Tap pressure on touchscreens for detection of emotion

Kevin Blom
University of Amsterdam
Science Park 904
xxxxxx@xxxxxx.xx

ABSTRACT

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User-centered design

Author Keywords
Human centered multimedia; emotion; tap pressure

INTRODUCTION

Affective computing as introduced by Picard[11] in 1995 lays a foundation for computers and technology to incorporate the recognition and expression of emotions. It can provide better performance when assisting humans or enhance the computers ability to make decisions. It does not have the goal of making computers more human-like, but it is more practical in nature; make computers function with intelligence and sensitivity towards its users[12]. According to Shah et al.[14] there are two general models to represent emotion; discrete and continuous. The discrete model represent emotions that are measurable and physiologically distinct like angry, sad, happy, etc. [3] The continuous model represents emotions on a two-dimensional scale, where one axis represents *valence* and the other *arousal* [13]. Mauss et al. [9] suggest that using a dimensional framework is a better option when capturing emotion, relative to discrete frameworks. Since, the measuring of emotion has been a subject of research and several different angles have been discovered to approach it.

Physiological detection

One angle uses physiological signals of the human body to measure and detect emotion. In a review by Wioleta[19], eight studies were collected that measure emotion using one or more physiological signals combined. These signals are *EEG, skin conductance, blood volume pulse, temperature, heart rate, blood pressure, respiration, EMG,* and *ECG*. Most of these physiological signals have the drawback that they need specialized sensors attached to the body, making unobtrusive measurements difficult. With the recent rise of smart wearables, heart rate is one of the signals that is more readily available to use in applications on smart devices.

CHI'16, May 07–12, 2016, San Jose, CA, USA

ISBN 123-4567-24-567/08/06…$15.00
DOI: http://dx.doi.org/10.475/123_4

Facial detection

Facial detection of emotion incorporates the measurement of facial muscle movement, voice or speech [17], and also includes the eye as point of detection, i.e. movement, blinking, and pupil dilation [16]. By connecting facial muscle movement to visual display of emotions, Ekman et al. [4] conclude with a basic set of six mutually exclusive emotions that could be recognized. Expanding, De Silva et al. [15] found that several emotions are expressed by either visual or auditory cues, or both, meaning that some emotions can be recognized by visual cues alone, auditory cues alone, or need a combination of both to be detected accurately.

Posture/gestures emotion detection

Other means of detection emotions involve the tracking and interpretation of posture and gesture. Wallbott et al. [18] concluded in 1998 that there are, in some cases, distinctive patterns of movement and postural behavior that have a strong correlation to emotions. In other cases, they mention that in absence of patterns there are still distinctive features from which emotion could be inferred. Coulson et al. [1] researched static body postures and the recognition of emotions from these body postures by participants. It showed that disgust is a tough emotion to recognize but anger and sadness had over 90% correct detection rates. Furthermore, happiness and surprise were two emotions that were often confused.

Practical applications

Looking at a more practical and applied side of emotion detection, Gao et al. [5] used touchscreen devices, where the application of gestures on touch screens was successfully linked to emotional states with the use of a game. The emotional states that were tested for are: excited, relaxed, frustrated and bored, and accuracy of detection reached at minimum 69%. However, the research of Gao et al. was limited to gestures and did not incorporate data from taps. Furthermore, Lv et al. [6] have created means to detect emotion from keyboard pressure using feature extraction. This indicates that the use of a keyboard on a touch screen could also be used as means of detecting emotion, but one must keep in mind that a regular keyboard is not fully comparable to a touchscreen keyboard. It lays flat on a desk, and is often typed upon with more than one or two fingers, which means that the pressure exerted on the keyboard is likely not directly correlated with the pressure on a touchscreen keyboard. Moreover, Lee et al [7] propose an unobtrusive way of detecting emotion by analyzing smartphone usage patterns (not unlike LiKamWa et al. [8]) and social network status updates. However, this required that the user would post status updates through independently developed social networking applications, that are not officially supported by the social networks themselves.

Research Question

From the related work can be concluded that most types of detection of emotions are invasive, either requiring constant monitoring, possibly with sensors attached to the body, or by constant recording of audio and visual data. The touch screen is a technology a lot of people interact with every day, where they deliberately choose to participate in those interactions. Using touch screen presses as indicators for emotional state could be an unobtrusive way of detecting emotion without the need for constant monitoring. With the introduction of pressure sensitive touchscreens in recent smart devices, an interesting new sensor is added to the plethora of sensors already available. Subsequently, this leads to the following research question:

Can pressure sensitive touch screen devices be used to tell more about the mood of the user?

METHODS

In order to test for the correlation between taps on a touch screen and emotion, there has to be a standardized way of eliciting different emotions. Fortunately, there exists a photo set that has been thoroughly tested for emotional response on a two dimensional scale that is called the Geneva Affective Picture Database (GAPED) [2]. Utilizing the emotional responses of this photo set as a baseline, touch screen taps and their pressure can be compared to emotional response. Participants were selected using a convience sampling process at an office. The participants varied in age, educational level, current line of work, and background.

Emotional elicitation

Using a standardized photo set that has been thoroughly researched for emotional response when showed to participants, a ground truth for emotion was set. The GAPED photo set uses the continuous model of representing emotions, i.e. the two-dimensional valence and arousal model. The photo set counts 730 pictures and is divided into 6 categories: Animal, Human, Neutral, Positive, Snakes, Spiders. From each of the categories, 10 pictures were randomly selected, resulting in a set of 60 pictures used for the experiment. Each participant was presented with the same 60 pictures, but in random order. Brown et al. [10] remark that 5 second exposure to pictures is often used for the International Affective Picture System photos. The GAPED photo set has been created because of two issues with the IAPS; extensive use decreases impact of the stimuli, and the limited number of pictures for specific themes. Both these issues are not exposure time related, so the choice of exposure time of the photo to the participant is 5 seconds.

Pressure detection

Taps were detected on an Apple iPhone 6s device with a 3D touch screen running iOS 10.3.1. The pressure of taps was registered on a floating point scale from 0.0 to 6.67 (Corresponding with 0 to ± 350 grams) and for every tap, several pressure measurements were registered in chronological order. Furthermore, the duration of a tap (in nanoseconds) was also registered.

Data collection

In order to collect a larger data set, 4 taps per photo were required to advance to the next photo. These taps are directed with the use of gray colored buttons that are randomly shown on a 4 by 4 grid on the screen (Figure 1). The random pattern of the buttons ensures that the position of the tap on the screen does not matter for the pressure measurement. The gray color is used because it is perceived as neutral. The buttons are random for every photo, and for every participant. In other words, no participant received the same grid for the same photo.

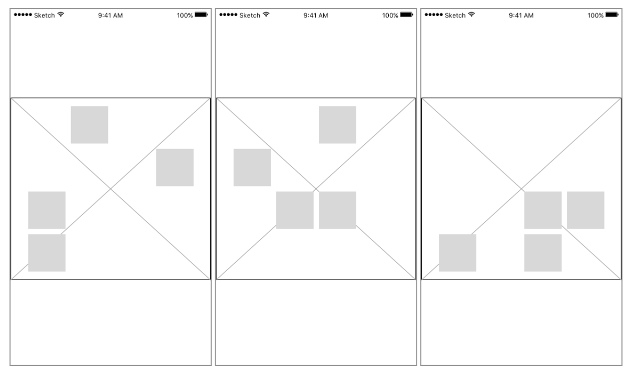

Figure 1. Three examples of the grid as presented over a picture.

All the data that was collected was anonymously and securely sent realtime to a Firebase[1] database. Firebase utilizes a JSON[2] tree structure that can be described as in Figure 2.

Experiment setup

Firstly, participants were told what the expirement entailed and were presented with a consent form. Subsequently, the participants continued the experiment on the smart device with test application. The test application is structured as follows:

1. Participant is presented with a screen that asks if they received and signed a consent form and if not, that they should contact the supervisor immediately. There is also a *start* button to start the experiment.
2. The participant is shown a picture.
3. After five seconds, four gray buttons are shown, overlaid on the picture in a random pattern (Figure 1).
4. When the participant pressed all the 4 buttons, the next picture is presented.
5. This process repeats untill al 60 pictures have been shown.

[1] http://firebase.google.com/
[2] http://www.json.org

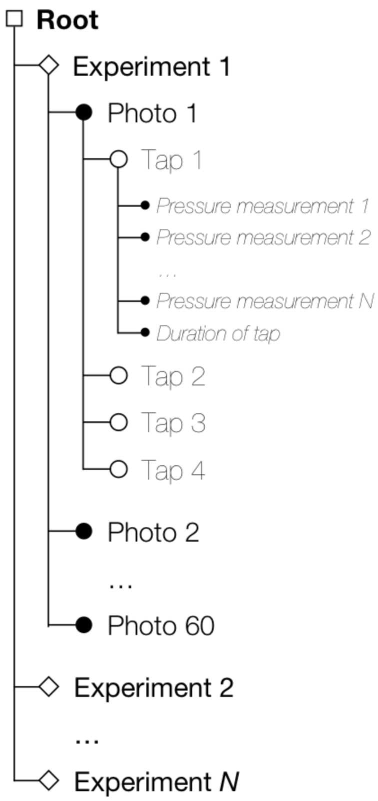

Figure 2. Database structure in simplified form.

6. The participant is presented with a conclusive screen that has a thank you message and refers to the supervisor if there are questions.

Data analysis

The collected data was exported as JSON from Firebase and subsequently mutated using Python[3] 2.7 on macOS[4] 10.12.4 in order to create an *.csv* file that was readable by SPSS 20.0[5]. Because of a suspicion that either maximum exerted pressure or average exerted pressure of a tap might be of influence, these two variables were manually added using averaging. For maximum pressure, the maximum pressure value of each tap was extracted, and for each photo this was averaged. Regarding average tap pressure, the average pressure of a tap was calculated and subsequently all the average tap pressures were averaged again per photo. These averages make it possible to compare means. The result is six variables in SPSS;

1. **Photo filename** - String, containing the photo filename for identification purposes.
2. **Valence** - Numeric, decimal value on a scale from 0.0-100.0.
3. **Arousal** - Numeric, decimal value on a scale from 0.0-100.0.
4. **Maximum tap pressure average** - Numeric, decimal value on scale from 0.0-6.67.
5. **Average tap pressure average** - Numeric, decimal value on scale from 0.0-6.67.
6. **Duration** - Numeric, decimal value in nanoseconds.

In other words, for every photo there is a value for valence, arousal, maximum tap pressure average, average tap pressure average and duration.

Multiple linear regression

Testing for any correlation was completed with a multiple linear regression method. The advantage of this method is that both *duration* and *pressure* can be used as independent variables to check if there indeed is a relation to *valence* or *arousal* as dependent variables and that if there is a relationship, it also immediately produces a model to predict the dependent variables. A disadvantage of is that the relationship can not be checked for *valence* and *arousal* simultaneously, only for the seperate variables.

Before proceeding to the results, several assumptions needed to be considered before concluding that the data could be analysed using multiple linear regression;

1. **Independence of observation** - Using Durbin-Watson to test for 1st-order autocorrelation.
2. **Linear relationships** - Visually inspecting a scatterplot of studentitized residuals and unstandardized predicted values, and partial regression plots can indicate a (non)linear relationship.
3. **Homoscedasticity of residuals** - Visually inspecting a scatter plot of studentitized residuals and unstandardized predicted values.
4. **No multicollinearity** - Inspection of correlation coefficients and Tolerance/VIF values for indication of correlation between independent variables.
5. **Unusual data points** - There should be no outliers, high leverage points or highly influential points.
6. **Normal distribution of errors** - Errors in prediction need to be normally distributed, otherwise determining significance can become problematic.

All these assumptions will be considered and results of any tests that come with it are presented in the next section.

What statistical methods were used, based on what principles and data..

- Overview of the research.
- Report of who took part and where.
- Report of what procedures were used.
- Report of what materials were used.
- Report of any statistical analysis used.

[3] https://www.python.org
[4] https://www.apple.com/lae/macos/sierra/
[5] https://www.ibm.com/analytics/us/en/technology/spss/

RESULTS

- Report of findings.
- Reference to any diagrams used.

Maximum tap pressure

There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.718. There was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There were no studentized deleted residuals greater than ±3 standard deviations, no leverage values greater than 0.2, and values for Cook's distance above 1. The assumption of normality was met, as assessed by Q-Q Plot. The multiple regression model did not statistically significantly predicted tap pressure, $F(2,57) = 2.033$, $p > 0.05$, adj. $R^2 = 0.034$. Non of the variables added statistically significantly to the prediction(i.e. $p > 0.05$), with $p_{valence} = 0.35$ and $p_{arousal} = 0.079$.

DISCUSSION

- Summary of main purpose of research.
- Review of most important findings.
- Evaluation of findings.
- Explanation of findings.
- Comparison with other researchers findings.
- Description of implications and recommendations.

REFERENCES

1. Mark Coulson. 2004. Attributing emotion to static body postures: recongition accuracy, confusion and view point dependence. *Journal of Nonverbal Behavior* 28, 2 (2004), 117–139. DOI: http://dx.doi.org/10.1023/B:JONB.0000023655.25550.be
2. Elise S Dan-glauser and Klaus R Scherer. 2011. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. (2011), 468–477. DOI: http://dx.doi.org/10.3758/s13428-011-0064-1
3. Paul Ekman. 1992. An argument for basic emotions. (1992). DOI: http://dx.doi.org/10.1080/02699939208411068
4. Paul Ekman, E Richard Sorenson, and Wallace V Friesen. 1969. Pan-Cultural Elements in Facial Displays of Emotion. (1969). DOI: http://dx.doi.org/10.1126/science.164.3875.86
5. Yuan Gao, Nadia Bianchi-Berthouze, and Hongying Meng. 2012. What Does Touch Tell Us about Emotions in Touchscreen-Based Gameplay? *ACM Transactions on Computer-Human Interaction* 19, 4 (2012), 1–30. DOI: http://dx.doi.org/10.1145/2395131.2395138
6. J. Dong H. R. Lv, Z. L. Lin, W. J. Yin. 2008. Emotion recognition based on pressure sensor keyboards. *2008 IEEE International Conference on Multimedia and Expo* (2008), 1089–1092. DOI: http://dx.doi.org/10.1109/ICME.2008.4607628
7. Hosub Lee, Young Sang Choi, Sunjae Lee, and I. P. Park. 2012. Towards unobtrusive emotion recognition for affective social communication. *2012 IEEE Consumer Communications and Networking Conference, CCNC'2012* (2012), 260–264. DOI: http://dx.doi.org/10.1109/CCNC.2012.6181098
8. Robert Likamwa, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. 2013. MoodScope: Building a Mood Sensor from Smartphone Usage Patterns. *MobiSys '13 Proceeding of the 11th annual international conference on Mobile systems, applications, and services* April (2013), 389–402. DOI: http://dx.doi.org/10.1145/2462456.2464449
9. Iris B Mauss and Michael D Robinson. 2009. Measures of emotion: A review. *Cognition and Emotion* 23, 2 (2009), 209–237. DOI: http://dx.doi.org/10.1080/02699930802204677
10. Human Neuroscience, Stephen B R E Brown, Henk Van Steenbergen, P Guido, H Band, Mischa De Rover, and Sander Nieuwenhuis. 2012. Functional significance of the emotion-related late positive potential. 6, February (2012), 1–12. DOI: http://dx.doi.org/10.3389/fnhum.2012.00033
11. Rosalind W. Picard. 1995. Affective Computing. *MIT press* 321 (1995), 1–16. DOI: http://dx.doi.org/10.1007/BF01238028
12. Rosalind W Picard. 1997. *Affective computing*. Vol. 252. MIT press Cambridge.
13. Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 3 (2005), 715–34. DOI: http://dx.doi.org/10.1017/S0954579405050340
14. Sachin Shah, J. Narasimha Teja, and Samit Bhattacharya. 2015. Towards affective touch interaction: predicting mobile user emotion from finger strokes. *Journal of Interaction Science* 3, 1 (2015), 6. DOI: http://dx.doi.org/10.1186/s40166-015-0013-z
15. Liyanage C D E Silva and I Tsutomu Miyasato. 1997. Facial Emotion Recognition Using. September (1997), 9–12. DOI: http://dx.doi.org/10.1109/SMC.2015.387
16. Mohammad Soleymani, Maja Pantic, and Thierry Pun. 2015. Multimodal emotion recognition in response to videos (Extended abstract). *2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015* 3, 2 (2015), 491–497. DOI: http://dx.doi.org/10.1109/ACII.2015.7344615

17. Dimitrios Ververidis, Constantine Kotropoulos, and Pitas Ioannis. 2004. Automatic Emotional Speech Classification. *Artificial Intelligence and Information Analysis Laboratory* (2004), 593– 596. DOI: http://dx.doi.org/10.1109/ICASSP.2004.1326055

18. Harald G. Wallbott. 1998. Bodily Expression of Emotion. *European Journal of Social Psychology* 28, 6 (1998), 879–896. DOI: http://dx.doi.org/10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

19. Szwoch Wioleta. 2013. Using physiological signals for emotion recognition. *Human System Interaction (HSI), 2013 The 6th . . .* (2013), 556–561. DOI: http://dx.doi.org/10.1109/HSI.2013.6577880