Appeared in Proceedings of Intelligent Building and Smart Home Conference, 2009.

A Prototype of Smart Interactive TV

Hung-Wei Lee
Department of Applied Psychology
Hsuan Chuang University
Hsinchu, Taiwan
spoon@hcu.edu.tw

Wen-Hung Liao
Department of Computer Science
National Chengchi University
Taipei, Taiwan
whliao@cs.nccu.edu.tw

Shwu-Lih Huang
Department of Psychology
National Chengchi University
Taipei, Taiwan
slh@nccu.edu.tw

Yu-Chieh Chang
Department of Psychology
National Chengchi University
Taipei, Taiwan
96752501@nccu.edu.tw

Ying-Szu Chen
Department of Computer Science
National Chengchi University
Taipei, Taiwan
g9710@cs.nccu.edu.tw

Tsai-Yen Li
Department of Computer Science
National Chengchi University
Taipei, Taiwan
li@nccu.edu.tw

***Abstract* — In the smart living space of the future, the interaction between humans and machines will become more frequent and diverse than ever. We will need more attentive and affective designs of human-computer interface in all living fields, including study rooms, bedrooms and living rooms. Our research is aimed to design a prototype of a new interactive TV system which implements both instrumental and psychological services. The instrumental services include new channel selection and volume adjustment functions, while the psychological services include program recommendations and other context-aware reminding functions. The prototype is not only designed and implemented from scratch but also validated by psychological experiments for better performances.**

Keywords- smart living space; interactive TV; prototype

I. INTRODUCTION

With the great progress of modern civilization, the roles played by information technologies in everyday life become more and more important. It is expected that in the smart living space of the future, information service systems that implement attentive and affective computing will be the mainstream, and the interaction between humans and machines will become even more frequent and diverse. Nowadays, in major living fields such as the study room, the bedroom, and the living room, information technologies are ubiquitous. Therefore, the design of the smart living space for the future should step further to focus on how to detect users' cognitive and affective states actively and provide pertinent services to satisfy their needs. The purpose of the present study is to design a considerate prototype of interactive television (iTV) for the creative imagination of a smart living room in the future.

The concept of iTV has already existed for a long time. For decades, numerous attempts and achievements have been proposed and accumulated both in the fields of academic research and practical business [1]. However, few previous exemplars have been designed to take both sides of the environmental context and users' psychological states into account. Therefore, in this work, we aim to design a prototype of iTV system to implement instrumental and psychological services. As for instrumental services, we mean new ways of operating the TV set to make it more useful and usable, such as new channel selection and volume adjustment functions. On the other hand, as for psychological services, we mean new ways of interacting with TV to enhance users' personal experiences and satisfy their psychological needs, such as program recommendations and other context-aware reminding functions. In addition to the accomplishments of these innovative functions, we also design several psychological experiments to validate the effectiveness and efficiency of the prototype for further refinement.

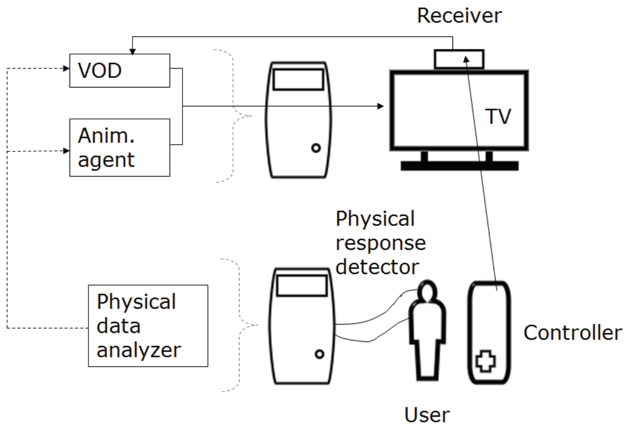

Figure 1. An overview of the system.

II. OVERVIEW OF SYSTEM

A. Framework of the system

In order to implement the prototype of such a smart iTV system, it is crucial to choose an appropriate development platform first. However, most of the commercial iTV platforms at present (e.g. Tivo [2]) do not allow designers to edit program contents and control hardware directly. Besides, they usually provide no effective tools to record the user's viewing and operating behaviors which are crucial for further psychological experiments and validations. Therefore, as shown in Fig. 1, we have developed an iTV platform which integrates the following major features. (1) Different types of TV-watching scenarios can be provided, inclusive of customized video-on-demand

This work was partially supported by the National Science Council under the INSIGHT project by NTU and the contract NSC 97-2627-E-004-001.

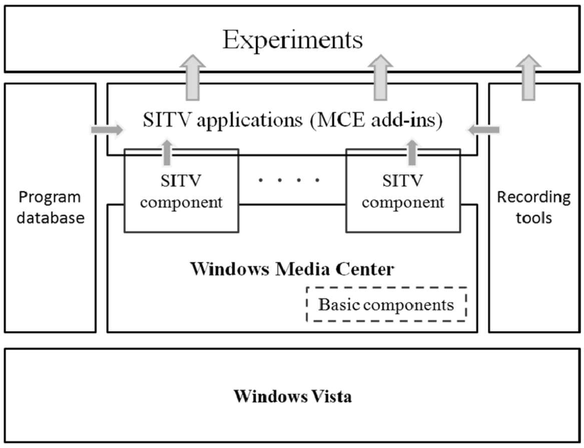

Figure 2. System framework

(VOD), emulated cable TV system, and real cable TV. (2) Direct access to the hardware and program contents can be provided, inclusive of controlling TV operations, scheduling or broadcasting TV programs, designing advanced TV functions, and creating innovative user interfaces. (3) Automatic collection and coding of the user's viewing and operating behaviors can be provided, inclusive of the operation of remote control and other behavioral and physiological signals.

The system is named Smart Interactive TV (SITV), and it contains three layers (as shown in Fig. 2). At the bottom layer, we use Windows Vista as the operation system. At the middle layer, there are several application components. Some of them are provided by Windows Media Center (MCE) in Vista, and others are customized SITV components. Combining these components with program database and recording tools, at the top layer, we can design and execute experiments to improve the services and usability of the SITV prototype. The prototype is programmed with the C# language and Microsoft MCE SDK while the user interface is designed with the MCML presentation language. Besides, we use SQL Server 2005 to build the database of TV programs, schedules, and users' information.

Although MCE has already provided designers with various basic components such as hardware control and program broadcast, in order to implement functions beyond hardware operation, i.e. psychological services, two more components specific to SITV are developed: *component for detecting user's preference* and *component for detecting user's identity and motion*.

B. SITV component for detecting user's preference

Program recommendation is one of the most popular topics in the field of iTV research [3][4]. The SITV component for program recommendation is based on the user preference which is learned by an artificial neural network (aNN). We have adopted a package of aNN [5] from an open source web site [6] to build up a learning network for user preference. The goal of using such a learning mechanism is that when the user needs program recommendation, the system will calculate a predicted score for each of all TV programs available in the current simulated broadcasting channels. Then the channels

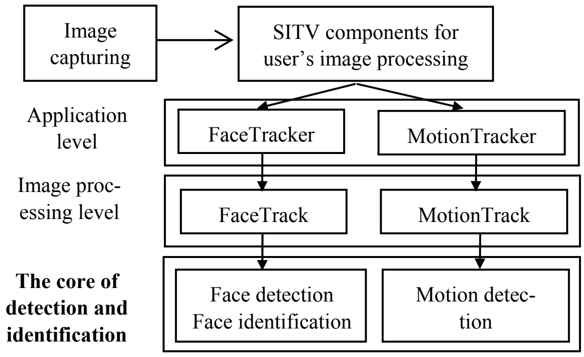

Figure 3. The structure of SITV component for user's image processing.

with highest scores will be recommended to the user. We use activation network model and adopt a sigmoid function which is shown below as the activation function.

$$f(x) = \frac{1}{1 + e^{-\beta x}} \quad (1)$$

In addition, we use a back-propagation learning algorithm to train the aNN. This algorithm adjusts the weights associated with nodes in the network by comparing difference between the target and the real values of each node.

As for the structure of the aNN, in addition to the input and output layers, a hidden layer is used to represent the nonlinear relationship within the network. There are no common rules to determine the number of nodes in the hidden layer. However, researchers believe that it is the number of input attributes, instead of the input nodes, that determine the number of nodes in the hidden layer. For this reason, we use four hidden nodes because there are four input parameters in the aNN, including program category, program language, short episodes or not, and user's emotional states.

C. SITV component for detecting user's identity and motion

Beside program recommendation, the other important SITV component is user detection and identification. This component is composed of several image processing modules, inclusive of one image capturing module, four internal image processing modules, and two external functions (as shown in Fig. 3).

The external image processing module contains two core functions at the bottom layer, i.e. face identification and motion detection. The face identification function includes two subcomponents: *face detection* and *face identification*. For the face detection component, we use a modified Adaboost [7] machine learning method and the Directional Edge Maps (DEM) to extract image features. The main idea of DEM comes from the concept of edge and Haar-like feature [8], but with information of "direction" added. Different operators represent different directions whose strengths are represented by calculated values. We use both strength and direction as the features for training. After training, we obtain a range of value for each feature that can be used to determine if a given test case is a face region or not.

Upon detecting the face region, SITV can identify the user for the next step. We collect several face pictures of single user and generate the most representative features for his/her face. When given a face at run time, we compare its features with the ones stored in the database and select the most matched one. Although the training method of face identification is similar to face detection, the training data for face identification need to be more homogeneous for better performances.

As for the motion detection function, we mark potential motion regions by inter-frame differences. The momentum and direction of a specific moving region are computed by means of connected component detection. Then we can use the motion information denoting the user's body movements to implement context-aware reminding services as described later.

III. INSTRUMENTAL SERVICES

Based on above-mentioned framework, SITV implements two instrumental services, including new volume adjustment function and new channel selection function.

A. *Volume adjustment function*

To make volume adjustment more efficient, we design and test four new ways of control. (1) Linear mode: The difference between the current volume and the ideal volume is divided into several equal steps. Each button-press adjusts the same amount but large steps go with large differences. (2) Degression mode: The serial amounts of volume adjustment for continuous presses decrease gradually for the user to fine tune when the ideal volume level is approached. (3) Frequency-dependent mode: The amount of adjustment depends on the duration between two consecutive presses. The faster the user presses the button, the greater the amount to be adjusted. (4) Semi-automatic mode: In this mode, the user can press a button to adjust the volume directly to a pre-defined ideal level and then fine tune it.

We conduct an experiment to evaluate five modes of volume adjustment, including four new modes just mentioned and the traditional one which adjusts one unit by every button-pressing. The result reveals that it takes more time and needs to press buttons for more times in the traditional mode than in the other four new modes, which means that the traditional mode is less efficient. As for the four new ways of volume adjustment, semi-automatic mode is the most efficient one because it takes less time and fewer button-pressing than others.

B. *Channel selection function*

With the fast expansion of the mass media market, the number and diversity of cable TV channels and programs are rapidly increasing. In order to cope with such an information explosion, we conceive a new way of channel selection that can help the user to retrieve the channels that he/she has visited. The new way of control is called Prior-N.

In Prior-N, the channels which are viewed by the user will be kept on a special list. By pressing a specific button on the remote control, the user can make the Prior-N list pop up (as shown in Fig. 4). We design and test three types of contents to be shown in the Prior-N list, inclusive of "channel number", "channel logo", and "representative scene of the program". The experimental result shows that most users prefer Prior-N than the traditional mode. When "channel logos" are used as the contents on the list, the users keep in using Prior-N most frequently. But the "channel number" mode seems the most efficient one because the users need less time to have necessary tasks done. However, according to the users' subjective experiences and reports, the "representative scene of the program" is the best design. We will need more theoretical backgrounds and empirical studies to clarify such seemingly inconsistent results.

Figure 4. The display of Prior-N list.

IV. PSYCHOLOGICAL SERVICES

In addition to instrumental services which relate to the operation of TV hardware, SITV also implements two kinds of psychological services, inclusive of program recommendation function and context-aware reminding function.

A. *Program recommendation function*

As mentioned above, program recommendation plays an important role in the contemporary iTV research. It brings about extensive discussions in both academic and practical communities [3][4]. But in SITV, we consider not only the user's personalized preferences, but also the influences of the user's emotional states on his/her own preferences and the needs for recommendations. Based on the rationale, we design a psychological experiment in which the participant's emotions are induced first, and then measure his/her preference scores for each TV program.

We use several pre-recorded TV programs as experimental materials and each of them belongs to one of the following five categories: TV show, movie, news, sport, and soap opera. The user's emotional states are induced to be positive, negative or neutral by listening to specific music. Three participants are recruited for the 10-day experiment. On the first day, music excerpts that can induce different emotions are listened and chosen by each participant. In the following nine days, one of the three target emotions is induced by listening to music first, and then the participants have to watch those programs and rate preference scores for each of them. Based on the experimental results, we underline two main points. (1) User's preference for TV program is fairly stable. (2) Therefore, emotional states do not significantly influence user's preference for TV program.

In addition to the psychological experiment aimed to reveal the influence of emotion on preference, we also use the above-mentioned empirical data to train an aNN and build up our own program recommendation function in SITV. Two-thirds of the data (the beginning 6 days for each participant) are used to train the network and the remaining one-third is used to validate its performance. The correlation between the predicted values from the aNN and the user's real preference scores can be used as an effective index. The result shows that the correlation coefficient is greater than zero which means that the aNN is good enough to give program recommendations. However, if we exclude the user's emotional states from the input layer of the aNN, no different results are found. Such a conclusion once again supports findings of the psychological experiment which claim that emotions have no significant influence on program preference.

B. *Context-aware reminding function*

In SITV, we can use customized components of face detection and motion detection to capture the user's identity, location, posture, viewing distance and other information in the context. Therefore, the following reminding/warning functions have been implemented.

- The power of TV will be turned on automatically while the user sits in front of it.
- A gentle reminding message will show on TV while the user leans against the sofa or tilts his/her head.
- A gentle reminding message will show on TV while the user gets too close to it.
- The program will be paused automatically while the user leaves and it will be resumed while he/she comes back.
- A gentle hint of using program recommendations while the user moves his/her body and/or switches channels frequently which may mean that he/she feels bored.

With the prototype of SITV, we have recruited ten users to give trials on those psychological services. According to their subjective experiences, the "automatically pause on leaving" function is the most popular one which is followed by the "hint to recommendation on frequent channel switching" function. However, several users mention that too many interruptions will be bothering, instead of helping. Therefore, we will raise the criteria of determining user's needs in later versions of SITV prototype. In other words, we would like to miss users' possible needs rather than interrupt them by false alarms.

V. Conclusion

In this paper, we focus on one of the most important fields of the future smart living space, i.e. the smart living room. We create an SITV prototype and make attempts to implement both attentive and affective computing. We enrich the SITV through instrumental services and psychological services for the users. Moreover, because our interdisciplinary research team includes both computer scientists and psychologists, we emphasize not only creative ideas for the new functions of SITV but also empirical validations and psychological backgrounds behind the user's TV-watching behaviors.

As mentioned from the beginning, with the great progress of modern civilization, the roles played by information technologies in everyday life become more and more important. As a result, in the smart living space of the future, we should try the best to concern and satisfy the natures and needs of those people who are going be served by any kind of new technologies.

References

[1] J. F. Jensen, "Interactive Television: A Brief Media History", Proc. of 6th EuroITV Conference, pp. 1-10, 2008.

[2] TiVo, http://www.tgc-taiwan.com.tw/

[3] J. Zaletelj, "Recommender System for the Multi-Channel TV Production", Proc. of 6th EuroITV Conference, pp. 102-106, 2008.

[4] R. Bernhaupt, D. Wilfinger, A. Weiss and M. Tscheligi, "An Ethnographic Study of Recommendations in the Living Room", Proc. of 6th EuroITV Conference, pp. 92-101, 2008.

[5] The Code Project: http://www.codeproject.com/

[6] The Code Project: Neural Network on C#, http://www.codeproject.com/KB/recipes/aforge_neuro.aspx

[7] Y. Freund and R E. Schapire "A Short Introduction to Boosting", in Proc. of Japanese Society for artificial Intelligence, vol. 14(5), pp. 771-780, 1999.

[8] P. Viola, and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", Proc. of Computer Vision and Pattern Recognition, vol.1, pp. I-511~I-518, 2001.