Received November 21, 2016, accepted December 2, 2016, date of publication December 19, 2016,
date of current version February 25, 2017.

Digital Object Identifier 10.1109/ACCESS.2016.2641480

INVITED PAPER

Emotion Communication System

MIN CHEN[1], PING ZHOU[1], AND GIANCARLO FORTINO[2]
[1]School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
[2]Department of Informatics, Modeling, Electronics, and Systems, University of Calabria, 87036 Rende, Italy

Corresponding author: M. Chen (minchen2012@hust.edu.cn)

ABSTRACT In today's increasingly rich material life, people are shifting their focus from the physical world to the spiritual world. In order to identify and care for people's emotions, human–machine interaction systems have been created. The currently available human–machine interaction systems often support the interaction between human and robot under the line-of-sight (LOS) propagation environment, while most communications in terms of human-to-human and human-to-machine are non-LOS (NLOS). In order to break the limitation of the traditional human–machine interaction system, we propose the emotion communication system based on NLOS mode. Specifically, we first define the emotion as a kind of multimedia which is similar to voice and video. The information of emotion can not only be recognized, but can also be transmitted over a long distance. Then, considering the real-time requirement of the communications between the involved parties, we propose an emotion communication protocol, which provides a reliable support for the realization of emotion communications. We design a pillow robot speech emotion communication system, where the pillow robot acts as a medium for user emotion mapping. Finally, we analyze the real-time performance of the whole communication process in the scene of a long distance communication between a mother–child users' pair, to evaluate the feasibility and effectiveness of emotion communications.

INDEX TERMS Emotion communications, emotion communication protocol, Markov state transfer.

I. INTRODUCTION

In our world, interconnected through the Internet, mobile phones and countless things, the seamless integration of physical world and information world has become the future development trend of the network [1]–[3]. In the increasingly rich material life today, people are transferring the attention from the physical world to the spiritual world, and human emotions will represent a direct reference to the spiritual world. Therefore, a variety of methods and technology for human emotion recognition are emerging and human-robot interaction also has come into being as a new research hotspot [4].

There are a lot of research results in the human-machine interaction (HMI) system, mainly based on the machine to identify human emotions. They can be divided into four categories [5], [6] according to different emotion recognition types: audio-visual information-based, physiological signals-based, tactile perception-based, and multi-modal form-based.

- The study of emotional HMI system based on audio visual information mainly uses machine learning algorithms to analyze the human voice, facial expression or body posture to recognize human's emotions. Anagnostopoulos *et al.* [7] survey statistics from 2000 to 2011 with different characteristics and classifiers of emotional verbal communication system. This system simply aims to improve the accuracy of emotion recognition, without taking the level of emotional interaction into account. The non-verbal communication system in [8] identifies human emotions with high accuracy by non-verbal cues such as expression recognition and head movement, but the mobility and freedom of users are greatly limited.
- Physiological signals based emotional HMI system research [9] collects some of the user's representative physiological signals, such as ECG, body temperature and heart rate, but it requires binding or pasting sensors in the human body which makes the human comfort limited. Kim *et al.* [10] solved the above-mentioned comfort problem by integrating all kinds of sensors into a portable device. However, the particularity of this kind of sensor device makes the freedom of human being restricted and the accuracy of emotion recognition low. For instance, when the user carries such equipment on the stairs, the emotional recognition would be interfered

TABLE 1. A Comparison of emotion recognition method for HMI.

Modality	Accuracy	Computational Complexity	Cost	Convenience	Robustness
AudioVisual-based[7,8]	Medium	Medium	Low	High	High
Physiological Signal-based[9,10,11]	Medium	Medium	Medium	Medium	Medium
Haptic-based[12,13]	High	High	Medium	Medium	High
Multimodal-based[14,15]	High	High	Medium	High	Medium

owing to the influence of physiological signals caused by external factors such as movement. There is also a class of research using bioelectric technology to replace sensor-based devices directly, such as EQ-Radio systems [11] that employ RF (radio frequency) based identification to collect user information for emotion recognition. The system collects the user's breathing and heartbeat signals without the use of sensors, but this approach makes the user's mobility and the user's range of motion greatly restricted.

- The research of emotional HMI system based on tactile perception is mainly divided into two types. One is that the robot feels the user's own touch, so as to analyze the user's emotion [12]. In this kind of system, the tactile perception would be strong only when the user's tactile performance is exaggerated. Its sensitivity is low, and the robot itself is limited by the interactive function. The other is through wearing feedback device, remotely triggering action for tactile comfort [13]. Such devices are more comprehensive considered the emotional comfort for users, but do not involve emotion recognition and communication, and the user's degree of freedom is limited by the tactile comfort device.
- The multi-modal based emotional HMI system [14], [15] is a combination of audiovisual information, physiological signals or tactile perception information. The computational overhead of this kind of multimodal identification is very high, and the delay of the robot feedback is long.

In summary, the above four types of systems focus on the identification of human emotions, as much as possible to improve the accuracy of emotion recognition, but rarely even never consider the degree of human-machine interaction. The simple feedback of robot is the same to all users. It cannot achieve the emotion communications between human beings and robots, nor do they regard the emotion as a mutual transferring information between two (or multi) parties. Emotion communications should be aimed at the individual and need personalized treatment. Different people having the same emotion need different communications, the same person feeling a different emotion also needs different communications.

The rest of paper is organized as follows. Section II proposes a taxonomy for emotion communications. In Section III the key technologies used are detailed. Section IV presents the emotion communication system architecture whereas in Section V the communication protocol enabling emotion communications is defined. Section VI analyzes the proposed system using a markov state transition-based model. An effective example of emotion communications is proposed and evaluated in Section VII. Finally, conclusions are drawn and future work is anticipated.

II. TAXONOMY OF EMOTION COMMUNICATIONS

Table 1 summarizes the representative research of traditional HMI systems based on the five features of emotion recognition method, that is, emotion recognition accuracy, computational complexity, practical cost, convenience and robustness:

- In terms of the accuracy of recognition methods, the haptic-based and multimodal-based approach have higher accuracy owing to that the haptic-based approach has real physical sensations, and multimodal-based approaches can take several performance factors into account. In contrast, methods based on physiological signals are generally medium, whereas the method that bind sensors directly on the human body are slightly higher, and the physiological signal based feedback device and the radio frequency based identification is relatively low.
- In terms of the computational complexity of recognition methods, the audiovisual-based and physiological signal based methods are computationally less complex because their data dimensions are usually small.
- In terms of the cost in practical, the widespread use of audiovisual sensors makes the cost of audiovisual-based methods lower than the other three, whereas physiologic-based methods tend to be costly because of the required number and accuracy of the sensors.
- In terms of the convenience for use, physiological signal based and haptic-based methods are less convenient than audiovisual-based, multimodal-based methods because of the need for body participation of user.
- In terms of the robustness, the certainty of audiovisual information and multimodal forms make them more robust to external factors.

The current research on human-machine emotional interaction system is aimed at improving the accuracy of human emotion recognition. In this paper, the difference between the existing HMI systems and the emotion communication system proposed is considered from the perspective of communication, as shown in Table 2. From the two modes of LOS communications and NLOS communications, the differences between the existing HMI system in the aspects of synchronization, user mobility, user freedom, feedback effect and communication QoS are compared.

TABLE 2. A comparison of HMI system from the communication perspective.

Modality			Synchronization	Mobility of User	Freedom of User	Feedback Effect	Communication QoS
LOS	AudioVisual-based	Verbal-based[7]	Low	Low	Low	Low	Low
		Nonverbal-based[8]	Low	Low	Low	Low	Low
	Physiological Signal-based	Body sensors[9]	Medium	Low	Low	Low	Low
		Portable wireless biofeedback device[10]	Low	Medium	Medium	Low	Medium
		EQ-Radio[11]	Medium	Low	High	Low	Medium
	Haptic-based	Haptic creature[12]	Low	Medium	Medium	Low	Medium
		Haptic jacket[13]	Medium	Medium	Low	Medium	Medium
	Multimodal-based	[14]	Medium	Low	Low	Low	Medium
		[15]	Medium	Medium	Medium	Medium	Medium
NLOS	Non-line-of-sight Distance Based Emotion Communications		Medium	High	High	High	High

- In terms of the synchronization of feedback and emotion recognition, high synchronization of feedback and emotion recognition means low latency. In the LOS mode, methods [7], [8], [10], [12] have low synchronization, which takes much time from signal collection, data storage and data analysis.
- In terms of the degree of user mobility, in the LOS mode, the distance limits the capture of face or voice information in [7], [8], [14], the binding of sensors on human body in [9], and the RF signal transmitting device needs to align the human chest in [11] so making low mobility of user.
- In terms of the degree of user freedom, in the LOS mode, the limitation of the methods [7]–[9], [14] mentioned above also makes the freedom of the user low. And the constraints of wearing Haptic Jacket in [13] limited user's freedom.
- In terms of the effect of feedback, in the LOS mode, all methods involved less interaction triggered by robot and the feedback effect is not satisfactory.
- In terms of communication QoS, in the LOS mode, due to restrictions of delay, user mobility, user freedom, feedback effect, all methods have low QoS for user.

The goal of the current HMI system is often the interaction between the human and the robot in the near distance, namely the interaction under the sight distance. The NLOS model proposed in this paper is to remove the restrictions of traditional HMI, not just to identify people's emotions and simple robot feedback, but the emotion as a kind of information that can be transmitted over a long distance. Robots are not only an interactive tool, but also a medium for end-user communications. Based on this, we propose the emotion communications, and put forward solutions from three aspects of emotion definition, generation and delivery to achieve emotion communications. In particular:

- **We define emotion as a kind of information that is similar to multimedia**: emotions can be produced, elaborated, transmitted, shared and stored. Emotion could affect the expression of other information [16]. For example, a person is happy, his/her mouth should be rising, the tone of voice should be high, the person even expresses the happiness by dancing with increased heart rate. We define emotion as a kind of special multimedia information, which will affect the expression of other information, and the expression of other information can also reflect emotions.
- **We define the generation of emotion from three categories**: physiological information, environmental information and social network information, respectively. The physiological information is mainly the body signal data of the user. Environmental information refers to the surrounding information of the user. And social network information refers to social networks personal posted data of the user.
- **We define two modes of emotion communications**: solo-mode and mutual-mode, respectively. Improving the existing HMI system based on their advantages. Under the *Solo-Mode*, the user communicates with the robot. The robot has its own emotions and interacts with the user. And under the *Mutual-Mode*, the robot itself does not have emotion, only acts as the medium of the emotion mapping. The robot incarnates the emotion of the far-end user and realizes the emotion communications between the remote users.

The real sense of emotion communications between users can be realized through the definition, generation and transmission of emotion. But the realization of emotion communications is facing great challenges. As the traditional telephone communications and social media communications, emotion is transmitted as information between users, but the generation of emotions is complex and time-consuming for its particularly nature. The transmission of emotion also faces the challenge of required communication quality of service of user. Therefore, this paper first gives the key technology of the generation and analysis of emotion, and then puts forward the emotion communication protocol to solve the communication synchronization challenge of the two sides in the communication process.

III. KEY TECHNOLOGY

We define an emotion communication system as: a cognitive system, which can collect human's emotion data, understand human's emotion, care human's emotion and interact with users emotionally. In order to accomplish the three tasks of gathering human emotion data, understanding human emotions and achieving emotion interaction, the key technologies involved in the system include: smart clothing, robot, machine learning and mobile communication technology (5G) throughout the whole system.

Smart Clothing: Carrier for collecting data. We proposed Wearable 2.0 [17]–[19], which is based on Smart Clothing as the carrier, a variety of sensors integrated in the clothes, the user cannot feel the presence of wearable equipment, achieves the physiological data collection. The Smart Clothing is proposed to solve the traditional wearable devices with low comfort, the user needs to carry, inconvenient operation, a single function, as well as the need for users to participate in the operation and maintenance of equipment and other shortcomings. At the same time, the physiological signals of the wearer can be uploaded to the intelligent device or the cloud in real time, and provide the monitoring of mobile health or exercise management at any time and any place in daily life [20]–[22].

Robot: Media for environmental perception and intelligent interaction. As an emerging communication medium, the rapid development of artificial intelligence has brought endless possibilities and advantages to robots, from the early entertainment function to the replacement of manual labor and then to today's care and perception. The robot we proposed is a new generation of intelligent perception and interactive robot [23]. Its advantages are: on the one hand, the humanoid appearance, the use of cute cartoon image gives users endless goodwill. On the other hand, it has the function of environmental perception and intelligent interaction: its integration of multiple sensors can sense the environment information and user information such as voice, facial expression, movement. Meanwhile, it can take emotion care for people in turn.

Machine Learning: Key for emotion analysis. In the emotion communication system proposed in this paper, the key of emotion communications is to understand and analyze each other's emotions, which is inseparable from the tools and methods of emotion analysis. Based on cloud platform, we use machine learning and deep learning algorithms to build self-learning mechanism and model in the cloud, and analyze the current status and emotion in real-time according to the user's data [24].

5G: Support for network communication. 5G mobile communication technology, including some vehicular telematics oriented communication technology [25], makes long-distance and low-delay communication possible [26]–[28]. User's emotion is subject to the impact of subjective and objective factors and very easily to instantaneous change, so timely analysis of the user's emotion has become the evaluation criteria of emotion communication accuracy. 5G is the support of the whole communication process of the emotion communication system [29], [30].

The characteristics of emotion communication system mainly include two aspects:

- **Emotion**: We have interconnected the physical world with the spiritual world and focus on people's emotions, to solve the problem of the emotion care for children, young people and elderly. We make communications and emotion care between children and distant parents, young people and distant friends, the elderly and distant relatives possible.
- **Communication**: Just like telephone communications and social media communications, the emotion communications discussed in this paper is concerned with the communication between two sides. Not only the simple static expression of text, voice and video, but from the forms of touch and sense to give users a real sense of feelings and experiences. At the same time, the two sides of the emotion communication, as well as other communication methods, should follow the established protocol to ensure the authenticity, validity and integrity of the system. Emotion communication can be integrated with other human decision based communication systems, such as cooperative networks [31], to enhance specific content-aware applications.

IV. ARCHITECTURE

As an emotion communication system, the communications between the two sides need to perform the following tasks: the collection of emotional signals, the transmission of emotional signals, the recognition of emotional signals and emotional feedback. Emotion is a kind of personal and private information, and its privacy protection is also very important in the whole communication process [32]. According to the tasks that need to be done, the system is divided into four layers, namely, data collection layer, emotion communication layer, cloud analysis layer and emotional feedback layer. The specific tasks for each layer are as follows:

- **Data collection layer**: The first step of emotion communications is to get the data that can express emotion. The depth and breadth of data acquisition is the key to identify the correct emotion. Therefore, in order to collect multi-dimensional user data, we collected user data from the perspective of physiology, environment and social networks simultaneously. Smart clothing can collect user's physiological indicators, including ECG, heart rate, body temperature, blood pressure and other parameters. Robot can capture the user's voice, facial features and some of the surrounding environment information, including the user's current location, ambient temperature, noise and so on. The user's mobile smartphone can extract the user's recent social information, including phone logs, SMS logs and social networking.
- **Emotion communication layer**: Considering the key to the emotion communication system is accurately and timely understanding the user's current emotions, we use 5G to support more computing tasks, only a small amount of work offloaded to the terminal to complete [26]. By using 5G mobile communication technology, the communication protocol between user

and robot, robot and cloud, cloud feedback to robot and robot feedback to user is proposed to improve user's QoS [33]–[35] and QoE [36]–[38].

- **Cloud analysis layer**: We use a powerful cloud platform to analyze and deal with the various data transmitted by the communication layer [39]. In the cloud analysis layer, firstly the data should be filtered to remove unnecessary data; secondly data aggregation, i.e. time-space label based integration of physiological data, environmental data and social networks data, to obtain multi-dimensional fusion data results is carried out [40], [41]; thirdly, the emotion analysis model is established, and the emotion of the users is analyzed according to the multi-dimensional emotion data.
- **Emotional feedback layer**: The key that users can truly feel their emotions in communications is to get real emotional feedback, in addition to including text, voice, video and other visual and auditory comfort, but also including the real comfort of sense and touch. In this layer, we design a humanoid robot. Through robot's soft features and perceptual characteristics to reflect the other party's true feelings from the cloud, and then present the features of the other party and comfort the current user through their feedback actions.

The architecture of emotion communication system is shown in Fig. 1. From the architectural point of view, support for the emotion communication layer and cloud analysis layer can have different concrete forms, that is the various modalities to support emotion communication layer and cloud analysis layer based on network $< N >$ and cloud platform $< CP >$. The $< Network, Cloud >$ is considered as the abstract level, that means the type of network and cloud are determined by related real platforms where to implement the abstract level. Based on the abstract level, the concrete level could have various modalities, such as:

- Mode 1: $< Internet, Cloud >$, the general Internet and cloud platform to support general communication and tasks.
- Mode 2: $< 4G\ LTE, Cloud >$, 4G LTE and cloud platform currently support high-speed computing tasks.
- Mode 3: $< 5G, Mobile\ Cloud >$, 5G and mobile could platform to support long-distance but low-delay communication and complex computational tasks in the future.

The three major advantages of the emotion communication system we created are:

- *Sensing and Analysis of Expressed Emotions*. With the smart clothing, 5G communication network and the powerful processing power of the cloud, this system can collect user data and analyze user emotion timely.
- *Robot-based Emotional Feedback*. With the help of robot media for perception and intelligent interaction, users can truly feel emotions and comfort of the other communication side.

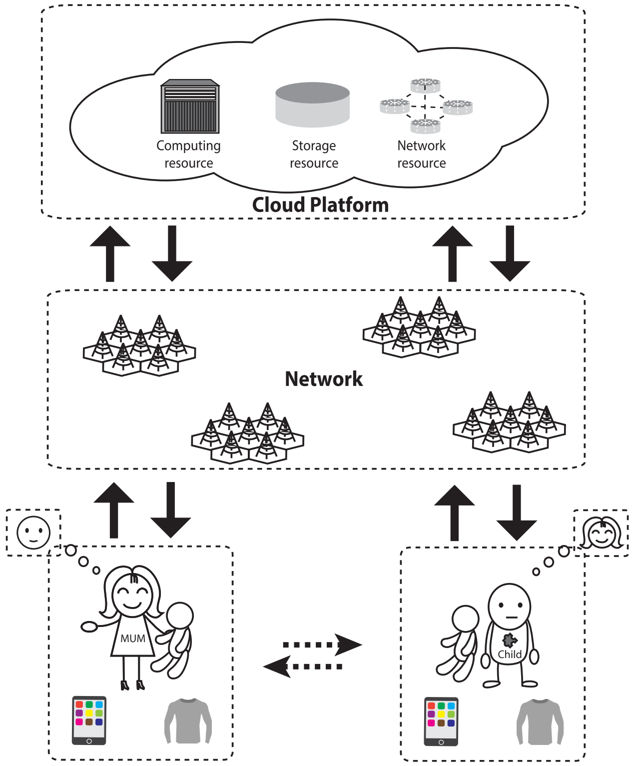

FIGURE 1. The architecture of emotion communication system.

- *Emotion Communications*. Long-distance transmission of the real emotion in accordance with the specific emotion communication protocol.

V. EMOTION COMMUNICATION PROTOCOL

Like telephone communications and social media communications, both sides of emotion communication also need to follow specific communication protocols, that is, the behavior obeyed throughout the whole communication process from communication link establishment, emotional data collection, emotion analysis, emotion push to emotional feedback. In the layer of emotion communication, it can be re-layered according to the different behavior modules, namely:

- *Object initialization layer*: Complete the initialization work of all the objects in the emotion communication.
- *Communication layer*: Conduct emotion communications after the establishment of communication link between two sides.
- *Emotional feedback layer*: Take feedback mutually to the two communicating parties after the end of a communication.

The flow chart of emotion communications is shown in Fig. 2. Emotion communication protocol is introduced in detail below from the perspective of objects, parameters, communication command set, and communication process.

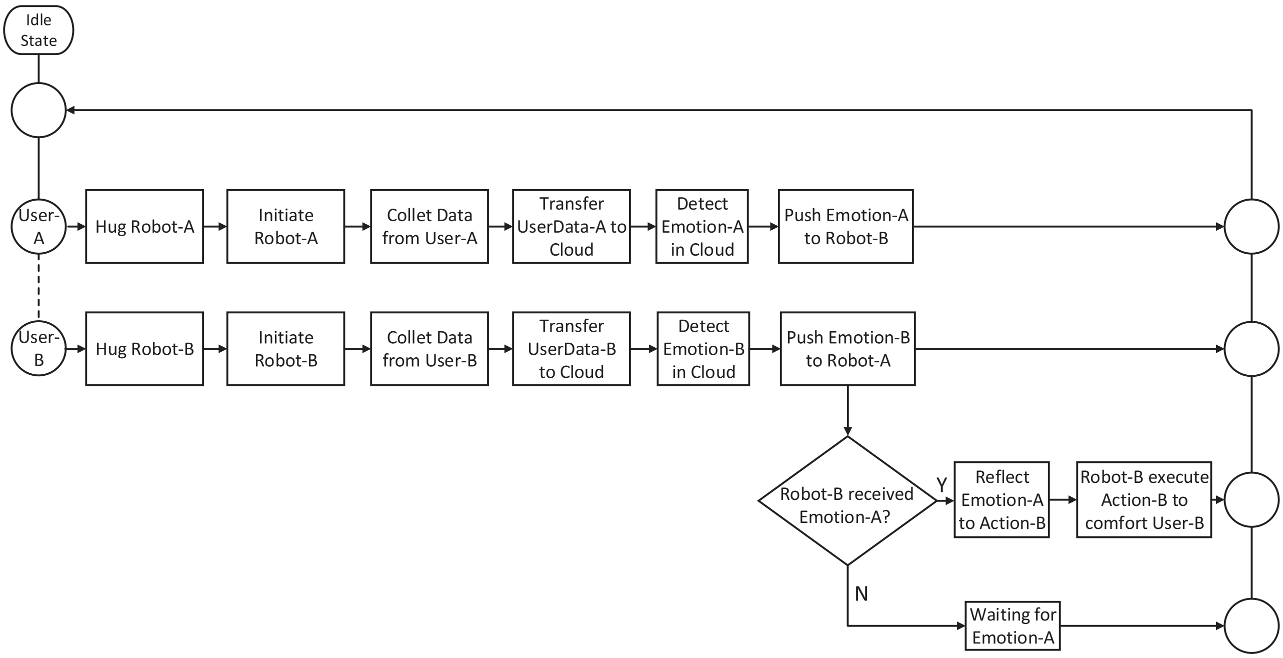

FIGURE 2. The flow chart of emotion communications.

A. OBJECTS

The objects in this protocol, i.e., entities involved in the emotion communication, include:

User: Users involved in emotion communication.

Emotion Communication Terminal Device (ECTD): Robot terminals that collect user data, receive emotions and perform emotional feedback.

Cloud: Cloud platform that inputs user data, for emotion analysis, and outputs emotions.

B. PARAMETERS

The parameters involved in the communication process are as follows:

Data set of user: UserData = { ECG signals, voice, location, call logs, application logs, social networks personal posted data };

Emotion set of user: Emotion = { anger , boredom, sadness, calm, surprise, happiness, fear };

Feedback action set of ECTD: Action = { embrace force, temperature, heartbeat, music, voice };

The elements in the parameter set can be defined according to the different applications, and the robot terminal takes different levels of feedback action in accordance with the emotion level.

C. COMMAND SET

1) CommunicationConnect for communication link establishment: the two parties establish a communication link. The instruction prototype is CommunicationsConnect( ) and applicable to user.
2) ECTDInitiate for robot terminal initialization: initialize the parameters of the robot terminal itself. The instruction prototype is ECTDInitiate( ) and applicable to ECTD.
3) DataCollect for user data collection: collect the physiological signal and data of the user after the initialization of robot terminal. The instruction prototype is DataCollect(User) and applicable to ECTD.
4) DataToCloud for data transmission to the cloud: transfer the collected data to the cloud. The instruction prototype is DataToCloud(UserData) and applicable to ECTD.
5) EmotionDetection for emotion Analysis in the cloud: analyze the emotion in the cloud when receiving user data. The instruction prototype is EmotionDetection(UserData) and applicable to cloud.
6) EmotionPush for emotion push: push a certain result of emotion set to the robot terminal after a comprehensive analysis by the cloud. The instruction prototype is EmotionPush(Emotion) and applicable to ECTD.
7) ActionInference for action inference: infer the mapping through robot's own emotion-action inference engine according to the emotion pushed by the cloud. The instruction prototype is ActionInference(Emotion) and applicable to ECTD.
8) EmotionFeedback for feedback execution: take the feedback action according to the mapping result of the inference engine. The instruction prototype is Emotion-Feedback(Action) and applicable to ECTD.
9) CommunicationBreak for communication link break: the two parties disconnect the communication link. The instruction prototype is CommunicationBreak( ) and applicable to user.

TABLE 3. A summary of the command set of the emotion communication protocol.

Command	Description	Applicability
CommunicationConnect()	The two parties establish a communication link	User
ECTDInitiate()	Initiate the parameters of the robot terminal itself	ECTD
DataCollect(User)	Collect user' s physiological signal and data	ECTD
DataToCloud(UserData)	Transfer the collected data to the cloud	ECTD
EmotionDetection(UserData)	Conduct a comprehensive analysis of the user's emotions	Cloud
EmotionPush(Emotion)	Push the final emotion result to the robot terminal	ECTD
ActionInference(Emotion)	Infer the mapping from emotion to feedback through robot's own inference engine	ECTD
EmotionFeedback(Action)	Take the feedback action according to the mapping result	ECTD
CommunicationBreak()	The two parties disconnect the communication link	User

The summary of communication command set is shown in Table 3.

D. COMMUNICATION PROCESS

First, a communication link is initiated by either of the two communicating parties, so as to establish a CommunicationConnect. A communication module of one of the robot terminals initiates a link signal to the opposite robot terminal and waits until a reply is received to indicate that the communication link has been established.

User-A and User-B embrace their respective robot terminals ECTD-A and ECTD-B, and the robot terminal performs its own initialization operation ECTDInitiate.

After its initialization is complete, the terminal ECTD-A through the DataCollect collects the data UserData-A of the User-A, and the terminal ECTD-B through the DataCollect collects the data UserData-B of the User-B.

After collecting the data, terminal ECTD-A and ECTD-B transmit UserData-A and UserData-B to the cloud through DataToCloud respectively.

After receiving the UserData-A and UserData-B, the Cloud performs the emotion analysis for both sides through EmotionDetection, including data preprocessing, feature extraction, emotion recognition and emotion fusion.

After the emotion fusion in the cloud, pushing the certain emotion result of emotion set to the other party's robot terminal through EmotionPush, that is, the emotion analysis result Emotion-A of UserData-A is pushed to ECTD-B, and Emotion-B of UserData-B is pushed to ECTD-A.

After knowing the emotion level of User-A, ECTD-B completes the mapping of emotion Emotion-A to action set Action-B according to its own inference engine.

According to its own reasoning results, ETCD-B takes corresponding feedback action through EmotionFeedback, to simulate human-like action, including the embrace, heartbeat, body temperature, play music and comfort of voice, etc.

An illustration of execution sequence of communication events between users, terminals, and cloud is shown in Fig. 3.

VI. THEORETICAL SUPPORT FOR EMOTION COMMUNICATION PROCESS

In this paper, the law of user's state transition is analyzed from the perspective of Markov state transition theory, that is, the state transfer rule of the terminal user from the beginning of communication to the end of receiving feedback is defined, as the support of the reliability of emotion transfer in the system.

Suppose the current state (i.e., the emotion, here called the state) of User-A is E_a ($E_a \in$, *where Emotion is the set of the considered emotions*). After receiving the state E_b of the User-B pushed from the cloud, the terminal ECTD-A deduces the corresponding feedback action according to its own emotion-action inference engine, and changes the state of User-A after performing the action from $E_a \rightarrow E_i$, where $E_i \in Emotion$. That is, after the action feedback, User-A can arrive at any of the state of the emotion set with the probability p_i from the initial state. For any kind of feedback action $Action = \{x_1, x_2, x_3, \ldots, x_n\}$, the state transition matrix is as follows:

$$\underset{action(i)}{P} = \begin{array}{c} \\ E_a \\ E_b \\ E_c \\ \vdots \\ E_n \end{array} \begin{array}{c} \begin{array}{ccccc} E_a & E_b & E_c & \cdots & E_n \end{array} \\ \begin{pmatrix} p_{aa} & p_{ab} & p_{ac} & \cdots & p_{an} \\ p_{ba} & p_{bb} & p_{bc} & \cdots & p_{bn} \\ p_{ca} & p_{cb} & p_{cc} & \cdots & p_{cn} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ p_{na} & p_{nb} & p_{nc} & \cdots & p_{nn} \end{pmatrix} \end{array} \tag{1}$$

Where n is the number of elements of the emotion set, and the constraint condition of the probability matrix is that the sum of each row of the matrix is 1, namely, $\sum_{j \in Emotion} p_{ij} = 1$.

It can be seen from the transition matrix that any state can be reached from the current state, whereas the emotion communication is a process of continuous state transition between link establishment and termination. The newly arrived state will again arrive at the next state with the new probability according to the feedback, but also get the new state transition matrix. Taking the seven states of emotion set as an example, the state transition diagram is shown in Fig. 4. The state S1 in the graph may reach any state of the emotion set after the execution of the feedback, the same is true for the states S2-S7. Then in an emotion communication process, the state transition law is as follows:

$$\begin{aligned} &Action(x_1, x_2, x_3, \ldots, x_n) \rightarrow P_1 \\ &Action(x_1, x_2, x_3, \ldots, x_n) \rightarrow P_2 \\ &\quad\vdots \\ &\quad\vdots \\ &Action(x_1, x_2, x_3, \ldots, x_n) \rightarrow P_n \end{aligned} \tag{2}$$

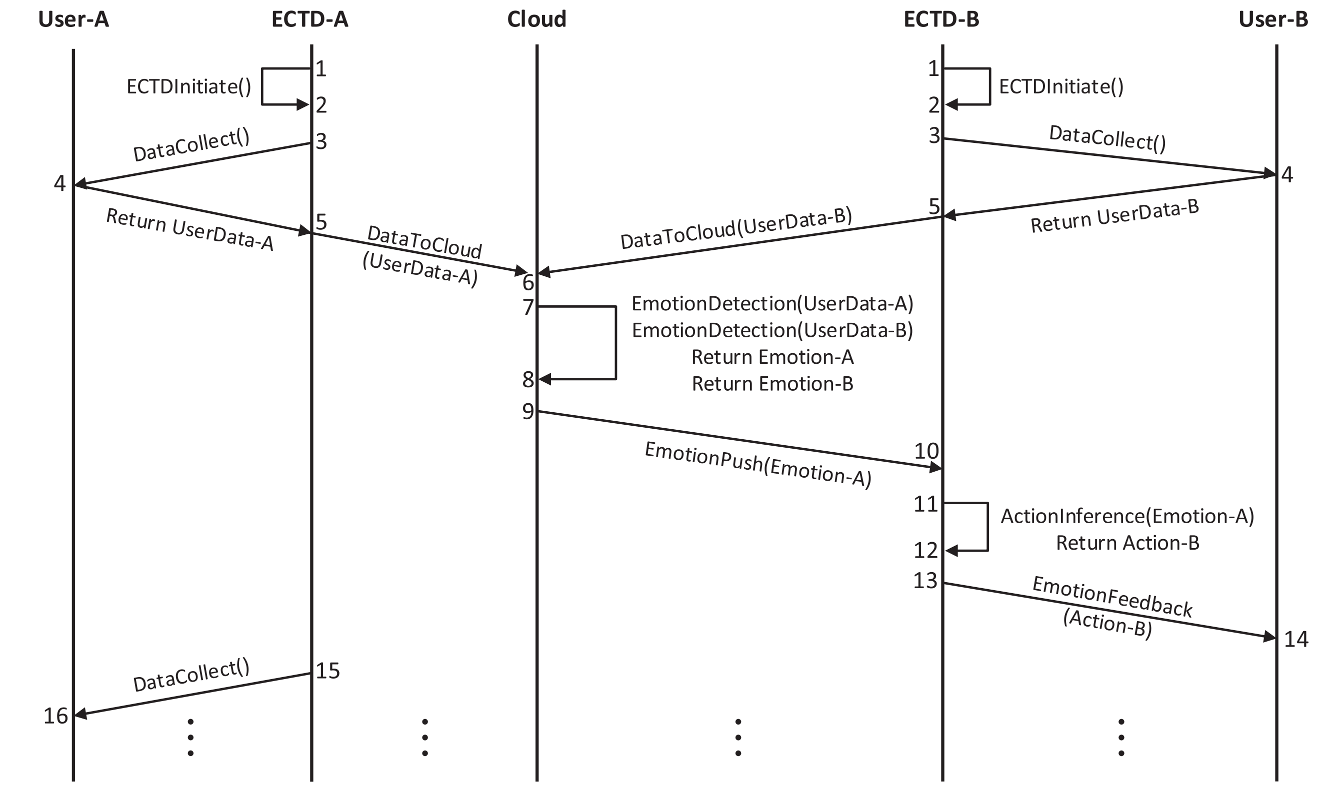

FIGURE 3. The execution sequence of emotion command set.

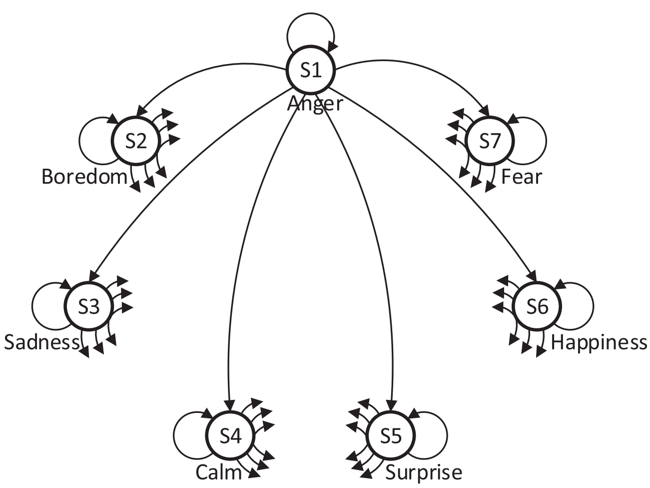

FIGURE 4. The illustration of state transition.

We establish the mapping from $Action = \{x_1, x_2, x_3, \ldots, x_n\}$ to the probability matrix, that is:

$$f(Action(x_1, x_2, x_3, \ldots, x_n)) \rightarrow P \tag{3}$$

According to the above derivation, the user's final status is:

$$p^T(n) = p^T(0)P_1P_2 \cdots P_n = p^T(0)\prod_{i=1}^{n} P_i \tag{4}$$

The above mentioned Markov state transition model indicates that the initial state in the system will be stabilized to a final state with the influence of external factors. The Markov state transition model is the theoretically support for the feasibility of our emotion communication system: the initial state of the user entering the emotion communication system is to arrive at any of the states of the emotion set with an uncertain probability under the feedback of the other side. Multiple times of feedback in a communication process can stabilize the user's state, that is, such a cycle process of analyzing user's emotion and taking emotion feedback possesses the theoretical basis, which also proves that the emotion communications of our research is reasonable.

VII. SPEECH EMOTION COMMUNICATION SYSTEM FOR PILLOW ROBOTS

We use the pillow robot [42] to carry out emotion communications, as for the two advantages of comfort and convenience. As can be seen in the Markov state transition law mentioned above, the user's emotions are in a dynamic change, high-delay emotion communications will give users a very bad experience. Considering a series of operations of user data collection, user data transmission, cloud platform emotion analysis and robot feedback during the whole communication process, this system regards real-time as an important reference index of the communication quality in order to guarantee the communication QoS [43]. The smart clothing, robot and cloud platform used in this system are all from the EPIC laboratory, and the cloud platform adopted is

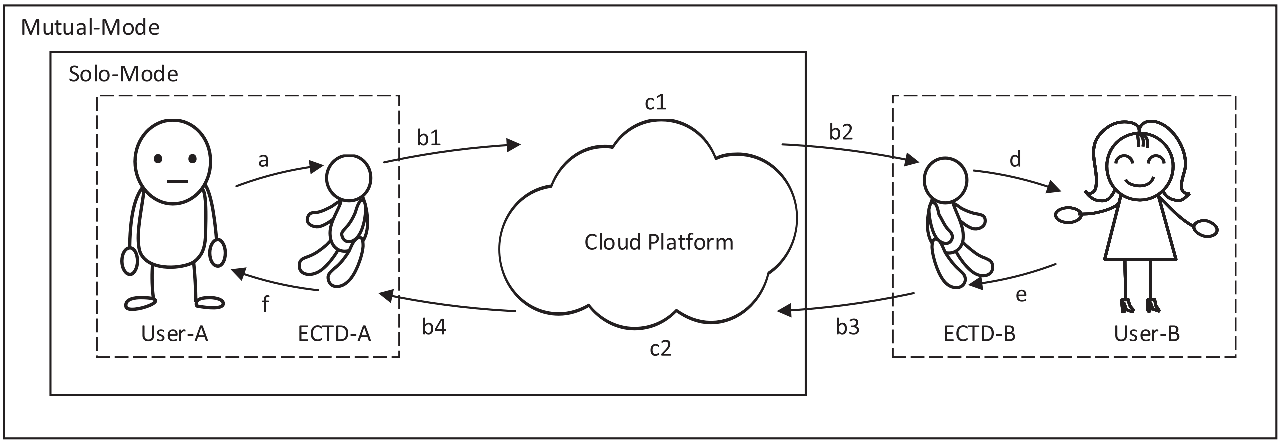

FIGURE 5. The two modes of emotion communication system.

Inspur SmartData Appliance [44] with the computing power of 84-core CPU and 336G RAM.

TABLE 4. Scenes for voice database recording.

Number	Description
1	The little silver lamp was on the table in the lab
2	There is a matrix theory class tomorrow afternoon
3	There is an important meeting to be attended tomorrow morning
4	Today I'm going to the library to borrow a book
5	I will go mountain climbing with Ace in the countryside this weekend
6	I waited for him half an hour in that place
7	Tomorrow I will go far away with him to do the experiment
8	The teacher in hearsay is coming to our lab
9	I'll take a train to Beijing on business tomorrow
10	I scored eighty in the exam last week
11	I have not received that email so far
12	I met Bob again when I having the lunch

A. VOICE DATABASE

Voice data can well reflect the user's current emotional state, the speech feature extraction based on specific speech emotion database can effectively identify the user's emotional state. In this system, the accuracy of user emotion recognition becomes the premise of emotion communication. Based on this, we built our own voice database. The basic situation of the voice database is, a total of 24 volunteers were selected from laboratory to participate in the voice recording, including 13 males, 11 females, aged between 20 to 30 years old. Based on seven states in the given set of emotion (see Section IV), recording was done in twelve scenes as shown in Table 4, each volunteer expresses each scene with seven kinds of emotions. The total database has 1176 voice recordings, the average voice length is 3s.

B. TESTING SCENARIOS

We designed an application scenario where the emotion communications between the child User-A at home and the mother User-B on business. The mother is away on business, the child at home misses his mother very much, and he wants to get as mother-like comfort by hugging pillow robot. By hugging their pillow robot respectively, the two sides complete the establishment of communication link, and pillow robot begins to collect the user's voice data, and real-time transmission to the cloud platform. The cloud platform learns the voice database through the machine learning model [42] and analyzes emotions based on the user's real-time voice data. The child's emotion Emotion-A is sent to the mother's pillow robot ECTD-B. When the mother understands the current emotion of the child, she usually adopts the positive voice content and tone to cheer the child up. The same mother's emotions through the cloud platform analysis are pushed to the child pillow robot ECTD-A, then ECTD-A triggers its own feedback behavior. Through the mapping of emotion to feedback, ECTD-A achieves comfort for the child. The specific comfort behavior in the action set is the simulation of the embrace, heart rate, body temperature and play music, etc. When the communication is completed, the communication process is interrupted by releasing the pillow robot. The scene in the two modes of the emotion communication system is shown in Fig. 5, in which the small solid line encloses the communication closed-loop in Solo-Mode mode, and the large solid line encloses the communication closed-loop in Mutual mode.

TABLE 5. The process of emotion transition of child.

The Number of Times	Emotion of Child	Emotion of Mother
1	Sadness	Calm
2	Boredom	Calm
3	Calm	Calm
4	Happiness	Calm

C. REAL-TIME ANALYSIS

We carried out experiments based on the above mentioned test scenarios, in which the pillow robot is equipped with a 4G LTE enabled communication module, to achieve the communication among the current user's robot and the remote user's robot and the cloud platform [45] by renting the LTE service of the telecom operator. The communication

TABLE 6. Average communication delay in two modes.

Mode	Data Volume	Voice Duration (V)	Transmission Time (T)	Analysis Time (A)	Command Sending Time (S)	Transmission + Analysis + Sending (T+A+S)	Total Time (V+T+A+S)
Solo-Mode	194 KB	2824ms	110ms	256ms	3ms	370ms	3194ms
Mutual-Mode	160KB+213KB	2553ms+3016ms	151ms+174ms	220ms+274ms	3ms	823ms	6393ms

bandwidth is 20Mbps. The first experiment is an emotion care process between child and robot with distance of 1 m. The second experiment conducts emotion communications between a mother and her remote child. In the second case, the emotion transfer results of the child being comforted is shown in Table 5.

We consider communication delay of the two modes in emotion communication system, that is, the communication delay in the Solo-Mode and the Mutual-Mode. Specifically, in the Solo-Mode, the closed-loop communication delay between a single user and the robot is counted. In the Mutual-Mode, the closed-loop communication delay between the two users is counted. Comparison of data volume and time-consuming in each stage in two modes, as shown in Table 6. The first column attribute in the table is the communication mode; The second column attribute is voice data volume collected by the pillow robot; The third column attribute is voice duration, where Solo-Mode corresponds to the **a** segment delay in Fig. 5, and Mutual-Mode corresponds to the **a**+**e** segment delay in Fig. 5; The fourth column attribute is the transmission delay of the data during the whole communication process, where Solo-Mode corresponds to the **b1**+**b4** segment delay in Fig. 5, and Mutual-Mode corresponds to the **b1**+**b2**+**b3**+**b4** segment delay in Fig. 5; The fifth column attribute is the delay of cloud emotion analysis, where Solo-Mode corresponds to the **c1** segment delay in Fig. 5, and Mutual-Mode corresponds to the **c1**+**c2** segment delay in Fig. 5; The sixth column attribute is the time in which the cloud sends command to the pillow robot, corresponding to the **f** segment delay in Fig. 5; The seventh column counts the sum delay of transmission, cloud analysis and command sending; The last column is the total delay of the communication process.

We carried out 30 experiments in two modes. The time delay in Solo-Mode is shown in Fig. 6, and the time delay in Mutual-Mode is shown in Fig. 7. From the two figures we can see that, the transmission delay, cloud analysis delay and the sum delay of transmission, cloud analysis and command sending are changed steadily in two modes. Analysis of average communication delay in Table 6 shows that the main factor that affect the real-time performance of the system communication is cloud analysis delay, which is 256ms in the Solo-Mode, and $220 + 274 = 494$ms in the Mutual-Mode. However, the overall communication link delay (T+A+S) remains in the order of 10^2ms, which is 370ms in Solo-Mode and 823ms in Mutual-Mode. And the standard deviation of overall communication link delay (T+A+S) is 24ms in

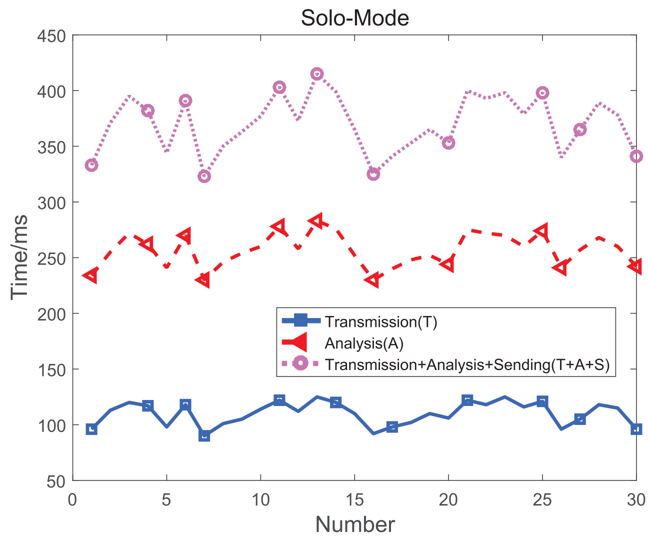

FIGURE 6. Time delay in Solo-Mode.

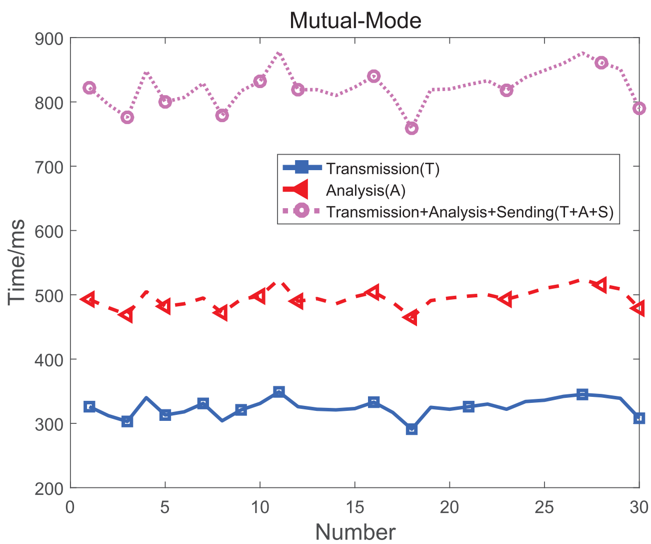

FIGURE 7. Time delay in Mutual-Mode.

Solo-Mode and 27ms in Mutual-Mode, which shows the feasibility and validity of the real emotion communication between two parties in Mutual-Mode.

VIII. CONCLUSION

In this paper, the current HMI system is summarized according to the different methods of emotion recognition. At the same time, the gap between the existing HMI system and the emotion communication system proposed in this paper is compared from the communication point of view. Then, this paper defines the emotion communication system, and the key technologies needed in the system are summarized. Based on the four tasks: collection, communication, analysis and feedback of emotion that need to be done in the emotion communication system, this paper puts forward the architecture of the emotion communication system. At the same time, in order to meet the synchronization requirements of the communication for both the two sides when the emotion is transmitted as a kind of multimedia information, this paper proposes an emotion communication protocol, which provides a high-level reliable support for the realization of emotion communications. Finally, this paper analyzes the real-time performance of a speech emotion communication system based on the pillow robot, and highlights the feasibility and effectiveness of realizing the emotion communications. In the future, we would consider the case of multi-party emotion communications in our system.

Acknowledgment

This work was supported by the National Science Foundation of China under Grant 61572220.

REFERENCES

[1] Y. Li, W. Dai, M. Qiu, and Z. Ming, "Privacy protection for preventing data over-collection in smart city," *IEEE Trans. Comput.*, vol. 65, no. 5, pp. 1339–1350, May 2016.

[2] K. Gai, M. Qiu, H. Zhao, L. Tao, and Z. Zong, "Dynamic energy-aware cloudlet-based mobile cloud computing model for green computing," *J. Netw. Comput. Appl.*, vol. 59, pp. 46–54, Jan. 2016.

[3] K. He, J. Chen, R. Du, Q. Wu, G. Xue, and X. Zhang, "DeyPoS: Deduplicatable dynamic proof of storage for multi-user environments," *IEEE Trans. Comput.*, vol. 65, no. 12, pp. 3631–3645, Dec. 2016, doi: 10.1109/TC.2016.2560812.

[4] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," *Found. Trends Human-Comput. Interaction*, vol. 1, no. 3, pp. 203–275, Feb. 2007.

[5] H. Yan, J. M. H. Ang, Jr., and A. N. Poo, "A survey on perception methods for human–robot interaction in social robots," *Int. J. Social Robot.*, vol. 6, no. 1, pp. 85–119, Jan. 2014.

[6] D. Mccoll, A. Hong, N. Hatakeyama, G. Nejat, and B. Benhabib, "A survey of autonomous human affect detection methods for social robots engaged in natural HRI," *J. Intell. Robot. Syst.*, vol. 82, no. 1, pp. 101–133, Apr. 2016.

[7] C.-N. Anagnostopoulos, T. Iliou, and I. Giannoukos, "Features and classifiers for emotion recognition from speech: A survey from 2000 to 2011," *Artif. Intell. Rev.*, vol. 43, no. 2, pp. 155–177, Feb. 2015.

[8] T. Williams, P. Briggs, N. Pelz, and M. Scheutz, "Is robot telepathy acceptable? investigating effects of nonverbal robot-robot communication on human-robot interaction," in *Proc. 23rd IEEE Int. Symp. Robot Human Interactive Commun.*, 2014, pp. 886–891.

[9] F. Agrafioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 102–115, Jan./Mar. 2012.

[10] D. K. Kim, J. Kim, E. C. Lee, M. Whang, and Y. Cho, "Interactive emotional content communications system using portable wireless biofeedback device," *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1929–1936, Nov. 2011.

[11] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *Proc. 22nd Annu. Int. Conf. Mobile Comput. Netw.*, 2016, pp. 95–108.

[12] S. Yohanan and K. E. MacLean, "Design and assessment of the haptic creature's affect display," in *Proc. Int. Conf. Human Robot Interaction*, Mar. 2011, pp. 473–480.

[13] S. K. A. Hossain, A. S. M. M. Rahman, and A. E. Saddik, "Haptic based emotional communication system in second life," in *Proc. IEEE Int. Symp. Haptic Audio-Vis. Environ. Games*, Oct. 2010, p. 1.

[14] C. Busso *et al.*, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proc. 6th Int. Conf. Multimodal Interfaces*, 2004, pp. 205–211.

[15] Z.-T. Liu *et al.*, "A multimodal emotional communication based humans-robots interaction system," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6363–6368.

[16] Z. Yang and S. S. Narayanan, "Analysis of emotional effect on speech-body gesture interplay," in *Proc. INTERSPEECH*, 2014, pp. 1934–1938.

[17] M. Chen, Y. Ma, J. Song, C.-F. Lai, and B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 825–845, Oct. 2016.

[18] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, "Wearable 2.0: Enable human-cloud integration in next generation healthcare system," *IEEE Commun.*, vol. 55, no. 1, pp. 54–61, Jan. 2017.

[19] G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, and R. Jafari, "Enabling effective programming and flexible management of efficient body sensor network applications," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 1, pp. 115–133, Jan. 2013.

[20] G. Fortino, S. Galzarano, R. Gravina, and W. Li, "A framework for collaborative computing and multi-sensor data fusion in body sensor networks," *Inf. Fusion*, vol. 22, pp. 50–70, Mar. 2015.

[21] H. Ghasemzadeh, P. Panuccio, S. Trovato, G. Fortino, and R. Jafari, "Power-aware activity monitoring using distributed wearable sensors," *IEEE Trans. HumanŰMach. Syst.*, vol. 44, no. 4, pp. 537–544, Aug. 2014.

[22] S. Galzarano, R. Giannantonio, A. Liotta, and G. Fortino, "A task-oriented framework for networked wearable computing," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 621–638, Apr. 2016.

[23] Y. Ma, C. H. Liu, M. Alhussein, Y. Zhang, and M. Chen, "LTE-based humanoid robotics system," *Microprocess. Microsyst.*, vol. 39, no. 8, pp. 1279–1284, Nov. 2015.

[24] L. Lei, Z. Zhong, K. Zheng, J. Chen, and H. Meng, "Challenges on wireless heterogeneous networks for mobile cloud computing," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 34–44, Jun. 2013.

[25] D. Tian, J. Zhou, Y. Wang, Y. Lu, H. Xia, and Z. Yi, "A dynamic and self-adaptive network selection method for multimode communications in heterogeneous vehicular telematics," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3033–3049, Dec. 2015.

[26] M. Chen, Y. Hao, M. Qiu, J. Song, D. Wu, and I. Humar, "Mobility-aware caching and computation offloading in 5G ultradense cellular networks," *Sensors*, vol. 16, no. 7, pp. 974–987, 2016.

[27] J. He, D. Wu, X. Xie, M. Chen, Y. Li, and G. Zhang, "Efficient upstream bandwidth multiplexing for cloud video recording services," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 10, pp. 1893–1906, Oct. 2016.

[28] Y. Zhang, M. Chen, D. Huang, D. Wu, and Y. Li, "iDoctor: Personalized and Professionalized Medical Recommendations Based on Hybrid Matrix Factorization," *Future Generat. Comput. Syst.*, vol. 66, pp. 30–35, Jan. 2017.

[29] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, no. 1, pp. 947–959, Sep. 2014.

[30] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys Tut.*, vol. 17, no. 3, pp. 1738–1760, 3rd Quart., 2015.

[31] D. Tian, J. Zhou, Z. Sheng, and Q. Ni, "Learning to be energy-efficient in cooperative networks," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2518–2521, Dec. 2016.

[32] M. Chen, Y. Qian, S. Mao, W. Tang, and X. Yang, "Software-defined mobile networks security," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 729–743, Oct. 2016.

[33] M. Qiu, Z. Ming, J. Li, K. Gai, and Z. Zong, "Phase-change memory optimization for green cloud with genetic algorithm," *IEEE Trans. Comput.*, vol. 64, no. 12, pp. 3528–3540, Dec. 2015.

[34] S. Mao and S. S. Panwar, "A survey of envelope processes and their applications in quality of service provisioning," *IEEE Commun. Surveys Tut.*, vol. 8, no. 3, pp. 2–19, 3rd Quart., 2006.

[35] M. Gong, S. E. Midkiff, and S. Mao, "Design principles for distributed channel assignment in wireless ad hoc networks," in *Proc. IEEE Int. Conf. Commun.*, Seoul, South Korea, May 2005, pp. 3401–3406.

[36] M. Chen, Y. Hao, S. Mao, D. Wu, and Y. Zhou, "User intent-oriented video QoE with emotion detection networking," in *Proc. IEEE Globelcom*, Washington, DC, USA, Dec. 2016, pp. 1552–1559.

[37] Z. He, S. Mao, and T. Jiang, "A survey of QoE-driven video streaming over cognitive radio networks," *IEEE Netw.*, vol. 29, no. 6, pp. 20–25, Nov./Dec. 2015.
[38] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, and L. Hanzo, "Quality-of-experience assessment and its application to video services in LTE networks," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 70–78, Feb. 2015.
[39] Y. Xu and S. Mao, "A survey of mobile cloud computing for rich media applications," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 46–53, Jun. 2013.
[40] M. Chen, Y. Zhang, Y. Li, M. M. Hassan, and A. Alamri, "AIWAC: Affective interaction through wearable computing and cloud technology," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 20–27, Feb. 2015.
[41] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Inf. Fusion*, vol. 35, pp. 68–80, May 2017.
[42] M. Chen *et al.*, "CP-robot: Cloud-assisted pillow robot for emotion sensing and interaction," in *Proc. Int. Conf. Ind. IoT Technol. Appl.*, 2016, pp. 81–93.
[43] J. Chen, K. He, Q. Yuan, G. Xue, R. Du, and L. Wang, "Batch identification game model for invalid signatures in wireless mobile networks," *IEEE Trans. Mobile Comput.*, to be published, doi: 10.1109/TMC.2016.2604820.
[44] M. Liu *et al.*, "Mobile cloud platform: Architecture, deployment and big data applications," in *Proc. Cloudcomp*, 2016.
[45] X. Deng, D. Wu, J. Shen, and J. He, "Eco-aware online power management and load scheduling for green cloud datacenters," *IEEE Syst. J.*, vol. 10, no. 1, pp. 78–87, Mar. 2016.

MIN CHEN (M'08–SM'09) was an Assistant Professor with the School of Computer Science and Engineering, Seoul National University (SNU), from 2009 to 2012. He was a Post-Doctoral Fellow with SNU. He was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of British Columbia. He is currently a Professor with the School of Computer Science and Technology, Huazhong University of Science and Technology (HUST). He is also the Director of the Embedded and Pervasive Computing Laboratory. He has authored over 260 paper publications, including over 120 SCI papers, over 50 IEEE Transactions/Journal papers, six ISI highly cited papers and one hot paper. He has authored a book on IoT, *OPNET IoT Simulation* (HUST Press, 2015), a book on 5G, Software Defined 5G Networks (HUST Press, 2016) and a book on big data, *Big Data Related Technologies* (2014) with Springer Series in Computer Science. His Google Scholars Citations reached over 8200 with an h-index of 45. His top paper was cited over 900 times. His research focuses on Internet of Things, mobile cloud, body area networks, emotion-aware computing, healthcare big data, cyber physical systems, and robotics. He received the Best Paper Award from the IEEE ICC 2012 and the Best Paper Runner-up Award from QShine 2008. He is currently a Guest Editor of the IEEE NETWORK and the *IEEE Wireless Communications Magazine*. He is the Co-Chair of the IEEE ICC 2012-Communications Theory Symposium and the Co-Chair of the IEEE ICC 2013-Wireless Networks Symposium. He is the General Co-Chair of the 12th IEEE International Conference on Computer and Information Technology and the Mobimedia 2015. He is the General Vice Chair of the Tridentcom 2014. He is a Keynote Speaker of CyberC 2012, Mobiquitous 2012, and Cloudcomp 2015.

PING ZHOU received the B.S. degree in computer science and technology from the University of Jinan, China, in 2015. And she recommended for admission to the Huazhong University of Science and Technology as a master candidate in September of the same year. Learning more than seven months later, she applied for pursuing Ph.D. degree in June 2016. She is currently a 1st year Ph.D. student of the Embedded and Pervasive Computing Laboratory for learning and scientific research. Her research interests include: emotional cognitive computing, multimedia communication.

GIANCARLO FORTINO (SM'12) received the Laurea (B.S. and M.S.) and Ph.D. degrees in computer engineering from the University of Calabria, Rende, Italy, in 1995 and 2000, respectively. He has been with the Department of Informatics, Modeling, Electronics, and Systems, University of Calabria, since 2006, where he is currently an Associate Professor of Computer Engineering. He holds the Scientific National Italian Habilitation for Full Professor, and is also an Adjunct Full Professor of Computer Engineering with the Wuhan University of Technology, Wuhan, China, in the framework of high-end foreign experts in China and an Adjunct Senior Researcher with the Italian National Research Council. He is the Co-Founder and the CEO of SenSysCal S.r.l., a spin-off of the University of Calabria, where he is involved in the advanced applied Research and Development of IoT systems. He has authored over 300 publications in journals, conferences, body area networks and books. His current research interests include distributed computing, wireless sensor networks, software agents, IoT technology, and cloud computing. He is currently an Associate Editor of the IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, the IEEE TRANSACTIONS ON Human-Machine, IEEE Sensors Journal, IEEE Access, *Information Fusion*, *Engineering Application of Artificial Intelligence*, the *Journal of Network*, and *Computer Applications*. He is the chair of the Italian Chapter of the IEEE SMC society.