International Journal of Engineering Sciences Paradigms and Researches: (Vol. 23, Issue 01) and
(Publishing Month: August 2015)

Multimodal Emotion Detection to Enhance Communication

Priyabrata Nayak[1], Subhashree Sukla[2] and Saswati sahoo[3]
[1,3]Associate Professor, Department of Computer Science Engineering, Gandhi Institute For Technology (GIFT), Bhubaneswar
[2]Assistant Professor, Department of Computer Science Engineering, Gandhi Engineering College, Bhubaneswar

Publishing Date: August 27, 2015

Abstract— In current society, there are numerous robots made for various purposes, including manufacturing, cleaning, therapy, and customer service. Other robots are used for enhancing H2H communication. In this research, we proposed a robotic system which detects the user's emotions and enacts them on a humanoid robot. By using this robotic avatar, users with motor disabilities are able to extend their methods of communication, as a physical form of expression will be added to the conversation.

I. INTRODUCTION

In the United States, approximately 17,000 new spinal cord injury cases occur each year [1]. When the number is calculated globally, there is a significant population of those unable to move their bodies. Non-verbal communication which includes facial expressions, body posture, gestures, eye movement, and so on, complements verbal communication in human to human interaction. People who cannot create gestures to express their thoughts and feelings are void of this crucial communication component. Thus, the way in which people with motor disabilities communicate could vastly change with the use of personal robotic avatars which act as replacements for their bodies. By sensing physical movement within the same atmosphere, the users' conversation partners are hypothesized to better understand the current state of the user, as physical engagement increases between the user with motor disabilities and conversation partner. On the other hand, according to research by Mehrabian [2], under the circumstances where three of verbal, vocal and body language communications are sending contradictory content, 55% of communication lies within body language. This indicates that the accuracy of matching avatars' body language with users' utterances and emotions is important.

Our goal was to design a robotic system which correctly and efficiently identifies the user's real-time emotion and creates abundant robotic expressions which can compensate for the user's lack of body expressions. We coined this system the "Ex-Amp Robot", short for Expression Amplification.

The requirements that meet the expectations for this system are summarized in the following list:

TABLE I

REQUIREMENTS AND RELATED PUBLICATIONS

	Autonomous	Emotion Detecting	Expressive Movements	Communication Medium	Avatar	Human-like Appearance	Multi-modal EmotionSensing
Affdex							
Facerig							
Orihime							
Jimmy							
Ex-Amp							

1) The user's changing emotions must be correctly detected and identified
2) Robot must perform expressive body movements and gestures that suit the identified emotions
3) The timing of robot reactions must not disturb the conversation
4) The conversation partner should be able to understand the robot's actions

II. RELATED PUBLICATIONS

Table I organizes the related publications according to the necessary requirements for our research. In this way, our system requires an efficiently merged combination of technologies. By taking away aspects of each requirement category from related publications, we sought to create a well rounded, highly sophisticated expressive robotic avatar.

A. Emotion Detection

Currently, the most common method of emotion detection is facial expression recognition. In addition to academic research [3], companies such as EYERIS [4] and Affectiva [5] provide their facial expression detection technologies to the public. Other emotion detection methods include speech frequency analysis and retrieve other speech parameters including pitch, duration, and energy [6] [7] in combination with facial cues. Other researchers have published their works on creating a deep learning model for speech emotion recognition [8]. In our work, we have decided to incorporate a commercial emotion detection system and sentiment analysis to create a higher level detection technology.

B. Robot Movement

For the Ex-Amp Robot to be effective, body language must be expressed by the robot in a natural, human-like way. Therefore, it is important that the robot can create a full range of motion. Both NAO and Pepper [9] are humanoid in shape, and can be very expressive with their 20 motors on the arms and lower bodies. Although most parts of their bodies move flexibly, the face has no motor for movement, and is invariably constant. There are previous works of robot heads manufactured to create facial expressions which were successful at creating complex expressions according to simple artificial consciousness [10]; however, such robotic facial expressions are unnecessary for the Ex-Amp Robot because our hypothetical users are capable of creating facial expressions for themselves. On the other hand, Whitney and the team at Disney research have created an enhanced hybrid configuration of fluid actuators that allow robot arms to create fluid motions [11]. Ideally, we would like to combine this technology to create movements that are as fluid as this Disney Jimmy Robot, to mimic the natural movements of the human body.

C. Robots as Communication Medians

OriHime eye [12] is a personal avatar that enhances H2H communication by creating emotion expressions with movement. This product is targeted towards users with intractable nerve diseases who are unable to move. While it is very similar to the overall idea of our research, it is different in two aspects. First, OriHime eye does not detect the user’s emotions autonomously. Whereas our system detects and performs actions on its own, OriHime eye moves only according to user commands. Our focus is for the robot to detect, understand, and create action without the users’ specific effort. The second aspect in which our system varies from OriHime eye is the use of a humanoid robot. We believe that using a robot that is human-like in form would allow both the user and their conversation partners to sense the reality of a personal avatar, as though the robot is the user’s body.

III. PILOT STUDY

The Ex-Amp Robot uses facial expression recognition to detect the user’s real-time emotions. This proposed system is specifically targeted towards people who are physically paralyzed from the neck down but have full capacity of speaking as well as creating precise facial expressions. The objective of the system is to enhance users’ communication with a conversation partner by expressing real-time emotion. Figure 1 shows the hypothetical target user and the scenario in which the Ex-Amp robot system is used.

We created our first prototype of the Ex-Amp Robot system in [13]. We connected Pepper to Affectiva’s facial expression detection engine [14] for detecting emotions. We used onomatopoeic expressions because they are natural and easy to understand for Japanese speakers, as Japanese is especially rich in phonetic expressions of sensory experiences [15]. We chose to use Pepper as our robotic body since its form and range of motions were most suitable within robots that were readily available to us.

For the experiment, we visited an adult day-care facility in Ayase, Tokyo. The elderly people at this retirement facility were chosen as experiment candidates because they were close to our ideal target users. As a result of the preexperiment interview, most answered that they felt unable to move their bodies freely to express themselves using body language as compared to before when they were younger and more active. Although there is a gap between our hypothetical users and experiment candidates, they were the closest prospects to our requirements.

A. Prototype System

Figure 2 shows the system configuration diagram of the system. Affectiva’s Affdex SDK for iOS allows detection of emotional metrics on a mobile platform. Its fast and lightweight system allows emotions to be analyzed all within the device [16]. The system connects this facial expression detection application and Pepper with a MQTT pub/sub protocol [17]. Its continuous and instant service allows for data signals to be transferred quickly enough to suit the speed of a natural conversation.

In short, the list below describes the system flow:

1) Camera device uses facial expression detection to identify users’ emotion
2) Detected emotion data is converted into emotion keyword
3) Keyword is published onto designated MQTT channel
4) The robot subscribed to the designated MQTT channel receives message protocol
5) Robot performs gestures according to the received keyword

For robot gestures, we utilized the pose and movement library for Pepper created by Yoshimoto Kogyo [18]. We used a wide variety of refined poses for each detected emotion. With Pepper's built-in tablet and LED lighting system, we extended the emotion expressions. We added ambient signs that represent emotion by setting the LED lights to the appropriate emotion color and displaying emoticons on the tablet. Pepper's human-like size allowed conversation partners to feel realism in the robot's expressions. In total,

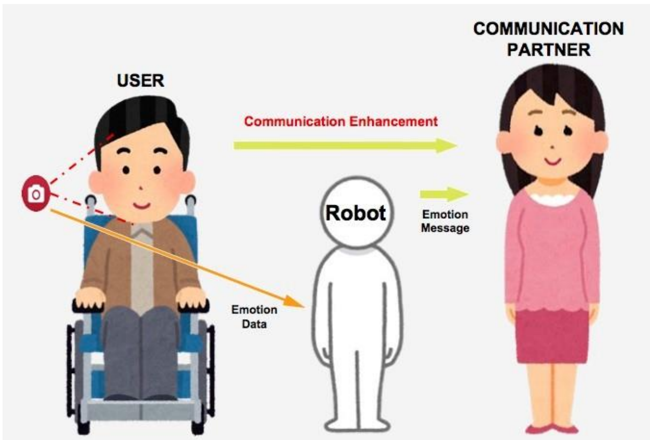

Fig. 1. Usage Scenario

we created 6 sets of different movements and speech for the 5 emotions.

B. Prototype Experiment at an Elderly Care Facility

For the experiment, we had two experimentees sit across from each other, one as the user and one as the conversation partner. Figure 3 displays the experimentation setup. We had the two converse freely for 5 minutes. Every minute, we gave the experimentees a topic to talk about, to facilitate an efficient conversation. A happy topic, sad topic, surprising topic, irritating topic, and a disgusting topic were suggested every minute to advocate the 5 emotions. After 5 minutes, we had the two experimentees answer their corresponding surveys. We did a total of 5 experiments across two days.

This experiment conducted at the elderly care facility was published as part of the 4th International Conference on Human-Agent Interaction [19]. Survey results showed that both the users and conversation partners believed joy was detected and expressed by the system at high ratings. Most conversation partners were able to distinguish the expressions for the two types of joy. Although results varied, most users were in agreement that the robot was able to correctly express their emotions. Most conversation partners agreed that they were able to understand the emotions that the robot expressed as well. All users and conversation partners agreed that they enjoyed their conversations using the robot system. However, it appeared that conversation partners enjoyed the conversation more, being the receivers of the robot expressions.

C. Findings from Prototype

From this experimentation, we realized that the design of the system needed improvement. The iPhone application which showed the user's face on the screen distracted some users. Oppositely, some had no interest in the iPhone camera, and kept going out of frame. Furthermore, there were several false recognitions, especially in disgust due to the low angle of the camera and the elderly people's facial features.

Experiment results with elderly members showed that the system was successful at amplifying users' happy emotions.

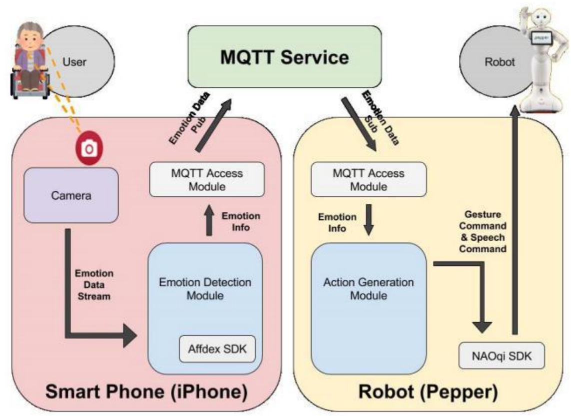

Fig. 2. System configuration diagram of prototype

However, during a sad conversation context, the system was inefficient at detecting the users' sad emotions. Not limited to just elderly users, people in general show smiles while conversing with others, even when the content of the conversation is negative. In such instances, we found that multi-modal features were necessary for emotion recognition to decrease the amount of false detections. If the system could understand the content of the conversation and adjust for the gap between users' facial expressions and what the user is saying, it would greatly improve the reliability of the robotic system.

IV. MULTI-MODAL EXTENSION

To improve the system, we sought to add a sentiment analysis feature that would pick up on the context of the user's conversation. We chose to use speech recognition and extract the polarity of the detected words, by matching them with a polarity dictionary. Understanding the context of the conversation is important for decreasing false recognitions. If the robot could filter and decide what facial expressions do not match the conversation context, the system could hypothetically decrease the amount of false actions and avoid negatively disrupting the conversation. In this way, our target with making the Ex-Amp Robot multi-modal was to improve the system to fit the user's conversations more naturally.

We also improved our original system by changing the location of the camera, while still using the Affdex software for facial expression detection. Since the iPhone in the previous design was a hindrance to the users, we refined the physical system design to have less distractions. We changed the camera source from an iPhone camera to a smaller external webcam. We used Affectiva's Mac OS X SDK for the facial recognition instead.

For conversational context recognition, we used Julius, an open source software for speech recognition [20]. This software recognizes Japanese vocabulary in real-time, at high speed. We matched the detected words with the Japanese

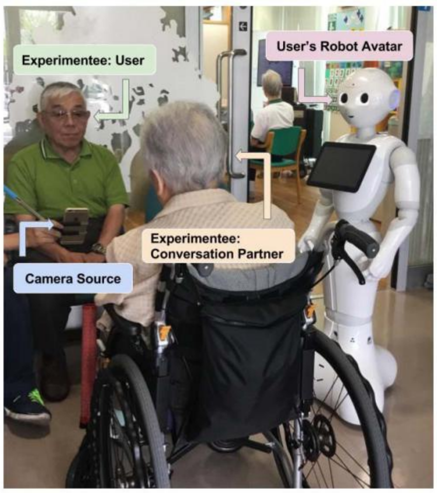

Fig. 3. Experimentation setup at elderly care center

Sentiment Polarity Dictionary by Inui and Okazaki Labratory at Tohoku University [21]. We used a verb polarity dictionary and a noun polarity dictionary for our system.

Figure 4 shows the new system configuration diagram. When the detected emotion exceeds a set threshold, an action command is sent to the robot. The connecting system design is the same as described in the previous prototype. For each emotion: joy, anger, surprise, sadness, and disgust, we created three sets of robot actions, totaling up to 15 available emotion actions. One of the three robot actions are chosen at random to express the emotion.

At the same time, speech recognition is carried out on a different program module, using a microphone, speech recognition software, and polarity dictionaries. Upon speech recognition, detected words are matched with the dictionaries to identify the polarity of the word. This information helps the system understand the positivity or negativity of the conversation, helping it to understand the context of the conversation.

Julius distinguishes an end of a detection period when there is a pause in speech. Polarity is extracted in the form of polarity points. Each detected word is matched with the dictionary, and the points are added up. The polarity points are used to manipulate the threshold of the emotion value. For example, when +1.0 polarity point is calculated, the threshold for joy decreases and the threshold for sadness, anger, and disgust which suggest negativity are increased. Polarity points are cumulative throughout the entire conversation, and can be either positive or negative. Understanding the negativity or positivity of a conversation helps the system to distinguish true emotion detection.

Based on the emotion information, the robot expresses emotions using pattern variations of LED colors, tablet display, movement, speech, and voice. While in the previous version, we displayed pictures of abstract faces expressing the emotions, we change the tablet display in the new version to solid colors to induce a more ambient effect. Figure 5 shows three of the emotions and robot reactions.

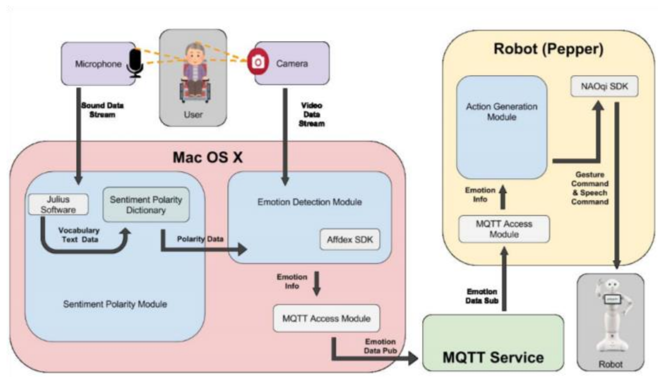

Fig. 4. Multi-modal system configuration diagram

V. COMPARATIVE EXPERIMENT AND DISCUSSION

We conducted a comparative experiment in which the experimentees conversed without the robot to determine the validity of survey answers. Moreover, we purposefully arranged the experiment so that the users were unable to move, and not just imagining the immobility.

A. Experimentation Method

We restricted the user experimentees from being able to move by taping duct tape around their arms and legs (with their approval). By enforcing immobility, they were made to stand in our hypothetical user's shoes and desire to express.

After a 7 minute conversation with no robot intermediate, we had experimentees converse with the Ex-Amp Robot for another 7 minutes. We extended the conversation time to 7 minutes to gather more data within the experiment. However, we minded a shorter set of time to sustain the level of conversation. Duncan's face-to-face interaction research also suggests that seven minute interactions are appropriate for a natural conversation period [22]. Figure 6 shows the experimentation scene with and without the robot.

After both sessions were finished, we had both the user and conversation partner answer survey questions. A total of 10 people cooperated in the experiment.

The post experiment questionnaire was designed in a 5 point likert scale, similar to our experiment at the elderly care center. A list of questions were asked to be answered on a scale of 1 to 5, where 1 was 'Strongly Disagree' and 5 was 'Strongly Agree'. We designed the questionnaire to elicit subjective, honest answers by anchoring and adjusting parts of the survey questions [23] [24]. Since it is proven that visual survey design influences user's tendencies to score higher or lower [25], we decided to provide only the questions, instructions, and a pen and paper.

B. Experimentation Results

Table II and III shows the survey questions and results for the user and conversation partner answered after an

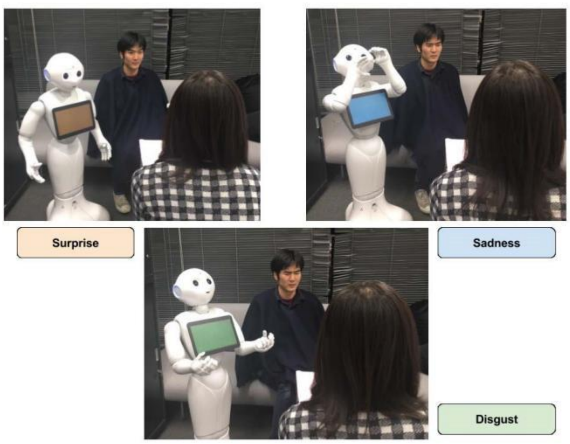

Fig. 5. Improved robot action examples

experiment without the system. Table IV and V summarizes survey questions and results answered by the user and conversation partner after using the Ex-Amp Robot. The average answer rating and standard deviations are calculated for each question.

From the post experiment questionnaire, we saw that users were able to express happiness with the robot. While only 30% of respondents strongly agreed that they were able to express happiness without the robot, 50% of respondents strongly agreed that they were able to express happiness with the robot. We also see an increase in high ratings for surprise. We could not see as big of changes in expressing anger and sadness within the conversation.

From the conversation partner's point of view, the robot seemed to have a significant effect on expressing joy. While only 20% of respondents strongly agreed that they were able to read the user's happy emotions without the robot, 60% of respondents strongly agreed that there were able to understand the user's emotions with the help of the robot system. Surprise was also better received with the robot. There was a subtle increase in reading anger and sadness.

Overall, users answered that they were able to expand the realm of their expression using the robot, and conversation partners were able to feel that the limitation of the user's expressiveness was reduced with the Ex-Amp system. The majority felt that the robot functioned when they wanted it to move, and did not disrupt the conversation.

C. Effect of Sentiment Polarity Analysis

We examined the sentiment polarity analysis (SPA) feature to see its effect on emotion detection. Three experimentees spoke 5 minutes using the Ex-Amp Robot with the SPA turned on, and another 5 minutes with the SPA turned off. Experimentees were told to speak on a positive topic for both turns. They were not notified of the purpose of the test. The results are as shown in Table VI. The numbers represent the amount of times a reaction of the emotion was made by the Ex-Amp Robot.

According to Human-Machine Interaction Network on Emotion (HUMAINE) [26], disgust, sadness, and anger are negative emotions, joy is positive, and surprise can have any valance. When we calculated using this as a measure, we can

Fig. 6. Experimentation scene with and without robot

see that, for Ex. 1, the amount of negative reactions decreased from 6 to 2. For Ex. 2, 7 negatives emotion reactions became 2, and for Ex. 3, 1 became 0. From this experiment, we can see that polarity sentiment analysis can have a significant effect on the emotion detection.

D. Discussion

The comparative experiment results showed that there was significant difference in the conversations with and without the robot, for both the users and the conversation partners. After using the robot, 60% of the users agreed that they were able to expand their realm of expressions (Table IV). This is most likely a result of the robot's undistuptive design which received good ratings on Q8 of Table IV. It was important to design the robot to react quickly and express simply.

Since the amount of body language use varies between cultures [27], we found that the amount of restriction necessary varied on the physical expressiveness of the person. Therefore, there was not a distinguishable difference in outcome of Q2 of Table II and Q6 of Table IV.

As with the last experiment, it is unusual that a person becomes extremely sad or angry within a 7 minute conversation. We did not force experimentees to express fake emotions deliberately just for the experiment to prove the system's use in a casual conversation. Therefore, we do not see high ratings for strongly agree on Q4 and Q5 in Table II, Q2 and Q3 in Table IV, Q2 and Q3 on Table III and on Q2 and Q3 on Table V. If the user were to use this system for a longer period of time, the range of emotions could expand. With that, the frequency of robot reactions may also need to be changed. Even if a user used the system for 7 minutes and did not feel that it was annoying, if the robot expressed as many emotions for 12 consecutive hours, the robot actions may be excessive. The current system allows for flexible personalization to match the specific user's needs. As a personal robot, it is configured so that the user can easily change parts such as the robot's verbal expressions and frequency. Moreover, we also realized that some users may not want their negative expressions to be amplified, and only amplify the good. For some, it may make them even more angry if the robot expressed anger besides them. On the other hand, it may alleviate their anger for other users. This system has more than one use, and it may match with certain users more than others.

Lastly, there were some opinions regarding whether or not the vocal expressions of the robot were necessary. There may be no need for the robot to speak if the users can speak for themselves. However, it was necessary with Pepper because a sudden movement of a silent large size robot was uncanny. If the system was made on a slightly smaller robot, structuring the system without vocal effects may be an option.

Key takeaways from the addition of verbal sentiment analysis feature to the Ex-Amp Robot system were the following: the system is able to recognize the conversation context, false recognition (over-detection) of user emotion decreased, and the timing of robot reactions were relatively non-disruptive to the conversation.

TABLE IISURVEY QUESTIONS AND THE RESPONSES ANSWERED BY USERS WITHOUT THE ROBOT ON A 1 TO 5 SCALE

Question	5	4	3	2	1	NA	Ave.	SD
1. Were you able to express yourself being physically restricted?	10	40	30	20	0	0	3.4	0.96
2. Did you feel stressed being unable to move?	40	50	10	0	0	0	4.3	0.67
3. Were you able to express happiness?	30	30	30	10	0	0	3.8	1.03
4. Were you able to express anger?	0	0	30	50	20	0	2.1	0.73
5. Were you able to express sadness?	0	20	20	30	30	0	2.3	1.15
6. Were you able to express surprise?	10	20	20	50	0	0	2.9	1.10

5: Strongly Agree, 4: Agree, 3: Neither Agree nor Disagree, 2: Disagree, 1: Strongly Disagree,
N/A: Non-Applicable. *Number ratings shown in percentages of total poll

TABLE III

SURVEY QUESTIONS AND THE RESPONSES ANSWERED BY CONVERSATION PARTNERS WITHOUT THE ROBOT ON A 1 TO 5 SCALE

Question	5	4	3	2	1	NA	Ave.	SD
1. Did you understand the user’s happy emotions?	20	70	10	0	0	0	4.1	0.56
2. Did you understand the user’s angry emotions?	10	10	30	40	0	10	2.9	1.05
3. Did you understand the user’s sad emotions?	10	10	50	30	0	0	3.0	0.94
4. Did you understand the user’s surprised emotions?	20	30	50	0	0	0	3.7	0.82
5. Did you feel a limitation on the user’s expressiveness?	30	60	10	0	0	0	4.2	0.63
6. Did the physical restriction make the conversation unnatural?	30	40	30	0	0	0	4.0	0.81

Question	5	4	3	2	1	NA	Ave.	SD
1. Were you able to express happiness?	50	40	10	0	0	0	4.4	0.69
2. Were you able to express anger?	0	10	20	40	20	10	2.2	0.97
3. Were you able to express sadness?	0	20	40	10	20	10	2.6	1.11
4. Were you able to express surprise?	20	20	40	10	10	0	3.3	1.25
5. Did the realm of your expression expand using the robot?	0	60	20	20	0	0	3.4	0.84
6. Did the robot lessen the burden you felt being physically restricted?	0	40	40	20	0	0	3.2	0.78
7. Did the robot move when you wanted it to move?	30	40	10	20	0	0	3.8	1.13
8. Did the robot disrupt the conversation?	0	10	20	30	40	0	2	1.05
9. Did you enjoy the conversation with the robot intermediate?	70	20	10	0	0	0	4.6	0.69

5 point scale indicates the same answer as Table II.

*Number ratings shown in percentages of total poll

TABLE IVURVEY QUESTIONS AND THE RESPONSES ANSWERED BY USERS WITH THE ROBOT ON A 1 TO 5 SCALE

5 point scale indicates the same answer as Table II. *Number ratings shown in percentages of total poll

TABLE V

SURVEY QUESTIONS AND THE RESPONSES ANSWERED BY CONVERSATION PARTNERS WITH THE ROBOT ON A 1 TO 5 SCALE

Question	5	4	3	2	1	NA	Ave.	SD
1. Did you understand the user’s happy emotions?	60	40	0	0	0	0	4.6	0.51
2. Did you understand the user’s angry emotions?	10	30	30	10	0	20	3.5	0.92
3. Did you understand the user’s sad emotions?	0	40	30	20	0	10	3.2	0.83
4. Did you understand the user’s surprised emotions?	40	40	20	0	0	0	4.2	0.78
5. Did the robot help amplify the user’s emotions?	30	60	10	0	0	0	4.2	0.63
6. Did the addition of physical movement in the conversation make the conversation more natural?	50	40	10	0	0	0	4.4	0.69
7. Were you able to enjoy the conversation with the robot intermediate?	70	20	10	0	0	0	4.6	0.69

5 point scale indicates the same answer as Table II. *Number ratings shown in percentages of total poll

The combined Ex-Amp Robot system collected good impressions on experimentees, and was a meaningful system to express emotions. Although the system is capable of expressing negative expressions such as anger and sadness, it is evident that most people try not to show anger and sadness in a conversation. Therefore, what is truly needed to improve this system may be a more abundant library for expressing finely classified positive expressions.

E. Research Limitations

Although we were able to collect valuable feedback from experiments, we were unable to conduct our experiment on our true hypothetical users who have motor disabilities. However, with this, we were able to see that this system is not limited to use by users with motor disabiities, but also those who desire an expressive personal robot.

In terms of system technicalities, cultural influences may have played a role in emotion detection via facial expressions. Although Affdex uses data points gathered from 3.9 million faces from over 75 countries [14], the result is averaged into a global expression standard. Studies show that Japanese facial expressions are limited compared to expressions made by American people [28]. This gap between a global system and Japanese users may have been an obstacle for our experiments.

For future work, we would like to improve the multi-modal emotion recognition system as well as create a bigger library of robot actions for specific emotions. This system can be used in a remote setting as well.

TABLE VI

SENTIMENT POLARITY ANALYSIS ACROSS THREE EXPERIMENTEES

Ex.1 Emotion	Without SPA	With SPA
Joy	22	20
Surprise	4	3
Sadness	0	0
Disgust	6	2
Anger	0	0
Ex. 2 Emotion	Without SPA	With SPA
Joy	17	20
Surprise	5	9
Sadness	0	1
Disgust	6	1
Anger	1	0
Ex. 3 Emotion	Without SPA	With SPA
Joy	14	16
Surprise	4	3
Sadness	0	0
Disgust	1	0
Anger	0	0

We, human beings, can recognize another person's emotions by looking at facial cues, body language, vocal tones, and speech content. However, upon research, methods other than facial expression recognition were still either systematically underdeveloped or unavailable for wide use. Combining radio signal emotion recognition technology [29], vocal tone analysis [30] or contactless vital biosensor [31] with our system could increase the accuracy of emotion detection. Just as a human being processes others' emotions with our 5 senses and knowledge, the system must detect emotion from various angles to extract correct information.

VI. CONCLUSION

In this research, we proposed a robotic system that can express detected emotions of users using visual sensing, sound recognition, and a humanoid robot. Through experimentation, we saw that users of the system held positive impressions on the speed and effect of robotic expressions within the conversation. We see that with improvements on the variations of robot expressions and emotion detection, the system can be used more widely to suit anyone who wishes to have their emotions expressed to aid communication.

REFERENCES

[1] National Spinal Cord Injury Statistical Center. The contagious leader: impact of the leader's mood on the mood of group members, group affective tone, and group processes.

[2] A Mehrabian. Silent messages: Implicit communication of emotions and attitudes (2nd ed.), Wadsworth Pub Co, June 1980.

[3] M. Shin, M. Kim, and D. S. Kwon. Baseline cnn structure analysis for facial expression recognition. In 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN), pp. 724-729.

[4] Emovu emotion recognition software.

[5] Affectiva. Emotion Recognition Software and Analysis. http://www.affectiva.com/.

[6] K. M. Kudiri, A. M. Said, and M. Y. Nayan. Human emotion detection through speech and facial expressions. In 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), pp. 351-356.

[7] R. V. Darekar and A. P. Dhande. Enhancing effectiveness of emotion detection by multimodal fusion of speech parameters. In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 3242-3246, March 2016.

[8] X. Zhou, J. Guo, and R. Bie. Deep learning based affective model for speech emotion recognition. In 2016 Intl IEEE Conferences on Ubiquitous Intelligence Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), pp. 841-846.

[9] SoftBank Robotics. Who is nao? https://www.ald.softbankrobotics.com/en/cool-robots/nao.

[10] J. Takeno, K. Mori, and Y. Naito. Robot consciousness and representation of facial expressions. 2008 3rd International Conference on Sensing Technology, 2008.

[11] John P. Whitney, Matthew F. Glisson, Eric L. Brockmeyer, and Jessica K. Hodgins. A low-friction passive fluid transmission and fluid-tendon soft actuator. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014.

[12] Orylab. Bunshin Robot. http://orihime.orylab.com/.

[13] Ai Kashii, Kazunori Takashio, Hideyuki Tokuda, Robot Avatar for Enhancing the Communication of Physically Disabled Humans, 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 26-31, Teachers College Columbia University, NY, USA

[14] White paper: Exploring the emotion classifiers behind affdex facial coding. Technical report. http://studylib.net/doc/13073324/whitepaper-exploring-the-emotion-classifiers-behind-affd...

[15] Chisato et al Asaga. Onomatopedia: Onomatopoeia online example dictionary system extracted from data on the web. Progress in WWW Research and Development, pp. 601-612, 2008.

[16] Affectiva. Mobile. http://www.affectiva.com/solutions/mobile/.

[17] CloudMQTT - A globally distributed MQTT broker. CloudMQTT Hosted broker for the Internet of Things. https://www.cloudmqtt.com/.

[18] Yoshimoto Robotics Labratory. http://www.yoshimoto.co.jp/yrl/en/.

[19] Ai Kashii, Kazunori Takashio, Hideyuki Tokuda, Ex-Amp Robot: Physical Avatar for Enhancing Human to Human Communication, International Conference on Human-Agent Interaction (IHAI) Singapore.

[20] Julius development team. What's julius? http://julius.osdn.jp/.

[21] Yuji Matsumoto. Masahiko Higashiyama, Kentaro Inui. Learning sentiment of nouns from selectional preferences of verbs and adjectives. Proceedings of the 14th Annual Meeting of the Association for Natural Language Processing, pp. pp.584-587, 2008.

[22] Starkey Duncan and Donald Winslow Fiske. Face-to-face interaction: research, methods, and theory. L. Erlbaum Associates, 1977.

[23] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse. The impact of eye gaze on communication using humanoid avatars. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01, pp. 309-316, New York, NY, USA, 2001. ACM.

[24] Hunter Gehlbach and Scott Barge. Anchoring and adjusting in questionnaire responses. Basic and Applied Social Psychology, Vol. 34, No. 5, p.417-433, 2012.

[25] Don A. Dillman Stern, Michael J. and Jolene D. Smyth. Visual design, order effects, and respondent characteristics in a self-administered survey. Survey Research Methods, Vol. 1, p. 128-38, 2007.

[26] Humaine emotion annotation and representation language (earl).

[27] Julius Fast. Body Language. M. Evans, 1970.

[28] Ayano Kimura. Kenkyuyo Nihonjin hyojo shigeki no sakusei to sono rinshoteki tekiyo. Kazama Shobo, 2013.

[29] Mingmin Zhao, Fadel Adib, and Dina Katabi. Emotion recognition using wireless signals. In Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking, MobiCom '16, pp. 95-108, New York, NY, USA, ACM.

[30] Watson developer cloud. http://www.ibm.com/watson/developercloud/t one-analyzer.html.

[31] MIO Corporation. Contactless Vital Biological Sensor.
https://www.mio-corp.co.jp/sensor/