Affective Communication System with Multimodality for the Humanoid Robot AMI

Hye-Won Jung[‡], Yong-Ho Seo[†], M. Sahngwon Ryoo[†], Hyun S. Yang[†]

Artificial Intelligence and Media Laboratory, Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, 373-1 Guseong-dong, Yuseong-gu, Daejeon
305-701 Republic of Korea[†]
{yhseo, hsyang}@paradise.kaist.ac.kr
http://mind.kaist.ac.kr

M1 Group, Mobile Multimedia Lab. LG Electronics Institute of Technology
16 Woomyon-dong, Seocho-gu, Seoul, Republic of Korea[‡]
hyewonj@lge.com

Abstract. Nonverbal communication is vital in human interaction. To interact sociably with a human, a robot must recognize and express emotions like a human. It must also speak and determine its autonomous behavior while considering the emotional status of the human. We present an affective human-robot communication system for a humanoid robot, AMI, which we designed to communicate multimodally with a human through dialogue. AMI communicates with humans by understanding and expressing nonverbal communication through channels such as facial expressions, voice, gestures and posture. Interaction between a human and a robot is made possible through our affective communication framework. The framework enables a robot to catch the emotional status of the user and to respond appropriately. As a result, the robot can engage in a natural dialogue with a human. It chooses appropriate conversation topics and behaves appropriately in response to human emotions. Moreover, because the user perceives the robot to be more human-like and friendly, the interaction between the robot and human is enhanced.

Keywords: Human-Robot Interaction; Sociable Robot; Affective Communication; Multimodal Interaction.

1. Introduction

Many researchers in robotics have been exploring affective interaction skills between humans and robots. Some remarkable studies on robots and agents have focused on emotion-based communication with humans. A sociable robot called Kismet conveys intentionality through facial expression and engages in infant-like interaction with a human caregiver [1]. AIBO, an entertainment robot, behaves like a friendly, life-like dog [2]. Cat Robot was developed to investigate the emotional behavior of physical interaction between a cat and a human [3].

Due to the limitations of recognition technology, most researchers in robotics have focused on developing passive robots that interact mainly by responding to users. We have adopted a new approach to building a sociable robot that interacts with humans by spontaneously leading a multimodal dialogue. Our research enables high-level dialogue similar to the conversational robot Mel, who illustrates the first attempts to have a robot lead a conversation [4]. However, Mel paid no attention to the emotions and multimodality of human-robot communication

We have therefore developed an affective human-robot communication system for a humanoid robot named AMI [7]. The system enables AMI to achieve high-level communication by leading a conversation. To make a robot lead interactions, we considered the following objectives. First, the robot must have a social personality to induce interactions, and it must first approach people to initiate interaction. Secondly, to ensure robust emotional interactions, the robot has to change the conversational topic on the basis of its own emotions as well as the user's emotions. Thirdly, the robot must store memories of previous interactions with people in order to lead the conversation more naturally with respect to previous events and emotions. Lastly, the robot must continue guessing the user's response on the basis of the context. Furthermore, the robot must not only lead interactions, but also communicate with people multimodally.

Accordingly, we designed and implemented an affective communication framework to enable AMI to successfully accomplish these objectives. The system comprises of five subsystems related to perception, motivation, memory, behavior, and expression. In the perception system, we implemented a bimodal emotion recognizer for recognizing emotions. We designed the other subsystems to enable the robot to use its own drive, emotions, and memory to determine the appropriate behavioral response to the user's emotional status.

This paper is organized as follows. In section 2, we give an overview of the system. In sections 3 to 7, we discuss the implementation of each subsystem. In section 8, we present the experimental results and conclusions.

2. System Overview

2.1. Design Concept

The affective communication system was designed to achieve affective communication with multimodality. Multimodality refers to multimodal communication channels such as dialogue, facial expressions, voice and gestures. The main goal of this system was to make the robot proficient at understanding the emotional expressions of a human partner and at transferring emotions back to the

human partner, by giving it a variety of social skills that foster communicative behavior with humans.
AMI's main tasks for leading affective communication with a human are as follows:

- To determine the most appropriate behavioral response to the human partner's emotions and to behave autonomously
- To recognize human emotions and to express emotions through multimodal channels
- To synthesize its emotions as an artificial life.

2.2. Affective Communication Framework

For a robot to communicate affectively, we designed an affective communication framework that includes the five subsystems shown in Fig. 1. Our framework is based on the creature kernel framework for synthetic characters [5] and the framework that was applied to the software architecture of Kismet [1].

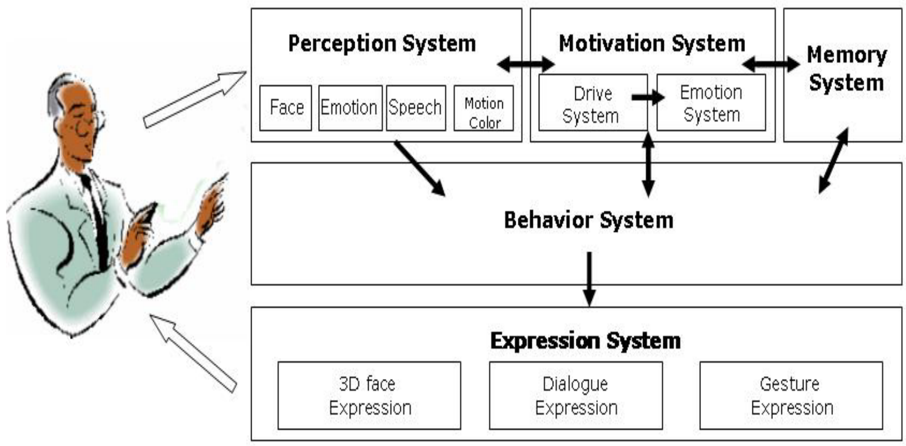


Fig. 1. Configuration of the proposed telepresence system

The arrows in Fig. 1 represent the information flow and influence of the subsystems. Well-coordinated communication of these subsystems is required for our robot to successfully function in a dynamic world. Our affective communication framework, however, has the following notable differences:

1) The internal design and implementation of each subsystem. Kismet was based on interaction with an infant-caretaker so that it could not talk with a human. Our system is mainly based on affective dialogue. Accordingly, the internal design and implementation differ because of the distinct goal of multimodal affective communication.

2) Memory system. We added a memory system to the referred framework. The memory system enables the robot to represent itself and to reflect upon itself and its human partners. Memory develops during the lifetime of a human being and is socially constructed through interaction with others. We consequently designed a memory system to enhance the robot's sociable ability and to foster communication with a human.

The main functions of each subsystem are summarized as follows. The perception system mainly extracts information about the outside world. In our system, it detects whether the human is there, who the human is, the status of the human's emotions, and what the human says. The perception system is composed of other subsystems such as face detection, face recognition, emotion recognition, and motion and color detection.

The motivation system is composed of a drive and emotion system. Drives are motivators; they include endogenous drives and externally induced desires. The robot has three basic drives for accomplishing its tasks and objectives: the drive to interact with a human, the drive to ingratiate itself with a human and the drive to maintain its own well-being. The emotion synthesis system produces the robot's artificial emotions. We modeled three emotions to give the robot synthetic analogues of anger, joy, and sorrow.

The memory system stores the most frequently occurring emotion of the latest interaction with the user, thereby influencing the robot's initial emotional status when meeting the user again. The behavior system then selects the action; that is, it chooses the most relevant behavior in response to the given perception, motivation and memory input. Lastly, the expression system plays expressions composed of 3-D facial the expressions, dialogue, and gestures according to the results of the behavior system.

3. Perception System

The perception system comprises face detection, face recognition, speech recognition, emotion recognition, and visual attention detection that enables AMI to detect objects and color. Accordingly, through the perception system, AMI can learn basic knowledge such as the identity of the human user, the user's emotional state, and what the user is saying and doing. Figure 2 shows the overall process of the perception system.

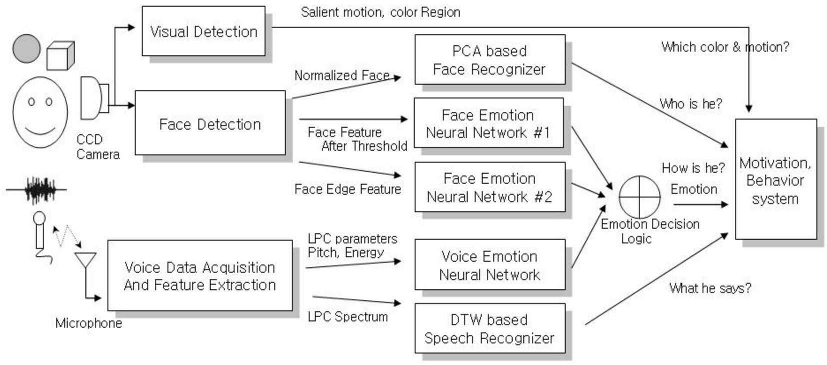


Fig. 2. Architecture of the perception system

3.1. Face Detection and Recognition

The face detection system finds human faces in an image from CCD cameras. For robust and efficient face detection, the face detection system used a bottom-up, feature-based approach. The system searches the image for a set of facial features such as color and shape, and groups them into face candidates based on the geometric relationship of the facial features. Finally, the system decides whether the candidate region is a face by locating eyes in the eye region of a candidate's face. The detected facial image is sent to the face recognizer and to the emotion recognizer.
The face recognizer determines the user's identity from the face database. To implement the system, we used an unsupervised PCA-based face classifier commonly used in face recognition.

3.2. Bimodal Emotion Recognition

1) Emotion recognition through facial expression
For emotion recognition through facial expression, we first normalized the image captured by the CCD camera. We then extracted from the normalized image the following two features, which are based on Ekman's facial expression features [8]:

- Facial image of lips, brow and forehead. After applying histogram equalization and the threshold of the standard distribution of bright, we extracted the parts of lips, brow and forehead from the entire image.
- Edge image of lips, brow and forehead. After applying histogram equalization, we extracted the edges around the regions of the lips, brow and forehead.

2) Emotion recognition through speech

For emotion recognition, we adopted a recognition method similar to the one used in life-like the communication agents called MUSE and MIC [6]. Of the system's two features, one is a phonetic feature and the other is a prosodic feature. We trained each feature vector using a neural network for the three emotions of happiness, sadness and anger. We chose these three emotions as classifiers because they are dominant in social interaction and, in the experimental results of systems that recognize emotions through speech, they have a higher recognition rate than other emotions such as surprise, fear and disgust.

3) Bimodal emotion recognition

For bimodal emotion recognition, we integrated the training results of two emotion recognizers by facial expression and one emotion recognizer by speech. To integrate these recognizers, we used decision logic to catch the user's emotion. The decision logic determines whether happiness, sadness or anger is the final emotion.

We calculated the final result vector of the decision logic (Rfinal) as follows:

$$R_{final} = (R_{face}W_f + R_{speech}W_s) + R_{final-1} - \delta t \qquad (1)$$

where Rface and Rspeech are the results vector of the emotion recognition through facial expression and speech, Wf and Ws are the weights of the two modalities, Rfinal-1 is the previous emotion result determined by decision logic, and δ is a decay term that eventually restores the emotional status to neutral. Consequently, the decision logic saves the human's emotional status in the final result vector, Rfinal. We made the final decision logic conduct a weighted summation. Accordingly, we optimized the weight valuables Wf and Ws by experiments. The overall bimodal emotion system yielded approximately 80 percent for each of five testers. By resolving some confusion, it achieved a better performance than facial-only and speech-only systems.

4. Motivation System

The motivation system sets up AMI's nature by defining its "needs" and influencing how and when it acts to satisfy them. The robot is designed to affectively communicate with humans and ultimately to ingratiate itself with them. The motivation system consists of two related subsystems, one that implements drives and a second that implements emotions. Each subsystem serves as a regulatory function that enables AMI to maintain its "well-being".

4.1. Drive System

For the drive system, we defined three basic drives for the system objective of communicating affectively with humans: the drive to interact with a human, the drive to ingratiate itself with a human, and the drive to maintain its well-being. In the current implementation, the three drives are as follows:

- *To interact with humans.* This drive motivates the robot to find, approach and greet a human. If the robot cannot find and greet a human through face detection and recognition of the perception system, the robot's activation intensity increases.
- *To ingratiate itself with humans.* This drive prompts the robot to make the human feel better. When the robot interacts with a human, it tries to ingratiate itself while considering the person's emotional state. When the person feels sorrowful, the robot attempts to console the person; when the person feels surprise or anger, the robot attempts to pacify the person. If the intensity of the human emotion that is recognized through the perception system is over a predefined threshold, the robot's activation intensity increases.
- *To maintain its well-being:* The third drive is related to the robot's maintenance of its own well-being with regard to psychological and physical fatigue. In the first case, when the robot has extreme anger or sadness, it stops interacting with the human; in the second case, when the robot's battery is too low to act any more, it takes a rest to recharge its battery. The psychological and physical condition is expressed as a value of robot energy. If the robot energy is too low, the robot's activation intensity increases.

4.2 Emotion Synthesis System

Emotions play a significant role in human behavior, communication and interaction [10]. Accordingly, the robot's emotions are important in our system. The robot expresses its emotional status by 3-D facial expressions, speech and gestures. The synthesized emotion then influences the behavior system and the drive system as a control mechanism.
To synthesize AMI's emotions of happiness, sadness and anger, we used the emotion model of the three dimensions of emotion [9]. This model characterizes emotions in terms of stance (open/close), valence (negative/positive) and arousal (low/high), thereby allowing the robot to derive emotions from physiological variables. Our system relies on an open stance because AMI is motivated to be openly involved in interaction with humans.

- *Arousal factor (CurrentUserArousal).* The "CurrentUserArousal" factor is determined by the human and the human's responses, and by factors such as whether AMI finds the human, and whether the human responds. If AMI fails to find the human, the arousal decreases. When AMI finds the human and asks the human something, the arousal decreases if the human says nothing to the robot for a long time. Low arousal increases the emotion of sadness. High arousal increases the emotions of happiness and anger by determining whether the human's response is positive or negative.
- *Valence factor (CurrentUserResponse):* The "CurrentUserResponse" factor is determined by whether the human responds appropriately to the robot's requests. When AMI waits for a yes or no answer, if the human says something unexpected that AMI can't understand, the user responds negatively.

A negative response increases the emotion of anger; a positive response increases the emotion of happiness.

Next, the synthesized emotion is influenced by the drive and memory system. In the drive system, the intensity of the unsatisfied drive increases the emotion of sadness. The emotion of joy, on the other hand, increases when the unsatisfied drive is activated. In the memory system, when the robot first meets the person, the robot's emotion is initiated by the most recently saved emotion of the person.
In summary, we used the following equation to compute the robot's emotional status (Ei(t)):

$$\text{If } t = 0,\ E_i(t) = M_i \quad (\ t = 0, \text{ when new face appeared }\ )$$

$$\text{If } t \neq 0,\ E_i(t) = A_i(t) + E_i(t-1) + \sum D_i(t) - \delta t\,, \qquad (2)$$

where Ei is the robot's emotional status; t is time; i is happiness, sadness, or anger; Ai is the emotional status calculated by the mapping function of [A, V, S] from the current activated behavior; Di is the emotional status defined by the activation and intensity of unsatisfied drives in the drive system; Mi is the emotional status of the human recorded in the memory system; and δt is a decay term that eventually restores the emotional status to neutral.

5. Memory System

Our system presents the memories required for more natural and intelligent affective communication with a human. The memories are as follows:

1) The robot's emotional response to the user. Humans often have feelings for those with whom they have communicated. When they meet someone again, they might be influenced by the remembered feelings and other memories of the person. By saving the most frequently occurring emotion in the latest interaction with the user, the robot influences its initial emotional status when it meets the person again.

2) The user's personality and preference. The robot saves features of the user's personality such activeness or passiveness. This information helps the robot to control the activity in subsequent interactions with the user. For active users, the robot suggests dancing or singing. For passive users, the robot plays quiet songs. The user's preference for such things as likes and dislikes of dialogue topics are also saved. If the robot talks about such information in the next interaction, the human may consider the robot to be more intelligent and the interaction may be more dynamic and interesting.

6. Behavior System

The behavior system organizes its goal into a cohesive structure, as shown in Fig. 3. The structure has three levels and branches that address the three drives: the drive to interact with a human, the drive to ingratiate itself with a human and the drive to take a task. Each branch has a maximum of three levels. As the system moves down a level, the specific behavior is determined according to the affective relationship between the robot and human.

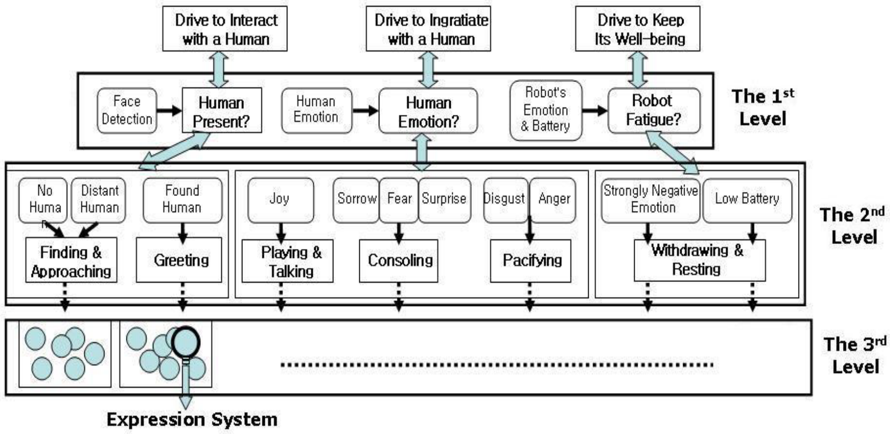


Fig. 3. Hierarch of the behavior system

6.1 First Level: Drive Selection

The behavior group of the first level determines which of the three basic drives should be addressed. The most important drive was designed to be addressed urgently and to effectively activate the appropriate behavior.

The drive to interact with a human is determined when the robot detects the human through the face detector of the perception system. The robot’s drive to ingratiate itself with a human is determined by the emotional state of the human as recognized through the bimodal emotion recognizer. The robot’s drive to preserve its own well-being is determined by the robot's energy; namely, the robot’s emotion and battery.

When the magnitude of the drive increases, there is more urgency to address the need and the drive makes a greater contribution to the activation of the behavior. However, the robot’s drive to preserve its own well-being has the highest priority for activation because this drive is related to the operating power of the robot.

6.2 Second Level: High-Level Behavior Selection

The second level decides which high-level behavior should be adopted according to the perception and internal information in the determined drive. In the first drive, if the human is far away or absent, finding and approaching behavior is adopted; however, if the human is detected, greeting behavior is adopted. In the second drive, one of three types of behavior is adopted depending on the emotional state of the human. If the human expresses sorrow, fear or surprise, consoling behavior is adopted. If the human expresses disgust or anger, pacifying is adopted. If the human is joyful, playing or talking is adopted to give the human more pleasure. In the third drive, if the robot is extremely angry or sad, withdrawing and resting behavior is adopted.

6.3 Third Level: Low-Level Behavior Selection

One type of high-level behavior includes several low-level types of behavior. Each low-level type of behavior is composed of dialogue and gestures, and is executed in the expression system. The low-level types of behavior in the same behavior group feature different dialogues but have the same behavior goal. A low-level type of behavior is therefore randomly selected in the behavior group.

7. Expression System

The expression system comprises three subsystems: a dialogue expression system, a 3-D facial expression system and a gesture expression system.

The expression system plays two important functions. The first function is to execute the behavior received from the behavior system. Each type of behavior consists of a dialogue between the robot and the human. Sometimes the robot uses interesting gestures to control the dialogue's flow and to foster interaction with the human. The second function is to express robot's emotion. The robot expresses its own emotions through facial expressions but it sometimes uses gestures to covey its intentions and emotions.

7.1 Dialogue Expression

Dialogue is a joint communicative process of sharing of information (data, symbols, and context) between two or more parties. In addition, humans use a variety of paralinguistic social cues (facial displays, gestures and so on) to regulate the flow of the dialogue [11]. There are three primary types of dialogue: low-level (prelinguistic), nonverbal, and verbal language. AMI communicates with humans through verbal language and appropriate gestures.

However, it is difficult to enable a robot to engage in natural dialogue with a human because current techniques for speech recognition and natural language processing are limited. We therefore predefined the dialogue flow and topics. Because AMI can recognize only a limited number of speech patterns, we constructed the dialogue as follows: First, to prevent AMI from misunderstanding the human speech, we enabled AMI to actively lead the dialogue by asking the user's intention in advance. Second, to avoid unnatural language, we enabled AMI to answer the most frequently used responses.

The dialogue expressions comprise the most commonly used sentences according to the selected behavior of the following behavior groups: finding and approaching, greeting, talking, playing, consoling, pacifying, withdrawing and resting. In the finding and approaching group, AMI calls the human. In the greeting group, AMI says hello to the human and asks the human's name and so on. In the talking group, the dialogue consists of various common topics such as hobbies, weather, movies and so on. In the playing group, AMI plays with the human through various activities such as jokes, an O-X quiz and a nonsense quiz. In the consoling and pacifying group, AMI asks what the human is angry about and then makes a joke to console or give pleasure to the human. Furthermore, AMI asks what the human is worried about when it recognizes that the humans is sad, and it listens like a friend or counselor to what the human says.

For AMI's speech synthesis, we used a text-to-speech program developed by Cowon Systems, Inc. The program produces natural Korean and English sentences.

7.2 Facial Expression

AMI's 3-D facial expressions show the internal emotional status synthesized in the robot's motivation system. These expressions make up for the limitations of the robot's mechanical face, which has difficulty displaying emotions. To implement these expressions, we used a 3-D OpenGL graphic library that shows various emotions through 3-D model faces with a 3-D heart that displays its heartbeat. Figure 4 shows AMI's facial expressions. Whenever AMI's emotions change, the facial expressions and the background image change dynamically.

Fig. 4. Facial expressions of AMI

7.3 Gesture Expression

AMI's gestures were designed to make humans feel as if AMI were human-like and friendly as shown in Fig. 5. AMI uses these gestures to express its own emotions and to make the dialogue more expressive. In designing AMI, we considered expressions that would attract the interest of humans. Accordingly, we developed various interesting gestures that matched the dialogue and robot's emotional status.

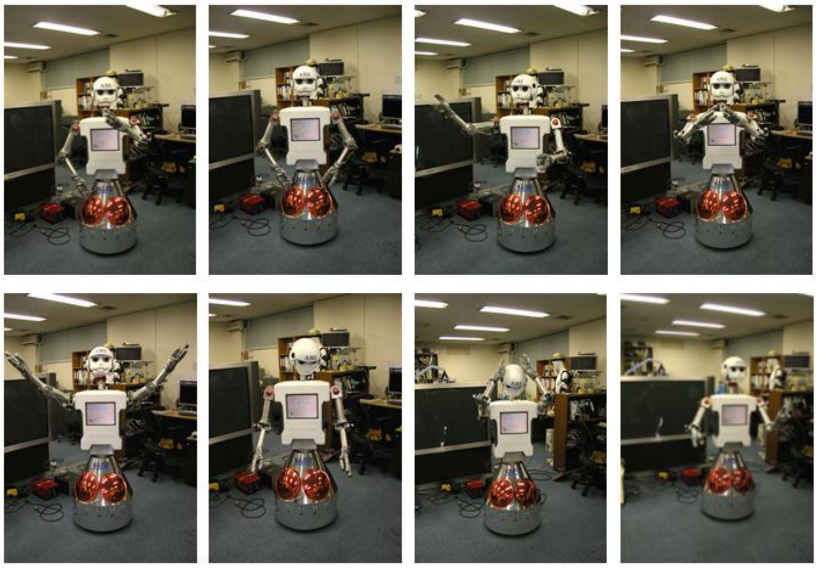


Fig. 5. Gesture expressions of AMI

8. Experimental Results and Conclusions

We checked the following factors on the basis of recorded internal parameters and the working status of each subsystem:

- Does the system recognize the human's emotional status? (The perception system.)
- Does the system match the robot's emotion to the user's emotion and interactive response? (The emotion system.)
- Does the system activate the right drive by considering the activation condition of each drive? (The drive system.)
- Does the system memorize the robot's previous emotional response to several users? (The memory system.)
- Does the system determine the right behavior according to the current activated drive and user's emotional status? (The behavior system.)

We confirmed that each subsystem works properly towards satisfying its objectives. Based on the evaluation results, we drew a graph in Fig. 6 which shows the subsystem's flow during a sample interaction. The graph also shows the behavior system (finding, greeting, consoling and so on), the motivation system (the robot's drives and emotions), and the perception system (the user's emotional status).

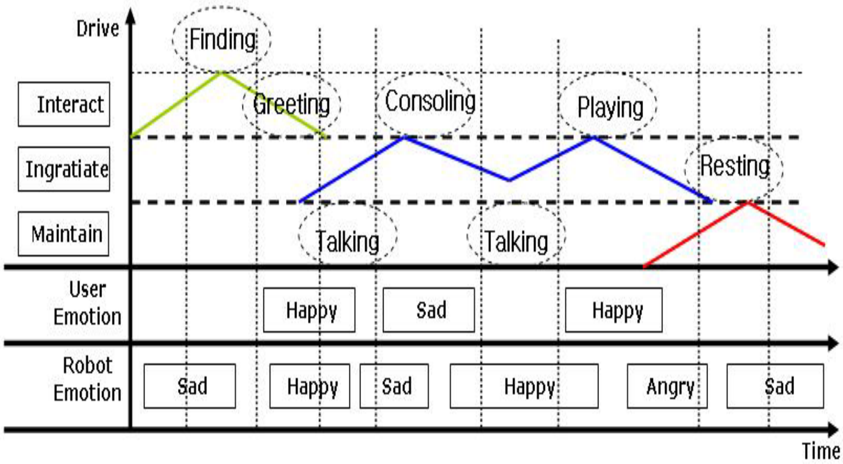


Fig. 6. The work flow of the system

We have presented an affective communication robot that is designed to lead human-robot interactions by recognizing the emotional status of a human, by expressing its emotions through multimodal emotion channels similar to those of a human, and by responding appropriately to human emotions.

In the future, we plan to extend AMI's dialogue expressions in various topics. AMI's current framework is designed to interact with people, but only in a limited number of conversational topics. Furthermore, the current memory system stores the memory of users and their emotions but not conversational topics. Our future system will be capable of memorizing previous conversational topics as well as expecting the user response by considering current conversational topics.

Fig. 7. Interaction with a human

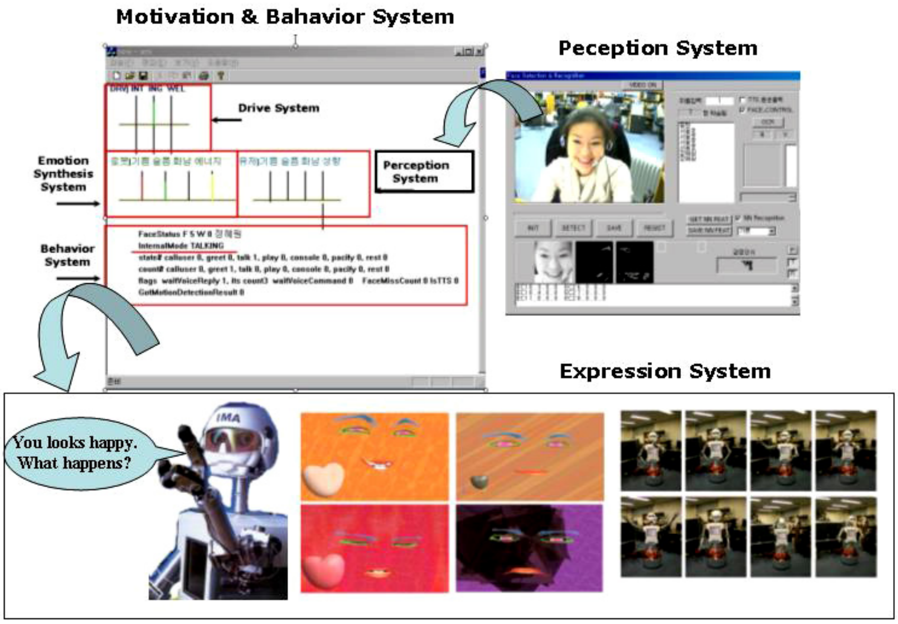


Fig. 8. The flow of the overall system

Acknowledgement

This research was partly supported by the Korea Ministry of Science and Technology through the Brain Science Research Center, and by the Digital Media Lab Project in Korea.

Appendix A: Sample conversation between AMI and users

Conversation 1

(GREETING)
AMI: Is there anyone who wants to talk to me?
User 1: Hello, AMI.
User 1 approaches. AMI finds a human face.
AMI: I heard someone calling me.
AMI: Ah, you are there. You're (User 1's name).
AMI: Let's shake hands.
User 1: OK.
User 1 and AMI shake hands.
AMI: How are you today?
User 1: Not very good. I had a pretty bad day.
AMI: I'm sorry to hear that.
(TALKING)
AMI: Is it cold today?
User 1: Yes.
AMI: I think so. I like warm and sunny days. How about you?
User 1: Well, I also prefer warm weather.
AMI: I thought so.
(CONSONLING)
AMI: Anyway, you look somewhat sad. It there anything wrong with you?
User 1: Yes.
AMI: Yeah, you look pretty unhappy. Can you tell me what happened to you today?
User 1: OK.
AMI: Tell me then.
User 1: It was a terrible day for me. I had a lot of homework to submit and I also had an exam today. I didn’t have enough time to study for the exam because of my homework. So my exam was terrible, and that's why I'm not feeling good now.
AMI: I really fell sorry about that. Why don't we listen to some music?
User 1: Good idea.
AMI plays some music.
... (The conversation continues until the user says, “AMI, let’s stop talking.”)

Conversation 2

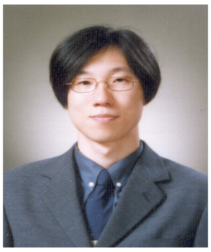
(GREETING)
AMI: Hello? Is anyone there?
User 2: Hi.
User 2 starts speaking.
AMI: Ah. I head a voice. Where are you?
User 2: Turn left.
AMI: I found you. Nice to meet you. Can you tell me your name?
User 2: (User 2's name).
AMI : You're (User2's name). Let's shake hands.
User 2: OK.
User 2 and AMI shake hands.
AMI: How are you today?
User 2: Fine.
AMI: I'm fine too.
(TALKING)
AMI: What did you have for lunch?
User 2: I had a hamburger, potato chips, and a salad.
AMI: Sounds nice. What's your favorite food?
User 2: Hmm. Nothing particular. I like everything.
AMI: Good.
(PLAYING)
AMI: You look fine today. Let's play a game.
User 2: OK.
AMI: Good. How about a common sense quiz. I'll give you some questions and you have to answer. Is that OK?
User 2: OK.
AMI: The first world cup was held in Greece.
User 2: Correct.
AMI: Sorry. Wrong. Next. The reason for VDT syndrome is too much exercise.
User 2: No!
AMI: You're right. Good job.
User 2: Had enough?
AMI: OK. Let's do it again later.
... (The conversation continues until the user says, “AMI, let’s stop talking.”)

References

[1]. C. Breazeal, "Designing Sociable Robots, MIT Press", Cambridge, MA, 2002.
[2]. Ronald C. Arkin, Masahiro Fujita, Tsuyoshi Takagi and Rika Hasegawa, "An Ethological and Emotional Basis for Human-Robot Interaction", Robotics and Autonomous Systems , 42 (3-4), March 2003.
[3]. T. Shibata, at al., "Emergence of emotional behavior through physical interaction between human and artificial emotional creature", Int. Conf. on Robotics and Automation, pp. 2868-2873, 2000.

[4]. Sidner, C.L.; Lee, C.; Lesh, N., "The Role of Dialog in Human Robot Interaction", International Workshop on Language Understanding and Agents for Real World Interaction, July, 2003
[5]. SY Yoon, RC Burke, BM Blumberg, GE Schneider, "Interactive Training for Synthetic Characters", AAAI 2000.
[6]. Naoko Tosa and Ryohei Nakatsu, "Life-like Communication Agent - Emotion Sensing Character "MIC" & Feeling Session Character "MUSE"", ICMCS, 1996.
[7]. Yong-Ho Seo, Ho-Yeon Choi, Il-Woong Jeong, and Hyun S. Yang, "Design and Development of Humanoid Robot AMI for Emotional Communication and Intelligent Housework," Proceedings of IEEE-RAS Humanoids 2003, pp.42.
[8]. P. Ekman and W. V. Friesen., "Facial Action Coding System: Investigator's Guide", Consulting Psychologists Press, Palo Alto, CA, 1978.
[9]. H. Schlossberg, "Three dimensions of emotion", Psychology Review 61, 1954.
[10]. C. Armon-Jones, "The social functions of emotions", R. Harre (Ed.), The Social Construction of Emotions, Basil Blackwell, Oxford, 1985.
[11]. M. Lansdale, T. Ormerod, "Understanding Interfaces", Academic Press, New York, 1994.

Hye-Won Jung received a BS degree from Korea University, Seoul, Korea, in 2001, and an MS degree from the Department of Electrical Engineering and Computer Science, the Korea Advanced Institute of Science and Technology (KAIST), in 2004. She is currently a researcher in the M1 Group, Mobile Multimedia Lab, LG Electronics Institute of Technology. Her research interests include affective computing, human-robot interaction and mobile multimedia.



Yong-Ho Seo received a BS and MS from the Department of Electrical Engineering and Computer Science, KAIST, in 1999 and 2001, respectively. He is currently pursuing a PhD at the Artificial Intelligence and Media Laboratory, in the Department of Electrical Engineering and Computer Science at KAIST. His research interests include humanoid robots, human-robot interaction, and robot vision.

M. Sahngwon Ryoo received his BS degree from KAIST in 2004. He has been a researcher in the Artificial Intelligence and Media Laboratory, KAIST, from 2003 until the present. His research interests include robotics, artificial intelligence, human-computer interaction and human-robot interaction.

Hyun S. Yang received a BS degree from the Department of Electronics Engineering, Seoul National University, Korea, in 1976, and an MSEE and Ph.D. from the School of Electrical Engineering, Purdue University, West Lafayette, IN, in 1983 and 1986, respectively. He worked with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, from 1986 to 1988 as an Assistant Professor. Since 1988, he has been with the Department of Computer Science at KAIST, Daejeon, where he is currently a full professor. Since 1990, he has been a director of the Computer Vision and Intelligent Robotics Lab of the Center for Artificial Intelligence Research, which is sponsored by the Korea Science and Engineering Foundation at KAIST. He chaired the Society of Artificial Intelligence Research of the Korea Information Science Society from 1991 to 1993. His current research interests include humanoid robotics, computer vision, interactive media art and multimedia.