2 reviews previous works that are relevant to this study. Section 3 proposes a framework for affective interaction between humans and robots. Section 4 provides specific procedures of music-aided affective interaction. Section 5 explains the experimental setup and results. Finally, Section 6 presents our conclusions.

2. Previous Works on Affective Interaction between Humans and Robots

An increasing awareness of the importance of emotions has been leading researchers to attempt to integrate affective computing into a variety of products such as electronic games, toys, and software agents [11]. Many researchers in robotics also have been exploring affective interaction between humans and robots in order to accomplish the intended goal of human-robot interaction.

For example, a sociable robot, 'Kismet', understands human intention through facial expressions and engages in infant-like interactions with human caregivers [12]. 'AIBO', an entertainment robot, behaves like a friendly and life-like dog that responds to either the touch or sound of humans [13]. A conversational robot called 'Mel' introduced a new paradigm of service robots that leads human-robot interaction by demonstrating practical knowledge [14]. A cat robot was designed to simulate emotional behavior arising from physical interactions between a human and a cat [15]. Tosa and Nakastu have concentrated on the technology of speech emotion recognition to develop speech-based robot interaction. Their early works, 'MUSE' and 'MIC', were capable of recognizing human emotions from speech and expressing emotional states through computer graphics on a screen. They have consistently advanced their research directions and developed more applications [16, 17].

3. Framework for Affective Interaction

In efforts to satisfy the requirements for affective interaction, researchers have explored and advanced various types of software functions. Accordingly, it is necessary to integrate those functions and efficiently manage systematic operations according to human intentions. The best approach for this is to organize a control architecture or a framework for affective interaction between a human and a robot.

Our research target is humanoid service robots that perform human-like operations and behaviors. Thus, we propose a new framework based on a model of the human brain structure developed by the cognitive scientist Joseph E. Ledoux [18]. This framework consists of three individual systems associated with one another, as demonstrated in Figure 1.

![Figure 1: Framework for affective interaction.](image)

The primary function of the perception system is to obtain human emotional information from the outside world through useful indicators such as facial expression and speech. The memory system records the emotional memories of users and corresponding information in order to utilize them during the interaction with humans. Finally, the expression system executes the behavior accordingly and expresses emotions of the robot.


In the conventional approaches to achieve affective interaction, both speech and facial expression have been mostly used as representative indicators to obtain human emotional information. Those indicators, however, have several disadvantages when operated in robots, as addressed in Section 1. In addition, most of the conventional approaches convey the robot's emotional states in monotonous ways, using a limited number of figures or synthesized speech. Thus, users easily predict the robot's reactions and can lose interest in affective interaction with the robot. To overcome these drawbacks, we adopt music information in the framework of affective interaction.

Music is an ideal cue for identifying the internal emotions of humans and also has strong influences on the change of human emotion. Hence, we strongly believe that music will enable robots to more naturally and emotionally interact with humans. For the music-aided affective interaction, the mood of the music is recognized in the perception system and is utilized in the determination of the user's emotional state. Furthermore, our expression system produces affective reactions to the user emotions in more natural ways by playing music that the robot recommends or songs that the user previously listened to while exhibiting that emotion. The music-aided affective reaction is directly supported by the memory system. This system stores information on the music the user listens to with a particular emotional state. This section describes further specific features of each system in the framework of music-aided affective interaction.

4.1. Perception System. The perception system recognizes human emotional states on the basis of various indicators. For multimodal emotion recognition, the proposed system utilizes the musical mood as a supplementary indicator along with speech and facial expression as primary indicators. Consequently, the perception system comprises three recognition modules: for musical mood, facial expression, and