACM, New York, NY, USA, pp. 11:1–11:8.
URL http://doi.acm.org/10.1145/1473018.1473033

Haladjian, J., Bredies, K., Brügge, B., 2016. Interactex: An integrated development environment for smart textiles. In: Proceedings of the 2016 ACM International Symposium on Wearable Computers. ISWC '16. ACM, New York, NY, USA, pp. 8–15.
URL http://doi.acm.org/10.1145/2971763.2971776

Hamilton, A., Nov. 2000. Metaphor in theory and practice: The influence of metaphors on expectations.

ACM J. Comput. Doc. 24 (4), 237–253.
URL http://doi.acm.org/10.1145/353927.353935

Haque, F., Nancel, M., Vogel, D., 2015. Myopoint: Pointing and clicking using forearm mounted electromyography and inertial motion sensors. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15. ACM, New York, NY, USA, pp. 3653–3656.
URL http://doi.acm.org/10.1145/2702123.2702133

Harms, H., Amft, O., Roggen, D., Tröster, G., Apr. 2009. Rapid prototyping of smart garments for activity-aware applications. J. Ambient Intell. Smart Environ. 1 (2), 87–101.
URL http://dl.acm.org/citation.cfm?id=1735835.1735837

Harrison, C., Benko, H., Wilson, A. D., 2011. Omnitouch: Wearable multitouch interaction everywhere. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11. ACM, New York, NY, USA, pp. 441–450.
URL http://doi.acm.org/10.1145/2047196.2047255

Harrison, C., Faste, H., 2014. Implications of location and touch for on-body projected interfaces. In: Proceedings of the 2014 Conference on Designing Interactive Systems. DIS '14. ACM, New York, NY, USA, pp. 543–552.
URL http://doi.acm.org/10.1145/2598510.2598587

Harrison, C., Lim, B. Y., Shick, A., Hudson, S. E., 2009. Where to locate wearable displays?: Reaction time performance of visual alerts from tip to toe. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09. ACM, New York, NY, USA, pp. 941–944.
URL http://doi.acm.org/10.1145/1518701.1518845

Harrison, C., Ramamurthy, S., Hudson, S. E., 2012. On-body interaction: Armed and dangerous. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction. TEI '12. ACM, New York, NY, USA, pp. 69–76.
URL http://doi.acm.org/10.1145/2148131.2148148

Harrison, C., Tan, D., Morris, D., 2010. Skinput: Appropriating the body as an input surface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10. ACM, New York, NY, USA, pp. 453–462.
URL http://doi.acm.org/10.1145/1753326.1753394

Heller, F., Ivanov, S., Wacharamanotham, C., Borchers, J., 2014. Fabritouch: Exploring flexible touch input on textiles. In: Proceedings of the 2014 ACM International Symposium on Wearable Computers. ISWC '14. ACM, New York, NY, USA, pp. 59–62.
URL http://doi.acm.org/10.1145/2634317.2634345

Hoffman, M., Varcholik, P., LaViola, J., 2010. Breaking the status quo: improving 3D gesture recognition with spatially convenient input devices. In: Proceedings of the 2010 IEEE Virtual Reality Conference. VR '10. IEEE Computer Society, Washington, DC, USA, pp. 59–66.

Huber, B., Lee, J. H., Park, J.-H., 2015. Detecting user intention at public displays from foot positions. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15. ACM, New York, NY, USA, pp. 3899–3902.
URL http://doi.acm.org/10.1145/2702123.2702148

Ilmonen, T., Reunanen, M., 2005. Virtual pockets in virtual reality. In: Proceedings of the 11th Eurographics Conference on Virtual Environments. EGVE'05. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 163–170.
URL http://dx.doi.org/10.2312/EGVE/IPT_EGVE2005/163-170

Ishii, H., 2008. Tangible bits: Beyond pixels. In: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction. TEI '08. ACM, New York, NY, USA, pp. xv–xxv.
URL http://doi.acm.org/10.1145/1347390.1347392

Jadidian, J., Katabi, D., 2014. Magnetic MIMO: How to charge your phone in your pocket. In: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking. MobiCom '14. ACM, New York, NY, USA, pp. 495–506.
URL http://doi.acm.org/10.1145/2639108.2639130

Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., Greenberg, S., 2010. A comparison of ray pointing techniques for very large displays. In: Proceedings of Graphics Interface 2010. GI '10. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, pp. 269–276.
URL http://dl.acm.org/citation.cfm?id=1839214.1839261

Jurmu, M., Ogawa, M., Boring, S., Riekki, J., Tokuda, H., 2013. Waving to a touch interface: Descriptive field study of a multipurpose multimodal public display. In: Proceedings of the 2Nd ACM International Symposium on Pervasive Displays. PerDis '13. ACM, New York, NY, USA, pp. 7–12.
URL http://doi.acm.org/10.1145/2491568.2491571

Karrer, T., Wittenhagen, M., Lichtschlag, L., Heller, F., Borchers, J., 2011. Pinstripe: Eyes-free continuous input on interactive clothing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11. ACM, New York, NY, USA, pp. 1313–1322.
URL http://doi.acm.org/10.1145/1978942.1979137

Kendon, A., 1994. Do gestures communicate? A review. Research on Language and Social Interaction 27, 175–200.

Kendon, A., 2004. Gesture: visible action as utterance. Cambridge University Press.

Kim, G. J., Han, S. H., Yang, H., Cho, C., 2004. Body-based interfaces. Applied Ergonomics 35 (3), 263–274.
URL http://dx.doi.org/10.1016/j.apergo.2004.02.003

Korozi, M., Leonidis, A., Margetis, G., Koutlemanis, P., Zabulis, X., Antona, M., Stephanidis, C., 2012. Ambient educational mini-games. In: Proceedings of the International Working Conference on Advanced Visual Interfaces. AVI '12. ACM, New York, NY, USA, pp. 802–803.
URL http://doi.acm.org/10.1145/2254556.2254722

Kuramoto, I., Ishibashi, T., Yamamoto, K., Tsujino, Y., 2013. Stand up, heroes!: Gamification for standing people on crowded public transportation. In: Proceedings of the Second International Conference on Design, User Experience, and Usability: Health, Learning, Playing, Cultural, and Cross-cultural User Experience - Volume Part II. DUXU'13. Springer-Verlag, Berlin, Heidelberg, pp. 538–547.
URL http://dx.doi.org/10.1007/978-3-642-39241-2_59

Lakoff, G., Johnson, M., 1980. Metaphors we live by. The University of Chicago Press.

Lausberg, H., 2013. Neuropsychology of gesture production. In: Body - Language - Communication. An International Handbook on Multimodality in Human Interaction. Vol. 1. Walter de Gruyter GmbH, Berlin/Boston, pp. 168–182.

Lee, S. C., Li, B., Starner, T., 2011. AirTouch: Synchronizing in-air hand gesture and on-body tactile feedback to augment mobile gesture interaction. In: Proceedings of the 2011 15th Annual International Symposium on Wearable Computers. ISWC '11. IEEE Computer Society, Washington, DC, USA, pp. 3–10.
URL http://dx.doi.org/10.1109/ISWC.2011.27

Leiva, L. A., Martn-Albo, D., Vatavu, R.-D., 2017. Synthesizing stroke gestures across user populations: A case for users with visual impairments. In: Proceedings of the 35th ACM Conference on Human Factors in Computing Systems. CHI '17.
URL http://dx.doi.org/10.1145/3025453.3025906

Liu, M., Nancel, M., Vogel, D., 2015. Gunslinger: Subtle arms-down mid-air interaction. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology. UIST '15. ACM, New York, NY, USA, pp. 63–71.
URL http://doi.acm.org/10.1145/2807442.2807489

Lopes, J. a. L., Souza, R. S., Geyer, C. R., Costa, C. A., Barbosa, J. V., Gusmão, M. Z., Yamin, A. C., 2012. A model for context awareness in ubicomp. In: Proceedings of the 18th Brazilian Symposium on Multimedia and the Web. WebMedia '12. ACM, New York, NY, USA, pp. 161–168.
URL http://doi.acm.org/10.1145/2382636.2382672

Lou, Y., Wu, W., Vatavu, R.-D., Tsai, W.-T., 2016. Personalized gesture interactions for cyber-physical smart-home environments. Science China Information Sciences 60 (7).
URL http://dx.doi.org/10.1007/s11432-015-1014-7

MacKenzie, I. S., Kauppinen, T., Silfverberg, M., 2001. Accuracy measures for evaluating computer pointing devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '01. ACM, New York, NY, USA, pp. 9–16.
URL http://doi.acm.org/10.1145/365024.365028

Malik, S., Ranjan, A., Balakrishnan, R., 2005. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In: Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology. UIST '05. ACM, New York, NY, USA, pp. 43–52.
URL http://doi.acm.org/10.1145/1095034.1095042

Markova, M. S., 2013. How does the tangible object affect motor skill learning? In: Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction. TEI '14. ACM, New York, NY, USA, pp. 305–308.
URL http://doi.acm.org/10.1145/2540930.2558150

Markova, M. S., Wilson, S., Stumpf, S., Jan. 2012. Tangible user interfaces for learning. International Journal of Technology Enhanced Learning 4 (3/4), 139–155.
URL http://dx.doi.org/10.1504/IJTEL.2012.051578

Marquardt, N., Ballendat, T., Boring, S., Greenberg, S., Hinckley, K., 2012. Gradual engagement:

Facilitating information exchange between digital devices as a function of proximity. In: Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces. ITS '12. ACM, New York, NY, USA, pp. 31–40.
URL http://doi.acm.org/10.1145/2396636.2396642

Marquardt, N., Diaz-Marino, R., Boring, S., Greenberg, S., 2011. The proximity toolkit: Prototyping proxemic interactions in ubiquitous computing ecologies. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11. ACM, New York, NY, USA, pp. 315–326.
URL http://doi.acm.org/10.1145/2047196.2047238

McKay, B., McKay, K., 2015. A man's pockets. http://www.artofmanliness.com/2015/05/20/a-mans-pockets/.

McNeill, D., 1992. Hand and Mind: What Gestures Reveal about Thought. The University of Chicago Press, Chicago and London.

Meschtscherjakov, A., Reitberger, W., Mirlacher, T., Huber, H., Tscheligi, M., 2009. Amiquin - an ambient mannequin for the shopping environment. In: Proceedings of the European Conference on Ambient Intelligence. AmI '09. Springer-Verlag, Berlin, Heidelberg, pp. 206–214.
URL http://dx.doi.org/10.1007/978-3-642-05408-2_25

Michelis, D., Müller, J., 2011. The audience funnel: Observations of gesture based interaction with multiple large displays in a city center. International Journal of Human-Computer Interaction 27 (6).
URL http://dx.doi.org/10.1080/10447318.2011.555299

Microsoft, 2013. Kinect for Windows: Human Interface Guidelines v1.8. http://go.microsoft.com/fwlink/?LinkID=247735.

Montero, C. S., Alexander, J., Marshall, M. T., Subramanian, S., 2010. Would you do that?: Understanding social acceptance of gestural interfaces. In: Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services. MobileHCI '10. ACM, New York, NY, USA, pp. 275–278.
URL http://doi.acm.org/10.1145/1851600.1851647

Moyes, P., 1997. Just Pockets: Sewing Techniques and Design Ideas. Taunton Press.

Müller, J., Alt, F., Michelis, D., Schmidt, A., 2010. Requirements and design space for interactive public displays. In: Proceedings of the International Conference on Multimedia. MM '10. ACM, New York, NY, USA, pp. 1285–1294.
URL http://doi.acm.org/10.1145/1873951.1874203

Müller, J., Bailly, G., Bossuyt, T., Hillgren, N., 2014. MirrorTouch: Combining touch and mid-air gestures for public displays. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices and Services. MobileHCI '14. ACM, New York, NY, USA, pp. 319–328.
URL http://doi.acm.org/10.1145/2628363.2628379

Nakamura, T., Igarashi, T., 2008. An application-independent system for visualizing user operation history. In: Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology. UIST '08. ACM, New York, NY, USA, pp. 23–32.
URL http://doi.acm.org/10.1145/1449715.1449721

Nancel, M., Pietriga, E., Chapuis, O., Beaudouin-Lafon, M., Aug. 2015. Mid-air pointing on ultra-walls. ACM Transactions on Computer-Human Interaction 22 (5), 21:1–21:62.
URL http://doi.acm.org/10.1145/2766448

Oppl, S., Stary, C., 2011. Towards informed metaphor selection for TUIs. In: Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems. EICS '11. ACM, New York, NY, USA, pp. 247–252.
URL http://doi.acm.org/10.1145/1996461.1996530

Orth, M., Post, R., Cooper, E., 1998. Fabric computing interfaces. In: CHI 98 Conference Summary on Human Factors in Computing Systems. CHI '98. ACM, New York, NY, USA, pp. 331–332.
URL http://doi.acm.org/10.1145/286498.286800

Patel, S. N., Pierce, J. S., Abowd, G. D., 2004. A gesture-based authentication scheme for untrusted public terminals. In: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology. UIST '04. ACM, New York, NY, USA, pp. 157–160.
URL http://doi.acm.org/10.1145/1029632.1029658

Perrault, S. T., Lecolinet, E., Eagan, J., Guiard, Y., 2013. Watchit: Simple gestures and eyes-free interaction for wristwatches and bracelets. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13. ACM, New York, NY, USA, pp. 1451–1460.
URL http://doi.acm.org/10.1145/2470654.2466192

Pfeiffer, M., Schneegass, S., Alt, F., Rohs, M., 2014. Let me grab this: A comparison of EMS and vibration for haptic feedback in free-hand interaction. In: Proceedings of the 5th Augmented Human International Conference. AH '14. ACM, New York, NY, USA, pp. 48:1–48:8.
URL http://doi.acm.org/10.1145/2582051.2582099

Plamondon, R., O'reilly, C., Galbally, J., Almaksour, A., Anquetil, ı., Jan. 2014. Recent developments in the study of rapid human movements with the kinematic theory: Applications to handwriting and signature synthesis. Pattern Recogn. Lett. 35, 225–235.
URL http://dx.doi.org/10.1016/j.patrec.2012.06.004

Profita, H. P., Clawson, J., Gilliland, S., Zeagler, C., Starner, T., Budd, J., Do, E. Y.-L., 2013. Don't mind me touching my wrist: A case study of interacting with on-body technology in public. In: Proceedings of the 2013 International Symposium on Wearable Computers. ISWC '13. ACM, New York, NY, USA, pp. 89–96.
URL http://doi.acm.org/10.1145/2493988.2494331

Randell, C., Muller, H., Jan. 2000. The shopping jacket: Wearable computing for the consumer. Personal Ubiquitous Comput. 4 (4), 241–244.
URL http://dx.doi.org/10.1007/s007790070012

Rico, J., Brewster, S., 2009. Gestures all around us: User differences in social acceptability perceptions of gesture based interfaces. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '09. ACM, New York, NY, USA, pp. 64:1–64:2.
URL http://doi.acm.org/10.1145/1613858.1613936

Rico, J., Brewster, S., 2010. Usable gestures for mobile interfaces: Evaluating social acceptability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10. ACM, New York, NY, USA, pp. 887–896.
URL http://doi.acm.org/10.1145/1753326.1753458

Roalter, L., Kranz, M., Möller, A., Diewald, S., Stockinger, T., Koelle, M., Lindemann, P., 2013. Visual authentication: A secure single step authentication for user authorization. In: Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia. MUM '13. ACM, New York, NY, USA, pp. 30:1–30:4.
URL http://doi.acm.org/10.1145/2541831.2541863

Ruiz, J., Li, Y., 2011. DoubleFlip: A motion gesture delimiter for mobile interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11. ACM, New York, NY, USA, pp. 2717–2720.
URL http://doi.acm.org/10.1145/1978942.1979341

Salvador, R., Romão, T., 2011. Let's move and save some energy. In: Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology. ACE '11. ACM, New York, NY, USA, pp. 86:1–86:2.
URL http://doi.acm.org/10.1145/2071423.2071527

Sanford, J. P., Tietz, A., Farooq, S., Guyer, S., Shapiro, R. B., 2014. Metaphors we teach by. In: Proceedings of the 45th ACM Technical Symposium on Computer Science Education. SIGCSE '14. ACM, New York, NY, USA, pp. 585–590.
URL http://doi.acm.org/10.1145/2538862.2538945

Saponas, T. S., Harrison, C., Benko, H., 2011. PocketTouch: Through-fabric capacitive touch input. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11. ACM, New York, NY, USA, pp. 303–308.
URL http://doi.acm.org/10.1145/2047196.2047235

Schneegass, S., Hassib, M., Zhou, B., Cheng, J., Seoane, F., Amft, O., Lukowicz, P., Schmidt, A., 2015. SimpleSkin: Towards multipurpose smart garments. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers. UbiComp/ISWC'15 Adjunct. ACM, New York, NY, USA, pp. 241–244.
URL http://doi.acm.org/10.1145/2800835.2800935

Schönauer, C., Mossel, A., Zaii, I.-A., Vatavu, R.-D., 2015. Touch, movement and vibration: User perception of vibrotactile feedback for touch and mid-air gestures. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (Eds.), Human-Computer Interaction INTERACT 2015. Vol. 9299 of Lecture Notes in Computer Science. Springer International Publishing, pp. 165–172.
URL http://dx.doi.org/10.1007/978-3-319-22723-8_14

Shaer, O., Hornecker, E., Jan. 2010. Tangible user interfaces: Past, present, and future directions. Foundations and Trends in Human-Computer Interaction 3 (1–2), 1–137.

URL http://dx.doi.org/10.1561/1100000026
Shimozuru, K., Terada, T., Tsukamoto, M., 2015. A life log system that recognizes the objects in a pocket. In: Proceedings of the 6th Augmented Human International Conference. AH '15. ACM, New York, NY, USA, pp. 81–88.
URL http://doi.acm.org/10.1145/2735711.2735788
Shoemaker, G., Tsukitani, T., Kitamura, Y., Booth, K. S., 2010. Body-centric interaction techniques for very large wall displays. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. NordiCHI '10. ACM, New York, NY, USA, pp. 463–472.
URL http://doi.acm.org/10.1145/1868914.1868967
Strachan, S., Murray-Smith, R., O'Modhrain, S., 2007. BodySpace: Inferring body pose for natural control of a music player. In: CHI '07 Extended Abstracts on Human Factors in Computing Systems. CHI EA '07. ACM, New York, NY, USA, pp. 2001–2006.
URL http://doi.acm.org/10.1145/1240866.1240939
Svanaes, D., Verplank, W., 2000. In search of metaphors for tangible user intefaces. In: Proceedings of DARE 2000 on Designing Augmented Reality Environments. DARE '00. ACM, New York, NY, USA, pp. 121–129.
URL http://doi.acm.org/10.1145/354666.354679
Tanase, C. A., Vatavu, R.-D., Pentiuc, S.-G., Graur, A., 2008. Detecting and tracking multiple users in the proximity of interactive tabletops. Advances in Electrical and Computer Engineering 8 (2), 61–64.
URL http://dx.doi.org/10.4316/AECE.2008.02011
van Doorn, M., van Loenen, E., de Vries, A. P., 2008. Deconstructing ambient intelligence into ambient narratives: The intelligent shop window. In: Proceedings of the 1st International Conference on Ambient Media and Systems. Ambi-Sys '08. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, pp. 8:1–8:8.
URL http://dl.acm.org/citation.cfm?id=1363163.1363171
Vatavu, R.-D., 2011. Reusable gestures for interacting with ambient displays in unfamiliar environments. In: Novais, P., Preuveneers, D., Corchado, J. (Eds.), Ambient Intelligence - Software and Applications. Vol. 92 of Advances in Intelligent and Soft Computing. Springer Berlin Heidelberg, pp. 157–164.
URL http://dx.doi.org/10.1007/978-3-642-19937-0_20
Vatavu, R.-D., April 2012a. Nomadic gestures: A technique for reusing gesture commands for frequent ambient interactions. Journal of Ambient Intelligence and Smart Environments 4 (2), 79–93.
URL http://dl.acm.org/citation.cfm?id=2350758.2350765
Vatavu, R.-D., 2012b. User-defined gestures for free-hand TV control. In: Proceedings of the 10th European Conference on Interactive TV and Video. EuroiTV '12. ACM, New York, NY, USA, pp. 45–48.
URL http://doi.acm.org/10.1145/2325616.2325626
Vatavu, R.-D., 2013. The impact of motion dimensionality and bit cardinality on the design of 3D gesture recognizers. International Journal of Human-Computer Studies 71 (4), 387–409.
URL http://dx.doi.org/10.1016/j.ijhcs.2012.11.005
Vatavu, R.-D., 2015. Audience silhouettes: Peripheral awareness of synchronous audience kinesics for social television. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video. TVX '15. ACM, New York, NY, USA, pp. 13–22.
URL http://doi.acm.org/10.1145/2745197.2745207
Vatavu, R.-D., 2017. Beyond features for recognition: Human-readable measures to understand users' whole-body gesture performance. International Journal of Human-Computer Interaction.
URL http://dx.doi.org/10.1080/10447318.2017.1278897
Vatavu, R.-D., Anthony, L., Wobbrock, J. O., 2012. Gestures as point clouds: A $P recognizer for user interface prototypes. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction. ICMI '12. ACM, New York, NY, USA, pp. 273–280.
URL http://doi.acm.org/10.1145/2388676.23887f32
Vatavu, R.-D., Anthony, L., Wobbrock, J. O., 2013. Relative accuracy measures for stroke gestures. In: Proceedings of the 15th ACM on International Conference on Multimodal Interaction. ICMI '13. ACM, New York, NY, USA, pp. 279–286.
URL http://doi.acm.org/10.1145/2522848.2522875
Vatavu, R.-D., Anthony, L., Wobbrock, J. O., 2014. Gesture heatmaps: Understanding gesture performance with colorful visualizations. In: Proceedings of the 16th International Conference on Multimodal Interaction. ICMI '14. ACM, New York, NY, USA, pp. 172–179.
URL http://doi.acm.org/10.1145/2663204.2663256
Vatavu, R.-D., Grisoni, L., Pentiuc, S.-G., 2009. Gesture recognition based on elastic deformation en-

ergies. In: Sales Dias, M., Gibet, S., Wanderley, M. M., Bastos, R. (Eds.), Gesture-Based Human-Computer Interaction and Simulation. Springer-Verlag, Berlin, Heidelberg, pp. 1–12.
URL http://dx.doi.org/10.1007/978-3-540-92865-2_1

Vatavu, R.-D., Mossel, A., Schönauer, C., 2016. Digital vibrons: Understanding users' perceptions of interacting with invisible, zero-weight matter. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '16. ACM, New York, NY, USA, pp. 217–226.
URL http://doi.acm.org/10.1145/2935334.2935364

Vatavu, R.-D., Pentiuc, S.-G., 2008. Interactive coffee tables: Interfacing tv within an intuitive, fun and shared experience. In: Proceedings of the 6th European Interactive TV Conference. EuroITV '08. Springer Berlin Heidelberg, pp. 183–187.
URL http://dx.doi.org/10.1007/978-3-540-69478-6_24

Vatavu, R.-D., Pentiuc, S.-G., Cerlinca, T. I., 2007. Bringing context into play: Supporting game interaction through real-time context acquisition. In: Proceedings of the 2007 Workshop on Multimodal Interfaces in Semantic Interaction. WMISI '07. ACM, New York, NY, USA, pp. 3–8.
URL http://doi.acm.org/10.1145/1330572.1330573

Vatavu, R.-D., Vogel, D., Casiez, G., Grisoni, L., 2011. Estimating the perceived difficulty of pen gestures. In: Proceedings of the 13th IFIP TC13 Conference on Human-Computer Interaction. INTERACT '11. Springer Berlin Heidelberg, pp. 89–106.
URL http://dx.doi.org/10.1007/978-3-642-23771-3_9

Vatavu, R.-D., Zaiţi, I.-A., 2014. Leap gestures for TV: Insights from an elicitation study. In: Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video. TVX '14. ACM, New York, NY, USA, pp. 131–138.
URL http://doi.acm.org/10.1145/2602299.2602316

Vogel, D., Balakrishnan, R., 2004. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. In: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology. UIST '04. ACM, New York, NY, USA, pp. 137–146.
URL http://doi.acm.org/10.1145/1029632.1029656

Vogel, D., Balakrishnan, R., 2005. Distant freehand pointing and clicking on very large, high resolution displays. In: Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology. UIST '05. ACM, New York, NY, USA, pp. 33–42.
URL http://doi.acm.org/10.1145/1095034.1095041

Walter, R., Bailly, G., Müller, J., 2013. StrikeAPose: Revealing mid-air gestures on public displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13. ACM, New York, NY, USA, pp. 841–850.
URL http://doi.acm.org/10.1145/2470654.2470774

Walter, R., Bailly, G., Valkanova, N., Müller, J., 2014. Cuenesics: Using mid-air gestures to select items on interactive public displays. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices and Services. MobileHCI '14. ACM, New York, NY, USA, pp. 299–308.
URL http://doi.acm.org/10.1145/2628363.2628368

Webb, A., 2003. Statistical pattern recognition, 2nd. Ed. John Wiley & Sons Ltd., West Sussex, England.

Weiser, M., Brown, J. S., 1995. Designing calm technology.
URL http://www.ubiq.com/weiser/calmtech/calmtech.htm

Wiese, J., Saponas, T. S., Brush, A. B., 2013. Phoneprioception: Enabling mobile phones to infer where they are kept. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13. ACM, New York, NY, USA, pp. 2157–2166.
URL http://doi.acm.org/10.1145/2470654.2481296

Wilson, A., Shafer, S., 2003. XWand: UI for intelligent spaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '03. ACM, New York, NY, USA, pp. 545–552.
URL http://doi.acm.org/10.1145/642611.642706

Wilson, A. D., Sarin, R., 2007. BlueTable: Connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In: Proceedings of Graphics Interface 2007. GI '07. ACM, New York, NY, USA, pp. 119–125.
URL http://doi.acm.org/10.1145/1268517.1268539

Wobbrock, J. O., Wilson, A. D., Li, Y., 2007. Gestures without libraries, toolkits or training: A $1 recognizer for user interface prototypes. In: Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology. UIST '07. ACM, New York, NY, USA, pp. 159–168.

URL http://doi.acm.org/10.1145/1294211.1294238
Wu, A., Mendenhall, S., Jog, J., Mazalek, A., 2011. Creativity in software development in an academic research lab. In: Proceedings of the 8th ACM Conference on Creativity and Cognition. C&C '11. ACM, New York, NY, USA, pp. 401–402.
URL http://doi.acm.org/10.1145/2069618.2069716
Zhang, E., Zhang, Y., 2009. F-Measure. Springer US, Boston, MA, pp. 1147–1147.
URL http://dx.doi.org/10.1007/978-0-387-39940-9_483
Zhou, Z., Cheok, A. D., Pan, J., September 2004. 3D Story Cube: An interactive tangible user interface for storytelling with 3D graphics and audio. Personal and Ubiquitous Computing 8 (5), 374–376.
URL http://dx.doi.org/10.1007/s00779-004-0300-0