Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 3 (2015) 1926 – 1931

Procedia MANUFACTURING

6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

Useofelectroglottograph (EGG) to finda relationship between pitch, emotionand personality

Lu Hui[a,]*, Liu Hui Ting[a], Swee Lan See[b], Paul Yaozhu Chan[b]

[a]*National Junior College, 37 Hillcrest Rd, Singapore 288913*
[b]*Institute for Infocomm Research (I²R), A*STAR, 1 Fusionopolis Way, Singapore 138632*

Abstract

To ensure a more effective and applicable emotion recognition system with the use of an electroglottography (EGG) device, we investigated the relationship between pitch (an average frequency value obtained in terms of vibration around the neck), and the emotion felt, and if personality plays a role in the identification of emotions. Some of the motivations of this work include lie detection and paediatric emotional health evaluation.In this study, data obtained from EGG and DISC personality assessment of 25 participants were analysed.We found that EGG can actually serve the function to identify emotions between neutral, happy and sad. Through the use of mixed model ANOVA, the results of happy, neutral and sad are statistically different at $p = 0.01$. As to whether personality affects emotionexpression, a significant main effect for personality was not found and only a weak interaction between pitch and personality was reported. Thus contrary to the hypothesis that extroverts and introverts would convey emotions with different intensity and accuracy, there is no noticeable effect of personality on the emotion recognition system.

*Keywords:*Emotion recognition system; Electroglottograph; Pitch; Emotion; Personality

*Corresponding author. Tel.: +65-64082000; fax: +65-67761378.
*E-mail address:*luhui1304@gmail.com

doi:10.1016/j.promfg.2015.07.236

1. Introduction

This research attempts to useElectroglottograph(EGG) as an emotion recognition system, by measuring differences in pitch. Our aim is to improve current emotion recognition systems that still have limitations. The goal is to find out if there is a relationship between the pitch, in terms of vibration around the neck, as to the emotion felt by a person, and if extroverts and introverts would have different pitches for the same emotion.

There are various applications for our proposed emotion recognition system. Firstly, it can be enhance lie detector to detect discrepancy between emotional and verbal messages. Secondly, it can identify social-emotional problems in toddlers. Thirdly, automatic emotion recognition can improve human-machine interactions and human-computer interactions, such as in humanoid robots, car industry, call centers, mobile communication, and computer tutorial applications [1]. This form of speech recognition can be used as a voice tag in different database access systems, used mainly in telephony shopping, ATM machine as a passcode for entering a particular account [2]. Specifically, education assistance software can identify the emotions of its users and choose suitable teaching courses. Other assistance systems, such as virtual babysitting systems or virtual psychologist systems can also use the information [1], and respond accordingly to suit the needs of its user.

1.1. Current emotion recognition systems and their limitations

Emotions play a vital role in our daily live. It changes the phonetic, prosodic, acoustic, and linguistic aspects of speech, allowing us to deduce the psychological state of a person [3]. As such, we can recognise one's emotion via their expressions, speech, biological reactions and mental states, which are the most common methods of human emotion recognition.

Two main types of emotion recognition systems currently use a microphone or facial expression detection [4, 5]. Nevertheless, they have various limitations. For speech emotion recognition systems, speech signals are captured using a close-talking microphone. In an uncontrolled environment upon deployment, background noise can potentially degrade the performance of speech emotion recognition system as current acoustic transducers are unable to differentiate between noise and speech on a fundamental basis, leading to inaccuracy in recognition [6]. People may feel intimidated by having a microphone in front of them, further leading to the inaccuracy of the results collected. Speech is also affected by the gender of the speaker, dialect, age, language, emotions, stress and various other factors [7]. Syntactic structure information is crucial for natural language processing in textual emotion recognition, but sometimes cannot be obtained, so an additional parser may be necessary [8]. As to facial expression recognition, the performance of the system is often affected by the sex, race and age of the participants, leading to inaccuracy of the results [9].

1.2. Electroglottograph

To improve emotion recognition systems, we propose using a pair of disc electrodes to record the vibrations around the neck instead of acoustic methods. Electroglottography signals record the time varying displacement of air particles at the glottis during the production of voiced sounds such as vowels, semi vowels, nasals, diphthongs and voiced consonants. The electrodes are placed, non-invasively, at either side of the larynx. A high-frequency electric current is applied, and due to variance in electrical impedance from the opening and closing of the glottis, an electroglottogram can be produced.

There are several advantages of using EGG, the most significant being to reduce background noise. In applications in areas like interrogation or consultation, EGG is able to isolate the emotion of the participant while a microphone would pick up voices of other people and other random background noises. By eliminating irrelevant signals, EGG can increase the accuracy in the identification of emotions. In addition, EGG can be easily disguised into clothes, like the collar of sweaters or jacket. Thus, it may seem less intimidating than the visible microphone right in front of the participants.

As glottal waveform is greatly affected by the emotional state [10], changes in emotions can be more easily recorded. It is suggested and that glottal models can be used for expressive speech characterization and that speaker bias was not observed under controlled recording environments [11].

1.3. Personality

Likewise, one's personality may affect their expressions, which may lower the accuracy of emotion recognition systems.Thus the effect of personality on pitch was investigated as well. There are four main types of personality types, D, I, S and C, for dominant, inspiring, stable and compliant respectively [12]. People with dominant D-typed or I-typed personality are extroverts while people with dominant S-typed or C typed personality are introverts [12]. Each person will display one or more personality types. Based on their DISC personality test results, the participants will be classified as either extroverts or introverts.

An introvert is usually an individual with an accompanying tendency to withdraw from social contacts [13]. Thus, we can assume that they may be more reserved and less expressive, leading to difficulty in identifying the correct emotions. On the other hand, extroverts have an accompanying tendency to make social contacts [13]. Therefore, it may be easier for them to express emotions more easily but it may backfire as they can be overly expressive. The results obtained would not be practical as more than half of the world's population is introverts. As such, we would like to find out if our participants' personality affects the difference between their pitches for the different emotions. This allows for a more accurate emotion recognition system using the EGG as it determines whether the emotion recognition system will apply for everyone, regardless of their personality types.

1.4. Hypothesis

Of the myriad of emotions one feels daily, the six main emotions are happiness, sadness, anger, fear, disgust and surprise [14]. Despite there being six main emotions, we would only be differentiating between happy and sad, using neutral as a control. In addition, it has been shown that using fundamental frequency as a feature improves speech recognition accuracy [15], even for neutral speech [7]. Therefore, we have decided to utilise fundamental frequency, also known as pitch, in our experiment. However, electroglottograph instead of microphone is used in this research.

Our hypothesis is that there is a relationship between the pitch contour of the participants and the emotion felt by the person, and that extroverts are more expressive when acting out the emotions. (i.e. their results would be more accurate). If possible, we would like to propose a system that would allow us to either use the pitch of a subject to determine the emotion s/he is experiencing or to determine if s/he is an introvert or extrovert. In the future, real-life applications of EGG can be developed due to its ability to reduce background noise, such as a wireless EGG integrated with clothes.

2. Methodology

25 participants were chosen with a variety of gender, personality types and age. The participants include17 females and 8 males, among them 17 were introverts and 8 were extroverts. Their age ranged from 15 to 40 years old. For our results to be representative of the population, each sample point should represent the attributes of a known number of population elements. In our experiment, we conducted random sampling, and choosing of a sample unit is based on chance so every element of the population has a non-zero probability of being chosen. Hence, we can obtain representative samples by eliminating voluntary response bias and guarding against undercoverage bias.

Typically, the larger the number of participants, the smaller the margin error is. A good approximation of the margin of error (or confidence interval) is given by $1/\sqrt{N}$, where N is the number of participants [16].This means that a 90% confidence level would have a 10% probability of the results differing from the actual population mean. For our experiment, our margin for error is 0.2, which means there can be a 20% deviation in the results obtained.

There were two parts to our data collection.To test the effectiveness of electroglottograph as an emotion recognition system, participants were invited to a professional recording studio. They were asked to wear a pair of EGG electrodes around their necks. The signal from the EGG was fed to a pre-amplifier and recorded on a computer. Since music has the ability to express and regulate emotions [17], participants were required to listen to songs selected to reflect the emotions before the experiment, to ensure that the participants felt the emotion they were supposed to express. Once they felt they are feeling the correct emotion that are supposed to act out, would we

start recording their reading out lyrics from the song with the required emotion. Afterwards, there would be a short break given for them, before we repeated the process twice for the other two emotions.

To verify the effect of personality on pitch, the second part of the data collection involved the DISC personality assessment. The participants were given a code to access this test through email and encouraged to complete it within seven minutes.

The pitch values of different participants and different emotions were analysed. The pitch values were obtained frame by frame and the average pitch value of each line was calculated. Eight lines of the lyrics were taken into consideration and the average and variance of these lines were further calculated. To statistically analyse the data, a 2x3 mixed model ANOVA from SPSS was used.

3. Results

3.1. Pitch and emotion

Graphs were plotted based on data set obtained. In all diagrams, the average value in the diagram was obtained by subtracting the average pitch value of all three emotions from the pitch value of the particular emotion. The variance value in the diagram was calculated similarly using the average variance. This is to minimize the impact of individual pitch difference.

As shown by Fig.1., pitch values extracted from the vibration around the neck has an relationship with emotions, which confirms our hypothesis. A clear distinction can be drawn between happy and sad as emotion happy has a higher average of pitch value and a higher variance compared to emotion sad.

However, when three emotions, happy, sad and neutral, were taken into consideration, there was overlapping of results between sad and neutral. Emotion happy is relatively distinguishable from neutral due to its high average pitch value and high variance.

3.2. Statistical analysis

A 2 x 3 mixed model ANOVA was used to investigate the impact of personality on the differences in pitch values between three emotions, neutral, sad and happy. In the current study, 25 participants with 17 introverts and 8 extroverts were asked to record pitch values through the use of electroglottograph. The SPSS was used to analyse the data collected. The Shapiro-Wilk, F*max* and Levene's test statistics were used to test the assumptions of normality and homogeneity of variance. From SPSS, these assumptions for a mixed model ANOVA were not violated.

A significant main effect for pitch was obtained $F(1, 46) = 25.102$, $p = 0.001$, partial $\eta^2 = .52$ with emotion happy ($M = 198.23$, $SD = 43.29$) being significantly higher than emotion neutral ($M = 189.11$, $SD = 42.91$), which is significantly higher than emotion sad ($M = 184.61$, $SD = 43.12$). A significant main effect for personality was not found $F(1, 23) = .014$, $p = .906$, partial$\eta^2 = .01$. A weak interaction between pitch and personality was reported, $F(1, 23) = .698$, $p = .412$, partial $\eta^2 = .03$.

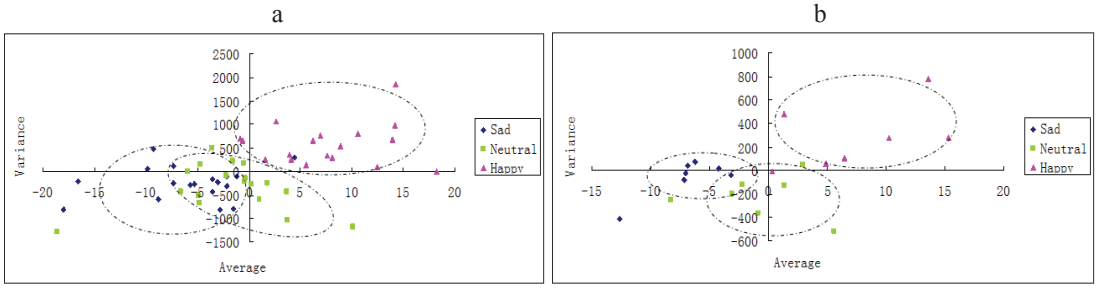

Fig.1.Effect of emotions on the pitch values on (a) female and (b) male participants.

Examination of the means indicated that although there is a large change in pitch values of introverts between emotion neutral (M = 189.21, SD = 10.61) to emotion sad (M = 183.40, SD = 10.67), there is no major change in pitch values of extroverts between emotion neutral (M = 188.90, SD = 15.50) and emotion sad (M = 187.18, SD = 15.56). This can be due to the small sample size of extroverts compared to introverts in this study. There is a large change in pitch values between emotion neutral (M = 189.21, SD = 10.61) to emotion happy (M = 197.20, SD = 10.72) for introverts, as well as between emotion neutral (M = 188.90, SD = 15.50) and emotion happy (M = 200.41, SD = 15.62).

4. Limitations

Firstly, the sample size of the participants is relatively small. This may affect the accuracy of data analysis as a minimum of 30 sample size is usually preferred in statistical analysis.Secondly, the results obtained in the study are partially dependent on the participants' acting skills and their emotional state at the time of recording. Hence, despite our many attempts to prevent the aforementioned situations, it was still possible that the participant felt calm when he was supposed to be sad or vice versa.

5. Conclusion

The relationship between pitch and emotion exists in existing emotion recognition systems. In this research, we managed to reconfirm this relationship between pitch and emotion in the use of EGG as a proposed emotion recognition system. It was also found that the pitch value of happy is more than the pitch value of neutral, which is more than the pitch value of sad. Thus, EGG could be an effective approach to recognize emotions. However, there was no noticeable relationship between personality and deviation in pitch values.

6. Further Work

We plan to have a larger sample size of adult participants to reconfirm the results in the future. The accuracy from use of EGG can be compared to that of microphone and facial expression. We also plan to extend the experiment to include other emotions –boredom, anger, fear, surprise and anxiety, so as to further analysed together with the studied emotions - happy, neutral and sad.

To further confirm our conclusion that the pitch is independent of a person's personality, we would repeat our experiment with a wider range of participants with the percentage of introverts and extroverts similar to that seen in the society, as well as to investigate whether there will be any possible effect between subjects of different age. Should a correlation exist, a system can be proposed to modify the pitch values collected according to the participant's personality, allowing higher accuracy in the emotion recognition system.

Acknowledgement

We would like to thank Mr. Adrian Lohfor his help in the process of obtaining IRBapproval, and the participants who contributed to the data collection. Last but not the least, we appreciate the Ministry of Education (MOE), A*STAR and National Junior College for providing this wonderful R&D opportunity. This project would not have been possible without the help and support from all of them.

References

[1] M. Bhargavaa, T. Polzehl. (2012) Improving Automatic Emotion Recognition from speech using Rhythm and Temporal feature. Srinivasa Ramanujan Institute of Technology, Anantapur. Retrieved April 7, 2015, from http://arxiv.org/ftp/arxiv/papers/1303/1303.1761.pdf.

[2] A. A., Khulage, and B.V., Pathak (2012) Analysis of speech under stress using Linear techniques and Non-Linear techniques for emotion recognition system. Retrieved April 7, 2015, from http://arxiv.org/ftp/arxiv/papers/1207/1207.5104.pdf.

[3] A.S., AlMejrad. (2010). Design of an intelligent system for speech monitoring and

treatment of low and excessive vocal intensity. Artificial Life and Robotics, 15(3), 320-324.
[4] G., Panagiotis, and P., Gerasimos. (2012) A Hierarchical Approach with Feature Selection for Emotion Recognition from Speech. Retrieved on 10 July 2014 from http://www.lrec-conf.org/proceedings/lrec2012/pdf/917_Paper.pdf
[5] I, Cohen. (2000) Emotion Recognition from Facial Expressions using Multilevel HMM. Retrieved July 15, 2014 from http://www.ifp.illinois.edu/~ashutosh/papers/NIPS_emotion.pdf
[6] P. T., Khiet, D. A.van Leeuwen, and M. A. Neerincx (2007). Unobtrusive Multimodal Emotion Detection in Adaptive Interfaces: Speech and Facial Expressions. Augmented Cognition, HCII 2007, LNAI 4565, pp. 354–363, 2007.
[7] D. Gharavian, M. Sheikhan, M. Janipour (2009) Pitch in Emotional Speech and Emotional Speech Recognition Using Pitch Frequency. Majlesi Journal of Electrical Engineering, Vol. 4, No. 1, March 2010. Retrieved April 7, 2015,from http://www.researchgate.net/profile/Mansour_Sheikhan/publication/50888741_Pitch_in_Emotional_Speech_and_Emotional_Speech_Recognition_Using_Pitch_Frequency/links/0c960514d4b632e01a000000.pdf
[8] Z, Chuang, C., Wu. (2004). Multi-Modal Emotion Recognition from Speech and Text. Computational Linguistics and Chinese Language Processing, 9:2, August 2004, pp. 45-62. Retrieved from http://www.aclweb.org/anthology/O04-3004.
[9] Z. Cohn, et al. (1998). Feature-point tracking by optical flow discriminates subtle differences in facial expression. Retrieved July 15, 2014 from http://www.ri.cmu.edu/pub_files/pub2/cohn_jeffrey_1998_3/cohn_jeffrey_1998_3.pdf
[10] K.E. Kummings. (1990). Analysis of Glottal Waveforms Across Stress Styles. Sch. of Electr. Eng., Georgia Inst. of Technol., Atlanta, GA. DOI: 10.1109/ICASSP.1990.115687
[11] L. T., Jaime, et. al. (2012). Towards Glottal Source Controllability in Expressive Speech Synthesis. Speech Technology Group, ETSI Telecomunicacion, Universidad Politecnica de Madrid, Spain, Department of Signal Processing and Acoustics, Aalto University, Finland ,Sound Analysis and Synthesis, IRCAM, Paris, France ,CSTR, University of Edinburgh, United Kingdom. Retreieved on 10 July 2014 from http://oa.upm.es/20409/1/INVE_MEM_2012_134451.pdf
[12] A. R., Robert. (2010) A Powerful Way to Understand People: An introduction of the DISC concept. Personality Insights Inc. Retrieved on 10 July 2014 from https://personality-insights.com/DISC_overview.pdf
[13] M., Freyd. (1924). Introverts and Extroverts. Psychological Review, Vol 31(1), 74-87. doi: 10.1037/h0075875
[14] Ekman P. (1999). Handbook of Cognition and Emotion. Retrieved July 15, 2014 from https://www.paulekman.com/wp-content/uploads/2013/07/Basic-Emotions.pdf
[15] H., Shah. (2007). Emotion Detection from Speech. Retrieved April 7, 2015, from
http://cs229.stanford.edu/proj2007/ShahHewlett%20-%20Emotion%20Detection%20from%20Speech.pdf
[16] R., Niles. (2006). "Robert Niles' Journalism Help: Statistics Every Writer Should Know," RobertNiles.com. Retrieved June 28, 2006 fromhttp://www.robertniles.com/stats/.
[17] G, Shan. (2007). Some reasons why songs are emotionally powerful. Retrieved July 15,2014, from http://www.martinandshan.net/userimages/Emotional_power.pdf