http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (14): 2641-2645, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.2641.2645

Multi-modal Based Human-computer Interaction System for Children

[1, 2]Fang Zhi-gang, [1]Yang Zhong-Hao and [1]Luo Wei-bin
[1]Institute of Electronic Circuit and Information System, Zhejiang University, Hangzhou, 310007, Zhejiang Province, China
[2]Zhejiang University City College, Hangzhou, 310015, Zhejiang Province, China

Abstract: In this article, we first analyze the advantages and disadvantages of the existing interactive systems for children; then from the perspective of design psychology and cognitive psychology, we propose child-centered principles for interaction design, based on which NIPad, a multi-modal human-computer interaction system for children, has been designed and implemented to evaluate the feasibility of our principles and system framework. In this system, gesture input and tangible input data are integrated by finite-statemachinebased semantic integration algorithm and a kind of multi-dimensional interactive task is generated to control Avatar£¬therefore, NIPad can achieve the purpose of what you get from what you see.

Key words: Children, multi-modal, natural Interaction

INTRODUCTION

As we all know, playing games is the nature of children. With the advancement of technology, computer games have become an essential part of them. However, few of the games are child-centered, since the majority of their Graphical User Interface (GUI) is based on Windows, Icons, Menus, Pointer selection (WIMP). Because of these features, some organs of children's body are over-used. Researches on children-oriented human-computer interaction have achieved a lot success both at home and aboard, such as MIT's Multi-jump, University of Waterloo's Solar Scramble, Northwest University's Encyclopedia and Zhejiang University's Kid Movie Creator, however, those software just only focus on a special technology such as tangible user interface, voice interaction, gesture interaction, high-level semantics based interaction and so on, instead of incorporating these technologies to present a unified framework for children-oriented human-computer interaction. In this article, we propose child-centered principles for interaction design which are based on design psychology and cognitive psychology and then design NIPad, a multi-modal human-computer interaction system for children to evaluate the feasibility of our principles and system framework.

CHILDREN'S COGNITIVE PSYCHOLOGY THEORY

Gestures and children's cognition: When people speak, they always unconsciously do the gestures (Feyereisen, 1991), as well as children. Research has shown that some gestures have positive effects on children's cognition. Babies have obvious synchronous reaction to adult's words and some parts of their bodies (such as hands, arms, or lips) will do the corresponding action according to the different words (Li, 1993). The research also has find that the children who make gestures spontaneously when learning can have a deeper impression on what they have learned than these who don't (Alibali and Goldin-Meadow, 1993).

Tangible user Interfaces and children's cognition: Tangible user Interfaces mainly use physical manipulative tools to control or represent the information of virtual world. (D'Amico *et al.*, 2010). Physical manipulative tools can help children in two ways in cognition: Offloading cognitive burden, understanding conceptual metaphors. McNeil and Jarvin (2007) show that physical manipulative tools can provide additional channels of sensory information to children, such as visual and touch sensory information. The visual and touch sensory properties can help children pay attention and also reduce the burden in memory load (Massarani and Gosselin, 2011).

CHILD-CENTERED PRINCIPLES FOR INTERACTION DESIGN

Considering the above analysis, we propose some basic principles for designing children-oriented interaction as following:

Corresponding Author: Yang Zhong-Hao, Institute of Electronic Circuit and Information System, Zhejiang University, Hangzhou, 310007, Zhejiang Province, China

- Non-invasive interaction should be the main technical. There is no need to wear any equipment since the body itself is a direct natural interface
- Focus on the application of the Tangible user Interfaces. Tangible Uuser Interfaces can connect the real object with the virtual world and show the mapping relationship
- Multi-channel and context-aware are integrated in this system. The recognition rate of a single channel is limited, but it can be highly improved with the data of multi-channel
- Correct and simple conceptual model is the basic content

CHILDREN-ORIENTED HUMAN-COMPUTER INTERACTION PROTOTYPE SYSTEM

In this article, We design a multi-modal human-computer interaction system for children called NIPad (Natural Interaction Pad) to evaluate the feasibility of our principles and system framework. We will describe the design and implementation of INPad in detail in the following part.

System framework of NIPad

NIPad consists of three modules: Gesture recognition module, tangible user interface module and 3D game module. The structure diagram of NIPad is shown in the Fig. 1.

Gesture input: Depth camera will catch the user's depth data to find the user from the background in this module. After this, it will identify user's whole body bone joints and then lock the user's hand joint to track and identify hand gestures.

Label input: This module is implemented by a portable Bluetooth card reader and this module can realize the function of wireless access and transmission.

Other input channels: We can add new channels since the NIPad we proposed is extensible, including touch screen, voice channel.

Interaction design of NIPad

Direct conceptual model: Children can control Avatar to interact with virtual reality with the help of NIPad. Waving their hands is the way to get the control of Avatar. Walk, Jump, and Idle are three different ways children can choose through the physical input. Finite-statemachinebased semantic integration algorithm is applied in the system and children can control Avatar to move to the specified location after he gets the control.

Multi-channel feedback: There is a radar view in the game to display the state of the game. Green means children have already got the control and red means they haven't. When children quit the game, it will be yellow. Children will be give different scores according to the objects Avatar have got. The system also gives children visual feedback by recording and displaying the final score as well as auditory feedback.

KEY TECHNOLOGIES OF NIPAD

Gesture input of the kinect: In this article, we catch the user's depth data with the help of Microsoft's Kinect and

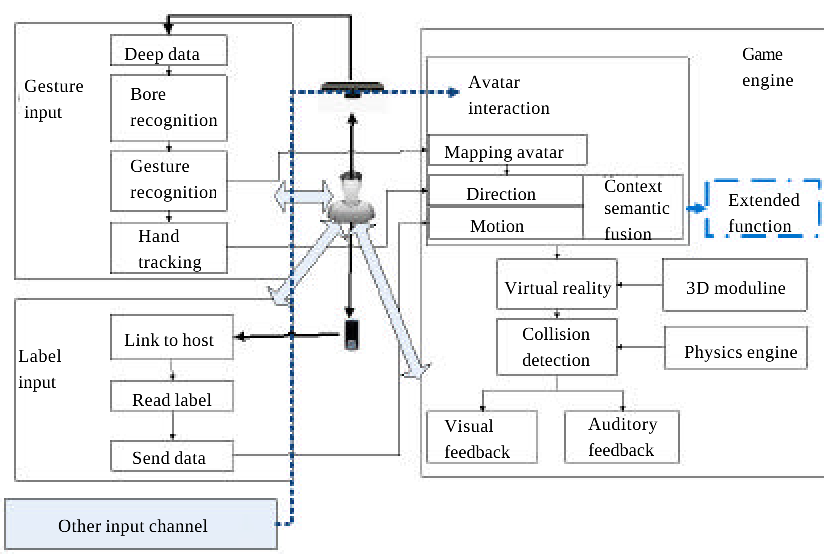

Fig. 1: Children-oriented human-computer interaction system framework

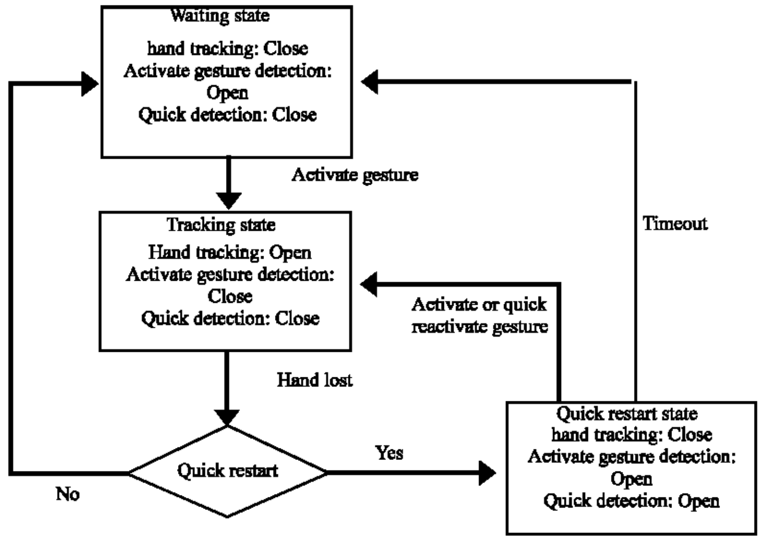

Fig. 2: Gesture interaction state transition model

then use the framework of the Open NI and NITE algorithm to identify 15 skeletal joints of the body. The state transition model of gesture interaction is shown in the Fig. 2.

Waiting state: The system will scan the scene and wait the gesture signal (Wave, Push, Raise hand and Swipe) to activate the system.

Tracking state: The system will track user's hand which has made a gesture signal to activate the system.

Quick restart state: When system can't find the hand of one user's, the user can still get the control back within the scope of time set by the system.

RFID physical input: The idea of physical input is to use physical manipulative tools to control or represent the information of the virtual world, thus combine the real life with the virtual world seamlessly. We use RFID and Label to achieve this function.

D modeling and character animation: We use Maya 2012 to create Avatar and use Unity3D to make a simple game scenario. Character animation production process includes: Role modeling, role bindings and key frame animation.

Semantic integration: There are two main types of multi-channel integration methods: One is characteristic quantity integration, the other is semantic integration. Compared with characteristic quantity integration, semantic has the following advantages: Can be individually trained; further training is not required after integration; existing recognizers can be used in a standard channel without modification. Considering the complexity and stability of the system, we use the method of semantic integration.

Interactive semantic: Interactive semantic means interactive task itself uses structured semantic representation. In the other word, interactive task is expressed as a set of a collection of structural components with semantic information. Interactive semantic commonly means a combination of semantics and syntax (Wang, 2008). The ultimate representation of interactive semantic is in the form of using semantic slot in the layer of semantic. As shown in Fig. 3, semantic slot includes task action, action object and auxiliary parameters of the action or action object.

Extraction of interactive semantic: According to the idea of the hierarchical model, this system is divided into four layers. They are the physical layer, the lexical layer, the syntactic layer and the semantic layer as shown in the Fig. 4.

The original bit stream of interactive input device channel is extracted in physical layer. The formats of bitstream are different if the interactive input devices are different.

What the lexical layer provides is the smallest logical unit:Input Primitives (IP) and the syntax layer represents the interactive operating rules between computer and users. The final interactive semantics are extracted in semantic layer, expressed in the way of using semantic slot.

Task action	Action object	Action parameter

Fig. 3: Semantic slot structure

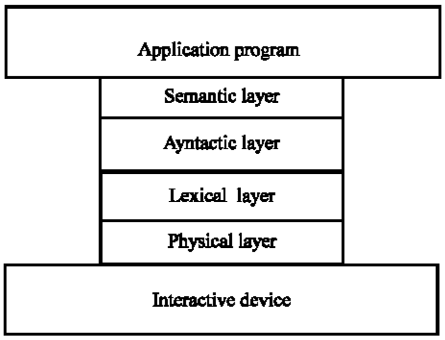

Fig. 4: Hierarchical model of the system

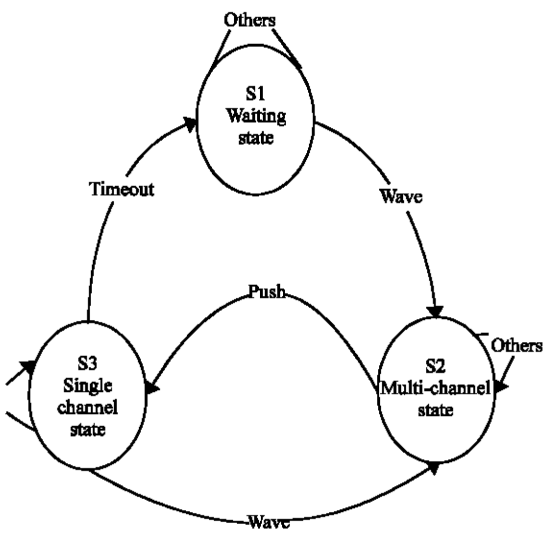

Fig. 5: Finite-statemachinebased semantic integration module

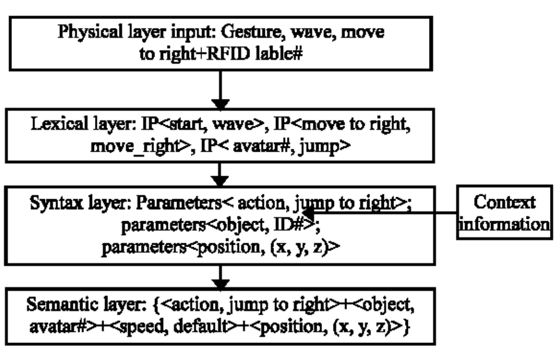

Fig. 6: Interactive semantic hierarchical integration process

Integration of interactive semantics: The gesture input and tangible input data are integrated by finite-statemachinebased semantic integration algorithm in this article to improve children's interactive experience. The model of finite-statemachinebased semantic integration is shown in the Fig. 5:

- **S1:** Waiting state. System can only recognize Wave gesture and will not respond to any input information if system hasn't recognized the Wave gesture
- **S2:**Multi-channel state. System will integrate the gesture input with the tangible input and a kind of multi-dimensional interactive task will be generated to control Avatar. System will turn to S3 state when detects Push gesture
- **S3:** Single channel state. Physical input is the only input channel and system will return to S2 state when recognizes Wave gesture and to S1 state if the pause time is more than set time

The process of finite-statemachinebased interactive semantic hierarchical integration is shown in the Fig. 6.

The kinect's depth data is the main input of the physical layer. User need to shift his hand to the right and read the label with the RFID after waving his hand. The action Wave is defined as the start of the interactive primitives in our system in lexical layer and shifting one's hand to the right means moving Avatar to the right in the virtual world. With the help of different labels read by RFID, Avatar can walk or jump. In syntax layer, we extract the intention of user's input. That is the parameters of the interactive semantic, including action, object and position. When there are multiple interactive objects, system determines to accept whose action according to the context information and provides the position of the action (x, y, z). Finally, system gets all the interactive information of the semantic slot in the semantic layer and the complete semantic structure is like this: (<action: Jump to Right>+<object: Avatar>+<parameter: Speed (default)+position (x, y, z)>).

CONCLUSION

We design a human-computer interaction system: NIPad for children in this article from the perspective of design psychology and cognitive psychology. NIPad has been fully tested according to the task flows and experiments prove that NIPad has the feasibility of practical application. But there still exist some deficiency to be improved: Only four gestures can be recognized by system which is not enough and the immersive experience of the game is not strong which also need to be improved.

ACKNOWLEDGMENTS

Our Research Project was fully sponsored by Bureau of Hangzhou city science and technology project, Hall of Zhejiang province science and technology public welfare project and Zhejiang University City College cross research foundation project with grant number 20120633B32, 2012C21038 and J-13027. Thanks for the sponsorships.

REFERENCES

Alibali, M.W. and S. Goldin-Meadow, 1993. Gesture-speech mismatch and mechanisms of learning: what the hands reveal about a child's state of mind. Cognitive Psychol., 25: 468-523.

Feyereisen, P., 1991. Gesture and Speech. Cambridge University Press, New York, pp: 106-113.

D'Amico, G., A.D. Bimbo, F. Dini, L. Landucci and N. Torpei, 2010. Natural Human Computer Interaction: Principles, Methods and Applications. In: Multimedia Interaction and Intelligent User Interfaces, Shao, L., C. Shan, J. Luo, M. Etoh (Eds.). Springer, London, ISBN: 978-1-84996-506-4, pp: 85-106.

Li, H., 1993. Research on baby's language and behavior. Psychol. Dev. Educ., 2: 39-43.

McNeil, N.M. and L. Jarvin, 2007. When theories don't add up: Disentangling he manipulatives debate. Theory Practice, 46: 309-316.

Wang, L., 2008. The Research of 3D Technology Based on Semantic Interaction in a Virtual Environment. USTC Press, UK., pp: 209-216.

Massarani, R. and P. Gosselin, 2011. French-speaking children's freely produced labels for facial expressions. Proceedings of the Meeting of Society for Research in Child Development, Montreal, Canada, pp: 305-309.