Helping People with Social Anxiety Disorder to Recognize Facial Expressions in Video Meetings

Jieyu Wang[1]*, Abdullah Abuhussein[2], Hanwei Wang[3], Tian Qi[4], Xiaoyue Ma[5], Amani Alqarni[6], Lynn Collen[7]
Department of Information Systems, St. Cloud State University, St. Cloud, USA[1, 2, 7]
University College London, London, UK[3]
University of San Francisco, San Francisco, USA[4]
CNBD Applied Research Institute Co. Ltd, Beijing, China[5]
Information Technology and Security Department, Jazan University, Jazan, Saudi Arabia[6]

***Abstract*—According to previous research on social anxiety disorder (SAD) and facial expressions, those with SAD tend to view all faces as portraying negative emotions; thus, they are afraid of chatting with others when they cannot understand the real emotions being communicated. The advancement of facial recognition technology has given people opportunities to get a more precise emotional estimation of facial expressions. This study aims to investigate the practical effects of apps that detect facial expressions of emotion (e.g., AffdexMe) on people with SAD when communicating with other people through video chatting. We conducted empirical research to examine whether facial emotion recognition software can help people with SAD overcome the fear of chatting with others in video meetings and help them understand others' emotions to reduce communication conflicts. This paper presents the design of an experiment to measure participants' reactions when they video-chat with others using the AffdexMe application and to interview participants to get in-depth feedback. The results show that people with SAD could better recognize the emotions of others using AffdexMe. This results in more reasonable responses and better interaction during video chats. We also propose design suggestions to make the described approach better and more convenient to use. This research shed a light on the future design of emotion recognition in video chatting for people with disabilities or even ordinary users.**

Keywords*—Social phobia/social anxiety disorder; video meeting; facial expression recognition; emotion recognition; empirical research

I. Introduction

In modern times, mobile facial recognition applications have become more and more popular. Many mobile phone manufactures have developed facial recognition capabilities, including companies such as Samsung, Huawei, and Vivo. Facial recognition is technology which has the capability to identify or verify a person using a digital image or a video frame from a video source [1]. This process is often completed by comparing facial features with images stored in a database. Emotional recognition has been broadly studied by the computer vision community. It has been developed based on the application of facial recognition technology and aims to define individuals' emotions by analyzing their facial expressions. The currently available applications of facial recognition include area access control systems, checking attendance systems, facial recognition phone unlock, and even logging onto banking apps and similar security applications. This widespread use of facial recognition has made many computing devices faster, safer, and easier to use in many ways. However, the application of emotion recognition using facial expressions is far less popular or widespread compared to other applications for facial recognition technology, in part because emotional recognition facial software faces challenges with recognizing moving objects, continuous detection of objects, unpredictable actions, and similar programming difficulties.

Emotion recognition is an interdisciplinary research field that is becoming very important for intelligent communication between humans and computers. It has been used in many areas such as gaming and entertainment, surveillance, robotics and many more. The potential for using video- and audio-aided emotion detection to identify and reduce the severity of psychological and mental disorders has recently gained attention from the research community. This is mainly attributable to advancements in artificial intelligence and video recording technologies.

According to the American Psychiatric Association [1], social phobia is a "persistent fear of one or more situations in which the person is exposed to possible scrutiny by others and fears that he or she may do something or act in a way that will be humiliating or embarrassing." It is estimated to be one of the most prevalent psychiatric disorders in the world [2], and is also known as social anxiety disorder (SAD). Generally, people with SAD feel embarrassed or fearful when interacting with others, especially strangers, and worry about being judged negatively [3]. They are more likely to misunderstand others' emotions in social situations.

Facial and emotional recognition tools and techniques can provide promising solutions for SAD patients. Many existing software tools can be used to analyze video content and recognize actions within them. These applications are being widely used in emotion recognition. AffdexMe [4] by Affectiva is one such app that analyzes and responds to facial expressions of emotion in real-time using the built-in camera on iOS or Android devices.

In this paper, we design and develop a solution that uses AffdexMe for emotion recognition to help people with SAD better communicate with others on video chatting platforms. Our proposed solution aims to enable SAD patients to

*Corresponding Author.

recognize the emotions of people with whom they communicate, and to help them to respond in appropriate ways. A usability testing experiment was conducted on participants with SAD for qualitative observation on approach effectiveness and usage behaviors.

The reminder of this paper is organized as follow. Section 2 presents the literature review and examines what others have done to tackle mental disorders using video-aided emotion recognition. Section 3 presents the study approach, experiment setup, and scenarios considered to demonstrate how AffdexMe is used to assist SAD patients. In Section 4, we present the evaluation method and results of the study. Discussions and study limitations are presented in Section 5. Finally, we conclude this work and lay out our future research plans in Section 6.

II. Related Work

A. Technologies of Facial Recognition Systems

Facial expressions are able to communicate certain emotions such as anger, fear, or anxiety that people may have when they encounter certain situations [5]. Different facial features of experiment subjects were collected and successfully analyzed for the emotions of task subjects. Under certain conditions, the 2D facial recognition system used to recognize a flat face might not be useful for real-time, video-based facial emotion recognition. However, traditional still image-based facial recognition systems are not accurate enough when analyzing motions, even though motions can help to better analyze the emotions behind individuals' facial expressions in terms of psychology and physiology studies [6].

Yet, not all video-based facial emotion recognition applications are accurate and mature enough to be effective for widespread use. For example, Hirt et al. (2018) applied FaceReader to examine students' emotions of interest and boredom by video recording their facial expressions while reading, and later compared these emotional recognition results to student self-reports [7]. They found that the results from student self-reports were significantly different from those of the facial emotion recognition software. They concluded that until this discrepancy could be better analyzed and accounted for, the emotion recognition technology could not be formally applied to educational practice.

However, educational researchers have also argued the importance of facial emotion recognition to people with disabilities, such as, for example, preschool children with autism [8]. They emphasized the importance of intervention tools such as facial emotion recognition games to help children with autism understand facial expressions, which may lead them to exhibit more appropriate social behaviors. Researchers presented a facial emotion recognition game to children with autism as an early intervention. They also presented suggestions for the future design of a facial expression detection game.

B. Facial Expression and Social Anxiety Disorder (SAD)

Facial expression, as a type of nonverbal communication, plays a significant role in conveying speakers' messages and involves at least 65% of the total meaning of a conversation [16], and yet people with social phobia have difficulties reading others' facial expressions and emotions. Researchers have found that such people have longer reaction times for almost all emotions and lower accuracy in emotion recognition [10,11]. Specifically, people with SAD tend to view various facial expressions as conveying negative feedback, whether they truly do or not.

Based on the research of Lange et al. [10], people with SAD tend to interpret negative feedback when they encounter the neutral faces of other people. They believe that a neutral face actually represents a negative emotion. With further questionnaires and experiments, researchers found that those with social anxiety disorder even viewed all facial symbols, including negative, positive, and neutral, as significantly threatening. Therefore, assistive technologies are required for this group of people so that they can communicate appropriately with others and actively join society instead of being gradually isolated from society.

C. Open-Source Facial Emotion Recognition Application: AffdexMe

Yu and Wang (2016) proposed an advanced, real-time, monocular, video-based facial expression recognition system that enabled more precise recognition of multiple variables [12]. An online statistical appearance model (OSAM) was used to track the facial movements of the participant, as shown in Fig. 1.

Based on different models created by the computer for analyzing the geometrically normalized facial mesh, the system was able to draw conclusions about participants' emotions when making different facial expressions. Examples are shown in Fig. 2 and 3.

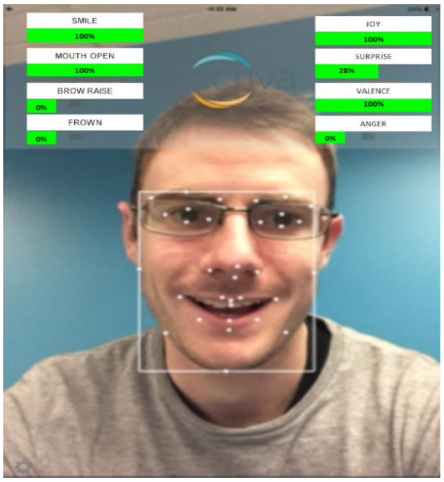

Fig. 1. A Sample Revised Interface of AffdexMe shows Six Emotions that can be Recognized by the Technology. Screenshots have been revised to Clarify the Font used in AffdexMe.

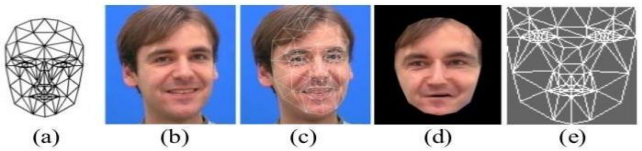

Fig. 2. The Process of Facial Recognition [5].

Fig. 3. Examples from the Facial Expression Recognition System [5].

McDuff et al. have also indicated that facial expressions can express certain emotions not only through still expressions but also through facial movement [13]. Their work provides the mechanism that enables facial expression recognition to work when analyzing users' emotions communicated by their facial movements. A Facial Action Coding System was used in their study to analyze facial expressions through analyzing the various actions according to the dataset from Affdex. This system automatically coded facial expressions when participants used the video chat platform Skype and analyzed the emotion of participants.

The traditional dataset for analyzing users' facial expressions has been manually collected; thus, it was difficult to classify diverse facial actions to indicate the emotions of users. However, the internet-based framework used by AffdexMe has enabled researchers to collect around 1.8 million online media videos of spontaneous facial responses [14]. Through active learning, the facial recognition system can learn 100 different scales of facial expression based on the movements of various facial features. This active learning method can help the computer system to recognize even more precise movements which are not similar to more common expressions such as smiles or closed eyes [14].

In sum, people with SAD tend to misunderstand facial emotions when they communicate with others, which makes it necessary to help them accurately classify the emotions behind different facial symbols such as positive or negative expressions. Therefore, a facial expression recognition system is needed for helping such people understand the true emotions behind facial features, and therefore reduce misunderstandings in communication. Our study aims to apply the AffdexMe facial emotion recognition technology to design a study to help people with SAD communicate with others appropriately. We also hope that our results can be generalized to the current global situation, in which people feel more isolated than ever before during the COVID-19 pandemic.

III. Experimental Setup and Study Approach

Our goal throughout this study is to help SAD patients to correctly interpret the emotions behind the facial expressions of others on video chat platforms. To achieve this goal, two groups of participants were asked to video chat with friends over Skype [15] for 10 minutes each. The first group included five ordinary participants (i.e., those not diagnosed with or symptomatic of SAD). The second group included five participants who self-reported themselves to have social anxiety disorder. Participants of both groups were undergraduate students at one university. Three of the participants with SAD were female and two were male students; all of these participants had never previously heard about AffdexMe, had not used similar emotion recognition software before, and seldom video chatted with others prior to the study. The five ordinary participants with whom the SAD participants chatted agreed to participate in the experiment (Table I). Also, participants' informed consent was obtained according to research protocols. This involved a statement of consent to participate in the study and to use their photos for research purposes. We reduced study bias by requiring the SAD-free participants not to intentionally control their facial expressions.

During the 10-minute Skype call, an observer (Fig. 4) holding a smartphone sat with a SAD patient to capture the call with the SAD-free participant. Observer's smartphones had AffdexMe installed and running during the 10-minute video call. The AffdexMe app was used to detect six real-time emotions including joy, sadness, fear, anger, contempt, and disgust. AffdexMe also showed the percentage of certainty for every emotion detected. When the AffdexMe detected emotions, the observer verbally said a short sentence that described the emotion to the SAD patient. For example, when the application detected "sadness," the observer said: "He/she shows sadness."

The observer orally describing the emotions detected by AffdexMe was used to simulate the new function we are proposing here. In the future, we will develop this feature to automatically play recorded sentences that describe the detected emotions. Fig. 5 presents screenshots of AffdexMe output presenting a participant showing joy, disgust, and anger, respectively.

We interviewed each of the SAD patients after the video chatting session and asked them about their experiences when using this application in order to receive their feedback, determine if they encountered any problems, and see if they had any suggestions for improvement.

TABLE I. Participants

	Total Participants (10)	
	Ordinary Participants	***Participants with SAD***
Video chat	Five	Five

Fig. 4. Observer using AffdexMe to Detect Emotions in Skype Video-chat with Ordinary Participants.

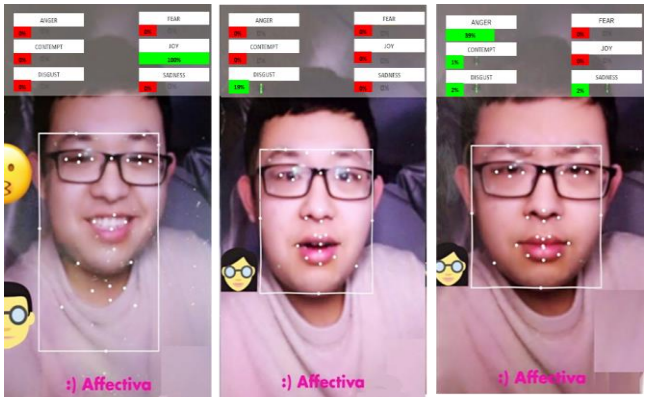

Fig. 5. AffdexMe Outputs on an Ordinary Participant to Help a Participant with SAD Understand Emotions. Screenshots have been Revised to Clarify the Font used in AffdexMe.

IV. Results

During participants' chat sessions, AffdexMe successfully and continuously detected emotions with high success rates. After the 10-minute Skype call, we interviewed the SAD-participants and asked them about their experiences. The following questions were asked of participants:

- How did you like the application?
- What changes will you suggest making the application more effective?
- Do you think you would use an application like this regularly in your video calls? Why?

All participants reported that this application was helpful and supported them in reducing misinterpretation of emotions during video calls. Some participants reported that before using this application they could not accurately recognize others' emotions, as they often thought others felt unpleasant or otherwise negative emotions. One participant also reported that the application enabled them to comprehend others' emotions better and that they do not feel as nervous as they did before when video chatting with others. One of the participants expressed that she always found difficulties in interpreting emotions and often mistakenly regarded others' facial expressions as uncomfortable or angry; however, using the application, she was told that the person with whom she was video chatting showed no or neutral emotions, rather than negative ones.

When considering the application's functionality, a participant mentioned that the observer sound describing the detected emotion was too long (Table II). For example, when the participant-caller showed joy, the observer said, "She shows joy (Fig. 6)," but, before the sound ended, the participant-caller had already turned to another emotion. Therefore, the observer sometimes could not report the emotion promptly. The participant suggested that the sound could be simple and short, like "joy," or could read faster. Another participant thought the sound was helpful because in that case, he did not need to look at the application all the time. Nevertheless, the participant said that the sound was a little disruptive, as it distracted attention when listening to what the caller was talking about. Participants reported that they had to pay attention to what the application said as well as what their conversation was about. All participants reported that it is more likely that they would download and regularly use a similar application if available for smartphones in the future.

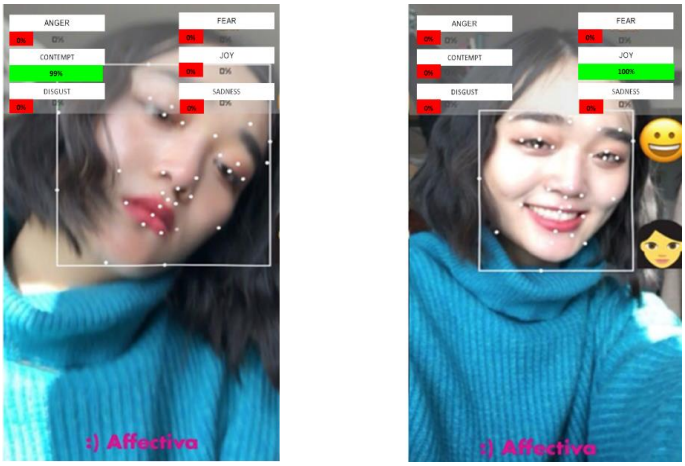

Fig. 6. AffdexMe Output showing One Emotion of a Person Called by One of the Participants. Screenshots have been Revised to Clarify the Font used in AffdexMe.

The designed solution to overcome the emotion recognition challenges encountered by SAD patients when communicating with others using video chatting apps was likable and beneficial according to our empirical study results. However, the study design may have subjected the results to bias. First, the number of participants interviewed in this work was relatively small and the findings may not be generalizable. However, we did try to overcome this limitation by providing more in-depth results about the subjects' perspectives on the research questions by interviewing participants rather than simply asking them to fill out a survey. This helped in being able to interpret the participants' facial gestures to get more reliable results. Additionally, the observer voice while describing the emotions detected by AffdexMe may have negatively impacted the cognitive performance of the SAD patients participating in this study. This issue can be resolved by recording different voices to describe the emotions and enable users to choose from a list of desired characters or possibly even languages.

TABLE II. Experimental Feedback

Participants' Feedback	
Emotion	***Functionality***
• Comprehended others' emotions better • reduced misinterpretation of emotions • calmed nervousness	• Should shorten the voice reminders' length • Voice reminders' distraction

V. Discussion and Limitations

Our study aims to help people with SAD to communicate effectively with others via facial emotion recognition technologies. Our observation and interview results suggested that this experimental design using facial emotion recognition helped the participants appropriately video chat with other participants. While only a limited amount of facial emotion

recognition research has been done to help people with SAD [7,8], the positive results of our study shed light on future related research design. As modern societies' environments are becoming more stressful due to factors including climate change disasters or pandemics such as COVID-19, this kind of individual assistive technology is needed to help people with SAD to adapt to the changing world. The results of our design can be generalized to other user groups, and even to ordinary people who need help when dealing with challenging communication methods.

Even though our design may help people with SAD to overcome their fears when interacting with others using video chat, there are some limitations in this study. For instance, we only recruited five participants with SAD to run the usability testing of this design. However, our sample size was based on the suggestion of Nelson (1993), who expressed that only five users can figure out 85% of the usability problems of a technology [9]. This study is only the first step to confirm that facial emotion recognition technologies can help people with SAD.

However, the practicability and conveniences of our design have many areas open to improvement. We suggest integrating the facial emotion recognition technology together with video social networks; in other words, inserting a facial emotion recognition function into video-chatting apps. Much existing social network software already has the function of video-chatting, such as Skype, WeChat, QQ, and Zoom. We plan to write codes to allow the emotion recognition function to operate during video-chatting, so that, when people are having video conversations, a small message box is placed at the corner on the screen. From the box, the user can read the computer's detection of emotion changes and the current emotion of the person being spoken to. We wish to design this function because video images can often be blurry, and a user may be able to see their communication partner but not be able to figure out the expressions on their faces, according to our experience of this study. With this new design, the user can better know the emotions of those with whom they chat and determine their own responses.

We also have specific suggestions for using facial emotion recognition to help people with visual disabilities. Such individuals cannot get any visual cues during communication, so they have more difficulty in interacting with others. We recommend adding a phonetic function to emotion recognition apps when people are video-communicating. When the talker shows strong emotion changes, the app will translate the emotion information into sounds. Therefore, visually-impaired people can get phonetic cues when communicating. This new function will allow them to better interact with others.

V. Conclusion

Our design demonstrated how a specific piece of facial emotion recognition technology helped people with SAD. Through experiments, observation, and interviews, we confirmed that facial emotion recognition as an assistive technology can help people with SAD to appropriately communicate with others via video chatting. This helps people with SAD better interact with others, allowing them to escape from isolation and be more involved in their societies. We also proposed design suggestions for further development of this technology. We recommended adding a small text box to show the emotion of the talker instead of only using time-consuming phonetic reminders, in accordance with the feedback of our participants.

In future studies, we first plan to create a prototype to integrate the facial emotion recognition technology within a video-chat software. Then we will run usability tests to see if users evaluate it as useful and effective. If possible, we will further program this integration and publish it to serve people who need emotion recognition aid when communicating with others.

Moreover, we believe that the sound function is even more significant for visually-impaired people who may also need emotion recognition help when video-chatting with others. They cannot see the faces of other participants, and so they often cannot figure out their emotions. With our future studies, visually-impaired people can freely video-chat with others. The integrated app will tell them the emotion of the other person so that they can better deal with their social interactions. Of course, the sound function is mainly designed for visually-impaired people, so ordinary people may feel this function is useless and distracting. Therefore, we will add a sound switch button on the small box. If some users need or prefer to have the sound function, they can click on it to make the sound function active (those with visual disabilities may require others' help at the beginning). If some users feel the sound is annoying, they can simply turn it off.

In sum, our study has shed light on future facial emotion recognition designs, since there has been only very limited research discussing how this technology can be applied to serve people up to this point. We hope this study can encourage more researchers to work on this topic and invent more promising designs and apps.

Acknowledgment

We thank all our participants for their contributions.

References

[1] Edition, Fifth. "Diagnostic and statistical manual of mental disorders." Am Psychiatric Assoc 21, 2013, pp. 591-643.

[2] Kessler, Ronald C., et al. "Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication." Archives of general psychiatry 62, 6, 2005, pp. 593-602.

[3] Lima, Laysa Karen Soares de, Erickson Duarte Bonifácio de Assis, and Nelson Torro. "Facial expressions and eye tracking in individuals with social anxiety disorder: a systematic review." Psicologia: Reflexão e Crítica 32, 2019.

[4] Experience It. Available online: https://www.affectiva.com/experience-it/ (accessed on 18 December 2020).

[5] K. Chengeta, "Comparative Analysis of Emotion Detection from Facial Expressions and Voice Using Local Binary Patterns and Markov Models: Computer Vision and Facial Recognition." Proceedings of the 2nd International Conference on Vision, Image and Signal Processing, New York, NY, USA, 2018, pp. 1–6.

[6] Yan, Yan, and Yu-Jin Zhang, "State-of-the-art on video-based face recognition." Encyclopedia of Artificial Intelligence. IGI Global, 2009, pp. 1455-1461.

[7] Hirt, F.S.; Moser, I.; Werlen, E.; Imhof, C.; Bergamin, P., "A Comparison of Students' Emotional Self-Reports with Automated Facial Emotion Recognition in a Reading Situation," In Proceedings of the Sixth International Conference on Technological Ecosystems for

Enhancing Multiculturality; Association for Computing Machinery: New York, NY, USA, October 24 2018, pp. 320–327.

[8] Christinaki, E.; Vidakis, N.; Triantafyllidis, G., "Facial Expression Recognition Teaching to Preschoolers with Autism: A Natural User Interface Approach," In Proceedings of the 6th Balkan Conference in Informatics; Association for Computing Machinery: New York, NY, USA, September 19 2013, pp. 141–148.

[9] Nielsen, Jakob, and Thomas K. Landauer. "A mathematical model of the finding of usability problems." In Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems, 1993, pp. 206-213.

[10] Oh, K.S.; Lee, W.H.; Kim, S.; Shin, D.W.; Shin, Y.C.; Lim, S.W., "Impaired Facial Expression Recognition in Patients with Social Anxiety Disorder: A Case-Control Study," Cognitive Neuropsychiatry 2018, 23, pp. 218–228.

[11] Bell, C.; Bourke, C.; Colhoun, H.; Carter, F.A.; Frampton, C.; Porter, R.J., "The Misclassification of Facial Expressions in Generalised Social Phobia:, Journal of Anxiety Disorders, 25, 2010, pp. 278–283.

[12] Yu, J. & Wang, Z. "A Monocular Video-Based Facial Expression Recognition System by Combining Static and Dynamic Knowledge", In Proceedings of the 9th International Conference on Utility and Cloud Computing; Association for Computing Machinery: New York, NY, USA, December 6 2016, pp. 353–357.

[13] McDuff, D.; Mahmoud, A.; Mavadati, M.; Amr, M.; Turcot, J.; Kaliouby, R. el AFFDEX SDK., "A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit", In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, May 7 2016, pp. 3723–3726.

[14] Senechal, T., McDuff, D., & Kaliouby, R., "Facial action unit detection using active learning and an efficient non-linear kernel approximation," In Proceedings of the IEEE International Conference on Computer Vision Workshops, 2015, pp. 10-18.

[15] Skype. Communication Tool for Free Calls and Chat. Available online: https://www.skype.com/en// (accessed on 20 December 2020).

[16] Wood, Julia T. "Interpersonal communication: Everyday encounters," Cengage Learning, 2015.