UDC 001.81:003.6:681.32

Human-machine Interface Using Humanoid Cartoon Character

●Satoshi Iwata ●Takahiro Matsuda ●Takashi Morihara
(Manuscript received June 6, 1999)

This paper describes our Cartoon Character Interface (CCIF). The CCIF is a graphical user interface that features a humanoid cartoon character that acts as an anthropomorphic agent. First, we report on the effectiveness of humanoid characters in human-machine interactions and how users accept such characters. Next, we describe a simulator for creating scenes for the CCIF for use in a sales terminal and describe the results of an experiment to evaluate how the CCIF affected the terminal's operability. We confirmed that the CCIF made the interface of the terminal easier to use. Finally, we introduce a cartoon animation composer for creating CCIFs.

1. Introduction

The use of animated humanoid characters to assist in interactions between humans and machines is quite common these days. Some examples are provided below.

1) Characters on terminals
   Animated characters are used on Automated Teller Machines (ATMs) and ticket vending machines.
2) Characters on PCs
   Some mailer systems use animated characters that indicate the status of mail delivery.
3) Characters in communication networks
   Users perform tasks in a virtual world by controlling a character (called an avatar).

There have been several attempts to use computer graphics (CG) characters as anthropomorphous agents on graphical user interfaces (GUIs). Hasegawa et al. designed multi-modal systems with an anthropomorphous agent.[1] Tani et al. used a character as the navigator of a digital archive.[2] Kawabata designed "Noddy," which is an agent that can recognize the speech of a user.[3] Ebihara et al. designed the humanoid agents "Shall we dance" on network applications, in which the agent dances with a partner on the network.[4]

These examples illustrate that humanoid-type characters can play important roles in interfacing between humans and machines.

On the other hand, in the case of terminal use, there are many occasions when people need to operate terminals with which they are unfamiliar. Some of these terminals are complicated to use and require many operations. The introduction of an anthropomorphous agent on the GUI would be very useful in these cases.

Our efforts are focused on terminal applications, so we developed an animated character interface and a new GUI with cartoon characters acting as anthropomorphic agents. We also made an experimental application of the character interface and then verified the effects of these interfaces.

We then developed an animation composition tool called the Cartoon Character Interface Composer (CCIC) for the animated character interface. In this paper we describe these developments and present the results of our experiments.

2. Our approach

2.1 Cartoon characters

We applied cartoon characters that were produced by connecting a series of cartoon animation cells created in advance to the animated character interface and called it the Cartoon Character Interface (CCIF).

It is sometimes difficult for users to view computer-generated characters as anthropomorphous or autonomous agents because the machine-rendered image limits the resolution of the characters, making it difficult to design features that move naturally.

We chose cartoon characters for the following reasons.

1) Most people are familiar with cartoon characters.
2) Although it is difficult to build a cartoon animation program, once built, it is relatively easy to use.

2.2 Effects of humanoid cartoon character

We investigated the effects of using a humanoid cartoon character. We sampled facial expression clips from well-known TV animation characters to acquire an understanding of animation methods and then categorized the clips according to expression. By joining the clips together and incorporating a voice, we made several experimental applications of interfaces that use a humanoid cartoon character. The applications included an Animation Narrator who acts as a weather reporter and as a guide on a web browser and ATM display, and an Interactive Tutor who acts out a simple Q & A scenario. The voices were created using a text-to-speech system designed by our company.[5]

We then asked 10 people between 20 and 30 years old to operate the interfaces and give their subjective evaluations. The subjects all use a personal computer in their daily work.

The subjects' comments after the experiment included the following.

1) The characters of the interfaces seem to give a personality to the system being operated.
2) The characters' voices are easier to catch with mouth movements than without.
3) The characters give the impression that a real human is assisting the user, which makes the terminal's responses seem autonomous.
4) The characters convey sufficient emotion and feeling to affect the user.
5) Repetition of the same character actions becomes boring and spoils the illusion of character autonomy.
6) Installing various gestures would be an improvement.

After considering the above comments, we came to the following conclusions with regard to the use of humanoid cartoon characters for the interface.

1) The humanoid cartoon characters attract users to the applications, so they can enhance the usability of practical GUIs.
2) The users are affected by the characters' movements and expressions, so the characters can be used to convey non-verbal information.
3) Natural looking movements increase the likelihood that the characters appear autonomous.

3. Cartoon Character Interface

3.1 Sales terminal

We used the CCIF simulator to simulate a character's expressions and gestures as described above for a sales terminal. **Figure 1** shows an example of a sales terminal produced by our company that is in current use. It has a display attached to a touch panel for input, a speaker, and a printer for issuing tickets, receipts, etc. It can handle various kinds of tickets and sales by catalogue and it can convey product information and sell software directly. Using this display, a user can select and purchase products and services.

Figure 1
Sales terminal.

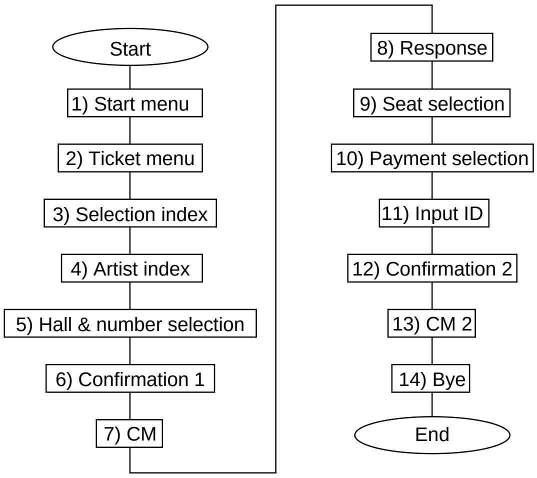
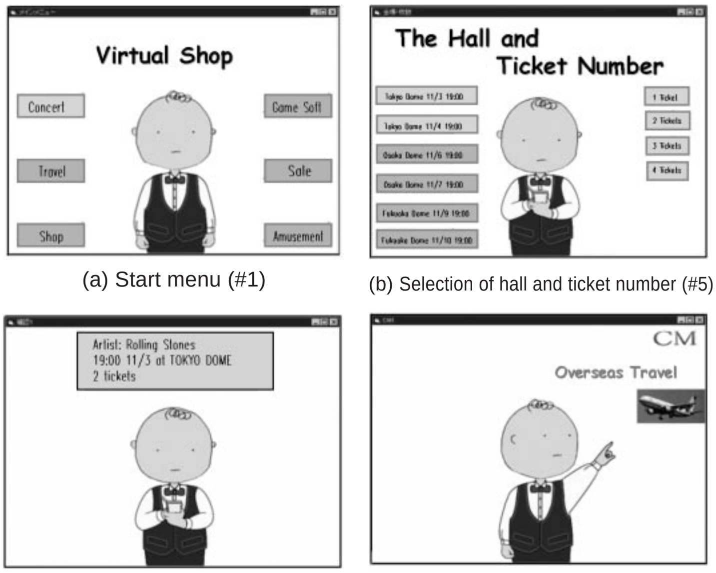


Figure 2
Example interaction sequence.



(a) Start menu (#1)
(b) Selection of hall and ticket number (#5)
(c) Confirmation (#6)
(d) CM (#7)

Figure 3
Example scenes in sequence.

Table 1
Instruction sheet.

Instructions
Please provide the following information:
Artists : Rolling Stones
Date : 11/3 or 11/4
Place : Tokyo Dome
2 Tickets
Seat : NA
Purchase by Credit Card.

3.2 CCIF simulator

3.2.1 Example interaction sequences

Figure 2 shows the sequence of interactions for the sales terminal, and **Figure 3** shows some example scenes in this sequence. The cartoon character is placed in the center of the display and explains the functions and confirms user inputs. This particular application is for concert ticket purchasing. In the sequence, users operate the sales terminal by following an instruction sheet (see **Table 1**).

The interaction flow of this terminal is as follows (the numbers correspond to the scene numbers in Figure 2).

1) Start Menu: The user selects the ticket.
2) Ticket Menu: The user selects the concert.
3) Selection Index: The user selects the Artist Index.
4) Artist Index: The user selects the "Rolling Stones."
5) Hall & Number Selection: The user selects

the hall name, date, and number of tickets.

6) to 13) The user completes the sequence.

14) Bye: The character thanks the user.

3.2.2 Gestures and facial expressions

After the sequence was determined, five categories of gestures were designed for this application. Categories 2) to 5) are gestures aimed at attracting users. In categories 2) and 5), five actions are displayed randomly to avoid repetition.

The five categories are as follows.

1) Pointing
Pointing gestures are intended to direct the user's attention to a button or space on the display.

2) Gestures made during speech
Hand movements improve communication. For a speaking character, five types of hand gestures are used randomly.

3) Operation
When the character is waiting for an input from the user, he brings out a pen and note pad. After the user performs the input, the character makes a note of the input and replaces the pen and note pad.
When the user inquires about seat availability, the character inputs the seat number on the display using the keyboard.
When the user inputs his or her ID and password, the character touches the touch panel on the display.

4) Emotional expression
If a seat is sold out, the character takes on an apologetic expression.

5) Unconscious movements
When the character is waiting for a user input, he makes various facial expressions and body movements unconsciously. These movements maintain the sense of autonomy.

3.2.3 Expression units

We designed various expression units that show mouth movements, body movements, neck movements, eye movements, and blinking. To connect two expression units, the end of one unit is joined to the beginning of the next unit via intermediate units which provide a smooth transition of the character's gesture and facial expression. **Figure 4** shows three of the expression units: bowing, speaking, and pointing to the upper right of

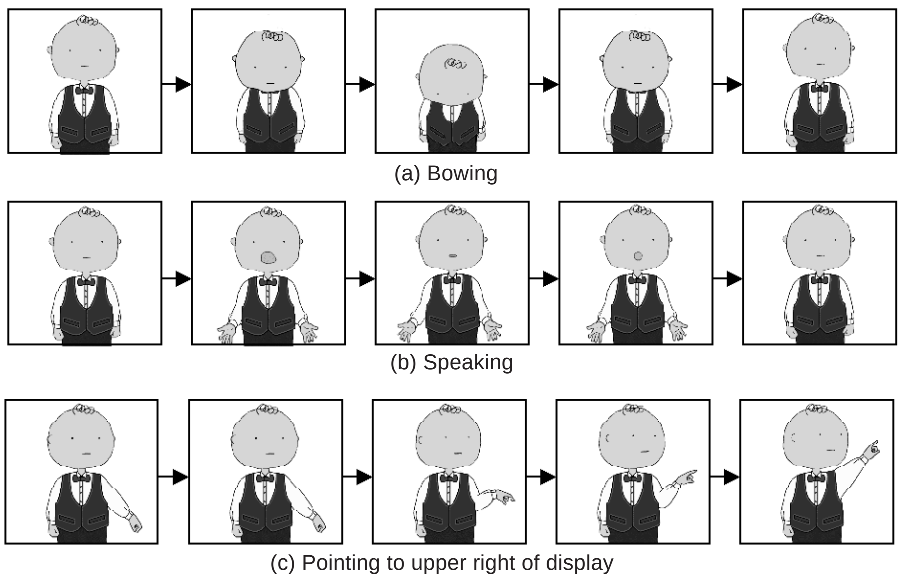


Figure 4
Example expression units.

the display. There are a total of 61 units. The scenes shown in Figure 3 were made by connecting various expression units with other graphics units and were linked with voice data.

3.3 Experiment using CCIF

We conducted an experiment using a display terminal installed with the example sequence.

The experiment is outlined below.

1) The system has 14 scenes (greeting, purchasing tickets, commercials, etc.).
2) The experiment consisted of two trials. In one trial, the character gave guidance during the scenes and explained the buttons by using his voice and making gestures.
   In the other trial, the character was not displayed on the screen and the user was given voice guidance only.
3) Nine of the scenes (Nos. 1 to 5 and 8 to 11) have menu buttons. These scenes are changed to the next scene by touching one or more buttons. The user can exit the guidance or explanation by touching these buttons. In this experiment, there were no branches, so the user could select only one button at a time. The other buttons were dummy buttons that had no responses.
4) The scenes do not change until the user selects the correct choice as indicated on an instruction sheet. If the user did not understand a scene, the user could not proceed to the next scene.
5) In the other five scenes, there were no buttons and the next scene was started automatically after the guidance had finished.

The 10 subjects involved in this experiment were between 20 and 30 years old and worked in the same laboratory. They attempted to follow the same operation sequence. The experiment was intended to make a comparison between a voice-only interface and the CCIF. We divided the subjects into two groups (see **Table 2**).

After the experiment, we investigated two points.

1) Scene display time
   This is the total amount of time during which a scene is displayed. This time is a measure of the user's performance in each scene. It varies with the complexity of the scene and the number of choices the user can make. Some scenes are intuitively understood after viewing them for only a short time.
2) Post-experiment interview
   We asked the users whether they thought the character made the sequence easier to follow and asked them for their subjective feelings about the character.

3.4 Experimental results

3.4.1 Evaluation of display times

Figure 5 shows the experimental results. The x-axis indicates the scene number and the y-axis indicates the average display time of each scene.

Figure 5 (a) shows the results of the first trials for group A and group B. In the 1st scene, the display time with the character displayed is twice as long as the display time without the character. This gap decreased gradually as the sequence moved to the 4th scene, which indicates that the users were interested and listening attentively to the character speaking. Then, after the users learnt the system operations scene by scene, the listening time decreased.

Figures 5 (c) and (d) show the results for both trials by group A and group B, respectively. In these figures, the results for the 5th scene are particularly interesting. The 5th scene is a complex scene in which users have many choices to make, for example, they must select a hall and ticket number. Group A's average time in their

Table 2
Experiment.

	1st trial	2nd trial
Group A	Instructions by voice and the character	Instructions by voice only
Group B	Instructions by voice only	Instructions by voice and the character

first trial on this scene, which was done with the assistance of the voice and the character, was almost the same as their average time in the second trial (voice only). However, group B were only given voice guidance in their first trial and then made a 6-second improvement when the character was added in the second trial. We attribute this 6-second improvement not only to the help the character gave to speed up the performance of the slower users but also to the braking effect it had on users who tended to make mistakes because they proceeded too quickly.

These results indicate that the character helped the users understand each scene and operate the terminal correctly.

3.4.2 Interview results

The users confirmed the effectiveness of the CCIF by making the following observations.

1) Intuitive interaction
   It is possible to understand the operations intuitively by watching the character performing them on the display.
2) Confirmation
   The user can visually check the information he or she inputs.
3) Voice assistance
   The character's mouth movements make the voice assistance easier to understand.
4) Attraction
   The users were attracted to the character and felt comfortable with the terminals.

Observations 1) and 3) suggest the potential of intuitive communication, which reduces the operation time of interfaces. Observation 2) suggests that the non-verbal communication given by the CCIF was effective.

4. Cartoon character interface composer

4.1 Configuration of composer

We designed an animation composer to help us compose the cartoon character animations for the CCIF. Using the composer, the designer of a terminal sequence can create cartoon character animations with no professional experience.

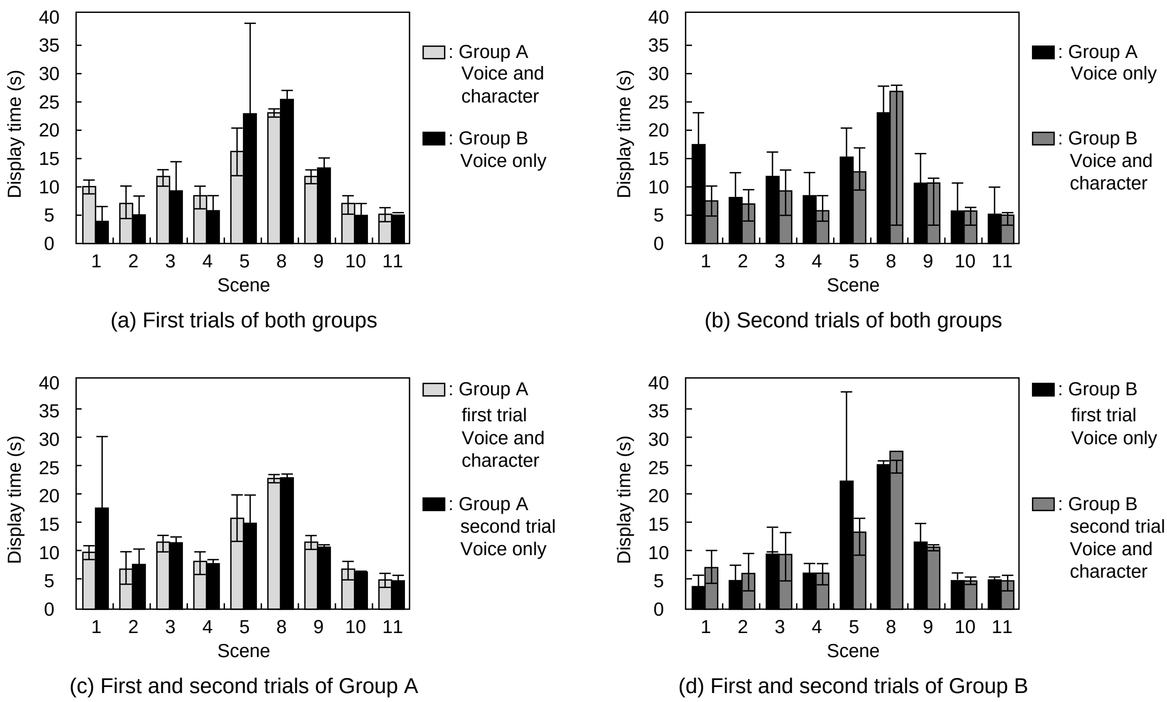


Figure 5
Evaluation results.

Figure 6 shows an outline of the composer. The composer makes character animation a resource of the Visual Basic developer tool. The principal elements are the animated character database, animated character engine, and composer GUI.

4.2 Animated character database

This database contains the properties of the expression units for the gestures and facial expressions of the character. The properties are as follows:

1) Unit ID
   Identifies an expression unit.
2) Repetition/talking
   Indicates talking or the repetition of a movement (e.g., operating the keyboard, writing with a pencil).
3) Frequency of use
   Used to create "natural behavior" and to avoid repetition.
4) Number of contiguous cells
   Indicates the number of image cells in an expression unit.
5) Playback order
   Indicates the order in which the image cells are displayed.
6) Principal cell ID
   Indicates the ID of a cell in the expression unit that has been chosen as the key image of the unit. The key images are used to retrieve the expression units.

We will add emotional information in the future.

4.3 Animated character engine

This engine is an executable component that is used to make the character's movements appear natural. It automatically gives the animation a "natural behavior" and is designed for application by the composer and the terminal.

The principal functions are as follows:

1) Connecting
   This function checks the IDs of two connecting expression units and inserts the correct intermediate unit for a natural-looking transition between them.
2) Talking
   This function inserts the appropriate number of speaking units by counting the words. (A speaking unit consists of image cells depicting two mouth movements.)
3) Idling
   When the terminal is waiting for a user input, this function selects parts automatically and provides animation. It prevents the display from remaining static and improves user-understanding.

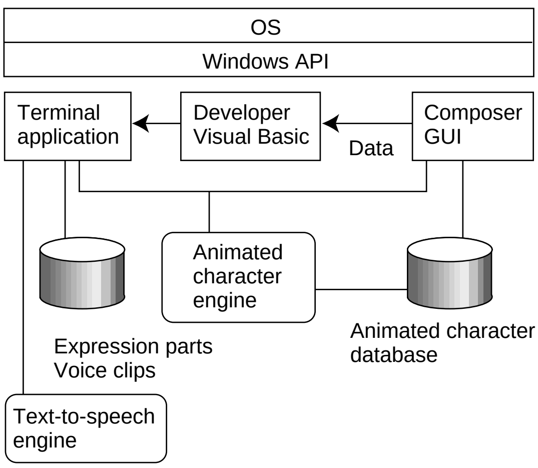


Figure 6
Configuration of character composer.

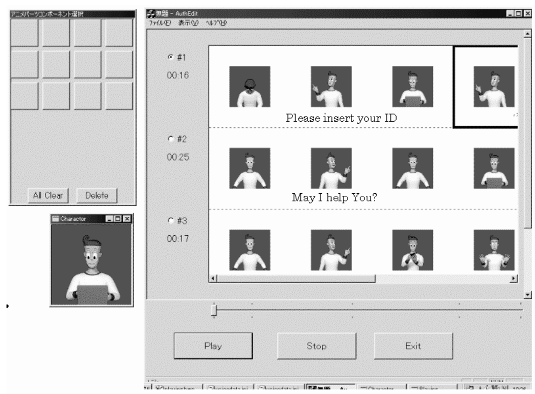


Figure 7
Character composer.

4.4 Graphic user interface of composer

Figure 7 shows the GUI of our composer.

The terminal designer refers to the animated character database to select sequences of expression units called animation parts and then places them at the appropriate spaces in the time chart on the composer GUI. There is one animation part for each set of communications between the character and the user. Some examples of animation parts are a greeting, a request for input, an instruction, and a confirmation.

Then, the animated character engine determines which animation parts should be connected together and makes the transition between them appear natural.

If the designer wants to use voice clips, they can be attached as text to the animation parts on the composer GUI. Then, the animated character engine selects the appropriate expression units for speaking.

The composed data is used to develop the scene. The animation parts described in the composed data are played back in the target application on the terminals.

This composer reduced the time required to create the CCIF described in Chapter 3 by about 99%.

5. Summary

Humanoid cartoon characters are being used more frequently. We studied the use of such characters for human-machine interface applications. First, we incorporated facial expression clips from well-known animations. We then made experimental applications using these clips and completed a subjective evaluation.

We found that users accept animated characters because they are familiar to them. We also found that an animated character can communicate non-verbally and that natural movement is necessary for users to view the character's behavior as autonomous.

After considering the above results, we designed a simulator for sales terminals so that the Cartoon Character Interface could be put to practical use. We conducted an experiment and completed an evaluation. We found that the CCIF made an existing sales terminal interface easier to use.

Finally, we introduced a cartoon animation composer we developed for producing CCIFs.

Acknowledgments

Some parts of this work were carried out in the Fujitsu Design Laboratory. The authors thank Mr. Kakiuchi and Mr. Watanabe for their contributions to this work and for the fruitful discussions we had with them.

References

1) O. Hasegawa: Computer Vision Inspired by Biological Visual Systems. (in Japanese), *Journal of the IPSJ* (The Information Processing Society of Japan), **39**, 2, pp.133-138 (1998).
2) M. Tani, T. Kamiya, and S. Ichikawa: User Interfaces for Information Strolling in a Digital Library. *NEC Research & Development*, **37**, 4, pp.535-545 (1996).
3) T. Kawabata: Spoken dialog system "Noddy." (in Japanese), *NTT R&D*, **47**, 4, pp.405-410 (1998).
4) K. Ebihara, L. S. Davis, J. Kurumisawa, T. Horprasert, R. I. Haritaoglu, T. Sakaguchi, and J. Ohya: Shall we dance? – Real time 3D control of a CG puppet –. SIGGRAPH'98 Enhanced Realities, conference abstracts and applications, July 1998, p.124.
5) N. Katae and S. Kimura: Natural Prosody Generation for Domain Specific Text-to-speech Systems. ICSLP'96, October 1996.

Satoshi Iwata received the B.S. and M.S. degrees in Electrical Engineering from Keio University, Yokohama, Japan in 1983 and 1985, respectively.
He joined Fujitsu Laboratories Ltd., Atsugi, Japan in 1985 and has been engaged in research and development of imaging technology for peripheral systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan and a member of the Institute of Electrical and Electronics Engineers (IEEE).

E-mail: iwata@flab.fujitsu.co.jp

Takahiro Matsuda graduated in Electrical Engineering from Kurume College of Technology, Kurume, Japan in 1991. He joined Fujitsu Laboratories Ltd. in 1991 and has been engaged in research and development of peripheral systems.

E-mail: tmatsuda@flab.fujitsu.co.jp

Takashi Morihara received the B.E. degree in Electronic Engineering from Iwate University, Morioka, Japan in 1980. He joined Fujitsu Laboratories Ltd. in 1980 and has been engaged in research and development of peripheral systems.

E-mail: morihara@flab.fujitsu.co.jp