

## Emotion recognition-based Interactive Theater -Romeo & Juliet in Hades-

Naoko Tosa, Ryohei Nakatsu

ATR Media Integration & Communication Research Laboratories  
2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan  
tosa,nakatsu@mic.atr.co.jp

### Abstract

Now a days many Interactive Theater research is appeared[1][2][3][4]. We have developed an emotion-based interactive theater system. The characteristic is the use of Interactive Computer Graphics technology and the generation of three-dimensional imagery to create autonomous actors and a cyberspace in which all participants obtain a feeling of immersion. We have developed a system that includes multi person participation, emotion recognition and gesture recognition through. In this system, participants can feel they are actually contributing to the development of the story in cyberspace. This is done a system that shows avatars on the screen as the alter ego of the participants. We have also produced an interactive story based on this system.

**keywords:** interactive cinema, emotion recognition, art & technology

### 1. Introduction

We selected "Romeo and Juliet" by Shakespeare as the base story. The main plot of the story is as follows. After their tragic suicide, the souls of Romeo and Juliet are sent to Hades, where they have no recollection of the past. Then, Romeo and Juliet each start on a journey to rediscover themselves and what relationship they shared with their characters, who are emotion-based autonomous actors.

### 2. Software that Realizes Conscious / Unconscious Interactions

(1) Participation in cyberspace

a) System for multiple players : Our initial efforts in making a system for multiple players allowed two players to participate in cyberspace in the development of a story. The ultimate goal was to create a multi-player system operating across a network. However the first step for the present study is to develop a prototype multi-player system that had two systems connected by a LAN.(Figure 6)

b) Avatar representation :

We used a system that showed avatars on screen as alter egos of the players(Figure 1). There were several

advantages to this system, as outlined below.

b-1)System for representing the avatar: The relationship of the player to the avatar and the relationship of the avatar to other characters in the movie can be controlled in various ways by changing the representational form of the avatar.

b-2)System for controlling the avatar:

The basic control system inputs player movement by using magnetic sensors and uses that movement to map avatar movement with a motion capture system.(Figure 1) Giving autonomy to avatar movement enables a complex movement that combines autonomous avatar movement with player movement. By varying the proportion of each movement based on time and circumstances, player movements can be used directly. In addition, the player can introduce desired movements that add diversity and depth to the relationship with cyberspace.



**Figure:1** Juliet's avatar and Romeo's avatar and People play Romeo by magnetic sensor

(2) Interaction

a) Introduction of interaction at any time:

To increase the frequency of interaction between the participants and the system, we devised a way for players to interact with cyberspace residents at any point in time. Basically, these impromptu interactions, called story unconscious interaction (SUI), occur between the players and the characters and generally do not affect story development. On the other hand, there are sometimes interactions that can affect the story development. These kind of interactions, called story conscious interactions (SCI), occur at branch points in the story. The results of such interactions determine the future outcome of the story.

b) Introduction of multimodal interaction :

The following interactive functions were added to the primary interactive function of voice recognition.

b-1) Emotion recognition: To make interaction at any time possible, an emotion recognition capability was introduced (Figure 3). When players utter spontaneous utterances, the characters in the story react with their own utterances and animations according to the emotion recognition results. Emotion recognition is achieved by using a neural-network-based algorithm. [5] People use a head-set microphone when communicating with character. For example, if one makes a happy high voice, character's feeling will be positive and he responds with excitement. If the speaker's voice is low and strong, character's feeling will be bad and he gets angry.

b-2) Motion capture: We used a motion capture system that is based on magnetic sensors attached to applicable parts of the player's body. This system allows avatar movement to reflect player movement. Data from the magnetic sensors are input into the system to move the computer graphic avatar. In this way, players get the feeling that they are controlling the movements of their avatar. This control can provide another form of interaction at any time.

b-3) Gesture recognition: We captured motion from the players with magnetic sensors and used an HMM to process the data from the sensors. This was done so that 3-D gestures, minute gestures and gestures under low-light conditions could be recognized. Ges-

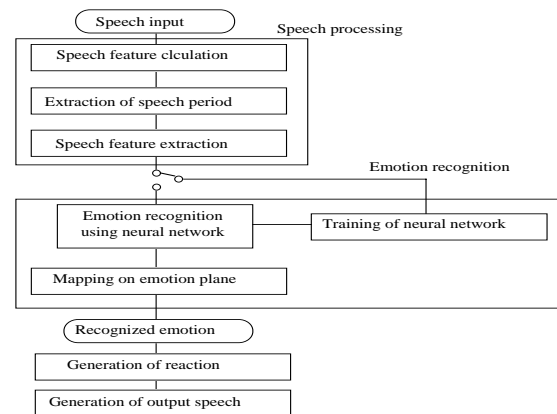


Figure 3: Emotion recognition diagram

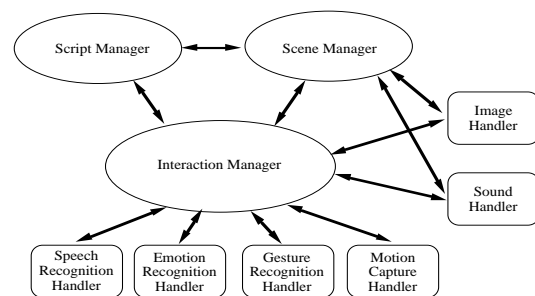


Figure 4: Software configuration of the system.

ture recognition results are used for SCI. Figure 4 shows an overview of the system, which was equipped with these functions.

3. Generation of Empathy by Interactive Story

The scenario in this study is devised by the people playing Romeo and Juliet, who select lines depending on their mood. How the drama unfolds depends on this selection process. If the people playing the roles of Romeo and Juliet speak words of anger or of happiness, depending on the context of the play, autonomous characters recognize the emotions from the tone of voice and express their emotions reactions by speech and animations. Emotional recognition technology allows a neural network to learn emotional speech, and through this learning, the network can create personalities, such as an angry character or a cheerful character.

Voice recognition is the function whereby the lines chosen by a player to go with the scene are said, and the autonomous character recognizes the meaning of these lines and reacts to these lines by their utterance

and gestures.[6]

Gesture recognition is the reaction of the character to action, contact or a pose in any location. Take for example if the impassioned Romeo was to try to kill his former friend, Macutio, with a pistol. Macutio would recognize this behavior, escape and condemn Romeo.

\* Follow interaction: This function gives the characters in the movie system to follow or accompany the avatar of a player in an appropriate situation. For example, if the player who play the role of heroin Juliet is in a serious situation, an autonomous agent Shin, an angel, recognize her mental situation by emotion recognition, accompany, and take care of her.

\* Background speech interaction: This function allows you to talk from the background to a character about your feelings and doubts concerning the performance running in the foreground.



Figure:5 One of Scene

\* Touch interaction: This function is used when, for example, the lead actor Romeo becomes very angry and hits his old enemy Paris. Paris then pulls back in fear. Figure 5 shows one of these scenes. In this scene, Paris who was once a Juliet's fiance is being killed by his enemy. Romeo & Juliet try to help him using voice and gesture recognition.

#### 4. Hardware system structure

Figure 6 shows the system's hardware structure, which is comprised of sub systems for image output, voice and emotion recognition, gesture recognition and sound output. Two workstations (Onyx Infinite Reality and Indigo 2 Impact) capable of generating computer graphics at a high speed are used to output the images. The Onyx is used to run the script manager, scene manager, interaction manager and all the image output software. The character images are stored on the workstations ahead of time in the form of computer graphic animation data so that computer graphics can be generated in real time. The background computer graphic images are also stored as digital data so that background images can also be generated in real time. The background images are real photographic images stored on an external laser disc. The multiple character computer graphics, background computer graphics and background photographic images are processed simultaneously through video boards on eoscopic vision control and projected onto a curved screen with two projectors.

#### 5. Results

Our system/work, "Romeo & Juliet in Hades" was selected for exhibition at Transmediale'99 which is a new media division of International Berlin Film Festival 1999. We have developed both Japanese and English versions. Of them we demonstrated the English version at the exhibition. More than five hundred people from various countries visited our site and about two hundreds of them played Romeo or Juliet by themselves. We asked them answer the following questionnaires.

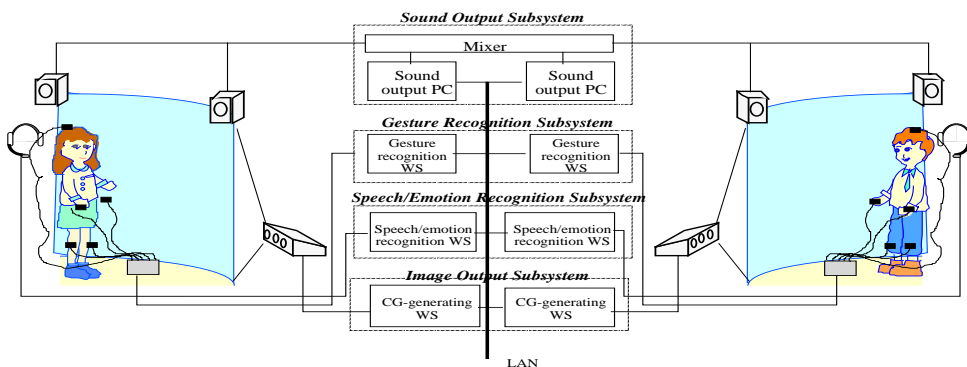


Figure: 6 Hardware configuration of the interactive theater system

- 1.How did you feel while playing Romeo or Juliet by yourself?
- 2.What is necessary for the system to have if you want to experiencedeeper empathy?
- 3.Others.

For the first question, almost people told that the experience was new and interesting. we were successful Romeo & Juliet in Hades.

For the second question certain amount of people said that they sometimes felt some difficulties for controlling their avatar and following the story.

And they indicated that it is necessary to introduce easier human interface. In our system, because people are both observers and players, they are a littlebit busy for watching the story on-going and playing their roles at the same time.At the same time young people felt no difficulties, because they are used to have the same kind of experience by playing video games. If we want this system used both by young and senior people, further research will be necessary to introduce more human-like interface.

Third question is some people want more touching feeling and pain feeling in the story. Also, people want more communicate with Romeo & Juliet each other.

## 6. Conclusion

It needed about three years for the construction of our Interactive Theater system. In the first year we mainly worked for the development of the authoring software for interactive narratives. In the second year we developed the emotion recognition software and combined it with the authoring system. Also a commercially available speech recognition software was combined with the software and the first prototype system was developed. In the third year, finally, we have developed the any time interaction algorithm, that is a key function for generating empathy, and added the function to the system. Also several improvement have been done. First, interactive story was revised to fully utilize the any time interaction capabilities. Vibration function was introduced to give participants sense of pains. Also other virtual reality technologies were introduced to give people the feeling of presence.

There are several topics that need further research. One of the most important issue is how to describe interactive situations in order to improve the real-to-life feel of each scene[8]. Here, a realism, that blurs the line between what is real and what is not, is not the key issue. Instead, what is important is the generation of unknown pleasant imagination that lays between people by utilizing interactive technology.

## Reference

1. Joseph Bates et al., "An Architecture for Action, Emotion, and Social Behavior," Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (1992).
2. Ken Perlin, "Real Time Responsive Animation with Personality," IEEE Trans. on Visualization and Computer Graphics, Vol.1, No.1, pp.5-15 (1995).
3. P. Doyle, B. Hayes-Roth, "Computer-aided exploration of virtual environments,"Proceedings of AAAI Workshop on AI and Entertainment, AAAI Press, Menlo Park, CA, (1996).
4. Pattie Maes et al., "The ALIVE system: Full-body Interaction with Autonomous Agents," Proceedings of the Computer Animation '95 Conference (1995).
5. Naoko Tosa and Ryohei Nakatsu, "Life-like Communication Agent - Emotion Sensing Character 'MIC' and Feeling Session Character 'MUSE' -," Proceedings of the International Conference on Multimedia Computing and Systems, pp.12-19 (1996).
6. N.Tosa, R.Nakatsu, "Interactive Poem System," The 6th ACM International Multimedia Conference Proc. pp.115-118 (1998).
7. N.Tosa, "Creating a Movie with Autonomous Actor for That Respond to Emotions" SIGGRAPH97Course Note title is "Interactive Movie, -Techniques, Technology, and Contents" (1997).
8. Oren Jacob, "Computer Graphics Story - A Personal Overview of Computer Animation in the Movies," ACM Computer Graphics, Vol. 31, No. 1, pp. 26-28 (1997).