Unlock Screen Application Design Using Face Expression on Android Smartphone

Rhio Sutoyo , Jeklin Harefa, Alexander, Andry Chowanda

Bina Nusantara University, Jakarta, Indonesia

Abstract. Nowadays, smartphone has been shifted as a primary needs for majority of people. This trend is making people do activities on their smartphone such as exchange information, playing games, paying bills, discussing ideas, etc. In almost every smartphone, security features (i.e. password, pin, face unlock) are already built in which are intended to protect the data inside the smartphone. Nevertheless, those security features are not perfect and have flaws either technically or by human errors. This research aim to provide an application that uses people' facial expressions to improve the security of smartphone, namely Android. We implement face detection and expression recognition in this application as an improvement from the existing unlock screen, which is face unlock. The results that are achieved from this research is an unlock screen application on Android devices that utilizes people' facial expression changes to open their smartphone. This application can eliminates the technical errors from existing face unlock screen application which can be deceived by photos.

Keywords: unlock screen, face expression recognition, biometric.

1 Introduction

The growing of globalization era makes smartphone as a part of today's life for most of the people. Not only adults, but young and elder people also generally have smartphone. One of the growing smartphone in the world is Android. Based on the pie chart in the Fig. 1 below, it can be seen that Android users was increasing significantly, starting in 2009 that only about 4% from market share, to 67% in 2012. In 2015, it is predicted that Android users are still on the first rank of the market share.
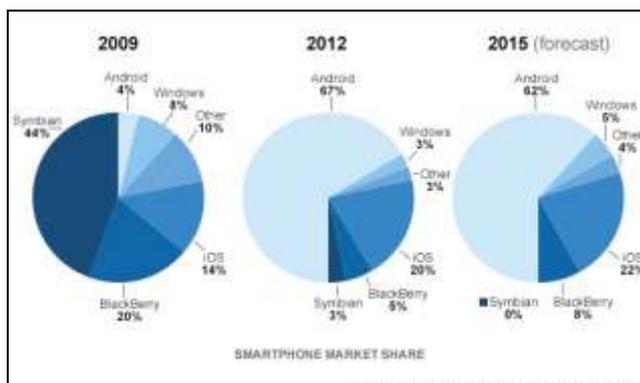

Figure 1. Statistic Diagram of Smartphone Sales

Based on that statistic, there is no wonder if many mobile vendors would prefer Android smartphone. Furthermore, the features in Android are growing as the increasing of Android demand and production. One of the feature is the security, which was originally just "a pin", now it has become more varied and sophisticated. For instance, pattern unlock, face unlock, and face and voice unlock. One of the most secure security system is a system that using biometric authentication [1]. Biometric is the study of the measurable biological characteristic. Biometric has a uniqueness of each individual. Android has implemented biometric on its face recognition lock. Face recognition has become an important topic in many application, such as security system, credit card verification, and identifying criminals. For example, face recognition is able to compare the face from the face database so it can assist in identifying criminals [2]. However, there is a weakness on that system which is if the users show photos of their face, the security system will detect that photos as the real user's face and the smartphone will be automatically opened. In conclusion, this flaw opens up strangers using users' photos to open their smartphone. Considering the weaknesses in the face recognition system, we believe we need a new kind of security system which can ensure that the faces shown in the system are not photos, but the real face of smartphone owners. In this work, we are building applications to eliminates those flaws by creating an unlock screen application with computer vision technology which uses people' face expression.

2 System Overview

Face recognition is a technology of artificial intelligence which allows computer to recognize people' face. It starts by capturing face photos through the camera. Then, the images will be processed so that face can be identified accurately by computer [3]. To the best of our knowledge, there are no face recognition tools that are perfectly able to distinguish between real faces and photos. As we have mentioned before, the goal of this study is to create a mechanism to overcome the inability of computer to recognize the real faces and photos by using facial expression changes. After successfully recognizes the face and the facial expression changes, the locked smartphone will be opened. To facilitate understanding of the process, a system overview will be presented in the following flow chart. There are two major processes, which are *Face Recognition Training* and *Unlock Screen*.

2.1 Face Recognition Training

The first major process is face recognition training. The purpose of this step is to acquire the data of users' face photos, process them, and stored them which will later be used for the unlock screen process. The overall system can be seen in the Fig. 2 below:

Figure 2. Face Recognition Training System

a. Face Detection

At the first step, face detection is done by using the front camera in the smartphone. The users must show their faces in the front of the camera and then the application will detect the face. In this process, we use the Viola Jones algorithm. For the classifier, we use data stored in XML files to decide how to classify each image location.

Figure 3. Users is Locating Their Faces in the Circle

In order to improve the accuracy, users are asked to locate their face in the shown circle on their smartphone (Fig. 3). Once the face is detected, the camera will take a few pictures of the face. Furthermore, the images will be cropped on the face area and the results will be resized to a size of 200 x 200.

b. Image Processing

Before we process the captured photos into the training step, they need to be treated first to decrease the training error rates.
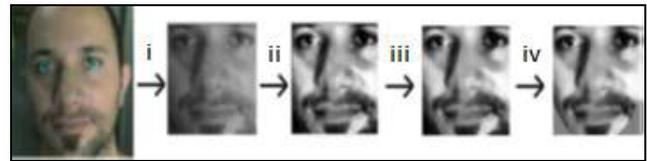

Figure 4. Face Image Processing

Therefore, the face images will be processed in the following process [4]:

   i. Geometrical Transformation and cropping.
     This process would include scaling, rotating, and translating the face image so that the eyes are aligned, followed by the removal, chin, ears and background from the face image.
   ii. Separate histogram equalization for left and right sides
     This process standardizes the brightness and contrast on both left and right sides of the face independently.
   iii. Smoothing
     This process reduces the image noise using a bilateral filter.
   iv. Elliptical mask
     The elliptical mask removes some remaining hair and background from the face image.

The example result of this process can be seen in the Fig. 4 where each photo is different and better than the previous photo.

c. Face Recognition Training

The processed photos from the image processing step will be used for data training in the current step. We conducted several test on various algorithms for the face recognition training such as: Eigenface, Fisherface, and LBPH (Local Binary Pattern Histogram). The result of this training process is stored in the form of an XML files which will be used later for identifying faces. For this step, we implemented the LBPH (Local Binary Pattern Histogram) algorithm because it has the best value of accuracy and FPS based on our test. The comparison of those algorithms will be discussed in the research result and discussion section.

2.2 Unlock Screen

Unlock screen is the second major process in the application. The purpose of this step is to compare the captured photos from the smartphone with the data training which is resulted from the face recognition training. If the photos are matched, the process will continue to the face expression detection. In that step, the users' face will be locked and they have to change their expression from natural expression to smile expression. The smartphone will be unlocked when those criteria are

completed. The overall system of unlock screen can be seen in the Fig. 5 below:

Figure 5. Application Framework when Opening the Screen

a. Face Detection and Image Processing

Both of these steps are similar with the processes in the face recognition training. These steps are done again in the unlock screen step to detect the users' face and process the captured photos before the system continue to the next step.

b. Face Recognition

The function of this step is to determine whether the captured and processed photos from the previous step is matched with the photos in the data training or not. The prediction will be determined by the number value of accuracy from the algorithm that we have implemented, which are LBPH (Local Binary Pattern Histogram).

c. Face Expression Recognition

Facial expression recognition consists of three major steps, including: pre-processing, feature extraction, and classification. The pre-processing step relates to face localization, tracking, and registration to remove the variability due to changes in head pose and illumination. The feature extraction aims to convert pixel data into a higher-level representation of shape, motion, color, and texture, which minimizes within class variations while also maximizing between class variations. The classification step aims to detect the action unit in the face [5]. In this work, we perform the face expression recognition after the face recognition step is succeed. The process is done by analyzing the facial point of users' face. When the expression recognition process is performed and the users change their expression (i.e. from natural to smile), the unlock screen function will be called. Furthermore, in this step the users' face will be locked after the face recognition step. Thus, if the face detection and face recognition step break (i.e. the users move their faces from the circle shown in the Fig. 3), the process will be back to face detection and face recognition again. We implement this method in order to avoid cheating (e.g. swapping photos of the users from natural expression to smile expression). In this step, we also have compared several algorithms such as Viola Jones, ASM and SVM, and Principal Component Analysis (PCA) in order to find the best algorithm for our application. Based on our tests, we choose the PCA algorithm. Our test results will be discussed in the research result and discussion section.

d. Unlock Screen

This function will be called when all of the constraints from the steps above are fulfilled. When this function is called, the system will unlock the phone screen and the users are able to use their phone normally.

3 Research Result and Discussion

We performed our tests with Samsung I9300 Galaxy S III that have Android OS, version 4.1 (Jelly Bean) installed. Moreover, it has Quad-core 1.4 GHz Cortex-A9 processor and have 1GB of RAM. The front camera specification is 1.9 MP, 720p@30fps. The resolution is 720 x 1280 pixels, 4.8 inches (~306 ppi pixel density) and it contains 5 MB of storage.

3.1 Face Detection

In this research, we apply Viola-Jones algorithm to detect the face in camera. There are five cascade classifiers that are tested, and one of them that most optimal both in speed and accuracy will be chosen. The results of each cascade is as follow:

Table 1. Comparison of Classifier Based on FPS and Accuracy

Classifier Name	FPS Range	Accuracy
Haarcascade_frontalface_alt_tree	~2-3 fps	96.296%
Haarcascade_frontalface_alt	~2-4 fps	96.296%
Haarcascade_frontalface_alt2	~3-4 fps	100%
Haarcascade_frontalface_default	~3-4 fps	100%
Lbpcascade_frontalface	~4-5 fps	100%

* *Higher frame per second (FPS) means faster speed*

Based on the table above it can be seen that the best classifier to perform face detection is *lbpcascade_frontalface*, which in this classifier, obtained the highest FPS with good accuracy rate when performing face detection.

3. 2 Face Recognition

There are three algorithms that we have tested, namely Eigenface, Fisherface, and LBPH (Local Binary Pattern Histogram). This process begin with an image training of the users' face. The output of this training is XML files that are used to identify the owner of the face. Once the training is done, the face image that want to be recognize will be tested and will gives prediction value as the output of the test. The prediction value will be used for determine the threshold value that serves as determinants whether the face recognized or not.

Table 2. Eigenface, Fisherface, and LBPH Algorithm Test Results

Comparison Subject	*Eigenface*	*Fisherface*	*LBPH*
Value prediction when testing with the same face	4633.81	318.59	29.32
Smallest value prediction when testing with different faces	2004.2	61.42	71.88
Biggest value prediction when testing with different faces	8360.78	2805.77	367.5
FPS Range	0.67	1.23	6.58

* *Lower prediction value means better face recognition*
* *Higher frame per second (FPS) means faster speed*

Based on the Table 2, it can be seen that the best algorithm to recognize face is LBPH algorithm. It obtained the highest FPS and accurate prediction value so that it can perform a better face recognition.

3.3 Face Expression Detection

At facial expression recognition, we compare Viola Jones with ASM (Active Shape Model) and SVM (Support Vector Machine), and also with PCA (Principal Component Analysis). The test conducted by calculating the accuracy and the FPS speed of the camera for each algorithm when the face expression detection is performed. We detail the result in the table below:

Table 3. Face Recognition Technique Test Results

Comparison Subject	Viola Jones	ASM and SVM	PCA
Accuracy	44 %	56.67 %	86.67%
FPS	1.74 – 4.36 fps	4.67 – 6.31 fps	4.13 – 5.57 fps

* Higher frame per second (FPS) means faster speed

Based on the tests, the PCA is the best technique for recognizing facial expression. It is because of PCA has the higher speed and the best accuracy among the others to perform facial expression recognition.

4 Related work

4.1 Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions

Tan and Triggs [6] tried to perform face recognition when light conditions are not controlled. They overcome this by combining powerful lighting normalization, local texture-based representation of the face, adjusting the distance transformation, kernel-based feature extraction, and the unification of many features. In their research, Tan and Triggs perform image processing which can eliminate the effects of light as much as possible but still be able to keep the details of the image that will be needed in the process of face recognition. Tan and Triggs also introduce a Local Ternary Patterns (LTP), a development of LBP which can distinguish the pixels with better and quite sensitive to noise in the flat areas. They showed that by changing the ratio of local spatial histogram transformations that have similar distance metrics, it will improve the accuracy of face recognition using LBP or LTP. Furthermore, they also improve their methods by using Kernel Principal Component Analysis (PCA) when extracting features and combines Gabor wavelets with LBP. A combination of Gabor wavelets with LBP produces better accuracy than the Gabor wavelets or LBP alone. The results of their research was tested on a dataset FRGC-204 (Face Recognition Grand Challenge version 2, experiments to-4) and produces an accuracy rate of 88.1% and 0.1% received error rate (False Accept Rate).

4.2 Fully Automatic Recognition of the Temporal Phases of Facial Actions

Valstar and Pantic [7] conducted a study in the facial expression recognition based on Action Unit. Action Unit is the smallest facial movements that can be seen visually. In their research, a video will be processed to find the position of the face in the image. Then after getting the position of the face in the picture, the system will finds 20 facial fiducial points automatically by using Gabor feature based classifier. Those facial points will be monitored to detect the expression changing. Facial movement data points will be classified by Support Vector Machine (SVM) and stored as an Action Unit. Valstar and Pantic has developed three versions to detect the movement of Action Unit, which are version 2005, 2006 and 2007. Valstar and Pantic successfully detect 22 Action Unit and detect automatically in 2007. Furthermore, for each Action Unit detected on the face, there will be a model of transient activation derived from several stages (neutral, onset, apex, offset). Action Unit at some stage is classified by the SVM to determine facial expressions. The results of research conducted by Valstar then tested and resulted in the introduction of Action Unit percentage of 95.3% on the image display facial expressions and by 72% in the image display spontaneous facial expressions.

4.3 A Dynamic Appearance Descriptor Approach to Facial Actions Temporal Modelling

In this study [5,8], there is a new approach to the analysis of the temporal dynamics explicitly produced by facial movements using a descriptor that can describe the dynamic changes in the face which is called Local Phase of Three Orthogonal Planes Quantisation (LPQ-TOP). Until now, most of the facial expression recognition system using only a static descriptor, where if there is a change in the movement of the face and other relevant information is ignored. This journal discusses these limitations by expanding LPQ static descriptor into Three Orthogonal Planes (TOP) which is inspired by a similar expansion in the Local Binary Patterns, combined with Gabor filter on the image histogram, LGBP-TOP [9]. These experiments have shown that the LPQ-TOP is very suitable for the detection problem Action Unit (AU) temporary. Moreover, it also can achieve detection of activation of Action Unit. However, when tested on a dataset GEMEP-FERA, Jiang et al concluded that the head does not pose an obstacle varies need to be resolved.

5 Conclusions and Future Work

In this paper, we have presented a novel approach to unlock phone with facial expression recognition method. Our approach utilizes the best algorithms for face detection, face recognition, and face expression detection based on our tests. These tests are important for the performance of our application because the application cannot detect and recognize the face correctly if the accuracy is bad and the FPS speed is slow. In the face detection tests, we notice that the Haar classifier has good accuracy but is too heavy to be implemented on Android

smartphone. On the other hands, LBP classifier has a pretty good FPS even though the accuracy is not as good as Haar classifier. Moreover, LBPH algorithm can be used to solve the problem of face recognition on Android smartphone. It is because the accuracy and FPS of LBPH is good enough to run on the smartphone. We also have implemented PCA algorithm for the face expression recognition by taking a picture of users' mouth and matched it with the existing facial expression. For the future work, we hope other general expressions such as surprised, angry, sullen, etc. can also be covered by the system. The reason for this is because not all people are able to perform good smiles which can easily captured by the system.

References

1. F. Ahmad, A. Najam, & Z. Ahmed, , Image-based Face detection and recognition: "State of the art", IJCSI International Journal of Computer Science Issues, **9**, issue 6, no. 1, 169-172, (2012).
2. G. Dashore, & V.C. Raj, An efficient method for face recognition using principal component analysis. International Journal of Advanced Technnology & Engineering Research, **II,** 2, 23-28, (2012).
3. P. V Saudagare & D. S. Chaudhari, "Facial expression recognition using neural network–An overview," International Journal of Soft Computing and Engineering (IJSCE), **2**, 1, 224-227, (2012).
4. D. L. Baggio, et al., Mastering OpenCV with Practical Computer Vision Projects. Packt Publishing, (2012).
5. B. Jiang, M. F. Valstar, B. Martinez, and M. Pantic, "A dynamic appearance descriptor approach to facial actions temporal modelling," IEEE Transactions of Systems, Man and Cybernetics-Part B, **44**, 2, 161–174, (2011 and 2014).
6. X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," IEEE Transactions on Image Processing, **19**, 1635-1650, (2010).
7. M. F. Valstar and M. Pantic, "Fully automatic recognition of the temporal phases of facial actions," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, **42**, 1, pp. 28-43, (2012).
8. B. Jiang, M. F. Valstar, & M. Pantic, "Action Unit detection using sparse appearance descriptors in space-time video volumes," IEEE Int'l. Conf. Face and Gesture Recognition (FG'11), 314-321, March (2011).
9. T. R. Almaev & M. F. Valstar, "Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition," Proc. Affective Computing and Intelligent Interaction, 356-361, (2013).