EMOTION BASED MUSIC PLAYER

By

TAN SIEW CHING

Dissertation submitted in partial fulfillment of
the requirements for the
Bachelor of Technology (Hons)
(Business Information System)

SEPTEMBER 2012

UniversitiTeknologi PETRONAS
Bandar Seri Iskandar,
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Emotion Based Music Player

By

Tan Siew Ching

A project dissertation submitted to the
Business Information System Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirement for the
BACHELOR OF TECHNOLOGY (Hons)
(Business Information System)

Approved by,

____

(Assoc. Prof. Dr. Baharum B Baharudin)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
September 2012

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the reference and acknowledgements, and that the original work contained herein has not been undertaken or done by unspecified sources or persons.

(TAN SIEW CHING)

ABSTRACT

The work presents described the development of Emotion Based Music Player, which is a computer application meant for all type of users, specifically the music lovers. Due to the troublesome workloads in songs selection, most people will choose to randomly play the songs in the playlist. As a result, some of the songs selected not matching the users' current emotion. Moreover, there is no commonly used music player which able to play the songs based on user's emotion. The proposed model is able to extract user's facial expression and thus detect user's emotion. The music player in the proposed model will then play the songs according to the category of emotion detected. It is aimed to provide a better enjoyment to music lovers in music listening. The scope of emotions in the proposed model involve normal, sad, surprise and happy. The system involves the major of image processing and facial detection technologies. The input for this proposed model is the .jpeg format still images which available online. The performance of this model is evaluated by loading forty still images (ten for each emotion category) into the proposed model to test on the accuracy in detecting the emotions. Based on the testing result, the proposed model has the Recognition Rate of 85%.

ACKNOWLEDGEMENT

First and foremost, the writer would like to take this opportunity to express her deepest gratitude and appreciation to project main supervisor, Assoc. Prof. Dr. Baharum B Baharudin, who continuously monitoring the progress of the project throughout the semesters. The constructive comments, advices, and suggestions given have leaded to the successful outcome of the project.

This gratitude also dedicated towards the individuals who participated in the survey. The author highly appreciates for the feedbacks and cooperation given by the participants which have eventually facilitate a lot in developing, improving and implementation of the system prototype.

Last but not least, the author would like to express her acknowledgement to family members and friends, for the support and verbal constructive opinions throughout the seven months.

Thank you very much.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

FYP	Final Year Project
RR	Recognition Rate

CHAPTER 1
INTRODUCTION

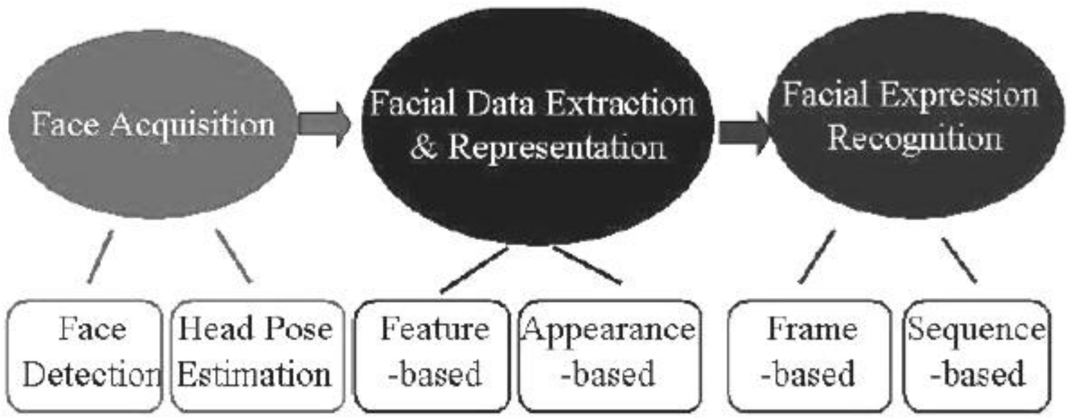
1.1 BACKGROUND OF STUDY

Emotions are the bodily feelings associated with mood, temperament, personality or character. Paul Ekman had developed the classifications of basic emotions which are anger, disgust, fear, happiness, sadness and surprise in 1972.

A facial expression can be expressed through the motions or from one or more motions, movements or even positions of the muscles of the face. These movements transmit of the emotional status of an individual. Facial expression can be adopted as voluntary action as individual can control his facial expression and to show the facial expression according to his will. For an example, a person can make the eyebrow closer and frown to show through the facial expression that he is angry. On the other hand, an individual will try to relax the face's muscle to indicate that he is not influence by the current situation. However, since facial expression is closely associated with the emotion, thus it is mostly an involuntary action. It is nearly impossible for an individual to insulate himself from expressing the emotions. An individual may have a strong desire or will to not to express his current feelings through emotions but it is hard to do so. An individual may show his expression in first few micro-second before resume to a neutral expression.

Since the work of Darwin in 1872, the behavioral scientists had actively involved in the research and analysis of facial expression detection. In 1978, Suwa et al. presented his early attempt on the idea of automatically facial expressions analysis by tracking the motion of twenty identified spots on an image sequence. After Suwa's attempt, there are lots progresses in developing the computer systems in order to help human to recognize and read the individual's facial expression, which is a useful and natural medium in communication.

Facial expression analysis includes both detection and interpretation of facial motion and recognition of expression. The three approaches which enabled the automatic facial expression analysis (AFEA) include i) face acquisition, ii) facial data extraction and representation, and iii) facial expression recognition.



Source from Ying-Li Tian, Takeo Kanade , and Jeffrey F. Cohn (2003)

Figure 1.1: Basic structure of facial expression analysis systems

The "Emotion Based Music Player" is a device developed aimed to detect the emotion of an individual, and play the lists of music accordingly. First, the individual will reflect his emotion through the facial expression. After that, the device will detect the condition of the facial expression, analyze it and interpret the emotion. After determined the emotion of the individual, the music player will play the songs which can suit the current emotion of the individual. The device will focus on the analysis of the facial expression only which does not include the head or face movement.

1.2 PROBLEM STATEMENT

The significance of music on an individual's emotions has been generally acknowledged. After the day’s toils and hard works, both the primitive and modern man able to relax and ease him in the melody of the music. Studies had proof that the rhythm itself is a great tranquilizer.

However, most people facing the difficulty of songs selection, especially songs that match individuals’ current emotions. Looking at the long lists of unsorted music, individuals will feel more demotivated to look for the songs they want to listen to. Most user will just randomly pick the songs available in the song folder and play it with music player. Most of the time, the songs played does not match the user’s current emotion. For an example, when a person is sad, he would like to listen to some heavy rock music to release his sadness. It is impossible for the individual to search from his long playlist for all the heavy rock music. The individual would rather choose the songs randomly or just “play all” for all the songs he had.

Besides, people get bored with this traditional way of searching and selecting songs. The method had been implemented since few years back.

1.3 PROJECT OBJECTIVE

1.3.1 General Objective

The main objective of this project is to develop the “Emotion Based Music Player” for all kinds of music lovers which aimed to serve as a platform to assist individuals to play and listen to the songs according to his emotions. It is aimed to provide a better enjoyment of entertainment to the music lovers.

1.3.2 Specific Objective

The Specific Objective for this project is specified as below:

i. To propose a facial expression detection model to detect and analyze the emotion of an individual.
ii. To accurately detect the four basic emotions, namely normal, happy, sad and surprise.
iii. To integrate the music player into the proposed model to play the music based on the emotions detected.

1.4 SCOPE OF STUDY

Currently there is no commonly used application or system which able to detect the emotion of individual and play music according to the emotion detected. This system will propose a new lifestyle to all music lovers which will ease them when searching for playlists. The target users will be the music lovers. English will be the main medium of language used in the proposed model and specifically aimed to detect some basic emotion such as normal, happy, sad or surprise. The evaluation of this system will base on the accuracy in detecting the correct facial expression as well as playing the right category of songs.

The scope of study will be as follow:

i. Study on the different method in expression detection.
With the improvement of technology in image processing, more and more experts did researches or introduced different technique in processing a specific area or small area on an image. All these techniques can be applied to the facial expression processing. Researches had to be done in order to understand each technique which will then be useful in the project development.

ii. Get information on the tools appropriate for the facial expression detection in order to build the proposed model for this project. Different tools (software and hardware) are studied on their feasibility and functionality as well as user-friendliness in order to figure out the most suitable and applicable tools to develop it.

1.5 FEASIBILITY STUDY

1.5.1 Within the scope frame

As stated above, the focus of this project will be entirely on the detection of facial expression and integrates it to the music player. As a prototype, the proposed model will detect only the basic emotion such as happy, sad, normal, and etc.

To understand the scope of the project in depth, massive research needs to be done in order to figure out the current technology in facial expression detection and the used of the technology. The studies include as well the technical aspect, especially the programming languages used in writing image processing and facial expression detection. Few analyses need to be done in order to find out the most user friendly programming language and application in order to develop this system. This is so that the development of the system is within the timeline given.

1.5.2 Within the time frame

The Final Year Project (FYP) course is divided into two parts, which are the FYP1 and FYP2. As in the syllabus given, FYP1 will focus more on the brain storming for FYP title, proposal writing, data gathering, and researches well as the report writing. On the other hand, the development of prototype, implementation and testing of the developed proposed model will be done in FYP2.

As the phase has been divided evenly between the two semesters which is equivalent to eight months, the project will be able to finish on time with the proper time management.

CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

In the year of 2009, Barbara Raskauskas had published an article stating the music is one of the widely accepted culture and language which can be accepted by any type of people. She mentioned that "music does fill the silence and can hide the noise. Music can convey cultural upbringing. Music is pleasurable and speaks to us, whether or not the song has words. I've never met a person who didn't like some form of music. Even a deaf friend of mine said she liked music; she could feel the vibration caused by music. Finding enjoyment in music is universal."

Emily Sohn (2011) stated that “People love music for much the same reason they're drawn to sex, drugs, gambling and delicious food, according to new research”. Through the actions and activities carried out by the people around, the statement mentioned is widely accepted by the public. Study had proved that human brain will release dopamine, a kind of chemical generated by body which involved addiction and motivation when people listen to harmony or melody that touch an individual.

Comparison with similar expression can be done in order to detect the facial expression of an individual. In the year of 2005, Mary Duenwald had published an article which summarizes that scientists had did several studies and researches and shown that facial expressions across the globe fall roughly into seven categories:

i. **Sadness**: The eyelids droop while the inner corners of the brows rise. When in extreme sadness, the brows will all push nearer together. As for the lips, both of its corners pull down and the lower lip may push up in a mope.

ii. **Surprise**: Both the upper eyelids and brows rise, and the jaw drops open.

iii. **Anger**: Both the lower and upper eyelids squeeze in as the brows move down and draw together. The jaw pushes forward, the upper and lower lip press on each other when the lower lip pushes upper a bit.

iv. **Contempt**: The expression appears on one side of a face: One half of the upper lip tightens upward.

v. **Disgust**: The individual’s nose wrinkles and the upper lip rise while the lower lip protrudes.

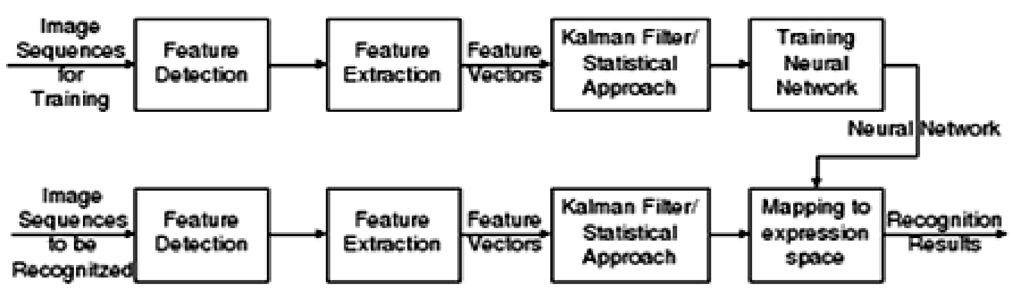
vi. **Fear**: The eyes widen and the upper lids rise. The brows draw together while the lips extend horizontally.

vii. **Happiness**: The corners of the lips lifted and shaped a smile, the eyelids tighten, the cheeks rise up and the outside corners of the brows pull down. [5]

2.2 FACIAL EXPRESSION DETECTION RESEARCH

Face features detection such as the mouth and the eyes is always one of the key issues in facial image processing as it involves wide and various areas such as the emotion recognition and face identification. Joseph C. Hager (2003) stated that face detection feature is used as one of the input to other image processing functions such as the face and emotion detection. Different researchers had studies on the different approaches in facial expression detection. Each approach can be applied effectively in different situation.

In the year of 2004, W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat had proposed a method of combining the feature detection and extraction with the facial expression recognition into an integrated system which can improve the recognition output in terms of the recognition process. With this system, the recognition process is not influenced by the subjective aspects and the bound of the areas are invariant during the sequence. W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat (2004) "We propose a method for facial expression recognition that uses integral projection, statistical computation, a neural network and kalman filtering. The face feature detection method uses multi-stage integral projection. Optical flow computation will be used on the detected feature namely, the eyebrows and the lip to extract movement."



Source from W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat. (2004)

Figure 2.1: Block diagram of the proposed facial expression recognition system

The integration projection enables the detection and location of small and precise features of the face, such as the eyebrows and lips. When the statistical approach on the optical flow field is applied, the overall movement of the features can be detected without the need to pinpoint the exact location of the features. The usage of this approach does not require the predefine setting such as the location of the head or the eyes. The predefine settings are preset using the normalized coefficients obtained using a facial image database.

Apart from the approach introduced by W.K. Teo and etc., in the year of 2010, Jagdish Lal Raheja and Umesh Kumar had introduced the Back Propagation Neural Network technique in human facial expression detection from captured image. The approach used is based on the usage of add-boosted classifier and finding and matching the token when detecting the facial expression while applying neural network. In face detection, the method proposed is by usign the Viola and Jones method. It is a better implementation when comparing to other techniques as it is feature based. Besides, it is able to perform the analysis relatively faster as compared to others. Edge detection, thinning, and token detection are carried out during the image processing process. Edge detection is aimed at identifying the points in a digital image at which the image brightness changes sharply or more formally has discontinuities. Thinning is applied in order to reduce the width of an edge, which is from multiple lines to single line. Token which generated after the thinning process divides the data set into smallest unit of information which needed for the following processes. After the three procedures, the recognition is performed. “It is a tedious task to decide the best threshold value to generate the tokens. So as a next process or the future work is to determine the best threshold value, so that without the interaction of user the system can generate the tokens.”

Besides above, Zhengyou Zhang (1998) had reported on the investigation on the feature-based facial expression recognition within an architecture based on the two-layer preceptron. Two types of factors are being derived from the face images during the investigation, i.e. the geometric positions of a set of fiducial points on a face as well as the set of set of multi-scale and multi-orientation Gabor wavelet coefficients at these points. Zhengyou Zhang (1998) “The recognition performance with different types of features has been compared, which shows that Gabor wavelet coefficients are much more powerful than geometric positions.” Secondly, the sensitivity of individual's fiducial point to the facial expression detection is examined. Through the sensitivity analysis, the author found out that the points on cheeks and forehead carry little useful information. Lastly, the author studied the importance of image scales in the facial detection process. The experiments show that facial expression recognition is mainly a low frequency process, and with a spatial resolution of 64 pixels X 64 pixels is probably enough to carry the process out.

Apart from that, in 2007, Eva Cerezo, Isabelle Hupont, Critina Manresa, Javier Varona, Sandra Baldassarri, Francisco J. Perales, and Francisco J. Seron presented their works on an automated real-time system for facial expressions recognition which functioned by tracking the facial features’ and the simple emotional classification method. The automatic feature extraction function enables the preamble of dynamic information in the classification system, which making the study of time evolution on the evaluated parameters as well as the categorization of user’s emotions through live video possible. The developed system had been embedded in the Maxine system, an engine developed by the group for managing 3D virtual scenarios and characters to enrich user interaction in different application domains to test its usefulness and real-time operation.

In addition, Caifeng Shan, Shaogang Gong, Peter W. McOwan (2008) reported on the Facial expression recognition based on Local Binary Patterns: A comprehensive study. Local Binary Patterns is the texture operator that tags the pixels of an image by thresholding the neighborhood’s pixel and represents the result in binary format. Result from the authors’ experiments show that Local Binary Patterns features perform stably and robustly over a useful range of low resolutions of face images. This means image with low resolution can as well be processed and accurately identified the emotion.

2.3 RELATIONSHIP BETWEEN MUSIC AND EMOTION RESEARCH

Many researchers had did research and studies on if the music can actually influence the emotion of individuals. Throughout the years, the results from the studies proved that different music style can actually influence individuals in different ways. For an example, in the year of 1994, Antoinette L. Bouhuys, Gerda M. Bloem, Ton G.G.Groothuis carried out a study in the relationship between the individuals' facial expression after listening to depression music. The results showed that depressing music bring on a major increase of depressed mood and significant decline if delighted mood. The study proved that music can actually influence individuals' emotions.

Besides the study mentioned above, Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou presented on their research in the used of music in manipulating the young tennis players' emotional state. A total of fourteen young tennis players were involved in this study. The research signifies that participants will often choose to listen to music in order to elicit various emotion states, increased arousal, and visual and auditory imagery. "Increasing the tempo and/or intensity of a musical excerpt may increase the magnitude of an effective response and concomitant action tendencies" (Frijda, 1986) such as increased motor behavior.

In 2001, Matthew Montague Lavy develops four basic assumptions regarding music lovers and their relationship to music. First, music is heard as sound. The constant monitoring of auditory stimuli will not be switched off once an individual listen to music but it will monitor and analyzed the music just like any other stimuli. Secondly, music is heard as human utterance. Everyone has the ability to identify and detect emotion in the contours and timbres of vocal utterances. Third, music is heard in context. Music is described as a wide and complicated network of knowledge, thoughts and environment. All these are the factors which contribute to an emotional experience. Forth, music is heard as narrative. Listening to music includes the integration of sounds, utterances and context.

2.4 EMOTION BASED MUSIC RETRIEVAL RESEARCH

Besides the study mentioned above which shows that emotions can be influenced by music, Wai Ling Cheung and Guojun Lu (2008) presented that automatic music emotion annotation is an important requirement to research music retrieval by emotion. Music emotion annotation is the task of embedding an emotional terms to musical terms. This research proposed a solution which automates a traditionally manual annotation task using a number of techniques from various disciplines, which is highly original. The author pointed out that through this research, the automatic music emotion annotation is possible and workable using hybrid sampling, data-driven detection threshold and synonymous relationships between emotional. "Our empirical result shows that training data size requirement is within reach for a workable annotation system. As music emotion description becomes readily available through automatic annotation, the development of a music research repository will be more attainable."

Maria M. Ruxanda1, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen (2009) studied on a number of dimensionality reduction algorithms, including both classic and novel approaches to test on the performance of applying multidimensional indexing over a dimensionally reduced audio feature space in audio extraction. "The proposed approach comprises of an effective methodology that projects the music into an audio feature space that captures the music emotion."

CHAPTER 3
METHODOLOGY

3.1 RESEARCH METHODOLOGY

Methodology is an essential element in the software development process. Methodology acts as a means of risk management during the different stages and processes of software development. It aims to improvise the management and control of the software development life cycle (SDLC). There are many different categories of project development methodologies and each one has its own unique structure to develop the project based on SDLC phase.

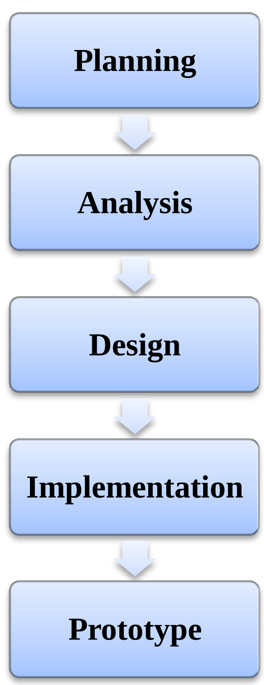


Figure3.1: The Waterfall Methodology

The methodology used to complete this project is a structured design. The waterfall model is a sequential design process, in which progress is seen to be flowing steadily downwards from the planning, analysis, design, implementation and maintenance phase. The project is done following one stage to the other from the first to the last stage.

First of all, during the planning stage, brain storming of ideas is carried out in order to find out the suitable field to focus on for FYP project. Few researches are carried out to determine the needs of innovation of systems in the specific field. After specified the field, some critical thinking is done in finding the problem arising from the field. Discussion is held with lecturers to determine the most suitable and executable project title.

The project is then moves to the analysis phase as we conduct a background studies, researches and data gathering process. All the data collected will be analyzed. The data consist of two main areas which are paper based research (e.g. current status of image processing, application of image processing and facial detection etc.) as well as the technical aspects (e.g. technology to be used in developing the system). Learning of new programming languages will start in this stage.

Once the analysis part is done, the project moves to the design phase, in which the analysis models as well as the interface design for the system will be determined. The development process starts with the design of the framework and interface of the system. The major focus in design phase is to write program in order to detect the facial expression. After developing the facial expression detection system, the system will be integrated with the music player.

Finally when the design of the project is done, the proposed model will then be tested to find out if there any bugs and to test its functionality and accuracy in facial detection for various emotions.

3.2 PROJECT ACTIVITIES

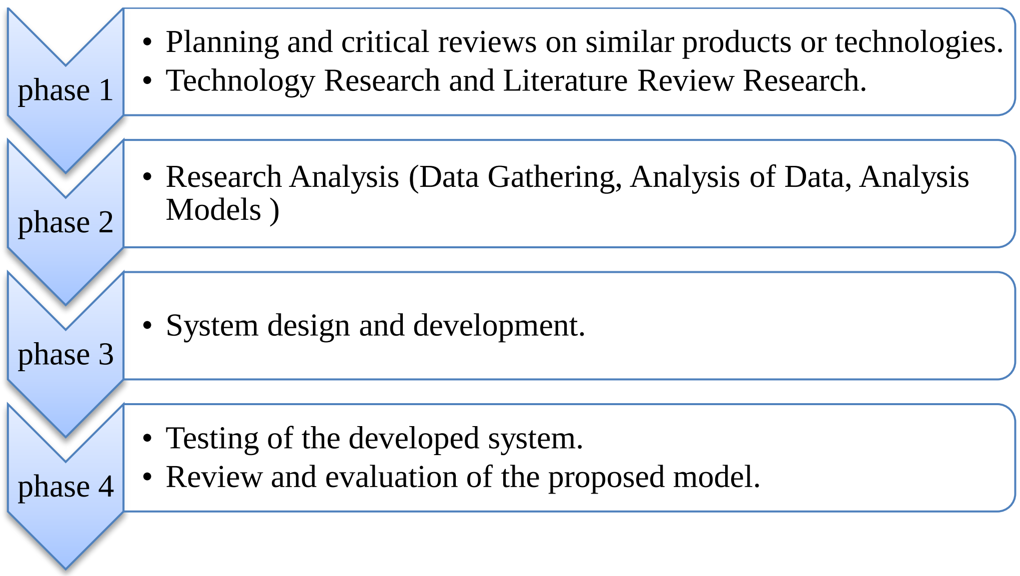


Figure 3.2: Project Activities

The project mainly consists of four stages. It starts with the planning and critical review on similar products and technologies available. In this stage, several of image processing and facial expression detection technologies are studied and analyzed.

In phase 2, the activity is then focuses on the research analysis. This includes data gathering process, the analysis of data as well as the development of analysis models for the system.

In Phase 3, the project will focus on the development of the system. The different functions of the system is developed and tested accordingly. This is to ensure the system able to detect and analyze the basic facial expression accurately.

The project will end with the last phase which is Phase 4, which is the review and evaluation of the system developed. The proposed model will be tested over different types of images to measure its accuracy in detecting the facial expression.

3.3 KEY MILESTONES

Some of the key milestones for this project are as shown in the table below:

Table 3.1: The Key Milestones of Final Year Project.

No	Deliverable / Activities	Schedule	
1	Title Selection	Week 1	FYP1
2	Briefing to Students and submission of proposal to the research cluster	Week 3	FYP1
3	Submission of extended proposal	Week 6	FYP1
4	Viva Presentation: Proposal defense and progress evaluation	Week 9	FYP1
5	Submission of interim report	Week 11	FYP1
6	Progress Update and Submission of Progress Report	Week 4	FYP2
7	Pre-Sedex Presentation	Week 11	FYP2
8	Dissertation	Week 11	FYP2
9	Viva Presentation	Week 12	FYP2
10	Final Dissertation	Week 14	FYP2

3.4 GANTT CHART

3.4.1 FYP1

Table 3.2: Gantt Chart for FYP1.

Title	Week											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Project title selection/proposal	■											
Proposal submission to research cluster		■	■									
Research on the current technology				■	■							
Extended proposal submission						■						
Data Collection: Images on different expressions from different individuals							■	■				
Data Collection: Collection of different genre of music/songs									■			
VIVA: Proposal defense and progress evaluation										■		
Prepare: Interim Report											■	
Submission of Interim Report												■

3.4.2 FYP2

Table 3.3: Gantt Chart for FYP2.

Title	Week													
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
System Development	■	■	■	■	■	■	■	■	■	■	■	■		
Progress Update and Submission of Progress Report				■										
Pre-Sedex Presentation											■			
Dissertation											■			
VIVA Presentation												■		
Final Dissertation														■

3.5 DEVELOPMENT TOOLS

- Software

i. Microsoft Visual C# 2010 Express

  - Main development software which includes interface design and coding.

ii. Microsoft Access Database:

  - To store the database for the facial expression features data.
  - To store the data extracted for each image processed.
  - To store the song lists

- Hardware

i. Personal Laptop (Dell Inspiron N4010)

  - Intel® Core™ i7-460M CPU @2.53GHz
  - 4Gb RAM

ii. Build in webcam/External Webcam

  - To capture the user's facial expression.

- Programming Language

i. C#

- Operating System

i. Windows 7 Home Premium

CHAPTER 4
RESULTS AND DISCUSSION

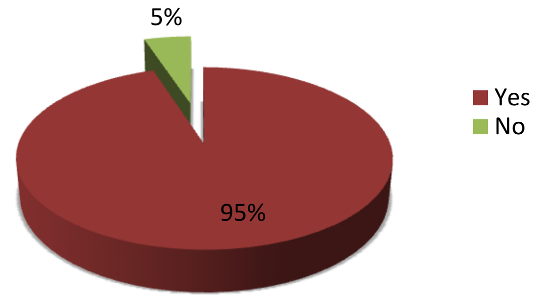
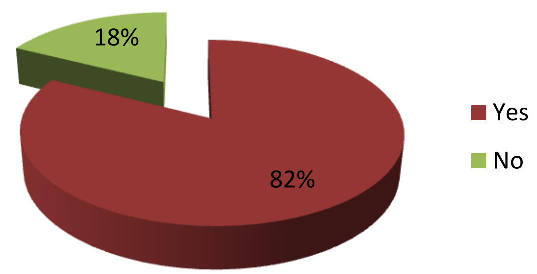
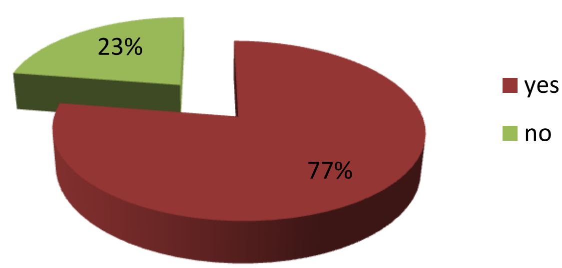
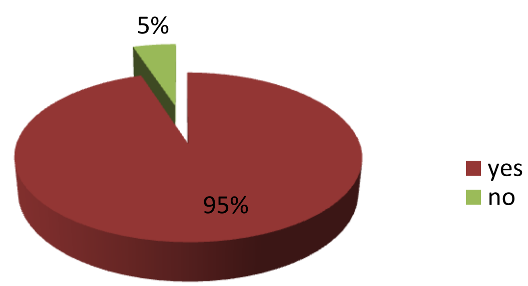
4.1 DATA GATHERING AND ANALYSIS

In this chapter, all the result and discussion will be briefly present and discussed. These results are not the final and complete yet – basically these are the result of the study obtained in order to see whether the system is really feasible.

In order to know if the idea of this project is acceptable by the public, questionnaire survey has been done in the early (planning stage) of this project. It is an online survey which completed through online survey form, www.surveymonkey.com. A total of four questions are asked in the survey in order to understand the passion of different individuals toward music, if they facing the problem of song selection as well as if the respondents like the idea of emotion based music player. A total of forty responds had been received within two days.

Below is the result for the questionnaire:

Table 4.1: Result for the survey

Question	Result
1. Do you love music?	
2. Do you listen to music/songs when your emotion distracted?	
3. Do you find it troublesome to search for songs that you want to listen from your music bank?	
4. Do you find it interesting If there is a system which able to detect your current mood and play the songs according to your mood/emotion?	

4.2 SYSTEM DESIGN

Design phase is also considered as one of the most challenging part in software development life cycle phase as it include the technical aspect of the project. The objective for this phase is to ensure the proper functions of the system and there is proper interaction between the system and the user.

4.2.1Proposed Model

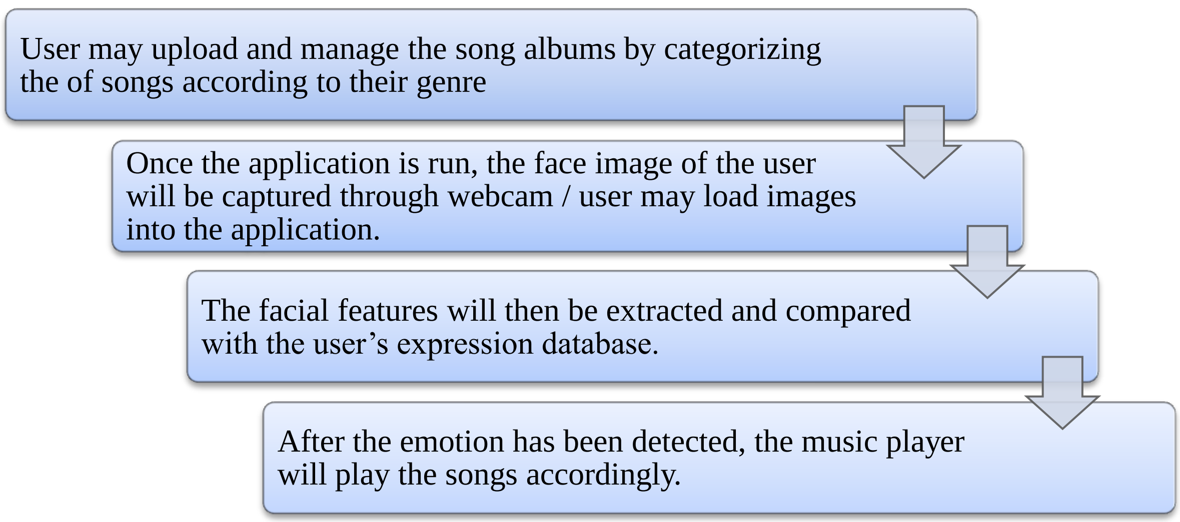


Figure 4.1: Working of the proposed model

From the developer side, as for this FYP, the proposed model will be focusing on two main functions, first is the expression detection and secondly the list of songs played for each category of emotion. As for the expression detection, the system is designed mainly to detect on the four major expressions, which are the happy, sad, normal and surprise. On the other hand, there will be songs ready available in each category. After the emotion of the individual is detected, the system will play the ten songs through music player.

Besides, there will be sets of still image with the four different expressions available in the database of facial expression detection. It will be used for the comparison purposes. After the image of the user is loaded, the features (lip and eyes) of the user will be extracted by the system. The system will then analyzed the condition of the features and do comparison with the sets of emotions in the database. The system will identify processed image for example as happy when the condition of the features are nearest to the "happy emotion" in the database.

As for the user side, the user will be able to customize the songs in each of the category according to their taste. Some might prefer sentimental music when she is sad but some might prefer some countryside music. There will be no limit of songs to be store in each category. User will have to launch the system in order to start the proposed model. Once the system is started, the user can choose to either select songs or to directly process the current emotion. A list of songs will be played automatically after the system done with the interpretation. User can choose to change the current emotion after the list of song is being played by repeating the image loading or capturing procedure.

4.2.2 Flow Chart

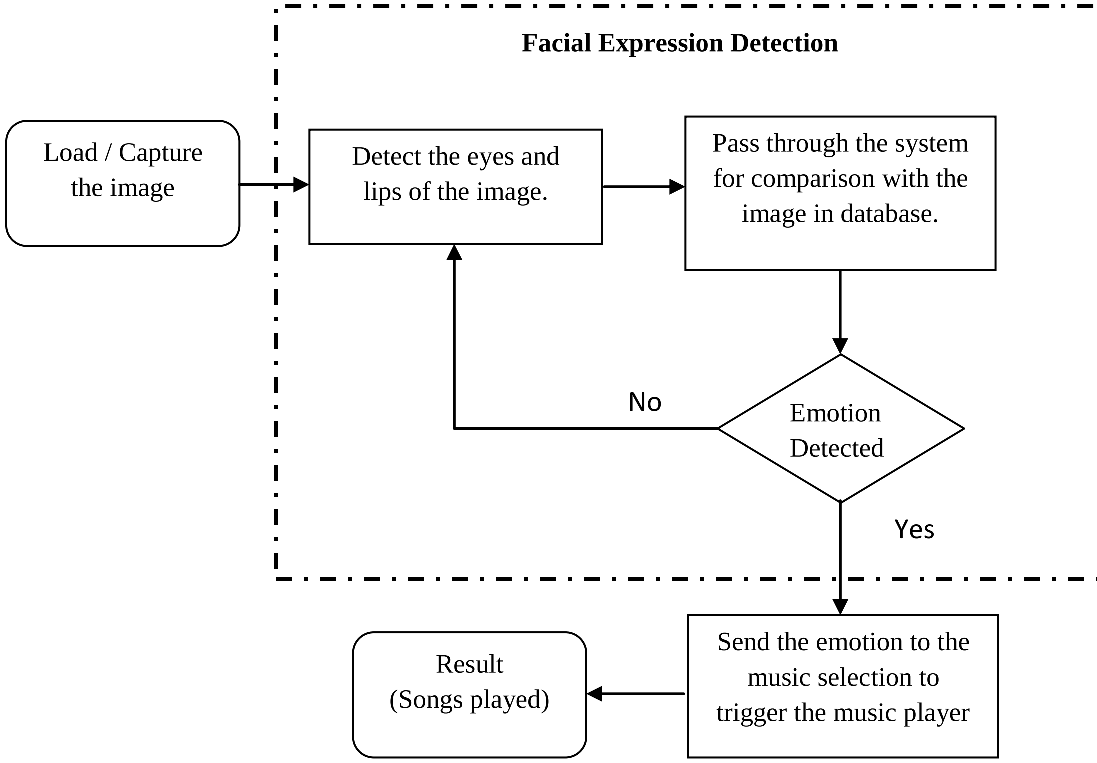


Figure 4.2: The Flow chart of the proposed model

4.2.3 As-Is-System and To-Be System

- **As-Is-System (User Process)**

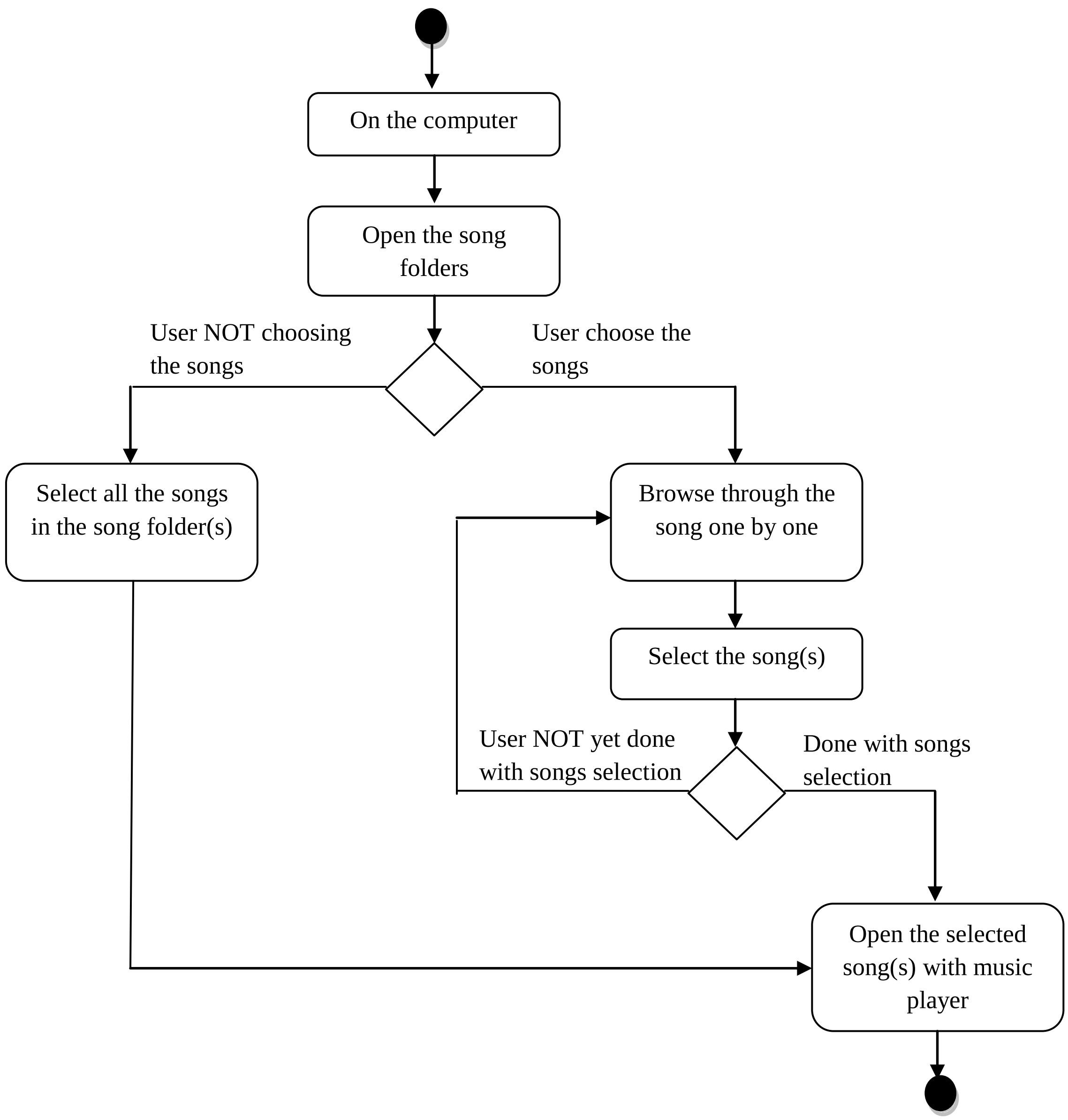


Figure 4.3: The As-Is user process Flowchart

➢ **To-Be-System (User Process)**

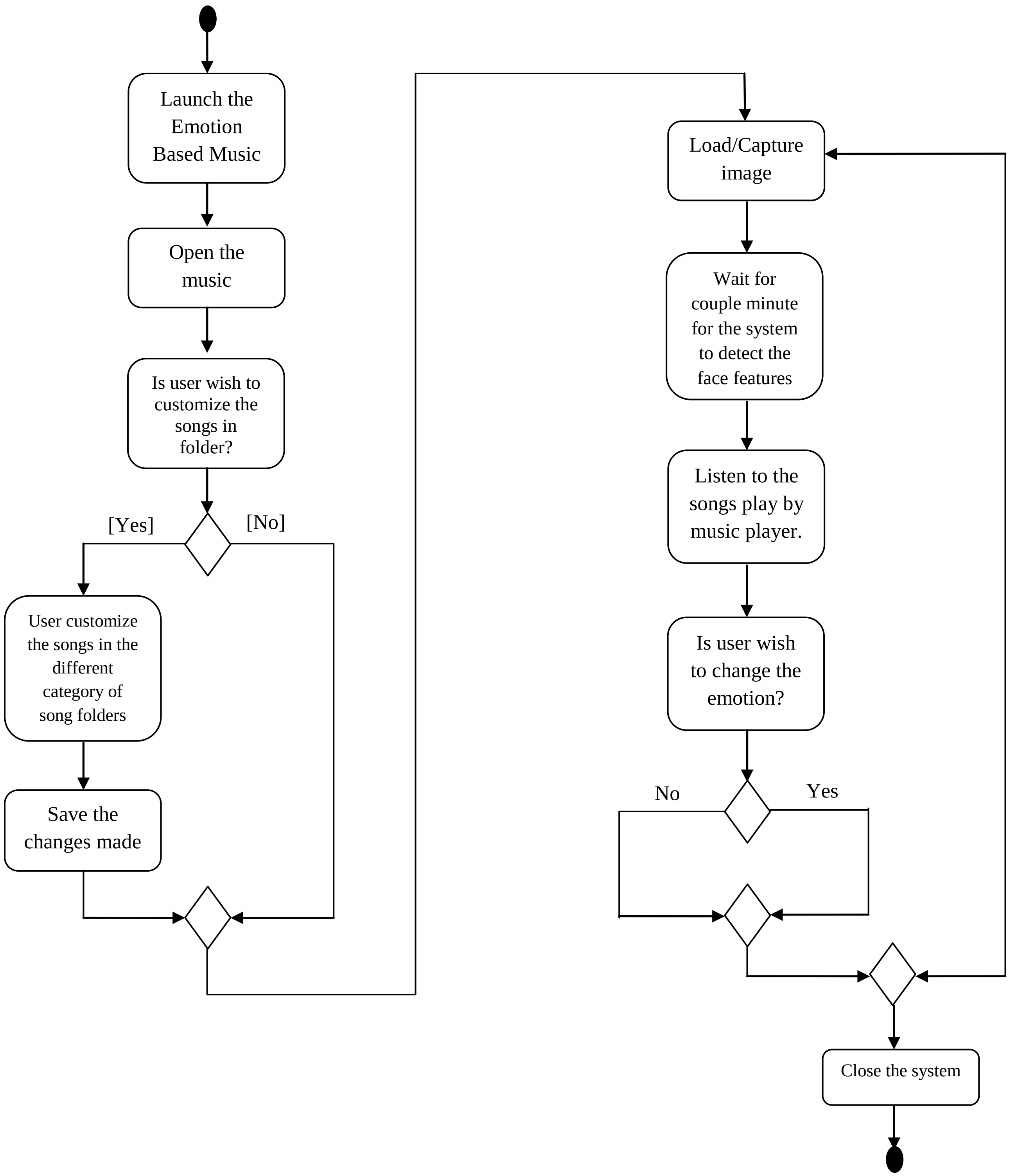


Figure 4.4: The To-Be user process flowchart

4.2.4 Interface Design (User Interface) and System Function

User interface refers to the interface users will see when running this proposed model. It basically either capturing the image or loading the still image and processes the image in order to detect and identify the user's emotion.

Once the proposed model is run, the below interface will appear to the user. Functions of each button in the first interface are as follow:

i.	"Browse" button	:	Enable the user to select available image from the local disk.
ii.	Webcam "Start" button	:	Automatically connect the user to the image capturing device to capture the current emotion of the user.
iii.	"Start" button	:	To start the image analysis process once the images is loaded.
iv.	"Restart" button	:	To restart the proposed model. It will start from the step of image selection.
v.	"Setting" button	:	To save the set of images into the database for comparison purposes.
vi.	"Music Library" button	:	Enable user to customize the songs list for each emotions.
vii.	"Exit" button	:	To exit the proposed model.

Figure 4.5: The first interface once the proposed model is launched

The proposed model detects the user's emotion based on the comparison method. Thus a set of emotion (normal, sad, surprise and happy) is saved in the database before the application is used by the user.

First of all, in order to save the set of images of different emotions in the proposed model, the "Browse" button is pressed. The user will have to choose an image which can represent the emotion. (Image selected in figure below is "Normal" emotion)

Figure 4.6: The "Normal" expression is loaded to the proposed model

Next, the “Start” button is pressed to indicate that the image selected is confirmed and the application can proceed with image processing.

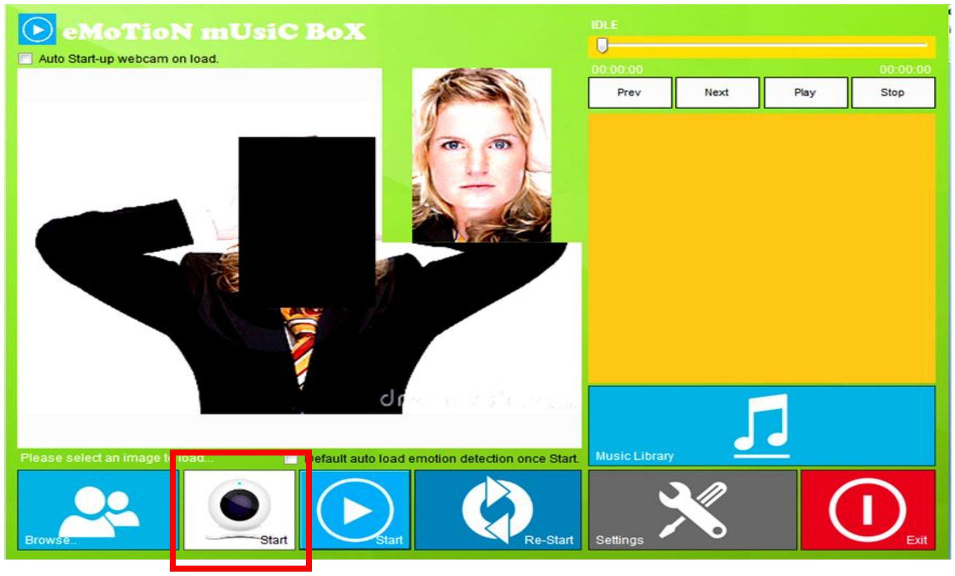


Figure 4.7: The interface after “Start” button is pressed

Then, the “Setting” button is pressed in order to set the loaded image as the dataset for comparison. After the proposed model done with the processing, the user can save the emotion type by clicking the “Save My Emotion Data” button. A message box will pop out once the emotion is successfully saved in the database.

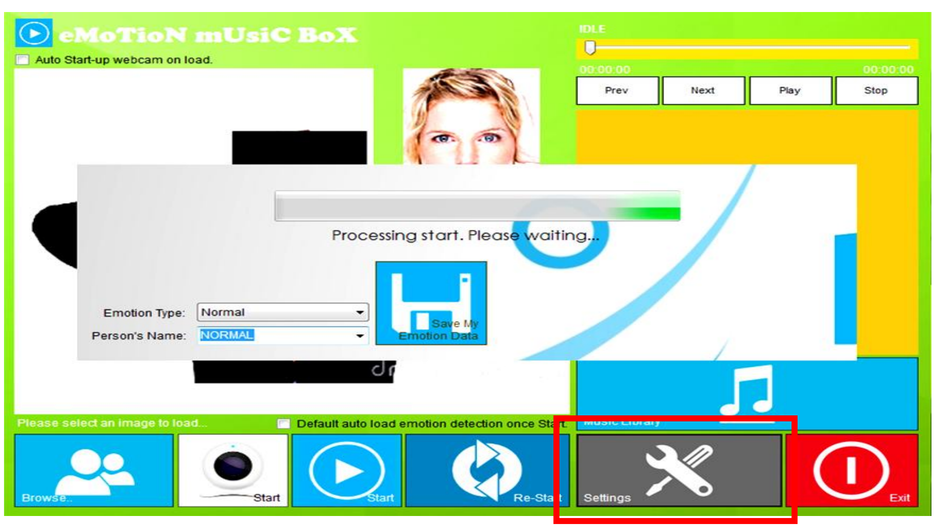


Figure 4.8: The interface after “Setting” button is pressed

After all the four emotions (normal, sad, surprise and happy) are successfully saved in the database, the application is now ready to detect the user's emotions. The proposed model is made to recognize the user's emotion according to the following steps:

i. Browse the image in local disk.
ii. Check the "Default auto load emotion detection once Start". The proposed model will automatically start the processing and detection procedure once the user clicks the "Start" button. (as shown in the image below)

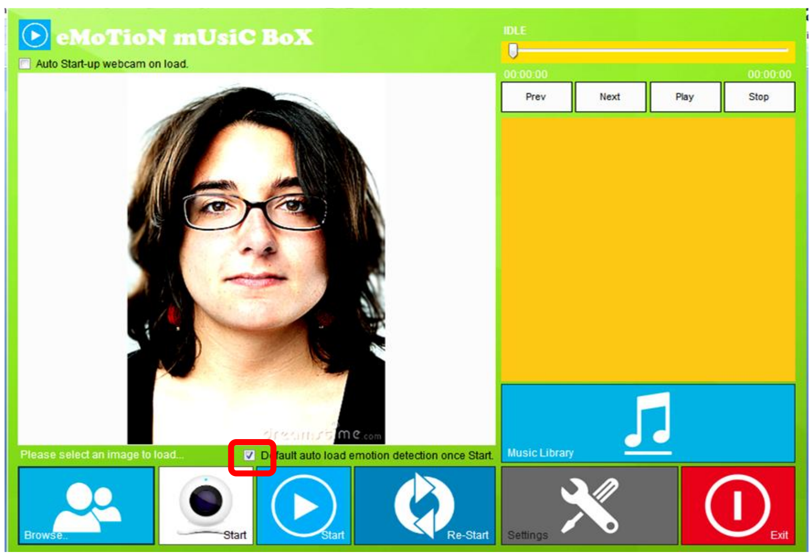


Figure 4.9: "Default auto load emotion detection once Start"

iii. Click the "Start" button.
iv. Click the "Play Song" button when the proposed model done with the image processing.

In order to have a clearer picture on how the proposed model work, the image below shows the condition of the interface in each step, from the loading of image until the song executed. (Emotion used in sample is "Normal" emotion.)

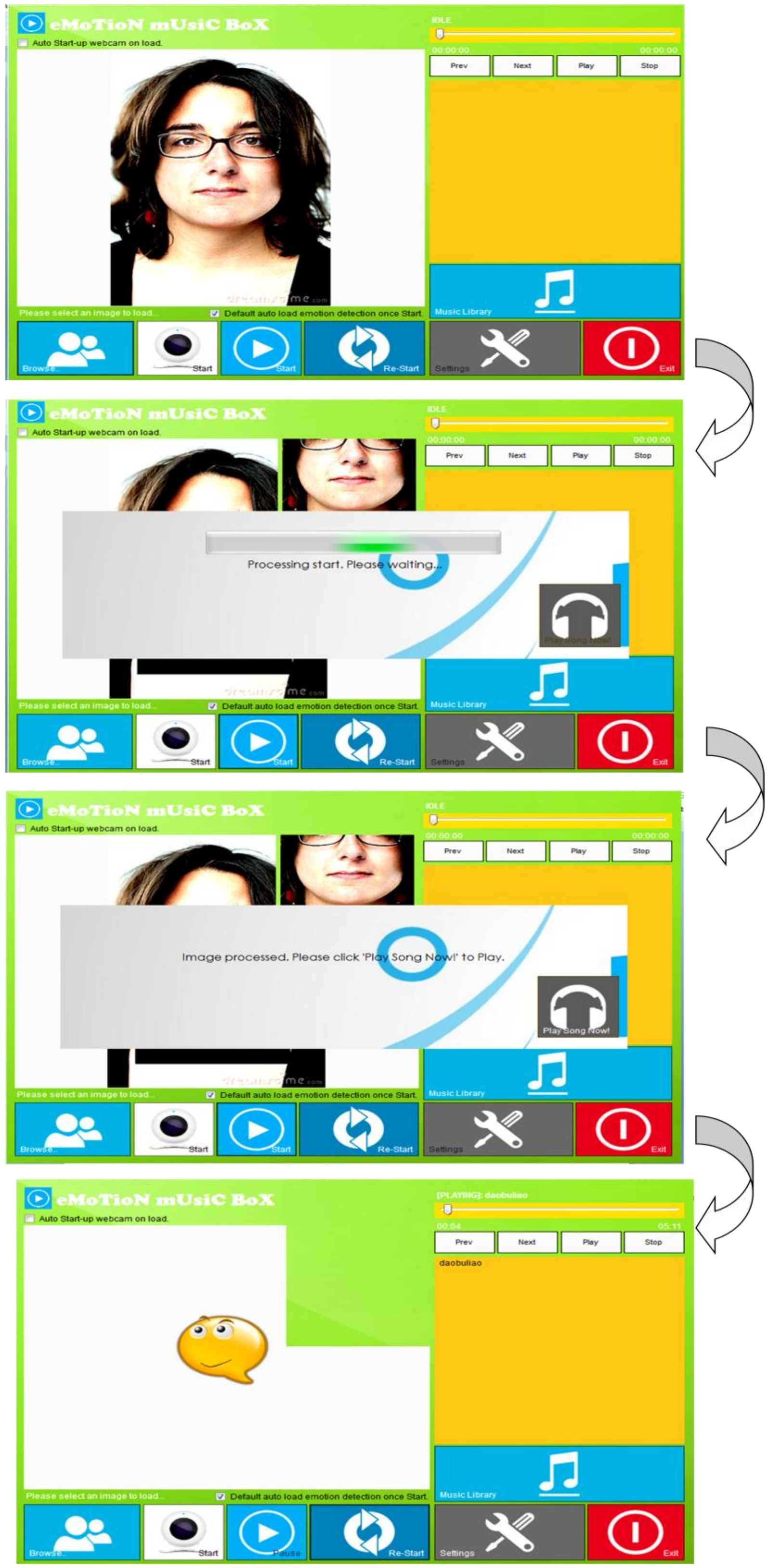


Figure 4.10: The step-to-step function of the proposed model

4.3 SYSTEM TESTING RESULTS

The user carried out system testing once the completion of the system development. The purpose of this testing is to check the functionalities system, whether if it is usable and well-functioned. The results from the functional testing can be seen in the table below.

Table 4.2: System functional testing results

Component	Expected Function	Testing Result	
		Positive	Negative
"Browse" button	Direct the user to the local disk.	✓	
	Enable the user to select image from the local disk.	✓	
Webcam "Start" button	Start the embedded webcam / external webcam	✓	
"Start" button	To start the image analysis process once it is pressed.	✓	
"Restart" button	Clear the previously loaded image.	✓	
	Stop the currently played songs.	✓	
"Setting" button	Enable user to save the set of images into the database.	✓	
"Music Library" button	Direct user to the local disk for song customization.	✓	
	Enable user to customize the songs list for each emotions, include deleting unwanted songs and adding new songs.	✓	
"Exit" button	Close the window of the proposed model.	✓	

4.4 EMOTION ACCURACY TESTING RESULTS

Set of images for the each emotions (normal, sad, surprise and happy) are saved in the proposed model for the comparison purposes. The newly load images will be compared with the saved dataset in order to detect the emotion of the users. Table below showed the set of images that saved in the proposed model.

Table 4.3: The dataset of images saved in the proposed model

Images	Emotion
	**Normal**
	**Sad**
	**Surprise**
	**Happy**

(Source from http://www.dreamstime.com)

The proposed model is tested with set of images of similar emotion to test on it accuracy in detecting the emotion. Ten images are tested for each category of emotions and the results are shown as tables below.

Table 4.4: The testing result for "Normal" Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	
	✓				✓

(Source from http://www.dreamstime.com)

Table 4.5: The testing result for “Sad” Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓				✓
	✓				✓
	✓				✓

(Source from http://www.dreamstime.com)

Table 4.6: The testing result for “Surprise” Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓				✓
	✓				✓

(Source from http://www.dreamstime.com)

Table 4.7: The testing result for "Happy" Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	

(Source from http://www.dreamstime.com)

Summary of Result are shown in the table below:

Table 4.8: The summary of the results tested

Emotion	No. of Samples	No. of Recognized Sample	RR
Happy	10	10	100%
Normal	10	9	90%
Sad	10	7	80%
Surprise	10	8	80%
Total	**40**	**34**	**85%**

In order to find the RR (Recognition Rate), the following formula is applied to the results collected as below:

$$RR = \frac{Classified\ Character}{Total\ Number\ of\ Character} \text{ x } 100\%$$

RR = 34/40 * 100

RR = 85%

Based on the result above, it shows that the proposed model has the recognition rate (RR) of 85%.

4.5 DISCUSSION

Based on the observation during the prototype evaluation period, there are several limitations which prevent the proposed model to perform accurately.

1. <u>Posture of the object in the image.</u>

   The proposed model will perform accurately if the object of the image is in upright posture and the individual's face is clearly exposed. Minor slant of individual's head is acceptable and detectable

   The images below are example of images that can be accurately detected.

(Source from http://www.dreamstime.com)

Figure 4.11: Image which is accurately detected (a)

The images below are example of images which fail to be detected by the proposed model.

(Source from http://www.dreamstime.com)

Figure 4.12: Image which cannot be detected accurately (a)

2. The quality of the image, either still image loaded or captured by webcam.

   The proposed model will detect the emotion more accurately when the images loaded or captured are in better resolution, brightness and contrast. There will be warning message prompted when the image loaded in is too fine or much noise. The best picture will be picture with upper part of the body, white color background and in high resolution.

(Source from http://www.dreamstime.com)

Figure 4.13: Image which is accurately detected (b)

(Source from http://www.dreamstime.com)

Figure 4.14: Image which cannot be detected accurately (b)

The same limitation applied to the on the spot image capturing through either webcam or external webcam. In Figure 4.16 it shows that once the “Start” button is clicked, the error message prompting out indicating the image is too fine to be processed.

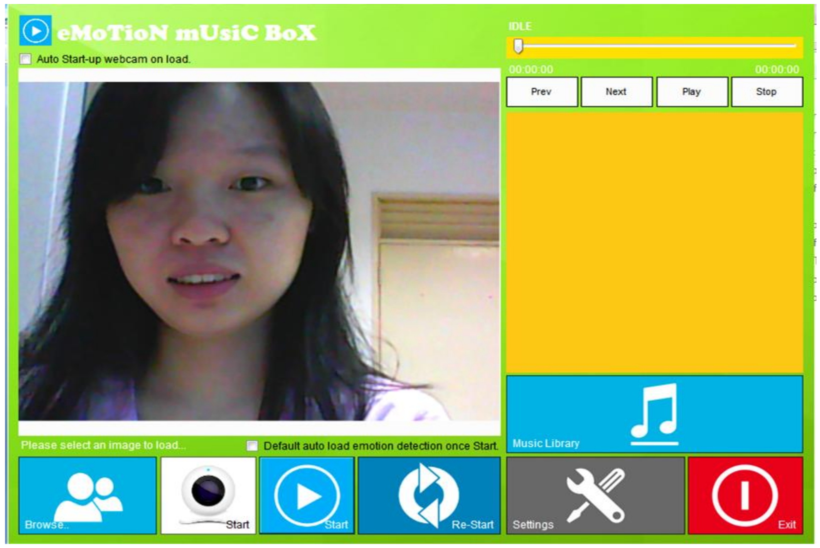


Figure 4.15: Image captured through webcam

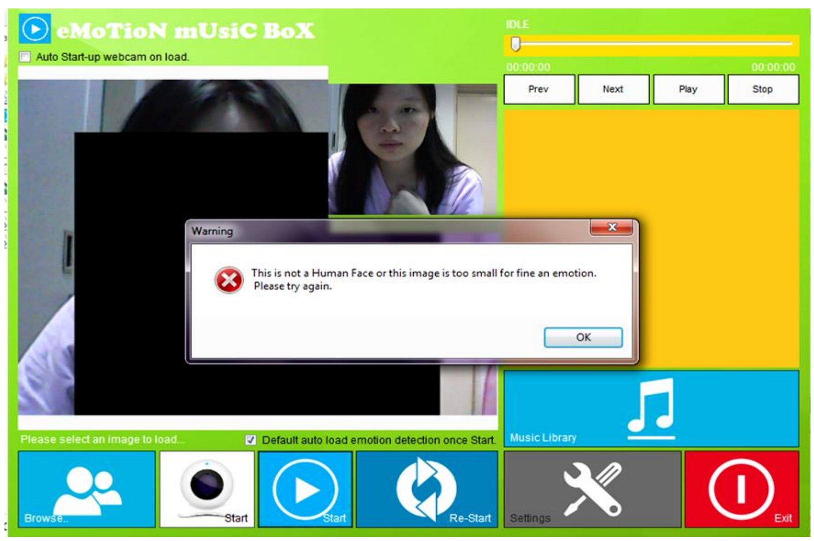


Figure 4.16: Error message

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

The significant of this project is the emotion detection of the images loaded into the proposed model. The main purpose is on its emotion detection functionality. Through the integration between emotion detection technology and music player, the proposed model is aimed to provide betterment in the individual's entertainment. The proposed is able to detect the four emotions i.e. normal, happy, and sad of the images loaded into it. Once the proposed model compared and detected the emotion of the user, the music player will play the song(s) accordingly.

As for the usability and accuracy, both system testing and emotion accuracy testing has been done to the proposed model and return a satisfying result. The proposed model able to recognized 34 out of 40 images loaded into it, which give a Recognition Rate of 85%. Besides, the proposed model is a computer application which can works well in all kinds of windows and computers.

Thus with this Emotion Based Music Player, users can have an alternative way of selecting songs, which is in a more interactive and simpler way. The music lovers will not have to search through the long list of songs for the songs to be played but to match the emotion in the songs selection.

5.2 RECOMMENDATION

Every applications or systems are subject to upgrades and improvement, so do the Emotion Based Music Player. This proposed model can still be enhanced in terms of its performances and features.

First and foremost is to reduce the limitation in the emotion detection. As mentioned before, there are several limitations in emotion detection. In order to improve accuracy, it can be done by increasing the numbers of facial features when doing extraction for comparison. Currently, the proposed extracts only the lips and eyes for comparison purposes. As for future work, other facial features such as eyebrows and movement of cheeks can be included in the comparison.

Besides, the future model can further enhanced by removing or minimizing the noise if the loaded or captured images. In future expansion, noise reduction software can be embedded in the model so that the noise for either still or captured image can be removed and thus increase the accuracy in emotion detection.

Apart from the above, the proposed model can be improved by having auto adjustment on the resolution or brightness and contrast of the images. The accuracy of emotion detection for the current application is greatly influenced by the quality of the images loaded. Hence by having the auto adjustment, the user can load in any quality of image or capture images with any kinds of webcam. The future model will be able to adjust the quality of the images which can be detected and processed.

In addition, for the better interactive between user and application, real time emotion detection technique can be applied to the model. The future model will detect and extract the facial feature once the application is launched and the emotion can be detected in real time.

REFERENCES

[1] Ekman, P. & Friesen, W. V. (1969): *The repertoire of nonverbal behavior: Categories, origins, usage, and coding:* Semiotica, Vol. 1: 49–98.

[2] Ying-Li Tian, Takeo Kanade, and Jeffrey F. Cohn. (2003): *Facial Expression Analysis*. Retrieved from www.ri.cmu.edu/pub_files/pub4/tian.../tian_ying_li_2003_1.pdf.

[3] Barbara Raskauskas (2009). Why Do People Listen to Music? Retrieved on October 9 2012 from http://voices.yahoo.com/why-people-listen-music-2608185.html.

[4] Emily Sohn(2011) Why Music Makes You Happy? Retrieved on October 9 2012 from http://news.discovery.com/human/music-dopamine-happiness-brain-110110.html.

[5] Mary Duenwald (2005). The Physiology of Facial Expressions. Retrieved on October 9 2012 from http://discovermagazine.com/2005/jan/physiology-of-facial-expressions

[6] W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat. (2004): *Facial Expression Detection And Recognition System:* Journal of The Institution of Engineers, Singapore. Vol 44.

[7] Jagdish Lal Raheja, Umesh Kumar. (2010): *Human Facial Expression Detection From Detected In Captured Image Using Back Propagation Neural Network:* International Journal of Computer Science & Information Technology (IJCSIT). Vol.2(1).

[8] Zhengyou Zhang. (1998): *Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiments With a Multi-Layer Perceptron:* Journal of Pattern Recognition and Artificial Intelligence. Vol. 13(6): 893-911.

[9] Eva Cerezo1, Isabelle Hupont, Critina Manresa, Javier Varona, Sandra Baldassarri, Francisco J. Perales, and Francisco J. Seron. (2007): *Real-Time Facial Expression Recognition for Natural Interaction:* In J. Martí et al. (Eds.),: *IbPRIA 2007, Part II, LNCS 4478,* (pp. 40–47). Springer-Verlag Berlin Heidelberg.

[10] Antoinette L. Bouhuys, Gerda M. Bloem, Ton G.G.Groothuis. (1994): *Induction Of Depressed And Elated Mood By Music Influences The Perception Of Facial Emotional Expressions In Healthy Subjects*: Journal of Affective Disorders. Vol. 33: 215-226.

[11] Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou. (2007): *A Grounded Theory of Young Tennis Players' Use of Music to Manipulate Emotional State:* Journal of Sport & Exercise Psychology: 584-607.

[12] Frijda, N.H. (1986): *The emotions*. New York: Cambridge University Press.

[13] Matthew Montague Lavy (2001). *Emotion and the Experience of Listening to Music A Framework for Empirical Research*. (Unpublished master's thesis). Jesus College, Cambridge.

[14] Wai Ling Cheung and Guojun Lu. (2008): *Music Emotion Annotation by Machine Learning.* doi: 10.1109/MMSP.2008.4665144.

[15] Maria M. Ruxanda1, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen. (2007): *Emotion-Based Music Retrieval On A Well-Reduced Audio Feature Space:* In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 181-184.

[16] Laughing On The Other Side Of One's Face [Image]. Retrieved on October 9 2012 from http://blog.inkyfool.com.

[17] Human Facial Expressions [Image]. Retrieved on October 9 2012 from http://www.flickr.com/groups/1876998@N24/pool/.

[18] Stock Photography [Image]. Retrieved on October 9 2012 from http://www.dreamstime.com

[19] Deposit Photos [Image]. Retrieved on November 10 2012 from http://depositphotos.com.

[20] Microsof Studio C# Programming. (n.d.) Retrieved from http://www.functionx.com/vcsharp/index.htm.

[21] C# Station. (2012). Retrieved from http://www.csharp-station.com/Tutorial.aspx.

[22] Joseph C. Hager (2003). Introduction To The DataFace Site: Facial Expressions, Emotion Expressions, Nonverbal Communication, Physiognomy. Retrieved 10 October 2012, from http://face-and-emotion.com/dataface/general/homepage.jsp

[23] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, May 2009.

APPENDIX

i. Questionnaire develops through www.surveymonkey.com.

Emotion Based Music Player Survey

1. Do you love music?

- ○ Yes
- ○ No

2. Do you listen to music/songs when your emotion distracted? (When you are sad/angry/happy)

- ○ Yes
- ○ No

3. Do you find it troublesome to search for songs that you want to listen from your music bank?

- ○ Yes
- ○ No

4. Do you find it interesting If there is a system which able to detect your current mood and play the songs according to your mood/emotion?

- ◉ Yes
- ○ No

Done

Powered by SurveyMonkey
Check out our sample surveys and create your own now!

ii. Part of the code in Filter.cs

```
public class BitmapFilter
    {
        public static bool Invert(Bitmap b)
        {
            // GDI+ still lies to us - the return format is BGR, NOT RGB.
            BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

            int stride = bmData.Stride;
            System.IntPtr Scan0 = bmData.Scan0;

            unsafe
            {
                byte * p = (byte *)(void *)Scan0;

                int nOffset = stride - b.Width*3;
                int nWidth = b.Width * 3;

                for(int y=0;y<b.Height;++y)
                {
                    for(int x=0; x < nWidth; ++x )
                    {
                        p[0] = (byte)(255-p[0]);
                        ++p;
                    }
                    p += nOffset;
                }
            }
            b.UnlockBits(bmData);

            return true;
        }
        public static bool GrayScale(Bitmap b)
        {
            // GDI+ still lies to us - the return format is BGR, NOT RGB.
            BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

            int stride = bmData.Stride;
            System.IntPtr Scan0 = bmData.Scan0;

            unsafe
            {
                byte * p = (byte *)(void *)Scan0;

                int nOffset = stride - b.Width*3;

                byte red, green, blue;

                for(int y=0;y<b.Height;++y)
                {
```

```
                        for(int x=0; x < b.Width; ++x )
                        {
                            blue = p[0];
                            green = p[1];
                            red = p[2];

                            p[0] = p[1] = p[2] = (byte)(.299 * red + .587 * green
+ .114 * blue);

                            p += 3;
                        }
                        p += nOffset;
                    }
                }

                b.UnlockBits(bmData);

                return true;
            }

            public static bool Brightness(Bitmap b, int nBrightness)
            {
                if (nBrightness < -255 || nBrightness > 255)
                    return false;

                // GDI+ still lies to us - the return format is BGR, NOT RGB.
                BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

                int stride = bmData.Stride;
                System.IntPtr Scan0 = bmData.Scan0;

                int nVal = 0;

                unsafe
                {
                    byte * p = (byte *)(void *)Scan0;

                    int nOffset = stride - b.Width*3;
                    int nWidth = b.Width * 3;

                    for(int y=0;y<b.Height;++y)
                    {
                        for(int x=0; x < nWidth; ++x )
                        {
                            nVal = (int) (p[0] + nBrightness);

                            if (nVal < 0) nVal = 0;
                            if (nVal > 255) nVal = 255;

                            p[0] = (byte)nVal;

                            ++p;
                        }
```

```
				p += nOffset;
			}
		}

		b.UnlockBits(bmData);

		return true;
	}

	public static bool Contrast(Bitmap b, sbyte nContrast)
	{
		if (nContrast < -100) return false;
		if (nContrast >  100) return false;

		double pixel = 0, contrast = (100.0+nContrast)/100.0;

		contrast *= contrast;

		int red, green, blue;

		// GDI+ still lies to us - the return format is BGR, NOT RGB.
		BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

		int stride = bmData.Stride;
		System.IntPtr Scan0 = bmData.Scan0;

		unsafe
		{
			byte * p = (byte *)(void *)Scan0;

			int nOffset = stride - b.Width*3;

			for(int y=0;y<b.Height;++y)
			{
				for(int x=0; x < b.Width; ++x )
				{
					blue = p[0];
					green = p[1];
					red = p[2];

					pixel = red/255.0;
					pixel -= 0.5;
					pixel *= contrast;
					pixel += 0.5;
					pixel *= 255;
					if (pixel < 0) pixel = 0;
					if (pixel > 255) pixel = 255;
					p[2] = (byte) pixel;

					pixel = green/255.0;
					pixel -= 0.5;
					pixel *= contrast;
					pixel += 0.5;
```

```
                        pixel *= 255;
                        if (pixel < 0) pixel = 0;
                        if (pixel > 255) pixel = 255;
                        p[1] = (byte) pixel;

                        pixel = blue/255.0;
                        pixel -= 0.5;
                        pixel *= contrast;
                        pixel += 0.5;
                        pixel *= 255;
                        if (pixel < 0) pixel = 0;
                        if (pixel > 255) pixel = 255;
                        p[0] = (byte) pixel;

                        p += 3;
                    }
                    p += nOffset;
                }
            }

            b.UnlockBits(bmData);

            return true;
        }

        public static bool Gamma(Bitmap b, double red, double green, double blue)
        {
            if (red < .2 || red > 5) return false;
            if (green < .2 || green > 5) return false;
            if (blue < .2 || blue > 5) return false;

            byte [] redGamma = new byte [256];
            byte [] greenGamma = new byte [256];
            byte [] blueGamma = new byte [256];

            for (int i = 0; i< 256; ++i)
            {
                redGamma[i] = (byte)Math.Min(255, (int)(( 255.0 * Math.Pow(i/255.0,
1.0/red)) + 0.5));
                greenGamma[i] = (byte)Math.Min(255, (int)(( 255.0 * Math.Pow(i/255.0,
1.0/green)) + 0.5));
                blueGamma[i] = (byte)Math.Min(255, (int)(( 255.0 * Math.Pow(i/255.0,
1.0/blue)) + 0.5));
            }

            // GDI+ still lies to us - the return format is BGR, NOT RGB.
            BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

            int stride = bmData.Stride;
            System.IntPtr Scan0 = bmData.Scan0;

            unsafe
            {
```

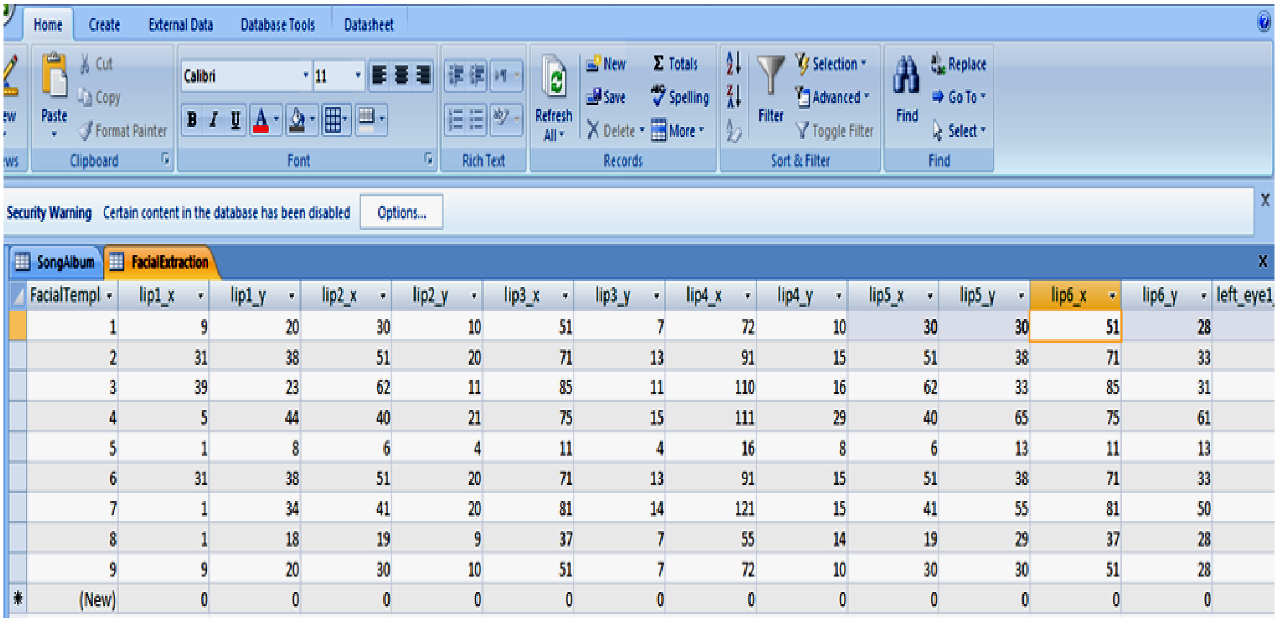
```
                byte * p = (byte *)(void *)Scan0;

                int nOffset = stride - b.Width*3;

                for(int y=0;y<b.Height;++y)
                {
                    for(int x=0; x < b.Width; ++x )
                    {
                        p[2] = redGamma[ p[2] ];
                        p[1] = greenGamma[ p[1] ];
                        p[0] = blueGamma[ p[0] ];

                        p += 3;
                    }
                    p += nOffset;
                }
            }

            b.UnlockBits(bmData);

            return true;
        }

        public static bool Color(Bitmap b, int red, int green, int blue)
        {
            if (red < -255 || red > 255) return false;
            if (green < -255 || green > 255) return false;
            if (blue < -255 || blue > 255) return false;

            // GDI+ still lies to us - the return format is BGR, NOT RGB.
            BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

            int stride = bmData.Stride;
            System.IntPtr Scan0 = bmData.Scan0;

            unsafe
            {
                byte * p = (byte *)(void *)Scan0;

                int nOffset = stride - b.Width*3;
                int nPixel;

                for(int y=0;y<b.Height;++y)
                {
                    for(int x=0; x < b.Width; ++x )
                    {
                        nPixel = p[2] + red;
                        nPixel = Math.Max(nPixel, 0);
                        p[2] = (byte)Math.Min(255, nPixel);

                        nPixel = p[1] + green;
                        nPixel = Math.Max(nPixel, 0);
                        p[1] = (byte)Math.Min(255, nPixel);
```

```
                        nPixel = p[0] + blue;
                        nPixel = Math.Max(nPixel, 0);
                        p[0] = (byte)Math.Min(255, nPixel);

                        p += 3;
                    }
                    p += nOffset;
                }
            }

            b.UnlockBits(bmData);

            return true;
        }

        public static bool Conv3x3(Bitmap b, ConvMatrix m)
        {
            // Avoid divide by zero errors
            if (0 == m.Factor) return false;

            Bitmap bSrc = (Bitmap)b.Clone();

            // GDI+ still lies to us - the return format is BGR, NOT RGB.
            BitmapData bmData = b.LockBits(new Rectangle(0, 0, b.Width, b.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);
            BitmapData bmSrc = bSrc.LockBits(new Rectangle(0, 0, bSrc.Width, bSrc.Height),
ImageLockMode.ReadWrite, PixelFormat.Format24bppRgb);

            int stride = bmData.Stride;
            int stride2 = stride * 2;
            System.IntPtr Scan0 = bmData.Scan0;
            System.IntPtr SrcScan0 = bmSrc.Scan0;

            unsafe
            {
                byte * p = (byte *)(void *)Scan0;
                byte * pSrc = (byte *)(void *)SrcScan0;

                int nOffset = stride + 6 - b.Width*3;
                int nWidth = b.Width - 2;
                int nHeight = b.Height - 2;

                int nPixel;

                for(int y=0;y < nHeight;++y)
                {
                    for(int x=0; x < nWidth; ++x )
                    {
                        nPixel = ( ( ( (pSrc[2] * m.TopLeft) + (pSrc[5] *
m.TopMid) + (pSrc[8] * m.TopRight) +
                            (pSrc[2 + stride] * m.MidLeft) + (pSrc[5 +
stride] * m.Pixel) + (pSrc[8 + stride] * m.MidRight) +
```

```
(pSrc[2 + stride2] * m.BottomLeft) + (pSrc[5 + stride2] * m.BottomMid) + (pSrc[8 + stride2] * m.BottomRight)) / m.Factor) + m.Offset);

            if (nPixel < 0) nPixel = 0;
            if (nPixel > 255) nPixel = 255;

            p[5 + stride]= (byte)nPixel;

            nPixel = ( ( ( (pSrc[1] * m.TopLeft) + (pSrc[4] * m.TopMid) + (pSrc[7] * m.TopRight) +
                (pSrc[1 + stride] * m.MidLeft) + (pSrc[4 + stride] * m.Pixel) + (pSrc[7 + stride] * m.MidRight) +
                (pSrc[1 + stride2] * m.BottomLeft) + (pSrc[4 + stride2] * m.BottomMid) + (pSrc[7 + stride2] * m.BottomRight)) / m.Factor) + m.Offset);

            if (nPixel < 0) nPixel = 0;
            if (nPixel > 255) nPixel = 255;

            p[4 + stride] = (byte)nPixel;

            nPixel = ( ( ( (pSrc[0] * m.TopLeft) + (pSrc[3] * m.TopMid) + (pSrc[6] * m.TopRight) +
                (pSrc[0 + stride] * m.MidLeft) + (pSrc[3 + stride] * m.Pixel) + (pSrc[6 + stride] * m.MidRight) +
                (pSrc[0 + stride2] * m.BottomLeft) + (pSrc[3 + stride2] * m.BottomMid) + (pSrc[6 + stride2] * m.BottomRight)) / m.Factor) + m.Offset);

            if (nPixel < 0) nPixel = 0;
            if (nPixel > 255) nPixel = 255;
            p[3 + stride] = (byte)nPixel;
            p += 3;
            pSrc += 3;
          }
          p += nOffset;
          pSrc += nOffset;
        }
      }
      b.UnlockBits(bmData);
      bSrc.UnlockBits(bmSrc);
      return true;
    }

    public static bool Smooth(Bitmap b, int nWeight /* default to 1 */)
    {
      ConvMatrix m = new ConvMatrix();
      m.SetAll(1);
      m.Pixel = nWeight;
      m.Factor = nWeight + 8;
      return  BitmapFilter.Conv3x3(b, m);
    }

    public static bool GaussianBlur(Bitmap b, int nWeight /* default to 4*/)
    {
      ConvMatrix m = new ConvMatrix();
```

```
            m.SetAll(1);
            m.Pixel = nWeight;
            m.TopMid = m.MidLeft = m.MidRight = m.BottomMid = 2;
            m.Factor = nWeight + 12;

            return  BitmapFilter.Conv3x3(b, m);
        }

        public static bool MeanRemoval(Bitmap b, int nWeight /* default to 9*/ )
        {
            ConvMatrix m = new ConvMatrix();
            m.SetAll(-1);
            m.Pixel = nWeight;
            m.Factor = nWeight - 8;

            return BitmapFilter.Conv3x3(b, m);
        }

        public static bool Sharpen(Bitmap b, int nWeight /* default to 11*/ )
        {
            ConvMatrix m = new ConvMatrix();
            m.SetAll(0);
            m.Pixel = nWeight;
            m.TopMid = m.MidLeft = m.MidRight = m.BottomMid = -2;
            m.Factor = nWeight - 8;
            return  BitmapFilter.Conv3x3(b, m);
        }

        public static bool EmbossLaplacian(Bitmap b)
        {
            ConvMatrix m = new ConvMatrix();
            m.SetAll(-1);
            m.TopMid = m.MidLeft = m.MidRight = m.BottomMid = 0;
            m.Pixel = 4;
            m.Offset = 127;

            return  BitmapFilter.Conv3x3(b, m);
        }

        public static bool EdgeDetectQuick(Bitmap b)
        {
            ConvMatrix m = new ConvMatrix();
            m.TopLeft = m.TopMid = m.TopRight = -1;
            m.MidLeft = m.Pixel = m.MidRight = 0;
            m.BottomLeft = m.BottomMid = m.BottomRight = 1;
            m.Offset = 127;
            return  BitmapFilter.Conv3x3(b, m);
        }
    }
}
```

iii. Screenshot of database for Facial Expression Table



FacialTempl	lip1_x	lip1_y	lip2_x	lip2_y	lip3_x	lip3_y	lip4_x	lip4_y	lip5_x	lip5_y	lip6_x	lip6_y	left_eye1
1	9	20	30	10	51	7	72	10	30	30	51	28	
2	31	38	51	20	71	13	91	15	51	38	71	33	
3	39	23	62	11	85	11	110	16	62	33	85	31	
4	5	44	40	21	75	15	111	29	40	65	75	61	
5	1	8	6	4	11	4	16	8	6	13	11	13	
6	31	38	51	20	71	13	91	15	51	38	71	33	
7	1	34	41	20	81	14	121	15	41	55	81	50	
8	1	18	19	9	37	7	55	14	19	29	37	28	
9	9	20	30	10	51	7	72	10	30	30	51	28	
(New)	0	0	0	0	0	0	0	0	0	0	0	0	

Emotion Based Music Player

Tan Siew Ching
Department of Computer and Information Sciences,
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Tronoh Perak, Malaysia
siewc0510@gmail.com

Assoc. Prof. Dr. Baharum B Baharudin
Department of Computer and Information Sciences,
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Tronoh Perak, Malaysia
baharbh@petronas.com.my

***Abstract* — The work presents described the development of Emotion Based Music Player, which is a computer application meant for all type of users, specifically the music lovers. Due to the troublesome workloads in songs selection, most people will choose to randomly play the songs in the playlist. As a result, some of the songs selected not matching the users' current emotion. The proposed model is able to extract user's facial expression and thus detect user's emotion. The music player in the proposed model will then play the songs according to the category of emotion detected. It is aimed to provide a better enjoyment to music lovers in music listening. The scope of emotions in the proposed model involve normal, sad, surprise and happy. The system involves the major of image processing and facial detection technologies. The input for this proposed model is the .jpeg format still images which available online. The performance of this model is evaluated by loading forty still images (ten for each emotion category) into the proposed model to test on the accuracy in detecting the emotions. Based on the testing result, the proposed model has the Recognition Rate of 85%.**

I. INTRODUCTION

Emotions are the bodily feelings associated with mood, temperament, personality or character. Paul Ekman had developed the classifications of basic emotions which are anger, disgust, fear, happiness, sadness and surprise. [1]

A facial expression can be expressed through the motions or from one or more motions, movements or even positions of the muscles of the face. Facial expression can be adopted as voluntary action as individual can control his facial expression and to show the facial expression according to his will. However, since facial expression is closely associated with the emotion, thus it is mostly an involuntary action. An individual may show his expression in first few micro-second before resume to a neutral expression. Facial expression analysis includes both detection and interpretation of facial motion and recognition of expression. The three approaches which enabled the automatic facial expression analysis includes i) face acquisition, ii) facial data extraction and representation, and iii) facial expression recognition.

The "Emotion Based Music Player" is a device developed aimed to detect the emotion of an individual, and play the lists of music accordingly.

First, the individual will reflect his emotion through the facial expression. After that, the device will detect the condition of the facial expression, analyze and determine the emotion of the individual. The music player will play the songs which can suit the current emotion of the individual. The device will focus on the analysis of the facial expression only which does not include the head or face movement.

What music lovers facing now is the difficulty in songs selection, especially songs that match individuals' current emotions. Looking at the long lists of unsorted music, individuals will feel more demotivated to look for the songs they want to listen to. For an example, when a person is sad, he would like to listen to some heavy rock music to release his sadness. The individual would rather choose the songs randomly or just "play all" for all the songs he had. Besides, people get bored with this traditional way of searching and selecting songs. The method had been implemented since few years back.

The main objective of this project is to develop the "Emotion Based Music Player" for music lovers which aimed to serve as a platform to assist individuals to play and listen to the songs according to his emotions automatically. It is aimed to provide a better enjoyment of entertainment to the music lovers. The specific objectives are to propose a facial expression detection model, to accurately detect the four basic emotions, namely normal, happy, sad and surprise and to integrate the music player into the proposed model.

The scope of study for this project includes the study on the different method in expression detection. With the improvement of technology in image processing, more and more experts introduced different techniques in processing and emotion detection. All these techniques are useful and fundamental for this project. The second scope is to get information on the tools appropriate in the facial expression detection when building the proposed model for this project. Different tools are studied on their feasibility as well as user-friendliness to figure out which is the most suitable and applicable tools.

II. LITERATURE REVIEW

A. INTRODUCTION

In the year of 2009, Barbara Raskauskas mentioned that "music does fill the silence and can hide the noise. Music can convey cultural upbringing. Music is pleasurable and speaks to us, whether or not the song has words. I've never met a person who didn't like some form of music. Even a deaf friend of mine said she liked music; she could feel the vibration caused by music. Finding enjoyment in music is universal."[2] Emily Sohn stated that "People love music for much the same reason they're drawn to sex, drugs, gambling and delicious food, according to new research". Study had proved that human brain will release dopamine, a kind of chemical generated by body which involved addiction and motivation when people listen to harmony or melody that touch an individual.[3]

In the year of 2005, Mary Duenwald had published an article which the facial expressions across the globe fall roughly into seven categories:

i. **Sadness**: The eyelids droop while the inner corners of the brows rise.
ii. **Surprise**: Both the upper eyelids and brows rise, and the jaw drops open.
iii. **Anger**: Both the lower and upper eyelids squeeze in draw together. The jaw pushes forward while lip pressed on each other.
iv. **Contempt**: The expression appears on one side of a face: One half of the upper lip tightens upward.
v. **Disgust**: The individual's nose wrinkles and the upper lip rise while the lower lip protrudes.
vi. **Fear**: The eyes widen and the upper lids rise. The brows draw together while the lips extend horizontally.
vii. **Happiness**: The corners of the lips lifted and shaped a smile, the cheeks rise up and the outside corners of the brows pull down. [4]

B. FACIAL EXPRESSION DETECTION

Face features detection such as the mouth and the eyes is always one of the key issues in facial image processing as it involves wide and various areas such as the emotion recognition and face identification. Joseph C. Hager stated that face detection feature is used as one of the input to other image processing functions such as the face and emotion detection. Different researchers had studies on the different approaches in facial expression detection. Each approach can be applied effectively in different situation. [5]

In the year of 2004, W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat had proposed a method of combining the feature detection and extraction with the facial expression recognition into an integrated. With this system, the recognition process is not influenced by the subjective aspects and the bound of the areas are invariant during the sequence. The integration projection enables the detection and location of small and precise features of the face, such as the eyebrows and lips. [6]

Jagdish Lal Raheja and Umesh Kumar had introduced the Back Propagation Neural Network technique in human facial expression detection from captured image. In the technique published, edge detection, thinning, and token detection are carried out during the image processing process. Edge detection is aimed at identifying the points in a digital image at which the image brightness changes sharply or more formally has discontinuities. Thinning is applied in order to reduce the width of an edge, which is from multiple lines to single line. Token which generated after the thinning process divides the data set into smallest unit of information which needed for the subsequent processes. [7]

Besides above, Zhengyou Zhang (1998) had reported on the investigation on the feature-based facial expression recognition within an architecture based on the two-layer preceptron. Two types of factors are being derived from the face images during the investigation, i.e. the geometric positions of a set of fiducial points on a face as well as the set of set of multi-scale and multi-orientation Gabor wavelet coefficients at these points. Secondly, the sensitivity of individual's fiducial point to the facial expression detection is examined. Through the sensitivity analysis, the author found out that the points on cheeks and forehead carry little useful information. Lastly, the experiments show that facial expression recognition is mainly a low frequency process, and with a spatial resolution of 64 pixels X 64 pixels is probably enough to carry the process out. [8]

Apart from that, Eva Cerezoand the team had presented their works on an automated real-time system for facial expressions recognition which functioned by tracking the facial features' and the simple emotional classification method. The automatic feature extraction function enables the preamble of dynamic information in the classification system, which making the study of time evolution on the evaluated parameters as well as the categorization of user's emotions through live video possible. The developed system had been embedded in the Maxine system, an engine which managed 3D virtual scenarios and characters to enrich user interaction. [9]

In addition, Caifeng Shan, Shaogang Gong, Peter W. McOwan (2008) reported on the Facial expression recognition based on Local Binary Patterns: A comprehensive study. Local Binary Patterns is the texture operator that tags the pixels of an image by thresholding the neighborhood's pixel and represents the result in binary format. Result from the authors' experiments show that Local Binary Patterns features perform stably and robustly over a useful range of low resolutions of face images. This means image with low resolution can as well be processed and accurately identified the emotion. [10]

C. RELATIONSHIP BETWEEN MUSIC AND EMOTION

Many researchers had did research and studies on if the music can actually influence the emotion of individuals. For an example, Antoinette L. Bouhuys, Gerda M. Bloem, Ton G.G.Groothuis carried out a study in the relationship between the individuals' facial expression after listening to depression music. The results showed that depressing music bring on a major increase of depressed mood and significant decline if delighted mood. The study proved that music can actually influence individuals' emotions. [11]

Besides the study mentioned above, Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou presented on their research in the used of music in manipulating the young tennis players' emotional state. A total of fourteen young tennis players were involved in the study. The research signifies that participants choose to listen to music in order to elicit various emotion states. [12] Frijda mentioned in one of his article "increasing the tempo and intensity of a musical excerpt may increase the magnitude of an effective response and concomitant action tendencies" such as increased motor behavior. [13]

In 2001, Matthew Montague Lavy develops four basic assumptions regarding music lovers and their relationship to music. First, music is heard as sound. Secondly, music is heard as human utterance. Everyone has the ability to identify and detect emotion in the contours and timbres of vocal utterances. Third, music is heard in context. Music is described as a wide and complicated network of knowledge, thoughts and environment. All these are the factors which contribute to an emotional experience. Forth, music is heard as narrative. Listening to music includes the integration of sounds, utterances and context. [14]

D. EMOTION BASED MUSIC RETRIEVAL

Besides the study mentioned above which shows that emotions can be influenced by music, Wai Ling Cheung and Guojun Lu (2008) presented that automatic music emotion annotation is an important requirement to research music retrieval by emotion. The author pointed out that through this research, the automatic music emotion annotation is possible and workable using hybrid sampling, data-driven detection threshold and synonymous relationships between emotional. [15]

Maria M. Ruxanda1, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen (2009) studied on a number of dimensionality reduction algorithms, including both classic and novel approaches to test on the performance of applying multidimensional indexing over a dimensionally reduced audio feature space in audio extraction. "The proposed approach comprises of an effective methodology that projects the music into an audio feature space that captures the music emotion." [16]

III. METHODOLOGY

A. Research Methodology

Numerous study efforts regarding the project background and related works are carried out through available resources which are accessible personally or from the Information Resource Centre (IRC). Sources involved are books, papers, journals, thesis reports, online source and etc. The data gathered consists of current status of image processing, application of image processing and facial detection.

B. Development Methodology

The waterfall model is a sequential design process, in which progress is seen to be flowing steadily downwards. The project is done following one stage to the other from the first to the last stage.

First of all, during the planning stage, few researches are carried out to determine the needs of innovation in the specific field. Discussion is held with supervisor to determine the most suitable and executable project title.

The project is then moves to the analysis phase where the programming languages and techniques are learnt and studied.

The project is then moved to the design phase. The development process starts with the design of the framework and interface of the system. The major focus in design phase is to write program in order to detect the facial expression and to integrate it with music player.

Finally the proposed model will then be tested to find out if there any bugs and to test its functionality and accuracy in facial detection for various emotions. The Figure 1 below shows the process and activities throughout the project period.

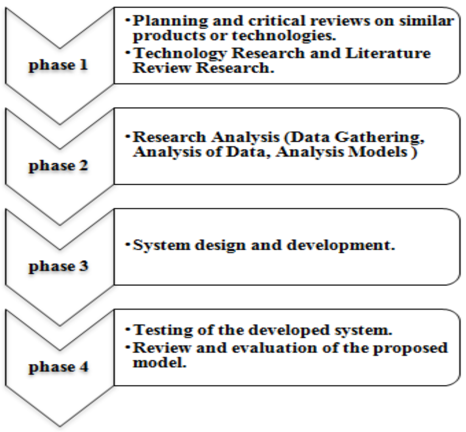


Figure 1: The Process Activities

IV. RESULTS AND DISCUSSION

A. Proposed Model.

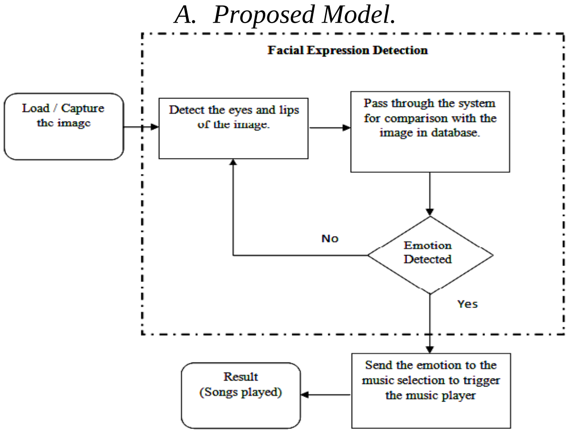


Figure 2: Flow Chart of the Proposed Model.

From the developer side, as for this FYP, the proposed model will be focusing on two main functions, first is the expression detection and secondly the songs played for each category of emotion. There will be sets of still image with the four different expressions available in the database for comparison purposes. User is able to customize the songs in each of the category according to their taste.

B. Interface Design and System Function

Figure 3: The user interface.

Functions of each button are as follow:

Table 1: Function of the buttons.

"Browse" button	Enable the user to select available image from the local disk.
Webcam "Start" button	Automatically connect the user to the image capturing device.
"Start" button	To start the image analysis process once the images is loaded.
"Restart" button	To restart the proposed model. It will start from the step of image selection.
"Setting" button	To save the set of images into the database for comparison purposes.
"Music Library" button	Enable user to customize the songs list for each emotions.
"Exit" button	To exit the proposed model.

The proposed model is made to recognize the user's emotion according to the following steps:

First of all, when an image is load, a conversion is done where to convert the colored image to gray scale image. Next, the model will get the cropped face height image and then divided into four parts. Eyes part divided into two parts (left and right). After the features are extracted, the thinning process will be carried out until the boundary of the eyes and lips are obtained. Lastly, the boundary is then converted into number and compared with the data in the dataset to classify the emotion of the data gained.

C. SYSTEM TESTING RESULTS

Functionality testing is done once the completion of the system development. The purpose is to check if the model is usable and well-functioned. The results are shown as table below.

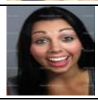
Table 2: System functional testing results

Component	Expected Function	Testing Result	
		True	False
"Browse" button	Direct user to the local disk and select image.	✓	
Webcam "Start" button	Start the embedded webcam / external webcam	✓	
"Start" button	To start the image analysis process once it is pressed.	✓	
"Restart" button	Clear the previously loaded image.	✓	
	Stop the currently played songs.	✓	
"Setting" button	Enable user to save the dataset into the database.	✓	
"Music Library" button	Direct user to local disk.	✓	
	Enable user to customize the songs for each emotion.	✓	
"Exit" button	Close the window of the proposed model.	✓	

D. EMOTION ACCURACY TESTING RESULTS

Set of images for the each category (normal, sad, surprise and happy) are saved in the model for comparison purposes. The model will compare the loaded image with the dataset in order to detect the emotion. Table below shows the dataset saved.

Table 3: The dataset of images in proposed model [17]

Images	Emotion	Images	Emotion
	**Normal**		**Sad**
	**Surprise**		**Happy**

The proposed model is tested with set of images of similar emotion to test on its. Ten images are tested for each category of emotions and the results are shown as tables below.

Table 4: The testing result for "Normal" Expression [17]


Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	
	✓				✓

Table 5: The testing result for "Sad" Expression [17]


Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓				✓
	✓				✓
	✓				✓

Table 6: The testing result for "Surprise" Expression [17]


Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓				✓
	✓				✓

Table 7: The testing result for "Happy" Expression [17]


Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	

Summary of Result are shown in the table below:

Table 8: The summary of the results tested

Emotion	No. of Samples	No. of Recognized Sample	RR
Happy	10	10	100%
Normal	10	9	90%
Sad	10	7	80%
Surprise	10	8	80%
Total	**40**	**34**	**85%**

In order to find the RR (Recognition Rate), the following formula is applied to the results collected as below:

$$RR = \frac{Classified\ Character}{Total\ Number\ of\ Character} \times 100\%$$

RR = 34/40 * 100

RR = 85%

Based on the result above, it shows that the proposed model has the recognition rate (RR) of **85%**.

E. DISCUSSION

Based on the observation during the prototype evaluation, there are several limitations which avert the proposed model to perform perfectly.

1. Posture of the object in the image.
   The proposed model will perform accurately if the object of the image is in upright posture and the individual's face is clearly exposed. Minor slant of individual's head is acceptable and detectable

Figure 4: Image which is accurately detected (a) [17]

Figure 5: Image which cannot be detected accurately (a) [17]

2. The quality of the image, either still image loaded or captured by webcam.

   The proposed model will perform better when the images loaded or captured are in better resolution, brightness and contrast. Warning message will prompt out when the image loaded in is too fine and the detection process stop.

   As shown in the images below, the image on the left can be processed and detected accurately as it has higher resolution compared to the right. The model able to detect the border of face features.

Figure 6: Image for emotion "Happy" [17]

The same limitation applied to the image captured through webcam. Figures below shows the error message prompting out as the image is too fine to be processed.

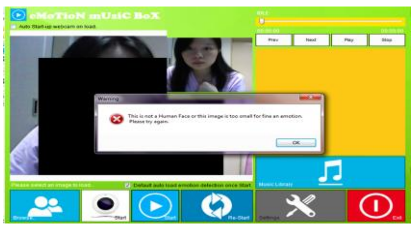


Figure 7: Error message

V. CONCLUSION AND FUTURE WORK

A. CONCLUSION

The significant of this project is the emotion detection of the images loaded into the proposed model. The main purpose is on its emotion detection functionality. Through the integration between emotion detection technology and music player, the proposed model is aimed to provide betterment in the individual's entertainment.

The proposed is able to detect the four emotions i.e. normal, happy, and sad. Once the proposed model detected the emotion, music player will play the song(s) accordingly. As for the usability and accuracy, both system testing and emotion accuracy testing has been done to the proposed model and return a satisfying result. The proposed model able to recognized 34 out of 40 images loaded into it, which give a Recognition Rate of 85%. Besides, the proposed model is a computer application which can works well in all kinds of windows and computers.

Thus with this Emotion Based Music Player, users can have an alternative way of selecting songs, which is in a more interactive and simpler way. It can help the music lovers to search and play songs according to their emotions automatically.

B. RECOMMENDATION

Every system is subject to upgrades and improvement, so do the Emotion Based Music Player.

First is to increase the accuracy in the emotion detection. It can be done by increasing the numbers of facial features used in emotion detection. Currently, the model extracts only the lips and eyes. As for future work, other facial features such as eyebrows and cheeks can be included. Besides, noise reduction software can be embedded in the model so that the noise for either still or captured image can be removed automatically.

Apart from the above, the proposed model can be improved by having auto adjustment on the resolution or brightness and contrast of the images. The accuracy of emotion detection for the current application is greatly influenced by the quality of the images loaded. Hence by having the auto adjustment, the user can load in any quality of image and the future model will be able to adjust the quality according to its requirement.

In addition, for the better interactive between user and application, real time emotion detection technique can be applied. The future model will detect and extract the facial feature once the application is launched and the emotion can be detected in real time.

ACKNOWLEDGMENT

First and foremost, the writer would like to express her deepest gratitude and appreciation to project main supervisor, Assoc. Prof. Dr. Baharum B Baharudin, who continuously monitoring the progress of the project throughout the semesters. The constructive comments and advices given have leaded to the successful outcome of the project.

This gratitude also dedicated towards the individuals who participated in the survey. The author highly appreciates for the feedbacks and cooperation given by the participants which have eventually facilitate a lot in developing and improvement of the system prototype.

Last but not least, the author would like to express her acknowledgement to family members and friends, for the support and verbal constructive opinions throughout the period.

REFERENCES

[1] Ekman, P. & Friesen, W. V. (1969): *The repertoire of nonverbal behavior: Categories, origins, usage, and coding:* Semiotica, Vol. 1: 49–98.

[2] Barbara Raskauskas (2009). Why Do People Listen to Music? Retrieved on October 9 2012 from http://voices.yahoo.com/why-people-listen-music-2608185.html.

[3] Emily Sohn(2011) Why Music Makes You Happy? Retrieved on October 9 2012 from http://news.discovery.com/human/music-dopamine-happiness-brain-110110.html.

[4] Mary Duenwald (2005). The Physiology of Facial Expressions. Retrieved on October 9 2012 from http://discovermagazine.com/2005/jan/physiology-of-facial-expressions

[5] Joseph C. Hager (2003). Introduction To The DataFace Site: Facial Expressions, Emotion Expressions, Nonverbal Communication, Physiognomy. Retrieved 10 October 2012, from http://face-and-emotion.com/dataface/general/homepage.jsp

[6] W.K. Teo, Liyanage C De Silva and Prahlad Vadakkepat. (2004): *Facial Expression Detection And Recognition System:* Journal of The Institution of Engineers, Singapore. Vol 44.

[7] Jagdish Lal Raheja, Umesh Kumar. (2010): *Human Facial Expression Detection From Detected In Captured Image Using Back Propagation Neural Network:* International Journal of Computer Science & Information Technology (IJCSIT). Vol.2(1).

[8] Zhengyou Zhang. (1998): *Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiments With a Multi-Layer Perceptron:* Journal of Pattern Recognition and Artificial Intelligence. Vol. 13(6): 893-911.

[9] Eva Cerezo1, Isabelle Hupont, Critina Manresa, Javier Varona, Sandra Baldassarri, Francisco J. Perales, and Francisco J. Seron. (2007): *Real-Time Facial Expression Recognition for Natural Interaction:* In J. Martí et al. (Eds.),: *IbPRIA 2007, Part II, LNCS 4478*, (pp. 40–47). Springer-Verlag Berlin Heidelberg.

[10] C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, May 2009.

[11] Antoinette L. Bouhuys, Gerda M. Bloem, Ton G.G.Groothuis. (1994): *Induction Of Depressed And Elated Mood By Music Influences The Perception Of Facial Emotional Expressions In Healthy Subjects*: Journal of Affective Disorders. Vol. 33: 215-226.

[12] Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou. (2007): *A Grounded Theory of Young Tennis Players' Use of Music to Manipulate Emotional State:* Journal of Sport & Exercise Psychology: 584-607.

[13] Frijda, N.H. (1986): *The emotions*. New York: Cambridge University Press.

[14] Matthew Montague Lavy (2001). *Emotion and the Experience of Listening to Music A Framework for Empirical Research*. (Unpublished master's thesis). Jesus College, Cambridge.

[15] Wai Ling Cheung and Guojun Lu. (2008): *Music Emotion Annotation by Machine Learning.* doi: 10.1109/MMSP.2008.4665144.

[16] Maria M. Ruxanda1, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen. (2007): *Emotion-Based Music Retrieval On A Well-Reduced Audio Feature Space:* In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 181-184.

[17] Stock Photography [Image]. Retrieved on October 9 2012 from http://www.dreamstime.com

[18] Human Facial Expressions [Image]. Retrieved on October 9 2012 from http://www.flickr.com/groups/1876998@N24/pool/.

[20] Microsof Studio C# Programming. (n.d.) Retrieved from http://www.functionx.com/vcsharp/index.htm