Gesture-Based Interaction in Domotic Environments: State of the Art and HCI Framework Inspired by the Diversity

Ana De Carvalho Correia, Leonardo De Miranda, Heiko Hornung

Gesture-based interaction in domotic environments: State of the art and HCI framework inspired by the diversity

Ana Carla de Carvalho Correia[1], Leonardo Cunha de Miranda[1], Heiko Hornung[2]

[1] Department of Informatics and Applied Mathematics
Federal University of Rio Grande do Norte (UFRN)
Natal, RN, 59078-970, Brazil
[2] Institute of Computing
University of Campinas (UNICAMP)
Campinas, SP, 13083-852, Brazil

anacarla@ppgsc.ufrn.br, leonardo@dimap.ufrn.br, heix@gmx.com

Abstract. Applications for the control and automation of residential environments (domotics) are an emerging area of study within Human-Computer Interaction (HCI). One of the related challenges is to design gestural interaction with these applications. This paper explores socio-technical aspects of gestural interaction in intelligent domotic environments. An analysis of literature in the area revealed that some HCI-related aspects are treated in a restricted manner that neglects socio-technical dimensions. We propose a framework for discussing related challenges in an integrated manner, considering the dimensions people, gestural mode of interaction, and domotics. Some of these challenges are addressed by literature outside the area of domotics. Many open research questions remain, e.g. how to design gestural vocabularies that minimize ambiguity and consider cultural and social aspects. The proposed framework might contribute to answering these questions thus to designing meaningful interaction that is intuitive and easy to learn.

Keywords: Gesture-Based Interaction, Home Automation, Smart Home, Domotics, Socio-Technical Framework.

1 Introduction

Home automation technology has emerged with the aim to facilitate activities in the household or at home, and to provide a more comfortable life for residents. Example tasks include programming the TV set, opening/closing window blinds, or controlling a home entertainment system. However, there are no user interface standards for these devices, often resulting in greater complexity of use. The same technology that simplifies life by enabling a greater number of features in a single device can also complicate our daily lives, making it more difficult to learn and use this technology.

For example, changing the sound volume with some controls is done by up/down buttons, with others by left/right buttons.

This is the paradox of technology, and the challenge of the area of Human-Computer Interaction (HCI) is to minimize these effects via interfaces that are better suited to the growing diversity of users with access to digital technology. Intelligent domotics can offer many benefits to the residents of a house in order to decrease the complexity of using technology for this purpose, providing greater autonomy, comfort and safety. However, the field of residential applications poses some challenges to designers, since this usage context refers to an intimate setting that involves multiple people with different behaviors and different levels of tolerance regarding the effects of technology. Thus, further HCI-related studies are required in order to understand the effects of home automation and promote its benefits, considering socio-technical aspects and the diversity of people regarding culture, gender, social backgrounds, psychological states, physical capabilities, etc.

As a result of making technology available to everyone and integrating multiple device types, challenges arise related to the forms and modes how people interact with these systems. Because of these challenges and advances in hardware and software, the use of gestural interaction has been explored in literature as a viable alternative. Domotics frequently provides solutions for elderly and people with special needs. Without intending to enter the discussion whether these solutions really empower these groups of people [3], we understand that any user, regardless of capabilities or preferences is also a potential user of home automation interfaces, including interfaces with more natural interaction based on gestures.

Literature in domotics frequently studies the development of applications which control lighting, temperature or television in the home, often focusing on technological aspects of gesture recognition. There are few works in HCI dedicated to the theme of human aspects of this form of interaction, e.g. [42]. Thus, studies are required to identify limitations of gestural interfaces for the domotic context. This work presents the state of the art of gestural interaction in domotics. Additionally, we devised a framework for identifying and discussing topics and challenges for research and development of solutions, considering socio-technical aspects.

The paper is organized as follow: Section 2 contextualizes domotics and gestural interaction; Section 3 presents the state of the art of gestural interfaces for residential environments; Section 4 presents a triadic framework of multimodal interaction and identifies challenges of gestural interaction in domotics; Section 5 discusses these challenges on the optic of HCI; Section 6 concludes the paper.

2 Domotics and gestural interaction

The words “home” and “house” are often used interchangeably in literature. In 1985, Dovey [10] discussed the differences of these two concepts. His conception of the subject resonates well with HCI-related perspectives on domotic environments, such as those of Saizmaa and Kim [42]. According to Dovey [10], the “house” is an object, and “home” constitutes an emotional and significant relationship between people and their houses, i.e. the house is the local where the experiences of the home take place. For a more accurate conception of the phenomenon living”, Dovey proposes to examine the house with respect to the concepts “order”, “identity” and “compliance”. The terms home automation, domotics, home computing, smart home and intelligent domotics have been employed in scientific papers in the area. However, there is still no consensus in literature regarding the use of these terms. Saizman and Kim [42] presented five scenarios of smart homes and analyzed common ideas between them. The conclusion was that applications use automated and intelligent computing in the context of the home. However, we believe that this description defines the “domotics” incompletely, since only the isolated meanings of each word are considered, and since it does not distinguish between what is, in fact, automated and what is intelligent.

For Aldrich [1, pp. 17], smart homes can be defined as a residence equipped with computing and information technology which anticipates and responds to the needs of the occupants, working to promote their comfort, security and entertainment through the management of technology within the home and connections to the world beyond. In this view, “smart homes” infer the needs of residents and (semi-)automatically execute them. On the other hand, “automated homes” require explicit commands from users to perform some action. In order to not to restrict ourselves to only either smart or automated homes, we chose the term “intelligent domotics” in our research, since its definition cover both cases. Our own preliminary definition of “intelligent domotics”, which is based on the current state of the art, is: intelligent domotics comprises the use of automated and smart applications in the home, aiming at improving aspects such as security, comfort, and health of residents.

For Cook and Das [9], smart environments are able to acquire and apply “knowledge” about the environment and its inhabitants in order to improve the inhabitants’ experience with the environment. Sadri [41] includes to the description of this environment the concepts “interconnection”, “adaptation”, “dynamism”, “intelligence”, and “integration”. In this view, the traditional means of a system’s input and output disappear. For environments controlled by technology, Sadri stipulates that the way of interaction should be the most intuitive and closest to the daily lives of residents. The use of gestural interaction, for being frequently used in everyday social life, is considered intuitive in human communication. Thus, joining gestures and intelligent domotics seems a topic that should be further explored.

In HCI, gestural interfaces can be studied within the context of Natural User Interfaces (NUIs). Studies in the area of NUIs are concerned with questions such as how the five senses of the human being can serve as a form of interaction with electronic devices. The basic idea is to approximate user experience to everyday contexts and dialogues without the need for complex learning. However, to Norman [33] NUIs are not “natural”. Norman states that the gestural vocabulary of applications with interfaces based on gestures is artificial from the time of its definition. Developers typically define gestures arbitrarily, and over time these settings can become “natural”, i.e. customary, for a group of people, but probably not for a user population with a great cultural diversity. To illustrate that many gestures are not intuitive or natural, Norman and Nielsen [32] cite the example of the zoom multi-touch interface, claiming that the “pinch to zoom” gesture is not natural: when reading a book and “zooming in”, we bow our heads closer to the book without thinking. We concur with Norman’s critique. When addressing the “naturalness” of interaction, i.e. an interaction that is "intuitive" and "easy to learn", it is indispensable to consider social and cultural aspects of a target audience when defining a gestural vocabulary. Since the adjective "natural" is now widespread in HCI literature of the area, we continue to use it in this text, however in quotes and as a synonym to "meaningful", i.e. when we write "natural gestures" we use it in the sense of "meaningful gestures that are intuitive and easy to learn".

3 State of the art of gestural interfaces for intelligent domotics

In this section we present the state of the art of gestural interfaces and synthesize the main findings. In concordance with the taxonomy in [19] we grouped the analyzed literature into two distinct forms of gestural interaction, perceptual and non-perceptual, whereas perceptual input allows the recognition of gestures without the need for any physical contact with an input device or any physical objects.

3.1 Perceptual technologies

While the Kinect has been a success in the games area, Panger [36] investigated the possibilities outside the living room. The author studied the problem of people who want to flip through a recipe book or select a tune to listen while cooking, even with sticky or oily fingers, or hands occupied with kitchen utensils. He proposed an application based on Kinect's depth camera that captures the user's joints for the recognition of movements that consist only in left, right, front and back. Another application that uses the Kinect is the Ambient Wall [21], a smart home system that can display the current status of the house through a projection on a wall, allowing the user to control the TV, check the room temperature, etc. Hands-Up [34] uses the Kinect device with a projector to project images onto the ceiling of the room. This projection location was chosen for usually being the least-used surface inside a house. Additionally, when people get tired of their jobs they often lie down on the bed or sofa and stare at the ceiling of the room. The Hands-Up application interface consists of a circular main menu, in which users can control various devices in the home.

The Kinect has also been used for applications that provide security to the user. Rahman et al. [38] mention a number of functions contained on a car dashboard that are controlled by touch interfaces, which increases the risk and distraction of drivers on the roads. To alleviate this problem, the authors developed and evaluated a purely gestural interface to control secondary functions of a car, that does not use a graphical interface, but audible and haptic feedback. Although this solution has been developed to control the sound system of a car, we deem it relevant for your presentation because it could be adjusted to the residential environment.

The need to always have a remote control on hand to interact with the devices was the main motivation of Solanki and Desai [43] to develop Handmote, an application which recognizes gestural movements to interact with various devices that use a remote control. Their Arduino-based solution recognizes images of the user's hand and processes them in real-time, sending infrared signals to the respective appliance.

Example gestures that are converted to infrared signals for a TV set include signalizing a cross for muting the TV, or turning the hand clockwise or counter-clockwise to change the volume level or TV channel.

Irie et al. [15] discussed a three-dimensional measurement of smart classrooms using a distributed camera system to improve the recognition of three-dimensional movements of the hands and fingers. Their solution allows controlling appliances, TV sets and room lighting through gestures, even when multiple users are present. The Light Widgets [11] application was developed to enable an interaction that is "transparent" and low-cost and that might be accessed on different surfaces, e.g. on walls, floors or tables. Based on the configuration of a surface as the locus of interaction, Light Widgets "reacts" when a user approximates his or her hand to the surface. Users are identified by their skin color. Yamamoto et al. [46] stated that various methods of gesture recognition using the recognition of the user skin color have limitations being sensitive to changes in illumination and certain colors of clothes. Furthermore, using single fixed cameras in narrow spaces, the gesture recognition is restricted to only one person. Thus, the authors use multiple cameras in the corners of the ceiling pointing downwards, to view the entire bodies of users and their faces. A distinguishing feature of this system is its ability to simultaneously process body movements, gestures and face recognition.

For Kim and Kim [20] a major concern is the recognition of gestures as a segment of a few significant gestures from a continuous sequence of movements, i.e., the question of how to detect the start and end points of an intentional gesture. This is a complex process because the gestures have two properties: ambiguity of recognition – due to the difficulty to determine when a gesture starts and ends in a continuous sequence of movements, and segmentation – since multiple instances of the same gesture vary in shape, length and trajectory, even for the same person. To solve these problems Kim and Kim [20] proposed a sequential identification scheme that performs gesture segmentation and recognition simultaneously.

Henze et al. [13] analyzed static and dynamic gestures as forms of interaction with a music application. Static gestures refer to the user's pose or spatial configuration, and dynamic gestures to his or her movement in a certain time interval. These authors performed a 3-step evaluation with twelve users with different profiles, five male and seven female. The results indicated that dynamic gestures are easier to remember, more intuitive and simpler for controlling a music application. Kleindienst et al. [23] discuss the HomeTalk platform that assists users in some domestic services via multimodal interaction. The core of the platform is a residential gateway that acts as a center of family communication. Through direct interaction with a home appliance it is possible to automate different services, and monitor their progress on a PDA. This application provides a greater level of security to residents by providing information about different locations in the home as well as by controlling the food cooking temperature and time, thereby avoiding possible fires.

As an attempt to design more intuitive interfaces in domotic environments, the system developed by Hosoya et al. [14] uses a technique of real-time self-imaging on a translucent in order to improve feedback to the performer of the gestural interaction. The system developed by the authors visualizes the objects in a local or remote room on a screen and superimposes a translucent image of the user. That way, a user can "touch" an object without making real contact – the user's translucent mirror image touches the object on the screen –, and manipulate or interact with objects such as the TV set, or a lamp. Objects in remote rooms need to be tagged with infrared tags.

3.2 Non-perceptual technologies

After having reviewed "perceptual technologies" for gestural interaction, we now present solutions that use non-perceptual technologies, i.e. solutions that enable gestural interaction via gloves, rings, wands or other physical artifacts.

Bonino et al. [5] mentioned that many domotic applications support interaction with devices by fixed touch panels, or by applications on desktop computers, tablets or smartphones. However the use of these technologies has limitations with respect to user interaction, e.g. regarding multi-purpose devices. Furthermore, in the case of mobile devices, there are situations where their use is not possible, e.g. during a shower. In order to circumvent these limitations, the authors chose to use a wristwatch-based solution they call dWatch [5]. Additionally to five other watch functions (time, alarm, temperature, motion, and list of favorite functions), gestural interaction is specifically responsible for controlling household appliances.

Rahman et al. [37] used a residential application to test the trajectory recognition of a user's gestural movements. Their glove-based system enables residential users to interact with the environment. This application consists of infrared cameras, infrared LEDs, and gloves, with the rationale to increase accuracy and to enable gesture recognition in the dark and at relatively low costs. To initiate interaction, the user presses a switch contained in the glove and then "draws" in the air. The system was built to control the lighting of the house, movies and music through movements that resemble some characteristic of the target object's interaction. For example, to control sound, the user "draws" the letter "S", to start media playback the sign ">".

Jing et al. [16] proposed a new physical interaction device called the Magic Ring (MR) which is intended to serve as a means of interaction with different electronics in a residential environment. A comparative evaluating of the use of a traditional remote control and the MR was performed. The results suggested that the use of MR has a smaller learning curve and provides the user with less fatigue than a traditional remote control. For Miranda et al. [28] the remote control in its current form is unsuitable for applications of interactive Digital Television (iDTV). For this reason, the authors proposed the use of Adjustable Interactive Rings (AIRs) to better interact with these applications. With a focus on diverse user capabilities and different contexts of use, the solution consists of three AIRs with distinct functionalities. According to its functionality, each ring has a different color, a single button and a Braille label.

XWand [45] is a multimodal application that enables input of speech and gestures to control various devices in the home. XWand is shaped like a wand that, when pointed at a device, can control it through speech or wand movement. To turn on a lamp, for example, the user has to point the wand at the lamp and say "connect". The system emits an auditory feedback when recognizing a target object. However, the interaction using the Magic Wand itself does not provide feedback to the user.

Carrino et al. [7] also described a multimodal approach based on deictic gestures, symbolic gestures and isolated words with the Wiimote control. The conceptual elements used in this study for three types of entries are: camera, accelerometer and microphone. The camera is attached to the arm or the hand of the user and dedicated to the recognition of deictic gestures using the method Parallel Tracking and Multiple Mapping (PTAMM). The accelerometer is used for recognition of symbolic gestures. The application provides auditory or haptic feedback, or gives feedback through the environment itself, e.g. the feedback of the successful execution of the command “turn on the lights” are turned on lights. The authors conducted a questionnaire-based usability evaluation of a prototype with ten participants using a Likert scale with questions of effectiveness, efficiency, experience and satisfaction. Another application that uses the Wiimote was developed by Neßelrath et al. [29] for gestural interaction with three appliances: kitchen hood, room lighting and TV. The main concern of the authors was to find a small set of commands for the application. They found that one way to decrease the set of commands is by gesture control in context, i.e., one gesture can activate functions of various applications. Another study that also is concerned with the gestural vocabulary is that of Kühnel et al. [24]. The application employs a smartphone and accelerometers as a means of gesture recognition, and can be used for controlling the TV, lamps, window blinds, as well as for interacting with an Electronic Program Guide (EPG).

Jung et al. [17] focused on the support of daily tasks of the elderly and people with special needs, especially those that require the use of wheelchairs. In order to increase comfort in posture and mobility, the authors developed an “intelligent bed”, an “intelligent wheelchair” and a robot for transferring a user between bed and wheelchair. The authors developed interfaces based on hand gestures and voice to control equipment and a health monitoring system that was used not only to assess the state of health, but also as a means to improve comfort by controlling the environment temperature and the transfer of the user between bed and wheelchair.

The work related to health is usually related to monitoring and health care of the residents of the house. The motivation for the creation of the “Gesture Pendant” [12] was the need to reduce the number and complexity of the vocabulary gestures used to interact with the appliances in the home. The Gesture Pendant has a camera surrounded by LEDs, which recognizes gestures also in dark environments. A user can interact through pre-defined control gestures, user-defined gestures or voice commands. Another feature of this application is the monitoring of the user for diagnosis, therapy or emergency services, such as reminding the user to take medication or notify family members.

3.3 A synthesis of the main findings of the literature review

Regarding HCI-related topics, the following were addressed in the works presented in the previous two subsections: accessibility, usability, personalization, privacy, ambiguity of gestures, gestural anthropology (i.e. the relation between gestures and culture), and gender-related issues.

Although many solutions address usability in some form, it can be noted that the "perceptual" solutions seem to put less focus on the other topics. This might be explained by the focus the "perceptual" solutions put on the quality of gesture recognition and related issues, i.e. before addressing accessibility, privacy or personalization, the underlying system of gesture recognition has to work at a satisfying level. On the other hand, when addressing the topics listed in Table 1, in the case of "non-perceptual" solutions this is not necessarily done considering gestural interaction per se but other components of the system. When addressing accessibility, this is often related to components of the system that provide visual feedback or related to specific solutions for people with specific special needs. Accessibility considering gestural interaction would mean to investigate how e.g. people with motor impairments would be able to interact with the system. Usability is arguably the most completely considered area, however, in the case of non-perceptual solutions, usability tests usually focus on graphical interfaces or different modes of feedback. Few works address personalization of gestural interaction, which is a complex problem, e.g. due to the complexity of guaranteeing fast and accurate recognition of personalized gestures. Privacy is often concerned with camera positions or data storage of user images, and not with issues regarding the privacy of performing gestures. Ambiguity in the reviewed solutions is only discussed regarding the relation between already habitually used gestures and gestures used for interacting with the system, however without regarding other contextual factors, such as characteristics of the target audience. Only one solution treated issues related to anthropology, e.g. questions such as cultural aspects that influence the understanding or appropriateness of gestures within a certain context. Regarding gender-related issues, none of the solutions of Table 1 reported that these were a concern during the earlier stages of development, and only three studies reported at least the number of male or female participants [12,17,24], however without stating whether gender had any influence on testing or subsequent design cycles.

Table 1. HCI-related topics addressed by the works.

	Perceptual												Non-Perceptual											
	[36]	Ambient Wall [21]	Hands-Up [34]	[38]	Handmote [43]	[15]	Light Widgets [11]	[46]	[20]	[13]	HomeTalk [23]	[14]	dWatch [5]	[37]	[16]	AIRs [28]	XWand [45]	Magic Wand [35]	[7]	[29]	[24]	SmartKom [44]	[17]	gesture Pendant[12]
Accessibility			X													X	X		X	X		X	X	X
Usability	X	X	X	X		X		X		X	X	X	X	X	X	X	X		X	X	X	X		X
Personalization																				X				
Privacy													X		X		X	X	X				X	X
Ambiguity					X				X				X	X				X	X		X			
Anthropology																					X			
Gender															X						X			X

We believe that some of these limitations are a result of considering the topics of Table 1 in an isolated manner regarding technology, gestural interaction, and individual/social aspects. For current solutions in the area of domotics this might be adequate to some extent, since the problems they address are of a relatively low complexity, e.g. functionalities in the areas of comfort or security such as turning on the lights. However, regarding the area of health, or more complex problems in the areas of comfort and security, or even new areas that are not yet addressed in domotics, we think a more comprehensive and particularly a socio-technical perspective is required, grounded on methods and frameworks of contemporary HCI. As a first step towards this direction, we thus propose in the next section a framework for conceptualizing research and design questions in an integrated manner.

4 Socio-technical aspects of gestural interaction: Framework and challenges

After analyzing the state of the art, we can observe that researchers and developers are more concerned with trying to offer "comfort" to users through some complex computational solutions that can control e.g. the lighting, the room temperature, TVs or home entertainment systems. However, besides of technical aspects of controlling appliances in the home, also aspects from the social sphere need to be considered.

In order to identify and discuss these HCI-related aspects and challenges of multimodal interaction, we propose a socio-technical framework of multimodal interaction in the context of intelligent domotics. The framework consists of the main dimensions technology, modes of interaction, and people. The concentric organization of these three dimensions symbolizes their interdependency in a triadic relationship. The main idea of this framework is to identify and discuss the challenges of different forms of interaction with technology considering socio-technical aspects in an integrated manner, i.e. we acknowledge that a challenge should not be considered isolatedly under a single perspective, but under a perspective that combines multimodal, technological, and social aspects and considers their interdependencies. It should be noted that in the context of this paper, we only consider the mode of "gestural interaction", and only technology in the context of domotics (Fig. 1). Possible users of the framework include researchers and developers, who can catalog socio-technical research challenges and discuss challenges that permeate the implementation of their solutions.

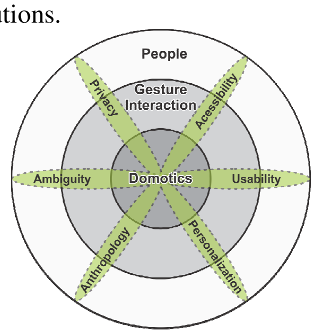

Fig. 1. Framework of socio-technical aspects of gestural interaction.

The aspects discussed in this section do not comprise an exhaustive list, but are the ones that emerged from the literature review presented in the previous section, namely accessibility, usability, personalization, privacy, ambiguity, and anthropology. Each aspect is represented by a dashed ellipse in Fig. 1 and has intersections with the three dimensions domotics, gestural interaction, and people.

A noticeable characteristic of the state of the art presented in the previous section is the high share of concepts or techniques based mainly on solutions that use graphical interfaces or non-perceptual technologies. These solutions, e.g. guidelines or methods, were developed and used for traditional GUI or Web applications and often do not meet requirements for interaction with domotics, i.e. an interaction that should be more "natural" in the sense of being meaningful and intuitive. In the following, we will discuss the challenges regarding the six aspects of gestural interaction in the domain of intelligent domotics presented in Fig. 1. We will analyze which challenges can be addressed by adapting work from related areas (e.g. gestural interaction in other contexts, or general literature in HCI), and which research issues remain.

Accessibility – Accessibility has generally been perceived as a necessary attribute of quality of software and hardware systems. Consequently, we believe it is essential to provide accessibility in residential solutions. Some studies proposed multimodal interfaces for gesture and voice, aiming at the inclusion of a greater diversity of users, e.g. [12,23,38,44,45]. However, none of these studies was concerned with analyzing e.g. the accessibility of the gestural vocabulary with respect to people with mental or physical impairments.

Computing solutions from various contexts have visual interfaces, whether in tablets, phones, computers or through projections. Often, researchers focus on accessibility in these conditions. However, for the context of intelligent domotics, considering technologies that allow gestural interaction and that have no visual interfaces, accessibility is poorly explored. Although there are already consolidated accessibility guidelines aimed at a content that is perceivable, operable, and understandable by a wide range of users, as well as compatible with a wide variety of assistive technologies, not all principles of accessibility are "compatible" with the context of home automation applications. Examples include some guidelines for Web applications or guidelines such as "making all features also available to keyboard use" or "provide alternative text to non-textual content". Changing the way people interact with the environment brings out new aspects of accessibility that require more research. There are several questions about how we can develop gestural applications that are more accessible to the diversity of the audience, and the challenge becomes even more complex by the lack of development methodologies for gestural applications, as well as evaluation methods for this particular type of interaction.

Kane et al. [18] reported a study on accessibility in gestural interfaces applied to touchscreen surfaces. They found that, given a gestural vocabulary, blind participants chose significantly different gestures than sighted participants. Blind participants showed strong preferences for gestures that were in the corners and edges of the screen, as well as for multitouch gestures. Kane et al. also discovered differences regarding the performance of gesturing, e.g. gestures produced by the blind participants were bigger, slower, and showed a greater variation in size than those produced by sighted participants. An important result of the study is that according to some blind participants, they did not know how to perform some of the gestures used in the defined gestural vocabulary, including letters, numbers and other symbols. However, there has been little research about the differences or peculiarities of gestures regarding so perceptual aspects and the residential context.

Another aspect not yet studied in literature is the accessibility depending on the gender of the users, since physical and psychological conditions differ for each gender. The structure of our framework considers the mutual dependencies of the mentioned challenges: accessibility may be interplaying with usability, personalization of gestures, privacy and the gestural ambiguity of the solution.

Usability – Most of the works presented in the previous section that discuss usability are concerned with efficiency and learnability, which are only two of Nielsen's [30] five main topics: learnability, efficiency, memorability, errors and satisfaction. One of the benefits of graphical interfaces is to aid the memorization of commands for interaction, because the information is organized graphically in windows and is represented by text, icons or other visual elements. The language of commands based on a menu structure has the cognitive advantage that commands can be recognized instead of being required to recalled. For solutions that use GUIs Lenman et al. [25] proposes the use of "marking menus". Learning the command set is performed gradually through "pie-menus" that indicate the direction of the movement that the user needs to perform. Another progressive form of learning gestures is through multimodal voice interaction, where the application supports the user by describing the movements required to perform an action. As the user learns to perform gestures, the application no longer provides the respective instructions.

For gestural interfaces, gestural ergonomics is also important, since the interface should not be physically stressing [31]. The comfort when interacting with the system is important, which is not achieved when a user has to "wear" technology, like a glove. With respect to the use of perceptual technology, discomfort and fatigue might arise when the user's main means of interaction are arms and hands interacting without a supporting surface. Nielsen et al. [31] describes ergonomic principles for building a good interface with gestures, e.g. relax muscles, avoid repetition or staying in a static position, avoid internal and external force on the joints.

Although usability guidelines are widely used in traditional solutions, new principles or the adaptation of existing ones are required for gestural interfaces [33], especially those without a GUI. Fundamental principles of interaction design that are independent of technology [33] might provide starting points for this investigation, e.g. visibility (related to affordances), feedback, consistency, non-destructive operations, discovery, scalability, and reliability.

Personalization – Many problems involving customization are related to the huge amount of information that needs to be managed simultaneously. To support versatility of gestural commands for different types of solutions it is essential that applications are customizable, as well as easy and fast to train. Achieving these properties leads to the problem of forming and recognizing gestures freely. In order for the application to learn new commands, users have to train it repeating the same command several times. This repetition might generate discomfort for the user. Liu et al. [27] aim to decrease the number of required repetitions by using discrete Hidden Markov Models (HMMs). However, the authors point out that there are several technical challenges for interaction based on gestures. Unlike many pattern recognition problems such as speech or writing recognition, gesture recognition lacks a commonly accepted standard or "vocabulary". Therefore, it is sometimes desirable and often necessary for users to create their own gestures. With customized gestures, it is difficult to gather a large set of training samples that is required to establish statistical methods. Furthermore, the gesture-based spontaneous interaction requires immediate engagement, i.e. the overhead of creating recognition instruments should be minimal. More importantly, the application platforms for specific custom gesture recognition are generally very limited in terms of cost and system characteristics, such as battery capacity or the buttons presented in [27].

A clear difference between customization of perceptual solutions of gestural interaction and "traditional" means of interaction is the difficulty of recognizing who interacts with the system, e.g. in a home, all residents are possible users of the application. Recognizing the user is indispensable for enabling customization. Thus challenges include how to recognize the interacting person, and how to achieve this unobtrusively, i.e. without e.g. requiring the person to utter his or her name or to look at a particular location for facial recognition. With regard to the integrated consideration of personalization within our framework, it is worth noting, that the technical feature "personalization", applied to the context of gestural interaction with domotics also has a strong social component, i.e. personalization should be consistent with the ideals of a natural, meaningful of interaction that respects the privacy of the users within the home. Furthermore, personalization may have a positive effect on usability and accessibility.

Ambiguity – In order to actually enable a "natural" interaction, the ambiguity of gestures, which is very present in the real world, needs to be reduced for interacting with the virtual world [33]. Gestures need to be cohesive and consistent in their meanings. For instance, if used as a command for interacting with a system, the movements used to express a farewell in the real world should to be used for the same purpose and with the same meaning, i.e. as a gesture of "farewell". Following this principle, the gestural vocabulary would become more intuitive, easing the learning curve of users. Some solutions acknowledge this principle [13,29], but do not take into account whether the gestures are ambiguous regarding the target population, since ambiguity is intrinsically related to the cultural aspects of the population. As an example of the problem of not considering the interdependency of gestural ambiguity with cultural aspects, the application in [39] uses only deictic gestures (pointing gestures). Gestures for some commands for this application resemble a firearm, which is probably not desired in a home or in war- or conflict-ridden regions of the world.

During social interactions, people use a large vocabulary of gestures to communicate in daily life. The gestures used vary according to contextual and cultural aspects [22] and are closely linked to other aspects of communication. A challenge that arises is that gestures for interaction with domotics must be sufficiently "natural", i.e. resemble to some degree the gestures used in everyday life. At the same time, these gestures must be recognized as intentional commands to the system, i.e. they must be distinguished from gestures of inter-person communication. This problem has also been called "immersion syndrome" [2], i.e. in a scenario in which all gestures are captured and can be interpreted by the system, gestures may or may not be intended for interaction with the system, and people can no longer interact simultaneously with the system and other people using gestures.

To clarify the idea of gestural ambiguity, we can draw an analogy to sign language. When the gestures of a domotics solution are established without a previous study, gestural ambiguity might occur on two levels, i.e. different gestures/signs might be required for asking a person or commanding the system to turn on the lights or draw the curtain, or the gesture for the command "draw the curtain" might have a different meaning in sign language. There are no studies on "gestural affordances", i.e. the problem of gesture discoverability. Another question related to gestural affordances is if there exist any universal gestures. A positive answer to this question might reduce ambiguity and cultural dependencies. From these last considerations, it also becomes clear, that ambiguity has a strong relation to anthropology.

Anthropology – Symbolic gestures, the meanings of which are unique within the same culture, are an example of a classification used to discriminate gestures depending on the anthropology of a certain population. An example is the "thumbs-up" gesture which signifies approval in Brazilian culture but can be an insult in some middle-eastern countries. Sign languages also fall into this category and vary significantly between countries.

Researchers are still trying to understand how the gestures are influenced by culture [26]. Due to this still largely unexplored area, applications often make use of deictic gestures [6,45], i.e. pointing gestures which have much less cultural dependency, but which are also limited since not every function in a domotic environment can be executed by "pointing at things". Furthermore, regarding the "naturalness" or "intuitiveness" of interaction in domotic environments, the use of only deictic gestures also imposes a limitation. Kühnel et al. [24] described anthropology as a requirement for computational solutions for domotic environments, not only with respect to gestural interaction, but in a broad context. One of the authors' concerns was the writing direction of the user. Although this detail might seem irrelevant for the definition of gestures, it might very well influence whether a certain gesture is considered appropriate in a certain cultural context.

Privacy – With the rapid advancement of technology, sensors and information storage devices are becoming increasingly integrated in the solutions. These devices provide various benefits such as accuracy of command recognition, and mobility in use. However, the context of the home requires a number of concerns about the privacy of those who utilize these technologies, because it refers to an intimate environment in which the lack of privacy can have negative effects on the social relationships among residents and result in failure of the domotic application.

Considering the difficulties in the development of applications that address the requirements of the involved stakeholders in a home environment, Choe et al. [8] aim to investigate the types of activities of which residents do not want to have stored records. To obtain these results, the authors analyzed the questionnaires of 475 respondents, with 71.6% female and 28.4% male participants. A total of 1533 different activities that respondents did not want to be stored was identified. The male respondents most frequently reported activities related to the category of intimacy and the use of media. Female respondents reported activities related to the category of self-appearance and oral expressions. Moreover, Choe et al. identified places that need more care regarding residents' privacy, e.g. bedrooms and bathrooms.

5 Discussion

Given the diversity of the population, designing purely gestural interfaces for residential environments might not be the most appropriate approach, because the sheer amount of gestures that would have to be memorized and performed would be exhaustive for the population as a whole, and especially for the elderly and people with special needs. As presented in Section 3, many applications choose to use multimodal interaction. Among the works mentioned, many authors use a combination of speech with gestures, in order to provide greater accessibility of the system, to facilitate recognition of user commands, or to decrease the complexity of the vocabulary of gestural applications.

However, as stated earlier, home automation has consequences far beyond the way we interact with the appliances in the home. Besides the technical, several other aspects about building those applications have to be addressed and analyzed under an integrated socio-technical perspective. These aspects include, but might not be limited to, accessibility, usability, ambiguity, privacy, anthropology and gender of users, since this context is closely related to physical, social, psychological, emotional and even spiritual concerns of each resident. Addressing these issues also increases the "naturalness" or "intuitiveness" of smart home automation applications.

Currently, the development of intelligent home automation applications often focuses exclusively on technological aspects, not taking into account what is actually necessary and desirable for users. Although affective, psychosocial and other aspects that cause an impact on residents are being explored more actively in the area of HCI, many open questions need to be studied, especially in the area of domotics.

Bardzell et al. [4] intended to explore issues of feminist thought intertwined with human-computer interaction. Both feminism and HCI have made important contributions to social science in recent decades. However, despite of the potential, there has been no engagement between the two areas until recently. A series of surveys, focused mainly in perceptual and cognitive tasks, revealed gender differences that can have implications for interactive systems design. However, believes still seem to be prevalent that gender does not have much influence on technology usage. Thus, one of the research objectives of Rode [40] is to show the importance of treating gender in HCI and to emphasize that it permeates all aspects of daily life, including domestic life. Many studies ignore important social aspects, in which the issues of gender occur daily. This point of view on users' gender is relevant for home automation applications, as in a home environment all people need to interact with the applications. Thus, applications must meet the needs of both genders and should be designed and developed for this.

Aiming to address more adequately the challenges presented above, we found it necessary to design a framework that specifically addresses the challenges contained in gestural interaction with domotics. Although used exclusively in the context of gestural interaction and domotics, the framework presented in this paper could be used for a similar discussion in other application contexts and regarding other or additional modes of interaction. Saizmaa and Kim [42] presented a framework that addresses some conceptual aspects similar to the framework presented in this paper. Saizmaa and Kim [42] identified important issues that are considered or omitted in the development of intelligent homes, and organized these issues into three dimensions, human, home and technology, highlighting the need to not only to see a house as a physical thing, with walls and ceiling, but also as a "home".

The holistic approach advocated by Saizmaa and Kim [42] elucidates the complexity of smart home automation, as well as draws attention to not limiting an analysis in this domain to technological aspects alone. Although similarities such as the problem domain of domotics exist between our framework and the framework of Saizmaa and Kim, there also exist significant differences. While Saizmaa and Kim discuss different aspects of interaction, they do not elaborate on the peculiarities of different modes of interaction. In this paper, we discussed gestural interaction, but in principle the discussion could be extended to other types or combinations of modes. Saizman and Kim discuss issues in the dyads home-technology, technology-human, and human-home, and assign research topics to each dyad. Maintainability, e.g., is assigned to the dyad home-technology, although we think that also personal, cultural, or social aspects might be relevant for this topic. Hence, our framework uses a triadic structure which enables to discuss topics in all three dimensions of people, domotics, and gestural interaction. Depending on the role of the framework user, it is thus also possible to investigate one or more topics in only one or two dimensions. For instance a sociologist in the research or development team might be interested in the "people dimension" of different aspects, while a developer might be interested in technical aspects of accessibility.

The contribution of this framework is to analyze an explicit context from its initial form of interaction until its possible social implications. When we addressed these challenges from the perspective of the framework, we did not take the traditional way of analyzing only parameters of a graphical interface. This framework enables the discussion on a triadic unfolding of the dimensions that make up the framework. Although only discussed for gestural interaction, it can be applied to other modes of interaction as well, for example, gestural, mobile, brain-machine sound, among other perceptual or non-perceptual interaction. Consequently the analyses of these challenges can provide insights to the particularities of each specific mode of interaction.

We understand that regardless of the utilized technology, one of the main questions is the definition of the gestural vocabulary of the application. This definition should not be driven by technology, but by a human-centered perspective that considers the "naturalness" and "intuitiveness" of gestural interaction. In order to formalize gestural interaction we consider it essential to create a grammar which is initially free of technological aspects and which considers both multimodal commands as well as the points already discussed in this paper.

6 Conclusion

This paper presented the state of the art of residential applications that use gestural interaction to communicate with various types of appliances in the home. An analysis of scientific literature in the area of gestural interaction with domotics revealed that aspects such as accessibility, usability, personalization, privacy, ambiguity, or anthropology of gestural interaction are often not considered, or only considered in a restricted, isolated way. We argued that aspects of gestural interaction with domotics need to be discussed in a more integrated, socio-technical way, and proposed a framework that permitted us to discuss these aspects within the triadic relationship consisting of the dimensions people, domotics, and gestural interaction. We identified open research questions and challenges. Literature outside the domain of domotics provides some pointers to these questions, e.g. literature about HCI-related aspects of gestural interaction in general, or literature from HCI and related areas. However, we conclude that there are no formalized design and evaluation methods in literature about gestural interaction with perceptual technologies, i.e. applications without GUIs and without physical artifacts of interaction. Furthermore, literature in the domain of gestural interaction with domotics that discusses aspects such as anthropology, accessibility, usability, gender, personalization and privacy is scarce.

Based on the results of this work, future work involves the human-centered conceptualization and formalization of multimodal interaction for this complex context of use.

Acknowledgments

This work was partially supported by Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), by the Brazilian National Council of Scientific and Technological Development (CNPq grant #141058/2010-2), and by the Physical Artifacts of Interaction Research Group (PAIRG) at Federal University of Rio Grande do Norte (UFRN), Brazil.

References

1. Aldrich, F.K.: Smart Homes: Past, Present and Future. Inside the Smart Home. pp. 17–39, Springer (2003)
2. Audel, T., Beaudouin-Lafon, M.: Charade: Remote Control of Objects Using Free-Hand Gestures. Communications of the ACM 36(7), 28–35 (1993)
3. Bannon, L.: Reimagining HCI: Toward a more Human-Centered Perspective. Interactions 18(4), 50–57 (2011)
4. Bardzell, S., Churchill, E., Bardzell, J., Forlizzi, J., Grinter, R., Tatar, D.: Feminism and Interaction Design. In: CHI'11, pp. 1–4, ACM (2011)

5. Bonino, D., Corno, F., Russis, L.: dWatch: A Personal Wrist Watch for Smart Environments, In: Procedia Computer Science 10, 300-307 (2012)
6. Cabral, M.C., Morimoto, C.H., Zuffo, M.K.: On the Usability of Gesture Interfaces in Virtual Reality Environments. In: CLIHC'05, pp. 100–108 (2005)
7. Carrino, S., Péclat, A., Mugellini, E., Khaled, O.A., Ingold, R.: Humans and Smart Environments: A Novel Multimodal Interaction Approach. In: ICMI'11, pp. 105–112, Alicante, Spain (2011)
8. Choe, E.K., Consolvo, S., Jung, J., Harrison, B., Kientz, J. A.: Living in a Glass House: A Survey of Private Moments in the Home. In: UbiComp'11, pp. 41–44 (2011)
9. Cook, D.J., Das. S.K.: How Smart are Our Environments? An Updated Look at the State of the Art. Pervasive Mob. Comput. 3(2), 53–73 (2007)
10. Dovey, K.: Home and Homeless. Home Environments. 36–61, Plenum Press (1985)
11. Fails, J.A., Olsen, D.: Light Widgets: Interacting in Every-Day Spaces. In: IUI'02, pp. 63–69, ACM (2002)
12. Gandy, M., Starner, T., Auxier, J., Ashbrook, D.: The Gesture Pendant: A Self-Illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. In: ISWC'00, 1–8, IEEE (2000)
13. Henze, N., Löcken, A., Boll, S., Hesselmann, T., Pielot, M.: Free-Hand Gestures for Music Playback: Deriving Gestures with a User-Centred Process. In: MUM'10, pp. 1–10, ACM (2010)
14. Hosoya, E., Kitabata, M., Sato, H., Harada, I., Nojima, H., Morisawa, F., Mutoh, S., Onozawa, A.: A Mirror Metaphor Interaction System: Touching Remote Real Objects in an Augmented Reality Environment. In: ISMAR'03, pp. 1–2, Tokyo, Japan (2003)
15. Irie, K., Wakamura, N., Umeda, K.: Construction of an Intelligent Room Based on Gesture Recognition: Operation of Electric Appliances with Hand Gestures. In: IROS'04, pp. 193–198, Sendai, Japan (2004)
16. Jing, L., Yamagishi, K., Wang, J., Zhou, Y., Huang, T., Cheng, Z.: A Unified Method for Multiple Home Appliances Control through Static Finger Gestures. In: SAINT'11, pp. 82–90, IEEE (2011)
17. Jung, J., Do, J., Kim, Y., Suh, K., Kim, D., Bien, Z.Z.: Advanced Robotic Residence for the Elderly/the Handicapped: Realization and User Evaluation. In: ICORR'05, pp. 492–495, Chicago, Illinois (2005)
18. Kane, S.K., Wobbrock, J.O., Ladner, R.E.: Usable Gestures for Blind People: Understanding Preference and Performance. In: CHI'11, pp. 413–422 (2011)
19. Karam, M., Schraefel, MC.: A Taxonomy of Gestures in Human Computer Interactions. Technical Report ECSTR-IAM05-009, University of Southampton (2005)
20. Kim, D., Kim, D.: An Intelligent Smart Home Control Using Body Gestures. In: ICHIT'06, pp. 439–446, IEEE (2006)
21. Kim, H., Jeong, K., Kim, S., Han, T.: Ambient Wall: Smart Wall Display Interface which can be Controlled by Simple Gesture for Smart Home. In: SA'11, pp. 1–2, Hong Kong, China (2011)
22. Kita, S.: Cross-Cultural Variation of Speech-Accompanying Gesture: A review. Language and Cognitive Processes 24(2), 145–167 (2009)
23. Kleindienst, J., Macek, T., Serédi, L., Šedivý, J.: Vision-Enhanced Multi-Modal Interactions in Domotic Environments. In: ERCIM'04 (2004)
24. Kühnel, C., Westermann, T., Hemmert, F., Kratz, S., Müller, A., Möller, S.: I'm Home: Defining and Evaluating a Gesture Set for Smart-Home Control. Int. J. Hum.-Comput. Stud. 69, 693–704 (2011)

25. Lenman, S., Bretzner, L., Thuresson, B.: Using Marking Menus to Develop Command Sets for Computer Vision Based Hand Gesture Interfaces. In: NordiCHI'02, pp. 239–242 (2002)
26. Lim, C.J., Pan, Y., Lee, J.: Human Factors and Design Issues in Multimodal (Speech/Gesture) Interface. In: JDCTA'08, pp. 67–77 (2008)
27. Liu, J., Zhong, L., Wickramasuriya, J., Vasudevan, V.: uWave: Accelerometer-Based Personalized Gesture Recognition and its Applications. Pervasive Mob. Comput. 5(6), 657–675 (2009)
28. Miranda, L.C., Hornung, H.H., Baranauskas, M.C.C.: Adjustable Interactive Rings for iDTV. IEEE Trans. Consum. Electron. 56(3), 1988–1996 (2010)
29. Neßelrath, R., Lu, C., Schulz, C.H. Frey, J., Alexandersson, J.: A Gesture Based System for Context-Sensitive Interaction with Smart Homes. Ambient Assisted Living, 1–12, (2011)
30. Nielsen, J.: Usability Engineering. Morgan Kauffmann, San Diego, California (1993)
31. Nielsen, M., Moeslund, T., Störring, M., Granum, E.: A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for Human Computer Interaction. In: Internation Gesture Workshop, Genova, Italy (2003)
32. Norman, D.A., Nielsen, J.: Gestural Interfaces: A Step Backward in Usability. Interactions 17(5), 46–49 (2010)
33. Norman, D.A.: Natural User Interfaces are Not Natural. Interactions 17(3), 6–10 (2010)
34. Oh, O., Jung, Y., Cho, Y., Hahm, C., Sin, H., Lee, J.: Hands-up: Motion Recognition Using Kinect and a Ceiling to Improve the Convenience of Human Life. In: CHI EA'12, pp. 1655–1660, Texas, USA (2012)
35. Ouchi, K., Esaka, N., Tamura, Y., Hirahara, M., Doi, M.: Magic Wand: An Intuitive Gesture Remote Control for Home Appliances. In: AMT'05, IEEE (2005)
36. Panger, G.: Kinect in the Kitchen: Testing Depth Camera Interactions in Practical Home Environments. In: CHI'12, pp. 1985–1990, ACM (2012)
37. Rahman, A.M., Hossain, M.A., Parra, J., Saddik, A.E.: Motion-Path Based Gesture Interaction with Smart Home Services. In: MM'09, pp. 761-764, ACM (2009)
38. Rahman, A.S.M.M., Saboune, J., Saddik, A.E.: Motion-Path Based in Car Gesture Control of the Multimedia Devices. In: DIVANet'11, pp. 69–76, ACM (2011)
39. Richarz, J., Plötz, T., Fink, G.A.: Real-Time Detection and Interpretation of 3D Deictic Gestures for Interaction with an Intelligent Environment. In: Int. Conf. on Pattern Recognition, pp. 1–4 (2008)
40. Rode, J.A.: A Theoretical Agenda for Feminist HCI. Interacting with Computer 23(5), 393–400 (2011)
41. Sadri, F.: Ambient Intelligence: A Survey. ACM Computing Surveys 43(4), 1–66 (2011)
42. Saizmaa, T., Kim, H.C.C.: A Holistic Understanding of HCI Perspectives on Smart Home. In: NCM'08, pp. 59–65, IEEE (2008)
43. Solanki, U.V., Desai, N.H.: Hand Gesture Based Remote Control for Home Appliances: Handmote. In: WICT'11, pp. 419–423, IEEE (2011)
44. Wahlster, W., Reithinger, N., Blocher, A.: SmartKom: Multimodal Communication with a Life-Like Character. In: Eurospeech'01, pp. 1547–1550, ISCA (2001)
45. Wilson, A., Shafer, S.: XWand: UI for Intelligent Spaces. In: CHI'03, pp. 545–552, ACM (2003)
46. Yamamoto, Y., Yoda, I., Sakaue, K.: Arm-Pointing Gesture Interface Using Surrounded Stereo Cameras System. In: ICPR'04, pp. 965–970, IEEE (2004)