6-2013

CaptureMyEmotion: Helping Autistic Children Understand their Emotions Using Facial Expression Recognition and Mobile Technologies

Valerie Gay
Faculty of Engineering and Information Technology, Australia, Valerie.Gay@uts.edu.au

Peter Leijdekkers
Faculty of Engineering and Information Technology, Australia, peter.leijdekkers@uts.edu.au

Johann Agcanas
Faculty of Engineering and Information Technology, Australia, integriti@gmail.com

Frederick Wong
Faculty of Engineering and Information Technology, Australia, frederick.k.wong@student.uts.edu.au

Qiang Wu
Faculty of Engineering and Information Technology, Australia, qian.wu@uts.edu.au

Follow this and additional works at: http://aisel.aisnet.org/bled2013

CaptureMyEmotion:
Helping Autistic Children Understand their Emotions Using Facial Expression Recognition and Mobile Technologies

Valerie Gay
Faculty of Engineering and Information Technology, Australia
valerie.gay@uts.edu.au

Peter Leijdekkers
Faculty of Engineering and Information Technology, Australia
peter.leijdekkers@uts.edu.au

Johann Agcanas
Faculty of Engineering and Information Technology, Australia
integriti@gmail.com

Frederick Wong
Faculty of Engineering and Information Technology, Australia
frederick.k.wong@student.uts.edu.au

Qiang Wu
Faculty of Engineering and Information Technology, Australia
qian.wu@uts.edu.au

Abstract

One of the main challenges for autistic children is to identify and express emotions. Many emotion-learning apps are available for smartphones and tablets to assist autistic children and their carers. However, they do not use the full potential offered by mobile technology, such as using facial expression recognition and wireless biosensors to recognise and sense emotions. To fill this gap we developed CaptureMyEmotion, an Android App that uses wireless sensors to capture physiological data together with facial expression recognition to provide a very personalised way to help autistic children learn about their emotions. The App enables children to capture photos, videos or sounds, and simultaneously attach emotion data and a self-portrait photo. The material can then be reviewed and discussed together with a carer at a later stage. CaptureMyEmotion has the potential to help autistic children integrate better in the society by providing a new way for them to understand their emotions.

Keywords: M-health, Emotion detection, Facial expression detection, Autism, Personalised technology, Affective computing

1 Introduction

The Autism Spectrum Disorder (ASD) describes a range of conditions typically characterized by social deficits, communication difficulties, stereotyped or repetitive behaviours and interests, and in some cases, cognitive delays (American Psychiatric association, 2000). The prevalence of ASD is estimated to be up to 1 in 88 children (Baio et Al, 2012). Autistic children, like other children of their generation, embrace mobile technology and enjoy using smart phones and tablets to learn and play. Most autistic children are visual learners and mobile technology appeals to them (de Urtiri et Al, 2011).

This research investigates how mobile technology, wireless sensors and facial expression recognition software can help autistic children and their carer understand emotions. Managing emotion is very challenging for the majority of autistic children and it is of major importance for their integration in classrooms and other social groups.

Websites like iAutism (2013) and Autism Speak (2013) are maintaining an impressive list of Apps designed for children with Autism. Apps exist for describing emotions or showing typical facial expressions and are used for emotion recognition exercises. Many of these apps have a fixed set of emotions where the user has to identify or understand the emotion. *Moody Me-Tracker* (Moody Me, 2012) app is of interest since it allows users to take pictures that make them happy or sad. It allows users to playback pictures to make them happy and it has a diary to track mood changes. *Wonkido Mood Journal* (Wondiko, 2013) is another app to help children recognize their emotions and journal their feelings. *Micro-Expressions* (micro expressions,) trains users to identify a facial expression using different characters. A similar app called *What's the Expression?* (Shanesh, 2013) teaches autistic children to understand and recognise expressions and emotions using a well-known cartoon character. Facial expression recognition is now feasible on smartphones due to improved facial expression recognition algorithms and increased processing power (Cho et Al, 2012) and (Yamaguchi, 2012). *FaceFetch* (Mahesh et Al, 2012) and *Beyond Touch* (Anand, 2012) are two smartphone apps that demonstrate the use of smartphones to recognize facial expressions.

Many apps only target the English speaking market and a few support 2 or more other languages. For an app that uses text and audio it is important that it addresses the autistic child in their mother tongue. Therefore an app should support as many languages as possible and should be configurable by the carer. Many apps use a fixed set of images, videos or audio, which might hinder a child to express its mood based on something very specific. A good example is the video in (Hardman, 2012) that shows an autistic child getting very excited when he sees the garbage truck emptying the bins every week. These types of situations are not covered in apps with a fixed set of scenarios and calls for apps that allow the child to record their own scenarios to communicate their feelings. Furthermore, none of the mobile apps use physiological data from a sensor to identify a particular emotion. They are based on self-reporting or on the observation of a carer. However, this can be highly unreliable since many autistics have difficulty communicating emotions (Autism speaks, 2010).

CaptureMyEmotion is different from existing apps since it creates a very personalised tool for a child to learn and discuss their feelings. It combines personally recorded photos, videos or sounds with wireless physiological sensors and facial expression recognition to determine

automatically the emotional state. Sensor data is used when the child takes a picture, as well as, when the child discusses the picture with a carer at a later stage.

The paper is organised as follows: Section 2 introduces Affectiva's Q sensor (Affectiva, 2013) and the Zephyr Bioharness (Zephyr, 2012) to detect the emotional state using physiological data, as well as, the facial expression detection software developed at University of Technology, Sydney. Section 3 presents *CaptureMyEmotion* and how the child and the carer use it. Section 4 discusses the challenges and opportunities.

2 Emotion detection using sensors and facial expression

Emotion can be described by two dimensions i.e. arousal (high versus low) and valence (positive versus negative) according to Schlosberg (1954). Based on these two dimensions, some basic emotions can be identified which exist for all human beings independent of race (Figure 1). For example, joy can be classified as high arousal and positive valence whereas sadness can be classified as low arousal but negative valence.
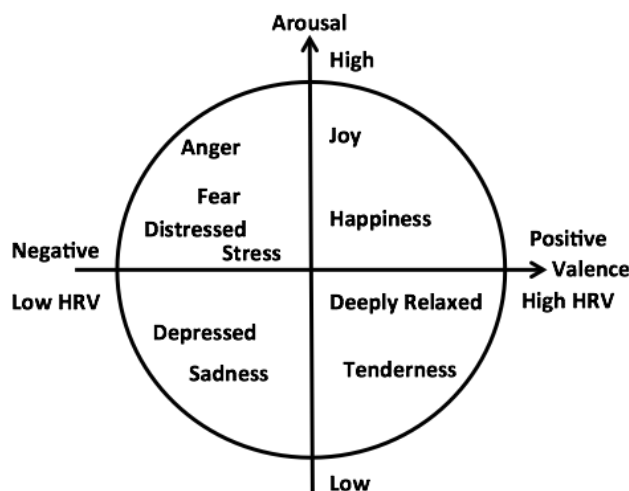

Figure 1: Emotions classified by arousal/valence levels

There is no perfect measure of arousal or valence. However, good approximations can be obtained by measuring arousal based on skin conductance and valence from facial expressions (Picard, 2008). Heart rate variability (HRV) is a good indicator of the stress level and can also be used to measure arousal. Combining the output from physiological sensors and facial expression detection software should provide insight into the emotional state of a person.

3 Arousal detection

For arousal detection, we use the Affectiva Q sensor, a Bluetooth enabled sensor that collects in real time accelerometer, skin temperature and skin conductive data (Figure 2, right). Skin conductance usually increases when a person is more aroused (engaged, stressed, or excited) and it tends to stay low or drop when a person is less aroused (disengaged, bored, or calm) (Picard, 2009). The sensor is easy to wear and is able to transmit the data in real time to a

mobile phone. The Q sensor can be worn for a long period of time to obtain a reliable baseline.

4 Stress detection

For stress detection, we use Zephyr BioHarness™, a small Bluetooth sensor worn on a strap that measures heart rate, heart rate variability (HRV), respiration, body temperature and movement (Figure 2, left). We are interested in heart rate and HRV since it relates to emotional arousal. HRV is the physiological phenomenon of variation in the time interval between heartbeats. It is measured by the variation in the beat-to-beat interval. Various emotional expressions produce different changes in autonomic activity. For example, anger or fear will increase the heart rate and lower the HRV whereas happiness or being relaxed will decrease the heart rate and increases the HRV.

Figure 2. Zephyr BioHarness™ (left) and Affectiva Q sensor (right)

5 Facial Expression detection

Six basic emotions (happiness, sadness, fear, disgust, surprise and anger) are defined in the Facial Action Coding System (FACS) (Ekman, 1978) and each emotion has a uniquely corresponding facial expression. Obtaining facial expression features can be divided into two categories: *spatial* (Song et Al, 2012), (Lajevardi et Al,2012) and *spatio-temporal* (Tie et Al, 2013), (Du et Al,2009). In the spatial approach, expression features are extracted from a static face image and used for expression classification. The spatio-temporal approach models dynamic features and computes the models' parameters through statistics of observations. The spatial-temporal approach is capable of capturing rich dynamic information of facial expression and is, in general, more suitable for identifying human facial expressions. The downside is that it is computationally expensive and often not suitable for real-time processing, especially on a smartphone with limited processing power.

The spatial approach is, however, an attractive solution and more suitable for *CaptureMyEmotion*. Many autistic children do not express explicitly their emotions or in a manner that varies widely from standard forms of expression. For example, the child may feel happy but barely expresses a smile. Thus a personalised approach catered for each child would be a desirable solution for facial expression recognition.

CaptureMyEmotion achieves this personalisation by utilising the OpenCV library (OpenCV, 2013). *CaptureMyEmotion* requires the child to 'train' the program by instructing the child to take multiple self-photos of their facial interpretation of each significant emotion. The spatial approach allows real-time processing of the facial expression since it does not involve complicated processes for acquiring dynamic facial expression. Once satisfied with the training photos, the child can initiate the algorithm that utilises the OpenCV real-time processes to recognise and extract the child's facial features. The child uses *CaptureMyEmotion* to take a photo of the object that induces an emotional response. The app has been designed in such a way that when a photo of the object is taken, simultaneously, a photo of the child's face is taken discreetly using the front camera on the phone. This allows for an accurate representation of the facial emotion that the child is displaying while viewing the object. The spatial algorithm will measure the facial expression in the input picture and compares it with the training photos. The most likely expression will be determined based on the closest similarity measurement.

6 CaptureMyEmotion

CaptureMyEmotion enables autistic children to take photos, videos or sounds and to comment on the emotion felt while taking the picture. Our target group is children from seven years and older who are able to use a smartphone to take pictures. *CaptureMyEmotion* let the children take their own pictures1 and is not based on an existing set of pictures. This increases the chance that a picture generates a genuine emotion. *CaptureMyEmotion* is developed for Android smartphones and tablets. The child uses the smartphone to take pictures and the carer would use either the phone or the tablet for evaluation together with the child. *CaptureMyEmotion* uses Dropbox (Dropbox, 2013) to share the data between several devices. When the child takes a picture, it is automatically uploaded to a Dropbox account via 3G/4G or WiFi. If the carer deals with several children (e.g. in a classroom situation), the carer can synchronize the data for the various children using Dropbox on the tablet.

7 Capture of the pictures and related context and emotions

The main screen for CaptureMyEmotion allows the child to select the camera, video or sound icon to capture what is shown in the viewer (Figure 3, left).

---

1 When we refer to pictures this also includes short videos, music or sounds.
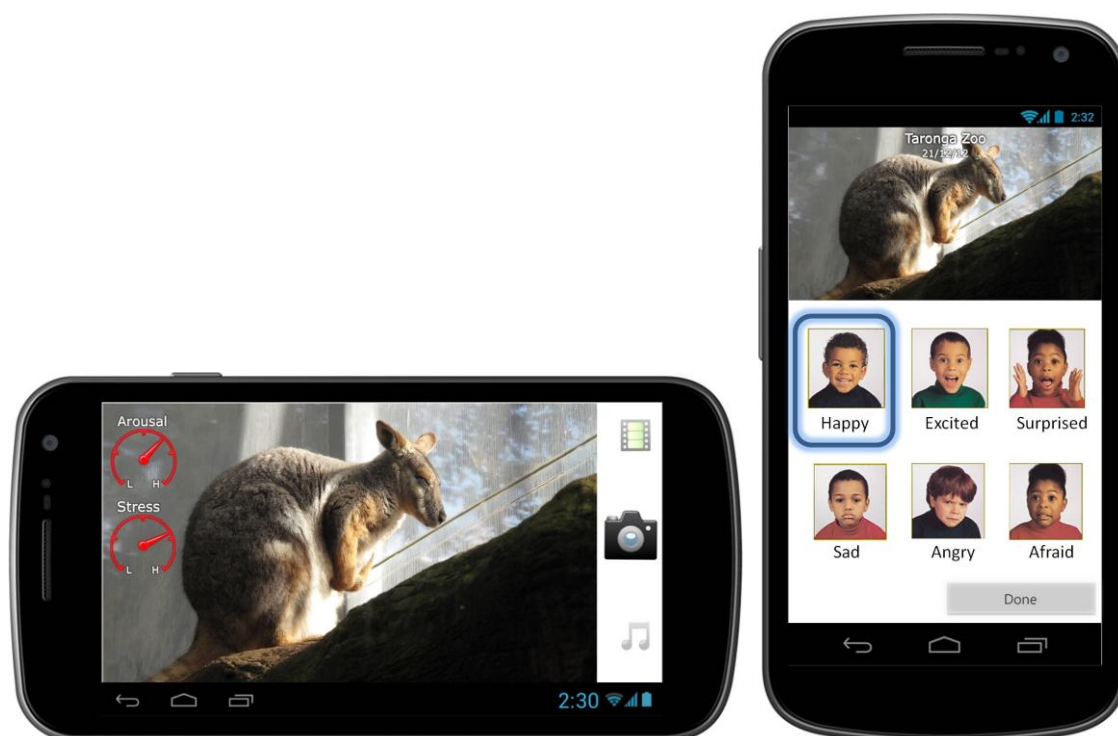
Figure 3: Main screen and screen for rating a picture

CaptureMyEmotion automatically connects to the Bluetooth sensors (if present) and gauge icons display in real-time the arousal and stress level experienced by the child. In practice, we expect the child to wear only one sensor so in that case the gauge shows either arousal or stress level. The gauge range is from low to high and for the arousal sensor, the range is determined for each user individually based on the baseline measurement.

Once the child has taken a picture, the front camera automatically takes a picture of the child's face. After the self-portrait has been taken, the app shows a screen where the child assigns an emotion to the picture (Figure 3, right). When the done button is tapped, the selected emotion is stored together with the sensor data and self-portrait picture. It is automatically uploaded to the Dropbox account if activated.

8 Facial expression recognition calibration

The facial expression recognition module requires the training of the app to the child's facial features. This is achieved through a training module, in which the child is instructed to take several photos of their personal interpretation of an emotion. Figure 4 shows a basic screen of the training module. The child is instructed through prompts on what kind of emotion to mimic until the training module is completed.
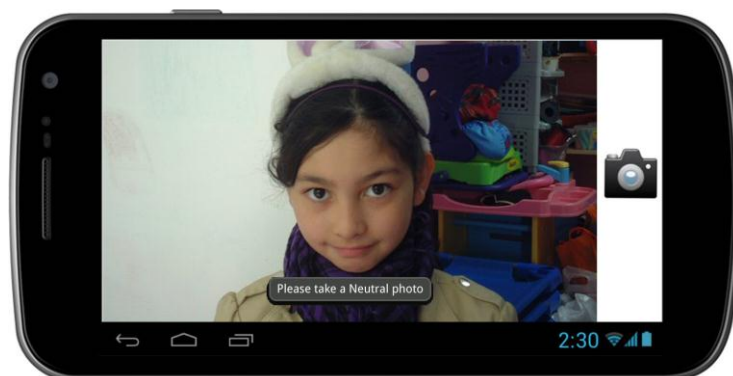
Figure 4: Training module

The child and the carer are given an option for each emotion to either accept or re-take the photos. Once all the photos are taken, the OpenCV training module is activated which operates by opening each image and 'recognising' each facial feature in relation to the subject's eyes. After each image is processed, the results are then averaged between the pictures for a final mapping of the emotional state.

9 Sensor calibration

The sensor operates as a service created by *CaptureMyEmotion*. This ensures that the data collection of the sensor is uninterrupted and allows for collection over a long period of time, which is required in order to obtain the baseline. The sensor data is stored in two formats: short-term and long-term medians. The short-term median shows the average arousal level of the child in a relatively short period of time, such as 10 minutes. The long-term median shows the average over 5 hours. This allows the calculation algorithm to compensate for short term but steady changes of arousal, such as periods of exercise. The gauge icons have their maximum nodes set to double of either of the medians, which in effect allows the needle to rest in the center when they are equal to the baseline, with variations either way correctly influencing the needle.

10   Feedback session with the carer

CaptureMyEmotion uses the pictures and collected emotional data as a basis to discuss emotions with the child and also obtain insight into what makes the child react. The feedback screenshot (Figure 5, left) shows the pictures organised on date, carer-rating or self-reported emotion. Each picture shows the arousal and/or stress range as recorded at the time the picture was taken. The pictures also show the previous rating given and number of times the user and carer reviewed the picture.
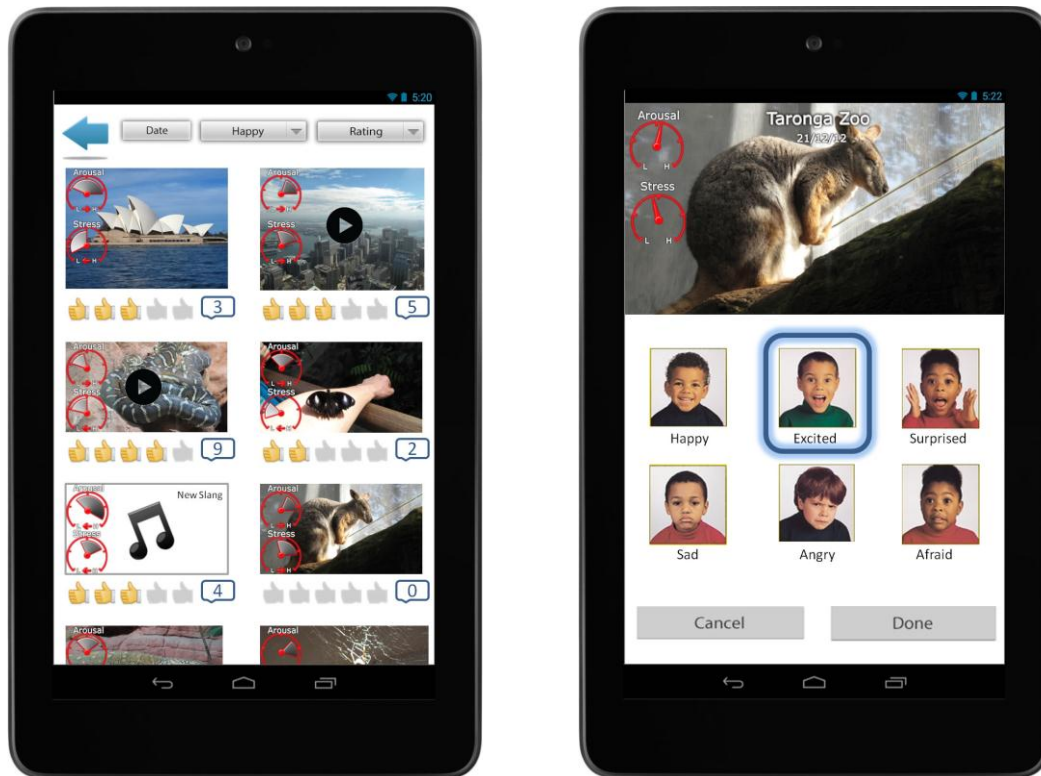
Figure 5: Review pictures

During the feedback session, the child and the carer discuss what the child did and felt using the child's pictures as a basis. When they select a picture, and the child is asked to identify his current emotion while looking at the picture (Figure 5, right).

The data related to that picture is then presented on a summary screen (Figure 6). The right hand side of the screen shows the arousal and/or stress level at the time the picture was taken. The range corresponds to the grey gradient in the gauge, and the general trend (e.g. low to high) is displayed at the bottom of the gauge. It also shows the emotion selected by the child ("You said"), and the result of the facial expression detection software ("Computer said"). Additionally, the carer will see the evaluation data showing the emotion selected by the child during the review. The carer may choose to hide or view all the emotion information on the right hand side by tapping the background picture.

The gauges on the left hand side of the screen (Figure 6) show the live arousal and stress level when the child wears (one or two) sensors. This gives the carer insight into what emotion the picture evokes, or can highlight that the child is getting tired of doing the review or being distracted by other external factors. This can be used in the discussion with the child, or as an indicator that the carer should switch to another activity. The carer may also use the sensor information as feedback to how effective their techniques of interacting with the child are.

The carer can comment and rate the discussion related to the picture as seen fit. The number of thumbs-up allows the carer to express the review result (e.g.: 0-no review, 1-very poor, 5-very good). Data accumulated from the carer and child can be used for the progress report (Figure 7). The progress report displays graphs detailing the number of thumbs, ratings or

pictures accumulated against time. The information viewed on these graphs can be filtered on emotions, locations, ratings or dates.
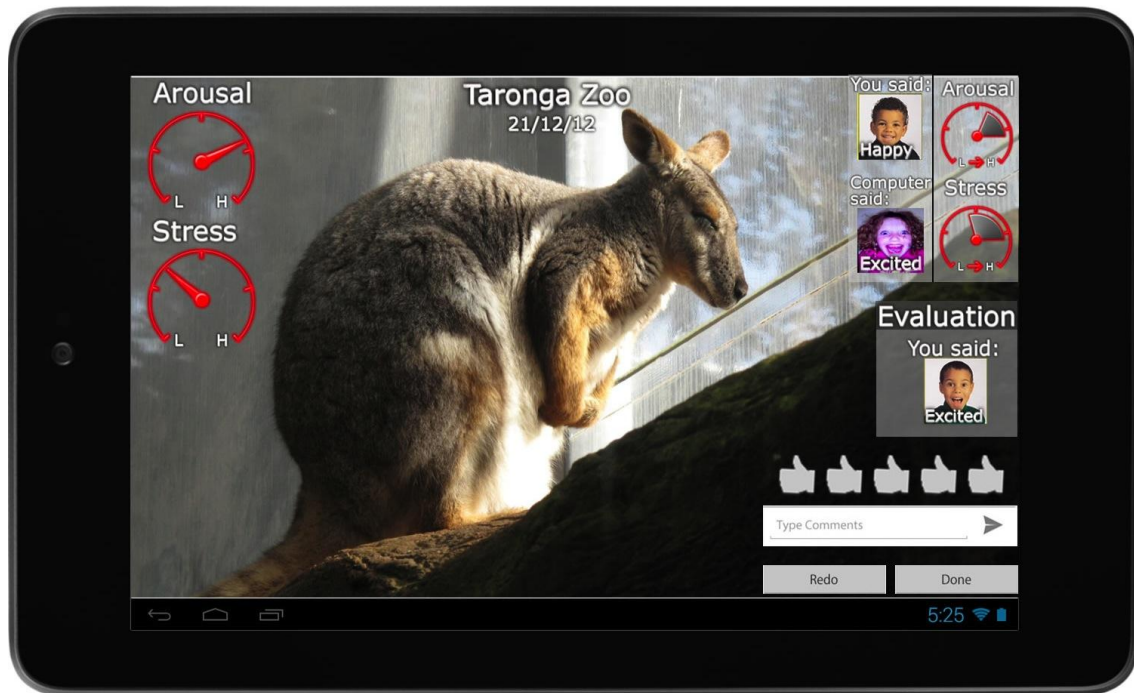

Figure 6: Summary screenshot

CaptureMyEmotion is a tool and (by choice) it allows the carer to choose what they want to do with the captured information. The frequency of use of the app, the type of rating and the number of feedback sessions highly depends on the child's interest and capabilities, as well as, on the carer's objectives.

Figure 7: Progress report screenshot

11   Conclusion

This paper presented *CaptureMyEmotion*, an app for smartphones and tablets. It aims to help autistic children understand their emotions using a smartphone, sensors and facial expression recognition software. The preliminary tests conducted on the facial expression recognition algorithm have been very positive. Utilising a variety of mid to high range smartphones and tablets, the analysis speed can be considered as 'real-time', while the lengthy processes of the

training algorithm varies from around one to five seconds depending on the device. This result is highly favourable to the spatial approach of facial expression recognition, and can only improve as devices continue to improve.

The prototype utilised the Affectiva Q Sensor during testing. While the sensor was functional and reasonably accurate in its function, the algorithm in interpreting an analysis needs to be further examined. The Q Sensor measures skin conductance, which is highly susceptible to external influences, despite long monitoring of a baseline. Testing has shown that spending a whole day in an air conditioned room, and then going out on a sunny day can lead to uncertain results, due to the heat and physical exertion varying thermal regulation and thus the skin conductance. This has been somewhat mitigated through the inclusion of a short-term median, which has met some success in curbing the variance.

Understanding and managing emotions is important for autistic children since it has an impact on how well they integrate into society. We believe that Apps such as *CaptureMyEmotion* have the potential to help them, and their carer in their daily lives.

References

Affectiva. (2013). Affectiva Q website, Retrieved 25/02/2013 from www.affectiva.com/q-sensor.

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association 4th edition.

Anand, B. Navathe, B.B. et Al. Beyond touch: Natural interactions using facial expressions (2012). IEEE Consumer Communications and Networking Conference, 255 - 259.

Autism speaks. (2011), 11 Myths About Autism. Retrieved 25/02/2013 from http://blog.autismspeaks.org/2011/11/21/11-myths-about-autismm-apps.

Autism speaks. (2013), retrieved 25/02/2013 from www.autismspeaks.org/autism-apps.

Baio et Al. (2012). Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008, MMWR 61(SS03). 1-19.

Cho, K.-S., Choi I.-H. and Kim Y.-G. (2012). Robust Facial Expression Recognition Using a Smartphone Working against Illumination Variation. Applied Mathematics & Information Sciences, 6 (1).

de Urturi Z.S., Zorrilla A.M., and Zapirain B.G. (2011). Serious Game based on first aid education for individuals with Autism Spectrum Disorder using android mobile devices. CGAMES USA, 223-227.

Dropbox. (2013). Dropbox website, Retrieved 25/02/2013 from www.dropbox.com.

Du R., Wu Q., He X., Jia W., Wei D. (2009). Facial expression recognition using histogram variances faces, Workshop on Applications of Computer Vision (WACV), pp1-7.

Ekman P. and Friesen W. (1978). Facial action coding system. Palo Alto, CA: Consulting Psychologists Press.

Hardman J. (2012) Garbage truck driver makes time for 11-year-old boy, wins a High 5. Retrieved 25/02/2013 from www.ksl.com/?nid=441&sid=19881661

iAutism. (2013). iPad, iPhone, Android... for Autism and Special Needs. Retrieved 25/02/2013 from www.iautism.info/en/category/application-lists.

Lajevardi, S.M. and Hong Ren Wu. (2012), Facial Expression Recognition in Perceptual Color Space," IEEE Transactions on Image Processing, vol.21, no.8, 3721-3733.

Mahesh M., Myunghoon B.S. and Balakrishnan P. (2012). FaceFetch: A User Emotion Driven Multimedia Content Recommendation System Based on Facial Expression Recognition. 2012 IEEE International Symposium on Multimedia (ISM).

Micro-Expression (2013) Micro-Expression Trainer website. Retrieved 25/02/2013 from http://www.microexpressionstrainingvideos.com/

Moddy Me. (2012). Moody Me-tracker developer's website. Retrieved 25/02/2013 from www.medhelp.org/land/mood-diary-app.

OpenCV. (2013). Retrieved 25/02/2013 from opencv.org/platforms/android.html

Picard R. W. (2009). Future affective technology for autism and emotion communication', Phil. Trans. R. Soc. B. 3575-3584.

Picard R.W. and Goodwin M.S. (2008). Innovative technology: the future of personalized autism research and treatment. Autism Advocate 1 (1), 32–39.

Schlosberg H. (1954). Three Dimensions of Emotion. The Psychological Review, Volume 61, No 2, 81-88.

Shanesh. (2013). Autism Intervention Portal. Retrieved 25/02/2013 from www.shanesh.com

Song K.-T. Song and Chien S.-C. (2012). Facial expression recognition based on mixture of basic expressions and intensities. IEEE International Conference on Systems, Man, and Cybernetics (SMC), 3123-3128.

Tie Y. and Guan L. (2013). "A Deformable 3-D Facial Expression Model for Dynamic Human Emotional State Recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol.23, no.1, pp.142-157.

Wondiko Mood. (2013). Wonkido Mood Journal developer's website, Retrieved 25/02/2013 from www.wonkidovideos.com/social/organizer/32.

Yamaguchi O. (2012). Face Recognition Technology and Its Real-World Application, Perception and Machine Intelligence Lecture Notes in Computer Science Volume 7143, 28-34.

Zephyr. (2012). Zephyr website. Retrieved 25/02/2013 from www.zephyranywherestore.com/corporate/dp/B009XDTN80