INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

MulTIS:
A Gesture Based Interaction Model for iDTV

Leonardo Cunha de Miranda
Heiko Horst Hornung
Maria Cecília Calani Baranauskas

Technical Report - IC-09-004 - Relatório Técnico

February - 2009 - Fevereiro

MulTIS: A Gesture Based Interaction Model for iDTV

Leonardo Cunha de Miranda, Heiko Horst Hornung, M. Cecília C. Baranauskas

Department of Information Systems
Institute of Computing – P.O. Box 6.176
University of Campinas – UNICAMP
13083-970, Campinas, SP, Brazil

professor@leonardocunha.com.br, heix@gmx.com, cecilia@ic.unicamp.br

Abstract

The existence of digital artefacts commonly used to interact with today's television systems does not guarantee that those devices are adequate to mediate the use of Interactive Digital Television (iDTV) and its new types of applications. This is especially true in the context of developing countries where iDTV is a promise to cope with the digital divide. This technical report reviews the state of the art of physical artefacts of interaction with iDTV. In addition it presents a gesture based interaction model for iDTV that could inform the design of a new interaction language based on gesture and physical artefacts of interaction with iDTV.

1 Introduction

Gawlinski [14], defines Interactive Television as a set of technological artefacts that allow establishing a dialogue between the user – or TV spectator – and a TV channel, program or service. Therefore, the appeal of interactive applications offered for a particular user profile as well as the adequacy of technology employed in the users' TV equipment can make this dialogue more or less effective.

The digitalization of terrestrial television broadcasting in Brazil and, consequently, the larger offering of interactivity on television establish a new paradigm of interaction with the TV spectators, which potentially has a great social impact for the Brazilian population.

According to last National Survey of Sample Households (PNAD in its Portuguese acronym) published by the Brazilian Institute of Geography and Statistics (IBGE in its Portuguese acronym), 93% of Brazilian households have television. This number is an indicator of the high importance – in different contexts, such as cultural, social, educational, strategic etc. – of TV to the Brazilian population. Television is the main source of information and entertainment for the Brazilian population [12,32,43].

To contextualize, some statistical figures illustrate the diversity of skills and competencies as well as the social reality of stakeholders of iDTV in Brazil. According to the latest population census of the IBGE, 14.5% of the population (≈ 24,6 million people) have at least one kind of impairment, and according to the latest National Indicator of Functional Literacy (INAF in its Portuguese acronym) published by the Paulo Montenegro Institute (IPM in its Portuguese acronym), 72% of Brazilians are not considered fully literate. Still, according to the latest Social Radar published by the Institute of Applied Economic Research (IPEA in its Portuguese acronym), 30.1% of the population live in poverty. We believe these are the user groups potentially more vulnerable to problems of interaction with the iDTV, and at the same time, they could be those benefited greatly from this technology.

The public TV in Brazil is a true instrument of national integration [43]. Even with so many cultural, social, and economical differences, the country is united through the information services and entertainment provided by TV. In the same way, Fortes [12] emphasizes that the deployment of digital television in Brazil is a highly strategic issue, because it involves questions of great impact for the Brazilians' lives. The Brazilian Computer Society (SBC in its Portuguese acronym) also recognizes the social relevance of the definition of a Brazilian System of Digital Television (SBTVD in its Portuguese acronym) and believes that the contributions and the participation of the Brazilian Academy are essential for technological independence of Brazil.

In Brazil, Interactive Digital Television (iDTV) is an emerging research area. At the end of 2003, with the creation of SBTVD by the Federal Government, research in this field in Brazil effectively was initiated.

Furthermore, the definition of an own standard for digital television – the Brazilian System of Digital Terrestrial Television (SBTVD-T in its Portuguese acronym) – seems to be advantageous over the adoption of existing digital television standards – e.g., the Japanese: Integrated Services Digital Broadcasting Terrestrial (ISDB-T), the European: Digital Video Broadcasting-Terrestrial (DVB-T), the North American: Advanced Television Systems Committee (ATSC), or the Chinese: Digital Multimedia Broadcast-Terrestrial/Handheld (DMB-T/H) – because it allows more flexibility with respect to the incorporation of technologies developed in Brazil, that address the peculiarities of television use of the Brazilian population.

At the end of 2007, the first transmissions of public Digital TV in Brazil were made in São Paulo city. However, so far, no interactive applications are available in public TV and only few in Pay-TV. In addition, there are few Brazilians who have, in fact, access to this technology. Regarding Digital TV we observe a focus on questions concerning the size of the screen and high definition image and audio, using Plasma Screens or Liquid Crystal Display (LCD) with 16:9 format (widescreen) and images in High Definition Television (HDTV), as well as sound systems with six speaker boxes (5.1 surround). Another phenomenon is the Brazilian market for set-top boxes that are sold although the necessary infrastructure for executing interactive applications does not even exist yet.

Thus, there are several reasons that motivate us in this research. First of all, we stress that the technologies created for developed countries meet the needs of users that are immersed in social contexts with an already established digital culture, where many people are already "digitally literate". This situation is very different from that present in Brazil. This moment of initial definitions of SBTVD-T being elaborated provides a good opportunity for discussion and creation of new options for the design of physical artefacts of interaction with iDTV.

The remote control, the main physical artefact of interaction with the television system, is not enough for the highly dynamic interaction with the iDTV, given the problems already identified and discussed by several authors [3,4,12,16,25,27,28,41]. In addition, users not familiar with the use of technology in their daily life can experience problems of interaction with the television system. Therefore, we emphasize the need to consider the development and/or adaptation of new physical artefacts of interaction with iDTV in Brazil, in order to not end up in an under-utilization of this rich new interactive media. We believe that a solution in our context could be applied to other contexts, given the huge diversity skills and competences of our population.

Despite the importance of physical artefacts for interaction of users with the television system, scientific publications that exploit the topic of interaction design in iDTV from the viewpoint of physical artefacts of interaction are still scarce, especially in Brazil. Papers being published in the context of iDTV have focused mainly on issues of interfaces – Graphical User Interfaces (GUI) – of interactive applications (e.g., accessibility [33], usability [40], navigability [3] and universal access [6]). This technical report discusses other hardware as an alternative to the remote control, that could be used as electronic mediators of communication between users and the television system. We argue that problems of accessibility, usability, navigability introduced into interactive applications may have the origin in unsolved questions of the interaction design based on the use of the remote control.

We still know very little about the Brazilian users of iDTV regarding their prospective interaction. Applying the classification of different levels of interaction proposed by Lemos [20] and extended by Montez and Becker [26], the great majority of Brazilian television users accesses only the terrestrial analog TV, and does not exceed the third level of interaction (the use of remote control to switch on/off the television, to adjust the sound volume, to change the channel etc.). Those who have the financial resources to access Pay-TV via cable or satellite, achieve a higher level of interaction (up to level 4 of interaction, already possible via today's remote controls, e.g., choosing the angle of vision or making a purchase online).

The level 7 of interaction will put a greater cognitive load onto TV spectators who potentially are not accustomed to interacting with digital interfaces (according to Montez and Becker, full interactivity is achieved at this level, when the distinction between TV spectators and broadcasters will be blurred due to the fact that the spectators will be able to actually create content).

Although Brazil is adopting the Japanese ISDB-T Standard as a technological basis for SBTVD-T, this does not mean that this standard meets all our needs regarding the possibilities of user interaction with the TV, as we face huge differences in the socio-economic, cultural, and regional dimensions, as well as in general access to technology and knowledge [2].

Within this scenario, we think it is essential to analyze and propose artefacts to facilitate interaction with iDTV, and thus maximize the use of this new media as an important tool for disseminating information and knowledge to a Brazilian citizen. Moreover, this research is in synergy with one of the great challenges of research in computing in Brazil for the decade [35], the SBC's Challenge nº 4: Participatory and Universal Access of Brazilian Citizens to Knowledge [2].

This technical report is organized as follows: Section 2 presents the state of the art regarding physical artefacts of interaction with iDTV and considerations about these artefacts on the perspective of Human-Computer Interaction (HCI) and our context. Section 3 presents the gesture based interaction model we are proposing. Section 4 shows our final considerations and indicates its continuity.

2 Physical Artefacts for Interaction with IDTV

Currently, TV is still a media presentation for collective use of audio-visual content. The provision of interactive applications will change the way the user understands and uses the television, mainly because with these applications he/she is no longer a TV spectator, but becomes an active subject in this new relationship that must be established with the television system. In both cases the interaction between the users and the television system occurs through an electronic mediator, usually the remote control.

Figure 1 presents an abstraction of this model of interaction between the users and the television system. The user or users group (a) interacts with the television system through sight, hearing and cognition, and usually with the aid of sight, directing their intentions of interaction – with the applications of iDTV, using the remote control as a physical artefact of interaction (b). The receiver of iDTV (c) responds to commands from the user through changes in the audio-visual content or through changes in the interface of the interactive applications (software).

The model shown in Figure 1 differs from the traditional model adopted from the current analog television systems in four aspects: 1) Digital transmission of audio-visual content; 2) The existence of interactive applications (software) – composed of text, graphics etc. – making the dialogue between the users and television system more dynamic and complex, exposing the users to a higher cognitive load; 3) Variety of physical artefacts needed for interaction, because of the new possibilities of the interactive applications; and 4) Increased diversity of technologies for transmitting data used for electronic communication between the physical artefact of interaction and the receiver, such as infrared (IrDA), bluetooth (IEEE 802.15), Wi-Fi (IEEE 802.11) etc.

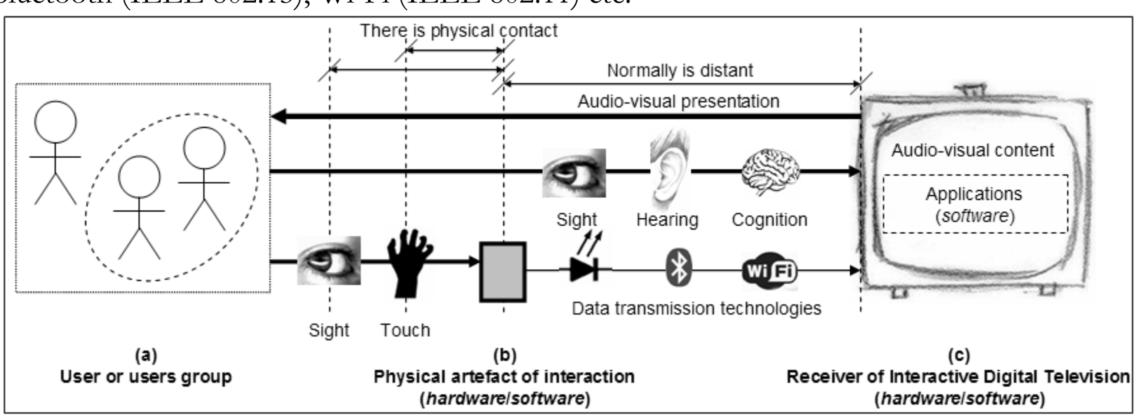

Figure 1: Interaction between the users and the television system through physical artefact of interaction

Different authors indicate that the remote control is the main physical artefact of interaction used for "data input" on television [3,32,34], which is why it is necessary to analyze its design and usage. Thus, given the specifics of this artefact, studies have sought to better understand the interaction design between the users and iDTV mediated by that device. The undeniable fact is that the invention of the electronic mediator promoted benefits/facilities to people, while complexity of use increased at the same time due to new devices with even more features.

It is a trend of the contemporary world that remote controls are used for all types of electronic devices for home consumption, such as television, videocassette recorder, DVD player, stereo system, ceiling fan, air conditioning etc. In a way, this leads to a problem, because in the same environment (e.g., living room, auditorium etc.), we usually find several consumer electronics, each with its own remote control.

For Nielsen [27,28] this problem is aggravated due to the existence of different designs for each remote control, e.g., buttons for switching on/off television on different positions. Barros [3] also noted that a task quite clearly and objectively defined, like how to turn on and turn off particular equipment, may not be as easy to be performed by the user when there exists more than one remote control in that environment. Furthermore, to Waisman [41], the lack of standardization of controls confuses the user with the different interfaces, different roles, different positions of the buttons, requiring a huge mental effort to remember the sequence of actions to perform certain function. For Freeman et al. apud Gomes et al. [16], the difficulty of using the remote control is related to the excessive number of buttons, problems of nomenclature of buttons, inconsistent terminology, symbols confused, among others.

According Berglund [4], the remote control as we know it is not appropriate for most services enabled by digital TV and for Smith apud Eronen [10] there may be a need to develop new technologies for hardware and software that have not yet been developed. On the same line of understanding, Piccolo and Baranauskas [32] said new agreements for the interaction design should be built so that the full potential of the iDTV is used. For Fortes [12], browsing the internet by iDTV still presents some challenges, and there still exist difficulties to provide the interaction via the remote control (for a better interaction the use of input devices such as mouse and keyboard is needed).

Several studies suggest alternatives to the remote control. In order to provide alternatives, research has adopted different methods like User Centered Design [30], Participatory Design [36] etc. seeking a better understanding of the design of interfaces for remote controls. Omojokun et al. [31], for example, examine how to integrate features of various equipments into a single device with the participation of users in the process of developing prototypes.

Considering the remote control a limited device for the interaction with applications of interactive iDTV, Lin and Chen [21] suggested an alternative solution, describing an architecture that allows the use of mobile devices that have a web browser – such as a Personal Digital Assistant (PDA), laptop, PocketPC or mobile phone – to control the set-top box of iDTV. Also for Sohn and Lee [37] iDTV requires more advanced methods of control. Thus, they proposed an ultrasonic pointing device – the SonarPen – that allows controlling the iDTV through an ultrasonic "pen". The electronic communication between these devices is established by means of infrared. According to these authors this approach offers greater precision than other similar studies.

As for the previous authors, Fujita et al. [13] believe that the features and greater complexity of operation with iDTV demand new alternatives to be developed. Thus, the authors incorporate a microphone to the remote control, enabling the user to control the television with speech recognition. One of the benefits cited by the authors in this new interface is the decrease in the number of necessary buttons on the remote control, despite the complexity of the systems for speech recognition.

Roibás et al. [34] consider the quality of access a major problem in the context of iDTV. In this study, they presented a summary and analysis of alternatives to the remote control. According to these authors a wireless keyboard facilitates the data entry by typing. However, they cite some problems of accessibility, like audible feedback. They suggest some options that allow more interactivity, e.g., adding pointing devices to remote controls, like mini-joysticks or mini-trackballs. Additionally, they comment on the possibility of carrying out commands to control iDTV through microphones and speech recognition. The work of Wittenburg et al. [42] and Ibrahim and Johansson [18] also follows the same type of proposal based on speech recognition.

Gomes et al. [16] propose an interface based on the reading of bar codes printed on paper to the specific context of T-Learning applications. Thus, they employ a bar code reader as physical artefact of interaction with iDTV. However, they indicate as the next step adding a bar code reader to the remote control. Finally, we like to point out the work of Enns and MacKenzie [9] who propose a touchpad based remote control (touch sensitive device). According to these authors, this solution enables the exploration of new methods of interaction, as for example recognition of characters "written" on the touchpad of the remote control [15]. Currently, this technique is used directly on the screens of some handhelds, smartphones, and Tablet PCs.

Table 1 summarizes a list of proposed technologies discussed in literature that can be used as alternatives to the use of remote control with iDTV.

Table 1: Alternative technologies and artefacts of interaction with iDTV

Reference	Technologies and Artefacts
[32]	Remote control and wireless keyboard
[13,42]	Remote control with speech recognition
[18]	Speech recognition
[34]	Wireless keyboard, remote control with pointing devices, such as mini-joystick or mini-trackball, and microphone with speech recognition
[9]	Touchpad based remote control
[21]	PDA, laptop, pocketPC and mobile phone
[37]	Ultrasonic "pen"
[16]	Bar code reader

2.1 Discussion

In this subsection, we discuss the physical artefacts of interaction proposed in the literature as alternatives to the use of the remote control from an HCI perspective, and based on previous results in the context of this research [23,24].

- **The "traditional" remote control:** People with special needs may have restrictions to make use of this artefact, and users with visual impairments may have difficulties to identify certain buttons, such as the colored buttons (soft keys). In addition, remote controls with many buttons add complexity to the language of interaction.
- **Remote control with speech recognition:** This artefact gives greater flexibility of interaction than the other solutions, but the use of voice commands in a collective scenario of using television in Brazil may not work properly because of the background noise of natural relaxed atmosphere where people assist TV. Furthermore, those systems need to be "trained" before they recognize the voice of the user and the user needs to keep the remote control near himself/herself.
- **Touchpad based remote control:** This artefact does not provide alternatives for interaction with the television system to users with visual and motor disabilities, and users with low literacy and no familiarity with the everyday use of computers may have problems of interaction with the artefact. Moreover, the writing with/on this artefact diverts the focus of user interaction, off the television screen to the writing area of the touchpad.
- **Wireless keyboard:** This wireless keyboard artefact does not provide alternatives for interaction with the television system for users with special needs.
- **Mouse:** This artefact does not provide alternatives for interaction with the television system to users with special needs, and users unfamiliar with the everyday use of digital interfaces may have difficulty identifying the mapping of movement of the mouse on a surface with the movement of the cursor in screen, as well as the different functions of the mouse buttons. Furthermore, the use of the television environment is different from the office environment, since most of the time the user is not sitting in front of a desk to use a television, which would be needed to support and move the mouse.
- **Mobile phone, handheld (PDA) and smartphone (PocketPC):** These artefacts do not provide alternatives for interaction with the television system to users with special needs, and users with visual impairments may find it difficult to use these devices since they are usually lacking auditory feedback.
- **Ultrasonic "pen":** This artefact does not provide alternatives for interaction with the television system to users with special needs. Users with visual impairment are likely to have problems to specify the exact location they wish to manipulate.
- **Bar code reader:** The proposed employment of this artefact in the context of iDTV requires its use in conjunction with the material where the bar codes are printed. This form of use may not be feasible in some types of interactive applications such as entertainment. Furthermore, users with visual impairment may encounter problems in identifying the exact location that the reader should apply.
- **Laptop:** Users with low literacy and no familiarity with the everyday use of computers may have problems of interaction with the artefact. Another factor is the cost of acquiring such equipment.

These solutions do not seem to be the most appropriate for addressing the diversity of users and other stakeholders that we seek to benefit from this research, as they were not designed based on the principles of Design for All, also known as Universal Design [5,7,19].

The formulation of solutions of hardware and software should be based on the characteristics of the target users. One should take into account the diversity of users in their differences, without discriminating them. The aim is therefore to create solutions that are used by all, to the maximum extent possible. The universal access, effective use and a more fluent dialogue in this new media will directly depend on the physical artefact of iDTV interaction.

Based on the understanding that the interaction with iDTV should be more natural, simple and intuitive, the physical artefact of interaction has to be more transparent to the user so that the focus of the interaction is not distracted from the audio-visual content or the interface of the interactive application. We are investigating the use of "gestures" to create a new interaction language with iDTV mediated by a physical artefact of interaction more "transparent" to the users.

3 MulTIS Model

Based on the previous discussion, a number of issues related to the use of gestures as an interface between the users and iDTV are raised: What kind of mapping of the user's gestures to consider (in three or two dimensions)? What gestures could represent the different functions of interaction with iDTV? Can sign language be used for interaction with iDTV?

Now, let's start thinking about some forms of interaction between people and some technological artefacts of the contemporary world as a way to illustrate some analogies to demonstrate the feasibility of bringing an interaction more natural, simple and intuitive for the prospective users of the iDTV. The first example is a situation of driving a car. In this case, the focus of the user should be on the cars around. Note that in this case the user does not need to use the sight to interact with the physical artefacts of interaction, which in this context is the march, steering wheel, pedals and so on. Thus, in this case, the physical artefacts of interaction are transparent to the user allowing him/her to focus on the traffic. For the second example, think of a situation where the focus of a user is on using the computer. Note that even using the mouse as a physical artefact of interaction, the focus of this user is directed towards the application on the screen and not towards the mouse. The last example is the context of using the Nintendo's Wii Video Game System [29] – that carries more transparent metaphors, directing the focus of the interaction in the screen game. These three examples show how the design of physical artefacts of interaction may allow a more intuitive interaction with, less training, and more focus in the task and not on the artefact. Then, one possible language interaction could bring the metaphors of real life through the everyday gestures to compose the language of interaction. The remote control does not bring this idea; everytime you need to look at the it diverts the focus of attention.

In the Brazilian-Portuguese Dictionary the word "gestures" refers to the movement of the body, especially the head and the arms, to express ideas or feelings, or to highlight the words. According to Kurtenback and Hulteen apud Mendes [22], humans make use of gestures to complete the sense of a spoken message, and bear alone the full message. Languages are codes, and we communicate through them. Thus, the communication codes are multiple and many, e.g., blind language, deaf language, and acronyms languages.

For Waisman [41], the remote control is seen as an extension of the human body, as it can touch the screen where the finger did not reach. In the context of this work, the proposal to define a new interaction language with iDTV based on gestures aims at enabling the interaction between the users and interactive applications in a more natural way. For this reason, it is necessary to make the physical artefact of interaction more transparent to the user, allowing to decrease the amount of used metaphors, and the cognitive load on the user.

Therefore, it is necessary to define an interaction language with iDTV by gestures that make more sense to the TV spectators. Aiming at considering the diversity of users, the language must be aligned with the new guidelines for physical artefacts of interaction previously presented in [23]. This language has to be designed to flexibly allow the users' interaction with the iDTV at all seven levels of interactivity proposed by Lemos [20] and extended by Montez and Becker [26].

After the analyses of artefacts for physical interaction with iDTV [23] the adoption of an interaction language based on gestures using the "reading" of movements of arms/hands/fingers has been identified as a possible approach. We could therefore adopt the concept of a multi-touch imaginary screen, where there is no real physical touch on the screen. This would be a 2D imaginary multi-touch screen, mapped via the X and Y coordinates, supported by computer vision for tracking the movement, digital image processing or capturing and tracking of equipment with infrared.

We propose to formalize a gesture based interaction model: Multi-Touch without physical touch on an Imaginary Screen computationally mapped by two dimensions. The definition of that model is based on issues related to skills and competencies of the target users, the usage environment, and the complexity of implementation (e.g., questions concerning the required hardware and software platform for motion recognition of arms/hands/fingers). One of the goals of this Model is to make the language accessible for all. The characteristics of this Model can be summarized as:

- **Multi-Touch:** We propose to use the concept of interaction based on multiple touches, i.e., the user visually indicates the area he/she wants to manipulate. This mapping allows the exploration of new forms of interaction with the latest digital interfaces;
- **without physical touch:** In our model there is no need to actually touch the TV screen which is consistent with the usage scenario, i.e., unlike with other artefacts with (multi-)touch screens (e.g., bank ATM, desktop computer, notebook, Tablet PC, mobile devices in general), the user generally keeps some distance from the TV set, which impedes a direct physical contact; and
- **on an Imaginary Screen computationally mapped by two dimensions:** We live in a three dimensional (3D) space and gestures made by people to express themselves are 3D, too. However, the specification of movements in three dimensions for representing certain actions with relation to the interaction with iDTV as well as the reproduction by the user is difficult. Moreover, the capturing and computational processing of moves in 3D space would be complex and would probably require a change in the environment, that in iDTV context can be quite varied, e.g., living room, dining room, bedroom, bathroom, kitchen, garden, garage etc., to have two cameras to capture the movements of users considering the three dimensions. Since the TV screens is 2D.

It is worth noting that this model is not aimed at supporting multimodal interaction, which combined two or more modes of communication (e.g., remote control combined with keyboard and/or mouse). The general idea of this Model is that the gestures made by users in 3D space will be mapped by two dimensions (2D), considering the movement of one or more points in two dimensions of a Multi-Touch Imaginary Screen (MulTIS). Figure 2 shows the components considered by MulTIS Model in an example with two points.

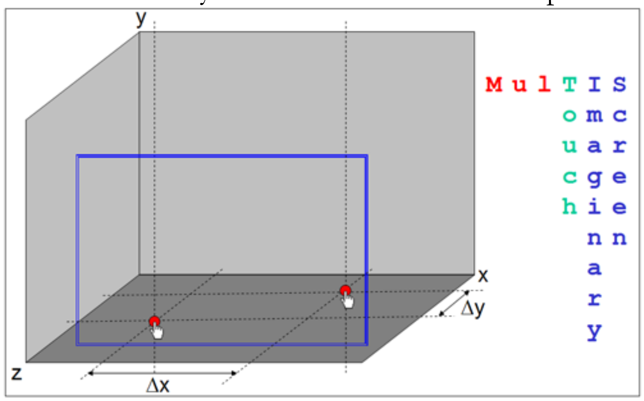

Figure 2: Two points mapping example of MulTIS Model

It is worth mentioning that the language will not be based on the Brazilian sign language or any other sign language because these languages are relatively complex and difficult to learn and use.

We propose to create an interaction language with iDTV based on gestures in order to formulate pre-defined movements between one or more virtual points in a specified period of time (different activities to be performed are thus represented by changes in location between those virtual points which are mapped to 2D in a specified period of time). Therefore, the recognition of the points in motion will be carried out using the physical artefacts of interaction. It is worth mentioning that due to the nature of the visual language of interaction and based on the principles of Design for All, extensions of the language of interaction will be considered, e.g., if a user wants to make use of “shortcuts”. Thus, it seems that the greatest benefits of these resources (shortcuts), without disregarding other users, would be people with visual and motor disabilities, and users with cerebral palsy. These shortcuts allow a direct interaction independent of the mapping of the imaginary screen and the television screen (real).

Thus, trying to make the dialogue as simple as possible between users and television system, movements that represent the different possible actions to be performed by the user to interact is being defined on the background of the following principles:

- Employ natural and spontaneous movements of the arms/hands/fingers;
- Use metaphors of people’s daily lives seeking to identify the most significant movements, such as the metaphors used in TV, movies, theatre etc.;
- Independence of any specific type of interactive application;
- Provide feedback of interaction actions done;
- Take into consideration the different usage environments of television in Brazil;
- Take into consideration the diversity of users without creating movements that exclude any user, e.g., do not define complex three dimensional movements;

- Consider previous experience of users with the use of technology;
- Consider the Brazilian culture;
- Allow this language to be extended to other contexts; and
- Keep the consistency of language, e.g., the “open” action is the opposite of “close”.

Regarding Computer Science literature that explores the topic of gesticulation interfaces, the majority of published work is located exclusively in the computational sphere, as the field on the tracking of “objects” – e.g., arms, hands and fingers – in real time still lacks effective results. Therefore, most of the work focuses on issues concerning the areas of computer vision for tracking the movements [11,38] and digital image processing [17]. An example of work in the field is presented by Dias and Leite [8], who propose a new method for the tracking of objects using fuzzy logic. Mendes [22] proposed a solution of a gesticulation hardware/software interface that made use of artificial neural networks and measurement of positions in three dimensions. Truyenque [39] proposed to use computer vision to capture gestures and create devices for interacting with computers faster and more intuitively. This work presents a study on the feasibility of using webcams as a device for interaction based on gestures of the hands, mapping certain actions to mouse and keyboard events, for example to go forward and back in a PowerPoint presentation.

Interaction with the iDTV by gestures is underexplored. As far as we know languages of interaction by gestures that takes into consideration the diversity of users in the context of iDTV is still an open problem for which we aim at seeking solutions.

4 Conclusion and Future Works

The majority of potential users of the Brazilian iDTV are not familiar with the everyday use of digital interfaces. Experiments carried out by our research team with representatives of target users in other application contexts have shown the difficulty of the majority of users in understanding digital interfaces. Ways towards an effective use and a more fluent dialogue in this new media will depend directly on the physical artefacts of interaction with iDTV.

The continuity of this research involves the application of the conceptual model proposed in this technical report as one of the pillars for the development of a new physical artefact as a mediator of the interaction with iDTV based on gestures and its evaluation with the users. The tests will be conducted on-site in the context of research projects that are already in progress.

Following are some inherent challenges regarding the formalization of the interaction language based on the MulTIS Model. Which movements with arm/hands/fingers are more significant to the audience regarding the basic interactions with iDTV as specified by the Brazilian norm ABNT NBR 15604:2007 [1]: 1) Turn On/Off television; 2) Adjust the sound volume; 3) switch from one channel to another and access it directly; and 4) access and use menu of options for the configuration of television (e.g., to adjust the brightness, contrast, sharpness of the image, among other possibilities). Other issues that arise in this context: How to handle the entry of text? What movement would be the analogue to the “Enter” key of the keyboard and/or the actions of clicking the mouse on the left? What movement represents the action of keeping the left mouse button pressed? These are some examples of questions to be answered. However, we must consider other actions to be carried out by the

MulTIS Model, e.g., selection, navigation and execution. Finally, we highlight the need to balance the physical artefact of interaction resources and the interaction language by gesture, making the interaction with iDTV more natural and simple mediated by a physical artefact of interaction more transparent to the users.

We believe that new forms of communication and interaction can drastically affect the way the user will interact with the television system. Thus, we expect this model could lead to a design that effectively promotes the iDTV as a tool for digital inclusion.

Acknowledgments

This research is partly funded by the National Council for Scientific and Technological Development (CNPq in its Portuguese acronym) through the Ph.D. scholarship from the first author of this technical report (141489/2008-1). The authors also thank the Microsoft Research – FAPESP Institute for IT Research (2007/54564-1), the Foundation of Development of University of Campinas – FUNCAMP in its Portuguese acronym (32/96) and the Foundation for the Coordination of Improvement of Higher Education Personnel – CAPES (DS-00014/07-9).

References

1. ABNT. NBR 15604:2007 - Televisão Digital Terrestre - Receptores, Versão C, 2008.
2. Baranauskas, M.C.C. e Souza, C.S. Desafio nº 4: Acesso Participativo e Universal do Cidadão Brasileiro ao Conhecimento. *Computação Brasil*, ano VII, nº 23, 7.
3. Barros, G.G. A Consistência da Interface com o Usuário para a TV Interativa, Set 2006. Dissertação, Universidade de São Paulo.
4. Berglund, A. Augmenting the Remote Control: Studies in Complex Information Navigation for Digital TV, Apr 2004. Ph.D. Thesis, Linköping University.
5. Bühler, C. Empowered Participation of Users with Disabilities in Universal Design. In *Universal Access in the Information Society*, 1(2):85-90.
6. Chorianopoulos, K. Interactive TV Design That Blends Seamlessly with Everyday Life. In *Proc. ERCIM UI4ALL Ws 2006*, Springer (2007), 43-57. DOI=http://dx.doi.org/10.1007/978-3-540-71025-7_4.
7. Connell, B.R., Jones, M., Mace, R., Mueller, J., Mullick, A., Ostroff, E., Sanford, J., Steinfeld, E., Story, M. and Vanderheiden, G. The Principles of Universal Design. http://www.design.ncsu.edu/cud/about_ud/udprinciplestext.htm.
8. Dias, F.A.S. and Leite, N.J. Fuzzy Data for Video Image Object Tracking, Technical Report IC-05-024, Institute of Computing, University of Campinas, 2005.
9. Enns, N.R. and MacKenzie, I.S. Touchpad-Based Remote Control Devices. In *Proc. CHI '98*, ACM (1998), 229-230. DOI=http://doi.acm.org/10.1145/286498.286717.
10. Eronen, L. User Centered Design of New and Novel Products: Case Digital Television, Nov 2004. Ph.D. Thesis, Helsinki University of Technology.
11. Forsyth, D.A. and Ponce, J. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
12. Fortes, R.S. Sistema Cossack: Uma Plataforma para Gerenciamento de Aplicações Hipermídia Distribuídas, Ago 2004. Dissertação, Instituto Tecnológico da Aeronáutica.

13. Fujita, K., Kuwano, H., Tsuzuki, T., Ono, Y. and Ishihara, T. A New Digital TV Interface Employing Speech Recognition. In *IEEE Transactions on Consumer Electronics*, 49(3): 765-769.
14. Gawlinski, M. *Interactive Television Production*. Focal Press, 2003.
15. Goldberg, D. and Richardson, C. Touch-Typing with a Stylus. In *Proc. INTERCHI '93*, ACM (1993), 80-87.
16. Gomes, F.J.L., Lima, J.V. e Nevado, R.A. Uma Interface Multimodal para Objetos de Aprendizagem Visualizados na TV Digital. In *Anais IHC 2008*, SBC (2008), 284-287. DOI=http://doi.acm.org/10.1145/1497470.1497508.
17. Gonzalez, R.C. and Woods, R.E. *Digital Image Processing*. Prentice Hall, 2008.
18. Ibrahim, A. and Johansson, P. Multimodal Dialogue Systems for Interactive TV Applications. In *Proc. ICMI 2002*, IEEE (2002), 117-122.
19. Iwarsson, S. and Stahl, A. Accessibility, Usability and Universal Design - Positioning and Definition of Concepts Describing Person-Environment Relationships. In *Disability and Rehabilitation*, 25(2):57-66.
20. Lemos, A.L.M. Anjos Interativos e Retribalização do Mundo: Sobre Interatividade e Interfaces Digitais. http://www.facom.ufba.br/ciberpesquisa/lemos/interativo.pdf.
21. Lin, C. and Chen, M. On Controlling Digital TV Set-Top-Box by Mobile Devices via IP Network. In *Proc. ISM 2005*, IEEE (2005), 1-8.
22. Mendes, G. Interface Gestual Ultrassônica, Dez 1999. Dissertação, Universidade Estadual de Campinas.
23. Miranda, L.C., Piccolo, L.S.G. e Baranauskas, M.C.C. Artefatos Físicos de Interação com a TVDI: Desafios e Diretrizes para o Cenário Brasileiro. In *Anais IHC 2008*, SBC (2008), 60-69. DOI=http://doi.acm.org/10.1145/1497470.1497478.
24. Miranda, L.C., Piccolo, L.S.G. and Baranauskas, M.C.C. Uma Proposta de Taxonomia e Recomendação de Utilização de Artefatos Físicos de Interação com a TVDI. In *Proc. Workshop on Perspectives, Challenges and Opportunities for Human-Computer Interaction in Latin America (INTERACT 2007)*.
25. Monteiro, M.S.M. TV Interativa e Seus Caminhos, Dez 2002. Dissertação, Universidade Estadual de Campinas.
26. Montez, C. e Becker, V. *TV Digital Interativa: Conceitos, Desafios e Perspectivas para o Brasil*. Editora da Universidade Federal de Santa Catarina, 2005.
27. Nielsen, J. Remote Control Anarchy. http://www.useit.com/alertbox/20040607.html.
28. Nielsen, J. Why Consumer Products Have Inferior User Experience. http://www.useit.com/alertbox/20040315.html.
29. Nintendo. Nintendo's Wii Video Game System. http://www.nintendo.com/wii/.
30. Norman, D.A. and Draper, S.W. *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associate Publishers, 1986.
31. Omojokun, O., Pierce, J.S., Isbell, C.L. and Dewan, P. Comparing End-User and Intelligent Remote Control Interface Generation. In *Personal and Ubiquitous Computing*, 10(2-3): 136-143. DOI=http://dx.doi.org/10.1007/s00779-005-0019-6.
32. Piccolo, L.S.G. e Baranauskas, M.C.C. Desafios de Design para a TV Digital Interativa. In *Anais IHC 2006*, ACM (2006). DOI=http://doi.acm.org/10.1145/1298023.1298025.

33. Piccolo, L.S.G., Melo, A.M. and Baranauskas, M.C.C. Accessibility and Interactive TV: Design Recommendations for the Brazilian Scenario. In *Proc. INTERACT 2007*, Springer (2007), 361-374. DOI=http://dx.doi.org/10.1007/978-3-540-74796-3_34.
34. Roibás, A.C., Sala, R., Ahmad, S.N.A.S. and Radman, M. Beyond the Remote Control: Going the Extra Mile to Enhance iTV Access via Mobile Devices & Humanizing Navigation Experience for Those with Special Needs. In *Proc. EuroITV 2005*, 133-141.
35. SBC. Grand Challenges in Computer Science Research in Brazil: 2006-2016. http://sistemas.sbc.org.br/ArquivosComunicacao/Desafios_ingles.pdf.
36. Schuler, D. and Namioka, A. *Participatory Design: Principles and Practices*. Lawrence Erlbaum Associate Publishers, 1993.
37. Sohn, M. and Lee, G. SonarPen: An Ultrasonic Pointing Device for an Interactive TV. In *IEEE Transactions on Consumer Electronics*, 50(2):413-419.
38. Trucco and Verri, A. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
39. Truyenque, M.A.Q. Uma Aplicação de Visão Computacional que Utiliza Gestos da Mão para Interagir com o Computador, Mar 2005. Dissertação, Pontifícia Universidade Católica do Rio de Janeiro.
40. Valdestilhas, A. e Almeida, F.A. A Usabilidade no Desenvolvimento de Aplicações para TV Interativa. In *Proc. Workshop de TV Digital e Interativa (SIBGRAPI 2005)*.
41. Waisman, T. Usabilidade em Serviços Educacionais em Ambiente de TV Digital, 2006. Tese, Universidade de São Paulo.
42. Wittenburg, K., Lanning, T., Schwenke, D., Shubin, H. and Vetro, A. The Prospects for Unrestricted Speech Input for TV Content search. In *Proc. AVI 2006*, ACM (2006), 352-359. DOI=http://doi.acm.org/10.1145/1133265.1133338.
43. Zuffo, M.K. TV Digital Aberta no Brasil - Políticas Estruturais para um Modelo Nacional. http://www.lsi.usp.br/~mkzuffo/repositorio/politicaspublicas/tvdigital/TVDigital.pdf.