Project FoSIBLE
Fostering Social Interactions for a Better Life of the Elderly

Deliverable

D5.3: **Report on new approaches of user-centred design research methodologies for ICT for the ageing society.** Delivery date: M18

Responsible
AIT (Lead Contractor)

Participants
AIT, CURE

Version: 1.0
Date: 10.11.2011
Dissemination level: (PU, PP, RE, CO): PU

Abstract

FoSIBLE fosters the application of mulit-modal input approaches like gesture recognition, smart furniture and remote control via touch-screen based devices. To find out, which gestures are acceptable and joyful for the target group of elderly people, a user experience study was carried out asking for gestures with low error rate, gesture joyfulness and acceptance of the method. The results show that "Hand Movement Tracking" is ranked best by the test persons and could performed rather fast with a low error rate. The overall good ratings of all menu types crowned with best values for joyfulness show that gesture control for TV like planned in FoSIBLE is considered as an alternative remote control for older people.

Table of Content

1 Introduction 4
1.1 Purpose of the Document 4
1.2 Definitions, Acronyms and Abbreviations 4

2 System description 5

3 User Experience of Gesture based Menu Interactions 7
3.1 Prototype 7
3.2 Test setup 9
3.3 Results 10
3.3.1 Performance Measures 10
3.3.2 User preferences and Acceptance 12
3.4 Implications for FoSIBLE 13

4 Design of an Integrated Environment for Social Interaction 14
4.1 Multi-modal input approach 14
4.1.1 Gesture recognition 15
4.1.2 Smart furniture 16

5 References 17

1 Introduction

1.1 Purpose of the Document

The purpose of this document is to clearly line out the approaches for user-centred design methodologies, which are applied in the FoSIBLE project. The document shall give an overview on how user-centred design is implemented in the current project.

1.2 Definitions, Acronyms and Abbreviations

Acronym	Description
FoSIBLE	Fostering Social Interaction for the Well-Being of the Elderly
IF	Interface
PC	Personal Computer
AE(R)	Address Event (Representation)
3D Sensor	Devices that delivers spatial (three dimensional 3D) information of a given scene , typical x,y and depth-z
SW	Software
HW	Hardware
OS	Operating System
BI	Behavioral Intention
OUT	Output Quality
PEOU	Perceived ease of use
PEC	Perception of external control
ENJ	Perceived Enjoyment

2 System description

The FoSIBLE System allows the interaction of elderly people via a social media platform. A central server will host a data broker, database and webserver which comprise the media platform. Several components are installed in the end-user's homes which run user-interfaces to allow the interaction with the FoSIBLE social platform.

Commands can be sent to the Smart-TV (HBBTV) using gestures or a Tablet PC. The user has the free choice to perform simple gestures for simple interactions with the TV (e.g. change channel, etc.) or use the Tablet PC for more severe actions.

The gesture Recognition Components are a Mediacenter PC running Microsoft Windows and the two 3D sensors UCOS and Microsoft Kinect. The two sensors use a different technological approach to generate 3D depth data. The parallel use of these technologies is primarily for evaluation.

The PC is used to run the Software, which will process the sensor data to recognize the gestures (Gesture Recognition Module). Probability values p are handed over to the Gesture Interpreter which will decide the meaning of the gesture and send a command to the correct receiver (TV or databroker).

The databroker is the central data handling component of the project. Data from all input devices (Tablet-PC, gesture control) and other sensors are collected in a central database. Part of this data are used as content for the Social media platform, which itself can be viewed either on the Smart-TV or the Tablet-PC.

A network connection (LAN/WLAN) is the primary communication media between the hardware components of the system.

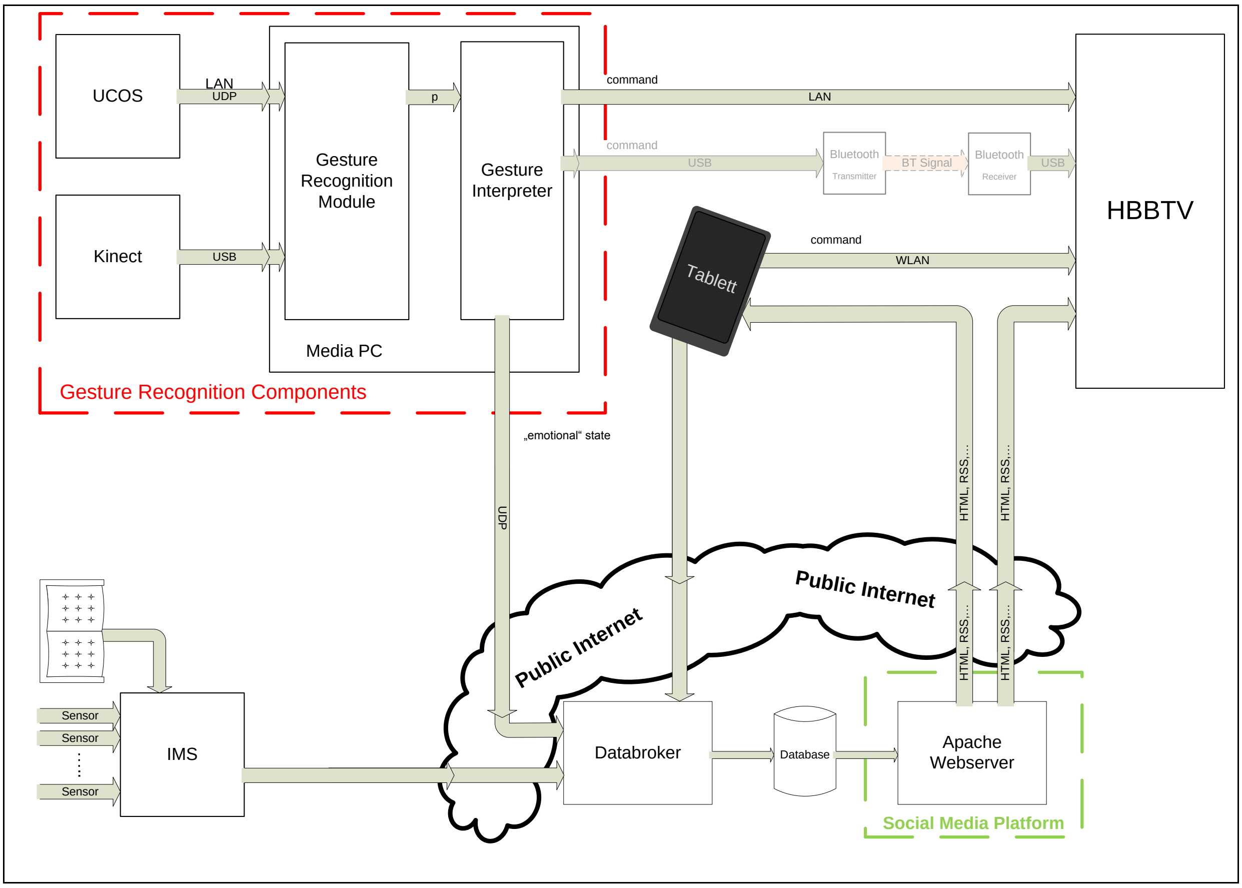


Figure 1: Overall FoSIBLE block diagram

3 User Experience of Gesture based Menu Interactions

Like introduced in Deliverable 5.1 we conducted a user experience study to gain knowledge which kind of menu is most feasible for gesture control in FoSIBLE. Gestures (manipulations and deictics) are generally considered to provide a very direct and natural way of interacting with a system. However, little research exists on how older people could benefit from gesture based interactions and if they meet the special needs and abilities of the elderly.

Older people tend to have particular problems with the interaction of new technology because devices are not designed to accommodate their special needs [3]. With the user experience study, conducted with representatives of our target group, we wanted to answer the following research questions:

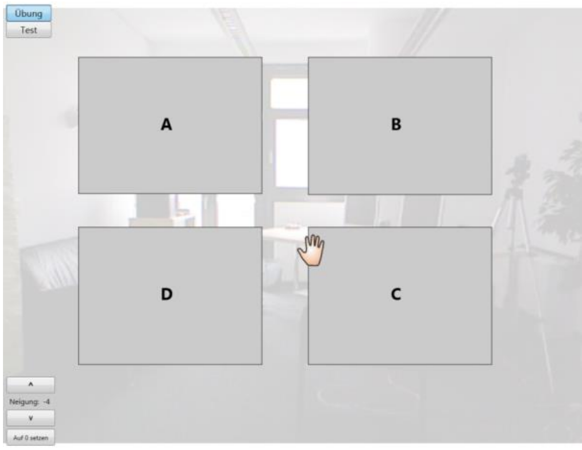
- Which kind of gesture control is the fastest for selection tasks and produces an error rate below 5%?
- Which kind of gesture control do older users prefer in terms of usability and joyfulness?
- Do older users accept gestures as input method?

In order to answer the research questions a comparative study with menu technique and amount of menu items as independent variables have been carried out.

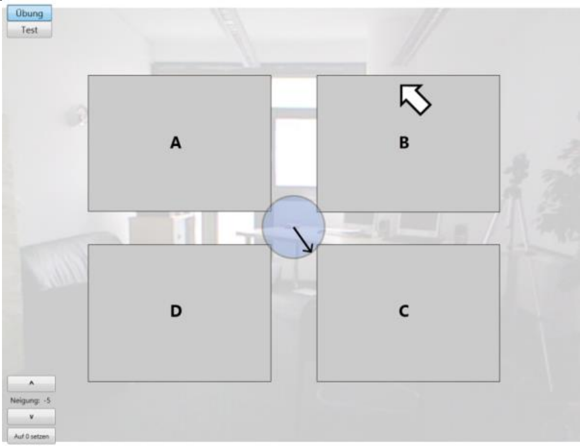
3.1 Prototype

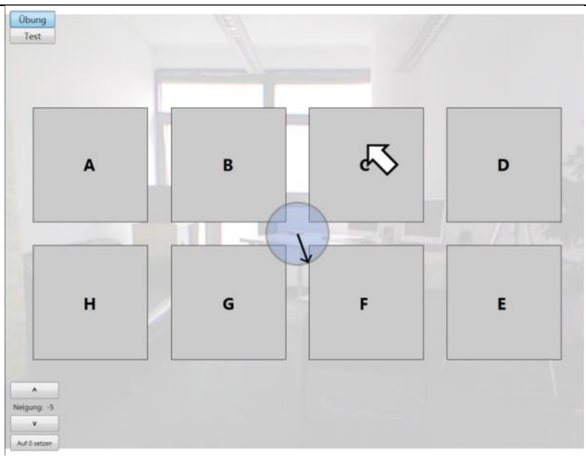
Subsequent to a literature analysis of simple gesture based menu interactions we selected four different types for testing them with older people: (i) hand movement tracking for cursor control, (ii) static hand positions for cursor control, (iii) hand strokes in a radial menu, and (iv) dial plate for a rotary menu. Every menu type has been implemented with a similar user interface with four or eight menu items respectively and is controlled with the right hand of the test persons. Table 1 introduces the used gesture-based user interfaces and provides screenshots of every menu type.

<u>Hand Movement Tracking</u>: The movement of the user's hand is tracked and controls a hand symbol. The cursor-like hand symbol is manipulated directly by the movements of the user.

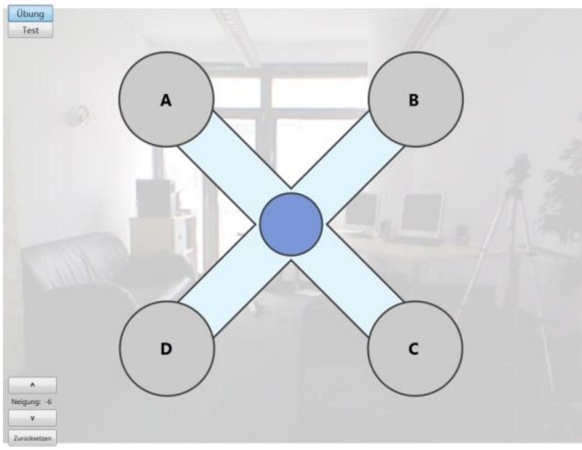


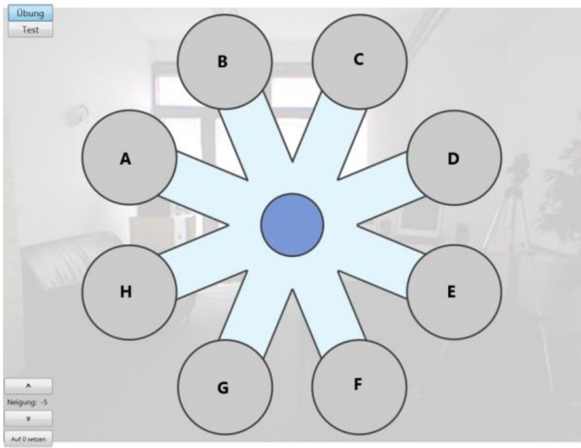
<u>Cursor Position Control</u>: The position of the (not moving) hand in relation to the neutral center of the screen makes the white cursor moving in the accordant direction. As long as user does not move his hand back in the neutral position the cursor remains moving. The black arrow in the center indictates the direction of the current movement.





<u>Hand Strokes</u>: Starting from the neutral position in the center the user just has to perform a stroke along one of the tracks to select the according menu element.





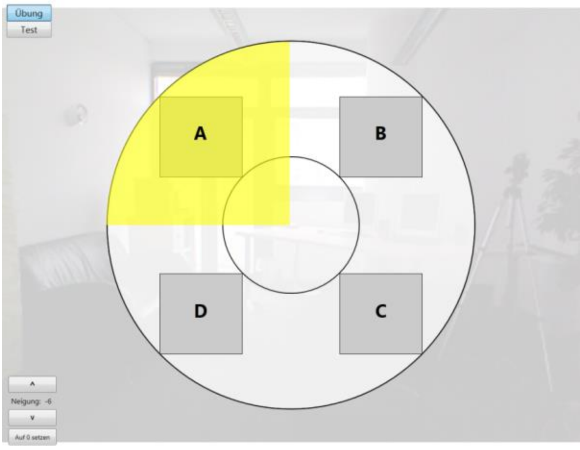
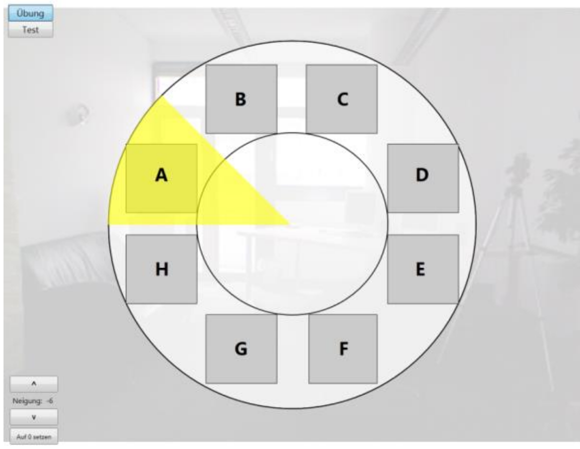
Dial Plate Circle: The user moves his hand along the circle to change the selection of the menu element.	
	

Table 1: This table presents the tested gesture-based menu interfaces. The adumbrated picture in the background is the view point of the camera and gives slight visual feedback about the current position of the hand. The buttons at the corners are for adjustments by the supervisor.

As it was neither feasible nor necessary to implement all menu types completely we used the methodology of Wizard of Oz [4]. This means in this case that we show an accordant TV user interface for every menu type to the invited test persons and give them visual feedback about their hand position. However, the initiation of the gesture recognition and the selection will be triggered by the test supervisor by hand. For the initiation the users were required to wave his hand shortly and for the selection the users need to perform a grab gesture (to form and release a fist).

The prototypes have been created with the Kinect for Windows SDK[1] to realize the tracking and with Microsoft Expression Blend[2] for the creating of the user interface. The output device was a 32” Samsung LCD TV with a refresh rate of 100 Hz.

3.2 Test setup

We recruited 24 right-handed test persons (thereof 12 women) between 65 and 73 years for conducting the tests. Before we turned to the actual testing of the gesture based menu types the test persons had to perform a simple motoric test. It was based on the standardised Box and Block Test [5] and the test persons had to put as many pieces as possible with one hand out of a box to an adjacent box crossing a 15 cm obstacle within 60 seconds. The motoric test was performed to better understand the performance results of the evaluation of the gesture based menu types.

[1] http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/

[2] http://www.microsoft.com/expression/

During the study every participant had to perform the same tasks for every menu type but in different order to exclude biases. A task consisted always out of: initiation, first selection and second selection. After the first selection the labels of the menu elements changed from letters to numbers. The instruction given to the test persons was before every task was e.g. "Please navigate to and select menu element B4." In case of a wrong selection the miss hit was counted as an error and the test person had to repeat the task. Only for the successful try the task completion time was measured.

Prior to the data collection an introduction to every technique and some practise time with each menu type were provided. Performance was measured through task completion time and the amount of selection errors. In addition the participants had to answer some selected items of the standardized questionnaire TAM3 [6] after every menu type to gain knowledge about the user experience and acceptance of the gesture based menu interactions. We utilised the following 7 items of the TAM3: BI1, BI2, PEC1, ENJ3, OUT1, OUT2, PEOU4 (see [6]). Furthermore we let every test person create a ranking wherewith they should order the four menu types according to their personal preferences, one for the menu types with 4 menu elements and one for the menu types with 8 menu elements.

3.3 Results

For the presentation of the results in this section we will use some abbreviations for the various gesture menus that raise the readability: H4 and H8 stand for Hand Movement Tracking, Cu4 and Cu8 stand for Cursor Position Control, S4 and S8 stand for Hand Strokes, Ci4 and Ci8 stand for Dial Plate Circle, each with 4 or 8 menu elements respectively. An influence of the motoric abilities on error rate, performance time and acceptance could not be detected.

3.3.1 Performance Measures

The error rates show overall quite good results for gesture based interactions (see Table 2). The lowest error rate of 0% could be reached with CU4. For H4, H8 and CU8 the error rates remain also beneath the aimed 5%. With 6% the error rate of S4 is also still acceptable.

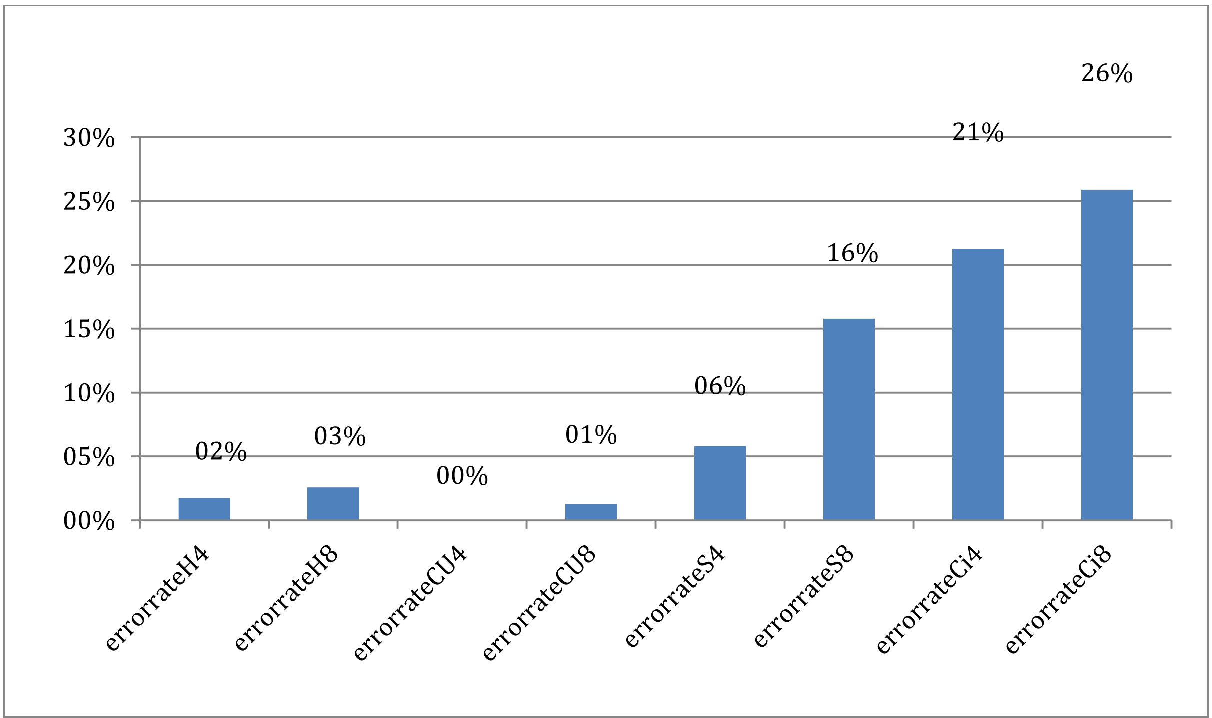


Table 2: Error rate for all menu types.

Rather easy to predict was that the error rates of the menu types with 8 menu items is always higher than their counterpart with 4 menu items. The big difference for the hand strokes can be explained with the nature of the interaction type itself and minor technical shortcomings. For four menu items the area of movement is twice as big and to aim towards the corners of the screen rather easy while for eights menu items the space for movement along one track is clearly smaller. The high error rates for the dial plate circle menus have two origins. At any time, one of the menu items is marked so every grab gesture of the user triggers a selection. Some users performed the selection gesture hastily which has for the other menu types no effect if nothing is marked. Besides limitations of the prototype let from time to time to short performance fall-offs and jerking so the false selection process was kind of supported.

Concerning the performance time the measured results of the menu items can be divided into three obvious groups (see Table 3). We call them: *very fast* (faster than 2 seconds), *fast* (around 6 seconds) and *slow* (between 12 and 14 seconds). The condition with the lowest performance time and only member of *very fast* is Ci4 with an average of 1,2 seconds. H4, S4 and H8 belong to the group *fast* while Cu4, Cu8, S8 and Ci8 show rather high performance times and belong therefore to the group *slow*.

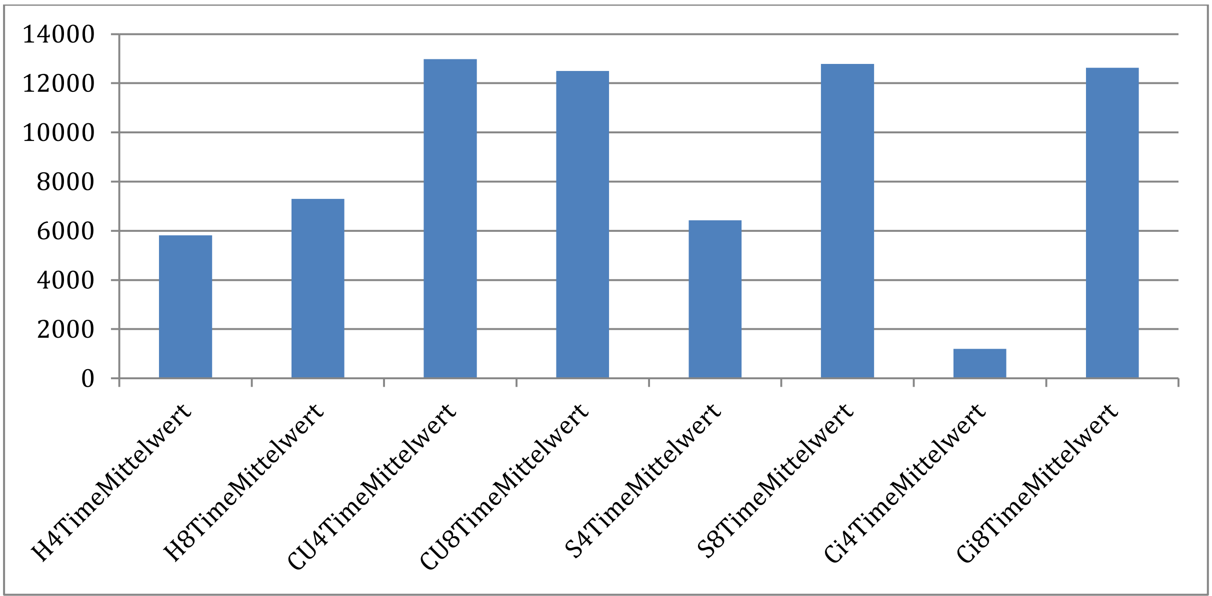


Table 3: Performance time for all menu types.

The very fast performance with Ci4 is caused by the nature of this menu type. In comparison to Hand Movement and Cursor Control there is no overshoot problem and the Hand Strokes need some time to position the hand in the centre area. So the participants could initiate the interaction and need to perform at most a semi-circle with their hand. In contrary to this stands the slow performance of Ci8. The reason for this is limitations of the prototype. It was not possible to transfer the circular movement directly onto one of the now smaller segments of the circle as the biggest distinguishable area was a quarter-circle. So the users had to move their hands further and the direct mapping between segment of circle and hand position was lost. The bad performance time of Cursor Control is caused by the constant movement speed of the cursor. The transfer from hand position to cursor controls only the direction of the movement and therewith a direct manipulation is not possible. Instead the user has to wait until the cursor reaches the desired position. The values of Hand Movement lay for both amounts of menu items in the group *fast* and are acceptable.

3.3.2 User preferences and Acceptance

The measured acceptance of the various gesture based menu types is rather high for every tested menu type. On the utilized 5-point Likert scale ranging from *Strongly disagree* to *Strongly agree* the lowest average value of the sample is 3.0 of OUT2 for S8, while the highest value is two times 4.2 of ENJ3 and PEC1 for H4. So differences between the eight menu types are rather small. However, the values of S8 decline a little bit (see Table 4). The highest overall score of 4.0 is measured for H4 and Cu4, the lowest of 3.1 is measured for S8. When focusing on usability only the values PEOU4, PEC1, OUT1 and 2 are relevant. However, there are almost no differences: the best scores receive Cu4 (4.1), H4 (4.0) and Cu8 (4.0); the worst receives S8 (3.1).

We found high scores on joyfulness (ENJ3) for all menu types. Except S8 (3.5), all tested menu types received at least a value of 4.0. So ENJ3 received the highest values of all items.

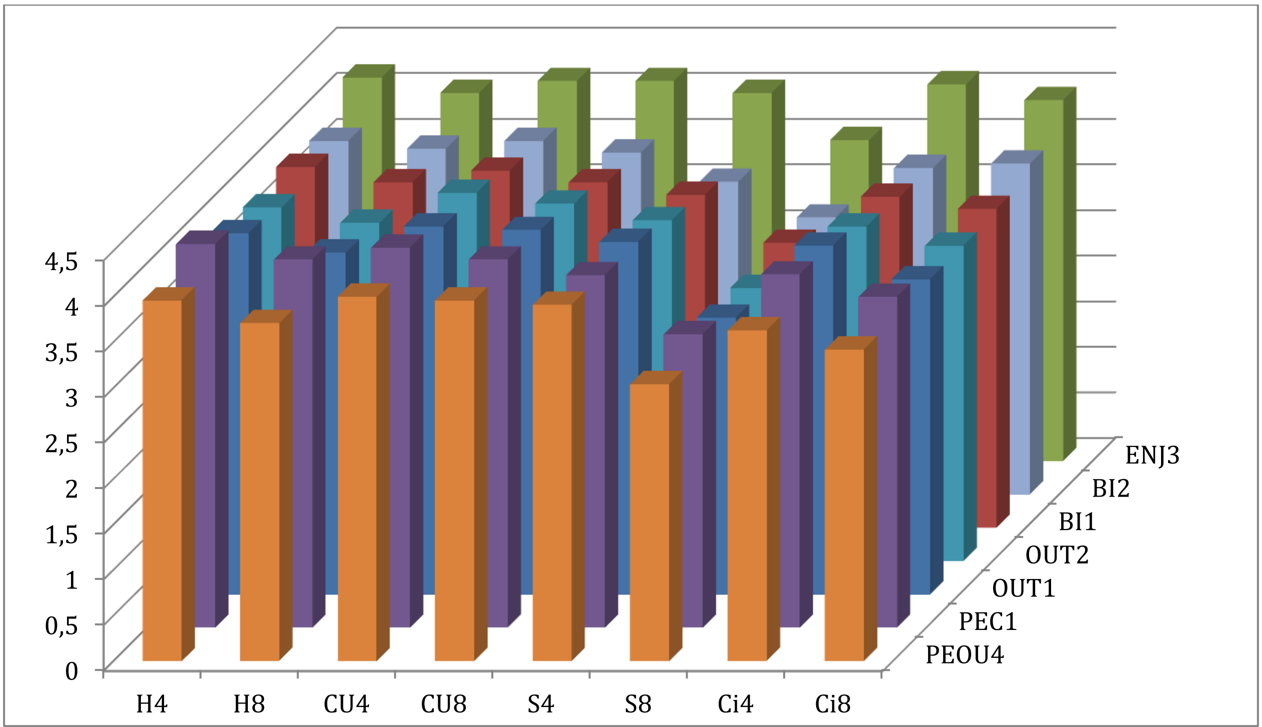


Table 4: The average acceptance values of the involved TAM3 items for the tested menu types.

The results of the personal rankings of the test persons (see Table 5) affirm the measured findings in some respects. Not surprisingly Hand Movement Tracking is ranked clearly on top as it performed well in objective and subjective test dimensions. The very low error rate and the good usability values suffice for Cursor Position Control for rank 2 although performance time is poor. The highest error rate of all menu types is probably the reason why Dial Plate Circle is ranked worse than Hand Strokes although S8 received the worst values in the acceptance questionnaire. The reason for the small differences types between 4 and 8 menu elements is that the majority of the test persons used the same ranking for both.

	Average rank for 4 elements	Average rank for 8 elements
Hand Movement Tracking	1,8	1,7
Cursor Position Control	2,5	2,3
Hand Strokes	2,8	2,8
Dial Plate Circle	2,9	3,1

Table 5: Rankings of menu types by test persons.

3.4 Implications for FoSIBLE

As Hand Movement Tracking is ranked best by the test persons and could performed rather fast with a low error rate we can clearly recommend to go for this gesture based menu type in the FoSIBLE prototype. The overall good ratings of all menu types crowned with best values for joyfulness show that gesture control for TV like planned in FoSIBLE is considered as an alternative remote control for older people.

4 Design of an Integrated Environment for Social Interaction

The central element of the application is a **Smart-TV system**, which is used to display messages, images and videos. A dedicated application runs on the Smart TV and provides chat functionality, as well as games and virtual libraries. To encourage the elderly to deal with this application, different input methods are supported. In addition to traditional manufacturer-specific remote controls or applications, gestures can be used for navigating through the menus, while a tablet-PC provides easy text input as well as the ability to navigate the system. The system is completed by sensor information and smart furniture to detect the presence and state of the user.

4.1 Multi-modal input approach

Multi-modal input approaches for controlling Smart-TV systems seem promising from a usability perspective, as active user input can be reduced to essential tasks. Based on context analysis and evaluation of measurements provided by visual, pressure and movement sensors, appropriate assistance can be provided, for instance by pre-selecting relevant menu options, automating sign-in processes and suggesting relevant information. In conjunction with a natural user interface, based on touch interactions on a tablet-PC as well as gesture recognition, applications can be designed to support heterogenic user groups, even without previous technology experience. To reach this goal, guidelines for user interface design need to be taken into account, supported by extensive end-user involvement through the whole development process.

Smart TV is a recent approach to integration of web-based interactive content into modern television sets. The application is designed as a widget that runs on a standard Smart TV platform and which is connected to our social community platform (Figure 1). It implements social features and functionalities identified during an end-user requirements process, including support for communication and awareness while watching TV, as well as a social media platform for exchanging information (eg on interesting TV programs or books) and playing games. A presence awareness implementation allows for further enrichments by making use of various sensors and smart furniture available in the future living room.

Figure 2: A prototype of the Smart TV widget for the FoSIBLE Social TV Community.

4.1.1 Gesture recognition

Gesture recognition promises to provide a simple to use, intuitive interface for interacting with Smart TV platforms. The appearance of Microsoft's Kinect device lead to a boom in the development and demonstration of gesture control interfaces. However, devices like Microsoft's Kinect regularly target only three-dimensional scene recognition and additional hardware and software are often necessary to perform the actual gesture recognition in real time. This results in increased system cost and complexity when deployed, because the device cannot be connected to a TV-set or set-top box directly. In FoSIBLE, we therefore targeted the development of a fully embedded system for hand gesture recognition based on a biology inspired "Silicon Retina" optical stereo sensor, which is ideally suited for the recognition of dynamic gestures. The sensor is only sensitive to light intensity changes and therefore tracks hand motion in its field of view robustly and with high accuracy. It can detect whether the user is in front of the TV set, and whether the user is in company or alone. The hand gesture recognition software (see Figure 2) for the embedded device is developed within the FoSIBLE project and will be evaluated against comparable gesture recognition software available for the Kinect device. The underlying gestures have been defined as a result of an end-user evaluation process, to ensure that the target group can cope with the defined gestures.

Figure 3: A sequence of still images provide an example of raw hand tracking data (left pane) and synchronous video (right pane) for a person performing a "circle" gesture.

Although gestures can be used for most navigation and selection tasks, not all inputs can be done using gestures alone. Furthermore, some users might not want to use gestures at all. Thus, additional input modes should be provided to users, which they can use to navigate the Smart TV system, but also to enter text (eg during chats or when writing a short article). As traditional input devices like keyboards and mice are quite cumbersome to use in a living-room environment, a tablet PC is used for this task. The tablet can also be used to display messages when the TV is switched off and is connected to the Smart TV system and the Social Community platform.

4.1.2 Smart furniture

Additionally, smart furniture is used to detect the presence and state of the user and to adopt the system to the users need. For this purpose, touch sensitive tables are used, as well as sensor equipped chairs and beds. By detecting the position of the user, context sensitive functionality can be provided, such as to start an application when the person sits in front of the TV and interacts with the couch-table. Furthermore, it is possible to adjust the TV screen and the camera, so they are directed towards the user.

The integrated FoSIBLE Social TV community system will be deployed in multiple home environments in France and Germany throughout the next year, delivering more results from the real-life usage and more insights about the end-user acceptance and usability for novel multi-modal interaction techniques for Smart and Social TV systems.

5 References

[1] Primesense, The PrimeSensor Reference Design 1.08 datasheet, Available at http://www.primesense.com/files/FMF_2.PDF as of 2011-03-15

[2] UCOS2 data sheet - http://www.ait.ac.at/research-services/research-services-safety-security/new-sensor-technologies/development-of-embedded-systems-for-customer-specific-solutions/smart-eye-ucos-universal-counting-sensor/?L=1 as of 2011-08-31

[3] Jetter, H.-christian, Gerken, J., & Reiterer, H. (2010). Natural User Interfaces : Why We Need Better Model-Worlds , Not Better Gestures. Natural user interfaces: the prospect and challenge of touch and gestural computing (Workshop CHI 2010) (pp. 1-4).

[4] Kelley, J.F. (1984). An iterative design methodology for user-friendly natural language office information applications. ACM Transactions on Office Information Systems, March 1984, 2:1 (pp. 26–41)

[5] Mathiowetz, V., Volland, G., Kashman, N., Weber, K. (1985) Adult norms for the Box and Block-test of manual dexterity. In: American Journal of Occupational Therapy, 1985, 39, 386-391

[6] Venkatesh, V. and Bala, H. (2008), Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences, 39: 273–315