A Face Based Real Time Communication for Physically and Speech Disabled People

Chapter · January 2011

DOI: 10.4018/978-1-60960-541-4.ch003

Ong Chin Ann
PSB Academy

Marlene lu
Swinburne University of Technology

Bee Theng Lau
Swinburne University of Technology

Assistive and Augmentive Communication for the Disabled:

Intelligent Technologies for Communication, Learning and Teaching

Lau Bee Theng
Swinburne University of Technology, Malaysia

Information Science
REFERENCE

Senior Editorial Director:	Kristin Klinger
Director of Book Publications:	Julia Mosemann
Editorial Director:	Lindsay Johnston
Acquisitions Editor:	Erika Carter
Development Editor:	Joel Gamon
Production Editor:	Sean Woznicki
Typesetters:	Jennifer Romanchak, Deanna Zombro & Michael Brehm
Print Coordinator:	Jamie Snavely
Cover Design:	Nick Newcomer

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com/reference

Library of Congress Cataloging-in-Publication Data

Assistive and augmentive communication for the disabled : intelligent technologies for communication, learning and teaching / Lau Bee Theng, editor.
p. cm.
Includes bibliographical references and index.
Summary: “This book provides benefits to professionals and researchers working in various disciplines in the field, such as special education, healthcare, computational intelligence and information technology offering insights and support to individuals who are concerned with the development of children and adults with disabilities”--Provided by publisher.
ISBN 978-1-60960-541-4 (hbk.) -- ISBN 978-1-60960-542-1 (ebook) 1. People with disabilities--Means of communication. 2. Communication devices for people with disabilities. 3. Assistive computer technology. I. Theng, Lau Bee, 1974-
HV1568.4.A83 2011
681’.761--dc22
2010054436

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 3
A Face Based Real Time Communication for Physically and Speech Disabled People

Ong Chin Ann
Swinburne University of Technology, Malaysia

Marlene Valerie Lu
Swinburne University of Technology, Malaysia

Lau Bee Theng
Swinburne University of Technology, Malaysia

ABSTRACT

The main purpose of this research is to enhance the communication of the disabled community. The authors of this chapter propose an enhanced interpersonal-human interaction for people with special needs, especially those with physical and communication disabilities. The proposed model comprises of automated real time behaviour monitoring, designed and implemented with the ubiquitous and affordable concept in mind to suit the underprivileged. In this chapter, the authors present the prototype which encapsulates an automated facial expression recognition system for monitoring the disabled, equipped with a feature to send Short Messaging System (SMS) for notification purposes. The authors adapted the Viola-Jones face detection algorithm at the face detection stage and implemented template matching technique for the expression classification and recognition stage. They tested their model with a few users and achieved satisfactory results. The enhanced real time behaviour

DOI: 10.4018/978-1-60960-541-4.ch003

monitoring system is an assistive tool to improve the quality of life for the disabled by assisting them anytime and anywhere when needed. They can do their own tasks more independently without constantly being monitored physically or accompanied by their care takers, teachers, or even parents. The rest of this chapter is organized as follows. The background of the facial expression recognition system is reviewed in Section 2. Section 3 is the description and explanations of the conceptual model of facial expression recognition. Evaluation of the proposed system is in Section 4. Results and findings on the testing are laid out in Section 5, and the final section concludes the chapter.

INTRODUCTION

Communication is a social process of exchanging information from one entity to another in verbal and non-verbal form. It defines our existence and it is an important instrument that connects people together. It comes naturally as a raw skill embedded in most people at birth and we acquired the ways of communication through cognitive learning. Communication is the basis, which drives the process of development in all the fields (Manohar, 2008) and it is the very core of our civilisation. The ability to communicate allows us to express emotion, feelings, convey our thoughts and ideas as well as to relate our experiences. It plays an important role in the dissemination of information and sharing of knowledge especially in the academic arena. Research has found that human started to learn how to communicate with each other since they are born not only through spoken and written languages but also body gesture, posture, facial expression and eye contacts (Busso, et al., 2004; Cohen, Grag & Huang, 2000).

Communication skill might come as a natural ability in majority of people. However, there are some people inflicted with some form of physical defects which affect their ability to communicate. One of the more severe disabilities is known as "cerebral palsy", a congenital disorder at birth which causes abnormality in their motor system. It affects their muscle movement and coordination, learning and speech abilities. Their malfunctioned motor system causes an uncontrollable and involuntary movement. They are unable to control their oral-facial muscles, thus affects their ability to perform facial expression appropriately.

Many assistive tools or formally termed as Alternative and Augmentative Communication (AAC) has been developed and employed to assist people with impaired communication skills. The term encompasses the whole combination of methods used for communication such as text to speech system, pointing gestures, facial expression and body language. Although these AACs have been widely used to assist the disabled, but it is not potentially effective because most AACs are text

Figure 1. Examples of alternative and augmentative communication (AAC) tools

YSpeak program (Softpedia, 2009)

TextAloud Splash (Macsolvers' Blog, 2009)

Talk Back 12 devices (Crestwood Communication Aids Inc, 2009)

STANDUP PUN user interface (Standup, 2006)

to speech and touch screen based applications, which are unsuitable for those with severe physical abilities. There are many kinds of AACs tools available in the market which is shown in Figure 1.

From the limitation of the existing tools reviewed (Novita, 2006; Macsolvers, 2009; Standup, 2006; Universiteit van Amsterdam, 2008; Crestwood, 2009; ScienceDaily, 2008), there is still a pressing need for more effective and efficient tools to alleviate this problem. One the possible methods is to implement a facial expression recognition system to predict or determine the emotional state of a disabled person through his expression projected on his face. The implementation of such method can be made possible through biometrics information systems. According to Gregory and Simon (2008), biometrics information system can be employed as a means to detect and classify the physiological aspect of a person in real time. Franco and Treves (2001) further support the notion that facial expression can be used for human computer interaction and usability enhancement.

Based on the problem statements deliberated above, we propose an improved real time behaviour monitoring application for the disabled by employing real time biometric information i.e. the facial expression recognition system. The aim to create a model that is capable of detecting user's emotion without engaging any physical action from the users. To increase the usability and interactivity of the tool, the emotion detected by the system will be sent to the care-taker's mobile phone in the form of SMS.

FACIAL EXPRESSION

Facial expressions recognition is an ability to recognize people by their facial characteristic and differentiate it with one another. Human is born with the ability to recognize other people easily by identifying their facial features such as shape, appearance, skin texture and skin complexion. Other than that, humans also have the ability to express, interpret and differentiate facial expressions. The regular recurring ones are happiness, anger, disgust, fear, surprise and sad (Ekman & Friesen, 1978). The six facial emotions stated above are important and play a major role in expressing emotion as well as recognising facial expression (Busso, et al, 2004).

In real life, inter personal human interaction are performed not only using speech or spoken language, but also non verbal cues for example hand gesture, body gesture, facial expression and tone of the voice. All these cues are sometimes being used for expressing feeling and give feedback (Busso, et al, 2004; Cohen, et al., 2000). We can see how human interact with each other using non-verbal cues everyday. For example a child cries in front of his mother because he is not happy or dissatisfied with something. Other people might interpret it differently thinking that the child might be in pain.

Facial expression interaction is relevant mainly for community social life, teacher and student interaction, credibility in difference contexts, medicine and so on. Besides, facial expression recognition is useful for designing new interactive devices which offers the possibility of new ways for human computer interaction - HCI (Franco & Treves, 2001). Cohen, et al. (2000) conducted survey on their users and noticed that they have been through traditionally HCI consists of the keyboard, mouse, joystick, trackballs, data gloves and touch screen monitors. The interaction can be improved and enhanced by introducing facial expression recognition that requires no direct contact from the user.

Facial Expression Recognition System (FER) has been a topic for research since Ekman and Friesen (1978) who pioneered this research and worked from the psychology perspective. In the past 20 years, many researchers have tried to adopt their idea and make improvement, innovation and modification on facial expression recognition by introducing different techniques, mainly concentrated on the improvement in term of accuracy, efficiency, mobility, and speed (Kotsia & Pitas, 2007). With all the enhancements on techniques for facial detection and recognition, the development of the facial expression recognition has also improved (Zhan & Zhou, 2007). The most active researches in computer vision and pattern recognition is face recognition in forensic identification, access control, user interface design (Wang, Plataniotis & Venetsanopoulos, 2005), emotion analysis, interactive video, indexing and retrieval of image and video database, image understanding and synthetic face animation (Zhan & Zhou, 2007).

In real world, humans are able to read complex communication where the synthesis of verbal and non-verbal communication is used to express feelings and opinions. Human can interpret and generate major facial expressions but a computer is not built with any facial recognition ability unless through the use of some software. It is even more complicated for the computer to interpret irregular facial expression, especially from those suffering from cerebral palsy. Due to their disorder, they do not have the ability to reflect their emotions like a normal typical person. Thus, a more natural and naive method has to be employed for the system to work by a manual labelling of the image captured with the emotion of the user.

FACIAL EXPRESSION RECOGNITION

The same concept of inter-human interaction can be applied for human-computer interaction in facial expression recognition. A computer uses microphone and camera to “see” and “hear” human expressions and learns to recognize it (Cohen, et al., 2000). The human face is captured by the camera attached to a computer and the captured images is stored for processing with some methods to recognize the identity and emotion of a user based on the features achieved. With automated facial expression recognition, the subjects do not need to operate the computer device or performing any actions in order to get a task done. Besides, facial expression recognition method has been recommended by many researchers to be a good technique to interpret a person’s emotion if compared with other recognition methods such as speech (Zhan & Zhou, 2007; Pantic & Patras, 2006; Lau, 2009).

There are myriads of attempts and research done to produce computer algorithms that serve as models to automate human face recognition function. One of the widely applied systems is known as Facial Action Coding System (FACS). Facial Action Coding System (FACS) was initiated by Ekman and Friesen (1978). It involves the analysis of facial muscle anatomy. It detects the changes to the facial muscle, the contraction and relaxation of a group of muscles to produce certain facial expression. Due to the method employed, it is not suitable to be used by a disabled person namely “Cerebral Palsy” candidates as they have difficulty controlling their muscles and their facial expressions are non-typical. Apart from that, Cohen, Sebe, Garg, Chen, and Huang (2003) stated that FACS processes were very time consuming and tedious.

Another technique is the Skin Color Model that extracts the skin colour from any region of the image as a matching based to detect the face. It will not be accurate if it is taken from other region like legs, arms, and neck appear in the image or video (Singh et. al., 2003). It is also easily affected by the lighting condition and camera settings.

There are other numerous methods developed over the years with their own technique and accuracy. We have chosen to integrate the Viola and Jones Face detection algorithm alongside template matching technique into our image processing engine.

Viola-Jones algorithm is one of the most widely used techniques for face detection and template matching is also the proven effective algorithm for facial expression classification and matching. Viola-Jones algorithm was used by Barlett, Littlewort, Fasel, and Movellan (2003) in their research to perform real time face detection and facial expression recognition process. Shan, Gong, and McOwan (2005) employed template matching techniques to perform person-independent facial expression recognition. They obtained recognition accuracy of 79.1% for 7-class recognition and 84.5% recognition accuracy for 6-class recognition. Refer to the Appendix for a summary of techniques and algorithms used in automatic facial expression recognition system.

We adopted Viola-Jones Face Detection algorithm (2004) and used the default Haar cascade for human face searching because this algorithm is considered the most common and useful algorithm in the field of facial expression recognition research (Cho et al, 2009; Brubaker et al, 2008; Zhan et al, 2006; Datcu and Rothkrantz, 2007; Bartlett, et al, 2003).

As for template matching, it had been a conventional method for object detection and pattern recognition especially facial features at the early stage of face recognition research (Chai, et al, 2009). The advantage of template matching for our proposed prototype is that it is simple, easy to implement, and does not take so much time and memory. We tried on other algorithm such as Support Vector Machine (SVM) as well as Principal Component Analysis (PCA) as both techniques is widely used and recommended by researchers (Dubuisson, et al., 2002; Chen and Huang, 2002; Bartlett, et al, 2003; Littlewort, et al., 2007; Ashraf et al, 2009) but we found that the expression training using these techniques is processor intensive and is consuming memory.

PROTOTYPE MODELLING

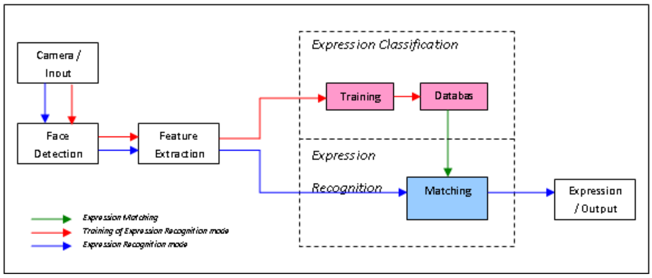
Generally, automated facial expression recognition system involves three major steps or stages. There are the face detection stage, feature extraction stage as well as expression recognition and classification stage (Zhang, Lyons, Schuster, & Akamatsu, 1998; Chibelushi & Bourel, 2002; Zhan et al, 2006). Figure 2 shows two major processes in an automated facial expression recognition system. The blue arrows indicate the process of expression recognition while red arrows indicate the training process of expression recognition. When performing expression recognition, a camera is used to capture the subject's face in video sequence and these video sequences or

Table 1. Summary of techniques and algorithms in automatic facial expression recognition system

Processes	Techniques and Algorithms	Integrated by
Face Detection	Viola and Jones Face Detection Algorithm	Cho, Mirzaei, Oberg, and Kastner (2009); Brubaker, Wu, Sun, Mullin, and Rehg (2008); Zhan, Li, Ogunbona, and Safaei (2006); Datcu and Rothkrantz, (2007); Bartlett, et al (2003);Viola and Jones (2004);
	Motion History Image (MHI)	Davis (2001); Valstar, Pantic and Patras (2004);
	Skin Color Model	Singh, Chauhan, Vatsa, and Singh (2003); Kovac, Peer, and Solina (2003)
	Convolution Neural Network (CNN)	Fasel (2002); Matsugu, Mori, Mitari, and Kaneda (2003); Graves, Fernandez, and Schmidhuber (2007);
Feature Extraction	Principal Component Analysis (PCA)	Dubuisson, Davoine, and Masson (2002); Chen and Huang (2002);
	Gabor Features, Filters and Wavelet	Zhan et al (2006); Tian, Kanade and Cohn (2002); Bartlett, et al (2003); Zhan and Zhou (2007);
	Optical Flow	Donato, et al (1999); Aires, Santana, and Medeiros (2008); Riaz, Mayer, Wimmer, Beetz, & Radig (2009); Su, Hsieh, and Huang (2007)
	Active Appearance Model (AAM)	Zhan and Zhou (2007); Datcu and Rothkrantz (2007); Tang and Deng (2007)
Emotion Recognition and Classification	Facial Action Coding System (FACS)	Zhang and Ji (2005); Ashraf et al, (2009)
	Template Matching	Shan, et al. (2005); Chai, Rizon, Woo, and Tan (2009); Xie and Lam (2009)
	Neural Network (Feed forward and Back propagation)	Ma and Khorasani (2004); Tian, et al. (2002); Franco and Treves (2001);
	Hidden Markov Model (HMM)	Cohen, et al (2003); Zhan and Zhou (2007)
	Linear Discriminant Analysis (LDA)	Lyons, Budynek, Plante, and Akamatsu (2000); Price and Gee (2005)
	Support Vector Machine (SVM)	Bartlett, et al (2003); Littlewort, Bartlett and Lee, (2007); Ashraf et al (2009)
	Dynamic Bayesian Network (DBN)	Cohen, et al (2003); Zhang and Ji (2005); Sebe et. al (2007)

images are used as input to perform the face detection. Once the subject's face is detected, his or her face or region of interest is extracted and temporary stored. The initial emotion of the subject is known after the last process which consist of the comparison and matching using the features extracted and the templates which are stored in the system database. Usually an automated facial expression recognition system needs to be trained before it can be used to recognize and classify human emotion. The process of face detection and features extraction in training stage is

Figure 2. Facial expression recognition system



the similar to the expression recognition process but the extracted features is used as training data for classifying subject's expression.

By implementing automated facial expression recognition system, we proposed a real time behavior monitoring for physical and communication disabled which is shown in Figure 3. The automated facial expression recognition system is mounted on the disabled wheelchair for expression recognition purposes. The system will detect the disabled face, extract the face and perform emotion or expression recognition. Critical or preset abnormal expression recognized will trigger notification module via Short Messaging System (SMS) gateway to notify the parents, teachers and the care taker for the disabled.

Figure 3. Real time behavior monitoring system

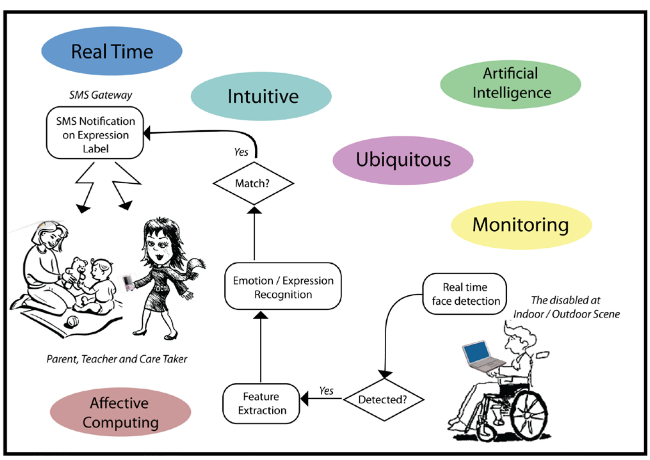
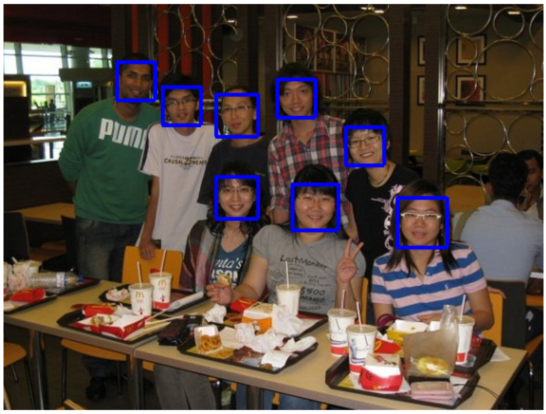


Figure 4. Face detection using Viola-Jones algorithm in still image



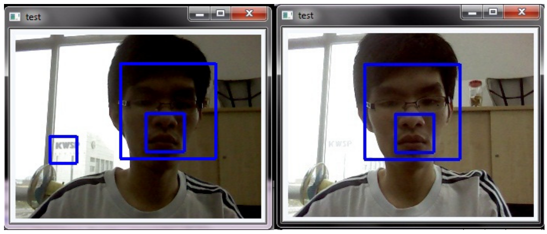
Software and Hardware Specifications

Our proposed automated behaviour monitoring model is built and tested on mobile machine with Intel Core 2 Duo Processor T6600 (2.2 GHz), 2GB DDR3 SD RAM and using 2 MP build in camera for the image capturing process running on PAL standard or 25 frames per second. In the implementation process, we develop our model using Microsoft Visual Studio 2008 C++ Version 9.0.21022.8 RTM with Microsoft .NET Framework Version 3.5 SP1 plug-ins. OpenCV 1.1 (Bradski, & Kaehler, 2008) was used as an assistive library in developing this model either in both face detection stage, feature extraction as well as facial expression training and recognition processes. While for Short Messaging System (SMS) notification, we employed SMS Gateway Development Kit for GSM Modem Type Q24-U-SGDK18 (Mobitek System, 2008) into our model and tested using DiGi as communication service provider.

Real Time Face Detection Algorithm

In face detection stage, we adopted Viola-Jones Face Detection algorithm (2004) and the Haar Cascade for human face searching. In the first few testing, we found that this algorithm could successfully and accurately detect faces in still images but when we tried it on real time video capturing, this algorithm seem to be less intelligent. Below is the snippet of pseudo-codes of the improved algorithm for face detection used in our prototype.

Figure 5. Misdetection of face using Viola-Jones face detection algorithm in real time video capturing



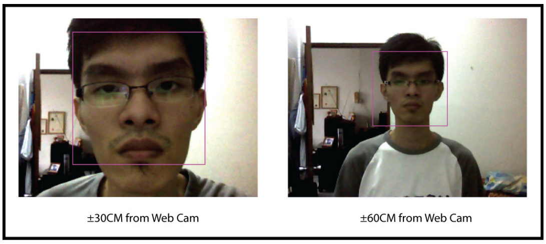
```
INITIALIZE face
WHILE face is equal
     GET face
     IF face is detected
          DRAW region of interest in image
     END IF
END WHILE
```

In Figure 4, we can see the result of Viola-Jones algorithm on still image. All the faces have been detected in still image and drawn with blue square boxes while the same algorithm did not work properly in real time video capturing which is shown in Figure 5. There are multiples detection indicated by the square blue boxes including an object that is not part of the face. This shows an inaccurate detection mechanism and will affect the viability of the facial expression recognition system.

Along the testing period, we found the causes and ways to solve this problem; at the same time we improved the detection accuracy. The Viola-Jones face detection algorithm scan through the whole image or frame using haar cascade classifier and return the face found regardless of the image size. We solve this problem by preset the video frame or resolution to 320 X 240 and minimum face detection or region of interest (ROI) of 100 X 100 pixels. With this scale, the distance between the subject and the camera has to be within 30cm to 60cm which is shown in Figure 6. Hence, small or unwanted objects are no longer detected as it is out of the preset range.

Other than misdetection problem, we also discovered another problem which affected our recognition result. Although the area of the detection or region of interest had been defined, the inconsistent light source is another factor that led to a failure in the detection process. To resolve this issue, we tested on several lighting conditions and tried to gauge the optimum condition suitable for the detection pro-

Figure 6. Face detection range



cess. Then we calculated average brightness value of the optimum lighting condition frame using HSV model (Bezryadin, Bourov, & Ilinih, 2007). When performing facial expression training or recognition, we standardized every captured frame by adjusting the average brightness to the optimum lighting condition value.

Figure 7 shows the result after brightness adjustment or brightness standardization based on the control value or optimum lighting condition. The image on the left is the original frame captured in real time using web cam while the image on the right has been edited to the level of brightness required. By implementing the brightness standardization or adjustment, we not only enhanced the original Viola-Jones face detection algorithm but also increased the recognition rate which will be discussed later under Facial Expression Recognition and Classification sub section.

Facial Feature Extraction Algorithm

Feature extraction is performed once the subject's face is correctly detected. In general, feature extraction can be performed in two approaches: Global feature extraction (Holistically) and local feature extraction. (Fasel & Luettin, 2003; Whitehill, 2006). In our model, we chose global features extraction as it does not consume so much computation time and power when performing facial expression recognition. We added a few image processing processes while performing feature extraction process to further increase the chances of recognition. Refer to the sample pseudo-codes below for an overview of the entire process.

Figure 7. Brightness adjustment

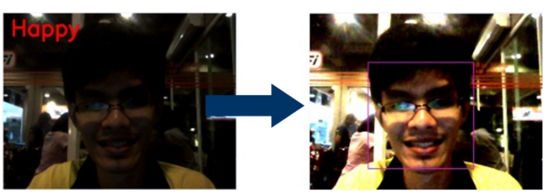
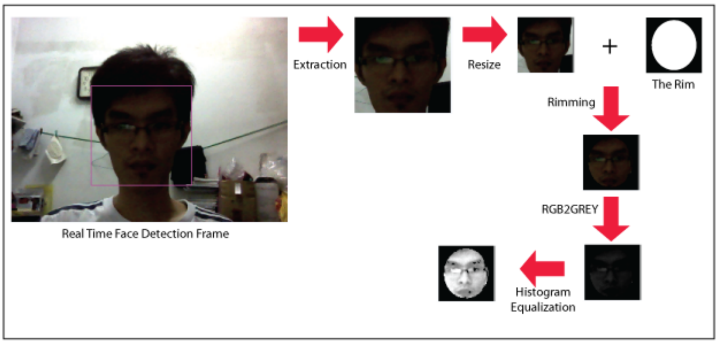


Figure 8. Image processing in feature extraction process

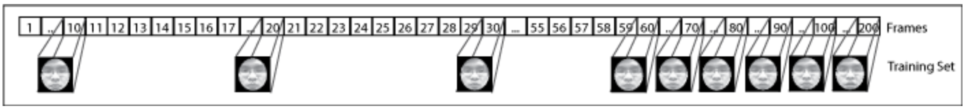



```
IF region of interest is drawn
     EXTRACT region of interest
     RESIZE extracted image
     ADD rim into extracted image
     SET grayscale to extracted image
     SET histogram equalization to extracted image
END IF
```

When the region of interest is detected, it will be extracted and will undergo several levels of modification and filtering. First, we resize the extracted image to 64 × 64 pixels resolution and set it as the standard image size to avoid mismatch or inconsistent template size in recognition process later. We tested it with other resolution such as 128 × 128 pixels but the outcome is similar with the previous resolution and it actually consumed more memory and took more time to process the image. When we tried on smaller resolution, 32 × 32 pixels, we found that the result was not as good as 64 × 64 pixels. Hence we set 64 × 64 Pixels as the standard extracted image size. After the image extraction, we trimmed the resized image to eliminate unwanted area. Then we convert the image from RGB to Greyscale and apply histogram equalization to obtain higher contrast and to increase the intensity of the image. The main purpose of executing all those processes is to reduce system memory usage and processing time when performing facial expression recognition. On top of that, it will simplify the comparison process especially in expression recognition process. The whole processes involved in feature extraction process can be seen in Figure 8.

At the end of the feature extraction process, the histogram equalized image will be used as a template for comparison purposes during the facial expression recognition process. It will be referred to as trained sets.

Figure 9. Neutral expression training set storing process



Training of Facial Expression

Before we run or use our automated facial expression recognition model, we need to feed the model with the user expression. As every human especially the disabled has different ways of expressing their emotions, our model requires the users to train the system to recognize their expression before using it. The procedures for the training function are expressed below.

```
INITIALIZE count
INITIALIZE number of train set
INITIALIZE train frame
FOR each of train frame
CALL face detection
     CALL feature extraction and image processing
IF count modules number of train set is equal to 0
     CALL feature extraction and image processing
     SAVE processed image
END IF
INCREAMENT count
END FOR
```

The purpose of the facial expression training process is to capture and store expression template for matching purposes. During the training session, our model captured 200 frames on each label or expression and stored only 20 frames where each frame is taken after every 10 frames interval. We have tried with fewer samples for each expression but the recognition accuracy seemed to drop compared with 20 samples per expression. Hence, we set a standard of 20 samples per expression, which will be stored in a template.

The attributes of the stored templates are 64 pixels × 64 pixels in resolution, trimmed, and histogram equalized to greyscale image which can be seen in Figure 9. Up to date, we have tested and trained 10 expressions or labels and the outcome in term of speed of storing is satisfactory. All the 20 trained images stored in the template will be used to match the real time captured image during the expression recognition phase.

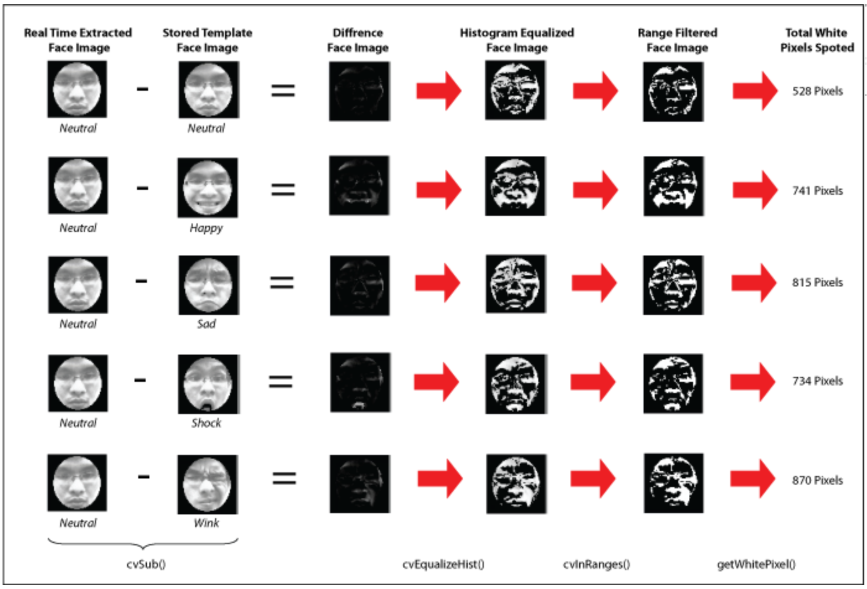
Facial Expression Recognition and Classification

In this final stage, we executed the automated facial expression recognition in real time with the entire enhanced Viola-Jones face detection algorithm. The real time captured image went through similar processes as in the training stage to obtain the histogram equalized image and it will be used to compare with the stored template images through template matching process. Before deciding this algorithm, we tried other algorithm such as Support Vector Machine (SVM) as well as Principal Component Analysis (PCA). Although both techniques are widely used and recommended by researchers (Dubuisson, et al., 2002; Chen & Huang, 2003; Bartlett, et al, 2003; Littlewort, et al., 2007; Ashraf et al, 2009), we found that performing expression training using these techniques required heavy processing power. By training 7 expressions, the system using SVM technique created 60 Megabytes of trained file while our proposed model only created 0.4 Megabytes of trained templates for the same number of expressions. Loading such a large trained file will be put a heavy load on the computer and will slow down the computer's processing speed. Hence we decided to adopt template matching algorithm for expression recognition, which is described below in methodological order.

```
INITIALIZE white pixel
INITIALIZE expression label
INITIALIZE frame capture from camera
IF frame is captured
     CALL face detection
     CALL feature extraction and image processing
FOR each expression label
     FOR each sample
          SUBSTRACT sample image from processed image
          SET histogram equalization to processed image
          SET range filter to processed image
          COMPUTE number of white pixel in processed image

IF number of white pixel in processed image is less than white
pixel
SET white pixel equal to number of white pixel in processed im-
age
               SET expression label equal to sample image label
```

Figure 10. Real time extracted face image and stored template face image matching



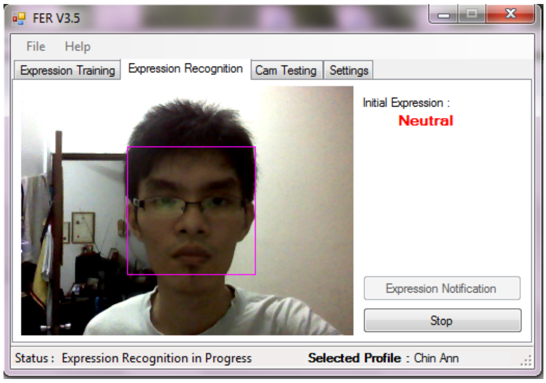

```
          END IF
     END FOR
END FOR
END IF
```

During the matching process, we applied several waves of processing which included image subtraction, another round of histogram equalization and image filtering. We assumed that all the expressions and labels have been fed into the model and the user was sitting within the detection range recommended when performing the facial expression recognition process. The real time captured frame went through the face detection stage followed by the feature extraction stage. The extracted image or processed image will then be used as input of facial expression recognition stage.

Figure 10 illustrated the facial expression recognition process in our model by matching the real time extracted face image with the images stored in the template. Each processed image was compared with every expression stored in the template. Each expression had 20 sample images and there were 5 expressions. Thus, each processed image had to be compared with 100 image templates.

For every sample matching process, our model performed image subtraction by subtracting the pixel value in the stored image template from the processed image to obtain the difference between both images. After that, histogram equalization is

Figure 11. Interface of proposed automated facial expression recognition model with personalization module



applied to the subtracted image so that the contrast and intensity can be distinguished clearly. Then we applied range filtering on the histogram equalized image to convert the image into dual tone color mode consisting of black and white pixels. The area populated with black pixels indicated no changes to the image while white pixel represented the difference between the input image and the image template. Finally our model calculated the total of white pixels and will select the image with the smallest amount of white pixels. Then it returned the emotional state labeled on the image.

Personalizing of User

To make our model more efficient and dynamic, we propose an additional feature called personalizing user's profile to cater for multiple users. It creates multiple user profiles and stores their expressions in their respective profile.

```
CASE of user task
     ADD Profile          : SAVE new profile
     GET profile          : SELECT profile
     REMOVE profile     : ERASE profile
END CASE
```

This feature is useful, increases the efficiency of the model and it saves time. The users are not required to keep retraining our model. The model will save the profile or the trained sets of first time user and retrieved in when needed. The model will automatically detect the user if it could find a match of the facial image from the

Figure 12. Type Q24-U-SGDK18 GSM Modem

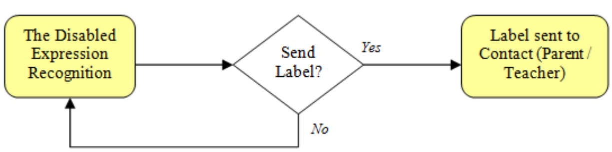
saved templates or profiles, thus automatically skipping the training stage. Figure 11 shows the interface of our proposed facial expression recognition model with personalization module.

Critical Expression Notification

As an enhancement to the facial recognition model, we added Short Messaging System (SMS) notification feature via SMS gateway into our model to notify the care taker, teacher and parent for unusual or critical expression such as anger, sad or shock. We employed SMS Gateway Development Kit for GSM Modem Type Q24-U-SGDK18 (Mobitek System, 2008) into our model which is shown in Figure 12.

```
GET initial expression label
FOR each expression label
IF initial expression label is checked
     SENT sms
END IF
     END FOR
```

Figure 13. Critical expression notification flow chart



User can choose to enable or disable this feature and the user is required to enter his contact number for the notification to work. As shown in Figure 13, when performing facial expression recognition, our model triggered the modem to send subject's initial expression which is marked as "send" to the preset contacts and notify the subject's parent, teacher or caretaker on their critical expression to get immediate attention. Through this, the disabled can be monitored remotely without a close supervision from the care-taker.

EVALUATION

We employed a heuristic approach through multiple sessions of evaluations and testings during the development of our prototype. The evaluations are done on the face detection rates and accuracy, numbers of samples in the template and time required for the training as well as the accuracy of the expression recognition process. A systematic approach is adopted into the testing process. Unit or module testing is done after the completion of each component. Testings were carried out on the face detection, recognition accuracy, response speed of our model, and real time field testing. For face detection module, we tested on an individual under different settings such as indoor environment, outdoor environment, day time as well as night time. For recognition accuracy, we tested on five different users during day time and indoor with white or plain background. We trained our model with 7 expressions for each user and each expression contained 20 sample images. While for response speed, we tested on the same users to examine our model's training time and matching time by training our model with 2 set of expressions per user. Set 1 with 20 samples stored per expression and set 2 with 10 samples stored per expression. For the field testing, we tested with a child who was suffering from physical and communication disabilities. Although we faced some problems and limitations, we managed to solve the major problems.

Face Detection

In the initial stage, we found that Viola-Jones face detection algorithm has some limitation and misdetection especially on the lighting issue. We tried a few ways to solve this problem and fortunately we managed to overcome this problem by adding in auto brightness adjustment based on HSV color model (Bezryadin, et al., 2007). We set a control value or optimum brightness value into our model and it increased or decreased the brightness of the image to the acceptable level before performing face detection. By adding the additional features, our face detection rate has increased from 75% to 95% which is shown in Table 2.

Table 2. Face detection frequency under different environments

Brightness Adjustment Environment		Before (%)	After (%)	Increment (%)
Indoor	Day Time	85	100	15
	Night Time	70	95	25
Outdoor	Day Time	85	95	10
	Night time	60	90	30
Overall Face Detection Frequency		**75**	**95**	**20**

Figure 14 shows the same images captured by webcam in real time. The image on left side was the original image but the face was not detected while image on the right was the brightened image and the face was successfully detected. Base on HSV color model, the average brightness value of the original image is 120 which is considered dark image and the image on the right had been adjusted to 200. It might be the appropriate value for this sample but after testing this setting with other sample images under different light intensity, the images were over-exposed. Therefore, we set the brightness value to the average level of 160.

After multiple testings, we obtained a satisfactory face detection accuracy using this setting under different lighting condition and environment such as indoor, outdoor, day and also night time which is shown in Figure 15. As mentioned before, a good quality of face detection result could increase recognition rates in the later process. The testings conducted proved that the enhancement on Viola-Jones face detection algorithm with automated brightness adjustment could help increase the recognition rate in our model.

Recognition Accuracy

Upon completion of our model, we conducted a mini testing on our model's expression recognition accuracy. We invited 5 normal people to participate in this mini test and have achieved satisfactory outcome when tested on 7 common expression

Figure 14. Brightness adjustment does increase face detection rate and accuracy

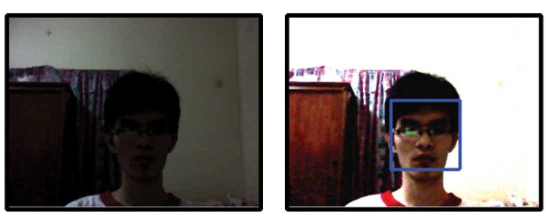


Figure 15. Findings of automated brightness adjustment and face detection

which were neutral, happy, sad, surprise, anger, disgust and confuse. We performed this test indoor during the day time with plain background or a wall, 160 brightness value and 20 sample frames per expression. The result is laid out on Table 3 with 91.4% of the overall expression recognition accuracy. We marked "Match" if our model returned the correct label by our participants and a red label indicated incorrect recognition.

As mentioned in the previous section, we captured 200 frames during the training stage and we stored 10% of the captured frames. To improve the efficiency of the training stage, we tried to reduce the captured frames to 100 frames and 10 frames were stored as the image templates. We found that the overall recognition

Table 3. First mini test matching result of facial expression recognition with 200 train frames and 20 frames stored per expression.

Expression / Participant	**User 1**	**User 2**	**User 3**	**User 4**	**User 5**	**Total**
Neutral	Match	Match	Match	Match	Match	**100%**
Happy	Match	Match	Match	Match	Match	**100%**
Sad	Confuse	Match	Match	Match	Match	**80%**
Surprise	Match	Match	Match	Match	Match	**100%**
Anger	Disgust	Match	Surprise	Match	Match	**60%**
Disgust	Match	Match	Match	Match	Match	**100%**
Confuse	Match	Match	Match	Match	Match	**100%**
Total	**71.4%**	**100%**	**85.7%**	**100%**	**100%**	**91.42%**

Table 4. Second mini test matching result of facial expression recognition with 100 captured frames and 10 frames stored per expression.

Expression / Participant	**User 1**	**User 2**	**User 3**	**User 4**	**User 5**	**Total**
Neutral	Sad	Match	Match	Match	Match	**80%**
Happy	Match	Match	Match	Disgust	Match	**80%**
Sad	Match	Neutral	Match	Match	Match	**80%**
Surprise	Match	Match	Match	Match	Match	**100%**
Anger	Sad	Match	Match	Disgust	Match	**60%**
Disgust	Match	Match	Match	Match	Match	**100%**
Confuse	Disgust	Match	Match	Match	Match	**80%**
Total	**57.1%**	**85.7%**	**100%**	**71.4%**	**100%**	**82.84%**

accuracy has dropped to 82.84% from the first mini test which was 91.42%. Refer to Table 4.

Response Speed

While performing the two mini tests, we monitored and recorded the time for training as well as template matching to compare the processing speed of our model which is shown in Figure 16.

Table 5 shows the results for the two tests conducted. In the first mini test, the average time consumption for training was 22.56 seconds and for template matching was 50 milliseconds. While for the second mini test, the average time consump-

Figure 16. ***Time consumption for training and template matching for mini test one***

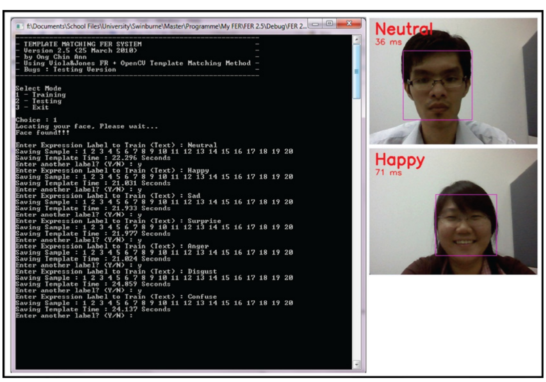


Table 5. Time consumption for expression training and testing session

Tests / Users	Test 1		Test 2	
	Avg. Train Time per Expression (Seconds)	Avg. Test Time (Milliseconds)	Avg. Train Time per Expression (Seconds)	Avg. Test Time (Milliseconds)
User 1	22.28	49	12.78	27.5
User 2	22.18	51.5	12.35	29.5
User 3	22.47	51.5	11.13	27
User 4	22.88	50.5	11.87	25.5
User 5	23.00	47.5	11.62	26.5
Average	**22.56**	**50**	**11.95**	**27.2**

tion for training was 11.95 seconds and average time consumption for template matching was 27.2 milliseconds.

From the result of the experimental testings conducted, the number of frames captured during the training phase will have a domino effect on other components as well. It will reduce the time required to execute the facial expression recognition component and at the same time reduced the recognition accuracy.

The results in both our testing showed the time for expression training and testing session for test 2 has been reduced by 47.03% and 45.6% respectively. But at the same time, the recognition accuracy suffered a decreased from 91.42% in test 1 to 82.84% in test 2.

Another test was conducted to determine the duration for the expression notification to reach the care-taker's mobile phone. We have tested it under normal weather condition by using DiGi as communication operator and the notification arrived within 15 seconds. There are multiple factors that will affect the delivery; the weather condition and the throughput. We did not manage to test this feature under those conditions specified.

Field Test

During the field testing, we tested our model with a cerebral palsy patient labelled as Student A, who suffered from physical and communication disabilities. Student A did not have the ability to walk and had difficulty holding an object with his hands. He had impaired communication skills and his speech was incomprehensible.

The setup of our model for Student A was different from the other users we tested previously. We tested 5 basic expressions with the student which include neutral, anger, happy, sad and surprise expressions. In the expression training phase, we configured our model to extract 5 out of 50 frames captured and saved them in the

Figure 17. Cerebral palsy patient field testing result

Images						**Overall Matching (%)**
Expressions	Neutral	Angry	Happy	Sad	Shock / Surprise	
Matching (%)	100	80	100	60	100	**88**

dataset for expression recognition matching process later. The reduced frame rates were due to inconsistency of the facial pattern while performing expression training.

We performed 1 set of expression training consisting of 5 expressions and 5 rounds of expression recognition during our field testing. We achieved an 88% of overall matching result while performing expression recognition with Student Refer to Figure 17 for the result. This has proven that our proposed model as an assistive communication tool is effective for physically and speech disabled patients.

FINDINGS

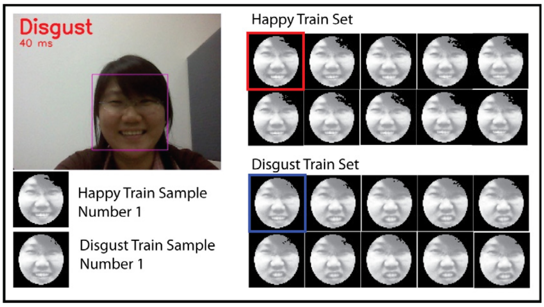
From the testings and observations conducted, we discovered a few potholes. In the face detection stage, we found that the original Viola-Jones face detection algorithm has a few limitations. It is affected by the lighting condition and it produced multiple face detections on similar area which led to misdetection. We have tried various ways to overcome this problem and we managed to solve it by adding an automated brightness adjustment features into an image before performing Viola-Jones face detection algorithm. With the added feature, our model could detect human face regardless of the environment and the lighting condition.

When training the user expression, users have to be consistent with their expression as our model is using pattern recognition or template matching method to perform facial expression recognition. This means that our model only returns the nearest expression matched.

In our second mini testing, our model returned the label “Disgust” while the subject is expressing a happy emotion which is shown in Figure 18. Initially we thought that our model did not work well and misrecognized the subject emotion. But when we explored further into the subject’s trained set, we found that the misrecognition was caused by the incorrect or inappropriate training conducted.

Based on Figure 18, we can see a set of happy expressions and a set of disgusted expressions. The first frames of both emotions reflect almost similar expressions.

Figure 18. Incorrect training causing misrecognition



When the real time captured imaged is compared with the stored template, it returned an incorrect result. The model detected a disgusted emotion while the participant was actually feeling happy. Thus, it is important that the training phase should be conducted accurately to obtain a valid result.

In term of processing and respond speed, our proposed model could respond within 100 milliseconds to our user in template matching stage. For recognition accuracy, our proposed model could achieve 91.42% accurate rate in distinguishing 7 types of expressions for the same person conducted in the mini test and 88% of recognition accuracy in the field test with our targeted user. Through our testings, we found that by collecting more samples per expression, it will increase the accuracy when performing expression recognition. At the same time, it will increase the duration required for the training and recognition phase.

CONSTRAINTS

In the face detection stage, we have enhanced the detection rate and accuracy by adding the auto brightness adjustment base on HSV colour model into the naïve Viola-Jones face detection algorithm. However, in some cases, our model still failed to detect human face in certain conditions as shown in Figure 19. When the

Figure 19. Backlight is too strong

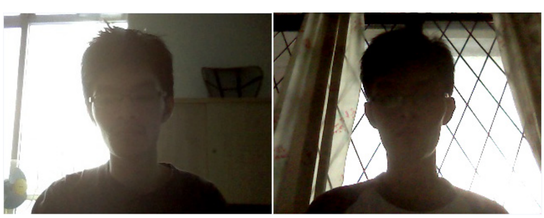
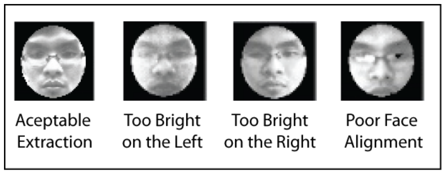


Figure 20. Face detection and extraction



backlight is too bright, it will overshadow the face turning it into a silhouette. The model will not be able to detect the underexposed face.

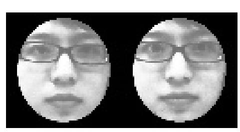
The second constraint on detection stage is the distance between the subject's face and the web cam. As we have preset our video or capture frame resolution to 320 × 240 pixels and minimum detection area is 100 × 100 pixels, the subject's head must be within 30 cm to 60 cm from the webcam. Otherwise our proposed model will fail to detect the face. It is important that the captured image and the extracted image have to be in good quality. We use the term "good quality" to indicate a proper alignment between the face and the camera in an upright position with proper lighting as shown in Figure 20. A poor face detection will reduce the recognition accuracy. Hence, users have to perform expression recognition in a controlled environment with suitable lighting.

In training stage, our model required a time frame of 20.024 seconds to 24.137 seconds to store a set of expression which consisted of 20 images extracted from 200 frames captured. Our users complained that the duration of the training time is too long. But to achieve a higher accuracy of recognition rate (91.42%), we have to capture at least 200 frames.

For critical expression notification via SMS, the delivery of the notification is affected by the weather condition, communication operators as well as their connection speed.

Up to date, we did not find any critical constraint for expression recognition stage as long as the input for facial expression training is properly conducted; we believe our model could obtain higher recognition accuracy than 91.42%. One issue that is worth discussing and taken into consideration is that certain people may appear indifferent. For example some people might feel sad even though their face showed

Figure 21. One face pattern may represent many expressions, Sad on the left; Neutral on the right



a neutral look which is shown in Figure 21. This is beyond our model's ability to distinguish from one expression to another because the facial pattern is the same.

FUTURE WORKS AND IMPROVEMENT

After a series of evaluations and continuous testing and probing on our proposed model, we have identified several limitations on a few stages.

The first unforseen issue emerged when we did the training session with a cerebral palsy boy who could not prop his head up and position his face directly in front of the camera. His head was constantly tilted to the right. The face detection mechanism failed to detect his face completely. Further exploration is required to improve the face detection mechanism taking this into consideration to give a wider angle of flexibility. There are rooms for improvements to increase the facial expression recognition rates. We are delving into ways to improve the quality of the face detection result as it will have a direct impact on the expression recognition process.

Other than that, we going to fine tune some settings when performing image processing in features extraction stage. With that, we hope we could refine the quality of the extracted face images for later processes. In the training stage, we will do more research specifically on how we could shorten the training time and reduce the number of templates within a train set while maintaining the recognition rate.

CONCLUSION

Communication is the very essence of our daily lives. The inability to communicate will hamper us from moving forward as an individual and as a community, robbing us from a normal life. Humans have tried ways to communicate since the beginning of time and communication has evolved throughout ages. It is a complex process triggering many senses and it comes naturally to most of us. Unfortunately there is a small group of people with physical and communication disabilities since birth. To assist these people in their communication needs, we propose an improved real time behaviour monitoring application for the disabled by employing a real time facial expression recognition system and Short Messaging System (SMS) to send notification to the third party for monitoring purposes. Our proposed automated real time behaviour monitoring has been designed and implemented with the ubiquitous and affordable concept in mind which suits the underprivileged who suffers from both physical and communication disabilities. In facial expression recognition system, we employed Viola-Jones face detection algorithm which is commonly used by researchers to detect the subject face and apply template matching methods for expression classification and recognition. The emotional state of the inflicted party will be sent to the care-taker in a form of SMS. From the testing result of our proposed model, the face detection rate was approximately 95% under different lighting condition and environment. At the same time, we achieved 82.84% to 91.42% recognition accuracy on average when conducting expression recognition testing. The processing speed for the expression training took 11.95 seconds to 22.56 seconds per expression and 27.5 milliseconds to 50 milliseconds for expression recognition process. While developing and testing our model, we identified the limitation of using naïve Viola-Jones face detection algorithm and the way to overcome this problem. Besides that, we discovered that by using template matching method for recognition, some expressions cannot be distinguished by the machine if the facial expression of one emotion has close similarities with the other. This would be another domain for further research on how a machine could distinguish different emotional state but reflected with the same facial expression. Apart from that, we are going to enhance and improve a few modules on our model especially the face detection stage as well as the feature extraction stage to increase the recognition accuracy. With the enhanced real time behaviour monitoring system, it could make the disabled life easier and assist them anytime and anywhere when needed. They can do their own tasks more independently without being physically monitored or accompanied by their care taker, teacher or even parent.

ACKNOWLEDGMENT

The author would like to thanks Noris, Valeria, N., Mandasari, V., Tran, C.T., Jain, N., Sim, S., Lau, T.C., Wee, S., as well as Liew, J. for participating in the proposed model testing as well as allowing the author to use their photos in Figure 4, Figure 17, Figure 18 and Figure 21.

REFERENCES

Aires, K. R., Santana, A. M., & Medeiros, A. A. D. (2008). Optical flow using color information: Preliminary results. *Proceedings of the 2008 ACM Symposium on Applied Computing.* Fortaleza, Ceara, Brazil: ACM.

Ashraf, A. B., Lucey, S., Cohn, J. F., Chen, T., Ambadar, Z., Prkachin, K. M., & Solomon, P. E. (2009). The painful face–pain expression recognition using active appearance models. *Image and Vision Computing*, *27*(12), 1788–1796. doi:10.1016/j.imavis.2009.05.007

Bartlett, M. S., Littlewort, G., Fasel, I., & Movellan, J. R. (2003). Real time face detection and facial expression recognition: Development and applications to human computer interaction. *Computer Vision and Pattern Recognition Workshop, 5,* 53.

Berhad, M. I. M. O. S. (2008). *National biometrics technology roadmap*. Ministry of Science, Technology and Innovation, Malaysia. Retrieved December 31, 2009 from http://www.mosti. gov.my/ mosti/ images/ stories/ DICT/ policy/ Biometrics% 20Technology% 20Roadmap% 20Public% 20Version.pdf

Bezryadin, S., Bourov, P., & Ilinih, D. (2007). *Brightness calculation in digital image processing*. Technologies for Digital Fulfillment, KWI International Inc. Retrieved March 25, 2010, from http://www.kweii.com/site/color_theory/2007_LV/BrightnessCalculation.pdf

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library.* O'Reilly Media, Inc. ISBN: 9780596516130

Brubaker, S. C., Wu, J., Sun, J., Mullin, M. D., & Rehg, J. M. (2008). Fast asymmetric learning for cascade face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*, 369–382. doi:10.1109/TPAMI.2007.1181

Busso, C., Deng, Z., Yildirim, S., & Bulut, M. Lee, C. M., Kazemzadeh, A., ... Narayanan, S. (2004). *Analysis of emotion recognition using facial expression, speech and multimodal information*. International Conference on Multimodal Interfaces.

Chai, T. Y., Rizon, M., Woo, S. S., & Tan, C. S. (2009). Facial features for template matching based face recognition. *American Journal of Applied Sciences*, *6*(11), 1897–1901. doi:10.3844/ajassp.2009.1897.1901

Chen, X., & Huang, T. (2003). Facial expression recognition: A clustering-based approach. *Pattern Recognition Letters*, *24*, 1295–1302. doi:10.1016/S0167-8655(02)00371-9

Chibelushi, C. C., & Bourel, F. (2002). *Facial expression recognition: A brief tutorial overview*. University of Edinburgh, UK. Retrieved February 4, 2010, from http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CHIBELUSHI1/CCC_FB_FacExprRecCVonline.pdf

Cho, J., Mirzaei, S., Oberg, J., & Kastner, R. (2009). *FPGA-based face detection system using Haar classifiers.* International Symposium on Field Programmable Gate Arrays.

Cohen, I., Grag, A., & Huang, T. S. (2000). *Emotion recognition from facial expression using multilevel HMM*. Neural Information Processing Systems.

Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). Facial expression recognition from video sequences: Temporal and static modeling. *Computer Vision and Image Understanding*, *91*, 160–187. doi:10.1016/S1077-3142(03)00081-X

Crestwood Communication Aids Inc. (2009). *Product: Talk Back12*. Retrieved October 14, 2009 from http://www.communicationaids.com/talkback12.htm

Datcu, D., & Rothkrantz, L. (2007). Facial expression recognition in still picture and video using active appearance models: A comparison approach. *ACM International Conference Proceeding Series, 285.*

Davis, J. W. (2001). *Hierarchical motion history images for recognizing human motion*. IEEE Workshop on Detection and Recognition of Events in Video, (pp. 39-46).

Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Classifying facial action. *Advances in Neural Information Processing Systems*, 8.

Dubuisson, S., Davoine, F., & Masson, M. (2002). A solution for facial expression representation and recognition. *Signal Processing Image Communication*, *17*, 657–673. doi:10.1016/S0923-5965(02)00076-0

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: Investigator's guide*. Palo Alto, CA: Consulting Psychologists Press.

Fasel, B. (2002). Multiscale facial expression recognition using convolutional neural networks. In *Proceedings of the Third Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2002).*

Fasel, B., & Luettin, B. (2003). Automatic facial expression analysis: A survey. *Pattern Recognition*, *36*, 259–275. doi:10.1016/S0031-3203(02)00052-3

Franco, L., & Treves, A. (2001). *A neural network facial expression recognition system using unsupervised local processing*. Image and Signal Processing and Analysis.

Graves, A., Fernandez, S., & Schmidhuber, S. (2007). Multi-dimensional recurrent neural networks. *Proceedings of the 17th International Conference on Artificial Neural Networks* (pp. 549-558). Porto, Portugal: Springer-Verlag

Gregory, P., & Simon, M. A. (2008). *Biometrics for dummies*. Wiley Publishing, Inc.

Kotsia, I., & Pitas, I. (2007). Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Transactions on Image Processing*, *16*(1), 172–187. doi:10.1109/TIP.2006.884954

Kovac, J., Peer, P., & Solina, F. (2003). Human skin colour clustering for face detection. In B. Zajc (Ed.), *EUROCON 2003—International Conference on Computer as a Tool*, Ljubljana, Slovenia, Sept. 2003.

Lau, B. T. (2009). Portable real time needs expression for people with communication disabilities. In D. Versick (Ed.), *Communication in computer and Information Science: Intelligent interactive assistance and mobile multimedia computing* (pp. 85-95). International Conference, Rostock-Warnemunde, Germany, Nov 2009. Berlin/Heidelberg, Germany: Springer-Verlag.

Lin,D. T.(2006). Facial expression classification using PCA and hierarchical radial basis function network.

Littlewort, G. C., Bartlett, M. S., & Lee, K. (2007). *Faces of pain: Automated measurement of spontaneous facial expressions of genuine and posed pain*. International Conference on Multimodal Interfaces. Oral session 1: Spontaneous behaviour (pp. 15-21).

Lyons, M., Budynek, J., Plante, A., & Akamatsu, S. (2000). Classifying facial attributes using a 2-D gabor wavelet representation and discriminant analysis. In *Proceedings of the 4th international conference on automatic face and gesture recognition*, (pp. 202–207).

Ma, L., & Khorasani, K. (2004). Facial expression recognition using constructive feed forward neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 34*(3). Macsolvers' Blog. (2009). *Testing the text to speech function*. Retrieved October 12, 2009 from http://macsolvers.wordpress.com/2009/03/24/text2speech/

Manohar, U. (2008). Why is communication important? *Buzzle.com*. Retrieved November 3, 2009, from http://www.buzzle.com/articles/why-is-communication-important.html

Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, *16*, 555–559. doi:10.1016/S0893-6080(03)00115-1

Mobitek System. (2008). *SMS gateway development kit*. Mobitek System. Retrieved February 24, 2010, from http://www.mobitek.com.my/SMS_Gateway/SMS%20 Gateway.html

Novita Children's Services. (2006). *Alternative and augmentative communication* (AAC). Retrieved October 23, 2009 from http://www.novita.org.au/Content.aspx?p=64

Pantic, M., & Patras, I. (2006). Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences. *IEEE Transactions on Systems, Man and Cybernetics—Part B, 36*(2).

Price, J. R., & Gee, T. F. (2005). Face recognition using direct, weighted linear discriminant analysis and modular subspaces. *Pattern Recognition*, *38*, 209–219. doi:10.1016/S0031-3203(04)00273-0

Riaz, Z., Mayer, C., Wimmer, M., Beetz, M., & Radig, B. (2009). A model based approach for expression invariant face recognition. In Tistarelli, M., & Nixon, M. S. (Eds.), *Advances in biometrics: Third International Conferences* (pp. 289–298). Alghero, Italy: Springer Berlin. doi:10.1007/978-3-642-01793-3_30

ScienceDaily. (2008). *Facial expression recognition software developed*.

Sebe, N., Lew, M. S., Sun, Y., Cohen, I., Gevers, T., & Huang, T. S. (2007). Authentic facial expression analysis. *Image and Vision Computing*, *25*(12), 1856–1863. doi:10.1016/j.imavis.2005.12.021

Shan, C., Gong, S., & McOwan, P. W. (2005). *Robust facial expression recognition using local binary patterns*. IEEE International Conference, Image Processing, 2005. ICIP 2005.

Singh, S. K., Chauhan, D. S., Vatsa, M., & Singh, R. (2003). A robust skin color based face detection algorithm. *Tamkang Journal of Science and Engineering*, *6*(4), 227–234.

Softpedia. (2009). *Y speak*. Retrieved October 20, 2009, from http://mac.softpedia.com/progScreenshots/Y-Speak-Screenshot-10550.html

Standup. (2006). *The STANDUP system.* Retrieved October 12, 2009 from http://www.csd.abdn.ac.uk/research/standup/software.php

Su, M. C., Hsieh, Y. J., & Huang, D. Y. (2007). Facial expression recognition using optical flow without complex feature extraction. *WSEAS Transactions on Computers*, *6*(5), 763–770.

Tang, F., & Deng, B. (2007). *Facial expression recognition using AAM and local facial features*. Third International Conference Natural Computation, 2007. ICNC 2007. (p. 632).

Tian, Y., Kanade, T., & Cohn, J. F. (2002). *Evaluation of Gabor wavelet based facial action unit recognition in image sequences of increasing complexity*. Fifth IEEE International Conference, (pp. 229–234).

Universiteit van Amsterdam. (2008). *eMotion: Emotion recognition software.* Retrieved October 12, 2009, from http://www.visual-recognition.nl/eMotion.html

Valstar, M., Pantic, M., & Patras, I. (2004). *Motion history for facial action detection in video*. International Conference on Systems, Man and Cybernetics.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision, January 10, 52*(2), 137-154.

Wang, J., Plataniotis, K. N., & Venetsanopoulos, A. N. (2005). Selecting discriminate eigenfaces for face recognition. *Pattern Recognition Letters*, *26*, 1470–1482. doi:10.1016/j.patrec.2004.11.029

Whitehill, J. R. (2006). *Automatic real-time facial expression recognition for signed language translation*.

Xie, X., & Lam, K. M. (2009). Facial expression recognition based on shape and texture. *Pattern Recognition*, *42*(5), 1003–1011. doi:10.1016/j.patcog.2008.08.034

Zhan, C., Li, W., Ogunbona, P., & Safaei, F. (2006). A real-time facial expression recognition system for online games. *ACM International Conference Proceeding Series, 207*.

Zhan, Y., & Zhou, G. (2007). Facial expression recognition based on hybrid features and fusing discrete HMM. In Shumaker, R. (Ed.), *Virtual reality, HCII 2007, (LNCS 4563)* (pp. 408–417). Berlin/ Heidelberg, Germany: Springer-Verlag.

Zhang, Y., & Ji, Q. (2005). Active and dynamic information fusion for facial expression understanding from image sequence. *Pattern Analysis and Machine Intelligence, 27*.

Zhang, Z., Lyons, M., Schuster, M., & Akamatsu, S. (1998). *Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron*. Third IEEE International Conference, (pp. 454–459).

ADDITIONAL READING

Agrawal, N., Cosgriff, R., & Mudur, R. (2009). Mood detection: Implementing a facial expression recognition system. *Stanford University*. Retrieved 31 January 2010, from http://www.stanford.edu/class/cs229/proj2009/AgrawalCosgriffMudur.pdf

Huang, X. & Lin, Y. (2008). A vision-based hybrid method for facial expression recognition. *1st international conference on Ambient media and systems*

Jemaa, Y.B. & Khanfir, S. (2009). Automatic local Gabor features extraction for face recognition. *International Journal of Computer Science and Information Security, Vol. 3*, No. 1.

Zhao, Y., Shen, X., Georganas, N. D., & Petriu, E. M. (2009). Part-based PCA for Facial Feature Extraction and Classification. *Haptic Audio visual Environments and Games, 2009*. HAVE 2009. IEEE International Workshop. 99 – 104.

KEY TERMS AND DEFINITIONS

Assistive Communication Tools: Tools used to assist physical and communication disabled peoples to communicate or interact with the others.

Biometrics: A method of verifying a person's identify by analysing a unique physical attribute of the individual for example fingerprints, face, DNA, etc.

Computer Vision: A process of taking a live raster image that represented as a matrix of numeric values and interpreting it into higher level data abstractions and symbolic objects such as humans, limbs, faces, props, poses, gestures, etc.

Facial Expression Recognition System: A tool used to identify and recognize human expression or emotion by providing human face image.

Image: Processing: Involve complex calculation and transformation of image(s) into another form, size, model and channel for example RGB, HSV, Greyscale and Black and White.

Template Matching: Involving the process matching between two or more images at the same time generate matching result.

Viola-Jones Face Detection: A common face detection algorithm used by researchers to search human's face pattern within an image.