An Assistive Communication Model for Children with Cerebral Palsy

Ong Chin Ann

A thesis submitted in fulfillment of the requirements for the degree of

Masters of Science (by Research)

Performed at

Swinburne University of Technology

2012

Abstract

Cerebral Palsy children with multiple disabilities have a combination of several disabilities that may include: speech, physical mobility, intellectual disability, visual, hearing and possibly others. Consequently, communication becomes a common difficulty of these children. Speech therapists and pathologies suggest the adoption of Augmentative and Alternative Communication (AAC) tool to help them in their communication. However, most of the suggested AAC tools like touch or text-to-speech based applications require good hand coordination skills. Thus, most of the AAC tools are not effective in improving their communication with others.

This research proposes an assistive communication tool (ACApp) with natural human computer interface to improve their communication. The proposed prototype recognizes their facial expressions and responses either by alerting with an alarm or sending messages through Short Messaging System (SMS) to the guardians' mobile phones. There were 21 children with Cerebral Palsy from a special education school being invited to participate in the prototype evaluation. These children were categorized as high and low functional groups according to their disabilities. The results showed that the proposed prototype recognize facial expressions and response with the accuracy of 82.30% for the high functional group and 81.89% for the low functional group respectively. The false positive detection was 5.65% and 18.22% for the high and low functional groups respectively. Evaluations were also conducted to investigate the effectiveness of the critical expression messages delivery to their guardians, 98.46% of SMS messages (with an average of 8.27 seconds of waiting time) were delivered successfully to the mobile phones. The proposed prototype is proven to be able to assist children with Cerebral Palsy in their communication.

Acknowledgement

The author would like to express his gratitude to both supervisors, Dr. Lau Bee Theng and Assoc. Prof. Dr. Henry Lee Seldon for their guidance, encouragement and patience. He would like to show his utmost appreciation to Mdm. Doreen Tie the principle of the special school as well as all the teachers, staff, parents and also students for their assistance on the collection of data. He would like to sincerely thank Tran Cao Thai for his help and discussions on the process of prototype development.

The author would also like to thank Serena Sim, Nia Valeria, Vivi Mandasari, Jofry Hadi Sutanto, Shane Wee Jo Nah, Lau Tiu Chung and all his other friends at Swinburne University of Technology, Sarawak Campus, who have made the years of his Master Degree candidature a part of his life that he will always cherish.

Finally, the author would like to thank his family members and relatives for their unconditional love and support.

Declaration of Originality

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university, and to the best of my knowledge contains no material previously published or written by another person, except where due reference is made in the text of the thesis. Work based on joint research or publications in this thesis fully acknowledges the relative contributions of the respective authors or workers.

ONG CHINANN

Signature: ______________________

Date: ________________________

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

This thesis describes a research on the development of a communication tool through the use of assistive technology. It proposes an assistive communication tool for inter personal-human communication and monitoring of the children with Cerebral Palsy. The proposed model developed by employing real time facial expression recognition and alert notification through sound and Short Messaging System (SMS) to monitor children with Cerebral Palsy and indirectly helping them to communicate. In this research, the targeted users are the children with Cerebral Palsy and the level of severity is categorized into high functional and low functional groups.

This research reviews the characteristic of children with Cerebral Palsy, the causes, their disabilities as well as their needs, the existing Augmentative and Alternative Communication (AAC) tools and its limitations, biometric information technologies in facial expression recognition and how it works. This includes the attributes as well as the existing or related techniques used to develop the facial expression recognition.

The study of the methodologies and models on existing facial expression recognition is conducted by analyzing the techniques employed to the facial expression recognition system. This includes the limitation and constraints of the systems, and system performance in terms of recognition accuracy, processing speed, efficiency and system mobility.

Along the research process, related products, models and applications are reviewed as references for the proposed model. The proposed model's design is based on the needs and problems faced by children with Cerebral. The proposed system is handy and easy to use, which fulfils the needs of children with Cerebral Palsy.

1.1 Research Background

Cerebral Palsy (CP) occurs in people who experience disability caused by brain damage during or before birth or in the first year of their lives which resulting in a loss of voluntary muscular control and coordination (Badawi et al. 1998; Columbia Electronic Encyclopedia 2008). The main causes of brain damage during or before birth includes natural or accidental occurrences i.e. birth injury or brain injury, drug and alcohol consumption, maternal infection, neonatal infection as well as premature birth (Krigger 2006 ;4MyChild 2009).

Generally, Cerebral Palsy can be classified into four groups which are Spastic (up to 80%), Athetoid or Dyskinetic (10% – 20%), Ataxic (5% – 10%) and mixed (Krigger, 2006; 4MyChild, 2009), with each group of Cerebral Palsy having its own specific disabilities (4MyChild 2009). According to Badawi et al. (1998), Krigger (2006) and also information obtained from 4MyChild (2009) and Farlex Inc. (2009), the common disabilities experienced by Cerebral Palsy patients include difficulty in walking around and doing simple exercises (Spastic), difficulty in communicating with others and having trouble sitting, walking, or speaking clearly (Athetoid or Dyskinetic), drooling all the time as they may have trouble controlling their facial muscles, and trouble tying their shoes, buttoning their shirts, cutting with scissors, and doing other tasks that require fine motor skills (Ataxic) .

As stated previously, most of the disabilities experienced by children with Cerebral Palsy are related to daily communication. For instance, children with Athetoid or Dyskinetic have problems speaking and they may produce incomprehensible speeches which are hardly to be understood by normal people (4MyChild 2009). Some of them do not only experience speech disability but also physical disability, limiting their body gesture and affecting communication skills.

Usually, children diagnosed with Cerebral Palsy cannot be cured or recover from the disability, but some treatment can improve a child’s capabilities. 4MyChild (2009) stated that some medical research reported that many patients could enjoy near-normal lives if their neurological problems were properly managed. There is no one standard therapy that works for all patients, but their capabilities to perform some tasks can be improved by identifying their unique needs and impairments. Based on that, an individual treatment plan can be created. Speech therapists and pathologies (Davis, Moore, and Storey 2003; Pennington, Goldbart, and Marshall 2004; Chapman 2009; Rummel-Hudson 2011) suggested for children with Cerebral Palsy to use AAC in order to overcome communication difficulties.

The common functionality of AAC includes natural communication methods such as pointing, gestures, facial expressions, and body language (AbleData 2010; ASHA 2010; ISAAC 2010). Sign language such as Makaton is particular to an individual to helps children to understand language easier and effectively.

Many types of AAC tools available in the market are recommended by speech pathologist (4MyChild 2009; Novita Children's Services 2006). They are also listed in Michigan's Assistive Technology Resource (2002), AbleData (2010) and some are suggested by the American Speech-Language-Hearing Association (ASHA 2010) as well as the International Society for Augmentative and Alternative Communication (ISAAC 2010) websites. A few examples of AAC such as Natural Reader by NaturalSoft (2010), YSpeak (Softpedia 2009), STANDUP PUN (2006), Talk Back 12 by Crestwood Communication Aids Inc. (2009), GoTalk Express 32 (Attainment Company Inc 2010), Auggie (Cooper and Associates 2010) and Dynavox (2010) were reviewed and most of these AAC tools are text to speech and touch screen based device or application.

Recent research has found limitations and ineffectiveness in using AAC in assisting and helping for children with Cerebral Palsy to communicate. Wilkins and Ratajczak (2009) explained that a user must be able to receive information or instruction from others, manipulate the information and then make correct decision or response in order to use high tech AAC devices effectively. Otherwise these AAC tools would be considered ineffective. Gonzales, Leroy, and Leo (2010) further elaborated that AAC tools could not be used in real life situation due to the lack of adaptability to the user and also their communication disability. In some cases, children with Cerebral Palsy do not understand the AAC tools that they employ (Bouck 2010) and have difficulties in learning or adapting to all the features (Gonzales, Leroy, and Leo 2010). Hence, AACs are no longer effective in assisting children with Cerebral Palsy to communicate because most AACs are text-to-speech and touch screen based applications which are not suitable for children with physical disabilities.

In short, a task oriented AAC i.e. touch screen based or text-to-speech based applications cannot solve communication difficulty faced by children with severe Cerebral Palsy as they cannot touch, click, or press a button, type a single word or make a speech for input to the AAC. Furthermore, children with Cerebral Palsy do not understand how to use the device due to their intellectual disability in some cases. As a result, these AAC tools are not able to assist them to communicate. In order to improve the efficiency of existing AAC application and devices, a new trend of AAC is required.

1.2 Research Problem

Children with Cerebral Palsy experience communication difficulty which is caused by intellectual, speech and physical disabilities (4MyChild 2009; Yale Medical Group 2011). Speech therapist and pathologies claimed that this difficulty could be solved by using augmentative and alternative communication tools (Davis, Moore, and Storey 2003; Pennington, Goldbart, and Marshall 2004; Chapman 2009; Rummel-Hudson 2011).

On the contrary, Wilkins and Ratajczak (2009), Gonzales, Leroy, and Leo (2010) and (Bouck 2010) disagreed that AAC could help children with Cerebral Palsy to communicate as most of the AACs were text-to-speech and touch screen based which required fine motor coordination to operate. Therefore, current AACs are less efficient in helping children with Cerebral Palsy to Communicate with people around them.

Consequently, this research would analyze the existing AACs, improve and design an assistive communication model that uses intuitive human computer interaction style for the children with Cerebral Palsy.

1.3 Proposed Solution

In order to overcome the research problems, this research adopted the facial expression recognition as the core module in the assistive communication tool. According to Branco and Encarnacao (2004), the human face is rich with communication information. This communication information called facial expression is always visible on human faces when they communicate with one and another. Sometimes, these facial expressions are also being used to give responses and feedbacks (Busso et al. 2004; Cohen, Grag, and Huang 2000; Ekman 1997).

Figure 1-1 briefly illustrates the overview on how the proposed assistive communication model could help children with Cerebral Palsy to communicate.

This communication model requires a portable computer like the laptop with a built-in video camera, and external GSM modem. Then the laptop is attached on the wheelchair so that the child with Cerebral Palsy can be monitored at any time and any place.

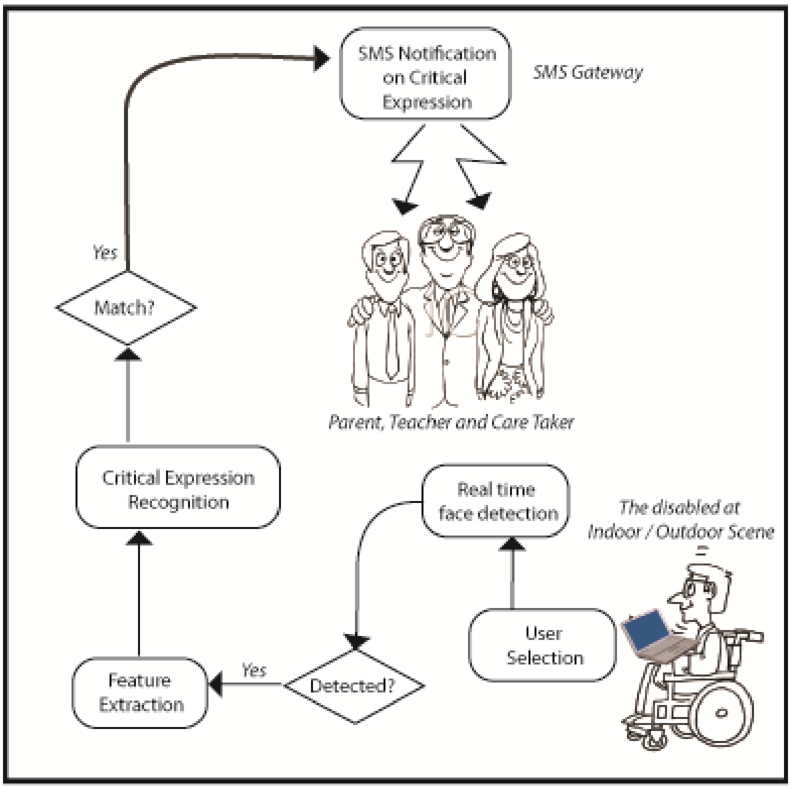


Figure 1-1 The proposed assistive communication model

During the monitoring session, the child's facial expression is captured and distinguished by the proposed model. A notification will be sent using Short Messaging System (SMS) to his or her guardian's mobile phone when pre-defined critical expressions are detected during the monitoring session. A further elaboration on the proposed solution will be discussed in section 3.3 later.

1.4 Research Goal and Objectives

The goal for this research is to develop an assistive communication tool for children with Cerebral Palsy and to overcome the limitations of the existing AACs.

Six research objectives are to be achieved. These objectives include investigating the disabilities and communication needs of children with Cerebral Palsy, investigating the existing assistive tools and technologies, analyzing the usage of facial expression recognition, designing an assistive communication tool, developing the prototype and evaluating the developed prototype.

a. To investigate the disabilities and communication needs of children with Cerebral Palsy

Understanding the disabilities and need of children with Cerebral Palsy is one of the main objectives to be obtained in this research. This objective is to be achieved by performing investigation on their disabilities and communication needs through literature review including reviewing published papers and journal articles from online databases. It is also achieved by observing their daily social communication through visiting a special school for children with Cerebral Palsy.

Important data and information related to these children is gathered and analyzed and the outcome of the investigation is produced by summarizing the collected data.

b. To investigate existing assistive tools and technologies

This objective aims to provide better solution to help and assist children with Cerebral Palsy by comparing each of tools and technologies obtained. The information on the assistive and related tools and technologies for Cerebral Palsy will be gathered by reviewing related products, models and applications from published articles as well as from the internet. The strengths and weaknesses of applications reviewed are identified and used as references for system design and development.

c. To analyze the usage of real time facial expression recognition

The objective strives to analyze the most suitable technique for the proposed model in this research. The use and application of the basic facial expression recognition system will be studied by reviewing published papers and journal articles from online databases. Throughout the reviews, common techniques and algorithms used to develop facial expression recognition system are further analyzed and compared.

d. To design an assistive communication tool

This objective aims to design a model for the proposed prototype. A real time communication tool will be designed based on the outcome and requirements obtained from previous investigation and analysis. The design requirement includes the disabilities and communication needs of children with Cerebral Palsy, the features in terms of the strengths and weaknesses of existing assistive tools and technologies as well as the used of current facial expression recognition systems.

e. To develop the prototype

This objective realizes the assistive communication prototype based on design model. This objective can be achieved by adopting a programming language (i.e. C++) and external APIs such as OpenCV, Mobitek SMS Gateway Development Kit, and .NET Frameworks for the user interface.

f. To evaluate the prototype

The proposed prototype is tested on children with Cerebral Palsy as evaluation the idea proposed in this research to cater the need of children with Cerebral Palsy. The feedback on efficiency, effectiveness and adaptability is evaluated and brought to discussion at the end of the testing.

1.5 Research Methodology

This is a 24-month research on assisting children with Cerebral Palsy in their communication, by employing an assistive communication tool. The research methodology is based on the following steps such as identification of the problem area, collection and organization of data, interpretation of data, action based on data and reflection. The aforementioned steps are carried out though the following tasks:

a. Literature Review

Articles on facial expression recognition, children with Cerebral Palsy and related products of assistive communication tools are reviewed. The purpose of the literature reviews is to understand facial expression recognition, children with physical and speech disabilities as well as the assistive communication tools available.

b. Related products and application investigation

Assistive communication tools and products for children with Cerebral Palsy are reviewed which to evaluate the strengths and weaknesses of every product. The strength of the product is adopted and kept for the proposed system. Modification is performed based on the problems and weaknesses of the reviewed products.

c. Gather and analyze facial expressions from children with Cerebral Palsy

The needs and problems of children with Cerebral Palsy are identified. The information on the children is gathered and studied by visiting schools for children with Cerebral Palsy.

d. Prototype modeling and implementation

A prototype is designed based on the needs and disabilities of children with Cerebral Palsy by developing an assistive communication prototype to help them communicate with people around them.

e. Prototype evaluation with the volunteers

The developed prototype is tested on children with Cerebral Palsy and the results are used for modification and amendment after conducting the analysis. The final testing result is brought into discussion and was used as supporting data for the proposed model in this research.

1.6 Research Schedule

The research is planned to be completed in 24 months which is shown in Table 1-1.

Table 1-1 Research schedule

Stages	Begin Date	End Date
Literature Review	17 August 2009	31 December 2009
Analysis	1 January 2010	30 March 2010
Prototype Design	1 April 2010	31 August 2010
Prototype Implementation	1September 2010	31 December 2010
Evaluation of prototype	1 January 2011	31 March 2011
Dissertation Wrap-Up	1 April 2011	31 August 2011

1.7 Overview of Thesis Structure

The whole thesis is divided into 6 chapters. This is Chapter 1 where the introduction and general information relating to this research is discussed. Chapter 2 discusses the collection of literature related to this research, which most of the key issues include Cerebral Palsy, human communication, existing Augmentative and Alternative Communication (AAC) tools, biometric information, and facial expression recognition system.

Chapter 3 discusses the conceptual modeling of the proposed prototype where a draft prototype is produced based on the requirement gathered in the literature review and site observation. Chapter 4 further elaborates the implementation of assistive communication application from stage to stage based on conceptual modeling with experiments.

The first part of Chapter 5 discusses the features of participants and experimental procedure or preparation for prototype evaluation while the second part discusses evaluation result. Lastly, Chapter 6 concludes the entire research.

Chapter 2: Literature Review

This chapter reviews all the related literature which is associated to five major key sections: Cerebral Palsy, human communication, Augmentative and Alternative Communication tools, biometrics information and facial expression recognition system (FERS).

The literature review begins by studying children with Cerebral Palsy and its related attributes i.e. characteristics of Cerebral Palsy, types of Cerebral Palsy, causes of Cerebral Palsy and also their disabilities. Human communication will be discussed later as it is one of the difficulties experienced by children with Cerebral Palsy caused by their speech and physical impairment. The review proceeds to Augmentative and Alternative Communication (AAC) tools as suggested by speech therapists and pathologies whereby these tools could overcome communication difficulties faced by children with Cerebral Palsy.

In a further review, AAC is found to be having limitations and thereby unable to assist children with Cerebral Palsy well up to a certain extent. Recent research proposed the adoption of biometric information utilizing human facial expressions as a solution to tackle the limitation of AAC and thereby overcoming communication difficulties experienced by children with Cerebral Palsy. Therefore, a facial expression recognition system was required. This research made a further study of facial expression recognition system processes, attributes, and techniques required.

The findings and outcome from literature reviewed are discussed in the critical review section and the conclusion for the entire chapter is located at the end of this chapter.

2.1 Cerebral Palsy

Cerebral Palsy (CP) is a term used for people who experience disability caused by brain damage during or before birth or in the first years of their lives, resulting in a loss of voluntary muscular control and coordination (Badawi et al. 1998; Columbia Electronic Encyclopedia 2008). Hyperbaric Service of the Palm Beaches (2009) uses the term "Cerebral Palsy" to describe a multitude of chronic disorder which affect body movements. This disorder is not caused by the functionality problems of the muscles or nerves, but rather of the brain's ability to adequately control the body.

Morris (2009) states that previous researches has struggles in coming to terms with the definition and classification for Cerebral Palsy for over 150 years, in his historical studies on Cerebral Palsy. He finally quotes Rosebaum et al.'s (2006) definition of Cerebral Palsy as "*Cerebral Palsy describe a group of permanent disorder of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of CP are often accompanied by disturbances of sensation, perception, cognition, communication behavior, by epilepsy and by secondary musculoskeletal problem."*

According to information provided by Krigger (2006) and 4MyChild (2009), Cerebral Palsy can be caused by facts or by accidents, for example birth injury, brain damage or brain injury, drug and alcohol consumption, maternal infection, neonatal infection as well as premature birth. From all the examples given, most of the occurrences causing Cerebral Palsy are related to pre and post born babies. A baby will have more chances of birth injuries in an incident of a prolonged labor period. When labor duration is longer than expected, it may cause the baby to be lodged in the birth canal and subsequently suffer oxygen deficiency (asphyxia).

Other than that, fever occurred to the mother during pregnancy will also cause brain injury to the baby and giving opportunities for Cerebral Palsy to develop. A baby may also suffer brain damage or injury during its pre-birth state if the baby's mother suffers from a blood clotting disorder. This disorder will stop the flow of blood from the mother to the baby. Brain damage may occur for a baby after birth when internal bleeding happens inside their brain.

A pregnant woman who takes drugs, consumes alcohol and who smokes regularly will also increase the percentage of the baby getting Cerebral Palsy (4MyChild 2009). The toxin from the mother's blood stream can be transferred to the baby's system, resulting in brain damage to the baby.

Lastly, pregnant women infected with viruses such as Rubella, Toxoplasmosis and Urinary Tract Infection (UTI) or kidney infection will have greater chances of getting a Cerebral Palsy baby. However, this infection is avoidable by taking precautionary steps.

2.1.1 Types of Cerebral Palsy

Cerebral Palsy can be classified into four groups, which include Spastic (up to 80%), Athetoid or Dyskinetic (10% – 20%), Ataxic (5% – 10%) and mixed (Krigger 2006; 4MyChild 2009). The classification for Cerebral Palsy is shown in Figure 2-1. Each of these groups of Cerebral Palsy has its own specific disabilities (4MyChild 2009).

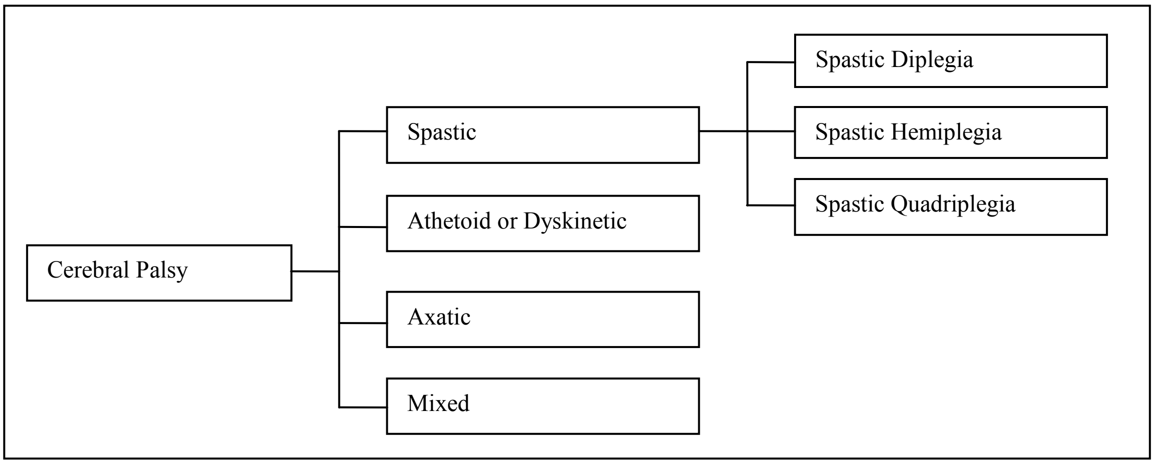


Figure 2-1 Classification of Cerebral Palsy

Information obtained from 4MyChild (2009) shows that a patient experiencing rigidity in his or her muscle is diagnosed as the spastic type of Cerebral Palsy. He/she may also experience difficulties in walking around and in doing simple exercises. This is the most common type of Cerebral Palsy and it can be divided into another three subtypes namely Spastic Diplegia, Hemiplegia and Quadriplegia.

Spastic Diplegia is the condition of a person who has tight muscles on both the legs and hip, with legs crossing at the knees and producing a “scissor” walking pattern. The second subtype of spastic Cerebral Palsy is Spastic Hemiplegia, a condition where half of the body is stiff. This means that the patient leg or hand may not develop normally, with paralysis on one side of the body and that will increase tendon reflexes and uncontrolled contraction in the affected muscles (Farlex Inc. 2009). The severest subtype of spastic Cerebral Palsy is Spastic Quadriplegia in which the patients will experience seizures and difficulties in walking or even talking. This group of patients may have mental retardation and bodily impairment.

Athetoid or Dyskinetic is the second most frequently diagnosed type of Cerebral Palsy. Patients infected with this disease will have normal intelligence, but his or her body will be totally affected by muscle problems. In this category, the body muscle tone can be weak or tight, and the patient might have trouble sitting, walking, or speaking clearly. Some of them may be drooling all the time as they have trouble controlling their facial muscles.

The least diagnosed type of Cerebral Palsy is called Ataxic. People who are infected with this type of disease will have trouble tying their shoes, buttoning their shirts, cutting with scissors, and doing other tasks that require fine motor skills. They might walk with their feet farther apart than normal and have trouble with their balance and coordination. Such patients may also suffer from "intention tremors", a shaking that begins with a voluntary movement.

For mixed Cerebral Palsy, the patient may be diagnosed with more than one type of Cerebral Palsy stated above.

2.1.2 Cerebral Palsy Recovery

Usually Cerebral Palsy cannot be cured. Neither can one recover from it, but some treatment can improve a child's capabilities. 4MyChild (2009) state that some medical researchers report that many patients can enjoy near-normal lives if their neurological problems are properly managed.

There is no one standard therapy that works for all patients. The capabilities of patients to perform certain tasks can be improved by identifying their unique needs and impairments. From there, an individual treatment plan can be created. The treatment plan includes drug consumption to control seizures and muscle spasms; special braces to compensate for muscle imbalances; surgery; mechanical aids to help overcome impairments, counseling for emotional and psychological needs, as well as physical, behavioral, and occupational needs; and speech therapy. In short, the earlier the diagnosis and treatment begin, the better chance a child has of overcoming developmental disabilities and of learning new ways to accomplish difficult tasks (Special Needs Hope 2006).

2.1.3 Discussion for Cerebral Palsy

This research has focused on Cerebral Palsy for discussion which is related to a further research later. This group of children experience disability caused by brain damage during or before birth. They have difficulties in verbal and nonverbal communication, social interactions and unusual repetitive or severely limited activities and interests.

From the reviews (Badawi et al. 1998; Krigger 2006; Columbia Electronic Encyclopedia 2008; Hyperbaric Service of the Palm Beaches 2009; Farlex Inc. 2009), the factors causing children to have Cerebral Palsy have been identified. Most of the causes are related to the infection of or damage to the human brain which in turn affects the motor skill of the child. It can happen during birth or when an accident has occurred to the child. Some of the causes can be prevented, for example, maternal infection, drugs and alcohol, but some are unavoidable.

Generally, the disabilities of children with Cerebral Palsy are summarized in Figure 2-2.

- Difficulty in walking around and doing simple exercises (Spastic).
- Difficulty in communicating with others (Mixed)
- Having trouble in sitting, walking, or speaking clearly (Athetoid or Dyskinetic).
- Constant drooling due to an inability in controlling facial muscles (Spastic).
- Trouble tying shoelaces, buttoning shirts, cutting with scissors, and doing other tasks that require fine motor skills (Ataxic).

Source: Badawi et al. 1998; Krigger 2006; 4MyChild 2009; Farlex Inc. 2009

Figure 2-2 Disabilities experienced by children with Cerebral Palsy

Most of the disabilities acquired affect their daily communication, for example, children with Athetoid or Dyskinetic will have problems communicating with others as they may produce incomprehensible speeches when talking (4MyChild 2009).

In short, children with Cerebral Palsy experience difficulties in their physical movements and also in communicating with other people. Some of them have problems in speaking, causing them to produce fragmentary speeches which are hard to be understood. The severe type of Cerebral Palsy has a worse plight as these children lack the ability to show any gesture to communicate with others.

2.2 Human Communication

Communication was the process of transferring information from one entity to another and it is a very important aspect of the human life. According to Manohar (2008), communication is the basis which drives the process of development in all fields. It helps human beings connect with each other as individuals and as independent groups. Besides, it is helpful in the aspect of information dissemination as it allows humans to send and receive information.

In Smith's (2004a) lecture notes, communication has various definitions taken from different perspectives. It could mean *1) the transmission of information with a focus on both the sender and the receiver parties, 2) the connection between persons caused by the understanding of verbal and nonverbal message transmitted, 3) the process of sharing human feelings and information via an exchange of verbal and nonverbal messages, 4) sharing of ideas through transmission of messages, and 5) the creation of shared understanding through participation and interaction in between two or more parties*. Tanova and Nadiri (2010) claims communication also helps in expressing emotion, feelings and ideas.

Thus, communication can be defined as the social interaction involving information exchange or information delivery between sender and receiver. Through communication, everyone can express their feelings and ideas in oral and written forms. It includes the ability in decision making as well.

In education and human development, communication plays a major role, for example, in the exchange of knowledge and information from one peer to another. Every human being has the right to express his/her needs, to connect with his/her community, to establish relationships with one another, to exchange information and to be involved with each other in all aspects of life. Without access to a means of communication, people are likely to experience marginalization, social isolation, discrimination, victimization, and inequitable compromised services and opportunities. Therefore, communication is very important to the human race as commented by Manohar (2008), and Tanova and Nadiri (2010).

2.2.1 Types of Communication

There are two types of communication, the verbal and non-verbal communication. Verbal communication is the transmission of information directly from one party to another party through a spoken language or speech (Smith 2004b). Nonverbal communication could be categorized as any type of communication other than speech (Mehrabian 2007). This type of communication involves written communication as well as communication transmitted through sign language, finger spelling, Braille, or other similar alternatives to verbal communication (Smith 2004b). From other perspectives, nonverbal communication may include body gesture i.e. hand and arm gestures, positions and movements of the human body, postures and facial expressions as well.

2.2.2 Communication Disorder relating to Cerebral Palsy

Communication disorder is impairment in the ability to receive or process a symbol system or a representation of a concept or symbol and to transmit it to the other party or receiver (Nicolosi, Harryman and Kresheck 2004). Collier and Blackstone (2009) state that communication disabilities may affect speaking, understanding, reading and writing. These defects are caused by disorders in the sensory (deafness or hearing loss, blindness or visual impairments), the motor (inability to speak or use gestures because of complete or partial paralysis), the intellectual or cognitive, and language-acquisition (conditions that can limit learning and understanding of spoken or written language, memory, organization and independent function).

In short, physical and speech disabilities acquired by children with Cerebral Palsy have similar attributes to the communication disorder defined by Nicolosi, Harryman and Kresheck

(2004).They are related to the disability to receive, send, process, and comprehend concepts and graphic symbol systems; and to conduct verbal or nonverbal communication. This is because both disabilities can be developmental or acquired and the causes are believed to be based on biological problems, the brain. This includes drug and alcohol consumption, an exposure to toxins during pregnancy, maternal infection, neonatal infection, premature birth, birth injury and abnormalities of brain development after birth. (Krigger 2006; 4MyChild 2009; Yale 2011).

As a result, children with Cerebral Palsy or communication disorder may have certain behavior and features that are different from normal people such as having a limited vocabulary, inability to speak at all, difficulty in understanding simple instructions or in naming objects, and difficulty in expressing abstract ideas (Yale 2011).

Throughout literature reviews for both Cerebral Palsy and communication, it is noted that the human daily life is very much depending on communication and children with Cerebral Palsy experience communication difficulties due to physical and speech impairment. In order to help and assist children with Cerebral Palsy in their communication, speech therapists and pathologies (Davis, Moore, and Storey 2003; Pennington, Goldbart, and Marshall 2004; Chapman 2009; Rummel-Hudson 2011) suggest the adoption of Augmentative and Alternative Communication (AAC) tool.

2.3 Augmentative and Alternative Communication for Children with Cerebral Palsy

Augmentative and Alternative communication (AAC) tools are devices that seek to "*increase, maintain, or improve the functional capabilities of individuals with disabilities*" (Davis, Moore, and Storey 2003). Novita Children's Services (2006) define Augmentative and Alternative Communication as "*a term used for all communication that is not speech, but is used to enhance or to replace speech.*" The AAC System encompasses the whole combination of methods used for communication, for example, gestures, eye pointing, vocalizations and pointing to symbols. It is also a form of communication that attempts to compensate for the impairment and disability of children with severe expressive communication disorders i.e. children with Cerebral Palsy, through the use of symbols, signs and devices.

The AAC system is very effective and useful for children with physical and speech disabilities (Hodge 2007). It helps the patients to improve their communication and socialization skills not only with their care takers and parents but with others as well. Hodge adds that the AAC system can train them in developing language, knowledge and in encouraging the subject to participate in any activities that are involved.

Overall, the advantage of using the AAC system is to be able to monitor patients and to have control over what happens to them from time to time. Some of the examples and stories of children who may need the AAC system have been documented by Novita Children's Services (2006) and it shows that the AAC system does help them to communicate well with the public.

2.3.1 Existing AAC tools

There are many types of AAC system tools available and are recommended by speech pathologist (4MyChild 2009; Michigan's Assistive Technology Resource 2002; AbleData 2010; Novita Children's Services 2006; ASHA 2010; ISAAC 2010). They also include natural communication methods such as pointing, gestures, facial expression, and body language.

Sign language such as Makaton, is particularly useful in helping children understand language easier and more effectively. Object symbols are objects that are small versions or parts of objects which represent an activity, object or person, for example, a key or a seat belt can represent a car or driving, a nappy or toilet paper can represent toilet, a spoon can represent eating and so on. Communication picture boards and displays are sets of photos, drawings, symbols or words that can be used by patients for communication. Speech generating device, displaying icons or symbols, generate a voice message when a symbol is pressed. Text-to-speech or spellings is the most commonly used device or software and is highly recommended by speech pathologists.

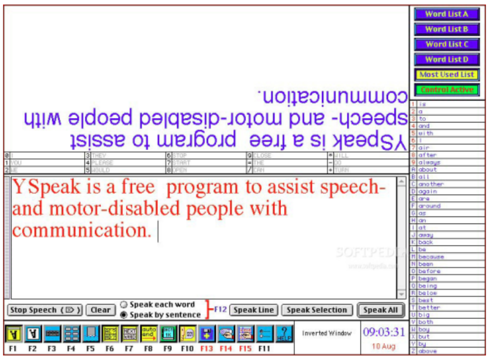
Many other AAC tools (a.k.a Speech Generating Devices - SGD) are available in the market. They are listed in Michigan's Assistive Technology Resource (2002), AbleData (2010) and some are suggested by the American Speech-Language-Hearing Association (ASHA 2010) as well as the International Society for Augmentative and Alternative Communication (ISAAC 2010).

A few AAC tools are found online including Natural Reader by NaturalSoft (2010), YSpeak (Softpedia 2009), STANDUP PUN (2006), Talk Back 12 by Crestwood Communication Aids Inc. (2009), GoTalk Express 32 (Attainment Company Inc. 2010), Auggie (Cooper and Associates 2010), Dynavox (2010), and there are many others which can be found in AbleData website.

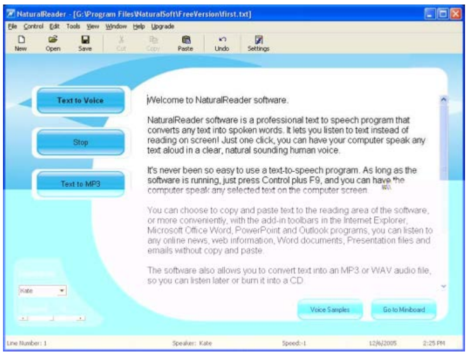
Due to the numerous AAC tools available in the market, it is impossible to review each and every single AAC's feature in detail. However, all these AACs can be further classified into two categories, which are the text-to-speech based AAC and also the touch screen based AAC.

Generally, the text-to-speech based AAC can be either applications or devices which generate speeches based on texts inserted by the user while the touch screen based AAC can be either applications or devices that show a list of pictures, icons or symbols where speech is generated when the user presses the buttons or screen. Samples and screenshots for both text to speech and touch screen based AAC is shown in Figure 2-3.

Text to Speech Based

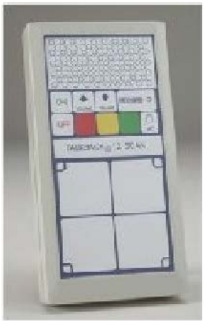


YSpeak



Natural Reader

Touch Screen Based



TalkBack 12

STANDUP PUN

Dynavox Meastro

GoTalk Express 32

Auggie

Figure 2-3 Sample of Text to Speech and Touch Screen based AAC

2.3.2 AAC Limitation

Based on reviews in section 2.3, most of the children with speech and physical disabilities, including children with Cerebral Palsy, experience communication difficulty (4MyChild 2009; Yale 2011) and their difficulty could be solved by using augmentative and alternative communication tools (Davis, Moore, and Storey 2003). However, recent research has found limitations of AAC tools in assisting children with Cerebral Palsy in their communication.

According to Wilkins and Ratajczak (2009), a user must be able to receive information or instruction from others, manipulate the information and then make correct decisions or responses in order to use the high tech AAC devices effectively. Otherwise, these AAC tools would not able to assist them at all. Children with multi disabilities such as speech, physical and intellectual impairment may find it hard to use touch screen based application or devices. They may get confused with all the buttons appearing on the screen and may find it hard to learn all the features on the device. The same scenario may become worse if the children try to use text-to-speech based AAC.

Gonzales, Leroy, and Leo (2010) claim that existing AAC tools cannot be used in real life situations due to a problem of adaptability of the users and also their communication impairment. Usually, children with Cerebral Palsy do not understand how to operate the AAC tools that they employ (Bouck 2010) and have difficulties in learning or adapting to all the features. Therefore, AAC is not effective in assisting children with Cerebral Palsy in their communication as most of AAC tools are text-to-speech and touch screen based applications that are not suitable for children with severe physical disabilities.

In short, task oriented AAC tools such as touch screen based or text-to-speech based application cannot solve communication problems faced by children affected with such disabilities as they cannot touch, click or press a button, type a single word or make a speech as input for the AAC. Furthermore, these children with Cerebral Palsy do not understand how to use the device due to their intellectual disability. Thus, AAC tools are deemed inappropriate to assist them to communicate with others. In order to improve the efficiency of existing AAC application and devices, a new trend or pattern of AAC is required.

2.4 Biometrics Information to Assist the Communication

A new model is needed to assist children with Cerebral Palsy to communicate and to overcome the limitations of existing Augmentative and Alternative Communication tools. According to Branco and Encarnacao (2004), the human face is rich with communication information. This communication information, called facial expression, is always visible on human faces when they communicate with one and another. Sometimes, these facial expressions can also be used to give responses and feedbacks (Busso et al. 2004; Cohen, Grag, and Huang 2000; Ekman 1997). Researchers (Jain, Ross and Prabhakar 2004; Rose and Jain 2003; Gregory and Simon 2008) claim that facial expression is one type of biometrics and it can be used as information for the biometric system.

2.4.1 Biometrics

A biometric based pattern-recognition system recognizes a person based on the feature vector which derived from a specific physiological or behavioral characteristic of that human being (Prabhakar, Pankanti, and Jain 2003; Jain, Ross and Prabhakar 2004). It has also been defined as the process of matching an input biometric to the stored biometric information (Kumar et al. 2002). The system captures or gets a sample from the human body, for example, fingerprints, and does a comparison or a match with the template sample which is stored in the database then generates the outcome or authentication, which normally comes out as true or false. This method of comparison or matching the scanned sample with the template sample is the function of the biometric system.

The report in MIMOS Berhad (2008) shows that there are two general modes of comparing or matching biometric and they are the identification mode and the verification mode. The identification mode is used to identify someone from a large number of other users, like “who?” or “is this person allow to access?” The captured biometric information will be used to match with a large number of samples that are stored in the database. Then, the result, for example, the user details is displayed.

In the verification mode, the system checks to see whether the person is really who he or she claims to be by comparing or matching the captured biometric sample with a template sample stored in the database of the person involved. The authentication result will only come our as either “accepted” or “rejected”. This mode is widely used for commercial purposes such as in the banking system involving credit cards and ATMs, computer network logon, electronic devices and security purposes.

2.4.2 Types of Biometrics

Many types of biometric technologies are being used worldwide recently. These technologies include the most common ones like fingerprints, palm prints, faces, facial expressions, iris, voices, hand geometry, veins, signature and keystrokes (Rose and Jain 2003 ;Gregory and Simon 2008; MIMOS Berhad 2008).

Basically, biometric is divided into two categories; the behavioral and the physiological which is shown is Figure 2-4 (Jain, Ross and Prabhakar 2004).

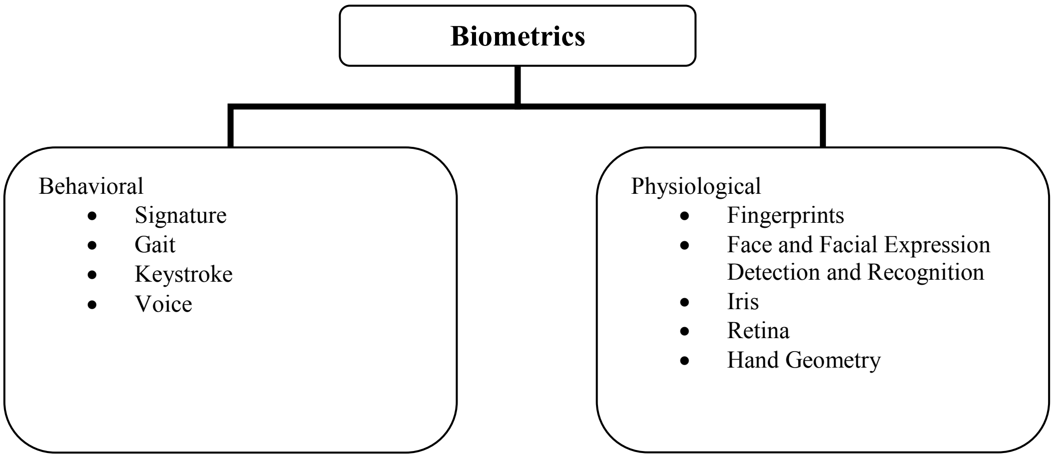


Figure 2-4 Types of biometrics information

Jain, Ross and Prabhakar (2004) further elaborate that a user only needs to provide his or her biometric sample to the system and the system will do the rest of the processes in the physiological biometric category. For example, the user only needs to place his or her finger on a fingerprint device and the system will begin the process of recognition and matching. Biometric that falls into this category includes fingerprint, face scan or detection, iris scan, retina scan as well as hand scan or hand geometry. MIMOS Berhad (2008) indicates that there are a number of biometric technologies in the developmental stage and they are facial thermography recognition, DNA recognition, odor recognition, nail-bed recognition, skin reflectivity as well as ear shape recognition.

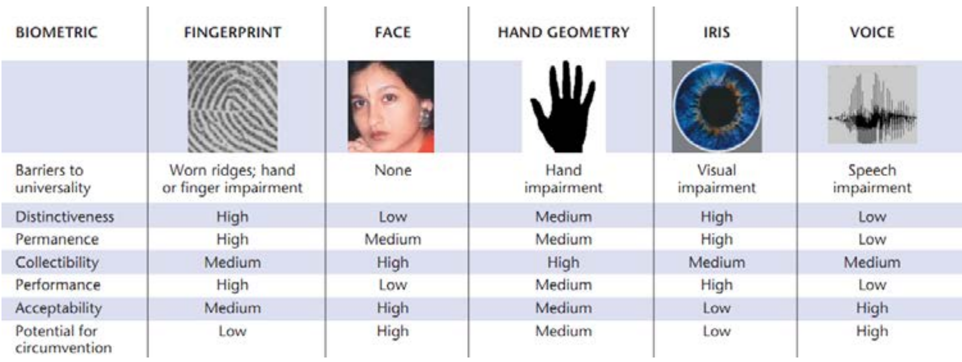
In the behavioral category of biometric, a user is required to do an action that he or she usually does to identify himself or herself. Take keystroke biometric recognition system as an example, the user is required to type a sentence using a keyboard for identification and the system will measure the typing speed, pattern and delay of the subject. Other biometric recognition such as signature, gait and voice are also placed in the category of behavioral biometric.

In the IT industry, biometric technology is one important area that contributes to its future development (Liu and Silverman 2001; Deriche 2008). Biometric provides security benefits across the spectrum, from IT vendors to end users, and from security system developers to security system users. Different types of biometrics have different usages. To authenticate a person, we can use the most common or simplest fingerprint technology to check and see whether the person is valid or not. To recognize a person's emotion, we can use facial detection and recognition. Multimodal of biometric recognition can be used for a higher level of authentication, especially for governmental, legal and commercial transaction such as for banking purpose.

Liu and Silverman (2001) add that some E-commerce developers are exploring the use of biometric and smart cards to more accurately verify a trading party's identity. Banks are bound to use this combination to better authenticate customers and ensure non repudiation of online banking, trading and purchasing transactions. Point-of-sales (POS) system vendors are working on the cardholder verification method, which would enlist smart cards and biometric to replace signature verification.

Biometric can help to obtain secure services over the telephone through voice authentication. The last interesting application is for covert surveillance. Using facial and body recognition technologies, researchers hope to use biometric to automatically identify known suspects entering buildings or traversing crowded security areas such as airports.

Each of the biometric technology has its own strengths and weaknesses (MIMOS Berhad 2008). The attributes of measurement on whether a biometric technology is good or not include the issue of universality (does everyone have it?), distinctiveness (how unique is the pattern?), permanence (how permanent are identifiers?), collectable (how well can the identifiers be captured?), performance (the speed and accuracy of matching), acceptability (willingness of the public to accept it), and circumvention (the foolproof) (Prabhakar, Pankanti and Jain 2003; MIMOS Berhad 2008). The comparison among the different types of biometric information is further illustrated in Figure 2-5.



BIOMETRIC	FINGERPRINT	FACE	HAND GEOMETRY	IRIS	VOICE
Barriers to universality	Worn ridges; hand or finger impairment	None	Hand impairment	Visual impairment	Speech impairment
Distinctiveness	High	Low	Medium	High	Low
Permanence	High	Medium	Medium	High	Low
Collectibility	Medium	High	High	Medium	Medium
Performance	High	Low	Medium	High	Low
Acceptability	Medium	High	Medium	Low	High
Potential for circumvention	Low	High	Medium	Low	High

Figure 2-5 Strengths and weaknesses for each type of biometrics information (Prabhakar, Pankanti and Jain 2003)

Through the comparison of the chart, (Prabhakar, Pankanti and Jain 2003), the face or facial expression has the potential to help children with Cerebral Palsy in communication. Facial expression technology is good in terms of universality as every human owns a unique face pattern from birth. Besides, the sample of the face in digital format can be collected easily by capturing it with a camera then storing it in database as image file. Face detection and recognition give a high level of security because of the high level of circumvention and its difficulty to be duplicated.

On the other hand, face detection and recognition could only achieve an average level of performance due to its accuracy level and the slow processing speed during the authentication and recognition process. Usually the image file size is larger compared with other biometric samples such as fingerprint, and hence it decreases the matching and processing speed. In addition, the features and shape of the human face may change from time to time and is not permanent, thus resulting in a lower level of accuracy rate in the matching for the long term unless the database is updated from time to time.

This technology is highly accepted by the public although the performance of this biometric is not as good as others. It requires no direct contact by the subject with the device used for sample collection. In general, face recognition and detection biometric technology has more strengths than weaknesses and it can be used for further development and enhancement in assisting children with Cerebral Palsy in their communication.

2.4.3 Facial Expression as a Medium of Communication

In real life, inter personal human interactions are performed not only by using speech or spoken language, but also in nonverbal cues for example hand gesture, body gesture, facial expression and tone of the voice. All these cues are sometimes being used to express feelings and give feedbacks (Busso et al. 2004; Cohen, Grag and Huang 2000) as human interact with each other every day. For instance, a child jumping up and down in front of his mother maybe expressing his feelings of unhappiness and dissatisfactions. On the contrary, this child might be happy and joyful because he sees his mother in front of him. A man shouting at someone and scolding him indicates that the man is in anger.

All the action or behavior that a person exhibits can represent a person's initial emotional state and hence it proves that inter-personal human interaction is not limited to the use of the spoken language, but to the body gesture as well and in this case, the facial expression.

Facial expressions are generated by contractions of facial muscles, which result in the deformation of facial features such as the eyelids, eyebrows, nose and lips, and also in the changes to their relative positions (Xie and Lam 2009). The face is indeed a rich, expressive communication member of the body that people have learned to interpret to the extent that they are not aware of doing it all the time (Branco and Encarnacao 2004).

To support the usefulness of facial expression as a medium of communication, Franco and Treves (2001) claim that facial expression interactions are relevant for community social life, teacher and student interaction, credibility in different contexts, medicine and so on. They add that, facial expression recognition is useful for designing new interactive devices which offer the possibility of new ways for human computer interaction (HCI).

Cohen, Grag and Huang (2000) conduct a survey on their users and notice that they have operated HCI in the traditional way which consist of only the using keyboard and mouse, interface and devices such as joysticks, trackballs, data gloves and touch screen monitors. However, this interaction can be improved and enhanced by introducing facial expression recognition that requires no direct contact by the user and the device or application, compared with the traditional interaction. On the other hand, Chen et al. (2008) adopted facial expression as part of their communication and interaction in health care and smart environment application.

Recent research shows that the adoption of facial expression recognition system in assisting people with communication disabilities is feasible, where it serves as a monitoring system to track the activities of the disabled (Lau, Tran, and Chiong 2009).

2.5 Facial Expression Recognition System

Facial expression recognition is the ability to recognize human emotion by their facial expression and to differentiate it one with another. A human being is born with the ability to recognize other homosapiens easily by identifying their facial characteristics such as face shape, appearance, skin textual and color.

According to Ekman and Friesen (1978), humans have the ability to express, interpret and recognize six main facial expressions which include happiness, anger, disgust, fear, surprise and sadness. The other facial expressions can be learned from the environment as well as from other humans (Datcu and Rothkrantz 2007). The six facial emotions stated are very important and they play a major role in the facial expression recognition methodology (Busso et al. 2004).

Facial Expression Recognition System (FERS) has been a hot topic in research ever since Ekman and Friesen (1978) pioneer this research and work from the psychological perspective. In the past two decades, researchers have tried to adopt their idea and to make improvement, innovation and modification on FERS by introducing different methods or techniques that can be found in the paper of Kotsia and Pitas (2007). These innovation and modification that are carried out mainly concentrated on the improvement of the system in terms of accuracy, efficiency, mobility, and speed. With all the improvement done by researchers, the development of the techniques for facial expression recognition has increased very quickly (Zhan and Zhou

2007). One of the most active researches in computer vision and pattern recognition is face recognition. This technique has been applied to identification, access control as well as human computer interaction (Wang, Plataniotis and Venetsanopoulos 2005).

As stated before, human being can interpret and generate six facial expressions but computers do not have the ability to do so. Hence, it is much easier for humans to communicate with each other than for the computer to communicate or interact with humans. In the real world, a human being can recognize the emotion of another human being via vision and speech which a computer cannot do. To improve interaction between human and computer, a more natural and intuitive way of human-computer interaction method is needed. This can be carried out by equipping the computer with the suitable software i.e. facial expression recognition system, to interpret human expression and trigger the appropriate responses.

The same concept of inter-human interaction can be applied to human-computer interaction in facial expression recognition. A computer will use a camera to “see” human expression and a microphone to “hear” the human voice, and then learn to recognize it (Cohen, Grag and Huang 2000).

During the recognition process, the human face is captured by the camera that is installed in a computer and the captured images will be saved and sent for processing, using a specific technique to recognize the face of a user based on the features achieved. With automated facial expression recognition system, human-computer interaction has improved as the subject does not need to coordinate with the computer device or perform any actions in order to get a task done since facial expression recognition falls into the physiological category of biometric (MIMOS Berhad 2008).

2.5.1 Processing in FERS

In order for a computer to recognize a person’s emotion, a few processes need to be performed. Figure 2-6 shows the processes required for a FERS. Generally, the stages involved in the process include face detection, feature extraction and lastly expression recognition (Zhang et al. 1998; Chibelushi and Bourel 2002; Zhan et al. 2008; Piatkowska 2010).

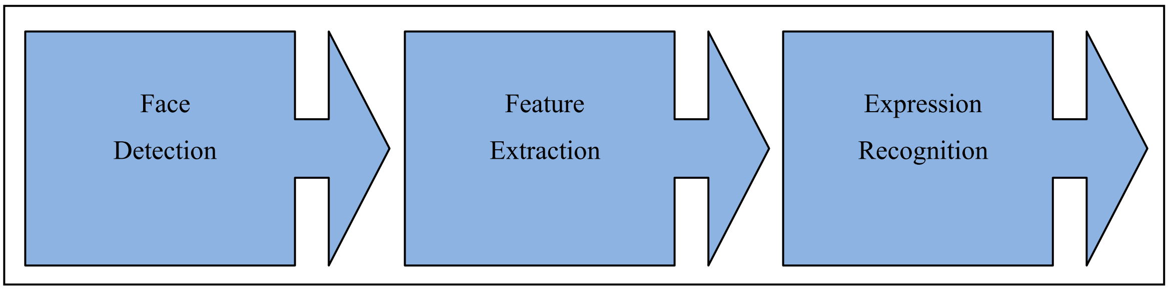


Figure 2-6 Processes involved in Facial Expression Recognition System (Zhang et al. 1998; Chibelushi and Bourel 2002; Zhan et al. 2008)

For the past few years, researchers have been introducing new and enhanced techniques to perform facial expression recognition after the invention of the Facial Action Coding System (FACS) a research done by Ekman and Friesen (1978). This system has since been improved by Donato et al. (1999). It is used for classifying facial action by combining the information from holistic spatial analysis, feature measurements as well as optical flow. Donato and colleagues claim that this model is widely accessible as a research tool as it would not only increase the speed of coding but also improve the reliability, precision, and temporal resolution of facial measurement.

A few years later, Franco and Treves (2001) propose their method of using unsupervised local processing of neural network for facial expression recognition as this technique shows that it is very effective in the task of reducing high dimensionality of the input data. This technique is constructed in two ways which are the hidden layer of neurons as well as the backpropagation neural network.

On the contrary, Cohen et al. (2003) state that FACS processes are very time consuming and have to be performed manually. Therefore, they apply the Bayesian Network Classifier to solve the problem and proposed the multi-level Hidden Markov Model (HMM) architecture as well as the Naive-Bayes Classifier for the Cauchy distribution assumption uses. They integrate the classifiers and the face tracking system to build a real-time facial expression recognition system which could speed up the process of facial expression recognition.

Barlett et al. (2003) propose real time face detection and facial expression recognition system by developing an application for human computer interaction using a face finder that trained is with Viola and Jones Face Detection algorithm. The Gabor representation is implemented to form the patch and finally SVM is used as a classifier in expression recognition. The performance of the recognition could be enhanced with a novel combination of Adaboost and SVM (Barlett et al. 2003).

Ma and Khorasani (2004) propose a facial expression system using the constructive feedforward neural network that uses two-dimensional discrete cosine transform (2D-DCT) over the entire face image as a feature detector while a constructive one-hidden-layer feedforward neural network is employed for the facial expression classifier.

In Lin's paper (2006), Yacoob and Davis use the optical flow to track the dynamic movement of facial features from video sequences and classify all the facial features representation into six expressions and boosted the recognition rate from 80 – 94%. Later, Zhan and Zhou (2007) introduce a model of a facial expression recognition system based on hybrid features and fusing discrete Hidden Markov Model. First, their model segments a face image in an image sequence and execute the gray scale normalization of circumrotation revision for the sub-face image. Then the Active Appearance Model (AMM) and the Gabor wavelet transformation are applied for feature extraction. Lastly the Hidden Markov Model is employed for recognizing emotion in their model as a complete set or model of facial expression recognition system. These studies show that the approaches used to develop facial expression recognition system are different among researchers in this field.

58 papers and articles ranging from year 2000 to 2010 related to techniques used, adopted or recommended for object detection, feature extraction and facial expression recognition system is reviewed. The techniques used by researchers are recorded and categorized according to three processing stages (face detection, feature extraction, expression recognition) which are shown in Table 2-1, Table 2-2, and Table 2-3.

Table 2-1 Techniques adopted by researchers for face detection stage

Techniques	Used / Adopted / Recommended by
Adaboost	Zhan and Zhou 2007
Deformable Template	Balasuriya and Kodikara 2000
Discrete Cosine Transform	Ma and Khorasani 2004
Dynamic Bayesian Networks	Zhang and Ji 2005
Facial Action Coding System	Wang and Lien 2009
Gabor Wavelet	Lyons et al. 2000
Local Binary Patterns	Feng, Pietikainen and Hadid 2005
Motion History Image	Davis 2001; Yilmaz, and Shah 2002; Zhou et al. 2003; Valstar, Pantic and Patras 2004; Harasse, Bonnaud and Desvignes 2005; Huang and Lin 2008
Neural Network	Fasel 2002; Matsugu et al. 2003
Piecewise Bezier Volume Deformation	Cohen et al. 2003; Sebe et al. 2007
Principal Component Analysis	Lyons et al. 2000; Franco and Treves 2001; Agrawal, Cosgriff and Mudur 2009; Niu and Qiu 2010
Skin Color Model	Yilmaz, and Shah 2002; Singh et al. 2003; Kovac, Peer and Solina 2003; Harasse, Bonnaud and Desvignes 2005; Huang and Lin 2008; Chai et al. 2009
Template Matching	Park and Park 2006
Viola - Jones Face Detection	Bartlett et al. 2003; Viola and Jones 2004; Datcu and Rothkrantz 2007; Littlewort, Bartlett and Lee 2007; Huang and Lin 2008; Wu et al. 2008; Zhan et al. 2008; Cho et al. 2009; Riaz et al. 2009; Zheng and Bhandarkar 2009; Piatkowska 2010

Table 2-2 Techniques adopted by researchers in feature extraction stage

Techniques	**Used / Adopted / Recommended by**
Active Appearance Model	Zhan and Zhou 2007; Datcu and Rothkrantz 2007; Tang and Deng 2007; Ashraf et al. 2009
Cadide Wireframe Model	Kotsia and Pitas 2007
Clustered Discriminate Analysis	Chen and Huang 2003
Dual-tree complex wavelet	Chen and Xie 2007
Gabor Filter	Zhan et al. 2008
Gabor Wavelet	Kepenekci 2001; Tian, Kanade and Cohn 2002; Zhan et al. 2004; Zhan and Zhou 2007
Geometry Features	Tian, Kanade and Cohn 2002
Local Binary Patterns	Shan, Gong and McOwan 2005; Piatkowska 2010
Optical Flow	Huang and Lin 2008; Lee, Chun and Park 2008; Riaz et al. 2009
Principal Component Analysis	Riaz et al. 2009

Table 2-3 Techniques adopted by researchers in expression recognition stage

Techniques	**Used / Adopted / Recommended by**
Dynamic Bayesian Networks	Cohen et al. 2003; Sebe et al. 2007
Facial Action Coding System	Zhang and Ji 2005; Littlewort, Bartlett and Lee 2007
Geometry Features	Xie and Lam 2009
Hidden Markov Model	Cohen, Grag, and Huang 2000; Cohen et al. 2003; Aleksic and Katsaggelos 2006; Zhan and Zhou 2007; Wang and Lien 2009
Linear Discriminate Analysis	Lyons et al. 2000; Dubuisson, Davoine and Masson 2002; Price and Gee 2005
Linear Programming	Feng, Pietikainen and Hadid 2005
Neural Network	Franco and Treves 2001; Tian, Kanade and Cohn 2002; Ma and Khorasani 2004; Valstar, Pantic and Patras 2004; Lin 2006
Principal Component Analysis	Balasuriya and Kodikara 2000; Moon, and Philips 2001; Dubuisson, Davoine and Masson 2002; Wang, Plataniotis and Venetsanopoulos 2005; Lin 2006
Support Vector Machine	Bartlett et al. 2003; Shan, Gong and McOwan 2005; Datcu and Rothkrantz 2007; Kotsia and Pitas 2007; Chen and Xie 2007; Sebe et al. 2007; Zhan et al. 2008; Huang and Lin 2008; Lee, Chun and Park 2008; Michel and Kaliouby 2008; Agrawal, Cosgriff and Mudur 2009; Ashraf et al. 2009; Niu and Qiu 2010; Piatkowska 2010
Template Matching	Balasuriya and Kodikara 2000; Mu et al. 2001; Zhan et al. 2004; Karungaru, Fukumi, and Akamatsu 2004; Shan, Gong and McOwan 2005; Ming et al. 2005; Park and Park 2006; Kato, Hirano, and Nakamura 2007; Chai et al. 2009; Yusoff et al. 2009; Xie and Lam 2009

Throughout the review, the common techniques used and recommended by researchers are identified. In the face detection stage, Viola-Jones's face detection algorithm is mostly used followed by Skin Color Model and Motion History Image. In feature extraction, Active Appearance Model is mostly used followed by Gabor Wavelet and Optical Flow. In expression recognition, the technique most used is Support Vector Machine followed by Template Matching and also Principal Component Analysis.

It is so obvious that a FERS can be developed in various ways by using different techniques as described in the tables above. Each technique has its unique characteristics and different computational algorithm. Due to time constraint and limited resources, only techniques recommended by most researchers are further reviewed. The summary of techniques mostly used by researchers based on the processing stages in the facial expression recognition system is shown in Table 2-4.

Table 2-4 Summary of mostly recommended techniques by researchers based on the processing stages in the facial expression recognition system

Stage	**Mostly Used / Recommended**
Face Detection	Viola-Jones Face Detection Algorithm
	Skin Color Model
	Motion History Image
Feature Extraction	Active Appearance Model
	Gabor Wavelet
	Optical Flow
Expression Recognition	Support Vector Machine
	Template Matching

2.5.2 Face Detection

The face detection stage encounters a very complex problem to solve due to high variability in faces (Wall 2005). This stage is very important as the features inside the model will try to detect and recognize a human face and to check whether the face is present in the captured sample or not. The captured samples of still images and video sequences in a real time face detection process is shown are Figure 2-7. Without this procedure, the rest of the process such as expression recognition cannot be preceded.

Figure 2-7 Face detection (Dan 2008; Fraunhofer Heinrich Hertz Institute 2009)

2.5.2.1 Viola- Jones Face Detection Algorithm

One of the most common methods used for face detection is the one done by Viola and Jones (2004). This technique was widely used by other researchers throughout literature reviews (Bartlett et al. 2003; Datcu and Rothkrantz 2007; Littlewort, Bartlett and Lee 2007; Huang and Lin 2008; Wu et al. 2008; Zhan et al. 2008; Cho et al. 2009; Riaz et al. 2009; Zheng and Bhandarkar 2009; Piatkowska 2010) as it achieves the best compromise between detection efficiency and speed (Zhan et al. 2008). This technique involves three parts (Viola and Jones 2004).

1. *Integral image* associated with Haar-like features which can be seen in Figure 2-8,
2. *Adaboost* and also
3. *cascade* classifier

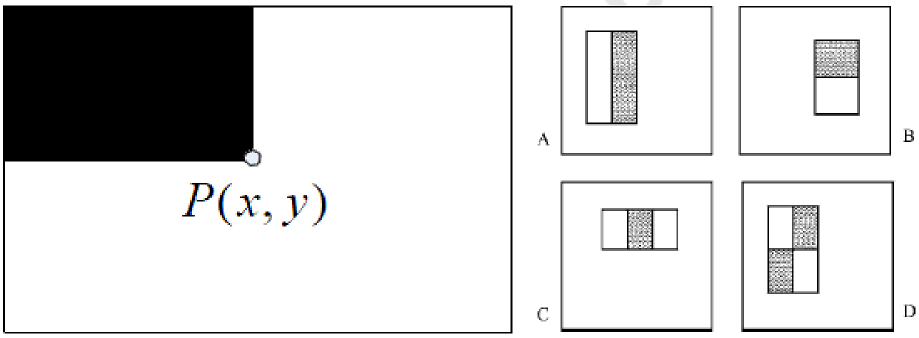


Figure 2-8 Detection window (Viola and Jones, 2004)

According to Viola and Jones (2004), the advantage of using this technique is that it simplifies the process of object detection. They claim that an integral image can be computed from an image using a few operations per pixel. Once computed, any of these Haar-like features can be computed at any scale or location in constant time.

Wu et al. (2008) produces significantly improved results in their work by combining the original Viola-Jones features into CART trees. Cho et al. (2009) claimed that rapid and robust face detection could be achieved by employing the Viola-Jones algorithm which is a variant of the AdaBoost algorithm.

Besides that, the Viola-Jones algorithm has lower false positive detection rates equivalent to the best published results (Bartlett et al. 2003) and it has been used in facial expression recognition or pedestrian detection (Viola, Jones and Snow 2003; Datcu and Rothkrantz 2007). The performance of the Viola-Jones face detection algorithm has shown to be very robust against illumination conditions and scale invariants. However, it is poor in performance against in-plan rotation of more than 15 degrees as well as out-plans rotation (Huang and Lin 2008). However, this is sufficient to perform face detection.

2.5.2.2 Motion History Image

In an earlier research on human motion tracking and recognition, Motion-History Images (MHI) is introduced to capture motion information in images (Davis 2001). This algorithm encodes, respectively, where a motion has occurred, and the history of motion occurrences in the image.

According to Huang and Lin (2008), MHI techniques are suitable to be used to perform human face detection, as this technique is fit for head pose and good for foreground segmentation by detecting moving objects in videos from stationary cameras. Davis (2001) claim that this technique is better than other techniques available as it remains a low computation consumption algorithm.

Valstar, Pantic and Patras (2004) have adopted and enhanced HMI in their work to Multilevel Motion History Image (MMHI). Instead of recording the motion history once for the entire image sequence and constructing a single MHI, they record the motion history at multiple time intervals and construct a Multilevel Motion History Image.

2.5.2.3 Skin Color Model

Many researchers (Yilmaz, and Shah 2002; Singh et al. 2003; Kovac, Peer and Solina 2003; Harasse, Bonnaud and Desvignes 2005; Huang and Lin 2008; Chai et al. 2009) employ the Skin Color Model for face detection and tracking in their work because color processing is much faster than other facial feature processing, as claimed by Singh et al. (2003) as well as Huang and Lin (2008).

By using the Skin Color Model, the skin regions can be separated from the rest of the image through a threshold process in both RGB and HSV color space (Huang and Lin 2008). This means that when performing face detection, the algorithm will detect and trace the same color space that has been defined and would proceed to mark the region as human face.

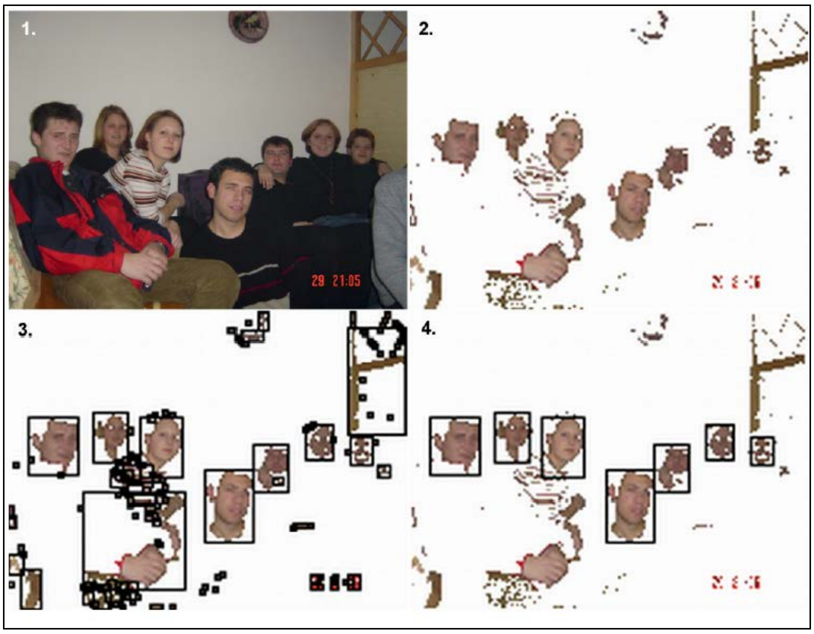


Figure 2-9 Misdetection using Skin Color Model (Kovac, Peer and Solina 2003)

Somehow, the Skin Color Model face detection technique may make misdetection on other parts of an image which is shown in Figure 2-9. This is because the misdetection part may have a similar color tone or color code to the skin color scheme that is available (Kovac, Peer and Solina 2003).

Besides, this technique is sensitive to light condition (Huang and Lin 2008) and it may affect or lower the accuracy of detection rate when this problem occurs. As a result, Skin Color Model can be used for other purposes but not for face detection because of all the limitations and problems stated above.

2.5.3 Facial Feature Extraction

Feature extraction is a special form of dimensionality reduction process (Zhu 2001) and is used to break a large input data into smaller portions based on different features by using techniques such as Independent Component Analysis (ICA), Linear Discriminate Analysis (LDA) Principal Component Analysis (PCA) or eigenface, and so on (Wang, Plataniotis and Venetsanopoulos 2005).

Feature extraction deals with the problem of matching features extracted from frame to frame in long sequences of images or video (Huang and Lin 2008). After the extraction process, each of the portions is processed separately in the next processing steps to reduce the complexity of a data as well as the computation time consumption.

In image processing and pattern recognition, feature extraction has become another important issue in the automatic facial expression recognition system after the face detection stage (Lin and Wu 1999). Basically, an original input image is projected to a lower dimensional features or sub image when the feature extraction process is conducted. For instance, when a human face image is captured and inserted, the technique will detect the basic features of the human face such as eyes, eyebrows, nose, mouth as well as cheek and extract them into small pieces of images (Gu, Su, and Du 2003). The extracted images or data will be used as resources for the expression recognition stages.

Figure 2-10 Global and local feature extraction (Whitehill 2006)

In general, feature extraction can be performed in two approaches: Global feature extraction (Holistically) and local feature extraction and both approaches are shown in Figure 2-10 (Fasel and Luettin 2003; Whitehill 2006).

Global feature extraction is shown in the first image which the chosen technique will detect the whole face by indicating it with a square box and extracting it. For local feature extraction, each of the features in the human face such as the mouth, eyes and eyebrows will be detected and extracted separately as shown in the second to the fourth image in Figure 2-10 above.

Besides the global and local feature extraction approaches, all the techniques can be classified in another four categories which include Knowledge-based Method, Feature-based Method, Template-based Method and Appearance-based Method (Jemaa and Khanfir 2009).

2.5.3.1 Active Appearance Model

Active Appearance Model (AAM) is one of the techniques used to perform feature extraction in a facial expression recognition system. Zhan and Zhou (2007) define AAM to be a technique used to build objects' shape and appearance models through a statistical study for objects' shape and texture. They add that AAM can obtain better detection result through training the shape and texture information of the model. This technique has been adopted by Datcu and Rothkrantz (2007) as well as Tang and Deng (2007) in their model to extract facial shape variety features.

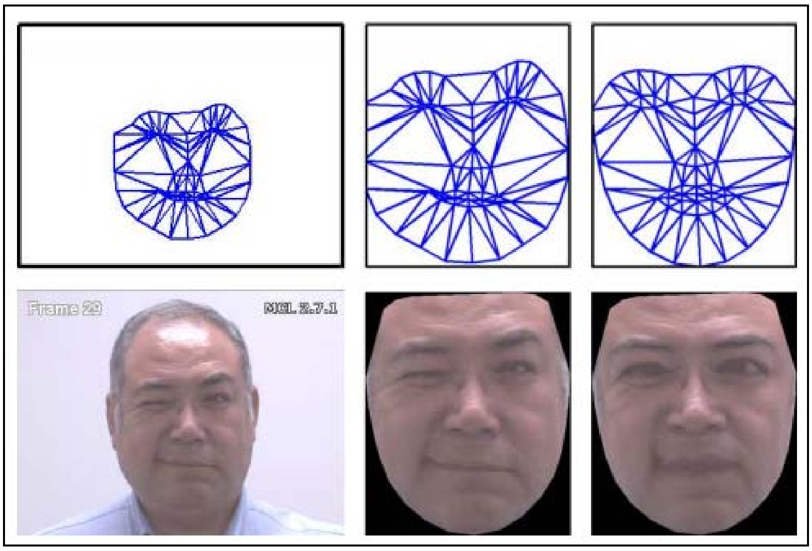


Figure 2-11 AAM derived representation (Ashraf et al. 2009)

Although AAM is capable of overcoming to some degree, the occurrence of occlusion, new samples containing partially occluded faces are introduced to increase the performance of the model fitting by plug-ins Facial Characteristic Point (FCP) technique which is shown in Figure 2-11, before performing AAM process (Datcu and Rothkrantz 2007).

2.5.3.2 Gabor Wavelet

The use of the 2D Gabor wavelet / filter in computer vision is introduced by Daugman (1988) in the late 1980s. According to Montillo, Metaxas and Axe (2004), Gabor wavelet was a complex sinusoid modulated by a 2D Gaussian function and it is an important tool for signal processing in space and frequency fields (Zhan et al. 2004). The 2D Gabor wavelet transformation has wide applications in such research areas as image processing and pattern recognition (Zhan and Zhou 2007). With properly chosen filter parameters, the Gabor filters have similar characteristics to multi-resolution techniques, such as the Gabor expansion and wavelets (Kyrki, Kamarainen, and Kälviäinen 2004). A simple example of facial image response to Gabor filters is illustrated in Figure 2-12.

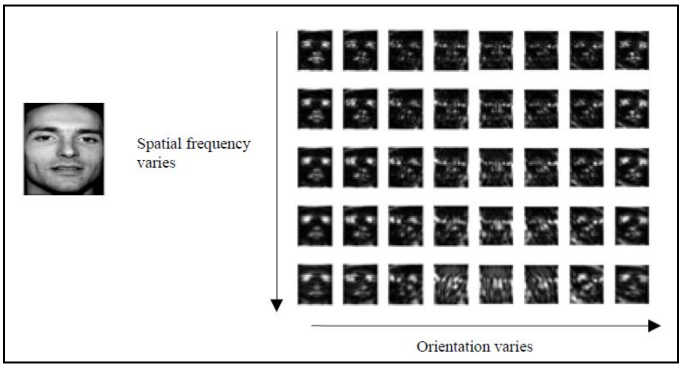


Figure 2-12 Facial image response to Gabor Filters (Kepenekci 2001)

In the pattern recognition and biometrics system, Gabor Wavelet has been used in feature extraction from image. Figure 2-13 shows the flow of the feature extraction process for facial images (Kepenekci 2001).

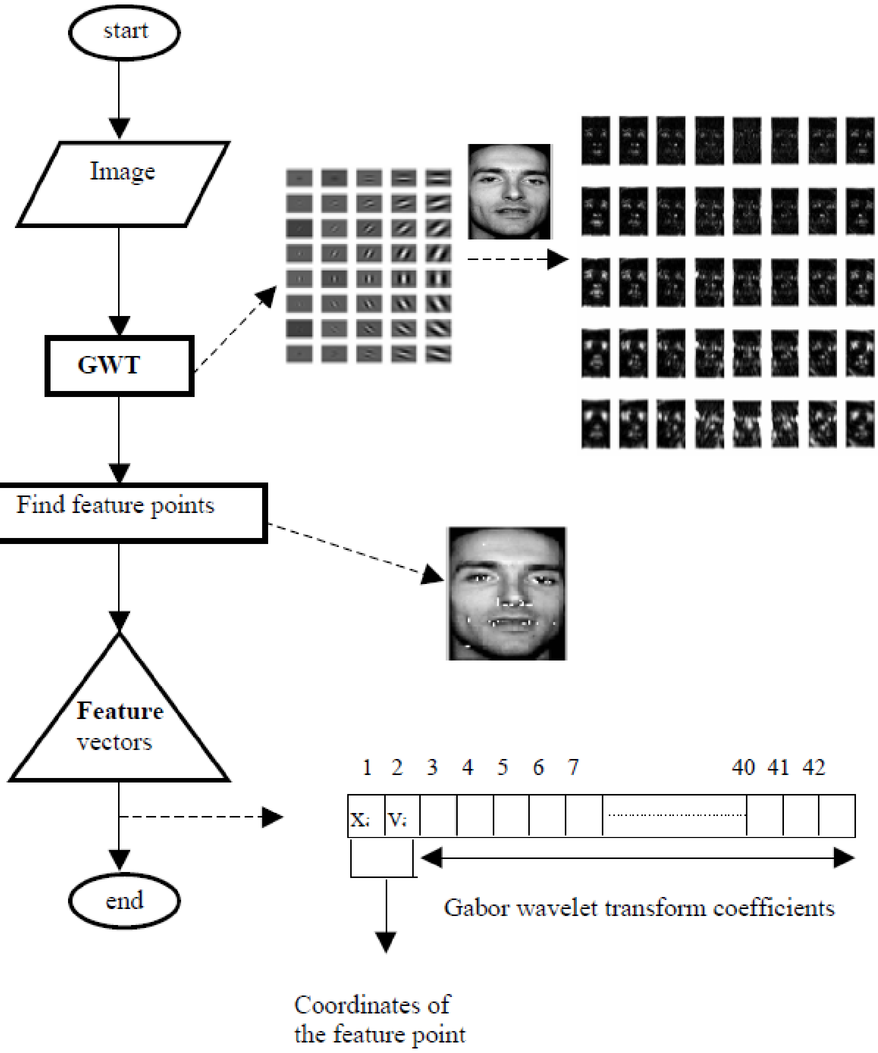


Figure 2-13 Gabor Wavelet / Filter for feature extraction in a face recognition system (Kepenekci 2001)

Researches (Kepenekci 2001; Tian, Kanade and Cohn 2002; Bartlett et al. 2003; Zhan et al. 2004; Zhan and Zhou 2007) included the usage of Gabor because Gabor filters are discriminative for facial expressions and robust against various types of noise than other methods in facial feature extraction (Zhan et al. 2008). Tian, Kanade and Cohn (2002) claimed that Gabor wavelets work well for single Action Unit (AU) recognition for the homogeneous subjects without head motion in their previous work, but surprisingly this approach returned poor recognition result for AU combination when image sequences included some non-homogeneous subjects with small head motions in their current work.

2.5.3.3 Optical Flow

In computer vision, optical flow is a velocity field associated with image changes. This effect generally appears due to the relative movement between object and camera or by moving the light sources that illuminate the scene (Aires, Santana, and Medeiros 2008). Apart from that, this algorithm can be used for a model to detect the human face in an image and to extract shape and texture information which is shown in Figure 2-14.

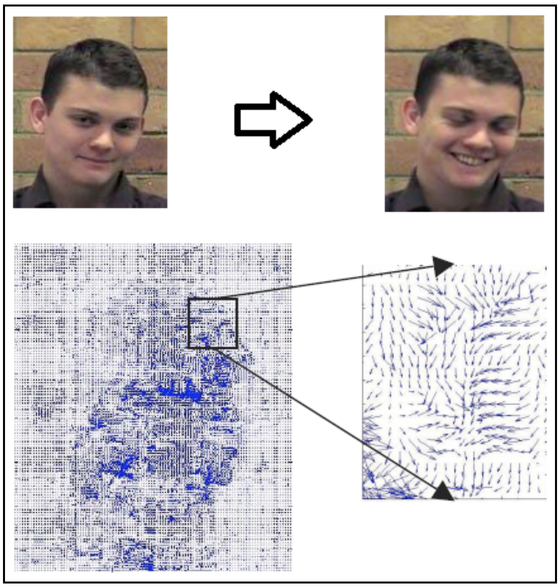


Figure 2-14 Optical flow of the transformation in human facial expression (Duthoit et al. 2008)

Generally, three steps are involved in computing optical flow (Huang and Lin 2008). 1) Detecting facial feature points in the first image according to the specific face model. 2) Matching the corresponding points in another image according to the pyramidal L. K.'s optical flow. 3) Calculate the displacement vectors of all these feature points

In a previous research, Donato et al. (1999) have adopted and used this technique in classifying facial action while Huang and Lin (2008), and Lee, Chun and Park (2008) adopted optical flow in their FERS. In the works of Riaz et al. (2009), they compute temporal parameters using optical flow to consider local feature variations after parameterized these extracted information from the image sequences using the Active Appearance Model (AAM) approach. Then, they combine these parameters to form a feature vector for all the images in their database.

2.5.4 Expression Recognition

Expression recognition or classification is the last stage in any facial expression recognition system. Classification is one of the most frequently encountered decision making tasks of human activity (Zhang 2000). A classification problem occurs when an object needs to be assigned into a predefined group or class based on a number of observed attributes related to that object. The purpose of expression recognition is to identify expression or emotion that is expressed by a person through the features such as eyebrow, eyes and mouth that have been extracted in the previous stage.

The features used for the expression recognition system are typically based on local spatial position or displacement of specific points and regions of the face (Busso et al. 2004), for example, the shape of the mouth, size of the eyes and width of the eyebrows.

Figure 2-15 Standard facial expressions template in JAFFE.

Figure 2-16 Yale Face Database (Franco and Treves 2001)

Usually a FERS needs to be trained by using a series of samples before the system can recognize human expression. The samples of the human facial expression are available in a standard template, for instance, Cohn-kanade database (Kanade, Cohn and Tian 2000; Datcu and Rothkrantz 2007; Pantic 2006; Zhan and Zhou 2007), Japanese Female Facial Expression database– JAFFE (Pantic 2006; Lin 2006) and Yale Face database (Franco and Treves 2001) shown in Figure 2-15 and Figure 2-16.

2.5.4.1 Support Vector Machine

Support Vector Machine (SVM) is one of the techniques used in expression recognition. This technique is also a set of related supervised learning methods used for classification and it performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories (DTREG 2005).

The inner computational algorithm for SVM is complex; however it can be used as “black box” tools (Huang and Lin 2008) for data classification. For instance, some sets of classified data is given or trained beforehand. When a new data is inserted, SVM will be able to predict which data set the new data should belong to. Figure 2-17 shows how data set can be classified by using SVM techniques.

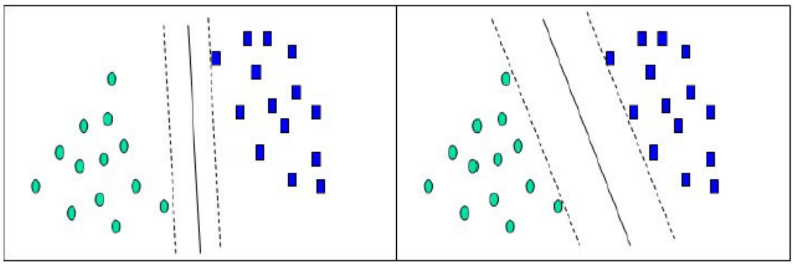


Figure 2-17 Simple classification using Support Vector Machine (SVM) (DTREG, 2005)

Bartlett et al. (2003) use SVM because it is suitable for this task as the high dimensionality of the Gabor representation does not affect training time for kernel classifiers. Besides, this technique is simple and fast in the process of classification (Taylor and Cristianini 2000), giving the best performance (Littlewort, Bartlett and Lee 2007) and it has been proven useful in many pattern recognition tasks including face and facial recognition (Ashraf et al. 2009).

In other words, SVM can perform better in expression recognition in terms of simplicity, accuracy as well as the processing. It is also robust in classification accuracy and is not limited even when only a modest amount of training data is available (Michel and Kaliouby 2008).

These features make SVM particularly suitable for a dynamic, interactive approach to expression recognition. However, Wu et al. (2008) claim classifiers such as SVM used in early work on FERS, i.e. face detection are complex and computationally expensive.

2.5.4.2 Template Matching

In computer vision, the Template Matching technique is widely used for object detection, classifying objects, as well as comparing portions of images against one another i.e. comparison between input image and template image (Brunelli 2009). A sample image (template image) may be used to recognize similar objects appearing in source image (input image).

Apart from that, Latecki (2005) claim that templates are most often used with characters, numbers, and other small or simple objects like the face. In other words, Template Matching involving the process of matching between two or more images at the same time generates matching results.

Usually, Template Matching is performed by comparing template image with input image using genetic algorithm on several locations around the input image and then generate matching result in percentage or similarity (Karungaru, Fukumi and Akamatsu 2004).

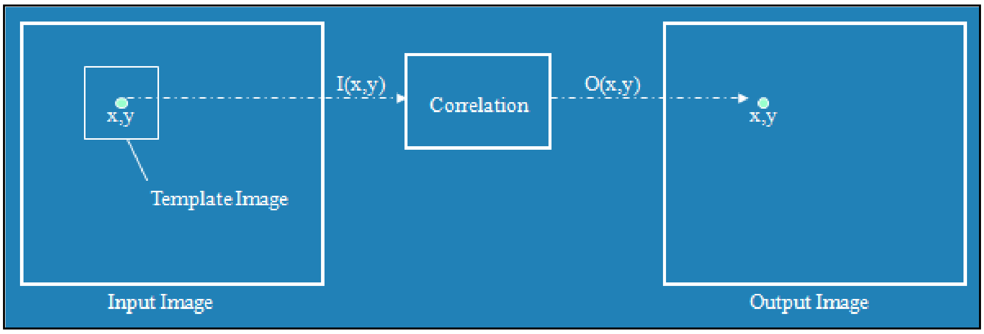


Figure 2-18 Template Matching techniques used for object detection (Latecki, 2005)

Figure 2-18 illustrates how Template Matching works. When performing Template Matching, a template image is shifted to all possible positions in a larger source image or input image and it then compute the percentage of similarity of each position (Latecki 2005). The matching process in each location is computed on a pixel-by-pixel basis. The location with the highest percentage of similarity will be considered a match with the template.

As claimed by Brunelli (2009), Template Matching has been widely used by many researchers in the computer vision field and also in the development in FERS (Balasuriya and Kodikara 2000; Mu et al. 2001; Zhan et al. 2004; Karungaru, Fukumi and Akamatsu 2004; Shan, Gong and McOwan 2005; Ming et al. 2005; Park and Park 2006; Kato, Hirano, and Nakamura 2007; Chai et al. 2009; Yusoff et al. 2009; Xie and Lam 2009) because Template Matching is simple (Shan, Gong and McOwan 2005).

Ming et al (2005) further elaborate that the Template Matching technique manages to solve the problems of time consumption and computational complexity experienced by other techniques i.e. the Support Vector Machine (Wu et al. 2008). In addition to that it can produce a high recognition rate (Xie and Lam 2009) and can manage to handle a large database.

Shubhangi, Preeti and Ajith (2010) compare it with several techniques in their paper and find that Template Matching is better than the Support Vector Machine (SVM) and Hidden Markov Model (HMM).However it is less efficient compared with Neural Network as well as Optical Flow to a certain extent.

2.6 Critical Review

The key issues on literature findings are discussed in this section. These issues include Cerebral Palsy and their difficulties, the limitation on existing Augmentative and Alternative Communication tools in overcoming Cerebral Palsy's difficulties, the usage of biometrics information as an enhancement to existing Augmentative and Alternative Communication tools, and also facial expression recognition system (FERS).

2.6.1 CP and their Difficulties

Generally, children with Cerebral Palsy are individuals who experience disability such as loss of voluntary muscular control and coordination. These disabilities are caused by brain damage before, during or after birth in the first year of their lives, brain injury, or birth injury resulted from drug and alcohol consumption, maternal infection, neonatal infection as well as premature birth (Badawi et al. 1998; Columbia Electronic Encyclopedia 2008; Krigger 2006 and 4MyChild 2009).

A majority of these children are diagnosed with severe physical disabilities making their condition worse. This includes difficulty in performing their daily tasks (4MyChild 2009). As a result, children with Cerebral Palsy have difficulties in communicating with others as they have problems controlling their mouth, hands and other parts of their body muscle (Krigger 2006). Therefore, these children need an assistive communication tool to help them to communicate.

2.6.2 Limitations of AAC to overcome CP's difficulties

Suggestions are found on how children with Cerebral Palsy can be helped by using Augmentative and Alternative Communication (AAC) tools (Davis, Moore, and Storey 2003; Pennington, Goldbart, and Marshall 2004; Chapman 2009; Rummel-Hudson 2011).

Many AAC tools are developed to fulfill the needs of children with Cerebral Palsy which are listed in websites such as Michigan's Assistive Technology Resource (2002), AbleData (2010) and some are suggested by the American Speech-Language-Hearing Association (ASHA 2010) as well as the International Society for Augmentative and Alternative Communication (ISAAC 2010). However, these tools are less efficient because most of them (NaturalSoft 2010; Softpedia 2009; Standup 2006; Crestwood 2009; Attainment 2010; Cooper 2010; Dynavox 2010) are text-to-speech and touch screen based which require good physical coordination.

These tools cannot be used in real life situations due to the lack of adaptability to different users and the patients' communication disabilities (Gonzales, Leroy, and Leo 2010). The limitations of these AAC tools are identified after a comparison is made among these tools (Gonzales, Leroy, and Leo 2010; Bouck 2010). The outcome of this comparison inspires this research to go towards biometric information as an alternative technology which is potentially enhancing the existing AAC tools (Branco and Encarnacao 2004; Busso et al. 2004; Cohen, Grag, and Huang 2000; Ekman 1997).

2.6.3 Biometrics Information as an enhancement for AAC

There are many types of biometrics information found in the biometrics industry (Prabhakar, Pankanti, and Jain 2003; Jain, Ross and Prabhakar 2004; Rose and Jain 2003; Gregory and Simon 2008; MIMOS Berhad 2008). According to Cohen, Grag and Huang (2000), facial expression is one of the best methods to adopt for Human Computer Interaction (HCI) purposes. This statement is then supported by Franco and Treves (2001) and Chen et al. (2008). Facial expression is suitable as an enhancement for AAC due to several factors:

1. **Information**–The human face is rich with communication information (Branco and Encarnacao 2004). On top of that, humans can communicate by expressing their initial mood or emotion via facial expressions. Therefore, facial expressions can be used to give responses and feedbacks as well (Busso et al. 2004; Cohen, Grag, and Huang 2000; Ekman 1997).
2. **Convenience** - Facial expression recognition is categorized as the physiological type in biometrics information. The subject only needs to provide his/her biometric sample to the system and the system will do the rest of the processes. Hence there is no other additional instruction required for the user or children with Cerebral Palsy to perform when using the system (Jain, Ross and Prabhakar 2004).
3. **Universal** - Facial expression technology is good in terms of universality as every human owns a unique face pattern from birth (Jain, Ross and Prabhakar 2004)
4. **Collectable** –A sample of the face in digital format can be collected easily by using a camera and then captured and stored in database as image files (Jain, Ross and Prabhakar 2004).

2.6.4 Adoption of FERS

Generally, a basic facial expression recognition system involves three stages which include face detection, feature extraction and lastly expression recognition (Zhang et al. 1998; Chibelushi and Bourel 2002; Zhan et al. 2008). The techniques used in each stage are different and all available techniques used for each stage is listed in Table 2-1, Table 2-2, and Table 2-3 on section 2.5.1.

2.6.4.1 Face Detection

A comparison has been made in a previous research by Lau, Tran, and Chiong (2009) on three techniques used for face detection (Viola-Jones Face Detection Algorithm, Motion History

Image, and Skin Color Model). From their findings, the best technique used for face detection related application is combining all three techniques.

However, when comparing each and every individual technique, it is discovered that the Viola-Jones face detection techniques is better than the other two techniques used. The results of the comparison show that these three techniques able to detect the human face. However the disadvantages among these techniques are varied.

For Motion History Image, it is normal that any moving object or object with motion will be detected. This could be a threat for any face detection application as there are higher chances for false positive detection. Anything that passes through the device will be treated as face and this may not be what a face detection application developer wants.

For Skin Color Model, anything that has similarity with the skin color will be detected (Kovac, Peer and Solina 2003) and this might lead to a higher false positive detection. Apart from that, this technique will detect side views of a human face. However, human expression is indistinguishable from this angle of view. So, it is redundant that the side of a human face is detected instead of the frontal part of the face.

For Viola and Jones Face Detection Algorithm, the disadvantages of this technique are sensitivity to the environmental lighting condition and its return of poor detection on rotated objects (Huang and Lin 2008). However, even though this technique produces poor detection for rotated objects, it will not affect expression recognition results.

Consequently, this research adopts the Viola and Jones face detection algorithm for the face detection part in the proposed prototype, after comparing these three techniques. IT is considered to be the most suitable technique for the purpose of this research.

2.6.4.2 Feature Extraction

From the review, Viola-Jones face detection algorithm is found to be sufficient for face detection and extraction. It extracts the entire detected human face found in an image as a global feature, without any further extraction into smaller features to avoid unnecessary memory and time consumption. Thus, this research adopted the global feature extraction which provides sufficient details for the proposed prototype.

2.6.4.3 Expression Recognition

The previous review has found several arguments and claims raised by researchers towards the use of the Support Vector Machine (SVM) and the Template Matching technique for expression recognition. Ashraf et al. (2009) claimed that SVM is highly useful in many pattern recognition tasks. This is due to the reason that SVM is simple and fast in the process of classification (Taylor and Cristianini 2000) and robust in classification accuracy (Michel and Kaliouby 2008).

On the other hand, Wu et al. (2008) oppose the effectiveness of SVM because this technique is complex and computationally expensive. According to Ming et al (2005), this problem can be solved by using the Template Matching technique because it is simple (Shan, Gong and McOwan 2005). Apart from that, Template Matching is widely used by many researchers in the computer vision field (Brunelli 2009) and is capable of producing a high recognition rate when dealing with a large database (Xie and Lam 2009).

To strengthen the effectiveness on Template Matching over SVM, an experiment was conducted in this research. This experiment was done by comparing SVM and Template Matching in terms of memory and storage consumption, processing speed as well as recognition accuracy which is shown in Table 2-5.

Table 2-5 Comparison between SVM and Template Matching in expression recognition

Attributes	**SVM**	**Template Matching**
Memory and storage consumption	60 MB	0.4MB
Load train file(s)	>= 5 seconds (depending on data size)	< 100 milliseconds
Recognition accuracy	82%	75%

For memory and storage consumption wise, a trained file with the size of 60 Megabytes was created by using the SVM technique. This trained file is the representation of binary value for every pixel in each of the sample image or train sets by training 7 expressions. However, when using the Template Matching technique, only 0.4 Megabytes of trained files were created for the same number of expressions. This shows the reduction of storage and memory consumption when Template Matching is used

When performing expression recognition, SVM loaded the same trained file and performed a comparison with a real time input image. This process took more than 5 seconds to load the trained file and start the recognition process. This could be a burden for a mobile machine to create and to load a file of such magnitude. Whereas for Template Matching, the trained file were loaded in less than 1 second upon request.

In terms of recognition accuracy, SVM (82%) performed slightly better compared with Template Matching (75%). However, this experiment has shown that when performing expression training using SVM, it requires high memory consumption and processing speed. Furthermore, it could slow down the processing speed of a mobile machine like the laptop.

Therefore, the literature reviews and experiments suggested that the Template Matching technique is more suitable for expression recognition as it consumes less memory, storage and processing time.

2.7 Conclusion

The disabilities of children with Cerebral Palsy were identified in the beginning of this chapter. They experience intellectual, speech and physical disabilities which cause them to have communication difficulty. Speech therapists and pathologies suggest the use of Augmentative and Alternative Communication (AAC) tools to overcome their difficulties in communicating with each other.

Research has found that the existing AAC tools suggested by speech therapists and pathologies are less efficient. Most of these tools are text-to-speech and touch screen based which requires good physical coordination. Furthermore, existing AAC tools have limited adaptability to different users and communication situations.

Previous researches suggest that biometrics information can be used to assist children with Cerebral Palsy to overcome communication difficulty. Researchers claim that the human face is rich with information for giving feedbacks and responses. Therefore, facial expression recognition technology is suitable to be used as a medium of communication for children with Cerebral Palsy.

Through comparison, Viola-Jones face detection technique is suitable in performing the face detection tasks as it is one of the most widely used algorithms in face detection tasks. For expression recognition, the comparisons of previous researches suggest that Template Matching is better when compared with other techniques.

In short, with the findings and outcomes from literature review, it is feasible to improve the communicational technique of children with Cerebral Palsy by adopting the biometric information that is facial expression and real time template matching technique.

Chapter 3: Conceptual Modeling

This chapter discusses an overview of the concept, requirements and design for the proposed prototype. The system development model used in this research is described in detail and the requirements are well defined for each stage in the development process.

Findings from site visits are discussed in association with literature review's findings. These findings have inspired this research to propose a solution to solve the research problem. The proposed solution is further discussed in detail and the requirement specification i.e. scope of system, functional and non-functional, constraints, other requirements and tools are defined based on findings obtained and needs for the proposed solution.

System architecture was designed according to the requirement specification followed by the use cases model. These use cases explained how the system worked and each cases in the system was further illustrated using activity diagrams. A summarization for the whole conceptual and design was elaborated at the end of this chapter.

3.1 System Development Model

The System Development Life Cycle (SDLC) model was adopted in this research for the prototype development. SDLC model is the simplest model to follow and safer compared with other software development model (Tayntor 2007). The requirements for every stage in this model are well defined before proceeding to the next stage and hence, the risk of errors and miscommunication are minimized (Lewis 2008).

Figure 3-1shows the basic SDLC model which involved five stages; these stages include requirement definition, system and software design, implementation and unit testing, integration and system testing as well as operation and maintenance (Sommerville 2011).

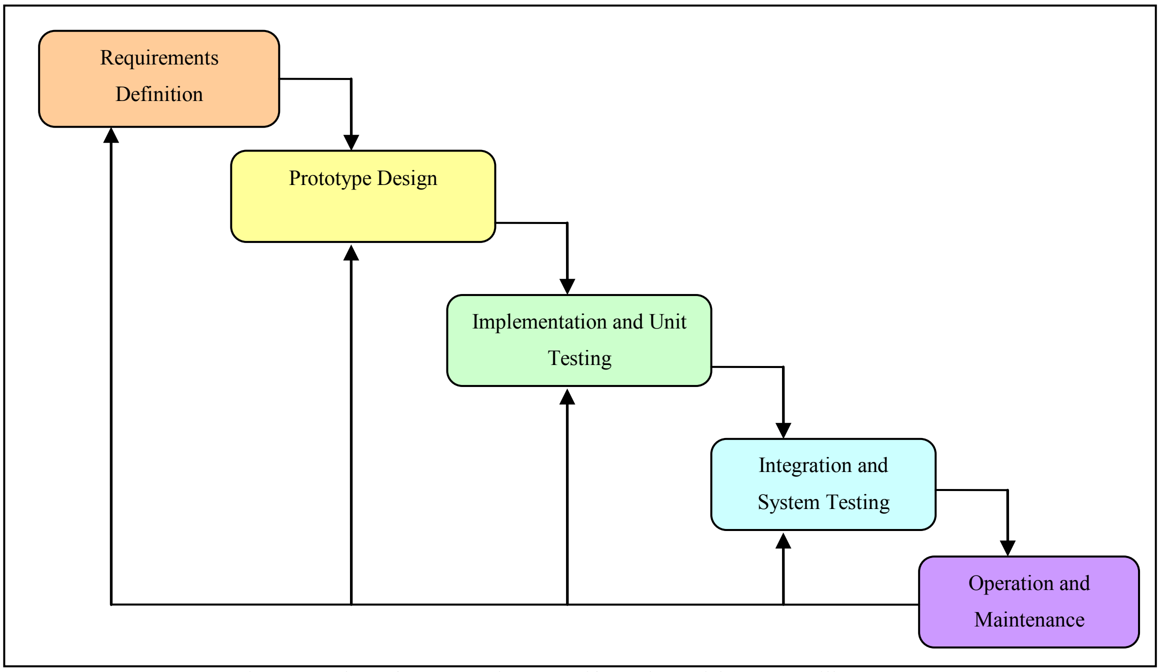


Figure 3-1 The Software Development Life Cycle Model (Sommerville 2011)

The domain of the research was defined and narrowed in the planning stage. This research developed an assistive communication tools to help children with Cerebral Palsy in their communication. The research timeframe, domain, objective, and scope were defined in this stage as well.

In the requirement definition or analysis stage, the research problem, definition, research goal and proposed solution were identified. The sources of requirement definition were collected from literature reviews on previous researches, visiting and collaboration with special school for mentally retarded children, and from online resources.

The special school was visited and staff or teachers in the school were invited to participate in this research by contributing more knowledge regarding the children's behavior, disabilities and their lifestyle.

The requirement and research goal were validated throughout the requirement definition or analysis stage based on information obtained from site visits and literature reviews. A solution was proposed within this stage to help children with Cerebral Palsy to fulfill their needs.

Every information and knowledge obtained in the requirement definition or analysis stage was forwarded to the System and Software Design stage for establishing the overall prototype plan. Each required hardware and software was organized and connected using software development diagrams and a description was added to avoid error.

Due to the many techniques involved, the prototype was divided into smaller modules and components. The proposed prototype was formed by combining each module or component. Basically, the user fed the prototype with input to the first module, and the output of the first module would be passed to the second module as input and the same applied to smaller components. By designing the prototype development using the "divide and conquer" concept (Gull et al. 2009), the debugging worked and the identifying of potential error occurring in the implementation stage later was much easier.

Prototype implementation starts once the system and software design is verified. The requirement specification is converted into executable prototype using C++ programming language with additional developmental tools i.e. OpenCV, a special API for image processing, and also .NET Frameworks for the Graphic User Interface (GUI).

All required experiments and tests were conducted within the development process to ensure that each module and components was verified and met with the required specifications. A few rollbacks to the previous stage for modification and improvement were required when a component did not meet the requirement stated before.

Preliminary and field testing were conducted in the integration and system testing stage. These tests were conducted to examine whether the proposed prototype had managed to solve the research problem.

Normal humans and a few children with Cerebral Palsy were invited to participate in the preliminary test which covered the prototype's functionality tests while field testing was conducted with a group of children with Cerebral Palsy in the special school. Some tests and experiments during prototype development were conducted in the lab because these test dealt with children with Cerebral Palsy who had difficulty in understanding or following the instructions given. Therefore, a series of experimental procedure were defined and followed along these two tests.

3.2 Site Visits and Collaboration

In addition to the reviews on children with Cerebral Palsy in the literature reviews, this research conducted multiple visits to the special school to enhance the understanding of the subject matter. It also helped in building social network and collaboration with specialists at the school. The collaboration with the specialists i.e. teachers, administrative staff, principal, and physio-therapists provided the researcher with a better understanding of the children. They have a lot of experience in handling and dealing with the children. Multiple informal interviews were conducted during the collaboration while notes taking were done in the conversation with these specialists upon visits.

Apart from collaboration with the staff and specialists, the researcher conducted field observation during the visits as well. These observations were done to strengthen the research understanding related to these children's activities, behavior, disabilities and needs.

Through the multiple site visits, this research found additional information related to children with Cerebral Palsy. The information includes:

1. Identifying the behavior, disabilities and alternative communication method of a child with Cerebral Palsy.
2. The plan on conducting evaluation or experimental procedure.

The outcome from the observation related to the behaviors and disabilities of children with Cerebral Palsy is summarized and shown in Table 3-1.

Table 3-1Behavior, disabilities and alternative communication method of children with Cerebral Palsy

Attributes	Description
Behavior and Attitude	A lack of concentration and focus in carrying out a task. Easily distracted by the presence of strangers. Drooling resulting from open mouth. Lack of passion.
Disabilities	Difficulty in holding items in the hand – physical disability. Inability to walk and stand well – physical disability.
Intelligence	Slow in comprehending and understanding given instruction. Difficulty in identifying complex words, items and options.
Communication	Difficulty in making coherent speeches – speech disability. Unresponsive towards request. The use of sign language to communicate (Makaton Sign Language).
Emotion	Difficulty to identifying and distinguishing some of their emotions and facial expressions. However, they have specific facial expression patterns when trying to express a specific emotion.

The difficulties experienced by children with Cerebral Palsy were verified throughout the site visit. Apart from their physical disability, the most difficult problem faced by these children was communication as stated in the literature review (refer to Figure 2-2).

The specialists in the special school suggested few ways for children with Cerebral Palsy to communicate and that included eye's gaze, Makaton sign language and even through their facial expression. Among all the suggested communication methods, Eye's gaze and Makaton sign language required participation and interaction between the children and teacher whereas the use of facial expression did not require such direct participation and interaction.

The teachers of the school claimed that their facial expression might not be their initial emotion. A child with Cerebral Palsy showing a sad face did not mean that he was really sad. It was only his inability to control his face muscle (4MyChild 2009). Only, the teachers, parents and care takers could distinguish their facial expressions because of close their relationship with them. The teachers commented that these children would show specific or unique facial patterns when trying to communicate or to express an emotion or a need.

3.3 Proposed Solution

In order to overcome limitations on existing AAC tools to be able to assist children with Cerebral Palsy to communicate, this research proposed an assistive communication tool which could monitor these children and indirectly monitoring them. This model utilizes the human facial expression as the human face is rich with communication information (Branco and Encarnacao 2004) and it is used to give responses and feedbacks (Busso et al. 2004; Cohen, Grag, and Huang 2000; Ekman 1997). Even the teachers in the special school agreed with the facts and explained that children with Cerebral Palsy could communicate their emotions and needs through their facial expression.

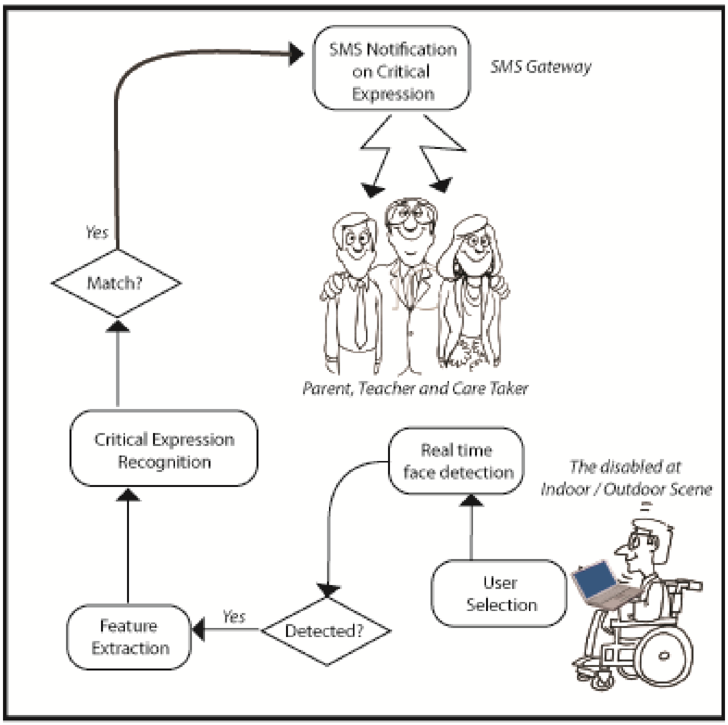


Figure 3-2 Proposed solution overview

Figure 3-2 illustrates the overview on how the proposed communication tool could help in the communication of Cerebral Palsy children efficiently, effectively and flexibly. The proposed communication tool was given a name, Assistive Communication Application, “ACApp”.

ACApp has a user interface to assist the interaction of computer and the children with Cerebral Palsy. Hence, these children do not need to coordinate any device when using ACApp.

ACApp is installed in a portable computer, like a laptop, with a built-in video camera. An external support device or GSM modem is attached to the laptop for sending critical expression messages. Then the portable laptop is mounted on the wheelchair so that the child can be monitored at any time and at any place.

During the monitoring session, each child’s facial expressions are captured and are distinguished by ACApp. A notification will be sent using Short Messaging System (SMS) to a parent’s or care taker’s mobile phone when critical expressions are detected during the monitoring session.

Generally, ACApp is used by two groups of users, the children with Cerebral Palsy as well as the guardians i.e. parents, teachers and caretakers. The children with Cerebral Palsy do not need to physically coordinate with any hardware or device when using the proposed prototype. Guardians will be the people who setup and train ACApp to recognize children’s facial expression.

ACApp runs in two modes, the training mode and the monitoring mode. In the training mode, ACApp will be trained to recognize the disabled child’s facial expressions and to distinguish them with labels. The guardians will train ACApp by creating a new profile or by selecting the existing profile of the children involved.

For each of the expressions trained, guardians can instruct ACApp to flag critical expressions that will be notified upon occurrence during the monitoring mode. Guardians can set his or her phone number so that ACApp could notify them by sending a SMS if any critical expression occurs during the monitoring session later. ACApp is ready to be used for monitoring children with Cerebral Palsy once the training is done. A GSM modem is required to send SMS.

3.4 Requirement Specification

The requirement specifications for the proposed prototype were defined after reviewing related literature as well as coordination with the special school. These requirements included scope of prototype, functional, non-functional, constraints, database, hardware and also developmental tools.

3.4.1 Scope of Prototype

The scope of the prototype was stipulated to define the boundary of the research and to ensure that it is on the right track to achieve the objective set. Any issue beyond the scope would not be discussed. The system scope included:

R1. The prototype is capable of performing the following functionalities: video capturing, images processing, template storing, template matching, and Short Messaging System (SMS) notification.

R2. The prototype collaborates with hardware devices i.e. laptop, webcam and GSM Modem only.

R3. The prototype uses only one language. i.e. English

R4. The prototype shall be able to run under various environments i.e. indoor, outdoor, day or night time.

R5. The prototype shall be able to store 100 profiles with 10 labels in each profile and 20 sample templates for each label (approximately 80 Megabytes in total of disk space on a laptop).

R6. The prototype is developed purposely to help and assist children with Cerebral Palsy in their communication.

R7. The stakeholders involve in this prototype includes the guardians (parents, care takers and teacher of the child with Cerebral Palsy), the child with Cerebral Palsy, and the prototype developer only.

3.4.2 Functional Requirements

A series of functional requirements are defined in this section to further elaborate on how the prototype shall behave and respond upon user's request. Functional requirements for the prototype include profile management, notification setting, face detection, expression training, expression recognition, and alert notification.

R8. Profile Management

The prototype shall allow guardians to manage children's profile (i.e. create, edit and remove profiles)

R9. Notification Setting

The prototype shall allow guardians to define phone numbers and to flag specific children's expression to be notified.

R10. Face Detection

The prototype must be able to search, extract and further process children's faces as in the image file. The face detection shall works under various environments and lighting conditions. It also shall works within the range of less than 80 centimeter between the subject and the camera.

R11. Expression Training

The prototype must be able to store the processed image in training mode as template images and label it properly for retrieval purposes in other processes later.

R12. Expression Recognition

The prototype must be able to recognize human facial expression in the monitoring mode. This process is done with template images retrieval, comparing template image with new images (matching), and returning the correct label.

R13. Alert Notification

The prototype must be able to send notifications to user-defined phone number of by producing an alarm when a flagged label or a critical expression occurs in the monitoring mode.

3.4.3 Non-Functional Requirements

The main focus on Non-Functional Requirements (NFRs) for this prototype includes usability, reliability, environmental conditions, maintainability, extensibility and portability.

R14. Usability

The prototype shall have a user friendly Graphic User Interface (GUI) as the guardians might only have moderate computer skills and literacy. The GUI has to be as simple as possible by minimizing the use of command buttons. Different processes GUI shall be placed, either using grouping or locating them, on different tabs for easy access.

On the other hand, the prototype also provides users with help and guidelines for troubleshooting i.e. error messages and codes for reference. The messages or responses shall be clear, using understandable and straight forward description to avoid user from getting confused when using the prototype.

R15. Reliability

Memory leakage is the main concern in this issue; therefore unused memory consumption in between processes shall be released immediately to avoid prototype from crashing.

The prototype must be able to process and return accurate result as it is trained to recognize human facial expression. In other words, the prototype has to deliver correct messages to the correct receiver based on user input and configured settings.

It must be able to store multiple user profiles with a large number of texts and images. Therefore, the prototype has to keep and organize data trained by users properly.

R16. Environment Conditions

The prototypes shall be installed on a portable computer attached to the wheelchair while the monitoring process is performed. The prototype also has to work under different environmental conditions i.e. day time, night time, indoor and outdoor without affecting the result or response.

R17. Maintainability

The prototype must be able to organize related data in an appropriate sequence or directory due to a large amount of data to be stored, processed, re-used and to be accessed by multiple users. Data backup shall be included to avoid data loss or accidental deletion.

R18. Extensibility

The prototype shall be able to be expended or enhanced in future to fulfill the user's changing needs and requirements in the future. All additional modules can be plugged in easily without restructuring the entire prototype or disturbing other retained modules.

Each error occurred during run time shall be recorded or logged to assist in debugging and for further enhancement purposes.

R19. Portability

The prototype shall be easily installed in all kinds of laptops. This prototype only works with personal computers i.e. the laptop. The laptop is the best solution in this case due to its light weight and mobility.

3.4.4 Constraints

Due to the limitation of the time frame and resources, three constraints were defined in developing the proposed prototype. These constraints included language, compatibility and service constraints.

R20. Language Constraints

For the development and testing stage, only one language is used–the English Language. The prototype's documentation, error messages, troubleshooting guides as well as the user manual are all written in the English language. Multiple languages maybe added in future works.

R21. Compatibility Constraints

This prototype is built and run on Windows Operating System platform i.e. Windows XP and Windows 7 only.

R22. Service Constraints

This prototype is associated with the Short Messaging System notification for users' immediate notification via the mobile phone. The SMS service through this prototype is provided by the local telecommunication service provider. Any failure or delay in sending or delivering SMS due to poor signal or service interruption will not be addressed in this research.

3.4.5 Data Requirement

This prototype does not store data in any professional database management systems like Oracle, MySQL, or Microsoft Access. The data is stored in text files within the laptop's local drive. A local drive with at least 80 Megabytes of free space is sufficient to store all required data. The attributes and relationship between entities are illustrated in Entity Relationship Diagram in Figure 3-3. Data dictionaries are listed in Table 3-2.

The user has attributes of the user name (PK –varchar2), guardian's name (varchar2), guardian's contact (number), sound notify (boolean) and SMS notify (boolean). While attributes for label include index (PK - integer), label name (varchar2), username (FK –varchar2), and also alert notification (binary). Template Image contains only two attributes which is sample no. (PK - number) and index (FK - number).

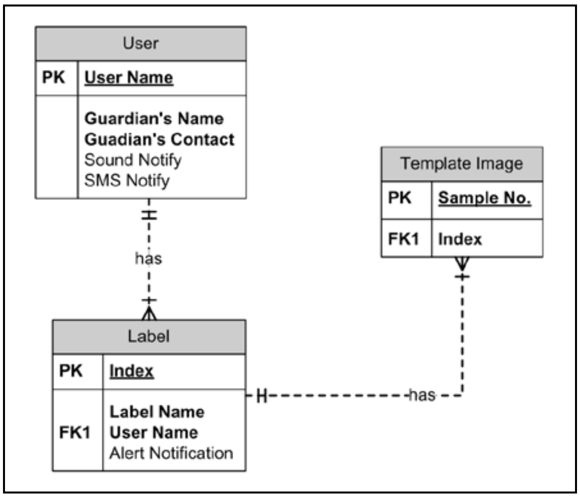


Figure 3-3 Entity Relationship Diagram

The user may have one or more labels while one label must belong to at least one user. One label must have at least one or more template images while one template image must belong to at least one label.

Table 3-2 Data dictionary

Table	Attribute Name	Data Type	Size	Key (PK / FK)	Null
User	Username	Varchar2	20	PK	NO
	Guardian Name	Varchar2	20	-	NO
	Guardian Contact	Number	10	-	NO
	Sound Notify	Boolean	1	-	YES
	SMS Notify	Boolean	1	-	YES
Label	Index	Number	3	PK	NO
	Label Name	Varchar2	20	FK	NO
	Username	Varchar2	20	FK	NO
	Alert Notification	Boolean	1	-	YES
Template Image	Sample Number	Number	3	PK	NO
	Index	Number	3	FK	NO

3.4.6 Hardware Requirement

The hardware listed below is required for the supporting device in operating the prototype. The requirement and functionality these hardware is described in detail in Chapter 4. List of required hardware:

1. Laptop
2. Webcam
3. GSM Modem

The prototype will be installed into a laptop with built in webcam. The GSM modem is connected through the USB 2.0 port.

3.4.7 Development Tools

The development of the proposed prototype requires a few development tools. A few open source APIs are used in the development which was highly recommended by other researchers. These development tools include Microsoft Visual Studio 2008, Intel Corp. OpenCV Ver. 1.1 Pre-Release, Microsoft .NET Framework 3.5 Service Pack 1, Microsoft Redistribution Package 2005 Service Pack 1 and Mobitek SMS Gateway Development Kit. All these tools are further elaborated in Chapter 4.

3.5 System Architecture

Figure 3-4 shows an overall architectural design for ACApp. This architecture was built based on requirements and specifications specified in Section3.4. ACApp was divided into three modules.

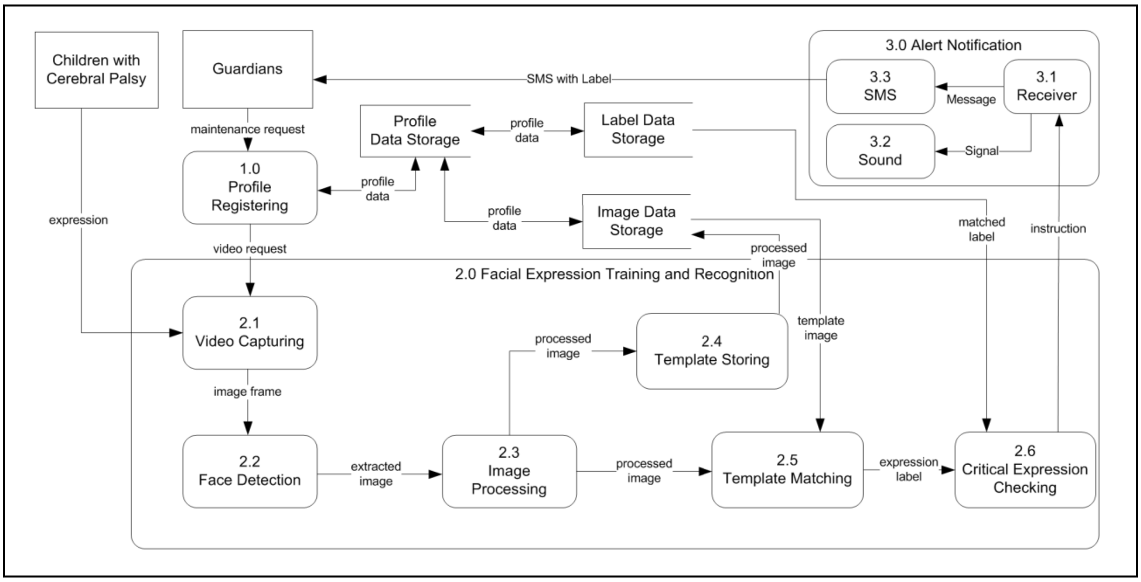


Figure 3-4 ACApp System Architectural Design

These modules include 1.0 Profile Registering Module, 2.0 Facial Expression Training and Recognition Module, and also 3.0 Alert Notification Module. Apart from these modules, ACApp uses three different data storages to store user profiles and images in local directories. These data storage include profile, image and label.

3.6 Software Process

This section describes the activities involving the users and ACApp in the main tasks. These tasks are elaborated based on the requirement specified in Section 3.4.

The use case for ACApp was finalized to seven (7) different tasks which are illustrated in Figure 3-5. These tasks includes 1) program startup 2) add new profile, 3) train expression, 4) flag or un-flag labels for notification (critical expression), 5) run monitoring, 6) delete expression and 7) lastly delete profile.

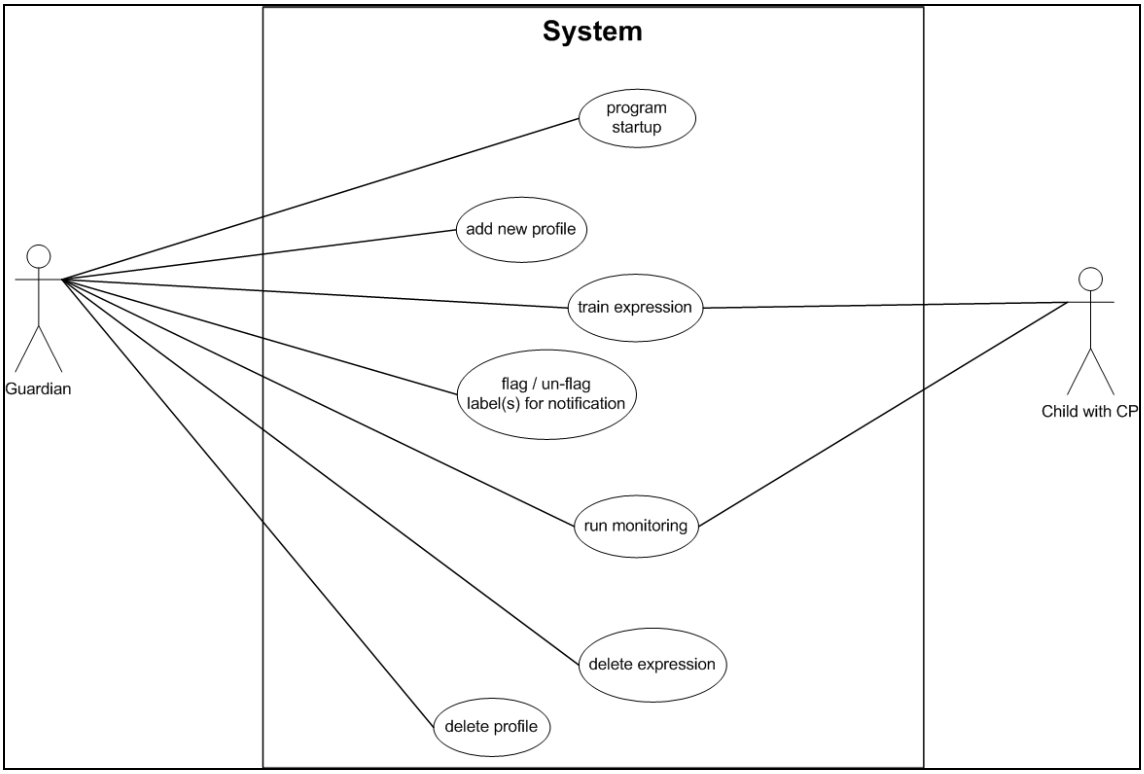


Figure 3-5 Use Case Diagram for the proposed prototype

Generally, guardians will perform all the seven tasks. In the initial stage, guardians will add a new profile into ACApp; the profile belongs to the children with Cerebral Palsy. Guardians can start to train ACApp by adding the children's expression for the prototype to recognize. This can be done by capturing the children's expression via a webcam. This expression will be labeled accordingly by the guardians to distinguish the different expression of the children. After the prototype has been trained with all the required expression the guardian can flag or unflag specific expressions to be notified upon occurrence during the monitoring session.

Guardian can then instruct ACApp to monitor the child after it is well trained. They will be notified by ACApp if any of the flagged expressions occurs during monitoring session.

Furthermore, guardians can remove or delete unused expression whenever necessary. Lastly, they are allowed to delete the child's profile and their trained expressions.

3.6.1 Program Startup

The goal for this process is to start up ACApp and to load the necessary data from the data storage which is shown in Figure 3-6. The process starts when the guardian executes the program, then the program will retrieve a list of profile from data storage and display it on the prototype interface upon receiving it. The process ends on the user's command.

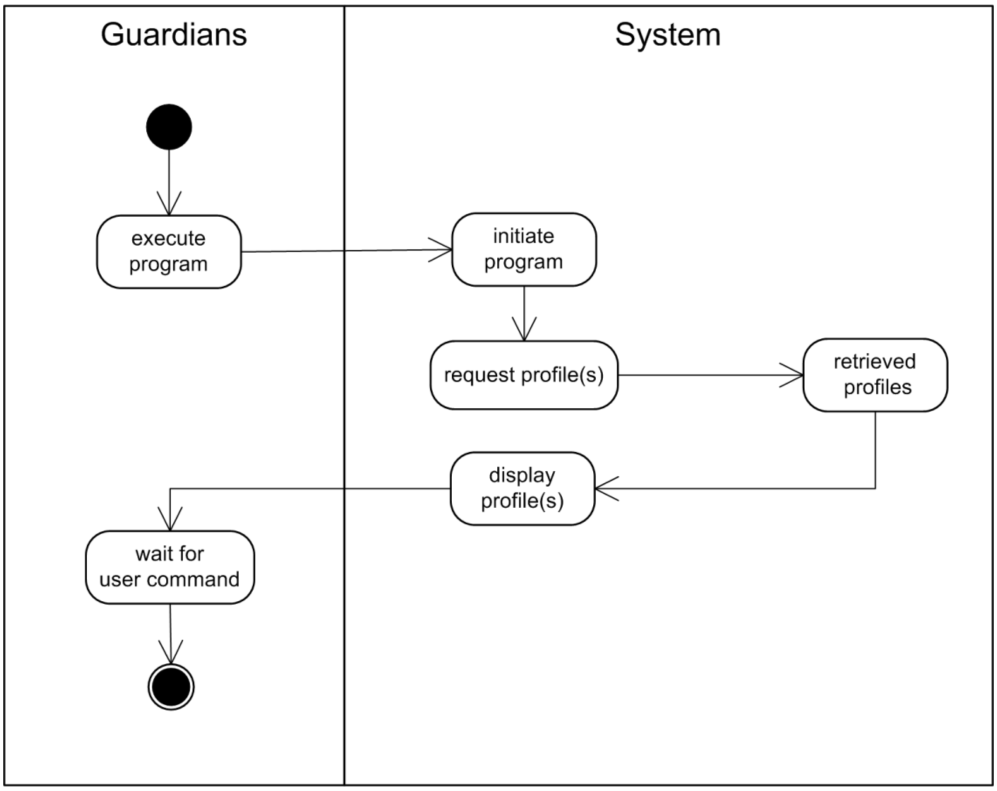


Figure 3-6 Activity Diagram for the process of Program Startup

3.6.2 Add New Profile

The goal for this process is to add new or registering new profiles into ACApp and the process is illustrated in Figure 3-7. The guardian is required to insert profile information such as profile name, and contact to be notified. The prototype will check if the information entered is complete or not. If the information is incomplete or duplicated, the prototype will ask the guardian to re-enter the profile information. If the entered information is complete, the prototype will add the new profile into the profile data storage. The process ends when a new profile is successfully added into the profile data storage.

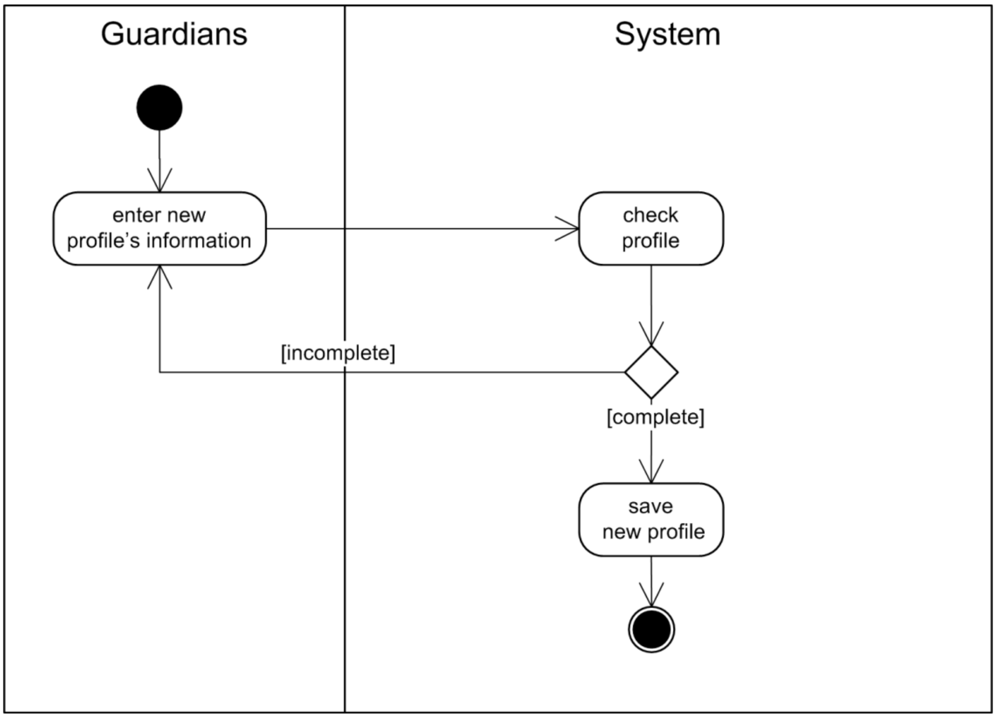


Figure 3-7 Activity Diagram for the process of Profile Registering

3.6.3 Train Expression

The goal for this process is to add a new label or expression into a specific profile in ACApp which is shown in Figure 3-8.

The guardian has to select a profile from the list displayed before adding a new expression and the prototype will request and display the selected profile data on the prototype interface. Later, the guardian is required to enter a label for the new expression that is to be trained. After that, the guardian can run the training.

When performing expression training, the webcam captures the child's expression as image sequence (frames) or video and then send each captured frame to the prototype for processing. The prototype will search for a human face for each frame received and will discard the frame if no face is found then request for another frame from the webcam. If human face is found in the search process, the prototype will extract it and proceed to image processing.

The processed image will be saved into the image data storage and the prototype will check if the training session is complete or not. The prototype will request subsequent frames from the webcam if the training is not yet complete. If the training is complete, it will update the profile data storage and add the new label into the label data storage.

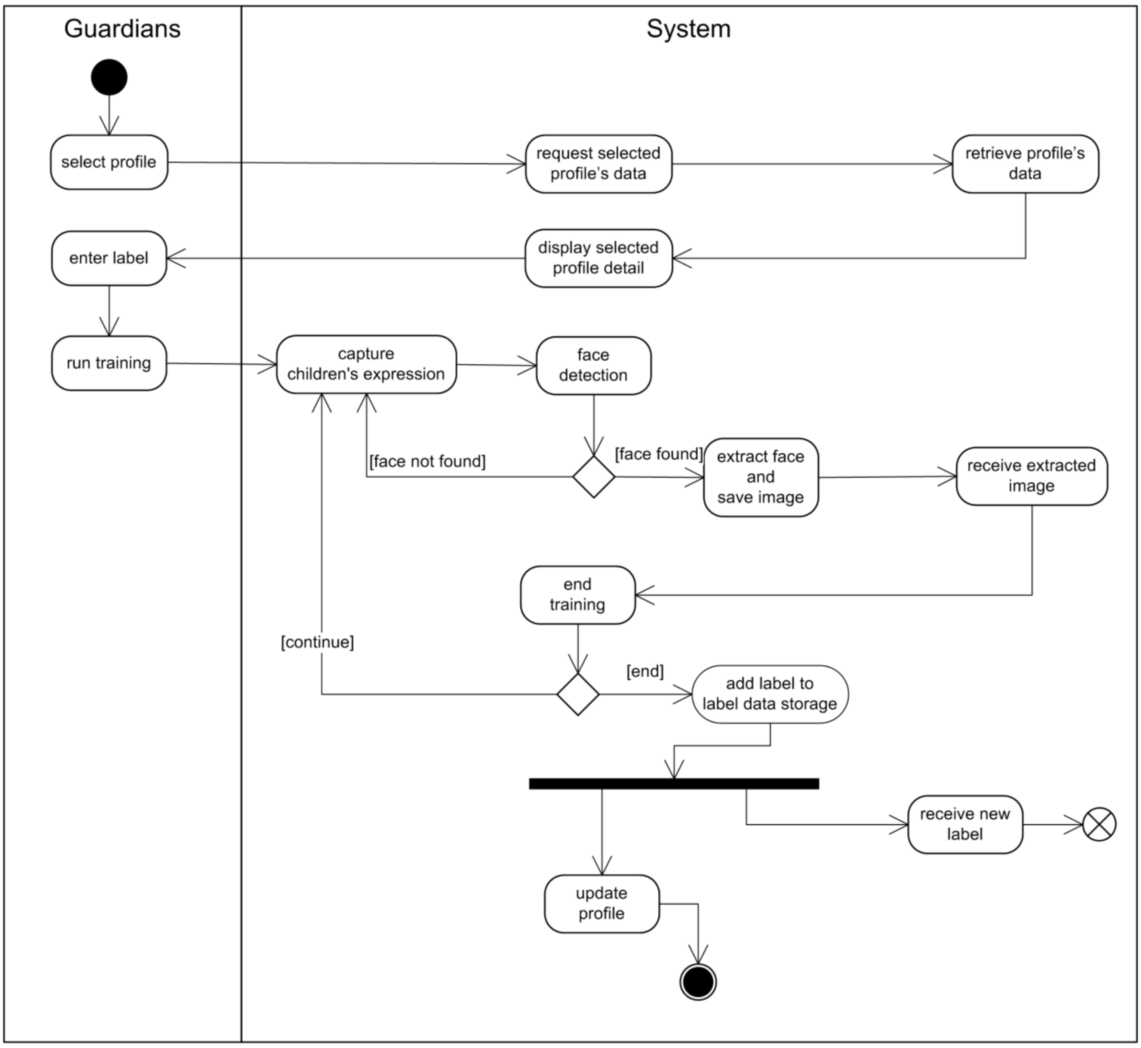


Figure 3-8 Activity Diagram for the process of Adding a New Expression

This process ends when both profile data storage is updates and the insertion of the new label into the label data storage is successful.

3.6.4 Flag or Un-flag Labels for Notification

The goal for this process is to flag or un-flag the children profile's labels for notification or cancels the notification and it is illustrated in Figure 3-9. When guardian selects a user profile, the prototype will request for the list of labels available in the data storage and display it on the interface. From there, the guardian can flag or unflag the desired labels to be notified or otherwise and confirm the updates. This process ends when the label data storage is successfully updated.

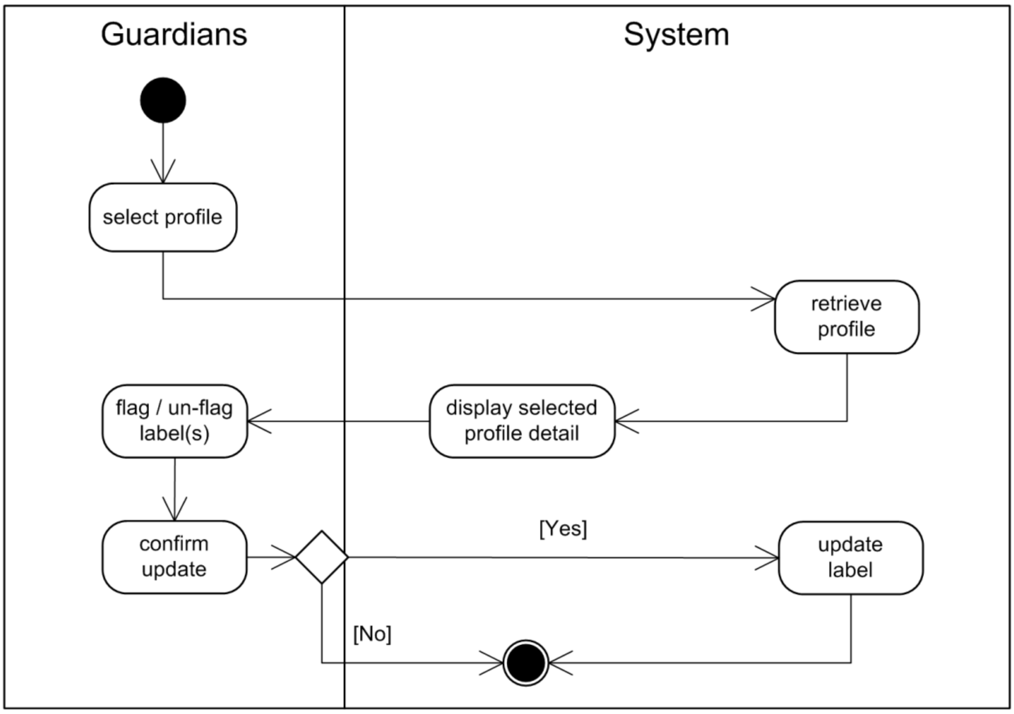


Figure 3-9 Activity Diagram for the process of Notification Setting

3.6.5 Run Monitoring

The goal for this process is to monitor children with Cerebral Palsy and to notify guardian if the expression flagged with the labels are detected during the monitoring session. The whole process of monitoring is illustrated in Figure 3-10.

The process starts once the guardian selects a profile and the prototype will retrieve it from the data storage and will then display the selected profile data on the prototype interface. When this has taken place, the guardian can then run the monitoring session. The webcam will capture the children's expressions as image sequence (frames) or video and then send each captured frame to the prototype for processing. The prototype will search for a human face for each frame received and will discard the frame if no face is found. It will then request for another frame from the webcam. If a human face is found in the search, the system will extract the facial region and perform further image processing.

The extracted image will be kept temporarily while the template image is requested from the image data storage. Once the template image is retrieved, the matching of the extracted image and the template image starts. The prototype will request a list of labels from the label data storage and will perform critical expression checking.

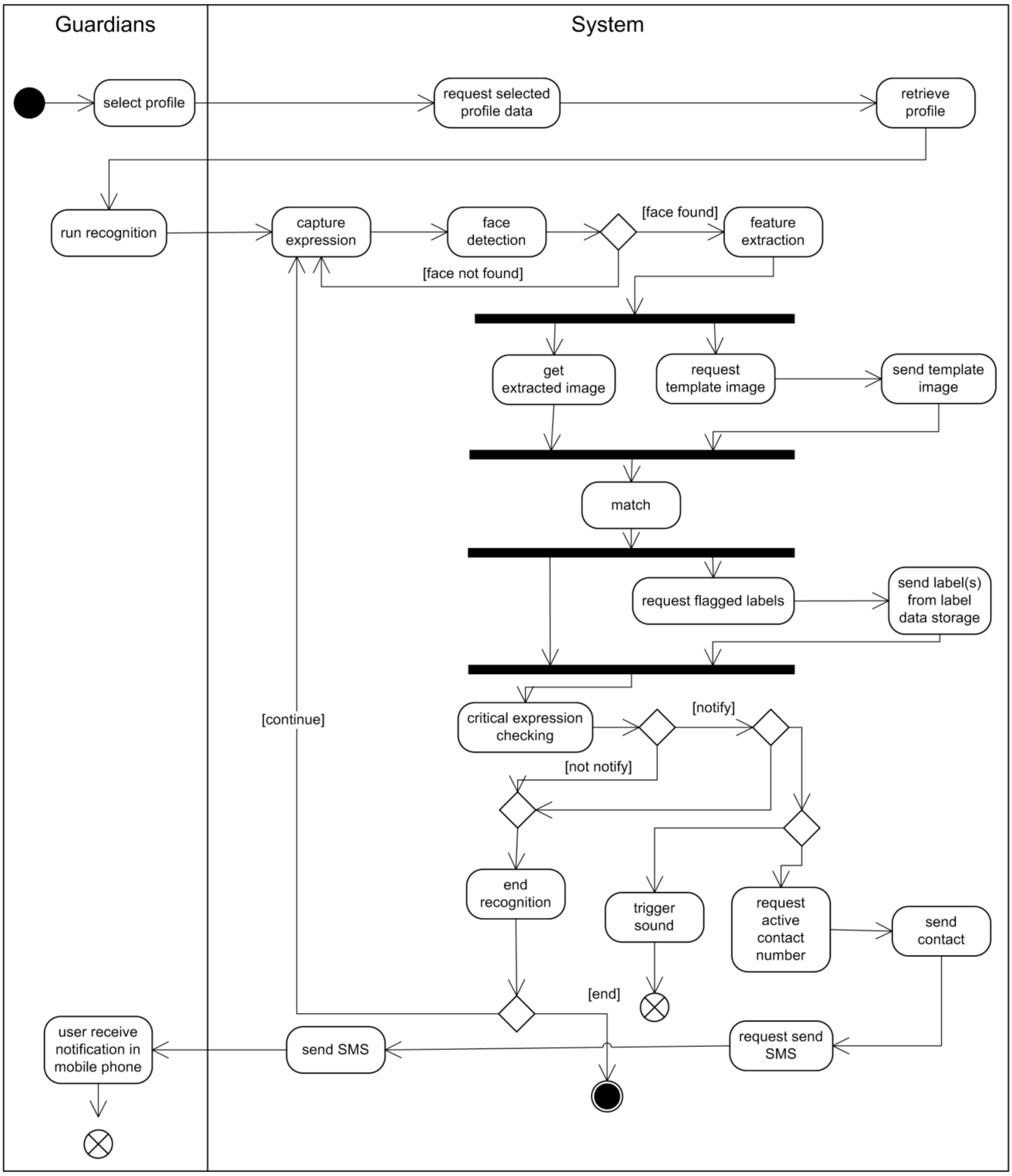


Figure 3-10 Activity Diagram for the process of Expression Recognition

If the result of the critical expression checking is flagged as "notify", the prototype will trigger a sound and at the same time, retrieve the guardian's contact number from the profile data storage. The prototype will gather the matched label (from matching) and the contact number (from profile data storage) and will then request the GSM modem to send an SMS.

The GSM modem will send an SMS with information of the critical expression, the name of child and the time of occurrence to the guardian. At this point, the process will continue with the "end recognition checking" state.

On the other hand, if the result of the critical expression checking is not flagged as “notify”, the prototype will check if the recognition process has ended. If not, it will continue to request for another frame from the webcam. The whole process ends if the “end recognition checking” process returns to an end.

3.6.6 Delete Expression

The goal for this process is to remove or delete a label as shown in Figure 3-11. When the guardian selects a user profile, the prototype will request for the list of labels available from the data storage based on the selected profile. The prototype will display the available labels on the interface upon retrieving them from the data storage.

At this point, the guardian can delete any label by confirming the deletion request. The prototype will also remove all template images of the selected label and then update the label data storage. This process ends when the label data storage is successfully updated.

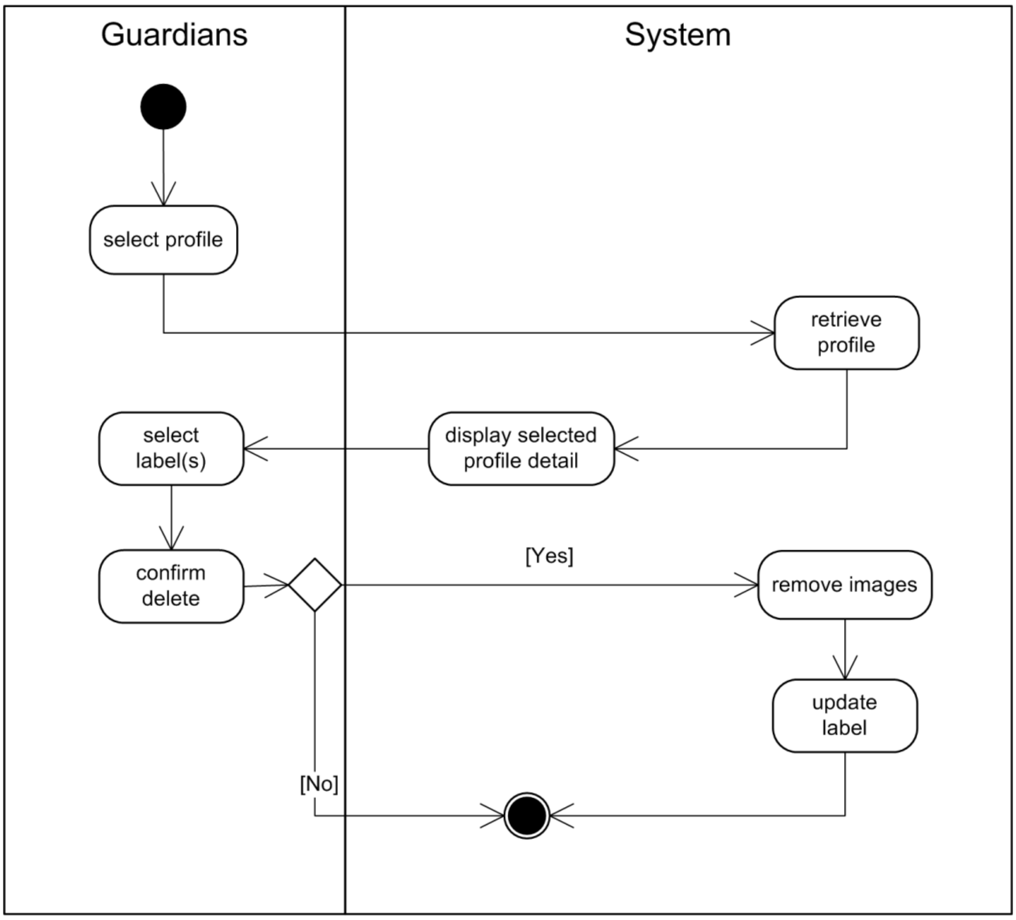


Figure 3-11 Activity Diagram for the process of Delete Label

3.6.7 Delete Profile

The goal for this process is to remove or delete a profile which is shown in Figure 3-12. The process starts after the user has selected a profile and the prototype will request for the selected profile data, retrieve it from the data storage and display it on the prototype interface. The guardian can confirm deletion request of the selected profile and the prototype will delete the profile.

In the deletion process, the prototype will delete the profile's template images from the image data storage as well as the profile's labels from the label data storage. The prototype will request for profile deletion once the template images deletion and the label deletion are complete and then the selected profile will be removed from the profile data storage. This process ends when the profile data storage is successfully updated.

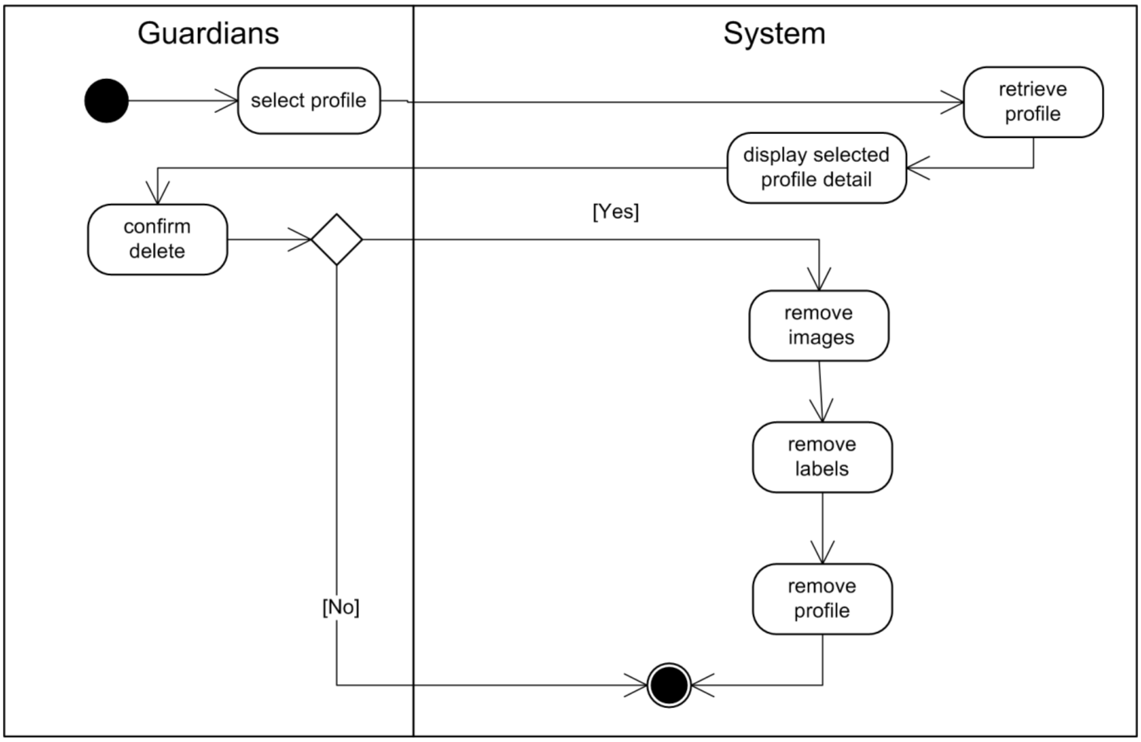


Figure 3-12 Activity Diagram for the process of Delete User Profile

3.7 Summary

A summary of findings from site visits in a special school were discussed. These findings were identified in association with the literature review related to the problems faced by children with Cerebral Palsy or specifically their communication difficulties. Speech pathologies and therapists suggested employing the Augmentative and Alternative Communication (AAC) tools to assist children with Cerebral Palsy in their communication. However, recent research has found that the suggested AAC tools were less efficient and were limited in real time usage.

Throughout the literature review and also the site visits, the problem of past research concerning communication difficulty experienced by children with Cerebral Palsy and the limitation of existing AAC tools in solving their difficulty were verified. These problems inspired the research to propose a solution utilizing an assistive communication tools (ACApp) to assist children with Cerebral Palsy to communicate with others.

The requirement and specification for ACApp were derived from the outcome on both the literature review as well as the site visits. These requirements included scope of prototype functional and non-functional, constraints, hardware, database, and also development tools.

An overall conceptual modeling of the proposed prototype was discussed in this chapter. The system development life cycle (SDLC) methodology has been chosen as a guideline for the prototype development. The overall idea and ACApp workflows were described in the system architecture. Generally, ACApp was divided into three major modules namely profile registering module, facial expression training and recognition module, and alert notification module. These modules were associated with three separated data storage, namely profile data storage, image data storage, and label data storage.

The functionality of ACApp was illustrated in the seven use cases format. These use cases included program startup, adding new profile, train expression, flag or un-flag expression, run monitoring, delete expression, and delete profile. Each use case was further explained by using activity diagrams.

The whole conceptual model and requirement specified in this chapter will be implemented in the next chapter.

Chapter 4: ACApp Implementation

In this chapter, the process of ACApp implementation will be discussed in details. Each stage of its implementation process was well defined and planned beforehand. This chapter will cover two major parts; 1) the pre implementation, and 2) the prototype implementation.

The pre-implementation process covers most of issues related to its preparation and the tools to be used in the prototype implementation. These issues include the usage of developmental tools usage, the available techniques for development, the comparison among available techniques, and the implementation plans.

The prototype implementation process covers the stages involved in developing ACApp. Generally, the prototype implementation is divided into four stages. Each stage represents an improved development or version of ACApp. The description, building, experiment, problems occurred and the outcomes are further discussed at the end of every stage.

4.1 Pre-Implementation

The implementation preparation is highly concern in this research to avoid unnecessary rollback from happening during the actual prototype implementation. These preparations include development tools, techniques, and the implementation plan.

4.1.1 Development Tools

Hardware and software or application programming interfaces (API) were required in order to develop an application. In this research, the hardware and software for developing the proposed prototype were meticulously defined and selected. However, in this program, the selection of the hardware and software were easier as it depends on the developer’s preference.

4.1.1.1 Hardware

The requirement and functionality for the hardware is described in details in the following section. The hardware required to operate ACApp:

1. Laptop – a device used to develop, run and test the developed prototype.
2. Webcam (Built-in / External) – a device used to capture video or sequence images for processing in the prototype.
3. GSM Modem – a device used to deliver SMS upon request by the prototype.

The proposed prototype does not have any physical appearance as it will be installed into a mobile PC or commonly known as a laptop. Hence, all other required hardware must be light weight to fulfill the efficiency requirement.

a. Laptop

The developed prototype was installed into a laptop (PC). A laptop was used in this research to develop, run and test the developed prototype. The specification for the laptop is shown in Table 4-1 below.

Table 4-1 Laptop specification

Processor	2.20GHz /800Mhz FSB/2MB L2 cache
Memory	2GB DDR2 1333MHz Memory
Hard Drive	320GB hard drive (5400RPM)
Camera	2.0 Megapixel
Power	6-cell Li-ion Battery
Ports	USB 2.0
Speaker	Yes
Operation System	Windows 7

b. Built in Webcam / Extended Webcam

An external webcam is needed for a PC that does not come with a built-in webcam. The built-in or external webcam should at least fulfill the specifications as shown in Table 4-2.

Table 4-2 Webcam specification

Photo Quality	> 1.3 Megapixels
Video Resolution	320 X 240 (QVGA) - 640 X 480 (VGA)
Video Capturing	> 15 frames per second
Cord	USB 2.0 (apply for external webcam only)

c. GSM Modem

A GSM modem is another mandatory hardware required in this research because this modem will be used as a depot for sending Short Messaging System (SMS) as notification upon request. Generally, a GSM Modem is a wireless modem or telecommunication device that uses GSM wireless network. It comes in various interfaces, such as PCMCIA Type II, USB, and Serial.

Figure 4-1 GMS modem type Q24-U-657f

Specifically SMS Gateway Development Kit for GSM Modem Type Q24-U-657f (Mobitek System 2008) which is shown in Figure 4-1 was used in this research to handle the task on delivering the SMS. The specification of GSM modem is listed in Table 4-3.

Table 4-3 GSM model specification

Brand / Model	Q24 / U-657f
Radio Frequency	GSM and GPRS Class 10
Quad Band	850 / 900 / 1800 / 1900 MHz
Core Module	Wavecom Q24 Plus
Cord	USB 2.0
Dimension	120 x 54 x 25mm (Length / Width / Height)
Weight	130g
API	COM/ACTIVEX Version 5.3 (MobitekSMSAPI5.dll)

4.1.1.2 Software for ACApp Development

The prototype development requires software and application programming interfaces (APIs) such as Microsoft Visual Studio 2008, Intel Corp. OpenCV Ver. 1.1 Pre-Release, Microsoft .NET Framework 3.5 Service Pack 1, Microsoft Redistribution Package 2005 Service Pack 1 and Mobitek SMS Gateway Development Kit.

a. Microsoft Visual Studio / C++ 2008

Microsoft Visual Studio / C++ 2008 is an integrated development environment (IDE) used by software developers to develop windows based applications. It can be written in both native code and managed code for all platforms. Usually, the developed application is supported by Microsoft Windows, Windows Mobile, Windows CE, .NET Framework, .NET Compact Framework and Microsoft Silverlight either in developed console and graphical user interface applications along with Windows Forms applications, web sites, web applications, and web services.

This research has chosen Microsoft Visual Studio 2008 to develop ACApp using C++ Windows Form Application template.

b. OpenCV

Open Source Computer Vision Library or OpenCV is a library of programming functions mainly aimed for real time computer vision which was developed by Intel. OpenCV is now supported by Willow Garage and it is free for use under the open source BSD license. This library is a cross-platform library which is compatible for the multiple operating system i.e. Windows, Unix, MacOS and newly Android OS. OpenCV focuses mainly on real-time image processing and other video related manipulations.

This research is using the OpenCV version of 1.1 pre-release as the core API for implementing major processes in ACApp.

c. Microsoft .NET Framework 3.5 Service Pack 1

The .NET Framework is a software framework for the Microsoft Windows operating system. It includes a huge library that supports several programming languages and allows language interoperability. Specifically, it is a plug-ins for IDEs and allows developers to wrap an application with Graphical User Interface (GUI).

The framework's Base Class Library provides user interface, data access, database connectivity, cryptography, web application development, numeric algorithms, and network communications. The class library is used by programmers, who combine it with their own code to produce applications.

Programs written for the .NET Framework execute in a software environment, known as the Common Language Runtime (CLR). The CLR is virtual machine so that programmers need not consider the capabilities of the specific CPU that will execute the program. The CLR also provides other important services such as security, memory management, and exception handling. The class library and the CLR constitute the .NET Framework. The .NET Framework was intended to be used by most new applications created for the Windows platform.

This research uses .NET Framework 3.5 as an additional library on major GUI parts and user interactivity features to fulfill user requirements.

d. Microsoft Redistribution Package 2005 Service Pack 1

Microsoft Redistribution Package 2005is a refined chunk of common library used on Visual Studio 2008 Service Pack 1 (SP1) and the .NET Framework 3.5 SP1. The refined works are a product of a combination of customer and partner feedback, as well as internal testing. These service packs offer Visual Studio and .NET Framework users improvements in responsiveness, stability and performance.

e. Mobitek SMS Gateway Development Kit (API)

SMS Gateway Development Kit (SMS GDK) is a collection of development kits for building SMS Gateway, SMS Server, SMS software, SMS application, and SMS solution. Software developer can utilize SMS GDK to build their own SMS Gateway to suit their needs.

This research has added SMS GDK into the application development process to drive or control the GSM Modem attached to the PC for the notification purpose of sending SMS.

4.1.2 Techniques Used

This research adopts the basic facial expression recognition system as one of the modules in the proposed prototype which has been stated in the chapter on conceptual model previously. In a basic facial expression recognition system, three processes were involved, the face detection, feature extraction and the facial expression recognition. However, no specific technique is used for feature extraction in this research. The entire detected face image found in the detection process will be used and processed without extracting any smaller features i.e. eyes, nose and mouth.

4.1.2.1 Viola-Jones face detection algorithm

In the face detection process, the Viola-Jones face detection algorithm was adopted in the proposed prototype after comparing with other techniques in chapter 2. The face detection technique is considered to be the most suitable technique compared with other techniques, due to its simplicity and its popularity with many researches in the object or face detection related field (Bartlett et al. 2003; Viola and Jones 2004; Datcu and Rothkrantz 2007; Littlewort, Bartlett and Lee 2007; Huang and Lin 2008; Wu et al. 2008; Zhan et al. 2008; Cho et al. 2009; Riaz et al. 2009; Zheng and Bhandarkar 2009; Piatkowska 2010).

The only disadvantage for the Viola-Jones face detection algorithm is its sensitivity to lighting. This disadvantage could be overcome by using brightness adjustment features to obtain the normal brightness or lighting condition on specific images.

Generally, the Viola and Jones face detection algorithm consists of three features (Viola and Jones 2004), an image representation called "Integral Image" which can be used by any system to compute quicker compared with other image representations; A simple and an efficient classifier which is built using Adaboost algorithm to select small features out of a larger set of features; and "cascade" which allows the background of an image to be discarded quickly. The details on how it works were described at 2.5.2.1 in Chapter 2.

4.1.2.2 Template Matching

This research has decided to adopt the Template Matching technique to perform expression recognition because it is widely used as well (Balasuriya and Kodikara 2000; Shan, Gong, and McOwan 2005; Chai et al. 2009; Yusoff et al. 2009; Xie and Lam 2009; Ong, Lu and Lau 2009). Besides, this technique is considered to be the simplest pattern recognition technique. For instance, a set of known images is stored in the template and when new images arrive, these images will be compared with all the images stored in the template. This is done by comparing each pixel value on both images. The new image is recognized when it matches an image available in the template. Specifically, the Template Matching technique is simple and convenient in term of memory and storage consumption.

4.1.3 Development Plans

The prototype development was divided into four stages or versions, due to the complexity of the proposed prototype. The prototype will be developed from the core module then it will slowly plug in with other modules and collaborate with the core module. It consisted of three major modules namely 1) profile registering, 2) facial expression training and recognition and 3) alert notification which is described in section 3.5 previously.

ACApp version 1.0 will start with module 2 – the facial expression training and recognition as this is the core module for the entire prototype. ACApp version 2.0 will refine and enhance ACApp 1.0. The module 1 – profile registering and module 3 – alert notification is developed and added in ACApp 3.0. Besides, this version will be done in GUI. ACApp 4.0 will be the final output with minor changes and transformation from ACApp 3.0.

Table 4-4 explains each prototype version which covers its own goal or expected outcome, its focus and a brief description on each version.

Table 4-4Prototype versions and implementation plans

Prototype Version	Goal / Expected Outcome	Focus	Description
1.0	Developing a basic facial expression recognition and identify potential problems	Module 2	Prototype development starts with the basic facial expression recognition using Viola-Jones face detection algorithm for the face detection and Template Matching techniques for the expression recognition. Experiment shall be conducted to identify potential problems by testing the techniques used for basic facial expression recognition development.
2.0	Improvement of the facial expression recognition and it readiness for the human expression recognition test	Module 2	Prototype improvement shall be made by overcoming problems identified in previous version's experiments. A mini test will be conducted to evaluate human facial expression recognition rate, responding time, speed and the efficiency of usage.
3.0	Developing module 1 and module 3, then combine the three modules and wrap the prototype with graphic user interface (GUI).	Module 1, Module 3 and Graphic User Interface	Developing module 1, profile registration and module 3, alert notification then combine with module 2 in prototype version 2 and collaboration with the GSM modem. Graphic user interface shall be added and minor changes or improvements on module 2 shall be done based on problems identified in version 2's experiments. A preliminary testing on targeted user will be conducted to evaluate the usability of the prototype before conducting the field testing.
4.0	Code clean up and fixing identified problems from previous version of prototype. Final prototype shall be ready for field testing and evaluation	Module 1, Module 2, Module 3 and Graphic User Interface	Final prototype shall be done by this version and all identified problems shall be tackled by performing minor changes on any of the related module. This version of prototype will be used for the real testing.

4.2 Prototype Implementation

ACApp was developed in Windows machine (laptop) using Visual C++ integrated development environment (IDE) with major application programming interface i.e. OpenCV, Microsoft redistributable package 2005 SP1 and also Microsoft .NET framework 3.5 SP1.

The technique used for face detection was Viola-Jones face detection algorithm while Template Matching technique was used in expression recognition. SMS notification was developed using SMS Gateway Development Kit (SMS GDK) which has been discussed before.

ACApp was developed according to the conceptual model which was defined in Chapter 3. The stages of development followed to the implementation plan which was located in the previous section.

4.2.1 ACApp V1.0

The main objective at this stage was to build the facial expression training and recognition module. A few problems or bugs occurred during the implementation process when conducting the experiments. The output of ACApp 1.0 was a CLR console based application.

Generally, ACApp 1.0 involves nine small processes. These processes are modulated into classes and they perform individual processes by receiving input then returning the output. These nine processes or classes are listed in Table 4-5.

Table 4-5 List of class or process in ACApp 1.0

Class	Input	Output	Description / Functions
FER()	void	void	main module to start up
Initialization()	void	vars	Initialize necessary variables
Capture()	void	frame	Capturing frame from webcam
FaceDetection()	frame	extImg	Searching face from frame, extract face and return extracted face image
ImageProcessing()	extImg	proImg	Perform image processing on image received and return processed image
TemplateStoring()	proImg	bool	Storing received image and return succeed or failed
Validation()	void	bool	Validating training process , and check if training process is end or not
LoadTemplate()	void	regVars	Load necessary template image and return
Matching()	proImg, regVars	label	Perform matching with the initial proImg and the templates loaded, and then return the label.

4.2.1.1 Implementation

Figure 4-2 shows the sequence of how ACApp 1.0 processes flow. The program starts with the variables initialization process. This process is required to initiate all necessary variables. The *Initialization* process includes printing messages which will be displayed on the console screen and waiting for the user’s command, i.e. run expression training or run expression recognition. The prototype startup screenshot is shown in Figure 4-3.

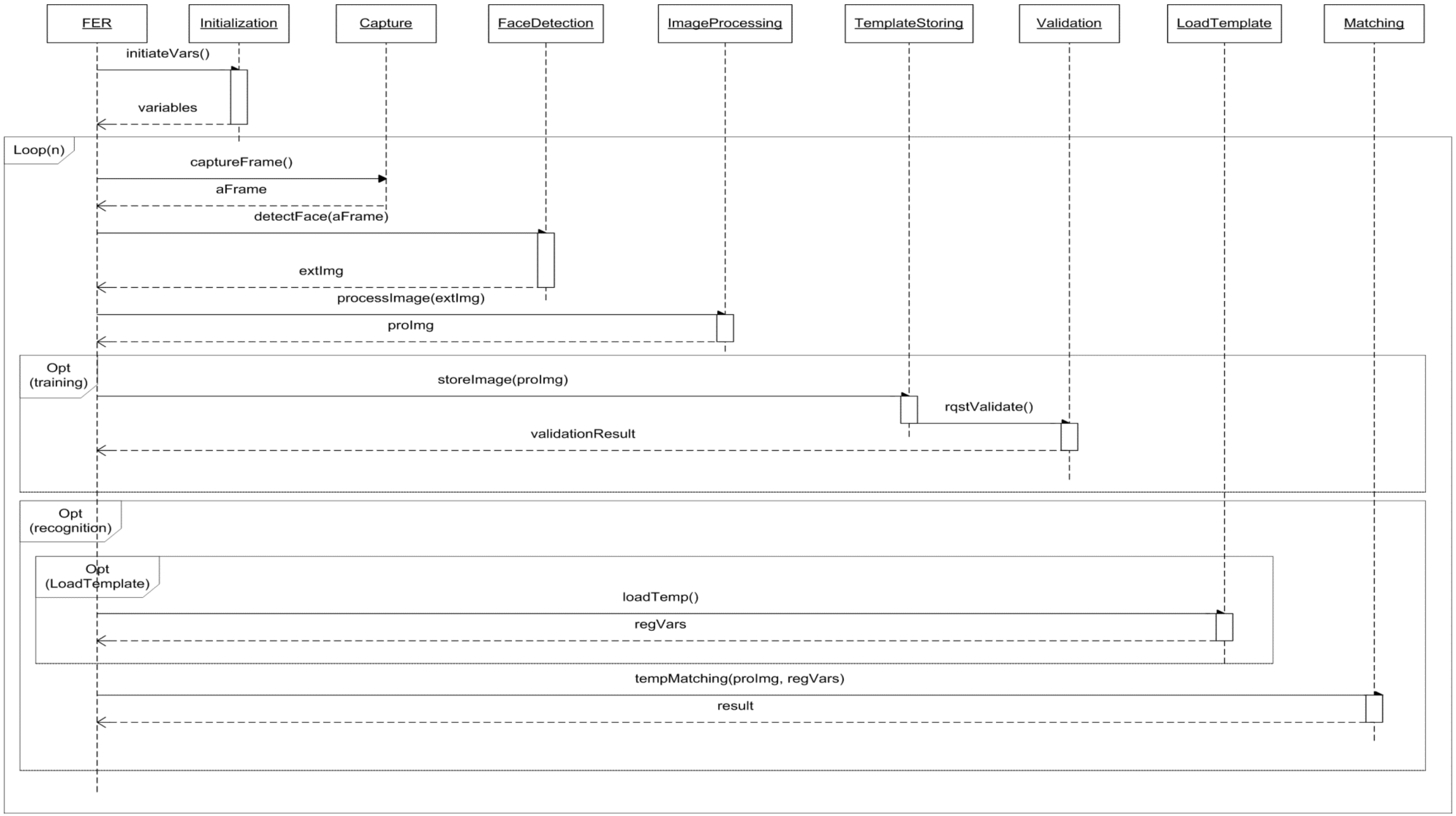


Figure 4-2 Sequence Diagram for ACApp 1.0

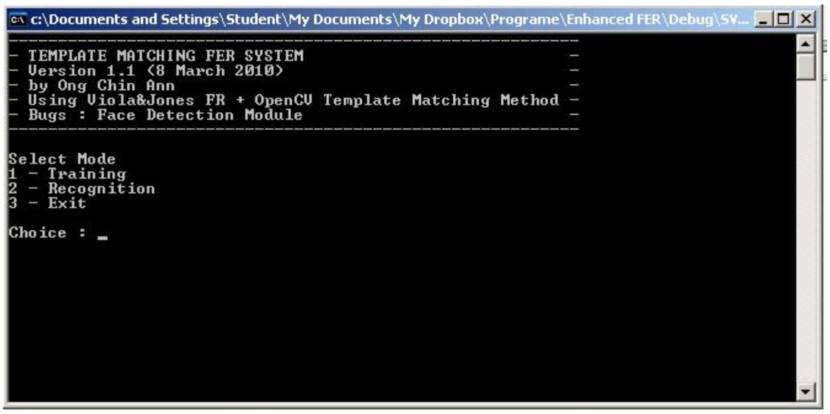


Figure 4-3 ACApp 1.0 startup screenshoot

Once the user's command is received (either training or recognition), *FER* will call *captureFrame()* and the *Capture* class will the trigger webcam to capture frames, and it will then return one frame as *IplImage* class object back to *FER*. *FER* will send the frame received to *FaceDetection* class for human face searching based on the Viola-Jones face detection algorithm and extraction.

The extracted image is returned back to *FER* and is sent to *ImageProcess* class for image resizing and color conversion. This process is performed to reduce the image size and to lighten the programs load in other processes later. The processed image is then returned to *FER*.

If the user selects expression training during the initialization process earlier, *FER* will send the processed image to *TemplateStoring* class where the image is stored in a local disk as a template based on the label inserted by the user done earlier on. Generally, the *TemplateStoring* process occurred once for every ten frames of the human face detected. *TemplateStoring* class will send validation request to the *Validation* class once the processed image is stored to check if the training process has reached the end.

The result of the validation will then be sent back to *FER* for logging and loop breaking purposes. If the training process is not yet ended, the program will loop back and call *captureFrame()* to request for the next frame, once variables are successfully reset to the initial state. The loop is broken when five template images are successfully stored into a local disk and the expression training process ends.

If the user selects expression recognition during the initialization process earlier, *FER* will request for stored variables such as the list of template images and labels from *LoadTemplate* class once. Then, *FER* will send the loaded variables together with the processed image to *Matching* class for template matching and identify the label of the processed image against loaded variables. The *Matching* class will send the identified label back to *FER* class and *FER* will display the matching results. The program will loop back and call *captureFrame()* to request for the next frame once the specific variables is successfully reset to initial state.

In ACApp 1.0, this loop will only break if the user terminates the program. A few prototype constants and attributes were predefined. These constants include limit range for face object (RANGE_LIMIT 100), extracted sample size (SAMPLE_SIZE 64) and total detected face frame (NUM_OF_DETECTED_FRAME 50)

a. Initialization Class

The *Initialization* process configures the prototype setting and also to initiates all necessary variables. The variables initializations from *highgui.h, cxcore.h, cvaux.h* and *cv.h* library provided by OpenCV include:

1. A variety of image reading and writing objects (*IplImage* class object),
2. video reading objects (*CvCapture* class object),
3. basic structure objects (*CvPoint , CvRect* class object),
4. Haar Feature-based Cascade Classifier for Object Detection (*CvHaarClassifierCascade* class object)

b. Capture Class

Figure 4-4 shows the process of *Capture* class in pseudo code. To start capturing images or frames from the webcam, the *cvCaptureFromCam()* function is called. The size of the frame is defined as QVGA size (320 Pixels for width and 240 Pixels for height) in this case as this size is sufficient to perform the face detection. A frame or captured image is retrieved every time when the function *cvQueryFrame()* is called. The retrieved frame is passed to the next steps for further processing.

```
INITIALIZE capture
SET capture to cvCaptureFromCam(webcam)
      GET frame from cvQueryFrame(capture)
```

Figure 4-4 Capture process

c. FaceDetection Class

Figure 4-5 and Figure 4-6 show the process of *FaceDetection* with face extraction. The captured image or frame from *Capture* class is received in this process. The received frame is sent to *cvHaarDetectObjects()* function with the Haar Cascade specified earlier and this function returns a list of sequence objects. The sequence object is a rectangular coordinate or *CvRect* class object which represents the location of the human face in the frame received.

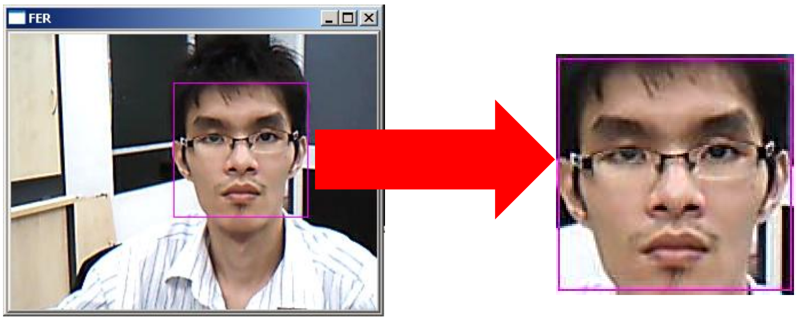


Figure 4-5 Face detection in initial frame and extracted image

A colored rectangle is drawn on the frame based on the sequence object, obtained if the width and height of the sequence object is greater than the RANGE_LIMIT, which was predefined earlier on. At the same time, it extracts the area where the human face is located as a separate image from the frame.

The range limit restriction is to avoid extracting a face image which has a smaller size and will thus cause difficulty in recognizing the facial expression later. The extracted image will be forwarded for image processing later.

```
GET face area from frame based on Haar Cascade
        IF width and height of face area is greater than RANGE_LIMIT
        THEN
                DRAW rectangle in frame with the size of face area
                obtained
                EXTRACT face area from frame
                RETURN the extracted image to main or FER
        ELSE
                DISCARD frame
ENDIF
```

Figure 4-6 Algorithm of FaceDetection

There were two problems encountered in the face detection process as discussed in Experiment 1-1 below.

d. ImageProcessing Class

Figure 4-7 and Figure 4-8 shows the flow of *ImageProcessing*. This process is performed when an extracted image is received and it reduces the size of the extracted image (in bytes) by resizing and gray-scaling the extracted image.

```
GET extracted image from face detection process
RESIZE extracted image to size 128 X 128
SUBTRACT resized image from rim
GENERATE gray scaled image by converting rimmed image
GERARATE histogram equalized image from gray scaled image
RETURN the histogram equalized image to main
```

Figure 4-7 Algorithm of ImageProcessing

The whole process start with resizing the extracted image to the predefined size using *cvResize()* function. Next, the rimming process is performed by downsizing the rim image and the resizing the facial image to remove unwanted areas. This procedure retains only the human face as the unwanted area can affect the recognition result.

A color conversion from RGB to gray scale imaging using *cvCvtColor()* function is performed once the rimming process is done. This conversion reduces the image's physical size (in bytes) which can speed up the prototype processes. The last process in image processing is histogram equalization by performing *cvEqualizeHist()* function on the gray scaled image. This process is carried out to increase the image contrast.

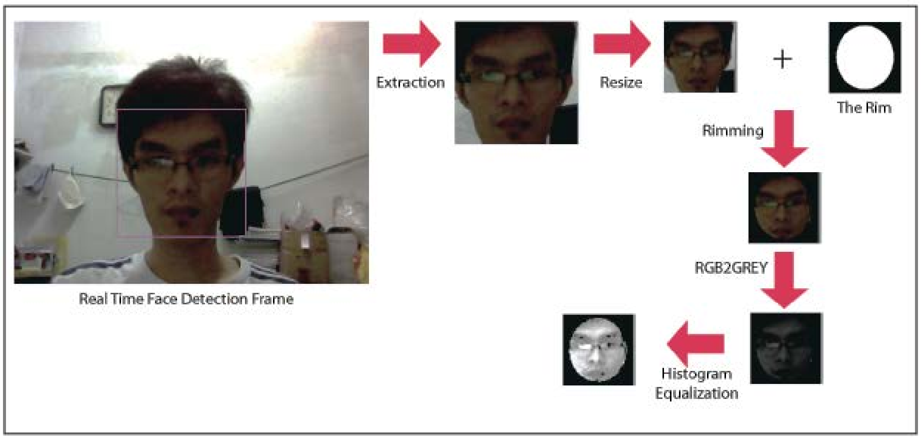


Figure 4-8 The flow of ImageProcessing

The output of *ImageProcessing* (processed image) is used for both *TemplateStoring* and *TemplateMatching* purposes later.

e. TemplateStoring Class

The *TemplateStoring* process which is shown in Figure 4-9 and Figure 4-10 is for the purpose of storing a processed image as template image. These stored images will be used in *TemplateMatching* for expression recognition later.

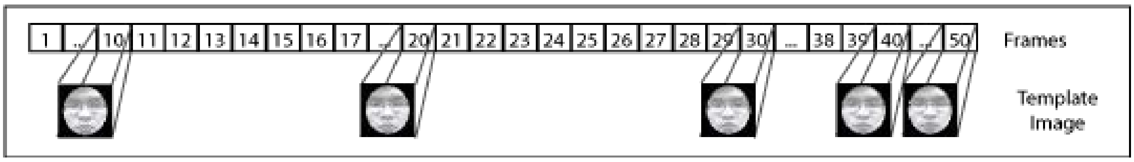


Figure 4-9 Frame extraction and storage

During the training session, ACApp will capture and process 50 frames for each expression as explained earlier, but only five templates are stored for each expression in ACApp 1.0. Hence, one processed image will be stored as template image for every ten processed images for the expression detected.

To effectively organize these template images, each of the template image is named according to this format *[label number]-[sample number].jpg* i.e. *1-3.jpg*which refer to the 3rd sample of image label number 1.

```
INITIALIZE count to 0
IF human face is detected in frame THEN
      EXTRACT face image from frame and perform image processing
      INCREASE count by 1
ENDIF
IF count modules by 5 equal to 0 and in training mode THEN
      SAVE processed image
ENDIF
```

Figure 4-10 Algorithm of TemplateStoring

f. Validation Class

Figure 4-11 shows the *Validation* process. This process is executed in the expression training mode and every time a processed image is stored. This process will validate and check whether the training process reached the end. If the training process has not or the number of human face detected frames is less than 50, the program will loop back to the *Capture* process, request another frame and then process it all over again. All other variables such as sample index number and frame count are updated while *Region of Interest* is reset.

```
WHILE exit program equal to false
    INITIALIZE count to 0
    INITIALIZE index sample to 0
    DO
        GET frame from capture
        IF human face is detected in frame THEN
            INCREASE frame count by 1
            PROCESS extracted image
            IF count modules by 5 equal to 0 and in training
            mode THEN
                INCREASE sample index by 1
                SAVE processed image
            ENDIF
        ENDIF
    WHILE count is less than NUM_OF_TRAIN_FRAME
    SAVE label
    WAITING user input
END WHILE
```

Figure 4-11 Algorithm of Validation

If the training process reaches the end or the number of human face detected frames exceed 50, the program will save the trained label into a text file and place it together with the stored template images.

g. LoadTemplate Class

The *LoadTemplate* process is performed when a user selects the recognition mode in the initial stage. The process details are shown in Figure 4-12. This process will load the trained template images which were done before and keep them as an image array for *Matching* purposes. This process happens once only when the user chooses to run the recognition mode. It starts with reading a text file from a local disk to obtain the number of the labels available and to retrieve the available labels. If the number of labels is equal to 0 or no text file is loaded, this means that no label has been trained previously and the program will terminate.

If there is a label available in the text file, the program will initiate a template image array and also a label array. For each index of the label, the program will read the text file and store the label based on the current index. At the same time, the program will load the template images and store them into the template image array as well based on the current index.

```
INITIALIZE number of sample to 5

GET number of labels from text file

    IF number of label is NULL or 0 THEN
        BREAK
    ELSE
        INITIALIZE template image array [number of label][number
        of sample]
        INITIALIZE label array [number of label]
        FOR each number of labels
            READ label from text file and store into label
            array [number of label]
            FOR each number of sample
                LOAD template image[number of label][number
                of sample] and store into template image
                array[number of label][number of sample]
            END FOR
        END FOR
    END IF
```

Figure 4-12 Algorithm of LoadTemplate

The reading of the label and the loading of the template images will loop until it reaches the maximum number of label multiplied by the number of samples available. All the variables i.e. number of label, label array and template image array will be passed to the next process, the *TemplateMatching*.

h. Matching Class

In the *TemplateMatching* process, the processed image will be used to match with all the template images obtained during training session. The final matching result is identified when the processed image matches with or closest to any of the templates available is found. In other words, the index of a template image will be marked as the matching result, if that particular template image achieves the least difference against the processed image.

In order to get a matching result, the template matching involves an additional image processing process which will produce a binary image as the outcome. These processes are illustrated in Figure 4-13 which include the subtraction of processed image from the template images, histogram equalization and range filtering.

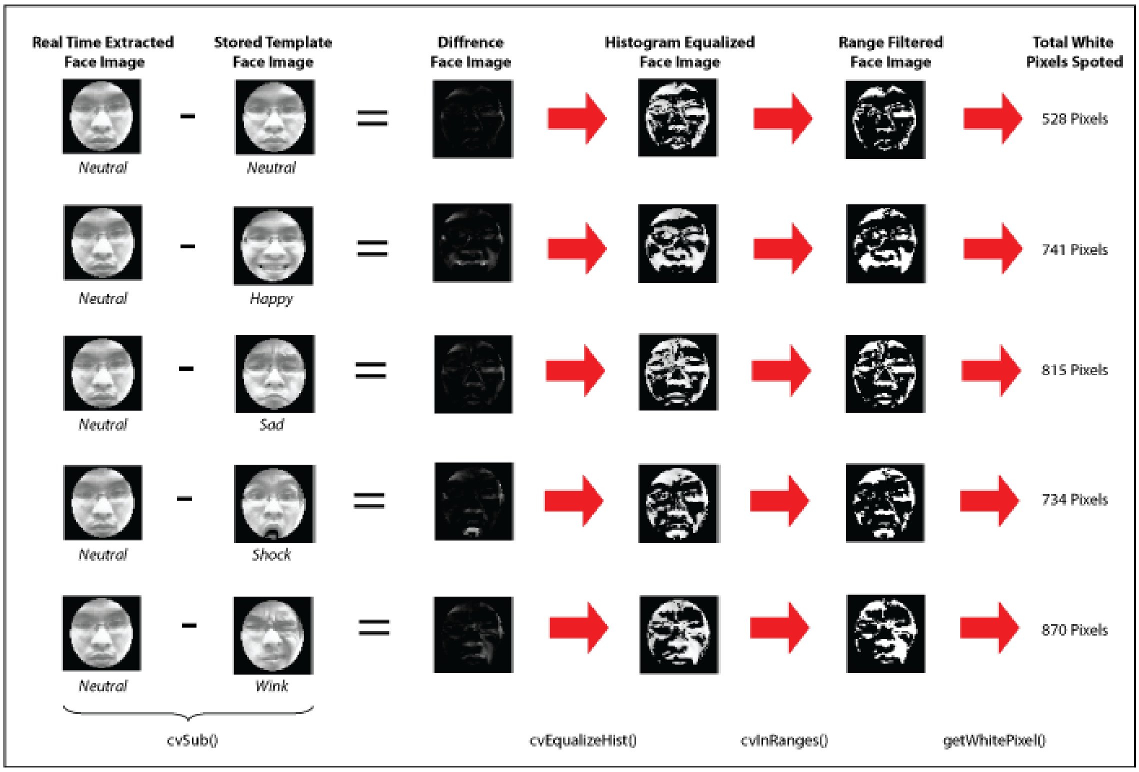


Figure 4-13 The flow of TemplateMatching()

The subtraction process is done purposely to get the differences between the processed images and template images. Histogram equalization is applied to the subtracted image so that the difference(s) can be distinguished better. A range filter is applied to the histogram equalized image so that only binary image is produced. The range filter converts a pixel value that equals or less than 10 to pixel value 0 (black) and the pixel value that is greater than 10 to pixel value 255 (white).

By performing range filter, the binary image will contain only black and white pixels. The black pixels represent no difference and the white pixels represent the difference between the processed and the template images. The white pixels contained in the binary images are calculated to measure the proximity of the template image to the processed image.

The number of rounds of matching with the processed image depends on the number of labels and samples for each label available. For instance, the available labels are four and the number of samples for each label is five, then the number of rounds to perform matching for one processed image will be equal to 5 X 4 = 20 rounds.

There must be one round of matching process producing the least difference among all the templates available. Then the program will identify the index number of that image and obtain the actual label from the label array at the position based on the index number. The flow of the template matching process is illustrated in the pseudo code in Figure 4-14.

Assume that all the expressions and labels were trained and loaded; initial frame was captured, human face was detected and processed image was produced.

```
INITIALIZE result equal to NULL
INITIALIZE white pixel to size of template multiply by size of
template
INITIALIZE result index equal to 0

FOR each number of label
      FOR each number of sample
            GENERATE different image by subtracting processed image
            with template image array[number of label][number of
            sample]
            GENERATE histogram equalized image by CALLING histogram
            equalization to different image
            GENERATE range filtered image by CALLING in ranges
            filter to histogram equalized image.
            COMPUTE image white pixel on range filtered image
            IF image white pixel is lesser than white pixel THEN
                  white pixel equal to image white pixel
                  result index equal to number of label
      END FOR
END FOR
result equal to label array[result index]
```

Figure 4-14 Algorithm of TemplateMatching

4.2.1.2 Experiments

A few experiments were conducted during ACApp 1.0 implementation. These experiments were conducted to measure both face detection and facial expression recognition rate.

a. Experiment 1-1 Face Detection

In the face detection experiment, a still image was fed to ACApp 1.0 to check if Viola-Jones face detection algorithm managed to search and detect a human face correctly or otherwise. Figure 4-15 shows that the Viola-Jones face detection algorithm did manage to detect human faces by drawing a blue square box around it.

Figure 4-15 Naïve Viola-Jones face detection on still image

However, the algorithm was less effective on a video or image sequences are applied to it. There were multiple detections with different boxes drawn in the frame. Furthermore, some detected objects were outside the perimeter of the face as shown in Figure 4-16.

This detection result could not be used in this research as the output of the extracted image can cause lower recognition rate. However, the causes and solutions to this problem were found in the follow up experiment and the face detection rate is improved.

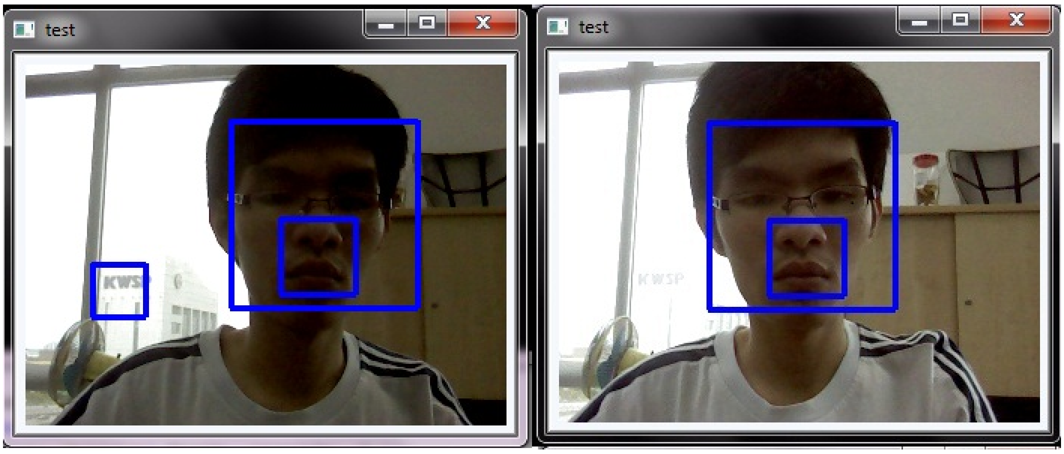


Figure 4-16 Poor face detection in video

The first problem was a major false positive detection. Viola-Jones face detection algorithm scanned through the whole image or frame using Haar Cascade and the image of the face was relayed back regardless of the size. This problem was solved by presetting the video frame or resolution to 320 X 240 and the minimum face detection or region of interest (ROI) of *100 X 100* pixels. With this scale, the distance between the subject and the camera had to be within 30cm to 60cm as shown in Figure 4-17. Therefore, small or unwanted objects could no longer be detected as it was out of the preset size.

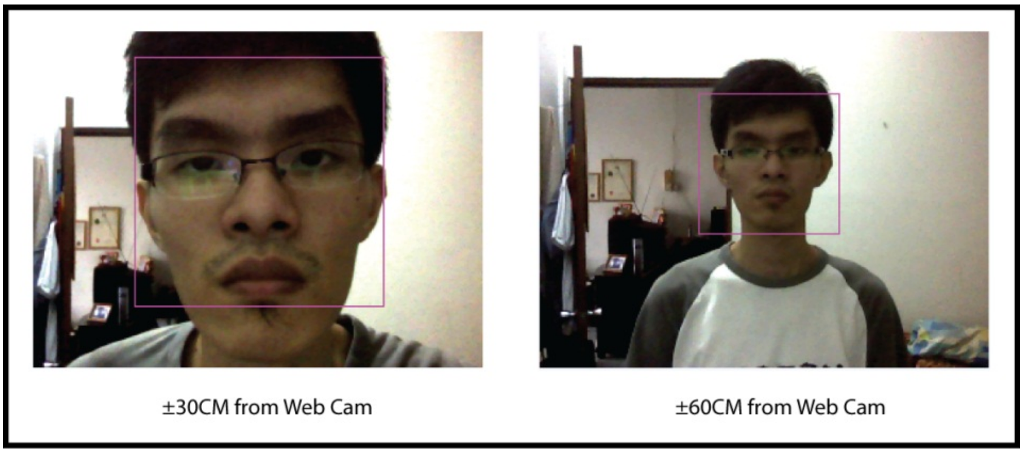


Figure 4-17 Distance between human face and webcam

The second problem of multiple face-likes detection was caused by the "Sequence Elements" obtained during face searching in a frame using Viola-Jones face detection algorithm. Multiple face-like objects were detected and the coordinate of detected object was stored in the Sequence Element list. Then box(es) were drawn on that frame based on the number of Sequence Elements found. To solve this problem, the prototype was forced to draw only the first Sequence Element found and to discard the rest of the other Sequence Elements during the implementation. Hence, only one box or ROI will be drawn on each frame and it solved the problem of multiple object detection.

Table 4-6 Face detection testing result

Environment		**Detection Rate (%)**
Indoor	Day Time	85
	Night Time	70
Outdoor	Day Time	85
	Night time	60
Overall Face Detection Rate		**75**

The prototype was then been tested under different environment and the overall face detection rate hit only 75% as shown in Table 4-6. Generally, the detection results obtained were not satisfactory as shown in Figure 4-18.

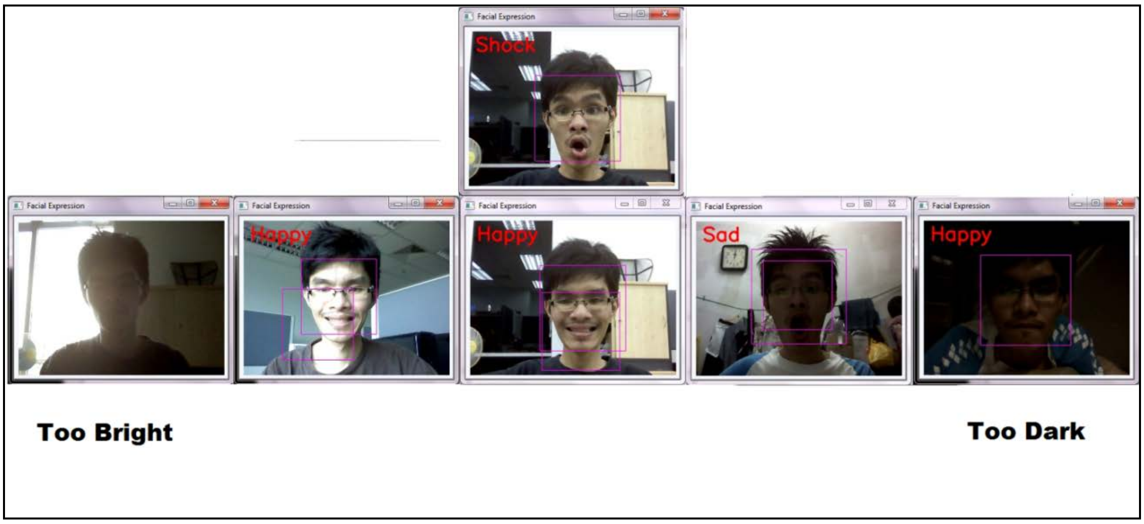


Figure 4-18 Illustration on face detection against frame brightness

The face detection rate was very much dependent on environment and lighting condition. This was a common problem for programs using the naive Viola-Jones face detection algorithm which had been discussed in pre-implementation and in the literature review section previously. This problem was unsolved in ACApp 1.0.

b. Experiment 1-2 Expression Recognition Ability

Another experiment was conducted once ACApp 1.0 implementation was completed. This experiment was conducted to test and to evaluate expression recognition in the prototype.

This experiment was participated by the developer. The developer trained the prototype with five different facial expressions or labels, with five samples for each label. One sample was stored out of ten face detected frames and the ACApp had to go through 50 face detected frames to complete training for one label. When five labels were trained, ACApp produced and stored 25 template images which would be used for template matching or recognition later.

The experiment continued with facial expression recognition which ran for four times under different environment including the research lab, café, outdoor, and bed room. The experimental results produced are shown in Table 4-7 below.

Table 4-7 Expression recognition result in Experiment 1-2

Environment	**Label 1**	**Label 2**	**Label 3**	**Label 4**	**Label 5**	**Recognition Rate**
Research Lab	Match	Match	Match	Match	Match	**100%**
Café	Match	Label 1	Match	Match	Match	**80%**
Outdoor	Label 3	Match	Match	Label 3	Match	**60%**
Bed Room	Label 3	Label 1	Match	Match	Match	**60%**
Recognition Rate	**50%**	**50%**	**100%**	**75%**	**100%**	**75%**

Based on the results obtained, ACApp only achieved 75% of overall recognition rate with one set of training images (5 labels with 5 samples per label) and this was performed under four different environmental conditions. A few labels were wrongly recognized in the experiment i.e. label 1 and label 2.

Based on observation during the matching process, there were two possibilities that caused the misrecognition and poor recognition results in Experiment 1-2. 1) Expression training was not well done resulting in incomplete templates. 2) Miscalculation in the template matching process.

4.2.1.3 Findings

Generally, there were two major problems found while developing ACApp 1.0. The first problem was related to face detection which was sensitivity to varying degree of lighting and environmental condition. This was considered a common problem faced by developers who adopt the naïve Viola-Jones face detection algorithm to perform a face detection task. Other face detection limitations were related to multiple object detection and non-face object detection but these problems were solved.

The second problem occurred was related to ACApp's facial expression recognition ability. The result or recognition rate produced from the experiment was 75%. The recognition rate was expected to exceed 85% or above by training only five labels and under controlled setting and environment. The causes leading to a lower recognition rate was identified in experiment 1-2.

Up to this stage, the limitations were yet unsolved and are being brought forward to the next developmental stage to improve ACApp 1.0.

4.2.2 ACApp V2.0

The main objective of ACApp 2.0 was to solve problems discovered in ACApp 1.0. Additional improvement was done at this stage and minor changes on settings and configuration on ACApp was made. The outcomes of ACApp 2.0 were used for further experiments to find out the readiness of ACApp for field testing on targeted users.

4.2.2.1 Enhancements

The changes, and improvements made in ACApp 2.0 included solving the false positive detection problem which was identified in ACApp 1.0, improving expression recognition rate and ACApp's structure.

a. Solving False Positive and Face Detection Problem

Previously, ACApp 1.0 obtained poor face detection rate when dealing with inconsistent lighting condition or environment. This problem could be solved by adding *Automated Brightness Adjustment*.

The *Automated Brightness Adjustment* process takes the initial frame, checks its average brightness, and adjusts the frame's brightness to a controlled level of brightness where the face can be detected better. This process was done by calculating the initial frame's average brightness value or the optimum lighting condition using HSV model (Bezryadin, Bourov and Ilinih 2007) then standardized the frame by adjusting the average brightness to the optimum lighting condition value or the control value.

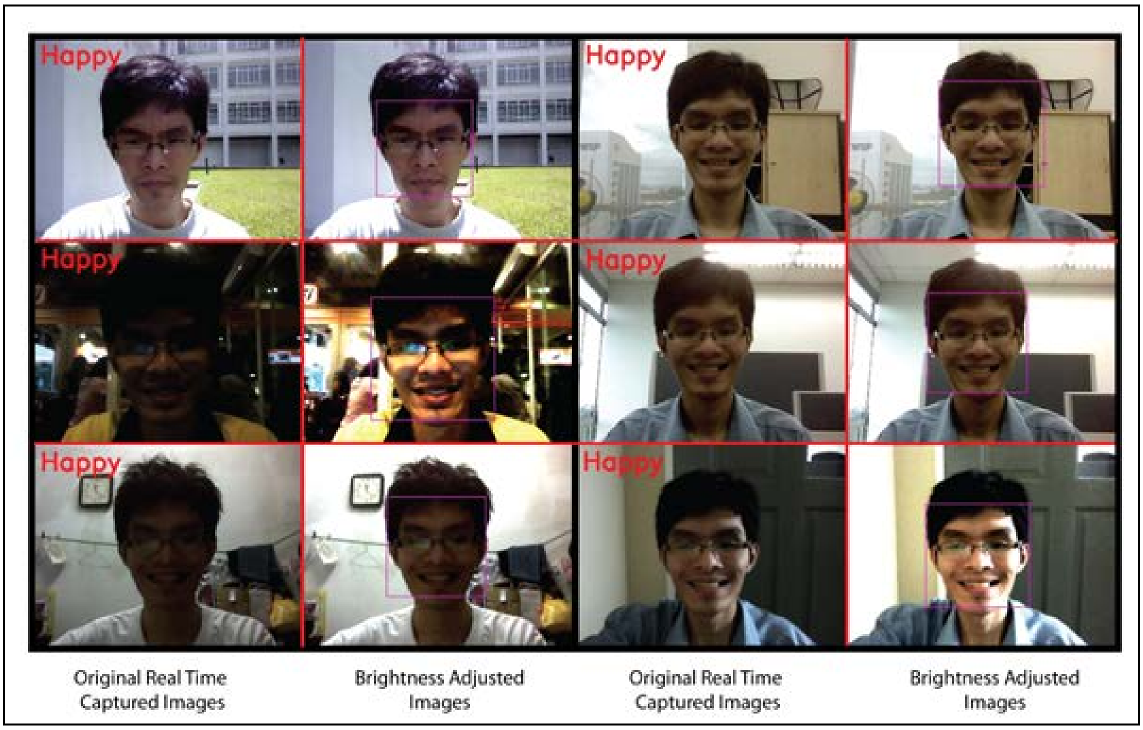


Figure 4-19 Face detection under various of lighting environment and condition

Throughout the mini experiment as shown in Figure 4-19, the optimum lighting condition value was defined and fixed at 160 where the brightness adjusted frame was not over or under exposed. The algorithm of *Automated Brightness Adjustment* is illustrated Figure 4-20.

```
INITIALIZE control value to 160
INITIALIZE total values to 0
INITIALIZE total pixels to 0

FOR each frame height
      FOR each frame width
            GET max pixel value from RGB
            ADD total value with pixel values
            INCREASE total pixels by 1
      END FOR
END FOR

GET average value by dividing total values with total pixels
GET scale by dividing average value with control value

FOR each frame height
      FOR each frame width
            SET each RGB pixel with initial value divides with
            scale value
      END FOR
END FOR
```

Figure 4-20 Algorithm of Automated Brightness Adjustment

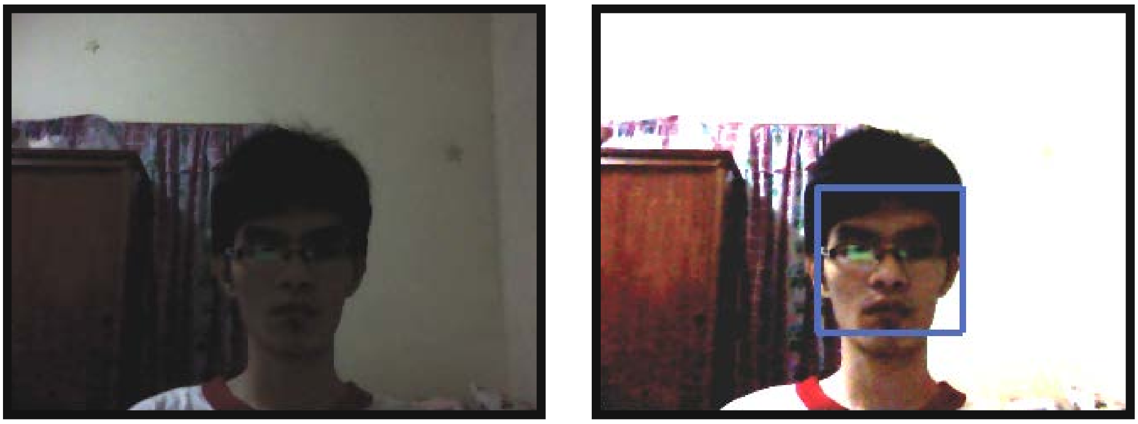


Figure 4-21 Before and after automated brightness adjustment process

Figure 4-21 shows the result after brightness adjustment, based on the control value or the optimum lighting condition value. The image on the left is the original frame captured in real time using a web cam while the image on the right is the brightness adjusted frame. ACApp 2.0 managed to detect human face under various lighting conditions by implementing *Automated Brightness Adjustment*.

b. Improving Expression Recognition Rate

As mentioned in the implementation of ACApp 1.0 previously, the possibility which caused ACApp to produce low expression recognition rate included the fact that the training sessions were conducted inappropriately and also a miscalculation in the *TemplateMatching* process.

After conducting experiments and observations on this matter, it proved that inappropriate training session produces poor results when performing expression recognition. For instance, when performing "neutral" expression training with ACApp and the user unintentionally shows a smiling face, this was actually conveying to ACApp that the "neutral" expression might look like a smiling face too. Consequently, ACApp may recognize a smiling face as a neutral face when performing real time expression recognition in future. Hence, this problem can only be solved by running an expression training carefully.

An experiment (Experiment 2-2) was conducted to find the most optimum number of samples per expression by testing expression recognition with several samples for each expression. The experiment result showed that the most optimum number of samples kept was 20. The experiment also indicated that training an expression with longer duration increases the recognition rate. Training an expression with 5 samples produced a 60% of recognition rate while training an expression with 20 samples increased the recognition rate to 92%.

c. Restructuring Process

Several minor changes had been made in restructuring ACApp 2.0 from ACApp 1.0. These changes included adding *LoadTemplate* process into *Initialization* process, adding *Automated Brightness Adjustment* process after acquiring frame from *Capture* process and adding *Timer* for experimental purposes.

The first change made on ACApp 2.0 was the code cleaning and restructuring. The source code was modulated into sub-functions so that any bugs occurred could be easily tackled.

Apart from modulating the source code, the *LoadTemplate* process was shifted and placed inside the *Initialization* process. All required variables were initialized and all required materials were loaded before expression training and expression recognition processes started. This reduced the time consumption for the expression training and the expression recognition sessions. It also eliminated the delay of template loading in the expression recognition mode.

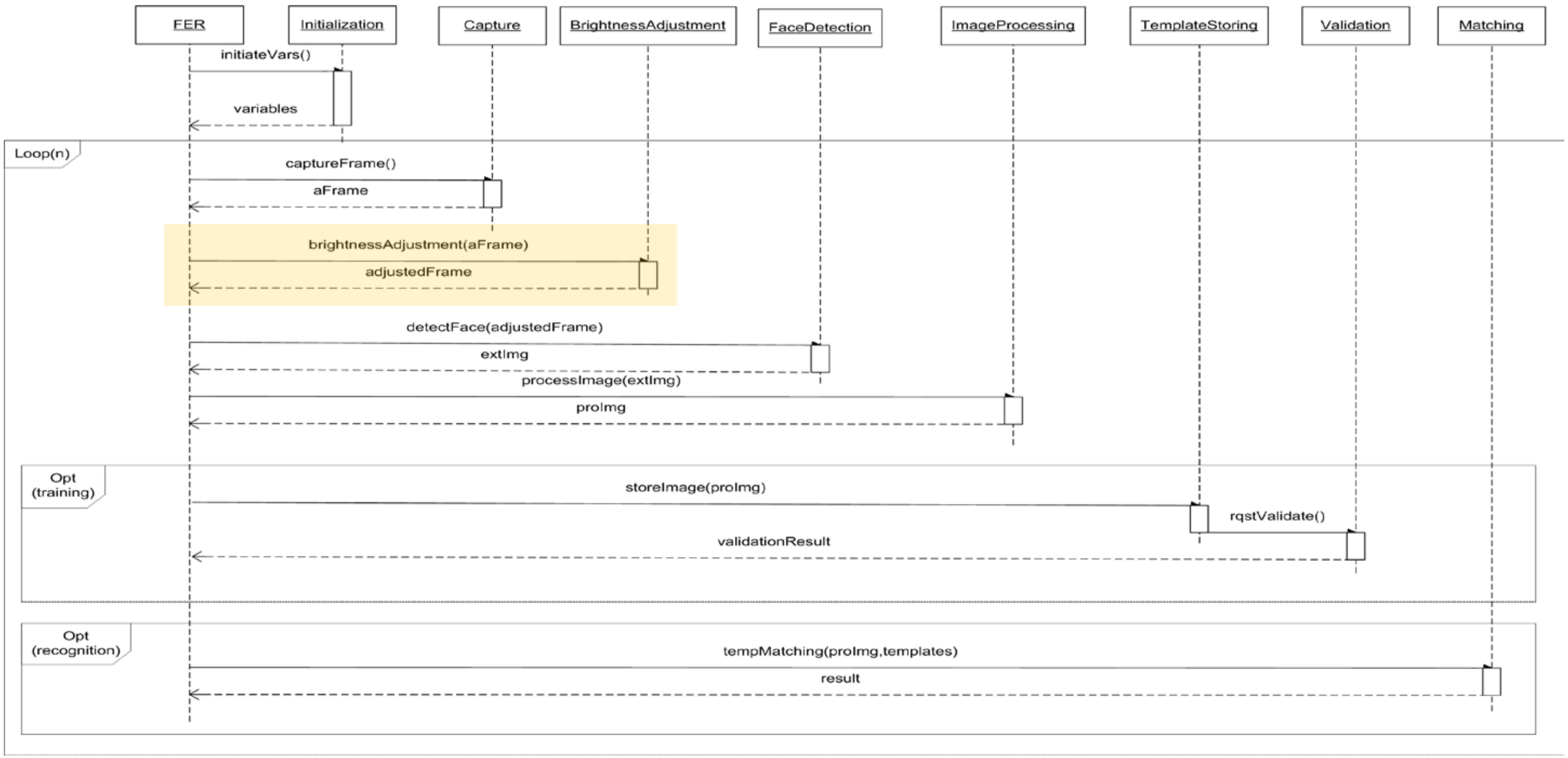


Figure 4-22 Sequence Diagram for ACApp 2.0 with automated brightness adjustment

The *Automated Brightness Adjustment* process was added into ACApp 2.0; this process was located after *FER* obtained a frame from *Capture* process which is shown in Figure 4-22. *FER* sends the frame received to *Automated Brightness Adjustment* process to adjust the frame's brightness.

Another change made in ACApp 2.0 was the addition of *Timer* for testing and evaluation purposes. This timer was added before a frame was captured to calculate the time needed to train an expression as well as the time needed to recognize the expression in real time.

4.2.2.2 Experiments

Three experiments were conducted in this stage i.e. the face detection with the *Automated Brightness Adjustment* process, facial expression recognition test with different quantity of template images stored per expression and also pilot testing to evaluate the improvement of ACApp 2.0.

a. Experiment 2-1 Face Detection with Automated Brightness Adjustment

An experiment on the *Automated Brightness Adjustment* process was conducted under the same lighting condition and environment as in Experiment 1-1. These environmental conditions included indoor and outdoor both during day and night time. 20 frames were captured by the *FaceDetection* process with the *Automated Brightness Adjustment* for each environment. The detection rate was recorded as shown in Table 4-8

Table 4-8 Result of face detection with Automated Brightness Adjustment

Environment		**Detection Rate (%)**
Indoor	Day Time	100
	Night Time	95
Outdoor	Day Time	95
	Night time	90
Overall Face Detection Rate		**95**

The previous ACApp had problems detecting the human face under varied environment and obtained only 75% detection rate which was believed to be caused by inconsistent lighting source received. The ability to detect the human face under varied environment was improved after implementing *Automated Brightness Adjustment* into ACApp 2.0. An overall result of 95% face detection rate was produced. The detection rate was 100% during day time in an indoor environment. 95% of the face was detected at night time in an indoor environment. For day time at an outdoor environment, the detection rate was 95%. Whereby, 90% of detection rate was achieved at night time in an outdoor environment.

b. Experiment 2-2 Enhancing Facial Expression Recognition

An experiment on improving facial expression recognition rate was conducted by running the recognition process with different numbers of samples for each expression. The purpose of this experiment was to find out the optimum sample size for each expression to optimize the recognition rate in ACApp. This experiment was conducted in a research lab with controlled lighting condition.

The experiment was conducted six times where the number of samples for each expression was set to 5, 10, 15, 20, 25, and 30. For each set of samples, one training session was conducted and went through five rounds of expression recognition. The matching result returned by ACApp was marked as match if it recognized the test image and mismatched if the matching result returned by ACApp did not match with the test image. The experiment results obtained are listed at Table 4-9.

Table 4-9 Expression recognition result

Number of Sample	Exp. 1 (%)	Exp. 2 (%)	Exp. 3 (%)	Exp. 4 (%)	Exp. 5 (%)	Overall Recognition Rate
5	80	40	60	80	40	**60%**
10	100	60	80	60	80	**76%**
15	100	80	60	100	80	**84%**
20	100	100	80	80	100	**92%**
25	100	100	100	80	80	**92%**
30	100	80	60	80	100	**84%**

The result obtained by training five samples per label had the lowest recognition rate among the six sets. Almost all of the expressions could be recognized however it was not consistent. For example, in expression matching with 5 samples training for each expression, expression 1 and 4 were matched four times, expression 2 and 5 were matched twice while expression 3 was matched three times. Overall, ACApp 2.0 achieved 60% of recognition rate with 5 training samples for each expression.

The same experiment method was applied to set 2 – 6 where ACApp managed to increase the overall recognition rate and hit 76% with 10 training samples for each expression. The results obtained improved greatly with the 15 and 20 training samples per expression, increasing to 84% and 92% respectively. The overall recognition rate remained at 92% with 25 training samples for each expression and the result dropped to 84% at sample set number 6 which trains 30 samples per expression.

Since by training ACApp with 20 and 25 training samples for each expression could produce the highest overall recognition rate which is 92%, ACApp 2.0 was fixed to train 20 samples for each expression because it consumed less time to train 20 samples compared with 25 samples.

c. Experiment 2-3 Pilot Testing with Normal People

An experiment on overall performance was conducted with five participants who were normal human in the research laboratory. These participants were instructed to act and express 7 different expressions during the training and expression recognition sessions. Each expression was tagged with a label. The expected outcome in this experiment includes ACApp recognition rate and the time consumption for these two processes.

The experiment was performed indoor during day time with a clear background where the lighting condition value was set to 160. 20 training templates were taken from 200 frames captured at the interval of 10 frames for each expression. Each participant went through one round of expression training and three rounds of expression recognition.

The participant's expression was marked as match if ACApp managed to recognize or return the correct label expressed by them at least for two times. A mismatch was marked to the participant's expression if ACApp returned incorrect label expressed by them for two rounds.

Table 4-10 Expression recognition result in Experiment 2-3

Expression / Participant	Par. 1	Par. 2	Par. 3	Par. 4	Par. 5	Recognition Rate
Expression 1	Match	Match	Match	Match	Match	100%
Expression2	Match	Match	Match	Match	Match	100%
Expression 3	Mismatch	Match	Match	Match	Match	80%
Expression 4	Match	Match	Match	Match	Match	100%
Expression 5	Mismatch	Match	Mismatch	Match	Match	60%
Expression6	Match	Match	Match	Match	Match	100%
Expression 7	Match	Match	Match	Match	Match	100%
Total	**71.4%**	**100%**	**85.7%**	**100%**	**100%**	**91.42%**

Table 4-10 shows the recognition result generated by ACApp 2.0 by testing five participants' expressions. The overall recognition rate of 91.42% was achieved which was better than the Experiment 1-2 conducted on ACApp 1.0 (75%) previously. This was due to the improvement made on the *FaceDetection* process and the *TemplateMatchin*g process.

In the face detection process, the *Automated Brightness Adjustment* helped ACApp to capture and the extract human face precisely; and with the proper human face extraction, no incomplete or void template was stored. In the facial expression recognition process, the incremental number of samples stored for each expression did help in increasing the recognition rate as ACApp had more choices and was able to get the closest template during the matching session.

Consequently, the overall expression recognition rate had increased in ACApp 2.0, however, the time consumed to perform both face detection and expression recognition had also increased. Table 4-11 shows the difference of time consumption for both the expression training and the expression recognition without the *Automated Brightness Adjustment* processes in conjunction with the difference in the number of sample used for each expression. For ACApp 1.0 that stored 10 samples on each expression, an overall of 11.95 seconds was needed to train one expression while an overall of 26.5 milliseconds was required to match one image with all the templates available. For ACApp 2.0 that stored 20 samples on each expression, it required an overall of 22.56 seconds to complete an expression training session and 50 milliseconds was required on average to match one image with all trained templates.

Table 4-11 Time consumption for training and recognition

Participant	20 Samples (ACApp 2.0)		10 Samples (ACApp 1.0)	
	Avg. Train Time per Expression (Seconds)	Avg. Test Time (Milliseconds)	Avg. Train Time per Expression (Seconds)	Avg. Test Time (Milliseconds)
1	22.28	49	12.78	27.5
2	22.18	51.5	12.35	29.5
3	22.47	51.5	11.13	27
4	22.88	50.5	11.87	25.5
5	23.00	47.5	11.62	26.5
Average	**22.56**	**50**	**11.95**	**27.2**

Obviously, when the number of samples stored increased, the time consumption also increased for both face detection and expression recognition process.

This experiment has proven that the recognition rate of ACApp 2.0 had increased when an *Automated Brightness Adjustment* was added in the *FaceDetection* process and following that there would be an increment in templates kept for each expression for *TemplateMatching* process. However, time consumption also increased due to the size of the samples.

4.2.2.3 Findings

Throughout the development and experiments for ACApp 2.0, the research managed to solve problems faced by ACApp 1.0 which included sensitivity to lighting condition for the *FaceDetection* process and the low recognition rate produced by ACApp 1.0.

The face detection problem was solved and improved from 75% to 95% by implementing the *Automated Brightness Adjustment* process into *FER*. While the recognition rate was also improved from 75.00% to 91.42% under controlled setting by fixing the number of template images stored for each expression to 20.

The ACApp 2.0 was not ready for real testing based on the feedbacks received from participants. One of the problems raised by the testers was related to the time required to train one label or expression which takes up to ±20 seconds is too long. Besides, the console based ACApp 2.0 was not user friendly at that moment.

At this stage, Module 2 (Facial Expression Training and Recognition) is considered complete with minor changes needed in the later version.

4.2.3 ACApp V3.0

The objective of ACApp 3.0 was to add in the Profile Registering module and the Alert Notification module to form the complete set of proposed ACApp as shown in Figure 4-23. The main focus at this stage was to implement the Profile Registering and the Alert Notification module, wrap ACApp with graphic user interface by using .NET Framework 3.5 as well as solving the problem encountered in ACApp 2.0.

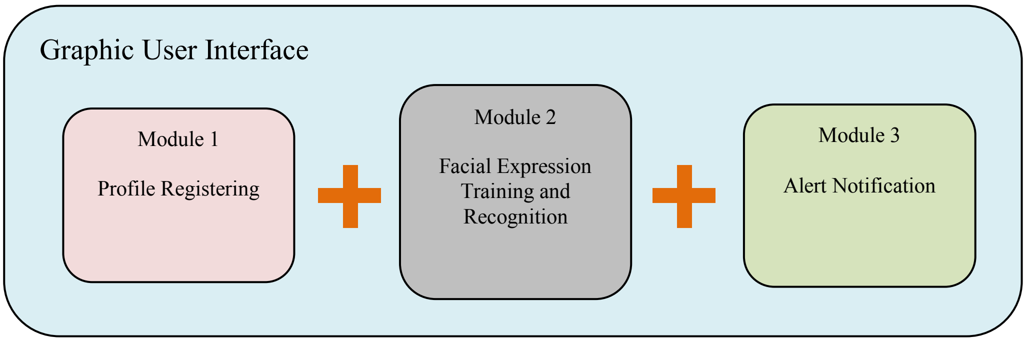


Figure 4-23 ACApp 3.0's components

4.2.3.1 Implementation

The change made on ACApp 3.0 started by migrating ACApp 2.0 from a "Console Application" into a "Windows Form Application".

In the initial stage for ACApp 3.0 development, the transformation or migration from ACApp 2.0 to ACApp 3.0 was done without adding additional modules. Only expression training and expression recognition for one user was available at that stage. The implementation of other modules i.e. Profile Registering, Alert Notification and other additional features were added after the migration works had completed.

a. ACApp Migration Works

Figure 4-24 shows the graphic user interface for the expression training mode in ACApp 3.0 after the migration works in the early stage. The user or specifically the guardian was allowed to train expression or remove trained expression from this interface under "Expression Training" tab (window A).

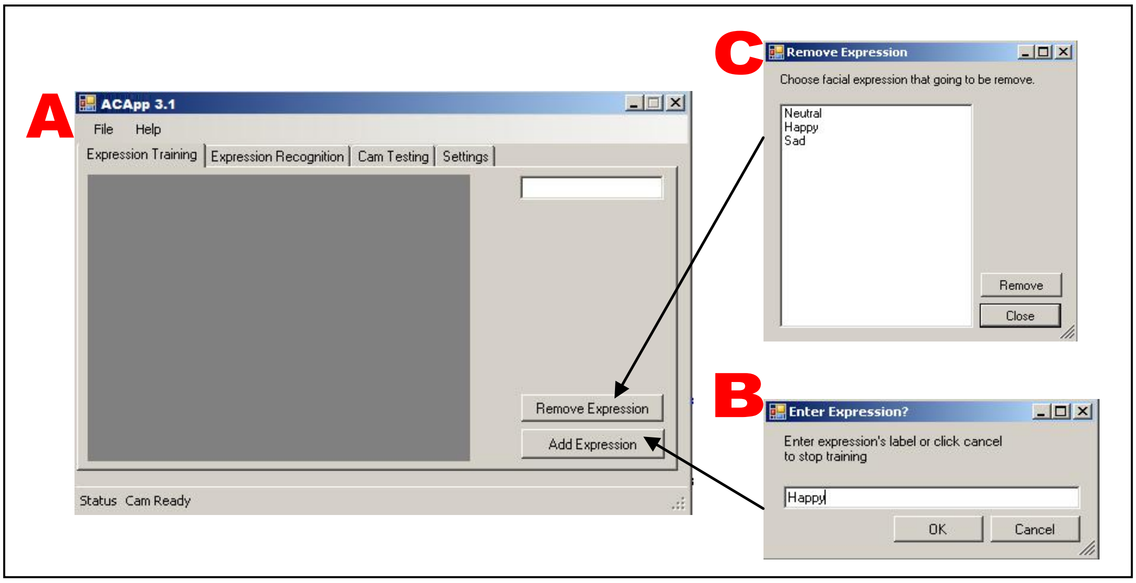


Figure 4-24 ACApp 3.0 GUI in the early stage – Expression Training Interface

When the user clicks the "Add Expression" button, a small window (B) will appear and the user to enter the label for the expression to be trained and followed by clicking on the "OK" button. The user is also allowed to remove existing or trained expression by clicking the "Remove Expression" button from interface (A) and a small window (C) will pop up with a list of trained expressions. The user can select multiple expressions that to be removed in the window (C).

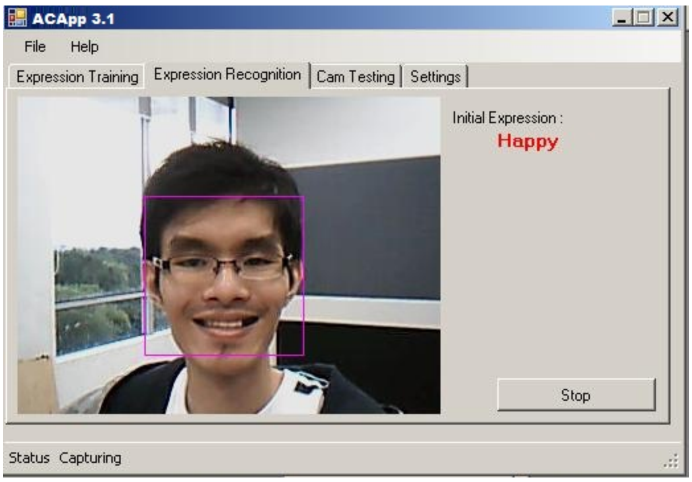


Figure 4-25 ACApp 3.1 GUI – Expression Recognition Interface

On the other hand, expression recognition can be performed by going to "Expression Recognition" tab in ACApp 3.0 interface (Figure 4-25). As usual, ACApp will be able to run expression recognition after it has been trained.

b. Module 1 - Profile Registering

The Profile Registering module was added based on the requirement gathered from section3.4.3. Guardians will be able to create the profile for each child with Cerebral Palsy and train individual expression(s).

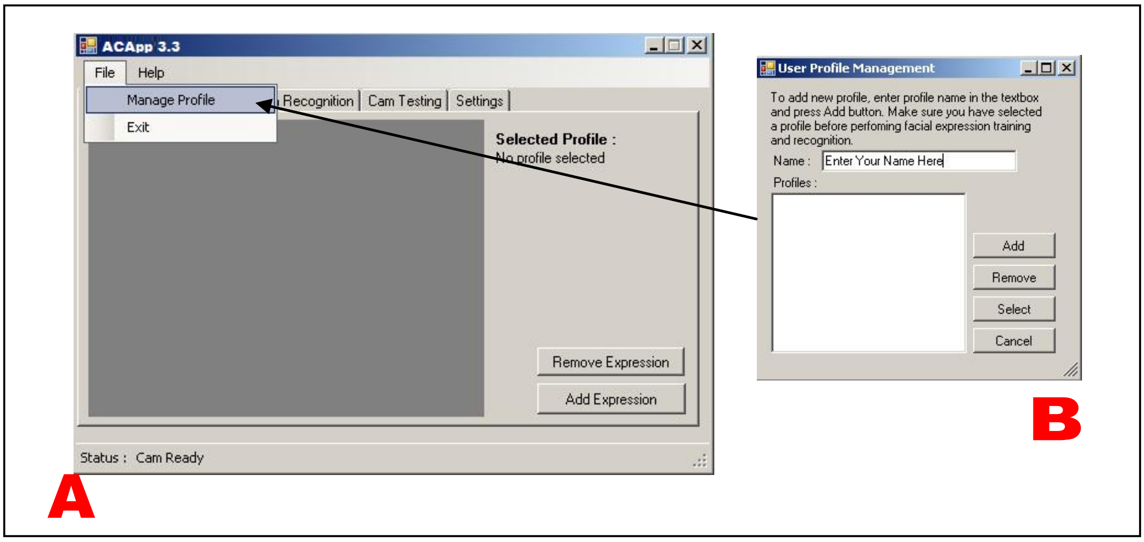


Figure 4-26 Prototype version 3 GUI – Profile Registering

Figure 4-26 shows the screenshot of profile management options. The user can manage the profiles by opening the User Profile Management window (B) then add, remove or select a profile.

Figure 4-27 shows the flow of the *ProfileRegistering* processes. When ACApp is executed, the user can select the "Profile Management" option by calling the *UserManagement* class. This class will request a list of user profile from database and display the available user profiles. Generally, the user can perform three commands; 1) Create Profile, 2) Remove Profile and 3) Select Profile.

A profile can be created and removed by sending *userAction(command)* to *AlterProfile* class. This class receives the user's command and performs changes, then requests data storage to be updated. Once the data storage has been updated, a new list of user profiles will be returned to the *UserManagement* class and it will display the new list.

If a user selects a profile and confirmed the selection, the *UserManagement* class will return the selected profile back to *FER* for further processing.

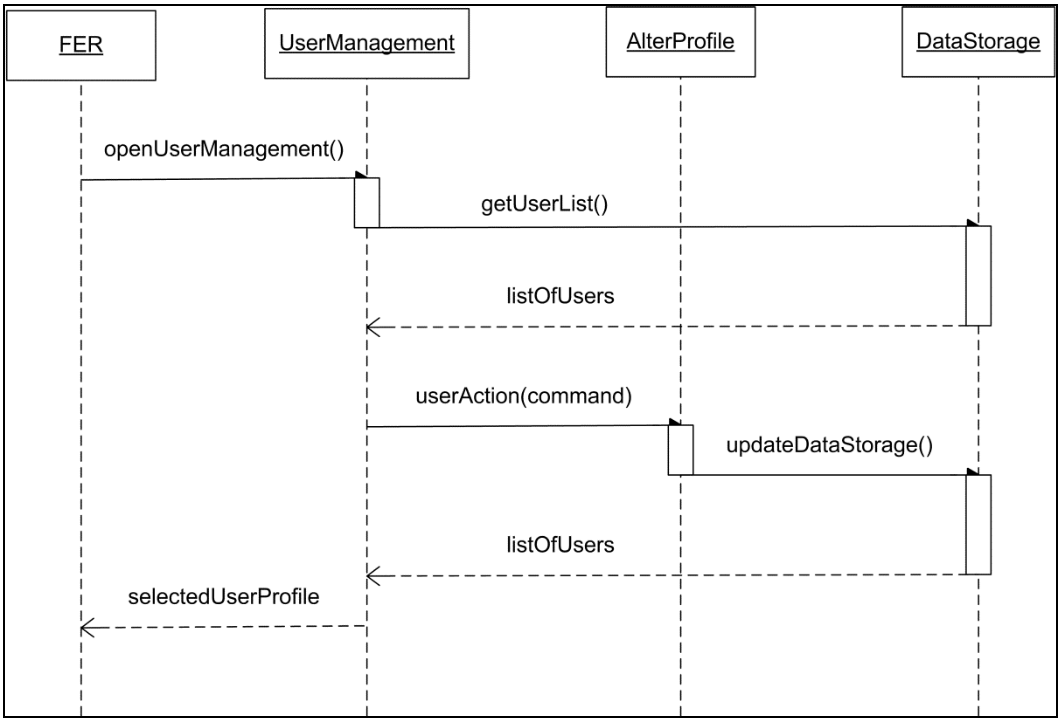


Figure 4-27 Sequence Diagram for creating, edits and deletes user profile process

A new folder with the profile name is created in data storage when a new user profile is added into ACApp. When performing expression training, the selected user profile templates will be stored into this folder and these templates will be used for the *Matching* process later. All templates and related documents within the profile folder will be deleted from data storage when a user profile is removed upon request.

c. Module 3 – Alert Notification

As mention in section 3.4.2, Alert Notification module was part of the requirement to notify the guardian if children with Cerebral Palsy showed critical or unusual expression during the monitoring session.

When a child's expressions are trained according to the selected user profile, the guardian can choose the way he/she wishes to be notified. The selections can be performed by clicking the "Expression Notification" button which is located under the "Expression Recognition" tab as shown in Figure 4-28 (A). When the "Expression Notification" button is pressed, a window (B) will appear which will allow the user to select or mark expressions to be notified.

By default only a sound alarm will be produced through a built in speaker when these marked expressions are detected during the monitoring session. The guardian can "tick" the checkbox with the label "Include Short Message System Notification (SMS)" and enter his or her contact number in Figure 4-28 (C) he/she wishes to be notified via SMS in his/her mobile phone.

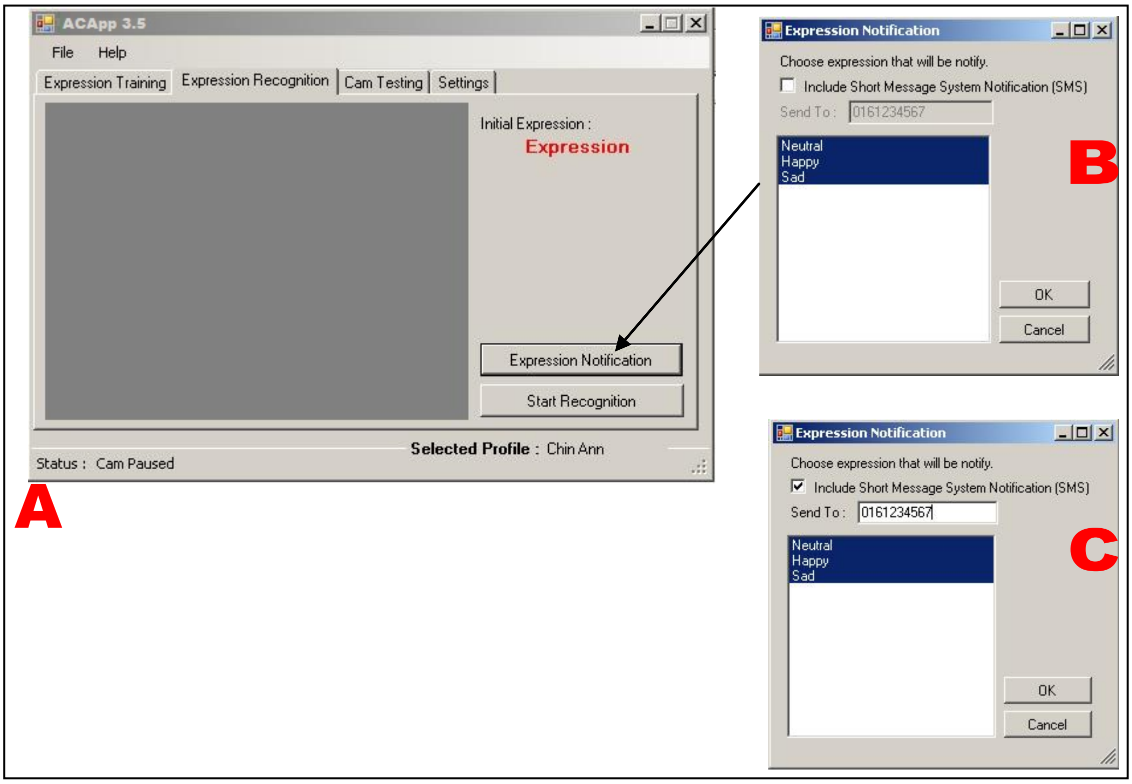


Figure 4-28 Prototype version 3 GUI – Alert Notification

Figure 4-29 shows the sequence of Alert Notification during the monitoring session. The diagram is divided into two parts, the top part represents the pre monitoring session which allows the guardian to mark or unmark expressions to be notified while the bottom part represents the process during the monitoring session.

Before the monitoring session starts, the guardian is allowed to mark specific expressions to be notified. By calling the *openAlertNotificationManagement(user)* a list of the selected profile's expression from data storage will be displayed upon request. The *AlertNotificationManagement* class will update data storage when the guardian completes the marked and unmarked expressions that are to be notified.

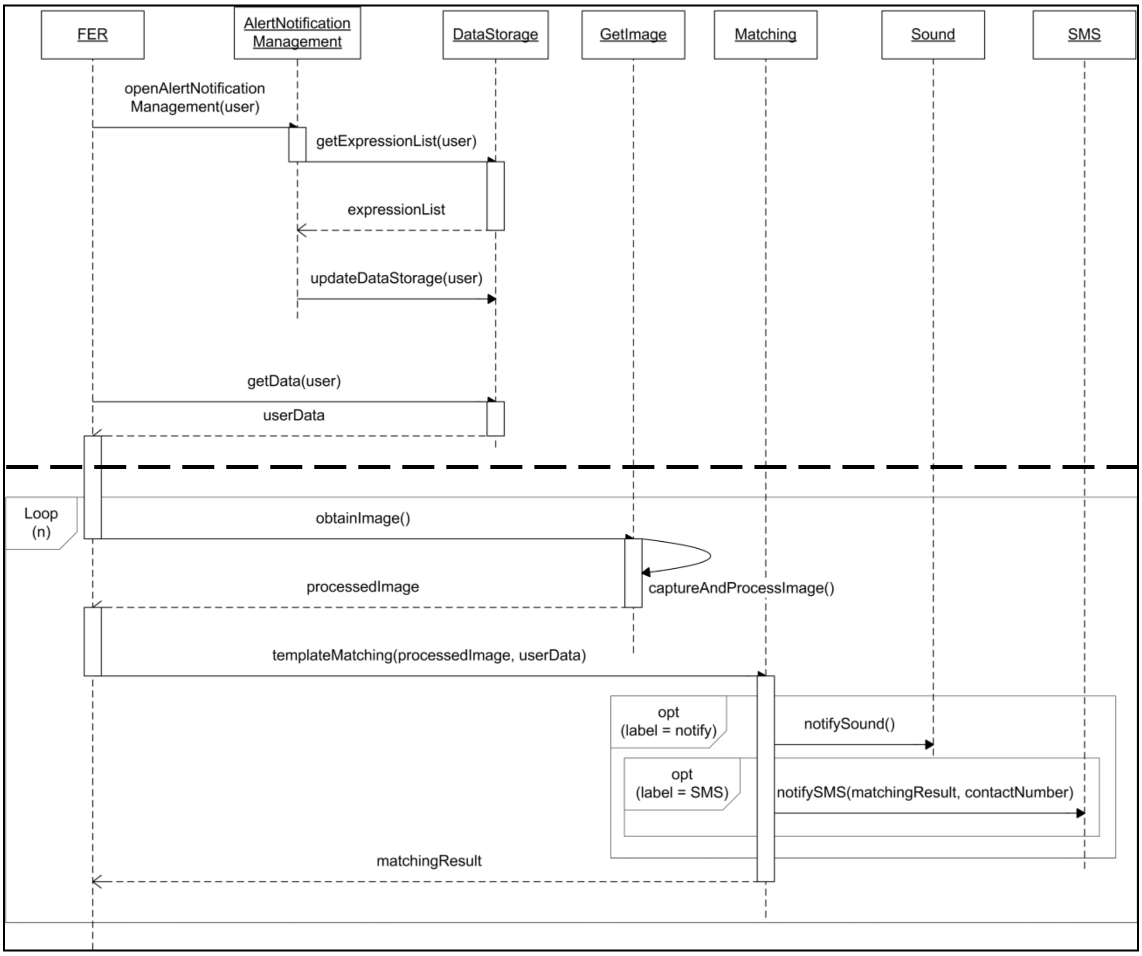


Figure 4-29 Sequence Diagram for Alert Notification process

When the monitoring session starts, ACApp will request for all the required data, based on the selected user profile and will initialize required variables. ACApp will then go into an expression recognition loop which will only terminate upon the guardian request.

Within the loop, ACApp will get an image which is the child's initial face image. This process is simplified by calling *obtainImage()* which is a combination of various processes i.e. *Capture*, *FaceDetection*, and also *ImageProcessing*. At the end of the *GetImage* process, ACApp shall be able to obtain a *processedImage* and send both *processedImage* and *userData* to the *Matching* class by calling *tempMatching(processedImage, userData)*.

The result of the matching will be compared with *userData* to ascertain whether notification is required or not. If notification is required, the *Matching* class will call *notifySound()* and instruct the *Sound* class to produce an alarm. If notification via SMS is required, the *Matching* class will call *notifySMS(matchingResult, contactNumber)*and instruct *SMS* class to send an SMS. After that, the *Matching* class will send the *matchingResult* back to *FER* for further processing.

d. Pre-Time Expression Training Mode

According to participants' feedback when performing Experiment 2-3 in ACApp 2.0, it was too long and tiring to train an expression in 20 seconds. To overcome this problem, a pre-time expression training mode was implemented in ACApp 3.0.

This is an additional concept added into ACApp 3.0 which will make it more efficient. Sometimes, it is hard to capture a human expression spontaneously and train ACApp in real time. Hence, the user can use external capturing devices i.e. cameras, cam coders, web cams, PDAs and even mobile phones to capture the expression of children with Cerebral Palsy and to save them as a video format file. The video format file will be fed into ACApp and will run the usual expression training.

The current ACApp 3.0 can only recognize the Audio Video Interleave (AVI) video format. However, the video recorder usually captures the file in other compressed format i.e. Windows Media Video (WMV), Moving Picture Expert Group (MPEG, MPEG-2, MP4), Third Generation Partnership Project (3GP), and QuickTime Movie (MOV). To overcome this problem, an automated video conversion was required before importing the video file into ACApp.

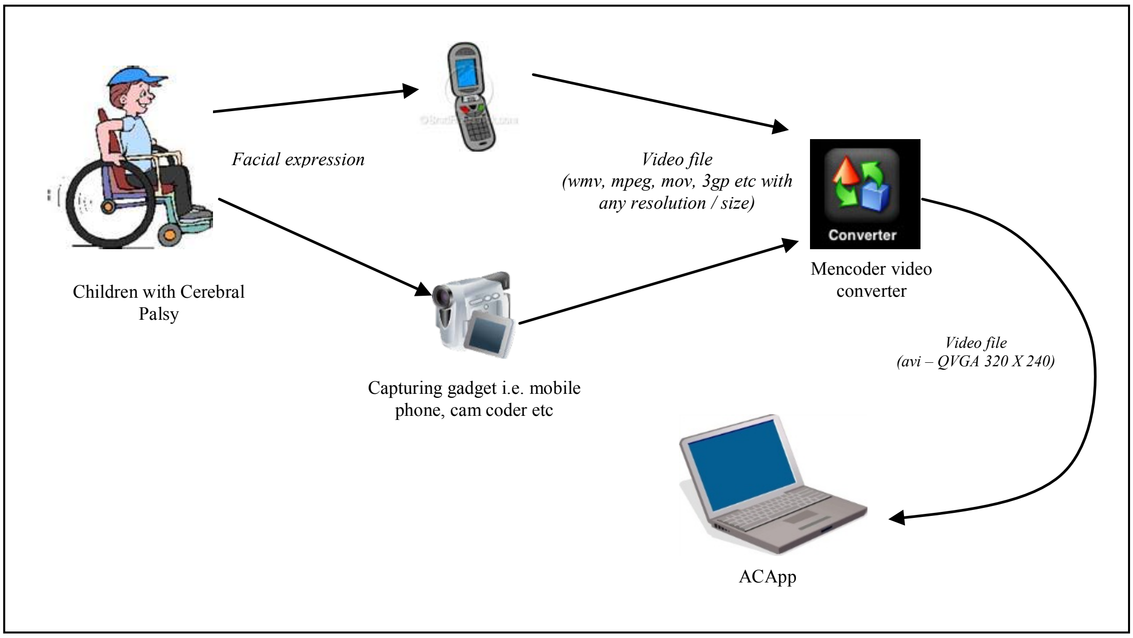


Figure 4-30 Illustration on video format conversion before feed into ACApp for expression training

The Mencoder (2010) was adopted as an additional API to perform video format conversion into AVI. Figure 4-30 illustrates how the expression training was done. Guardians can capture the facial expression of children with Cerebral Palsy at anytime with their mobile phone or other video recording devices. Then, the obtained video file can import into ACApp by going through a video converter - Mencoder.

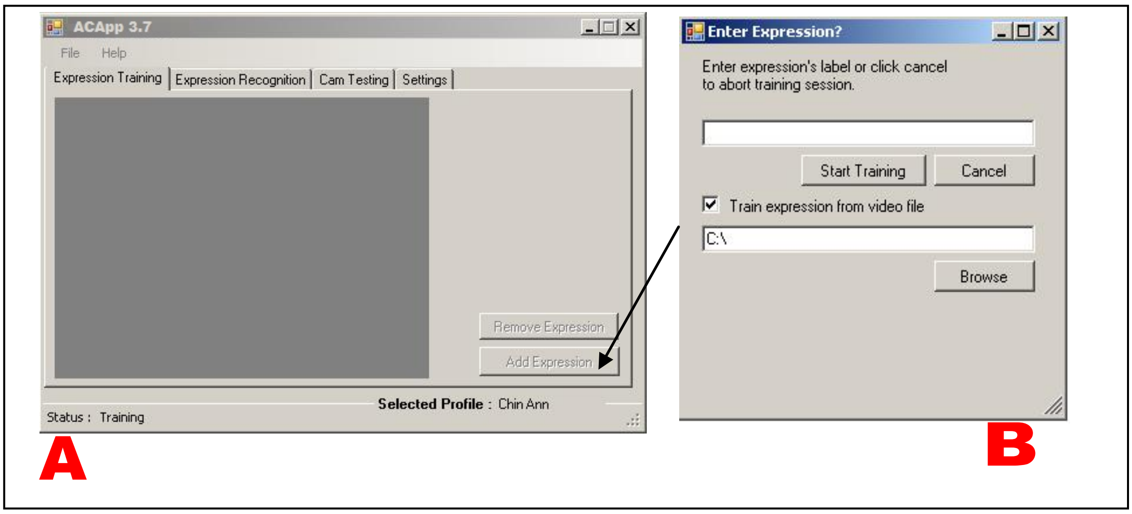


Figure 4-31 Prototype version 3 GUI – Pre-time expression training

The automated video conversion was added in ACApp 3.0 to fulfill the requirement and is accessible through the additional option in the “Add Expression” window (B) as shown in Figure 4-31. When adding a new expression, the guardian is required to enter the expression label in the first text box in window B. The captured video file can be inserted by checking the “Train expression from video file” checkbox and click on the “Browse” button to insert the video file. When the video file path is entered and the “Start Training” button is pressed, ACApp will check the file format. If the file is the AVI format, the expression training process will start immediately. However, if the file is not the AVI format, ACApp will send the file to Mencoder for video conversion. The output file (AVI format) will then be used for expression training.

e. Improving Face Detection Rate

One major improvement and two amendments were made when to improve face detection in ACApp. These improvement and amendments included 1) rotated face detection - to detect the human face, 2) replacing the “default Haar Cascade” to “alt_tree Haar Cascade” and 3) resizing the region of interest.

Previous observation has found that children with Cerebral Palsy have difficulty in controlling their head position. It was impossible for ACApp to detect faces when the subject’s head was tilted.

The naïve Viola-Jones face detection algorithm only managed to perform detection on a frontal face at a radius of less than 15 degrees, in-plane rotation tilted to the left or right and 45 degrees out of plane rotation toward profile side views. Hence the detection rates dropped in such a situation. Figure 4-32 shows a child with Cerebral Palsy with a tilted head position. To improve the naïve Viola-Jones face detection algorithm for tilted heads, the *FrameRotation* process was added.

Figure 4-32 Children with Cerebral Palsy with tilted head positions

Figure 4-33 illustrates the face detection with *FrameRotation*. A frame acquired by *Capture* class (with or without *Automated Brightness Adjustment*) will go through the face detection process immediately for the first face searching.

If the face is detected, the next process i.e. *ImageProcessing* will take place. However if the face is not found, this frame will be rotated at 30 degrees to the right and will go through the second *FaceDetection* process. If a human face is still not found, the frame will be rotated at 30 degrees again to the right and perform the third *FaceDetection* process. If unsuccessful again, the frame will be rotated for another 30 degrees to the right, then perform the fourth *FaceDetection* process.

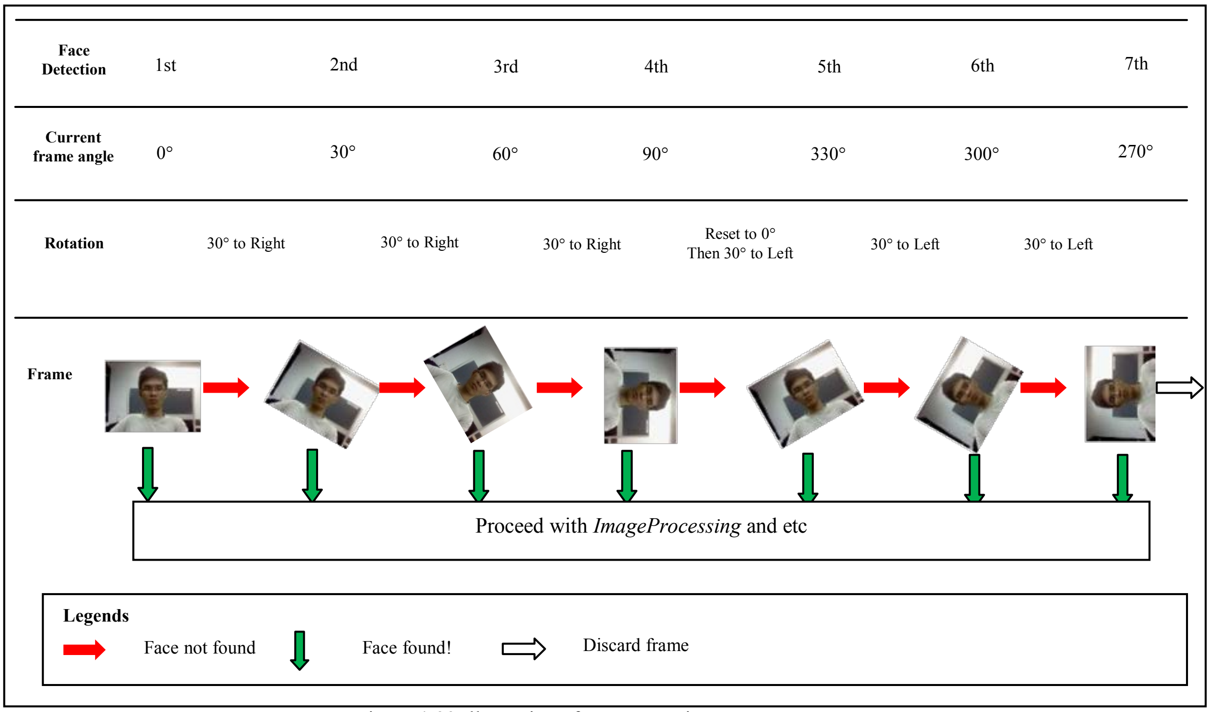


Figure 4-33 Illustration of FrameRotation process

If all else fouls at this stage, the rotation angle will be reset to 0 degree. Then a 30 degree rotation to the left will be performed for the fifth *FaceDetection* process. The frame will keep on rotating at 30 degrees to the left if the face is not found until the rotation angle reaches 270 degrees. If it fouls again, the frame will be discarded and another frame from *Capture* class will be obtained.

In total, a frame will be going through six rotations (three 30 degree to the right and three 30 degrees to the left) and seven *FaceDetection* processes before a frame is discarded if human face is not found.

Experiment 3-1 has shown significant improvements on face detection against a tilted head by implementing *FrameRotation* into ACApp. The detection rate increased from 36.4% without *FrameRotation* to 91% with *FrameRotation*. This was where the major improvement on ACApp 3.0 was made on the *FaceDetection* process. Apart from this improvement, two other amendments were also made to further improve the *FaceDetection* process.

The first amendment made was replacing the "default FrontalFace Haar Cascade" to "Alt Tree FrontalFace Haar Cascade". Experiment 3-2 has shown the 96.92% out of the total faces detected (74.79%) was true positive detection frame when processing by using the "Alt Tree FrontalFace Haar Cascade" compared with the 3 other Haar Cascades whose true positive detection rates were 64.06%, 78.40% and 50.33% for "FrontalFace Alt", "FrontalFace Alt2" and "FrontalFace Default" respectively.

The second amendment made was to resize the region of interest (ROI) from *100 X 100* pixels to *75 X 75* pixels while performing the face detection. Experiment 3-3 showed that improvement on true positive detection rate had increased after resizing ROI. The rate had increased from 73.0% to 93.34% throughout all total positive detection. This was another improvement achieved in the *FaceDetection* process.

4.2.3.2 Experiments

Five experiments were conducted to evaluate the effectiveness and efficiency of the current version of ACApp. These experiments included rotated face detection, different frontal face Haar Cascade, different sizes of ROI, alert notification and a pilot test.

a. Experiment 3-1 Rotated Face Detection

This experiment was conducted in the research lab with different head rotation positions to the left and right for face detection as shown in Figure 4-34. Two tests were performed with five rounds of face detection. There were 100 frames captured for both tests. *FrameRotation* was not used in Test 1 but it was used in Test 2.

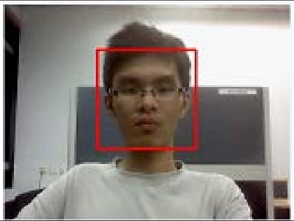
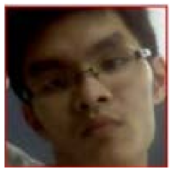
Description	Input Frame	Rotated Frame	Extracted Image
Head Position : Straight **Rotate frame** : none **Num. Rotation / Detection:** 0 / 1			
Head Position : bend to the left **Rotate frame** : 30^0 rotation to the right **Num. Rotation / Detection:** 1 / 2			
Head Position : bend to the right **Rotate frame** : 60^0 rotation to the left **Num. Rotation / Detection:** 5 / 6			

Figure 4-34 Rotated face detection description and frames screenshot

Experiment results in Table 4-12 shows that ACApp could perform better in the *FaceDetection* process for the tilted head position after the incorporation of the *FrameRotation* with the naïve Viola-Jones face detection algorithm.

Table 4-12 Comparison face detection result between non face rotation and rotated face

Round	Test 1 : Frame captured – 100 | Without Frame Rotation				Test 2 : Frame captured – 100 | With Frame Rotation			
	Frames				Frames			
	True Positive Detection	False Positive Detection	Not Detected	Total	True Positive Detection	False Positive Detection	Not Detected	Total
1	47	12	41	100	73	16	11	100
2	35	2	63	100	83	4	13	100
3	30	0	70	100	97	3	0	100
4	26	0	74	100	74	6	20	100
5	30	0	70	100	99	0	1	100
Total	**168**	**14**	**318**	**500**	**426**	**29**	**45**	**500**
Percentage	**33.6%**	**2.8%**	**63.6%**	**100%**	**85.2%**	**5.8%**	**9.0%**	**100%**

The experiment results show that detection rate has increased from 36.4% (33.6% true positive detection and 2.8% false positive detection) in Test 1 to 91.0% (85.2% true positive detection and 5.8% false positive detection) in Test 2.

b. Experiment 3-2 Comparing Haar Cascade

OpenCV API comes with four basic frontal face Haar Cascade which includes "haarcascade_frontalface_default.xml" (HFD), "haarcascade_frontalface_alt.xml" (HFA), "haarcascade_frontalface_alt2.xml" (HFA2) and also "haarcascade_frontalface_alt_tree" (HFAT). This experiment was conducted to compare the four Haar Cascades and to obtain the best Haar Cascade which will produce the highest face detection rates.

Five pre-captured video footages on children with Cerebral Palsy were used to compare the performances of various Haar Cascade. The experiment results are shown in Table 4-13.

From the experiment with five samples of video footages, the face detection using HFA achieved an overall of 69.99% of detection rate but only 64.06% of true positive detection was achieved. An overall detection rate of 93.80% was achieved by using HFA2 with the true positive detection rate increased to 78.40% compared with the previous Haar Cascade.

The detection rate by using HFAT dropped to 74.79%.However, the true positive detection rate increased to 96.92% of the total detected rate. The HFD achieved the poorest result as the detection rate was 86.81% and the true positive detection was only 50.33% of the total detected rate.

Table 4-13 Comparison of face detection results with various frontal face Haar Cascade

FrontalFace Alt Cascade (HFA)						
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	3	10	12	25	52.00%	23.08%
2	3	6	11	20	45.00%	33.33%
3	21	12	9	42	78.57%	63.64%
4	11	15	8	34	76.47%	42.31%
5	44	3	1	48	97.92%	93.62%
Total / Average	**82**	**46**	**41**	**169**	**69.99%**	**64.06%**
FrontalFace Alt 2 Cascade (HFA2)						
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	15	6	4	25	84.00%	71.43%
2	15	2	3	20	85.00%	88.24%
3	39	3	0	42	100.00%	92.86%
4	15	19	0	34	100.00%	44.12%
5	43	5	0	48	100.00%	89.58%
Total / Average	**127**	**35**	**7**	**169**	**93.80%**	**78.40%**
FrontalFace Alt Tree Cascade (HFAT)						
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	14	0	11	25	56.00%	100.00%
2	16	0	4	20	80.00%	100.00%
3	39	0	3	42	92.86%	100.00%
4	17	4	13	34	61.76%	80.95%
5	40	0	8	48	83.33%	100.00%
Total / Average	**126**	**4**	**39**	**169**	**74.79%**	**96.92%**
FrontalFace Default Cascade (HFD)						
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	3	13	9	25	64.00%	18.75%
2	7	8	5	20	75.00%	46.67%
3	25	17	0	42	100.00%	59.52%
4	12	21	1	34	97.06%	36.36%
5	30	17	1	48	97.92%	63.83%
Total / Average	**77**	**76**	**16**	**169**	**86.81%**	**50.33%**

By looking at the result table, HFAT can perform better compared with the other three Haar Cascades as it obtained 74.79% of detection rate and almost all (96.92%) detected frames were true positive. Hence, the HFAT had replaced the HFD which was used in ACApp 2.0.

c. Experiment 3-3 Resizing Region of Interest (ROI)

This experiment was conducted to find out the best true positive detection rate by comparing different regions of interest (ROI) sizes in the face detection process. The initial ROI size was *100 X 100* pixels in ACApp 1.0 and ACApp 2.0. The distance between subject face and camera was 30 cm to 60 cm. 2 tests were conducted in this experiment where the ROI size for Test 1 was *100 X 100* pixels while the size of ROI for Test 2 was *75 X 75* pixels. The video samples used for both tests were the same.

Table 4-14 and Table 4-15 show the face detection results with different sizes of ROI. An overall detection rate of 32.24% was obtained and the true positive detection rate was 73.0% out of the total detection rate for Test 1 with ROI size of *100 X 100* pixels. In Test 2 with ROI size of *75 X 75* pixels, the overall detection rate increased to 68.1% and the true positive detection rates increased to 93.43% from the total positive detection. This experiment proves that not only the detection rate increase by resizing the ROI from *100 X 100* pixels to *75 X 75,* pixels but the true positive detection rates rapidly increased as well.

Table 4-14Theface detection result Region of Interest of 100 X 100 pixels

	Test 1 : Region of Interest Range Limit – *100 X 100* pixels					
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	26	0	24	50	52.0%	100.0%
2	18	7	17	42	59.5%	72.0%
3	3	4	18	25	28.0%	42.9%
4	1	1	11	13	15.4%	50.0%
5	2	0	32	34	6.3%	100.0%
Total	**50**	**12**	**102**	**164**	**32.24%**	**73.00%**
Percentage	**30.49%**	**7.32%**	**62.19%**	**100%**		

Table 4-15Theface detection result Region of Interest of 75 X 75 pixels

	Test 2 : Region of Interest Range Limit – *75 X 75* pixels					
Videos	**True Positive (*P*)**	**False Positive (*FP*)**	**Not Detected (*N*)**	**Total (*T* = *P* + *FP* + *N*)**	**Detection Rate** $\left(\frac{P+FP}{T}\right)$	**True Positive Rate** $\left(\frac{P}{P+FP}\right)$
1	38	0	12	50	76.0%	100.0%
2	39	0	3	42	92.9%	100.0%
3	14	0	11	25	56.0%	100.0%
4	6	1	6	13	53.8%	85.7%
5	17	4	13	34	61.8%	81.0%
Total	**114**	**5**	**45**	**164**	**68.1%**	**93.34%**
Percentage	**69.51%**	**3.05%**	**27.44%**	**100%**		

This experiment also showed that unwanted or background object which is non-human face can be eliminated from detection by scaling down the ROI size. Thus, it increases the detection and true positive detection rate. With the experiment results obtained, ACApp has changed the ROI size from *100 X 100* pixels to *75 X 75* pixels.

d. Experiment 3-4 Alert Notification Test via SMS

This experiment was conducted to measure the reliability of SMS delivery when a critical expression is detected during the real time monitoring session. The experiment was conducted by triggering the GSM modem to send messages randomly at different hours (morning, evening and night time) in various places (urban areas and housing estates).

Figure 4-35 shows a sample message sent by ACApp via a GSM modem to a receiver's mobile phone to inform the receiver of a detected critical expression during monitoring session where Melvin is the receiver and James is the child with Cerebral Palsy under monitoring.

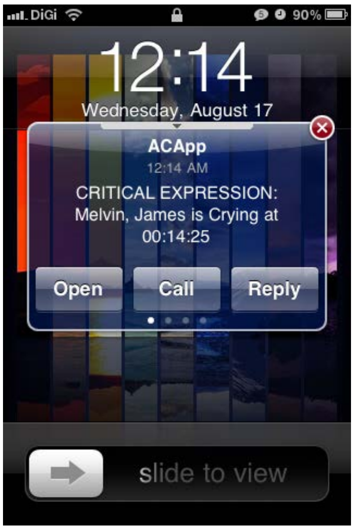


Figure 4-35 Alert notification message delivered to receiver's mobile phone

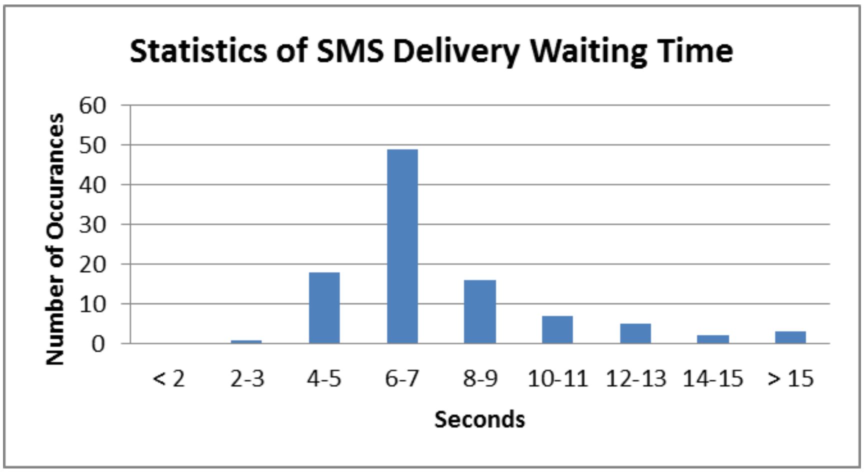


Figure 4-36 Statistics of number of occurrence for SMS delivery versus waiting time

Experiment results listed in Table 4-16 were obtained after two days of experiment. 96.19% out of 105 SMS sent were delivered (Figure 4-36) and it took 7.752 seconds on average to deliver a message to a receiver's mobile phone. In some cases, it took longer (39 Seconds – occurred once) which could be caused by service interruption or recipient's mobile phone delay.

Table 4-16 Alert notification testing result

Description	Result
SMS sending attempt	105
Number of SMS delivered	101
Succeed delivering rate	96.19%
Min. waiting time	3 Seconds
Max. waiting time	39 Seconds
Average waiting time	7.752 Seconds
Waiting time Standard Deviation	4.075%

Generally, the experiment result of 96.19% successful delivery rate with 7.752 seconds waiting time on average showed an accepted efficiency of alert notification via SMS during the monitoring session.

e. Experiment 3-5 Pilot Test with a Child with Cerebral Palsy

A pilot testing on ACApp 3.0 was conducted before running the real field testing. A child with Cerebral Palsy was invited to participate in this test. The child suffered from both physical and speech disabilities. The physical disability acquired caused him to be unable to walk properly and to have difficulties in holding an object with his own hands. He also had impairment in his speech making his words incomprehensible. The purpose of the testing was to evaluate the usability of ACApp before conducting the field test.

In the test, the child with Cerebral Palsy went through one session of expression training with five expressions. Later, the child went through another five sessions of real time monitoring. The result returned by ACApp was recorded in Table 4-17.

Table 4-17 Result of expression recognition with one child with Cerebral Palsy

Labels	1	2	3	4	5	Overall Recognition Rate
Matching	100%	80%	100%	60%	100%	**88%**

An overall expression recognition rate of 88% was obtained from the pilot test. This result showed that ACApp was ready for the real testing.

4.2.3.3 Memory Consumption Issue

In ACApp 3.0, all required features i.e. *Automated Brightness Adjustment* and *FrameRotation* were enabled whenever expression training or expression recognition sessions were conducted. This was a waste of computing resource i.e. memory consumption, especially in an environment with balance brightness and the head position of the child was not tilted.

Apart from that, ACApp 3.0 seems to have unreleased memory allocation while running which increased the memory consumption. This problem was identified by the developer while conducting several experiments which is shown in Figure 4-37. Other than that, no other bug was found in ACApp 3.0.

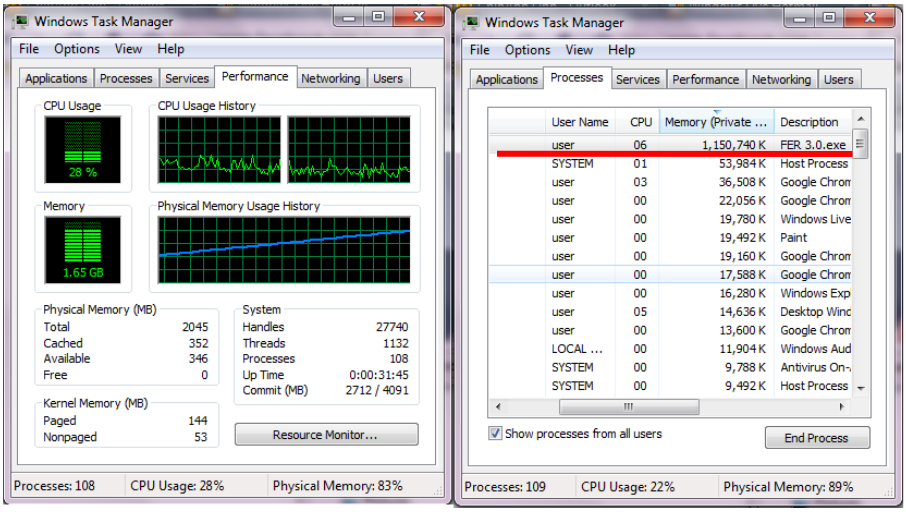


Figure 4-37 Memory leakage

4.2.3.4 Findings

Several major objectives were achieved throughout the developmental and experimental stages of ACApp 3.0. These achievements included migrating ACApp from Console based application to Windows Form Application, the adding of Profile Registering module and Alert Notification module implementation, and the integrating of the 3 modules i.e. the Profile Registering module, the Facial Expression Training and Recognition module and the Alert Notification module. This improvement solved the problems encountered during the experiments in ACApp 2.0 i.e. a more user friendly interface was needed to increase the usability aspect of the prototype.

The second problem discovered from Experiment 2-3 was related to the time required to train an expression but this has also been solved by implementing the pre-time expression training in ACApp 3.0. The expression training could be done separately where captured video file could be fed into ACApp for expression training.

The *FaceDetection* process was improved in ACApp 3.0 as well. This improvement involved rotated face detection, Haar Cascade replacement and resizing the ROI. Experiment results showed that the improvement increased the detection rates and most importantly it also increased true positive detection rate.

The results on Experiment 3-4 and 3-5 had strengthened the readiness of ACApp to venture into field testing. An overall expression recognition rate of 88% was obtained when conducting the pilot test on a child with Cerebral Palsy. Furthermore, experiments on Alert Notification via SMS with a 96.19% successful delivery rate at 7.752 seconds waiting time on average showed an accepted efficiency of alert notification in ACApp during the monitoring session.

Minor bugs on ACApp 3.0 were identified during the experiments. The developer found that a switch to enable and disable *Automated Brightness Adjustment* and *FrameRotation* was required to avoid the waste of resources. Besides, ACApp 3.0 had memory leakage problem in this stage. The two problems would be solved in the next version of ACApp before it could be used for field testing.

4.2.4 ACApp V4.0

The main tasks of the ACApp 4.0 implementation included codes cleaning and problems solving which were identified in ACApp 3.0. The finalized ACApp is expected to be ready for field testing.

4.2.4.1 Improvements

The major improvement involved redesigning ACApp's GUI to overcome the usability problem discovered in ACApp 3.0. Apart from that, switches were also added to the new interface to control *Automated Brightness Adjustment* and *FrameRotation* processes to reduce memory consumption while running ACApp.

a. Redesigned Graphic User Interface

ACApp 4.0 GUI was polished and re-designed to increase its usability. Previously, ACApp 3.0 had multiple tabs and windows to control its functionalities. The major control and command buttons were placed on one control panel in the redesigned ACApp GUI.

i Main ACApp Interface

Figure 4-38 shows the ACApp 4.0 GUI with most of the major features' control being placed within this panel. The explanation of this GUI has been divided into 10 portions with each portion was labeled with a number.

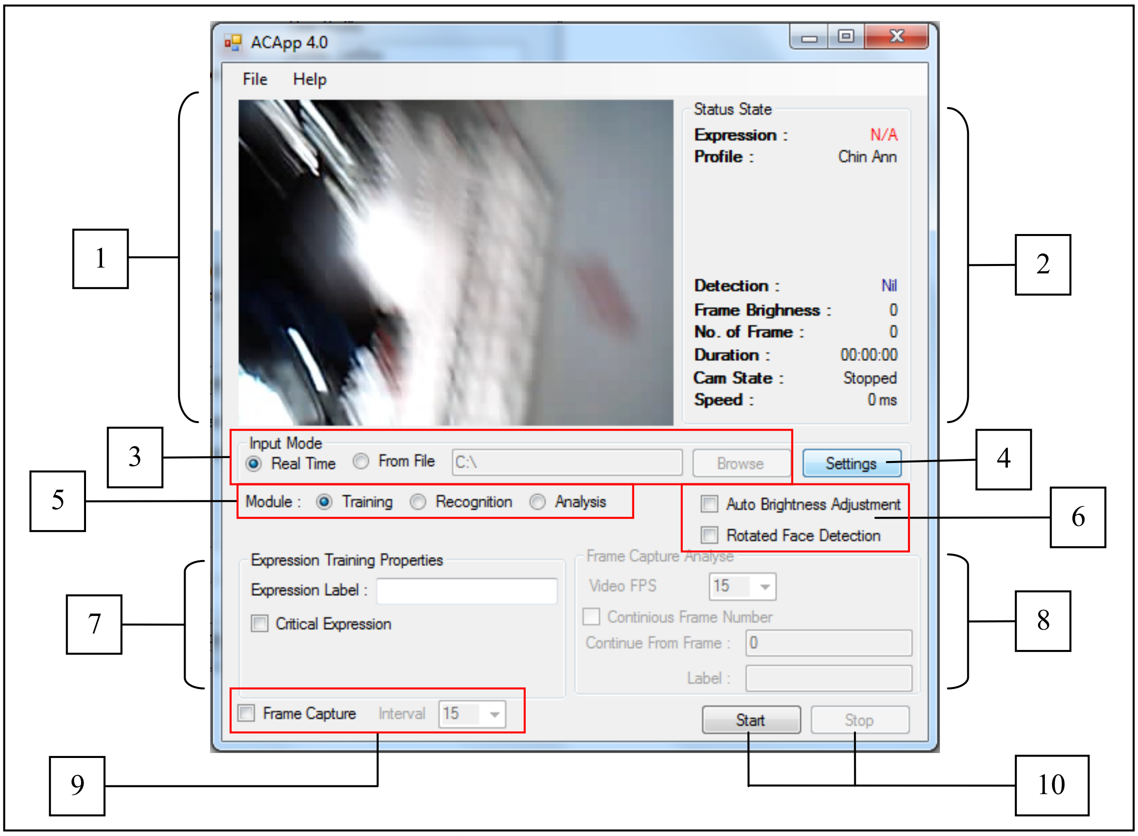


Figure 4-38 ACApp 4.0 GUI

Each frame captured in real time or prerecorded video footages will be displayed in Portion 1. Portion 2 will display the status of the current process during expression training or expression recognition. This information includes the current selected user profile, the trained or recognized expression, frame brightness value, index of frame, duration, web cam status, and processing speed.

Portion 3 shows the “Input Mode” control where the user can choose to run expression training or recognition from real time or prerecorded video footages. The “Browse” button is inactive when the real time input mode is selected and it will be activated when “From File” input mode is selected. Portion 4 indicates ACApp setting or specifically the user profile setting. When the setting button is pressed, an extended menu will be shown on the right side and will disappear if the setting is pressed for the second time.

Portion 5 shows the control of the operation mode. Three modes are available currently which include expression training, expression recognition and frame analysis (for testing purpose). Portion 6 shows the switch to enable or disable the *Automated Brightness Adjustment* and *FrameRotation* process. The expression training properties are located at Portion 7 in the GUI. This portion allows the user to enter the expression’s label that is to be trained and this label can be marked as a critical expression by checking the “Critical Expression” check box. Portion 8 will be activated when the “Analysis” mode in Portion 5 is selected. Portion 9 serves the same purpose as Portion 8.This portion is used to save frames captured in real time or from prerecorded footages. The Start and Stop switches for all operation modes are located in Portion 10.

ii Extended ACApp Interface

The extended ACApp 4.0 GUI which is shown in Figure 4-39 can be displayed and hidden by pressing the “Setting” button. The list box in Portion 11 indicates a list of available user profiles. A new profile can be added by entering a new name and by pressing the “Add” button. Portion 12 displays the list of expressions available for a selected profile.

The expressions listed can be selected and they can be removed from the expression list. On the other hand, the selected expression(s) can be marked as critical expression by pressing the “Add to Critical” button. The expression which is marked as a critical expression will be listed in the “critical expression list” located in Portion 13. The guardian will be notified either by sound, SMS or both (depending on the profile preference) when a critical expression is detected during the monitoring session. A critical expression can be unmarked by removing it from the “critical expression list”.

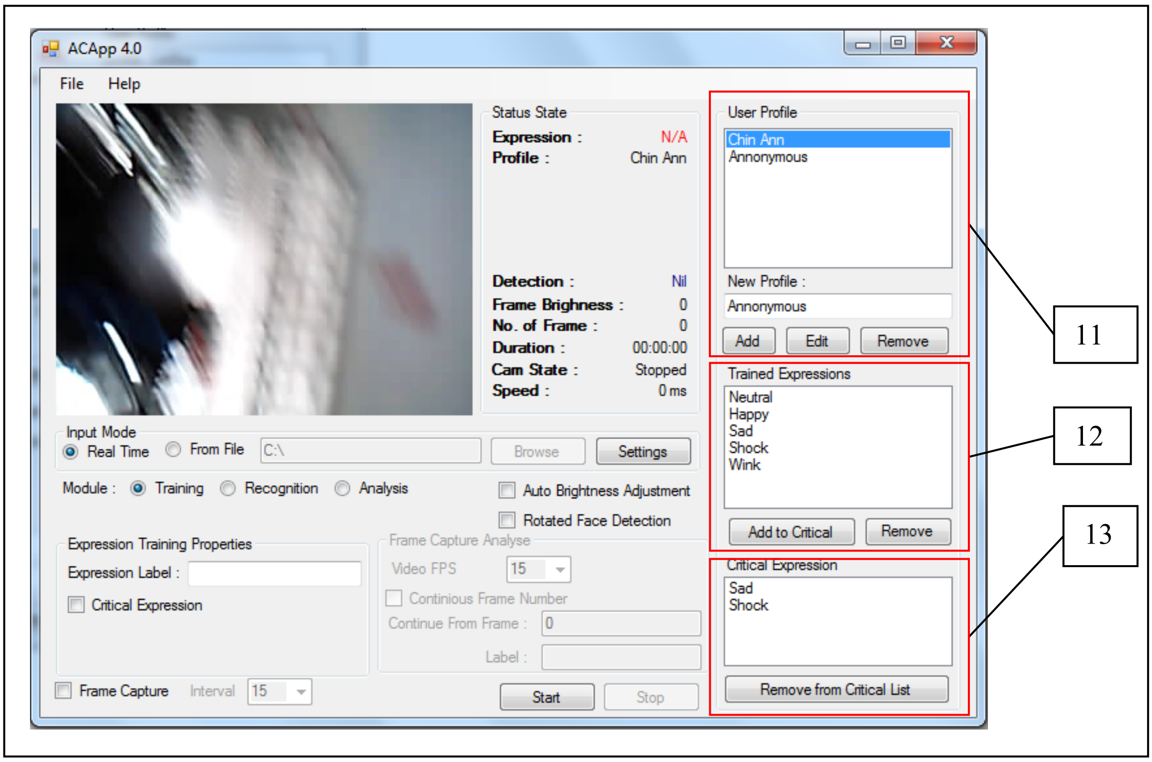


Figure 4-39 Extended ACApp 4.0 GUI

iii Edit Profile Interface

Figure 4-40 shows the GUI for profile editing in ACApp 4.0. Guardians are allowed to edit a child's profile by pressing "Edit" button in the extended interface.

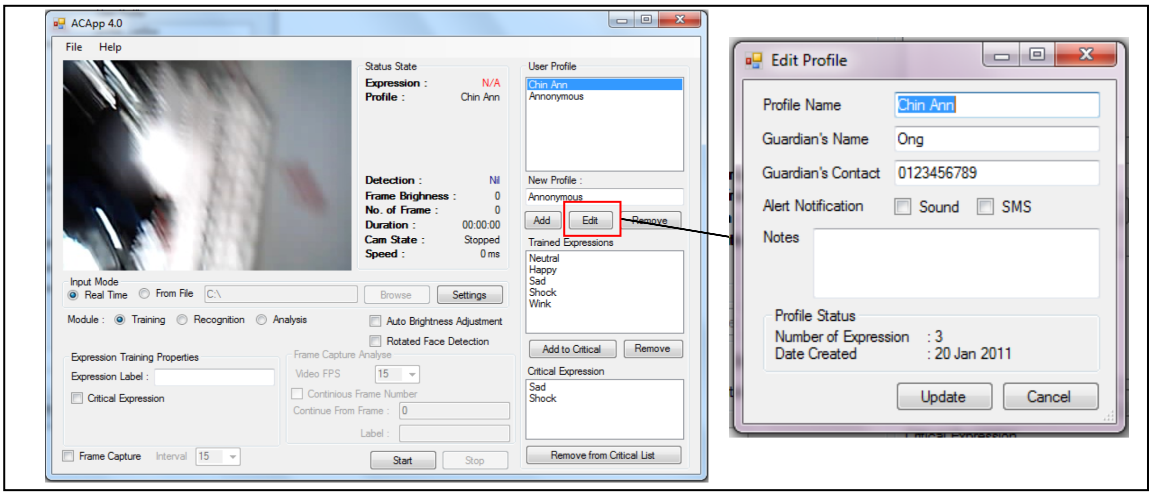


Figure 4-40 ACApp 4.0 Edit Profile GUI

This editing process includes renaming the user profile, defining the guardian's name, defining the guardian's contact number (for SMS notification purposes), alert notification selection (sound, SMS or both), and adding notes to the selected profile.

b. Switches for Optional Features

Previously, unnecessary memory consumption occurred when ACApp 3.0 was running under sufficient lighting condition. Besides, unnecessary memory consumption also happened when running ACApp 3.0 on children with Cerebral Palsy did not have a serious tilted head position.

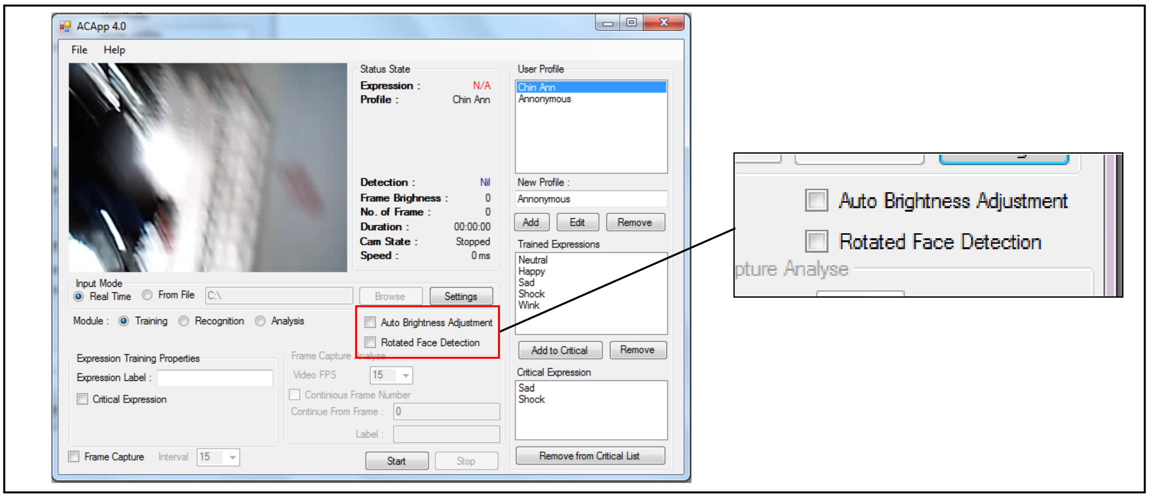


Figure 4-41 Automated Brightness Adjustment and Rotated Face Detection (FrameRotation) switch

In order to solve the problem, switches for *Automated Brightness Adjustment* and *FrameRotation* processes were added into ACApp 4.0 interface as shown in Figure 4-41. The switches allowed the guardian to enable *Automated Brightness Adjustment* and *FrameRotation* process when needed.

c. Memory Leakage and Code Clean Up

Another minor problem was discovered in Experiment 3-5 where ACApp 3.0 experienced memory leakage when it is used for a period of time. This problem was solved by releasing the memory allocation for unused variables in the program. Apart from that, some code cleanup was made to improve the program maintainability.

d. Special Control on GUI for Testing Purpose

A special feature named "Frame Capture" is added into ACApp 4.0 for field testing purposes in the evaluation stage. The frame capture allows initial the frame displayed on ACApp screen to be saved into local storage where further analysis and evaluation can be done. This feature can be enabled by checking the "Frame Capture" checkbox which is displayed in Figure 4-42.

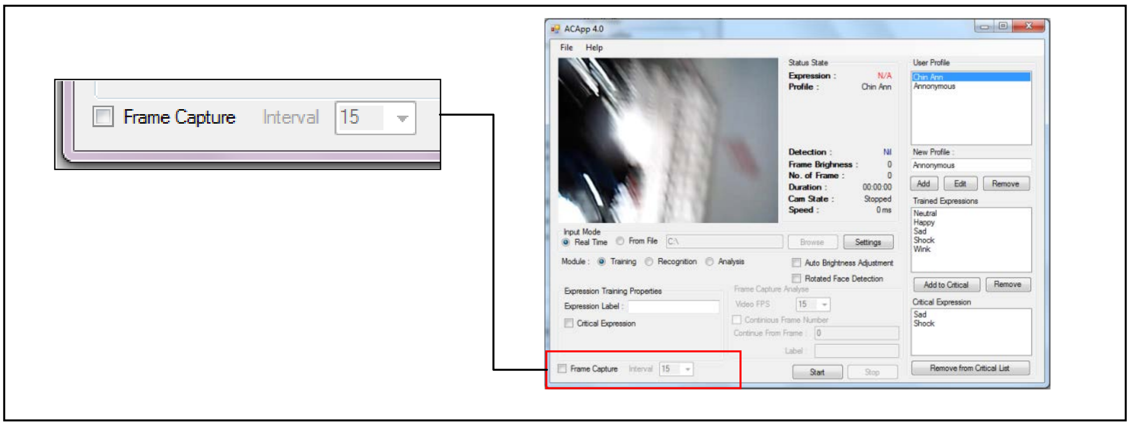


Figure 4-42 Frame capture switch for testing purpose

A sample of saved frame is shown in Figure 4-43. Additional information which includes current recognized expression, current frame number and current timeline are printed in the frame as well for evaluation purposes.

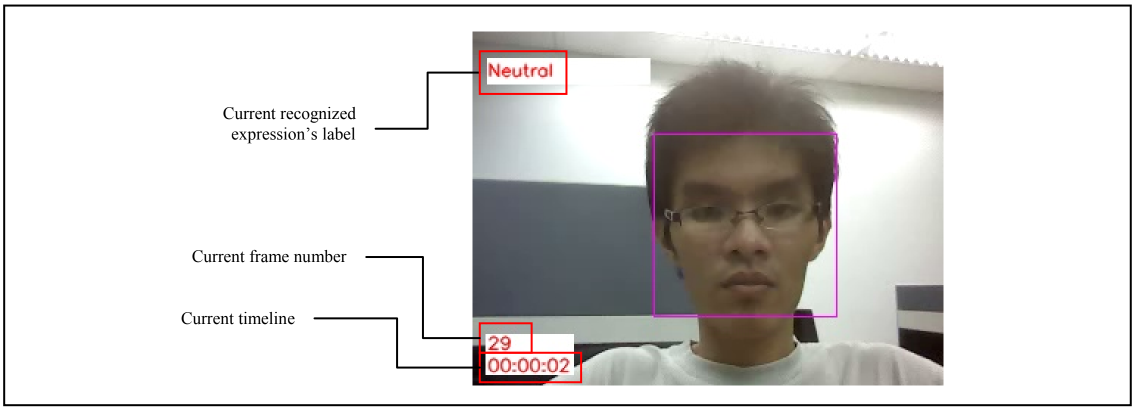


Figure 4-43 Sample of saved frame with information

4.2.4.2 Experiments

Two experiments on processing speed were conducted with ACApp 4.0. These experiments were tested with two methods. 1) Speed test by processes and 2) overall speed test while running ACApp.

a. Experiment 4-1 Process Speed Test

Generally, there are 4 cores and 2 optional processes that require memory consumption and might cause a delay in overall ACApp operation. The core processes includes *FaceDetection*, *ImageProcessing*, *TemplateStoring* and *TemplateMatching* while the optional processes are *Automated Brightness Adjustment* and *FrameRotation*.

All these processes were divided and tested separately to obtain the time required for each process. For each process (except *FrameRotation* – 100 frames), 150 frames were captured for processing. A timer was set in the processes and it started counting before the process started and stopped time counting when a process ended. Table 4-18 shows the result of time required for each process.

Based on experiment result, an average of 34 milliseconds were required to perform *FaceDetection* for each frame. 133 milliseconds were required to perform *Automated Brightness Adjustment* for 1 frame. In *FrameRotation*, 155 milliseconds were required to perform 6 rotations for 1 frame while 37 milliseconds for each rotation. For *ImageProcessing*, less than 1 millisecond was required for each frame. An average of 0.8 milliseconds was needed to store one template. Lastly, an average of 12.4 milliseconds were required to perform *TemplateMatching* for an image that to be matched with 100 templates (5 labels with 20 samples per label).

Table 4-18 Table of time required for ACApp 4.0 processes

Process	Tests	Total (ms)	Average (ms/frame)	Min (ms)	Max (ms)	Num. frame / Remarks
Face Detection (FD)	1	5564	37	28	65	150
	2	4810	32	23	50	150
	3	5014	33	24	54	150
	4	4916	32	22	55	150
	5	5485	36	24	62	150
	AVG	**5157**	**34**	**24.2**	**57.2**	**150**
Automated Brightness Adjustment (ABA)	1	18150	121	101	239	150
	2	23007	153	105	210	150
	3	18928	126	103	232	150
	4	21728	144	102	209	150
	5	18247	121	102	232	150
	AVG	**20012**	**133**	**102.6**	**224.4**	**150**
Frame Rotation (FR)	1	15565	155	123	222	100 (with 6 rotations)
	2	259	26	37	21	100 (with 1 rotation)
Image Processing (IP)	1	95	< 1	< 1	2	150
	2	96	< 1	< 1	2	150
	3	92	< 1	< 1	1	150
	4	96	< 1	< 1	2	150
	5	89	< 1	< 1	1	150
	AVG	**93.6**	**< 1**	**< 1**	**1.6**	**150**
Template Storing (T) – Expression Training	1	219	1	2	84	150
	2	83	< 1	3	36	150
	3	160	1	2	69	150
	4	166	1	2	45	150
	5	168	1	3	50	150
	AVG	**159.2**	**0.8**	**2.4**	**56.8**	**150**
Template Matching (R) – Expression Recognition (*5 labels and 20 samples per label*)	1	2103	14	10	18	150
	2	1932	12	11	21	150
	3	1410	9	8	16	150
	4	2229	14	11	19	150
	5	2099	13	11	18	150
	AVG	**1954.6**	**12.4**	**10.2**	**18.4**	**150**

b. Experiment 4-2 ACApp speed test

Figure 4-44 shows the processes with the average time required for each process while running ACApp. Assuming that an expected time required is calculated by combining the processes depending on the needs. For example, to perform basic facial expression training, 3 processes are required. These processes include *FaceDetection* (FD), *ImageProcessing* (IP) as well as *TemplateStoring* (T). In this case, the expected time required to perform the basic facial expression training is 34.8 milliseconds for each frame by combining average time required in each process.

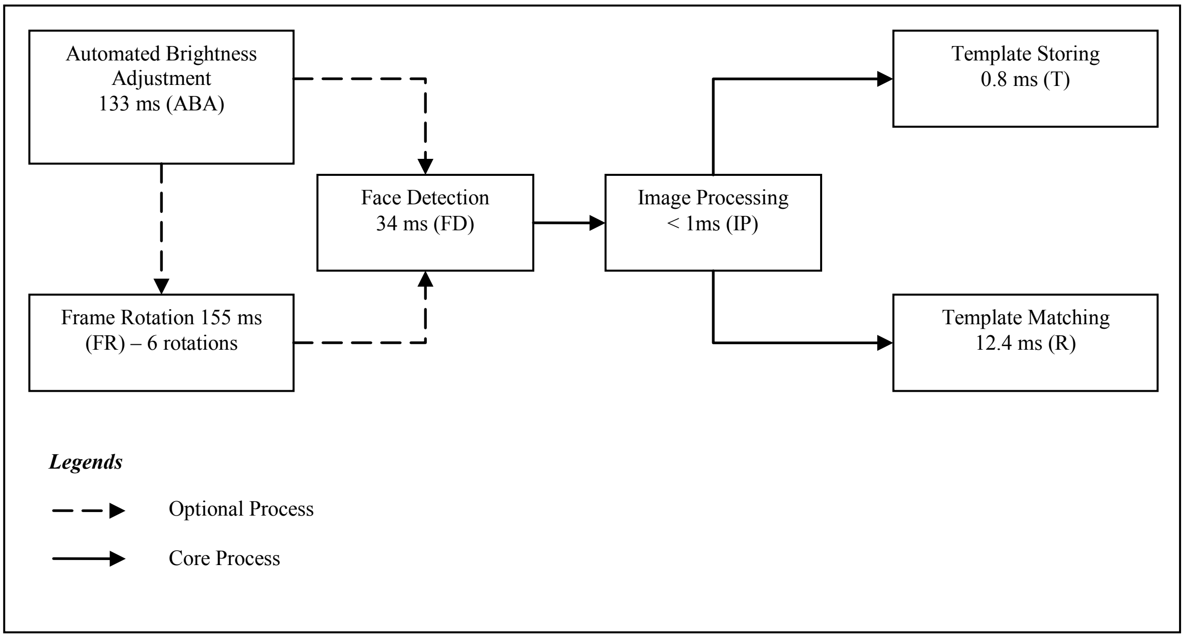


Figure 4-44 Process and expected completion time

An overall processing speed test was conducted and the actual test took up an average of 48 milliseconds for each frame to perform the basic facial expression training process. The difference between expected duration and real overall processing test is only 13.2 milliseconds. Hence, the other overall process is also done in this way as shown in Table 4-19.

Table 4-19 Overall process speed test

Overall Processes	Expected Avg. Duration (ms)	Overall Process Test (150 frames)				
		Total (ms)	Average (ms/frame)	Min (ms)	Max (ms)	Difference (Average - Expected)
FD + IP + T (Basic)	34.8	7240	48	24	1049	13.2
FD + IP + R (Basic)	46.4	8016	53	33	1037	6.6
ABA + FD + IP + T	167.8	29706	198	173	1223	30.2
ABA + FD + IP + R	179.4	24875	165	133	1208	-14.4
FR + FD + IP + T	189.8	32553	217	187	1421	27.2
FR + FD + IP + R	201.4	32711	218	179	1374	16.6
ABA + FR + FD + IP + T	332.8	51342	342	210	1477	9.2
ABA + FR+ FD + IP + R	334.4	53698	358	202	1867	23.6

This experiment showed that ACApp is efficient in its processing speed. The most complicated set of processes (ABA + FR + FD + IP + R) only required an average of 358 milliseconds (less than 1 second) to complete the expression recognition process with *Automated Brightness Adjustment* and *FrameRotation*.

4.2.4.3 ACApp's Overall Functionality Checklist

An overall functional checklist based on functional and non-functional requirement stated in Chapter 3 was created to evaluate ACApp's readiness before conducting the field testing. The overall functional checklist is shown in Table 4-20.

Based on the functional checklist above, ACApp 4.0 has reached a milestone where it is ready to bring forward for field testing. The Major requirements have been fulfilled and completed. The only requirement where it is only partially completed is the extensibility. This requirement can be fulfilled after the final evaluation is made.

Table 4-20 ACApp overall functionality checklist

Requirements	Description	Status	Remarks
Functional Requirement	R8. Profile Management	Completed	This requirement was fulfilled with module 1 in ACApp 3.0 onwards – Profile Registering.
	R9. Notification Setting	Completed	This requirement was fulfilled with module 3 in ACApp 3.0 onwards – Alert Notification.
	R10. Face Detection	Completed	This requirement was fulfilled in ACApp 1.0 and enhanced in ACApp 3.0.
	R11. Expression Training	Completed	This requirement was fulfilled in ACApp 1.0 and enhanced in ACApp 3.0
	R12. Expression Recognition	Completed	This requirement was fulfilled in ACApp 1.0 and enhanced in ACApp 2.0
	R13. Alert Notification	Completed	This requirement was fulfilled in ACApp 3.0 and enhanced in ACApp 4.0.
Usability	Having graphic user interface Simplified interface Provides user with help and guideline for troubleshooting	Completed	This requirement was fulfilled in ACApp 3.0 and enhanced in ACApp 4.0.
Reliability	Unused memory consumption in between processes shall be released immediately Able to process and return accurate result as it is trained to recognize human facial expressions Able to store multiple users profiles with large number of labels	Complete	Unused memory consumption requirement was fulfilled in ACApp 4.0 where "Solving Memory Leakage Bug and Code Clean Up" was done. The recognition accuracy and storage requirement is fulfilled in ACApp 2.0 onwards.
Environmental Conditions	Able to be attached to the wheelchair and carry around while monitoring is performed Able to work under different environmental condition	Completed	This requirement was fulfilled in ACApp 2.0 and enhanced in ACApp 4.0.
Maintainability	Able to organize related data in an appropriate sequence or directory Data backup shall be included to avoid data loss or accidental deletion	Completed	This requirement was fulfilled in ACApp 3.0 and enhanced in ACApp 4.0.
Extensibility	Able to expand or be enhanced in future to fulfil user's extended needs and requirements All additional enhancement modules shall be able to be added in easily without restructuring the entire prototype. Each error occurred during run time shall be recorded or logged	Partially completed	The prototype was built using C++ language and developed using Window Form Application template. Additional features can be plugged in easily in future
Portability	Ready to run and easy to install in most of the laptop with compatible Operating System.	Completed	The prototype was compiled into an .exe application and ready to be used in any machine according to the specification stated earlier in this chapter.

4.2.4.4 Findings

Major objectives were achieved especially in the improved GUI. The new GUI makes ACApp easier to use where most of the commands and buttons are now located within the panel. Hence, the user does not need to perform multiple events in order to get a task done. On top of that, the memory leakage problem which was discovered in ACApp 3.0 was solved.

Experiments were conducted to evaluate the processing speed for the facial expression training, and the facial expression recognition modes. When performing basic expression training, only 48 milliseconds on average were needed to process a frame while 53 milliseconds were needed for basic expression recognition. An average of 358 milliseconds was needed to process a frame while performing expression recognition with *Automated Brightness Adjustment* and *FrameRotation*.

An overall prototype's functionality checklist was created at the end of the ACApp development to confirm its readiness for the field testing. Up to this stage, no critical or major problem was found in ACApp 4.0. Hence, it was confirmed that ACApp 4.0 had meet the requirements stated in Chapter 3 and was ready for field testing.

4.3 Summary

Generally, this chapter had discussed two major parts concerning the ACApp implementation where the pre-implementation process was discussed in section 4.1 and the prototype implementation process in section 4.2.

In pre-implementation, the required developmental tools and techniques were discussed. From there, the best alternative tools and techniques were selected and adopted in the prototype implementation as well. The developmental tools consisted of hardware and software. The hardware used in this research included a personal computer (laptop), webcam, and a GSM modem while the software used included Microsoft visual studio 2008, OpenCV, Microsoft .NET Framework 3.5 SP1, Microsoft redistribution package 2005 SP1 and mobitek SMS gateway development kit.

The potential techniques used for prototype development were adopted after doing some comparison. In face detection, Viola-Jones face detection algorithm was chosen as it was the common method used in any face detection related application. In expression recognition, the Template Matching technique was used due to its light weight processing and algorithm simplicity.

In the prototype implementation, the prototype development was divided into 4 stages where 4 versions of ACApp were developed. Each stage had its own goal, focus, implementation, experiments and outcome. In stage one, ACApp 1.0 was developed with the basic facial expression training and expression recognition functionality. Two experimental results were obtained and showed that ACApp achieved 75% of face detection rate and 75% of expression recognition rate. Two limitations were found in this version of the prototype while conducting the experiment. The limitations found were related to face detection and expression recognition where the accuracy obtained was below expectation.

Improvement on ACApp 1.0 was done in stage two and an improved ACApp 2.0 was developed. The limitations discovered in ACApp 1.0 were solved in stage two development. Experiments on face detection and expression recognition rates for ACApp 2.0 were conducted where 95% of detection rate and 91.42% of recognition rate were obtained. According to the participants' feedback, ACApp 2.0 seemed to be less usable as it was a console based application. Furthermore, the duration required to perform the expression training was 22.56 seconds per expression on average.

In stage three, the ACApp 2.0 which was a console based application was transformed into a GUI based application. The ACApp 3.0 Profile Registering module and the Alert Notification module were implemented and added into ACApp 3.0 as well. In response to feedback received from the participants previously, a pre time expression training mode was added as well to make the expression training simple and to shorten the training time. The *FrameRotation* process was added as well to improve the *FaceDetection* process.

In the final stage, ACApp 4.0 was developed. Major improvements on GUI were made in this stage where all commands and buttons were located within a panel. Memory leakage problem was also solved in this stage and operation speed test experiments were conducted. Experiments found that ACApp 4.0 managed to perform basic operations i.e. facial expression training and facial expression recognition within a shorter period, which are 48 milliseconds and 53 milliseconds on average respectively for each frame.

Lastly, an overall functionality checklist was created. Based on the checklist, ACApp 4.0 was ready for field testing as it has fulfilled all major requirements stated in Chapter 3.

Chapter 5: Testing and Evaluation

This chapter evaluates the effectiveness on utilizing facial expression as a medium of communication for children with Cerebral Palsy, by conducting field testing with these children. A group of children with Cerebral Palsy, who received special education at a NGO school was invited to participate.

A series of experimental procedures consisting of the Pre-Capturing Phase, the On-Capturing Phase and the Post-Capturing Phase were defined as guidelines to conduct the evaluation. In the Pre-Capturing Phase, all required materials for evaluation purpose were prepared and gathered. The participants were invited to a watch collection of short video clips in the On-Capturing Phase while their facial expressions were recorded. These recorded videos were used for ACApp testing in the Post-Capturing Phase.

At the end of the evaluation, three testing result were produced which included the false positive detection rate, the recognition rate and the alert notification rate.

5.1 Participants

21 children with Cerebral Palsy with different years of birth ranging from the year 1996 to 2002 (aged 9 to 15 in year 2011 - Figure 5-1) from the special school were invited to participate in the ACApp evaluation. The children were recruited through a proper procedure, that what with concern of the parent. Subsequently, two forms were distributed to each child and his/her parents. They were "The Research Child's Assent Form" and the "Informed Consent Form" (refer to Appendix 1 and Appendix 2).

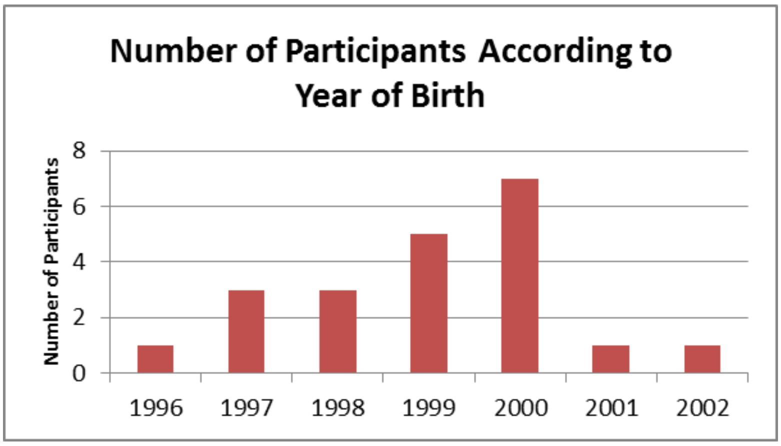


Figure 5-1 Number of participant according to year of birth (n = 21)

Based on the information obtained from the school, some of the children were diagnosed with a single disorder and some were diagnosed with multiple disorders.Table 5-1 shows the list of disabilities diagnosed for 21 participants. Three of them were diagnosed with general or mixed disorder. One was diagnosed with Spastic Diplegia and five were diagnosed with Spastic Quadriplegia. Another three were diagnosed with severe intellectual disability and one each for the other disorders which include Severe Birth Asphyxia, learning difficulty (Bladder Drainage), Epilepsy with Global Development Delay, Epilepsy with Dysmorphic facies, delay milestone and hypotonic, Global Development Delay, Microcephaly with Intellectual Disability, Physical Handicapped, Tuberous Sclerosis with Global Development Delay, and Arthogynphosis

Table 5-1Number of participant according to disabilities diagnosed

Types of Cerebral Palsy	Number of Participants
1) General / Mixed Disorder	3
2) Spastic Diplegia	1
3) Spastic Quadriplegia	5
4) Intellectual Disability	3
5) Severe Birth Asphyxia	1
6) Learning difficulty (Bladder Drainage)	1
7) Epilepsy with Global Development Delay	1
8) Epilepsy with Dysmorphic facies, delay milestone and hypotonic	1
9) Global Development Delay	1
10) Microcephaly with Intellectual Disability	1
11) Physical Handicapped	1
12) Tuberous Sclerosis with Global Development Delay	1
13) Arthogynphosis	1
Total	**21**

Generally, this research has classified the participants into two groups, the high functional participants and the low functional participant as shown in Figure 5-2. Among all 21 participants, 15 (11 boys and 4 girls) of them were classified as high functional where they only experience physical and speech disabilities with mild global development delay. The other 6 participants (6 boys and 0 girls) were classified as low functional participants as they have multiple disabilities including severe intellectual disability.

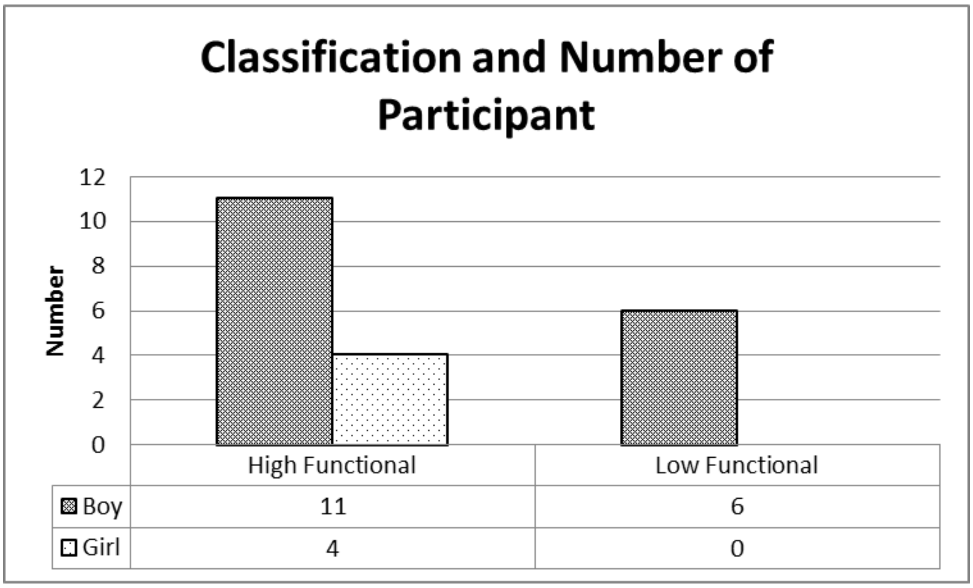


Figure 5-2 Classification and number of participants (n = 21)

5.2 Experimental Procedure

Testing on ACApp was needed to measure the effectiveness of real time monitoring of children with Cerebral Palsy. A series of experimental procedures were defined as illustrated in Figure 5-3 in order to conduct the evaluation smoothly with children with Cerebral Palsy. This procedure was specified to avoid repeating the tests involving participants who might not be able to co-operate well. Through this procedure, the participants' natural facial expressions could be obtained and used for retest later, if necessary, to produce better results to ensure that the evaluation result is accurate.

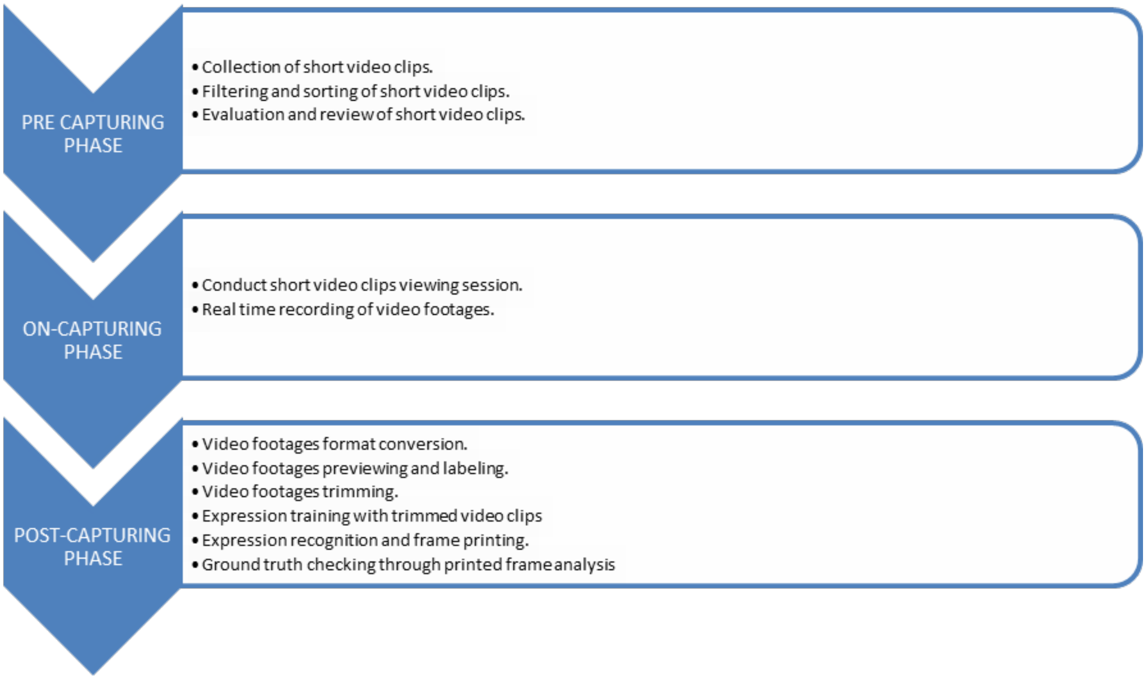


Figure 5-3 Experiment procedure checklist

The experiment procedure was divided into three phases namely the Pre-Capturing Phase, the On-Capturing Phase and the Post-Capturing Phase where each phase consisted of other smaller tasks. Generally, the Pre-Capturing Phase involved most of the preparation and material gathering tasks for the whole testing, done in a research lab. The On-Capturing Phase task was performed in the field where participants were invited to watch a series of short video clips. Their facial expressions during the viewing of the video were captured. The captured video footages were used to train ACApp and to create templates for each child accordingly. Another 2 sets of video footages were used to test the expression recognition rate of ACApp in the Post-Capturing Phase.

5.2.1 Pre-Capturing Phase

In the Pre-Capturing Phase, 36 short video clips were collected from the Youtube where each video was expected to trigger the emotion of the audience like boredom / neutrality, happiness, sadness, fear and disgust. All the short video clips were then sorted out and distributed to volunteers for review and rating for the relevancy of emotional elements contained in each video. Ten volunteers in the research lab aged from 19 years to 40 years with normal intelligence, were invited to review and to rate these short video clips.

Table 5-2 List of selected and rated short video clips

Expected Emotional Elements	Clip Title / Description	Duration	Average Rate of Relevancy
Neutral	Financial stock analysis presentation	00:04:20	100%
	NTV7 News	00:02:57	100%
	Campus voice debates	00:02:58	100%
	Financial presentation slides	00:02:03	100%
Happy	Tom and Jerry (Bell)	00:02:18	90%
	Tom and Jerry (Dog House)	00:02:40	90%
	Casper	00:06:15	90%
	Funny cats and dogs	00:03:13	100%
Sad	The Tale of Mari and 3 Puppies	00:06:07	100%
	A girl playing violin	00:03:57	90%
	Lonely old traveler lady	00:02:54	100%
	Horse sinking	00:03:16	90%
Fear	Sister kidnapped by UFO	00:02:57	80%
	Crocodile attack	00:01:38	100%
	Jaws attack	00:02:17	90%
	Shark attack	00:02:47	100%
Disgust	Fear Factor - Eating spider	00:01:55	100%
	Man vs. Wild - Eating Larva	00:01:25	100%
	Human bot fly	00:04:02	100%
	Leech vs. Earthworm	00:03:22	90%

Table 5-2above shows 20 out of 36 short video clips were reviewed and rated by ten volunteers. 20 clips were selected and they were rated by the volunteers with an average of 80% to 100% relevancy rating. These videos were rearranged into 4 sets of clips, each set consisted of 5 short video clips that had 3 or more different emotional elements as shown in Table 5-3. 3 sets of the short video clips were played to gather expressions of children with Cerebral Palsy during the On-Capturing Phase. Set 4 was reserved as a backup in case unexpected events happened in the On-Capturing Phase later. The total duration for all 4 sets of short video clips was about 60 minutes with approximately 15 minutes for each set of short video clips.

Table 5-3 Rearranged short movie clips playing sequence.

Set	Clip Title	Expected Emotional Element	Duration	Total Duration
1	Financial presentation slides	Neutral	00:02:03	00:16:42
	Tom and Jerry (Dog House)	Happy	00:02:40	
	Casper	Happy	00:06:15	
	Shark attack	Fear	00:02:47	
	Sister kidnapped by UFO	Fear	00:02:57	
2	Funny cats and dogs	Happy	00:03:13	00:13:33
	Man vs. Wild - Eating Larva	Disgust	00:01:25	
	Fear Factor - Eating spider	Disgust	00:01:55	
	Human bot fly	Disgust	00:04:02	
	Campus voice debates	Neutral	00:02:58	
3	The Tale of Mari and 3 Puppies	Sad	00:06:07	00:16:52
	Lonely old traveler lady	Sad	00:02:54	
	Horse sinking	Sad	00:03:16	
	NTV7 News	Neutral	00:02:57	
	Crocodile attack	Fear	00:01:38	
4 (Reserve)	Tom and Jerry (Bell)	Happy	00:02:18	00:16:14
	Financial stock analysis presentation	Neutral	00:04:20	
	Jaws attack	Fear	00:02:17	
	Leech vs. Earthworm	Disgust	00:03:22	
	A girl playing violin	Sad	00:03:57	

5.2.2 On-Capturing Phase

This phase was performed in the field, where participants' facial expressions were recorded while they were watching short video clips in their classroom, during their free time in between lessons on normal school days. Each participant was invited to watch the short video clips (approximately 45 minutes for each participant) individually where 1 set of short video clips was shown in each session.

During the recording session, the participant sat in front of a laptop which showed the prepared short video clips as shown in Figure 5-4. The distance between the participant and the webcam was approximately 60 centimeters. Another laptop was used to connect and control the webcam for starting and stopping the video recording. The webcam used was Logitech QuickCam E2500 and the application used was QuickCam V 11.9. Hence, the recorded video footages file format was in windows media video (wmv). 61 video footages were successfully recorded in 21 normal school days with about 2 to 3 video footages being recorded during recess time.

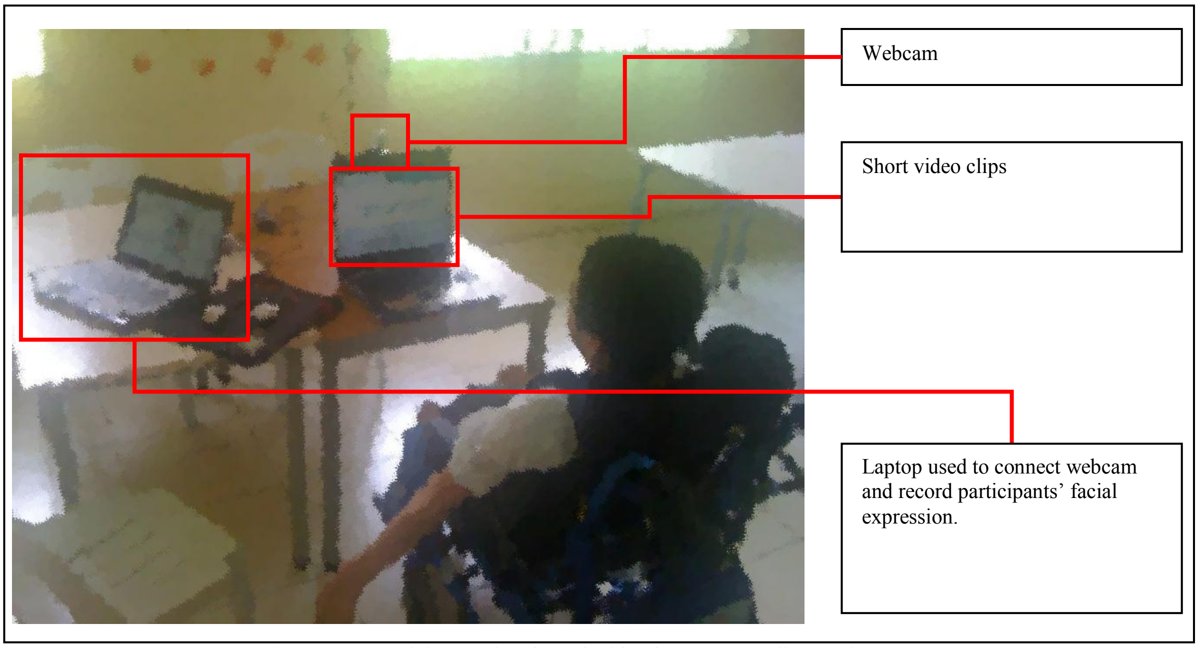


Figure 5-4 A participant going through video footages recording session

5.2.3 Post-Capturing Phase

In the Post-Capturing Phase, the testing on ACApp was conducted in the research lab. The tasks involved in this test included video footages conversion, footages labeling, footages trimming, expression training, expression recognition and frame printing, and lastly frame analysis.

The obtained video footages for each participant went through a video conversion process with the Mencoder (2010) to convert windows media video (wmv) to audio video interleave (avi).

After the video format conversion, the videos were previewed to label the five expected unique expressions of the participants. Each of the unique facial expressions had the duration of 1 to 8 seconds. The labeled portions were trimmed for ACApp expression training. When running expression training for each participant, 5 or more trimmed, short videos were fed into ACApp.

After expression training was done, all 3 converted videos for each participant, which was recorded in three sessions during the On-Capturing Phase, would go through an expression recognition process. In this process, ACApp would recognize or classify a participant's facial expression and save 1 true positive detection frame per second (with the interval of 15 positive detected frames) into a PDF file with labels of expression, frame number and timeline as shown in Figure 5-5. 1 set of PDF file containing printed frames was produced for each set of converted videos. The PDF files were analyzed for ground truth checking.

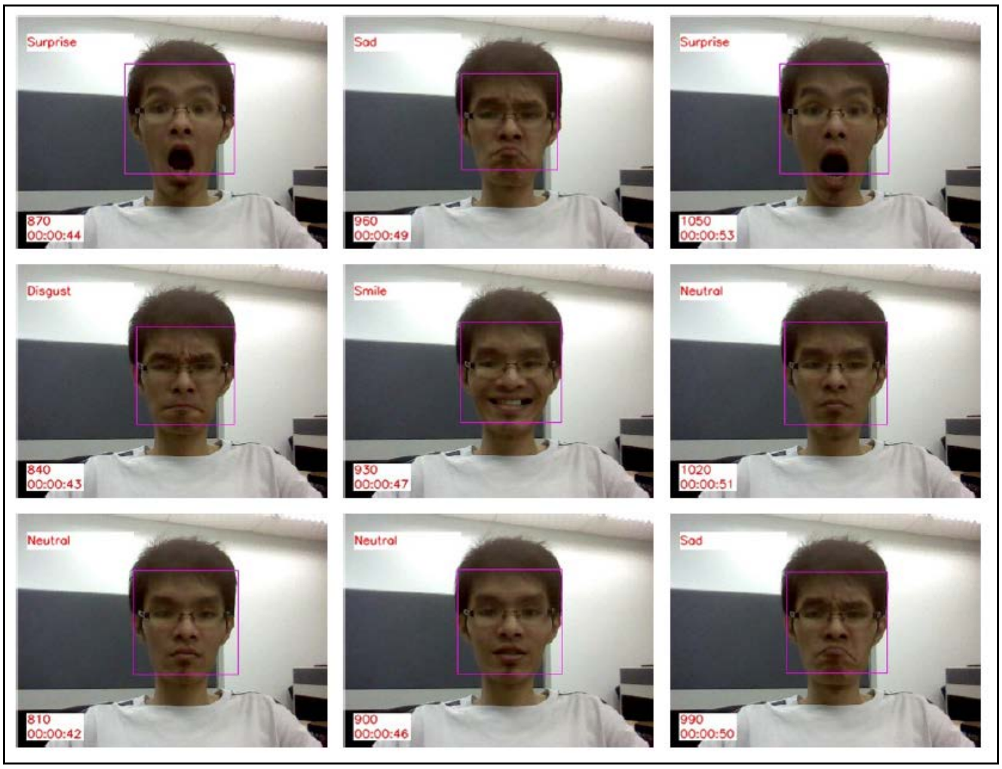


Figure 5-5 Sample of compiled frames with expression label and other information

Figure 5-6 shows the process of analyzing a frame, with each printed frame (positive detection) being classified as true positive frame or false positive frame. The true positive frame indicated that the face was correctly detected while the false positive frame indicated that a non-human face was detected, such as hands and background objects. The true positive frame was then classified into "match" or "not match". The frame that showed the correct label as expressed by the participant was marked as "match" while the frame that showed an incorrect label as expressed by the participant was marked as "not match".

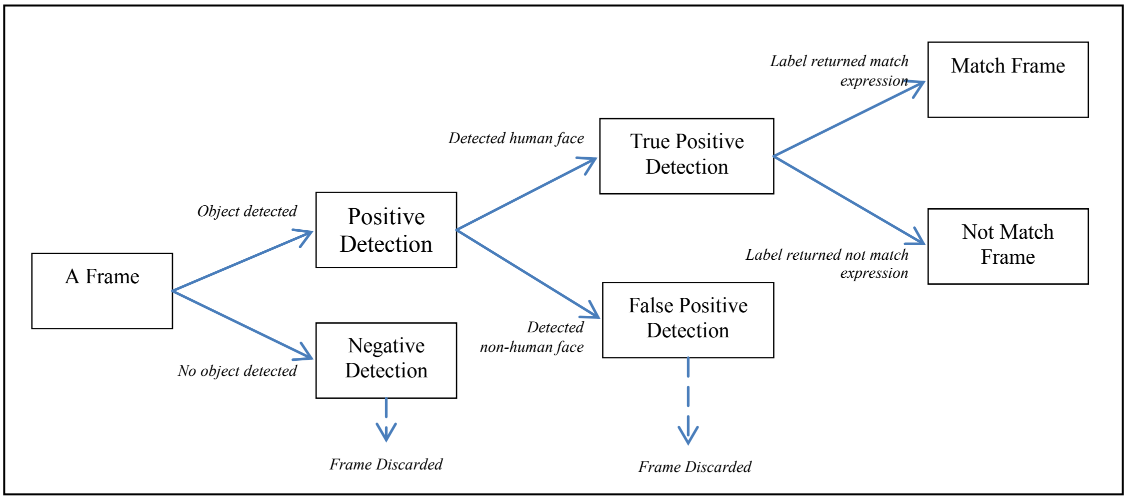


Figure 5-6 Frame analysis chart

5.3 Testing Result

Two sets of classified results for the high functional and low functional groups were produced. Each set of results produced was divided into two parts for the false positive detection rates and the recognition rates. The false positive detection rate indicated the number of frames detected which was not the human face, while the recognition rate indicated the result returned by ACApp that matched with the participant's expression in the true positive detected frame. The alert notification result was produced in another testing.

5.3.1 True and False Positive Detection Rate

The positive detection rate measured the accuracy of face detection by using ACApp. False positive detection means the object detected in a frame was not a face where as true positive detection indicated that a face object had been detected. The true and false positive results were obtained through the ground truth checking with the printed frames in the PDF file for each participant. This testing also checked on the low false positive detection rate to prove ACApps's effectiveness in detecting the face of children with Cerebral Palsy.

Table 5-4 Face detection results for the participants from the high functional group

High Functional Participant	Frames			Detection Accuracy	
	True Positive	False Positive	Total Positive Detected	True Positive Rate	False Positive Rate
1	1839	102	1941	94.74%	5.26%
2	1162	98	1260	92.22%	7.78%
3	1933	76	2009	96.22%	3.78%
4	752	120	872	86.24%	13.76%
5	1179	273	1452	81.20%	18.80%
6	2106	158	2264	93.02%	6.98%
7	2704	49	2753	98.22%	1.78%
8	2141	33	2174	98.48%	1.52%
9	1428	85	1513	94.38%	5.62%
10	1269	101	1370	92.63%	7.37%
11	1545	28	1573	98.22%	1.78%
12	1923	123	2046	93.99%	6.01%
13	2810	19	2829	99.33%	0.67%
14	2059	29	2088	98.61%	1.39%
15	1053	25	1078	97.68%	2.32%
Min	**752**	**19**	**872**	**81.20%**	**0.67%**
Max	**2810**	**273**	**2829**	**99.33%**	**18.80%**
Average	**1726.87**	**87.93**	**1814.80**	**94.35%**	**5.65%**
Standard Deviation				**5.02%**	

Table 5-4 shows the true and false positive detection results obtained from the 15 high functional participants. The average of the true positive detection rate obtained was 94.35% with the lowest being at 81.20% (Participant 5) and the highest at 99.33% (Participant 13). The average of false positive detection rate obtained was 5.65% with the lowest at 0.67% (Participant 13) and the highest at 18.80% (Participant 5).

The true and false positive detection results for the 6 low functional participants are listed in Table 5-5. The average true positive detection rate obtained from these participants was 81.78% with the lowest at 63.64% (Participant 3) and the highest at 96.35% (Participant 4) while the average false positive detection rate obtained was 18.22%, the lowest being 3.65% (Participant 4) and the highest being 36.36% (Participant 3).

Table 5-5 Face detection results for the participants from the low functional group

Low Functional Participant	Frames			Detection Accuracy	
	True Positive	**False Positive**	**Total Positive Detected**	**True Positive Rate**	**False Positive Rate**
1	518	102	620	83.55%	16.45%
2	251	50	301	83.39%	16.61%
3	168	96	264	63.64%	36.36%
4	686	26	712	96.35%	3.65%
5	301	82	383	78.59%	21.41%
6	115	20	135	85.19%	14.81%
Min	**115**	**20**	**135**	**63.64%**	**3.65%**
Max	**686**	**102**	**712**	**96.35%**	**36.36%**
Average	**339.83**	**62.67**	**402.50**	**81.78%**	**18.22%**
Standard Deviation				**10.67%**	

The overall false positive detection results showed that ACApp could detected faces better on the high functional group (5.65% false positive detection rate) compared to the low functional group (18.22% false positive detection rate). This was due to the inability of the low functional group to focus and to express their emotions in the short video clips viewing session. The standard deviations for the high functional and the low functional group in the false positive detection were 5.02% and 10.67% respectively.

5.3.2 Recognition Rate

This test could be used as the measurement of the effectiveness of ACApp in utilizing the facial expression of children with Cerebral Palsy to communicate with people around them. The recognition rate was obtained by the number of expressions returned by ACApp that matched with the participants' actual expressions.

Table 5-6 Matching results for high functional group

High Functional Participant	Set	Frames			Recognition Accuracy			
		Match	Not Match	Total	Match Rate	Not Match Rate	Average Recognition Rate	Average Not Recognition Rate
1	1	412	75	487	84.60%	15.40%	86.76%	13.24%
	2	475	38	580	81.90%	18.10%		
	3	724	28	772	93.78%	6.22%		
2	1	345	216	561	61.50%	38.50%	72.21%	27.79%
	2	199	48	247	80.57%	19.43%		
	3	264	90	354	74.58%	25.42%		
3	1	645	138	783	82.38%	17.62%	81.16%	18.84%
	2	374	92	466	80.26%	19.74%		
	3	553	131	684	80.85%	19.15%		
4	1	202	28	230	87.83%	12.17%	88.25%	11.75%
	2	178	28	206	86.41%	13.59%		
	3	286	30	316	90.51%	9.49%		
5	1	623	168	791	78.76%	21.24%	78.89%	21.11%
	2	95	25	120	79.17%	20.83%		
	3	211	57	268	78.73%	21.27%		
6	1	609	129	738	82.52%	17.48%	81.61%	18.39%
	2	449	115	564	79.61%	20.39%		
	3	665	139	804	82.71%	17.29%		
7	1	834	135	969	86.07%	13.93%	78.92%	21.08%
	2	620	173	793	78.18%	21.82%		
	3	683	259	942	72.51%	27.49%		
8	1	440	266	706	62.32%	37.68%	71.72%	28.28%
	2	549	108	657	83.56%	16.44%		
	3	539	239	778	69.28%	30.72%		
9	1	493	24	517	95.36%	4.64%	91.53%	8.47%
	2	307	42	349	87.97%	12.03%		
	3	513	49	562	91.28%	8.72%		
10	1	383	74	457	83.81%	16.19%	77.68%	22.32%
	2	276	106	382	72.25%	27.75%		
	3	331	99	430	76.98%	23.02%		
11	1	312	51	363	85.95%	14.05%	87.66%	12.34%
	2	393	54	447	87.92%	12.08%		
	3	655	80	735	89.12%	10.88%		
12	1	461	129	590	78.14%	21.86%	80.69%	19.31%
	2	457	137	594	76.94%	23.06%		
	3	643	96	739	87.01%	12.99%		
13	1	899	111	1010	89.01%	10.99%	81.57%	18.43%
	2	636	174	810	78.52%	21.48%		
	3	764	226	990	77.17%	22.83%		
14	1	618	88	706	87.54%	12.46%	85.91%	14.09%
	2	571	90	661	86.38%	13.62%		
	3	580	112	692	83.82%	16.18%		
15	1	221	8	229	96.51%	3.49%	89.92%	10.08%
	2	340	46	386	88.08%	11.92%		
	3	373	65	438	85.16%	14.84%		
Min		**95**	**8**	**120**	**61.50%**	**3.49%**	**71.72%**	**8.47%**
Max		**899**	**266**	**1014**	**96.51%**	**38.50%**	**91.35%**	**28.28%**
Average		**471.11**	**104.51**	**604.93**	**82.30%**	**17.70%**	**82.30%**	**17.70%**
Standard Deviation							**6.00%**	

Table 5-6 shows the recognition rate obtained from ACApp for the high functional group. An average recognition rate of 82.30% was obtained with the lowest being 71.72% (Participant 8) and the highest at 91.53% (Participant 9).

The recognition rate for the low functional group is listed in Table 5-7. The average recognition rate was 81.89%. The highest recognition rate was obtained by Participant 4 (90.49%) and the lowest recognition rate was obtained by Participant 3 (75.24%).

Table 5-7 Matching results for low functional group

Low Functional Participant	Set	Frames			Recognition Accuracy			
		Match	Not Match	Total	Match Rate	Not Match Rate	Average Recognition Rate	Average Not Recognition Rate
1	1	132	26	158	83.54%	16.46%	84.81%	15.19%
	2	127	14	141	90.07%	9.93%		
	3	177	42	219	80.82%	19.18%		
2	1	15	3	18	83.33%	16.67%	76.89%	23.11%
	2	14	7	21	66.67%	33.33%		
	3	171	41	212	80.66%	19.34%		
3	1	58	18	76	76.32%	23.68%	75.24%	24.76%
	2	25	9	34	73.53%	26.47%		
	3	44	14	58	75.86%	24.14%		
4	1	332	30	362	91.71%	8.29%	90.49%	9.51%
	2	91	11	102	89.22%	10.78%		
	3	201	21	222	90.54%	9.46%		
5	1	78	45	108	72.22%	27.78%	79.55%	20.45%
	2	61	8	67	91.04%	8.96%		
	3	95	29	126	75.40%	24.60%		
6	1	97	20	115	84.35%	15.65%	84.35%	15.65%
	-	-	-	-	-	-		
	-	-	-	-	-	-		
Min		**14**	**3**	**18**	**66.67%**	**8.29%**	**75.24%**	9.51%
Max		**332**	**42**	**362**	**91.71%**	**3.33%**	**90.49%**	24.76%
Average		**107.38**	**20.06**	**127.44**	**81.58%**	**18.42%**	**81.89%**	18.11%
Standard Deviation							**5.71%**	

The standard deviations obtained for both the high functional and low functional groups in the recognition rate were 6.00% and 5.71% respectively. This shows that the recognition rate obtained by each participant for both the low functional and the high functional groups were close to the mean value.

5.3.3 Alert Notification

The alert notification testing was conducted to measure the accuracy of ACApp in notifying users of any critical expression occurrences during the monitoring session. In this testing, the Alert Notification module was triggered when the flagged expression's label (critical expression) occurred during the monitoring session. For the sound notification, it would produce a beep sound every second when a critical expression occurred. As for SMS notification, it triggered GSM modem to send the critical expression's label in text messages to the preset mobile phone.

This testing was conducted to measure the success rate of SMS delivery and the waiting time when a critical expression occurred during the monitoring session. The results are shown in Table 5-8.

Table 5-8 Alert notification (SMS) testing result

Tests	SMS Sending Attempt	SMS Successfully Delivered	Successful Delivery Rate	Average waiting time (Seconds)
1	5	5	100.00%	8.40
2	13	12	92.31%	6.75
3	9	9	100.00%	7.89
4	6	6	100.00%	11.33
5	7	7	100.00%	7.00
Average	**8.00**	**7.80**	**98.46%**	**8.27**
Standard Deviation				**1.84%**

Throughout the test, an average of 98.46% SMS sending attempts were successfully delivered to the recipient's mobile phone at 8.27 seconds on average. The average waiting time for Test 4 was longer than normal (11.33 Seconds) while one attempt failed in Test 2. The standard deviation for the SMS waiting time was 1.84%.

5.3.4 Discussion

Table 5-9 shows the overall testing results which included false positive detection rate, recognition rate as well as alert notification rate. The false positive detection rate and the recognition rate results were further classified into high functional and low functional groups.

Table 5-9 Average and standard deviation for false positive detection rate, recognition and alert notification

Tests	Participants	Result	Standard Deviation
False positive detection rate	High Functional	5.65%	5.02%
	Low Functional	18.22%	10.67%
Recognition rate	High Functional	82.30%	6.00%
	Low Functional	81.89%	5.71%
Alert notification rate (SMS delivery waiting time)	Both	8.27 Seconds	1.84%

5.3.4.1 False Positive Detection Rate

The false positive detection rate shows that there was a 12.57% difference between the results produced by the two groups. The false positive detection rate for high functional group was 5.65% and 18.22% for the low functional group. This showed that ACApp could perform face detection better on the high functional group than with the low functional group. The high false positive detection rate for the low functional group was related to the participants' behavior during the On-Capturing Phase. This group had difficulty focusing on the short video clips. Hence, they did not show any expression asthey turned their heads around all the time. This was also related to the multiple disabilities that they had. A few of them ignored the videos by looking in the other direction and some were looking at the floor or ceiling.

A few of them got overly excited by swinging their hands all the time. Hence, their faces were covered by their hands, causing ACApp to detect their hands or other objects in the background. One special case happened on a low functional participant (Participant 6). He did not finish watching the short video clips. This participant lost his patience and moved away. Hence, only one set of video footages with the length of 11 minutes and 7 seconds was captured.

Generally, the false positive detection rate produced by both groups of participants (< 20%) was deemed acceptable.

5.3.4.2 Recognition Rate

The average recognition rate produced by both the high functional group and the low functional group were very close, with the high functional participants producing 82.30% and the low functional participants 81.89% (a difference of 0.41%). This result showed the ACApp had managed to detect and recognize the participants' facial expression based on the trained expression.

The testing was incomplete for a few low functional participants as they could not provide full cooperation due to their multiple disabilities. Due to time constraint, the researcher was unable to study their specific individual needs to design another method of conducting the testing.

Furthermore, the researcher had not been granted permission to conduct testing with participants outside their school premises. The environment at the school was less conducive for video recording due to the cluttered background and the participants getting distracted easily by their friends. The low quality video footages contributed to the poorer results in the test.

Regarding the time element, the testing processes (On-Capturing Phase and Post-Capturing Phase) for a participant could take up to several days and even weeks. The opportunities to conduct video recording relied on the participants' attendance at school and some of them were frequently absent. Thus, the testing with one participant may take up to a few weeks.

Due to the above mentioned problems encountered during testing, only 82.30% of the recognition rate was achieved.

5.3.4.3 Alert Notification

Generally, the results obtained in the alert notification showed that ACApp was suitable to be used as a communication tool for children with Cerebral Palsy to communicate with people around them. The sound notification was immediate, which meant that the recipient would be notified by the alarm (sound) whenever a flagged critical expression occurred. As for SMS notification, the recipient would be notified via SMS with only a few seconds delay compared with real time when it would take longer because of SMS sending time.

Throughout the alert notification evaluation, 98.46% of SMS sent was successfully delivered to the preset number of the recipient at 8.27 seconds of waiting time for each SMS. The guardians of children with Cerebral Palsy would get immediate notification from the children that needed assistance.

5.4 Summary

In this chapter, the evaluation results which were used to measure the effectiveness of ACApp in assisting with the communication of children with Cerebral Palsy were discussed. 21 children with Cerebral Palsy from an NGO school for special children were involved in the ACApp testing. The 21 children or participants were further divided into two groups which were the high functional group (11 boys and 4 girls) and the low functional group (4 boys).

A set of experimental procedure was defined in order to conduct the evaluation. The procedure involved three phases which were the Pre-Capturing phase, On-Capturing phase and also the Post-Capturing Phase, where each phase was involved in carrying out a few tasks.

In the Pre-Capturing phase, 20 out of 36 short video clips were selected which was used to trigger the participants' facial expressions. These short video clips, obtained from the youtube were reviewed by ten volunteers who rated the video clips according to the emotional elements contained.

In the On-Capturing phase, 61 video footages containing 21 participants' facial expressions were recorded while they were watching three sets of short video clips. Each participant went through three sessions where five short video clips were played in each session. All 61 video footages were used for ACApp testing in the Post-Capturing phase.

The Post-Capturing phase consisted of executing a few tasks which included footage video format conversion, footage previewing and labeling, footage trimming, expression training, expression recognition and frame printing, and lastly frame analysis for ground truth checking.

The testing results in the false positive detection rate, the recognition rate and the alert notification rate were discussed. The false positive detection result showed that ACApp could detect human faces effectively for the high functional group as the average false positive detection rate obtained was very low (5.65%). However the false positive detection result produced by the low functional group was higher (18.22%).

The expression recognition results obtained for both high functional and low functional groups were really close which was 82.30% and 81.98% respectively. On top of that, the results produced in the alert notification test had shown that 98.46% of the SMS sending attempts were successfully delivered to preset recipients at 8.27 seconds waiting time on average.

Chapter 6: Conclusion

This chapter concludes the entire research with four key issues, the research background and problems, the research findings and limitations, the research contributions, and future improvements.

6.1 Research Background and Problems

This research proposed an improved assistive communication model for children with Cerebral Palsy in assisting them to communicate with others. The research problems were identified throughout the background study on these children by literature review as well as by site visits.

Generally, children with Cerebral Palsy experience communication difficulties due to physical and speech disabilities which are caused before, during or one year after birth. Speech therapists and pathologists suggested for these children to use Augmentative and Alternative Communication (AAC) tools in assisting them to communicate with others. AAC tools are devices that seek to increase, maintain, or improve the functional capabilities of individuals with disabilities. In short, AAC is used for all forms of communication that is not speech, but is used to enhance or to replace speech.

Regrettably, recent researchers found that most of the AAC tools were less efficient as most of them were text to speech and touch screen based tools. Users were required to have fine motor skills and physical ability in order to use these AAC tools. Sadly, most of children with Cerebral Palsy who experience communication disability also experience physical disability.

Researchers also claimed that the AAC tools were not suitable for these children due to a lack of adaptability. Besides, some of these children who had lower intelligence also found it hard to operate these tools. Hence, a new trend or pattern of AAC was required in order to improve the efficiency of the existing AAC tools to help these children to communicate.

From this background study, two research problems were identified. 1) Children with Cerebral Palsy have difficulties in communication and; 2) The ineffectiveness of the existing AAC tools in assisting children with Cerebral Palsy to communicate.

A further research had found that the effectiveness of current and existing AAC tools could be improved by utilizing facial expressions (one type of biometrics information). Researchers claimed that the human face was rich with communication information. In addition, this information could be used to give feedback and responses. This statement was supported by statements and claims obtained during observation and site visits to a special school for children with Cerebral Palsy.

Hence, an assistive communication tool by utilizing human facial expressions (ACApp) was developed to overcome these research problems. Basically, ACApp was developed by adopting the basic facial expression recognition system and the recognition results would be used as alert signals to guardians on behalf of children with Cerebral Palsy by either displaying it on screen or notifying their guardians via sound and SMS.

This system was installed in a portable computer like a laptop with a built-in video camera and an external support device or GSM modem attached, for sending SMS purposes. The portable laptop would then be mounted on the wheelchair so that the child could be monitored at any time and at any place. During the monitoring session, the facial expression of each child would be captured and been distinguished by ACApp. A notification would be sent using the Short Messaging System (SMS) to the parent's or care taker's mobile phone when critical expressions were detected.

6.2 Research Findings and Limitations

Along the research journey, three major findings were achieved which included the experimental procedure of conducting facial the expression recognition test with children with Cerebral Palsy, the successfulness of utilizing facial expressions as a medium of communication for children with Cerebral Palsy as well as the overcoming of obstacles or limitations faced throughout the entire research.

The first finding in this research showed that it was impossible to conduct evaluation with children with Cerebral Palsy by using the common facial expression recognition testing method which utilized the standard template i.e. Cohn-Kanade databases and Japanese Female Facial

Expression (JAFFE) database. This was because these children had their own ways of expressing their emotions. Therefore, a series of experimental procedures were defined in this research which was specifically used as guidelines to run the expression recognition evaluation with children with Cerebral Palsy. By using the defined experimental procedure, the risk of test error and distraction was minimized, at the same time strengthening the evaluation results.

Second, the idea of utilizing facial expression as a medium of communication for children with Cerebral Palsy was shown to be working and could potentially replace the existing AAC tools proven through a series of prototype evaluations. 21 children with Cerebral Palsy from a special school were invited to participate in the evaluation. An overall result of 82.09% recognition accuracy (82.30% for high functional participants and 81.89% for low functional participants) has shown that ACApp could distinguish the children's facial expression and was capable of predicting these children's needs through their facial expressions. The additional Alert Notification module which obtained 98.46% of SMS successful delivery rate to presets recipient at 8.27 seconds waiting time on average has strengthened ACApp reliability. These children could attract their guardian's attention easily when they needed assistance during monitoring session.

The third finding was the limitations the research faced when conducting the same evaluation process. ACApp recognition accuracy could be up to 90% in this research if the limitations as follow could be resolved 1) the level of cooperation provided by participants, 2) time constraint and 3) environmental factors. The evaluation result was less convincing while dealing with low functional participants where they could not provide full cooperation while taking part in the evaluation. Due to time constraint, the aforementioned limitation was not resolved as the researcher was unable to make a further study on their needs and to design other ways to conduct separate tests individually. Other than that, the environment to conduct the research was limited as well. The researcher had not been granted permission to conduct evaluation or testing on participants outside their school premises. The initial environment in school was less conducive due to the cluttered background and the distractions to the participants. A re-test was unable to be conducted due to time constraint as the whole experimental process took several school days for each participant. On top of that, the opportunity to conduct testing was very much depending on participants' availability as some of them were frequently absent.

6.3 Research Contributions

Generally, the entire research has contributed to the community in various areas which include the idea of using biometrics information (facial expression) as a medium of communication for the disabled, the development of ACApp as assistive communication tools, the enhancement on face detection by adding Automated Brightness Adjustment and also the defined experimental procedure for conducting similar evaluation for children with Cerebral Palsy.

These contributions could be used as references for future improvement in any research fields which include information technology, social science, human communication, and also the welfare of children with Cerebral Palsy.

a. The idea of using biometrics information as a medium of communication for the disabled

The main contribution of this research was the proposal of utilizing biometric information as a medium of communication for the disabled or specifically children with Cerebral Palsy. Throughout literature review, existing Augmentative and Alternative Communication tools did not serve these children well due to their limitations. Hence, this research proposed an improved assistive communication tool for children with Cerebral Palsy to assist them in their communication. The evaluation conducted in this research has shown that biometric based assistive communication works in assisting children with Cerebral Palsy to communicate with others.

b. The assistive communication prototype

The second contribution of this research was related to the developed prototype, ACApp. This prototype worked as assistive communication tools for children with Cerebral Palsy and managed to distinguish different facial expression patterns expressed by specific children. The prototype could notify a third party i.e. guardians, when a critical expression occurred. Although ACApp has been developed and is functioning, the researcher would welcome suggestions for innovation and improvement on the current prototype (ACApp 4.0) in future research

c. Experimental Procedure

Another contribution of this research was related to the findings on the enhancement of face detection accuracy of the naïve Viola Jones face detection algorithm. The face detection enhancement could be done by employing "Automated Brightness Adjustment" on videos or specifically image sequences before performing the face detection using the Viola Jones face detection technique. This is useful for children with Cerebral Palsy who stay indoor most of the time. They are usually placed in the classroom and the house without much mobility.

As far as facial expression recognition is concerned, most researchers conduct expression recognition evaluation by training the system with standard templates or train sets i.e. Cohn-Kanade database, Japanese Female Facial Expression database– JAFFE and Yale Face database. However, this research has found out that the method stated above did not work well when conducting expression recognition evaluation on children with Cerebral Palsy, due to their disabilities. These children have limited ways to express their emotions and their facial pattern is unique and most of the facial expression recognition systems were unable to distinguish their expressions by training standard templates as stated above.

An experimental procedure for facial expression recognition evaluation on children with Cerebral Palsy was defined in this research. The evaluation on ACApp in this research was completed by following the defined experimental procedures and this has enabled the research to reach to this present stage. Thus, the experimental procedure can be used as a guideline for other similar researches, saving time and effort.

6.4 Future Improvement

Due to time and resource constraints, this research has some improvements for future works.

The first improvement is enhancing ACApp for mobile phone which has the same performance or even a better performance rate compare with the current version. It is beneficial as the smaller or lighters the device, the better it serves its purposes in terms of mobility and simplicity. However, a mobile version of ACApp in the current situation is impractical because of the reasons below:

1. Ordinary types of mobile phone are not capable of performing the task required by ACApp due to the limitation of the hardware processing ability
2. Most of the high end mobile phones i.e. iPhone and Android Phones are expensive and not economical for some families.

As for the second improvement, the researcher suggests an improved experimental procedure to evaluate facial expression recognition on the low functional group. Due to limitation of time and environmental constraints in this research, the evaluation results produced can be further improved. Further researches in detail on a better experimental procedure to conduct facial expression recognition with children with Cerebral Palsy will produce better result for the communication tool evaluation.

References

4MyChild 2009, *Types of Cerebral Palsy*, viewed 27 October 2009 <http://www.cerebralpalsy.org/types-of-cerebral-palsy/>.

AbleData 2010, *Augmentative and Alternative Communication Products: Communicators, viewed* 28 December 2010 <http://www.abledata.com/abledata.cfm?pageid=19327&top=10837&deep=2&trail=22,10825>.

Agrawal, N, Cosgriff, R, & Mudur, R 2009, 'Mood Detection : Implementing a facial expression recognition system', Stanford University, California. <http://www.stanford.edu/class/cs229/proj2009/AgrawalCosgriffMudur.pdf>.

Aires, KRT, Santana, AM, & Medeiros, AAD 2008, 'Optical flow using color information: preliminary results', *Symposium on Applied Computing archive. Proceedings of the 2008 ACM symposium on Applied computing*.

Aleksic, PS & Katsaggelos, AK 2006, 'Automatic facial expression recognition using facial animation parameters and multistream HMMs', *Information Forensics and Security, IEEE Transactions on*, vol.1, no.1, pp. 3- 11, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1597130&isnumber=33588>.

American Speech-Language-Hearing Assosiation (ASHA) 2010, *Augmentative and Alternative Communication* viewed 5 August 2010, <http://www.asha.org/slp/clinical/aac.htm>.

Ashraf, AB, Lucey, S, Cohn, JF, Chen, T, Ambadar, Z, Prkachin, KM et al. 2009, 'The painful face - Pain expression recognition using active appearance models', *Image and Vision Computing*, vol. 27, no. 12, pp. 1788-1796.

Attainment Company Inc. 2010, *GoTalk Express 32* viewed 8 January 2011 <http://www.spectronicsinoz.com/product/gotalk-express-32>.

Badawi, N, Watson, L, Petterson, B, Blair, E, Slee, J, Haan, E et al. 1998, 'What constitutes Cerebral Palsy?' *Developmental Medicine & Child Neurology*, vol. 40, no. 8, August 1998, pp. 520–527.

Balasuriya, LS, & Kodikara, ND 2000, 'Frontal view human face detection and recognition', BSc thesis, University of Colombo, Sri Lanka, May 2000, pp. 1–108.

Bartlett, MS, Littlewort, G, Fasel, I, & Movellan, JR 2003, 'Real time face detection and facial expression recognition: Development and applications to human computer interaction', *Computer Vision and Pattern Recognition Workshop,* vol. 5, p. 53.

Bezryadin, S, Bourov, P, & Ilinih, D 2007, 'Brightness Calculation in Digital Image Processing', *Technologies for Digital Fulfillment*, March 2007, <http://www.kweii.com/site/color_theory/2007_LV/BrightnessCalculation.pdf>.

Bouck, EC 2010, 'Technology and students with disabilities: Does it solve all the problems' in A Rotatori (ed), *Current Issues and Trends in Special Education: Research, Technology, and Teacher Preparation (Advances in Special Education, Volume 20)*, Emerald Group Publishing Limited, pp.91-104, viewed 20 December 2010 <http://books.google.com.my/books?hl=en&lr=&id=UmmZccoQRLcC&oi=fnd&pg=PA91&ots=K9M6ztpjar&sig=UKP4kshnVDfmlLgnu4K22G7ibcE#v=onepage&q=limitation&f=false>.

Branco, P & Encarnacao, LM 2004, 'Affective Computing for Behavior-Based UI Adaptation', *Proc. of Intelligent User Interface2004 Conf.*, Ukita, <http://www.research.ibm.com/iui-workshop/papers/BrancoEncarnacao.pdf>.

Brunelli, R 2009, *Front matter in template matching techniques in computer vision*, John Wiley & Sons, Ltd.

Busso, C, Deng, Z, Yildirim, S, Bulut, M, Lee, CM, Kazemzadeh, A et al. 2004, 'Analysis of emotion recognition using facial expressions, speech and multimodal information', *Proceedings of the 6th international conference on Multimodal interfaces*. State College, PA, USA: ACM.

Chai, TY, Rizon, M, Woo, SS & Tan, CS 2009, 'Facial features for template matching based face recognition', *American Journal of Applied Sciences*, vol. 6, pp. 1897-1901.

Chapman, E 2009, 'Using Alternative and Augmentative Communication to Facilitate Language Development in Children with Developmental Apraxia of Speech', *Speech Pathology and Audiology Capstone Anthology*, vol.1.

Chen, GY & Xie, WF 2007, ‘Pattern recognition with SVM and dual-tree complex wavelets’, *Image and Vision Computing*, vol. 25, pp. 960-966.

Chen, Q, Cordea, MD, Petriu, EM, Varkonyi-Koczy, AR, & Whalen,TE 2008, ‘Human computer interaction for smart environment applications using hand gestures and facial expressions’ *in the proc. of the 3rd IEEE international workshop on Medical Measurements and Applications*, Canada.

Chen, XW & Huang, T 2003, ‘Facial expression recognition: A clustering-based approach’, *Pattern Recognition Letters*, vol. 24, pp. 1295-1302.

Chibelushi, CC & Bourel, F 2002, ‘Facial expression recognition: a brief tutorial overview’, *CVonline: On-Line Compendium of Computer Vision*.

Cho, J, Mirzaei, S, Oberg, J, & Kastner, R 2009, ‘Fpga-based face detection system using Haar classifiers’, *Proceeding of the ACM/SIGDA international symposium on Field programmable gate arrays*. Monterey, California, USA: ACM.

Cohen, I, Grag, A, Huang, TS 2000, ’Emotion recognition from facial expression using multilevel HMM’, In: *Proceedings of NIPS workshop on affective computing*, Colorado.

Cohen, I, Sebe, N, Garg, A, Chen, LS, & Huang, TS 2003, ‘Facial expression recognition from video sequences: temporal and static modeling’, *Computer Vision and Image Understanding*, vol. 91, no. 1-2, pp.160–187.

Collier, B & Blackstone, S 2009, ‘Accessibility for People with Communication Disabilities’, *International Society for Augmentative and Alternative Communication (ISAAC)*, <http://www.isaac-online.org/en/MicrosoftWord_Accessibility_fact_sheet_for_home_page.pdf>.

Columbia Electronic Encyclopedia 2008, *Cerebral Palsy*, The Columbia Encyclopedia, 6th edn. 2008, Encyclopedia.com, viewed 19 January 2011 <http://www.encyclopedia.com/doc/1E1-cerebral.html>.

Cooper, RJ & Associates 2010, *Auggie (Augmentative Communication Device)* viewed 29 December 2010, <http://www.rjcooper.com/auggie/auggie/index.html>.

Crestwood Communication Aids Inc. 2009, *Product: Talk Back12* viewed 14 October 2009, <http://www.communicationaids.com/talkback12.htm>.

Dan 2008, *Face Detection on Digital Cameras: Empty Marketing or Must-Have Feature?* Viewed 2 October 2009, <http://www.cheapshooter.com/2008/09/18/facedetection-on-digital-cameras-empty-marketing-or-must-have-feature/>.

Datcu, D & Rothkrantz, L 2007, 'Facial expression recognition in still pictures and videos using active appearance models: A comparison approach', *Proceedings of the 2007 international conference on Computer systems and technologies*. Bulgaria: ACM.

Daugman, JD 1988, 'Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression', *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1169-1179.

Davis, AB, Moore.MM & Storey, VC 2003, 'Context-Aware Communication for Severely Disabled Users', *SIGCAPH Comput. Phys. Handicap.*, pp. 106-111.<http://delivery.acm.org/10.1145/960000/957224/p106-davis.pdf?key1=957224&key2=5474265921&coll=DL&dl=ACM&CFID=6979479&CFTOKEN=62365117>.

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

Deriche, M 2008, 'Trends and Challenges in Mono and Multi Biometrics', *Image Processing Theory, Tools and Applications, 2008. IPTA 2008*, 23-26 November 2008, pp.1-9.

Donato, G, Bartlett, MS, Hager, JC, Ekman, P, & Sejnowski, TJ 1999, 'Classifying Facial Action', in press: *Advance in neural Information Processing System 8*.

DTREG 2005, *Introduction to Support Vector Machine (SVM) Models*. [Online] viewed 27 December 2009, <http://www.dtreg.com/svm.htm>.

Dubuisson, S, Davoine, F & Masson, M 2002, 'A solution for facial expression representation and recognition', *Signal Processing: Image Communication*, vol. 17, pp. 657-673.

Duthoit, CJ, Sztynda, T, Lal, SKL, Jap, BT & Agbinya, JI 2008, 'Optical flow image analysis of facial expressions of human emotion: forensic applications', *Proceedings of the 1st international conference on Forensic applications and techniques in telecommunications, information, and multimedia and workshop*. Adelaide, Australia: ICST.

Dynavox 2010, *Product - DynaVox Maestro*.<http://www.dynavoxtech.com/products/maestro/>.

Ekman, P, & Friesen, WV 1978, 'Facial Action Coding System', *Investigator's Guide, Consulting Psychologists Press,* Palo Alto, CA.

Ekman, P 1997, 'Should we call it expression or communication?',*Innovation : The European Journal of Social Science*, vol. 10, pp. 333-344.

Farlex Inc. 2009, *Spastic Hemiplegia.* Mosby's Medical Dictionary, 8th edn, Elsevier.

Fasel, B, & Luettin, B 2003, 'Automatic facial expression analysis: a survey', *Pattern Recognition*, vol.36. pp. 259-275.

Fasel, B 2002, 'Multiscale Facial Expression Recognition Using Convolutional Neural Networks', *Tech. Rep*., IDIAP, 2002.

Feng, X, Pietikäinen, M & Hadid, T 2005, 'Facial expression recognition with local binary patterns and linear programming', *Pattern Recognition and Image Analysis,* vol. 15, no. 2, pp. 546–548.

Franco, L, & Treves, A 2001, 'A neural network facial expression recognition system using unsupervised local processing'. In: *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on, 2001*,pp. 628-632.

Fraunhofer Heinrich Hertz Institute 2009, *Face Detection in Images and Videos,* viewed 12 October 2009, <http://www.hhi.fraunhofer.de/en/departments/interactive-media-human-factors/overview/facedetection/?0>.

Gonzales, CH, Leroy, G, & Leo, GD 2010, 'Augmentative and alternative communication technologies' viewed 11 December 2010, <http://web.cgu.edu/faculty/leroy/content/Papers/2009-Heins-Gonzales-eHealth-chapter.pdf>.

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

Gu, H, Su, G, & Du, C 2003, 'Feature points extraction from faces', *Image and Vision Computing,* pp. 154-158.

Gull, H, Azam, F, Butt, WH, & Iqbal, SZ 2009, 'A new divide & conquer software process model', *World Academy of Science, Engineering and Technology*, vol. 60.

Harasse, S, Bonnaud, L & Desvignes, M 2005, 'Multiple faces tracking using local statistics', *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium*, 15-17 September 2005,pp.18-23.

Hodge, S 2007, 'Why is the potential of augmentative and alternative communication not being realized?', *Exploring the experiences of people who use communication aids. Disability & Society*, vol. 22, pp. 457-471.

Huang, X & Lin, Y 2008, ‘A vision-based hybrid method for facial expression recognition’, *Proceedings of the 1st international conference on Ambient media and systems (Ambi-Sys '08)*.ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, Article 4,p. 7.

Hyperbaric Services of the Palm Beaches 2009, *Hyperbaric Oxygen Therapy for Neurological Disorders* viewed 28 October 2009, <http://hbotxofpalmbeach.com/neuro.html>.

International Society for Augmentative and Alternative Communication (ISAAC) 2010, *AAC Technologies,* viewed 14 September 2010 <http://www.isaac-online.org/en/aac/technology.html>.

Jain, AK, Ross, A & Prabhakar, S 2004, ‘An introduction to biometric recognition’, *Circuits and Systems for Video Technology, IEEE Transactions*, vol. 14, pp. 4-20.

Jemaa, YB & Khanfir, S 2009, ’Automatic local Gabor features extraction for face recognition’, *International Journal of Computer Science and Information Security*, vol. 3, no. 1.

Kanade, T, Cohn, JF, & Tian, Y 2000, ‘Comprehensive database for facial expression analysis’, *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, vol. 4, Grenoble, 2000,pp. 46-53.

Karungaru, S, Fukumi, M, & Akamatsu, N 2004, ‘Face recognition using genetic algorithm based template matching’, *Communications and Information Technology, 2004. ISCIT 2004. IEEE International Symposium*, vol.2, 26-29 October 2004, pp. 1252-1257.

Kato, Y, Hirano, T & Nakamura, O 2007, ‘Fast template matching algorithm for contour images based on its chain coded description applied for human face identification’, *Pattern Recognition*, vol. 40, pp. 1646-1659.

Kepenekci, B 2001, Facial expression using gabor wavelet transform. Thesis, Middle East Technical University, <http://www.eee.metu.edu.tr/~alatan/PAPER/MSburcu.pdf>.

Kotsia, I, & Pitas, I 2007, ‘Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines’, *Image Processing, IEEE Transactions*, vol. 16, pp. 172-187.

Kovac, J, Peer, P, & Solina, F 2003, ‘Human skin color clustering for face detection’, *EUROCON2003*.

Krigger, KW 2006, ‘Cerebral Palsy: An Overview’, *American Family Physician*, 1 January, 2006, vol. 73, 1.

Kumar, BVKV, Savvides, M, Venkataramani, K, & Xie, C 2002, 'Spatial frequency domain image processing for biometric recognition', *Image Processing. 2002. Proceedings. 2002 International Conference*, vol.1, pp. 53-56.

Kyrki, V, Kamarainen, JK &, Kälviäinen, H 2004, 'Simple Gabor feature space for invariant object recognition', *Pattern Recognition Letters*, vol. 25, pp. 311-318.

Latecki, LJ 2005, 'Template Matching'. *Lecture Slide*. Temple University, Philadelphia.<http://www.cis.temple.edu/~latecki/Courses/CIS601-03/Lectures/TemplateMatching03.ppt>.

Lau, BT, Tran, CT & Chiong, R 2009, 'A real time communication assistant', in the *Proceedings of International Conference on Soft Computing and Pattern Recognition, IEEE Systems, Man and Cybernetics Society*.

Lee, B, Chun, J & Park, P 2008, 'Classification of facial expression using SVM for emotion care service system',*Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing. SNPD '08. Ninth ACIS International Conference*, 6-8 August 2008,pp. 8-12.

Lewis, J 2008.*SDLC 100 Success Secrets: Software Development Life Cycle (Sdlc) 100 Most Asked Questions, Sdlc Methodologies, Tools, Process and Business Models*. Emereo Publishing.

Lin, CH, & Wu, JL 1999, 'Automatic facial feature extraction by genetic algorithms', *Image Processing, IEEE Transactions*, vol. 8, pp. 834-845.

Lin, DT 2006, 'Facial expression classification using pca and hierarchical radial basis function network,' *J. Inf. Science and Eng*, vol. 22, pp. 1033–1046.

Littlewort, GC, Bartlett, MS & Lee, K 2007, 'Faces of pain: automated measurement of spontaneous all facial expressions of genuine and posed pain', *Proceedings of the 9th international conference on Multimodal interfaces*. Nagoya, Aichi, Japan: ACM.

Liu, S, & Silverman, M 2001, 'A practical guide to biometric security technology, *IT Professional*, vol 3, no. 1, pp. 27-32.

Lyons, MJ, Budynek, J, Plantey, A, & Akamatsu, S 2000, 'Classifying facial attributes using a 2-d gabor wavelet and discriminate analysis', *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000. IEEE Computer Society*.

Ma, L, & KhorasanI, K 2004, 'Facial expression recognition using constructive feedforward neural networks', *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, vol. 34, pp. 1588-1595.

Manohar, U 2008, *Why is Communication Important?*Buzzle.com , viewed 3 November 2009 <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*, vol. 16, pp. 555-559.

Mehrabian, A 2007, *A semantic space for nonverbal behavior in nonverbal communication*. Transaction Publisher.<http://books.google.com.my/books?hl=en&lr=&id=Xt-YALu9CGwC&oi=fnd&pg=PA1&dq=verbal+and+nonverbal+communication&ots=5wPbRc3ghq&sig=j9jFfw7j2wyzEN0d4LCsvkM36EA>.

Mencoder 2010, Mencoder Download, viewed 17 June 2010, <http://www.mplayerhq.hu/design7/dload.html>.

Michel P & Kaliouby REl 2008, 'Facial expression recognition using Support Vector Machines', *The 10th International Conference on Human-Computer Interaction*, Crete, Greece.

Michigan's Assistive Technology Resource 2002, *Voice Output Communication Devices Under $1,000*, viewed 14 October 2009, <http://www.cenmi.org/Portals/4/Documents/Materials/Voice_Output_Communication_Devices_Under_1000.pdf>.

MIMOS Berhad 2008, *National Biometrics Technology Roadmap*, Malaysia.

Ming, A, Ma, H, Zhang, H, & Xiong, W 2005, 'A Grid-based face recognition approach using general template matching', *Semantics, Knowledge and Grid, 2005. SKG '05. First International Conference*, 27-29 November 2005,p. 43.

Mobitek System 2008, *SMS Gateway Development Kit*. *Mobitek System,* viewed 24 February 2010, <http://www.mobitek.com.my/SMS_Gateway/SMS%20Gateway.html>.

Montillo, A, Metaxas D, & Axe, L 2004, 'Extracting tissue deformation using Gabor filter banks', *Proc. SPIE*, vol. 5369, 1.

Moon, H & Phillips, PJ 2001, 'Computational and Performance Aspects of PCA-Based Face Recognition Algorithms', *NIST Interagency/Internal Report (NISTIR) – 6486*, vol 30, pp.303 – 321.

Morris, C 2009, 'Definition & classification of Cerebral Palsy: a historical perspective', *Developmental Medicine and Child Neurology*, vol. 49, pp. 3-7.

Mu, X, Artiklar, M, Hassoun, MH & Watta, P 2001, 'Training algorithms for robust face recognition using a template-matching approach', *Neural Networks. Proceedings IJCNN '01. International Joint Conference, 2001*,vol.4, pp.2877-2882.

NaturalSoft 2010.*NaturalReader 10*. Viewed 22 April 2010, <http://www.naturalreaders.com/personal.htm>.

Nicolosi, L, Harryman, E, & Kresheck, J 2004, *Glossary of speech and language terms in definitions from terminology of communication disorders,* 3rd edn,<http://www.hwdsb.on.ca/instructionalservices/support/Hearing/G118-131.pdf>.

Niu, Z & Qiu, X 2010, 'Facial expression recognition based on weighted principal component analysis and support vector machines', *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference*, 20-22 August 2010, pp. 174-178.

Novita Children's Services 2006, *Alternative and Augmentative Communication (AAC)*.Viewed 23 October 2009 <http://www.novita.org.au/Content.aspx?p=64>.

Ong, CA, Lu, MV, & Lau, BT 2011, 'A Face Based Real Time Communication for Physically and Speech Disabled People', in BT Lau (ed), *Assistive and Augmentive Communication for the Disabled: Intelligent Technologies for Communication, Learning and Teaching*, Hershey: IGI Global, 2011, pp. 70-102.

Pantic, M 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: Cybernetics*, vol. 36, no. 2.

Park, CW & Park, M 2006, 'Fast template-based face detection algorithm using quad-tree template', *J. Applied Sci*., vol. 6, pp. 795-799.

Pennington, L, Goldbart, J, & Marshall, J 2004, 'Speech and language therapy to improve the communication skills of children with Cerebral Palsy', *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD003466.

Piatkowska, E 2010, 'Facial expression recognition system', Technical Reports (Master Thesis). Paper 17, College of Computing and Digital Media, DePaul University, <http://via.library.depaul.edu/tr/17>.

Prabhakar, S, Pankanti, S, & Jain, AK 2003, 'Biometric recognition: security and privacy concerns'. *Security and Privacy, IEEE*. pp. 33 - 42.

Price, JR & Gee, TF 2005, 'Face recognition using direct, weighted linear discriminant analysis and modular subspaces', *Pattern Recognition*, vol. 38, pp. 209-219.

Riaz, Z, Mayer, C, Beetz, M & Radig, B 2009, 'Model based analysis of face images for facial feature extraction', *Proceedings of the 13th International Conference on Computer Analysis of Images and Patterns*. Germany: Springer-Verlag.

Rose, A & Jain, A 2003, 'Information fusion in biometrics', *Pattern Recognition Letters*, vol. 24, no. 13, September 2003, pp.2115-2125.

Rosenbaum, P, Paneth, N, Leviton, A, Golstein, M, & Bax, M 2006, 'A Report : The definition and classification of Cerebral Palsy'.

Rummel-Hudson, R 2011, 'A Revolution at Their Fingertips', *Perspectives on Augmentative and Alternative Communication*, vol. 20, 19-23 April 2011.

Sebe, N, Lew, MS, Sun, Y, Cohen, I, Gevers, T & Huang, TS 2007, 'Authentic facial expression analysis', *Image and Vision Computing*, vol. 25, pp. 1856-1863.

Shan, C, Gong, S, & Mcowan, PW 2005, 'Robust facial expression recognition using local binary patterns', *Image Processing. ICIP 2005. IEEE International Conference*, 11-14 September 2005.

Shubhangi, G, Preeti, B & Ajith, A 2010, 'Emotion recognition system using connectionist models', <http://www.softcomputing.net/iccns09.pdf>.

Singh, SK, Chauhan, DS, Vatsa M, & Singh, R 2003, 'A robust skin color based face detection algorithm', *Tamkang J. Sci. Eng*. vol. 6, no.4 , pp. 227–234.

Smith, RD 2004a, 'Unit 1. Communication and Language', *Introduction to Language and Communication*, Undergraduate information syllabus, College of Humanities and Social Science, United Arab Emirates University, viewed 5 January 2011. <http://faculty.buffalostate.edu/smithrd/UAE%20Communication/Unit1.pdf>.

Smith, RD 2004b, 'Unit 2. Verbal and Nonverbal Communication', *Introduction to Language and Communication*, Undergraduate information syllabus, College of Humanities and Social Science, United Arab Emirates University, viewed 5 January 2011, <http://faculty.buffalostate.edu/smithrd/UAE%20Communication/Unit2.pdf>.

Softpedia 2009, *Y Speak Screenshots* viewed 20 October 2009, <http://mac.softpedia.com/progScreenshots/YSpeak-Screenshot-10550.html>.

Sommerville, I 2011, 'Chapter 2: Software Process' in I Sommerville (ed) Software Engineering, 9th edn. Pearson Publishing.

Special Needs Hope 2006, *What Are The Early Signs of Cerebral Palsy?,* viewed 14 October 2009 <http://www.special-needs-hope.com/signs-of-cerebral-palsy.html>.

Standup 2006, *The STANDUP System*. viewed 12 October 2009, <http://www.csd.abdn.ac.uk/research/standup/software.php>.

Tang, F & Deng, B 2007, 'Facial expression recognition using AAM and local facial features', *Natural Computation, 2007.ICNC 2007. Third International Conference*, 24-27 August 2007,pp. 632-635.

Tanova, C, & Nadiri, H 2010, 'The role of cultural context in direct communication', *Baltic Journal of Management,* vol 5, no. 2, pp 185 – 196.

Taylor, JS & Cristianini, N 2000, Support Vector Machines and other kernel-based learning methods', *Cambridge University Press*.

Tayntor, CB 2007, 'Chapter 16 : Introduction to the Traditional SDLC', in CB Taynotor (ed), *Six Sigma Software Development. Florida*, New York: Boca Raton, Auerbach Publications, cop. 2007, pp. 209-217.

Tian, YL, Kanade, T & Cohn, JF 2002, 'Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity', *Automatic Face and Gesture Recognition, 2002.Proceedings. Fifth IEEE International Conference,* vol. 21, pp. 229-234.

Valstar, M, Pantic, M, & Patras, I 2004, 'Motion history for facial action detection in video', *Systems, Man and Cybernetics, 2004 IEEE International Conference*, vol.1, pp. 635-640,<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1398371&isnumber=30409>.

Viola, P & Jones, MJ 2004, 'Robust Real-Time Face Detection', *International Journal of Computer Vision*, vol. 52, no. 2, 10January 2004, pp. 137-154.

Viola, P, Jones, MJ & Snow, D 2003, 'Detecting pedestrians using patterns of motion and appearance', *Proceedings of the Ninth IEEE International Conference on Computer Vision* –vol. 2. IEEE Computer Society.

Wall, H 2005, 'Context based algorithm for face detection', MSc thesis, Linkoping University, Sweden.

Wang, J, Plataniotis, KN, & Venetsanopoulos, AN 2005, 'Selecting Discriminate Eigenfaces for Face Recognition', *Pattern Recognition Letters,* vol. 26, pp. 1470–1482.

Wang, TH, & Lien, JJJ 2009, 'Facial expression recognition system based on rigid and non-rigid motion separation and 3D pose estimation', *Pattern Recognition*, vol 42, Issue 5, May 2009, pp. 962-977. <http://www.sciencedirect.com/science/article/pii/S0031320308004160>.

Whitehill, JR 2006, 'Automatic real-time facial expression recognition for signed language translation', MSc thesis, Universiti of the Western Cape.

Wilkins, J, & Ratajczak, A 2009, 'Developing Students' Literacy Skills Using High-Tech Speech-Generating Augmentative and Alternative Communication Devices', *Intervention in School and Clinic*, vol. 44, January 1, 2009, pp. 167-172.

Wu, J, Brubaker, SC, Mullin, MD, & Rehg, MJ 2008, 'Fast Asymmetric Learning for Cascade Face Detection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, March, 2008, pp. 369-382.

Xie, X, & Lam, KM 2009, 'Facial expression recognition based on shape and texture', *Pattern Recognition*, vol. 42, no. 5, May 2009, pp. 1003-1011.

Yale Medical Group 2011, *Communication Disorders*. Yale School of Medicine. Viewed 15 January 2011, <http://www.yalemedicalgroup.org/stw/Page.asp?PageID=STW026917>.

Yilmaz, A, & Shah, M 2002, 'Automatic feature detection and pose recovery for faces', *Proceedings of the Fifth Asian Conference on Computer Vision*, pp. 284–289.

Yusoff,FH, Rahmat, RWOK, Sulaiman, MN, Shaharom, MH & Majid, HSSA 2009, 'Tracking Facial Expression Using Template Matching and Optical Flow - A Relook at Looping Technique', *European J. of Scientific Research*, vol. 33, no. 2, July 2009, pp.313-329.

Zhan, C, Li, W, Ogunbona, P, & Safaei, F 2008, 'A real-time facial expression recognition system for online games.*ACM International Conference Proceeding Series*, vol. 207.

Zhan, Y & Zhou, G 2007, 'Facial expression recognition based on hybrid features and fusing discrete HMMs', *Proceedings of the 2nd international conference on Virtual reality*. Beijing, China: Springer-Verlag.

Zhan, YZ, Ye, JF, Niu, DJ & Cao, P 2004, 'Facial expression recognition based on Gabor wavelet transformation and elastic templates matching', *Image and Graphics. Proceedings Third International Conference*, 18-20 December 2004,pp. 254-257.

Zhang, GP 2000, 'Neural networks for classification: A survey', *IEEE Transaction on System, Man, and Cybernetics—Part C: Application and Reviews,* vol. 30, no. 4, November 2000.

Zhang, Y & Ji, Q 2005, 'Active and dynamic information fusion for facial expression understanding from image sequences' ,*Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol. 27, pp. 699-714.

Zhang, Z, Lyons, M, Schuster, M, & Akamatsu, S 1998, 'Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron', *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference*, 14-16 April 1998,pp. 454-459.

Zheng, W & Bhandarkar, SM 2009, 'Face detection and tracking using a boosted adaptive particle filter', *Journal of Visual Communication and Image Representation*, vol. 20, pp. 9-27.

Zhou, X, Collins, RT, Kanade, T & Metes, P 2003, 'A master-slave system to acquire biometric imagery of humans at distance', *First ACM SIGMM international workshop on Video surveillance*. Berkeley, California: ACM.

Zhu,M 2001, 'Feature Extraction and Dimension Reduction with Applications to Classification and the Analysis of Co-occurrence Data', Ph.D. dissertation, Stanford University.

Further Readings

Ong, CA & Lau, BT 2011, 'Biometrics based Assistive Communication Tools for Children with Special Needs', *Int. Proceedings of the 7th International Conference on IT in Asia*, 12 - 14 July 2011 Kuching, Malaysia, pp 250 – 255.

Ong, CA & Lau, BT 2011, 'A Study on the Effectiveness of Biometrics Based Alternative Communication Tool', *8th International Conference on Information, Communications and Signal Processing,* 13 – 16 December 2011, Singapore.

Ong, CA & Lau, BT 2012, 'Helping the disabled people: the potentials of biometric information', *Interactive Technology and Smart Education*.

Glossary

This abbreviated glossary covers only the most commonly encountered terms. There may be exceptions to some of the thesis definitions - The main concern of this glossary is to provide general concepts relative to commonly used. Abbreviations appear alphabetized at the beginning of each respective letter; only the most common abbreviations are included.

Term	Descriptions
ACApp	Assistive Communication Application
Assistive Communication Tool	Assistive tools to help children with Cerebral Palsy to communicate better.
Children with CP	People who experience disability caused by brain damage during or before birth or in the first years of their lives, resulting in a loss of voluntary muscular control and coordination.
Critical Expression	An expression expressed by children with Cerebral Palsy which require immediate attention from their guardians.
Detection Rate	A measurement unit used to measure human face detection on face detection system.
Expression Recognition	A mode performed by ACApp to read and distinguish human facial expression.
Expression Training	A mode used to train ACApp to recognize human facial expression.
Face Detection	A process to locate human face in a frame and extract the human face out .
Facial Expression	A unique pattern displayed on human face depending on initial emotion.
False Positive Detection Rate	Object detected in a frame was not human face.
FERS	Facial Expression Recognition System.
Frames	Captured image sequences.
Guardians	Parent, caretaker and teacher.
GUI	Graphical User Interface.
Image Processing	A process to refine extracted image.
Label	The name of facial expression.
Matching	The process of comparing two images and obtains the percentage of similarity.
Monitoring Session	A session where children with Cerebral Palsy is under ACApp monitoring using Expression Recognition mode. Any critical expression detected during this session will trigger ACApp's alarm either producing sound, send SMS or both.
OpenCV	An API integrated to develop ACApp.
Participants	ACApp's Experimental and Evaluation users.
Pre-Implementation	A stage before ACApp implementations or preparations.
Processed Image	Extracted and Processed Image

Prototype	Referring to ACApp.
Term	**Descriptions**
Recognition Rate	Outcome of Expression Matching Similarity
Standard template	The Readymade database e.g. JAFFE, Cohn Face, Yale etc…
Template Matching	A process of matching Unidentified Image with all Template Images stored
Templates Image	Stored Processed Image
Training Sample	Short Video Clips
Training Session	A session where ACApp is under training to recognize human facial expression.
True Positive Detection Rate	Object detected in a frame was human face.
Unidentified Image	Newly extracted and processed image from frame for recognition purpose
Video Footages	Recorded video files

Appendices

Appendix 1: Assent Form

<u>Child's Assent Form</u>

Introduction

Greetings, I am Ong Chin Ann, a postgraduate researcher from the School of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus and I would like to invite you to participate in this research evaluation. This research goal is to test on the effectiveness of information communication and technology to assist you in your communication by using an assistive communication tool developed by me. At the same time, this research is also contributing to the completion of my postgraduate studies (Master of Science). Your participation in this research is needed very much.

Procedures

You will be going through this experimental procedure if you agree to participate in this research.

1. You are required to watch 15 short movie clips in one session about 45 minutes during your recess or free time in your school.
2. You only need to react naturally while watching these short movie clips without any physical touch on any device or interface element.
3. Your facial expression will be captured using a built in camera during the session.

4. The captured footages will be evaluated and processed in the research lab in Swinburne University of Technology, Sarawak Campus.

Risks/Discomforts

You will only be involved in this research during your recess or free time in your school. You might miss a few minutes of your class time as the recess time is only 30 minutes while the capturing session requires about 45 minutes.

Benefits

You will be able to use this system to communicate through your facial expression with other peoples especially your parents, care takers and your teacher in future if this research is successful.

Confidentiality

All your personal information including your name, your pictures and your captured video will be kept and will remain confidential. I will not use your information as materials for any of my publication or presentation. Your information will be completely deleted and will not be use for other purposes at the end of this research.

Participation

Your participation contributes highly to this research. You are free to reject this invitation if you are uncomfortable with the experimental procedure. Even if you are already participating in this research, you are allowed to withdraw at anytime.

Question about the Research

You can ask and contact me at (+60)16-8084405 or email: cong@swinburne.edu.my anytime if you have questions or doubts related to this research. You can contact my immediate supervisor, Dr. Lau Bee Theng at (+60)19-8861016 or 082-260686, email: blau@swinburne.edu.my.

I have read, understood, and received a copy of the above consent form and I will of my own free will, participate in the study.

Signature: ____________________________ Date: ______________________

Name: _______________________________ IC Number:__________________

Appendix 2: Informed Consent

Informed Consent

Consent to Take Part in Research

(for parents of child with Cerebral Palsy)

Introduction

This research study is being conducted by Dr. Lau Bee Theng and Mr. Ong Chin Ann (postgraduate researcher) from the School of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus. This project also contributes to the completion of the postgraduate studies. We aim to study the effectiveness of information communication and technology to assist children in their communication. Your child is selected because he/she is currently studying at the PERKATA School.

Procedures

Participant with Cerebral Palsy having difficulty in both movement and verbal language:

We invite your child to participate in a real time behavior monitoring through facial expressions recognition using a laptop (with built in camera). During the evaluation sessions (approximately 45 minutes per session), he/she carries on with his classroom interaction or activities as usual. His/her facial expressions are detected naturally from where he is seated / placed at school. The accuracy of the detection will be recorded by the system and also by the researcher for ground truth verification. The automatic facial expression recognition operates continuously without any physical touch on any device or interface element by your child.

We will also provide you the demonstration on how to use our software to monitor critical expressions shown on the faces. Demonstrations of software are to be done at school as well. You can obtain a copy for usage upon request.

Risks/Discomforts

Your child will only be involved during his/her recess or free time during normal school hours. We will work closely with all teachers to make sure that research activities do not conflict with normal educational activities.

Benefits

It is hoped that the research will help educators find new channels to improve communication of your child.

Confidentiality

Your child's participation will be confidential. No child's name/pictures will be associated to the research materials and your child's name/pictures will never be used in connection with any presentation/publication of this research.

Participation

Participation is voluntary. If you give permission to include your child in the study, he/she will also be asked if he/she would like to participate. Even if you give consent, your child may withdraw at any time without penalty. Also, you may withdraw him/her at any time.

As a result of your participation, we will discuss with you on your child's response towards the usage of our software and we will be able to evaluate the effectiveness of using information, communication and technology to assist these children in their communication.

Question about the Research

If you have any questions concerning the study, please do not hesitate to contact the principal researcher, Lau Bee Theng at (+60)19-8861016 or 082-260686, email: blau@swinburne.edu.myor the postgraduate researcher, Ong Chin Ann at (+60)16-8084405, email: cong@swinburne.edu.my. You may also extend your queries or concerns to The Manager, Research and Consultancy, Assoc. Prof. Clement Kuek, 082-416353, email: ckuek@swinburne.edu.my.

I have read, understood, and received a copy of the above consent form and of my own free will allow my child to participate in the study.

Signature: ____________________________ Date: ______________________

Name: _______________________________ IC Number:__________________