Using Intelligent 3D Animated Character as the Interface for Interactive Digital TV System

Ying-Szu Chen[1], Tsai-Yen Li[1], Shwu-Lih Huang[2], and Hung-Wei Lee[3]

[1] Computer Science Dept. National Chengchi Univ., Taiwan
[2] Dept. of Psychology, National Chengchi Univ., Taiwan
[3] Dept. of Applied Psychology, Hsuan Chuang Univ., Taiwan
{97753010,li}@nccu.edu.tw, slh@nccu.edu.tw,
spoon@wmail.hcu.edu.tw

Abstract. In this research, we propose to incorporate intelligent 3D character into interface design to enhance the user experience on the interactive digital TV system. Based on the smart interactive digital TV system, call SITV, developed in our previous work, we consider several interactive scenarios in the design of the intelligent 3D animated character. We propose to design our intelligent 3D character with the concepts of mobility and expressiveness on body motion such that appropriate emotions can be presented according to TV watching scenarios and character configurations. We have designed experiments to evaluate different types of user interface design for various scenarios. We have found that although the interface with animated character can attract the attention of the user and enhance interactivity, the text-based dialog box interface is still the easiest to understand. We also have found that when the freshness of the new interface with animated character decays after some time of use, it is crucial for an animated character to have expressive motions.

Keywords: Interactive TV, Animated Character, Emotion, User Interface.

1 Introduction

With the emergence of digital TV, the interaction between TV and the audience is becoming richer [4]. In an ideal interactive TV, the audience can not only passively watch broadcast TV programs but also actively inquire program information and retrieve desired programs in a timely manner. The functions of a TV are now extended to general media services as well as web accesses. Most interactive TV today can provide customized services such as stock information, news, forecasting, game, etc. However, most interactions happen between the remote server and the audience, and TV only serves as a means for display and a command receiver.

Most digital TV's are equipped with a microprocessor possessing computing power of various degrees. It is becoming more feasible to design better user interface and richer functions at the client side of a TV system (set-top box or the TV itself) to facilitate better interactivity. The authors in [1] suggested that virtual characters can play the role of assistant in consumer electronics to provide operational and informational aids. In this work, we propose to design a context-aware animated

M. Chang et al. (Eds.): Edutainment 2011, LNCS 6872, pp. 300–308, 2011.

character with emotive body expression to provide better visual effects in various scenarios of an enhanced digital TV with new interactive functions.

In order to study the effects of the proposed interface design on interactive TV, we have adopted a procedural animation system and an experimental platform for interactive TV, called Smart Interactive TV (SITV), implemented in our previous work [3,2]. On SITV, several context-aware functions based on remote TV controller and web camera have been implemented to provide convenient user experiences. However, in the previous work, only textual dialog boxes were available for delivering messages to the user. In this work, we attempt to incorporate an animated character in the user interface and define the scenarios and corresponding actions that an animated character should take to respond to a given event. In this paper, in addition to describing the system design, we also focus on the psychological experiments that we have designed to study and compare the effects of three different interfaces: regular dialog box, animated character without emotive expression, and animation character with emotive expression.

2 Related Work

Due to the development of computer technology and digital TV in recent years, there exist several application framework and solutions for digital TV or interactive TV. For example, TiVo [11] is a popular service that can provide customized TV watching experience. In our previous work on SITV [3], we have designed new instrumental services such as providing customized volume adjustment, channel recommendation functions and psychological services. We have used this system in this work to design and evaluate new interface in the emulated TV system.

For an animated agent to be believable, it is crucial for it to be able to express appropriate emotion in a given scenario. In addition to facial animation that is commonly used in character animation, body gesture is also important for emotion expression. Diederiks [7] repositioned the role of animated character in a smart environment and pointed out that nonverbal expression should be better utilized in addition to verbal expression when designing a virtual agent for user interface.

As the advances of character animation, adopting animated virtual agent as an interface for delivering messages is becoming popular. For example, the authors in [8] developed a system taking textual contents and gesture commands for a speech to generate corresponding animations and synchronized voice for a virtual weather forecast reporter. Thalmann and Kalra [5] have also designed a virtual actor for TV program. Seron [10] developed an animation engine that can generate animations of a virtual presenter delivering a speech with the voice of the author and interacting with the audience.

3 System Design and Implementation

Our system consists of two subsystems, whose results are composed through a video compositor to overlay the character animation on the TV program. The TV program runs on the SITV system [3] which emulates a live TV program broadcasting environment. This subsystem was customized based on the Window MCE edition and provided controllable content delivery for experimental uses. The character animation system was designed based on a procedural animation system implemented in Java in our previous work. This subsystem can generate character animation in real time with a given set of parameters specifying the desired animation. The two subsystems are connected through sockets for communication. We have designed a simple protocol for the SITV subsystem to send scenario description to the animation subsystem for the synchronous generation of appropriate animation.

3.1 Problem Definition for Motion Generation

We assume that our 3D virtual character moves in a 3D space that can be projected onto the TV screen. Since the dialog box and menus are displayed at different locations of the screen, in order to give the user a better sense of depth on a 2D screen, we assume that the character moves on a slanted surface such that it moves higher on the screen when it moves into the screen. In our interactive TV system, the character should be able to move appropriately and expressively according to the context of a given scenario. Therefore, we consider the problem of motion generation by taking the following factors into account: *motion*, *scenario*, and *character* as described in more details below.

Motion: An animated character is given a set of motion abilities, each of which is described with two attributes: *expressiveness* and *mobility*. First, it is important for our animated character to be able to express its emotion for a given scenario. However, if a virtual character on the screen is not large, facial expression may be difficult to tell. As pointed out by Meeren [6], body expressions have higher priorities than facial expressions when perceived by a viewer. Therefore, in this work, we use full-body gestures to express emotion when delivering a message.

Since different motions may carry different emotive meanings, we specify the range of emotion expressiveness for each motion procedure that we have designed. We adopted the basic emotion model that is commonly used in affective computing and select four common ones: *happiness*, *sadness*, *anger*, and *fear* in our implementation. Second, in order to attract the attention of the user to a message delivered by the system, the animated character in our system may move to a target location to play a selected motion. Therefore, our character needs to have mobility. For each motion defined in our system, we specify the degree of their mobility by the speed and striding ability of the given motion.

Character: In addition to motion ability, the model for each character also contains variables about its current status such as position and tiredness. The current position will be used as an input to the motion generation program to generate feasible motions moving the character to the goal. The tiredness parameter is used to adjust the strength of the emotion parameter when the motion generation problem is specified. In addition, when the accumulated tiredness reaches a threshold, the character may voluntarily decide to leave the screen to take a rest and come back later.

Scenario: In a smart interactive TV system, different interactive functions may be triggered based on different interaction scenarios and user needs. In SITV, we have implemented several instrumental and psychological services to enhance user experience. Each of these services entails a usage scenario requiring delivery of different types of messages such as warning and reminder. For example, when a user (could be a kid) sits too close to the screen, a warning message will be issued. When a user leaves the scene temporarily, a gentle reminder to suspend the TV program will be issued. We have classified the scenarios into three categories: self-initiated move by animation character (e.g. leaving the screen for a rest), SITV functions (e.g. sitting too close), and interaction initiated by the user (e.g. frequent channel switches).

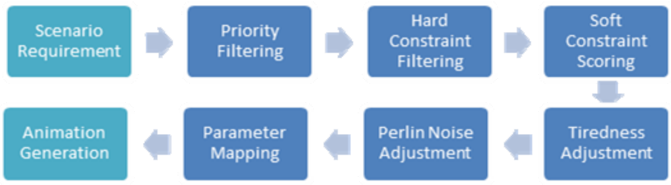

Fig. 1. Procedure for generating character animation

3.2 Generation of Character Animation

In this work, we have used the procedural animation system developed in our previous work [3] to generate character animation. The animation system takes motion descriptions from an xml file with a wide range of parameters specifying how to express different emotions with different strengths. In this system, we have implemented various motion procedures to generate common locomotion such as walking, running, and jumping. We also have implemented some customized ending motions such as waving, pushing, head shaking, pointing, listening, and boxing.

The system is able to generate expressive animation by specifying parameters at the emotion level. In our previous work [2], we proposed a hierarchical representation of motion parameters including four layers: procedure, motion, style, and emotion. Appropriate mapping between these layers through experiments has been explored. We found that the style parameters of stiff-soft, slow-fast, and expanded-extracted are more effective in expressing emotions. For example, happiness can be better expressed with softer, faster and more expanded motions; anger can be better expressed with more expanded motions; fear can be better expressed with stiffer, slower, and extracted motions. These styles can be specified with some quantitative values such that we can express the motions with finer granularity. Emotion expressiveness can also be affected by the type of motions in addition to the way it is expressed. Therefore, we have also defined the range of expressiveness of each motion for each basic emotion such as anger, fear, etc.

3.3 Animation Specification and Generation

The main procedure for generating character animation is depicted in Fig. 1. According to a given scenario, we first prioritize the tasks for the animated character and decide which action to take. The list of motions suitable for the given scenario will then be filtered according to the specified hard constraints and then scored according to the criteria of soft constraints. In order to add variations to the animation, the chosen motion with the highest score will be adjusted according to the tiredness of the character and with some random noise [9] on the parameter setting. Then the parameters specified at the emotion layer were mapped to the style layer and then to the motion and procedural layers to generate the desired animation.

(a) (b)

Fig. 2. Snapshot of SITV incorporating animated character

In the priority filtering module in Fig. 1, each motion is given a priority which is used to resolve conflicts when two requests of motions are overlapped. In the hard constraint filtering module, the available motions are filtered according to the list of hard constraints, such as required final position, desired time, and emotion, for a given scenario. We also have taken into account soft constraints, usually indicating preferences on the given scenario in the soft constraint scoring module. The tiredness adjustment module and noise adjustment module are used to add variations to a given motion to make it look more natural.

4 Experimental Designs

In order to understand how users perceive 3D intelligent characters as the interface of interactive TV, we have used the aforementioned character animation system to design two characters (normal and intelligent) in addition to the original interface based on textual dialog boxes. We will describe the three types of interfaces and our experiments on the evaluation of these user interfaces.

4.1 Description of Three User Interface Types

In the original design of SITV, dialog boxes are used to prompt messages related to the events that are happening. In this work, in addition to the dialog messages, we also have designed two types of animated character as the interface. The appearance of the two characters is the same but they differ in the emotion expressiveness of the generated motions. The three types of interfaces for comparison are as follows:

1) Interface A: dialog box with text
2) Interface B: regular animated character along with dialog box but without emotive expression (see Fig. 2a).
3) Interface C: emotive animated character, the same as Interface B, except for that it uses emotive expression (see Fig. 2b).

We have designed two experiments to study the effect of the three user interfaces on various interactive TV scenarios.

4.2 Experiment 1: Preliminary Comparative Study of User Interface

The objective of this experiment is to understand the impression of the subjects on these interfaces from an observer's point of view. We prepare nine videos for three scenarios and three interfaces in advance and show the videos to 30 subjects who were asked to provide their feedbacks by filling a survey at the end. The three scenarios that have been used in the experiment include login, getting too close, and temporarily leave and return. We have adopted the within-subjects design for the experiment. The independent variables are the three user interfaces and the three scenarios. The only dependent variable is the scores (-100 to 100) filled in by the users.

Table 1. The mean scores on five aspects for 3 (interfaces: A, B, and C)×3(scenarios) conditions in Experiment 1

	Login			Getting Too Close			Leave/ Return		
	A	B	C	A	B	C	A	B	C
Amusement	16.43	30.07	39.23	16.10	31.37	52.33	22.20	24.67	41.70
Novelty	14.37	41.77	44.40	25.13	40.10	50.30	27.00	35.40	45.67
Fondness	32.33	14.77	22.60	53.47	- 9.90	30.07	47.60	- 9.30	17.43
Understandability	56.50	41.33	36.87	64.67	47.67	58.33	60.77	40.27	53.93
Willingness to use	36.03	22.03	21.77	53.30	-19.20	27.00	47.77	-10.23	12.03

We asked the subjects questions about the following five aspects of the system: *amusement*, *novelty*, *fondness*, *understandability*, and *willingness to use*. The mean scores on these five aspects for three interfaces in each scenario are listed in Table 1. Two-way ANOVA on each of the five aspects separately shows that the main effects of interfaces are all significant ($F(2,58)$=12.64, 22.10, 17.97, 7.05, 18.75, respectively, $p<.05$). Post hoc comparisons with Tukey test show that the user interface with emotive animated character (Interface C) is significantly better than the interface with dialog box (Interface A) in amusement and novelty. However, Interface A is significantly better than Interface C in the aspects of fondness, understandability, and willingness to use. The interface with regular animated character (Interface B) has the worst overall performance. In other words, although the interface with animated character is more interesting to see, it is also less understandable compared to the dialog box interface.

4.3 Experiment 2: User Experience Assessment

With the result of the first experiment, we would like to further understand how the users assess the interfaces when they are placed in the real scene and operate the interactive TV with these interfaces. We have installed the SITV system in a living lab as shown in Fig. 3 for our second experiment. The experiments were conducted in two days separated by one week for ten subjects. In the first day, each subject was asked to operate the three interfaces in sequence and then choose any interface to watch the TV for the next 30 minutes. In the second day, the order was reversed, free watching and then exercising the three interfaces in sequence. At the end of each day, the subjects were asked to fill in a survey form and receive an interview.

Fig. 3. Living lab with the SITV system for the second experiment

Table 2. The mean scores on eight questions for the 3 (interfaces: A, B, and C)×2(day) conditions in Experiment 2

	Day 1			Day 2		
	A	B	C	A	B	C
Amusement	52.00	60.90	55.30	38.70	48.10	57.60
Novelty	48.80	61.80	57.60	43.30	48.80	52.40
Fondness	41.90	42.70	37.50	46.10	24.20	51.70
Understandability	22.30	48.70	42.80	58.40	62.90	67.10
Willingness to use	38.70	44.90	40.10	41.20	27.60	50.50
User friendliness	47.60	60.20	61.30	54.10	45.30	57.00
Good design	47.00	50.60	45.50	42.40	39.90	49.70
Desire to have	46.70	49.40	43.10	38.50	18.40	45.10

The experiment uses the within-subjects design by having the three interfaces and experimental days as the independent variables. The dependent variables are the scores (-100 ~ 100) that the users enter for the survey questions and the times for different types of interfaces. We asked the subjects to score the interfaces according to eight questions including the five in the previous experiment. The mean scores on eight questions for the three interfaces in Day 1 and Day 2 are listed in Table 2. The experimental data reveal that in Day 1, the most differences among the three interfaces are on amusement, and novelty, and understandability. The dialog box interface (Interface A) has the lowest scores in these questions. But the simple effects of ANOVA show that the effects of interfaces in Day 1 are all insignificant except for understandability ($F(2,18)=5.52$, $p<.05$). Tukey test indicates that Interface B is significantly more understandable than Interface A. Interface C is also higher than Interface A on understandability, but only marginally significant. In Day 2, we found that the differences on novelty and understandability have decreased. Simple effects of interfaces in Day 2 are significant on amusement, fondness, and desire to have ($F(2,18)=6.96, 4.37, 4.22$, respectively, $p<.05$). Tukey test indicates that Interface C is significantly superior to Interface A on amusement. Interface C is also significantly superior to Interface B on amusement and desire to have. By comparing the data in the two days, we found that the scores for Interface B have dropped significantly and becomes the worst one. In other words, once the user become familiar with the interface containing an animated character, the expressiveness of the character becomes more important. Improper or dull character behaviors will make the interface worse than a pure text-based interface after the freshness disappears.

5 Conclusions

In this work, we have used the SITV platform as a basis to develop new interactive TV functions and a new user interface with animated character. We believe that this study is the first work on realizing an emotive animated character based on the scenarios of using an interactive TV. We have defined the problem of how to generate expressive character animation for a set of given scenarios on interactive TV. The expressive animation with emotion is generated in real time by taking advantage of our previous work on procedural animation for emotive motions. A planner is used to search for the most appropriate motions in terms of mobility and expressiveness from a library of animation procedures. In order to assess the new user interface with animated character, we have designed two experiments to compare the new interface with the base-line dialog box interface. We have found that although the interface with animated character can attract the attention of the user and enhance interactivity, the text-based interface is still the easiest to understand. We also have found that when the freshness of the new interface decays after some time of use, it is crucial for an animated character to have expressive motions.

Acknowledgments. This work is supported in part by National Science Council under contract NSC-98-2221-E-004-008.

References

1. Diederiks, E.M.A.: Buddies in a Box - Animated Characters in Consumer Electronics. In: Proc. of the 8th Intl. Conf. on Intelligent User Interfaces, pp. 34–38 (2003)
2. Lin, Y.-H., Liu, C.-Y., Lee, H.-W., Huang, S.-L., Li, T.-Y.: Verification of Expressiveness of Procedural Parameters for Generating Emotional Motions. In: Proc. of the Eight Intl. Conf. on Intelligent Virtual Agents (2008)

3. Liu, P.-Y., Hsu, S.-W., Li, T.-Y., Lee, H.-W., Huang, S.L.: An Experimental Platform for Smart Interactive TV in Digital Home. In: Proc. of 2007 Symp. on Digital Life and Internet Technologies, Tainan (2007)
4. Ludwig, B., Mandl, S., Mammen, S.V.: What's on tonight - User-centered and Situation-aware Proposals for TV Programs. In: Proc. of International Conference on Intelligent User Interfaces, p. 258–260 (2006)
5. Magnenat-Thalmann, N., Kalra, P.: The Simulation of a Virtual TV Presenter. In: Proc. of Pacific Graphics 1995. World Scientific, Singapore (1995)
6. Meeren, H.K.M., van Heijnsbergen, C.C.R.J., de Gelder, B.: Rapid Perceptual Integration of Facial Expression and Emotional Body Language. Proc. of the National Academy of Sciences (2005)
7. Nijholt, A.: Where computers disappear, virtual humans appear. In: Proc. of Computers and Graphics (2004)
8. Noma, T., Zhao, L., Badler, N.I.: Design of a Virtual Human Presenter. IEEE Computer Graphics and Applications 20(4), 79–85 (2000)
9. Perlin, K., Neyret, F.: Flow Noise. SIGGRAPH Technical Sketches and App. (2001)
10. Seron, F., Baldassarri, S., Cerezo, E.: Maxineppt: using 3D virtual characters for natural interaction. In: Proc. of the 2nd Intl. Workshop on Ubiquitous Computing & Ambient Intelligence (2006)
11. TiVo, http://www.tgc-taiwan.com.tw/