Applying the Kinect Device for Human Identification System in the ATM Environment

تطبيق جهاز كنيكت لنظام التعرف على الاشخاص في بيئة الصراف الالي

Preparation

Ali Akram Abdul Kareem

Supervisor

Dr. Sadeq Al-Hamouz

A Thesis Submitted in Partial Fulfillment of the Requirements for Master Degree in Computer Science

Department of Computer Science

Faculty of Information Technology

Middle East University

Amman, Jordan

January, 2016

Authorization Statement

I, Ali Akram Abdul Kareem, authorize Middle East University to supply hardcopies and electronic copies of my thesis to libraries, establishments, or bodies and institutions concerned with research and scientific studies upon request, according to the university regulations.

Name: Ali Akram Abdul Kareem

Date: 9/1/2016

Signature:

Examination Committee Decision

This is to certify that the thesis entitled "Applying the Kinect Device for Human Identification System in the ATM Environment" was successfully defended and approved on January, 9th, 2016.

Examination Committee Member	**Signature**
(Supervisor) Dr. Sadeq Al-Hamouz *Head of Computer Information System* *Middle East University*	
(Chairman of Examination Committee) Dr. Sharefa Murad *Department of Computer Science* *Middle East University*	
(External Committee Member) Prof. Ghassan Issa *Professor* *Dean Faculty of IT* *University of Petra*	

Acknowledgements

"In the name of Allah, the beneficent the merciful". Praise is to Allah lord of the world praise of the thankers. And peace be upon Muhammad, the messenger of God, the Al-Mighty, master of messengers of God and prophets, and those who followed him in charity until the Dooms-Day and day of resurrection. Afterwards, I thank the Al-Mighty God for he realized to me what I was always seeking to obtain the Master Degree in computer science.

And I submit great thanks to my supervisor, Dr. Sadeq AL-Hamouz for his efforts and beneficial science. He was the soul of granting help and assistance in my study and laid to me the correct plan for scientific research.

My special thanks and respect to members of the committee of discussion of this dissertation. I am indebted to their valuable comments and their reinforcing suggestions.

I also deliver my worn greetings and thanks to all the reverend members of the teaching staff in the Middle East University, with gratitude.

I also offer special thanks to Mr. Nabeel Shafeeq Aabbas for his continuous support to me during the period of study.

And I want to express my deep love and thanks to my parents, the father and the mother and brother and sister for their unlimited encouragement and their continuous support during my studies.

My thanks and gratitude to all my friends for their support and encouragement along the period of my study.

Dedication

My father my intimate archetype with whom I boast all world. I am fond of you! My father, because you did make me a man!

My mother, you are the gift of the AL-Mighty God, you are the top of warm-heartedness. All words of commendation cannot meet your heavenly rights. Thank you; you are the spring of giving and donation.

You reverend Professors, from you I have gained my lure, Knowledge, information and science! Thank you!

To all those who lent a hand to me and stood beside me in my toil of life and study, I offer this humble work!

List of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
3D	Three-Dimensional image
2D	Two-Dimensional image
RGB	Red, Green, and Blue
RGB-D	Red, Green, Blue, and Depth
ATM	Automated Teller Machine or, Automatic Teller Machine
PCA	Principal Component Analysis
IR	Infrared
HCI	Human–Computer Interaction
CoG	Center Of Gravity
BPNN	Back-Propagation Neural Network
SVM	Support Vector Machines
MKD	Multi-Key point Descriptors
MKD-SRC	Multi-Key point Descriptors, Sparse Representation-based Classification
GTP	Gabor Ternary Pattern
SDKs	Software Development Kit
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
LBP	Local Binary Patterns
PittPatt	Pittsburgh Pattern Recognition
CUDA	Compute Unified Device Architecture
AdaBoost	Adaptive Boosting
IPTS	Interactive Physical Therapy System
TOF camera	Time-Of-Flight camera

Applying the Kinect Device for Human Identification System in the ATM Environment

Preparation

Ali Akram Abdul Kareem

Supervisor

Dr. Sadeq Al-Hamouz

Abstract

Biometrics techniques are from the most attracting concern of research issues, for people were used, through these systems, to recognize each other. Researchers attempt to develop cameras to imitate the human brain in distinguishing between people by many techniques, were mentioned in the literature. Distinguishing between the human beings is being done by classical method in the image picked up by the visible light cameras but these cameras do not provide enough amount of information. So, the integrative nature of the depth information and RGB information, of which the Kinect camera is distinguished assists researchers in obtaining tangible results from cameras development. This thesis submits a model for face detection and recognition on the basis of the Kinect technique to some basic problems in the computer vision. This model was applied in the environment of ATM machine: firstly, to prove the reliability of the Kinect outputs. Secondly, detection about the depth of human face by using maps drawing to distinguish the real human face, and get rid of the fraud processes, from which technique of face detection and recognition suffer. Finally, we use the tracking algorithm that represents one of the system stages to provide the biggest amount of security. And in the end, tests are done by using our data base obtained from RGB camera in Kinect in our lab: for the system was tested with in a hundred systematic attempts to enter the system. The results were 91% acceptance to enter the system, and 98% were rejected in the return 100 attacking attempts against the system. And averages of human face depth detection are 93% in return for 100 fraud attempts at photographic form, were the Kinect camera in the most of detection experiments about surfaces was able to distinguish between the fraud process in the form of the photographic picturing, and the correct entering process (the correct user). And these results indicate superior abilities of Kinect camera.

Keywords*: Kinect Camera, User authentication, Face detection, Face recognition.*

تطبيق جهاز كنيكت لنظام التعرف على الاشخاص في بيئة الصراف الالي

إعداد

علي اكرم عبدالكريم

إشراف

الدكتور

صادق الحموز

الملخص

تقنيات القياسات الحيوية هي من المسائل البحثية الأكثر إثارة للاهتمام، حيث اعتاد الناس من خلال هذه الأنظمة التعرف الى بعضهم البعض. ويحاول الباحثون تطوير الكاميرات لتقليد الدماغ البشري في التمييز بين الناس بتقنياتٍ عديدةٍ ذكرت في الأدب. ويتم التمييز بين البشر بطريقة تقليدية في الصور التي يتم التقاطها بواسطة كاميرات الضوء المرئي، ولكن هذه الكاميرات لا توفر قدراً كافياً من المعلومات. لذلك فأن الطبيعة التكاملية لمعلومات العمق ومعلومات RGB التي تتميز بها الكاميرا Kinect تساعد الباحثين في الحصول على نتائج ملموسة من تطور الكاميرات.

وتقدم هذه الرسالة نموذجاً للكشف والتعرف الى الوجه على أساس تقنية Kinect لحل بعض المشكلات الأساسية في رؤية الكمبيوتر. وتم تطبيق هذا النموذج في بيئة الصراف الآلي: اولاً، لإثبات موثوقية مخرجات Kinect، ثانياً، والكشف عن عمق وجه الإنسان باستخدام رسم الخرائط لتمييز الوجه الانساني الحقيقي والتخلص من عمليات الخداع التي تعاني منها تقنيات الكشف والتعرف على الوجه، واخيراً نستخدم خوارزمية التتبع التي تمثل واحدة من مراحل النظام لتوفير قدر أكبر من الأمن.

وفي النهاية، يتم اجراء الاختبارات بأستخدام قاعدة البيانات الخاصة بنا التي تم الحصول عليها من الكاميرا Kinect في مختبرنا، اذ تم اختبار النظام في غضون 100 محاولة نظامية لدخول النظام، وكانت النتائج هي 91٪ قبولاً لدخول النظام، وتم رفض 98٪ مقابل 100 محاولة هجوم ضد النظام، ومعدلات كشف عمق وجه الانسان هي 93٪ مقابل 100 محاولة احتيال بصورة فوتوغرافية، حيث

كانت الكاميرا Kinect في معظم تجارب الكشف عن الأسطح قادرةً على التمييز بين عملية الخداع في شكل التصويرالفوتوغرافي، وعملية الدخول الصحيحة (المستخدم الصحيح)،وتشيرهذه النتائج إلى القدرات الفائقة للكاميرا Kinect.

الكلمات المفتاحية: *الكاميرا كنيكت، مصادقة المستخدم، كشف الوجه، التعرف على الوجه*

Chapter One

Background of the study and its importance

1.1 Introduction

The image processing is one type of signals processing, where inputs are an image, such as: a photo or, video frame; the output may be either an image or, a grip of characteristics or, parameters associated with the image.

Most image-processing techniques cover two-dimensional picture, and applying standard signal-processing techniques. The digital image processing consists of a number of basic steps such as image acquisition, image enhancement, pre-processing, edge detection, segmentation, representation, description, matching and recognition. The output of these steps is either an image or, an image attributes. The digital image is made out of a limited number of components, each of them has a specific location and value. These elements are referred to as image elements or pixels, which are going to be processed image for converting into a digital form. There are many steps in the digital image area. The most important steps in this area are edge detection, segmentation, image matching and recognition processes. Multi-work has been done in different areas of image processing, where it focused on the biometrics field, because it is a good way for person identification. The Biometric fields are of the most interesting research subject matters in the last years. The human brain used to employ this field recognize any person through different traits and attributes such as: the face, the body, the hand, the weight, the sound and so on (Sadeghi, M., Maghooli, K., &Moein, M. S. 2014) (Mohammed, T. S., Al-Azzo, W. F., & Al Mashani, K. M. 2013).

The face attribute is considered the most important features to recognize people, because it gives so much information about the person and personality. Also simple accessibility and possibility, easy usage and ability of better acceptance among persons. The researchers attempt to develop the computer system that can exactly imitate the human brain in identifying and recognizing human faces by the computer technology that identifies human faces in digital image processing. The human face is processed in two procedures: face detection procedure, and face recognition procedure (Mohammed, T. S, et al., 2013).

Face detection is an interdisciplinary field that integrates different techniques that coordinate distinctive methods such as (image processing, pattern recognition, computer vision, computer graphics, physiology, and evaluation approaches). Face detection of targets is the first step in most automatic vision systems, for example, content-based image retrieval, video coding, video conferencing, crowd surveillance, and intelligent human–computer interfaces. Nonetheless, it was not as of not long ago that the face discovery issue got impressive consideration among analysts (Hjelmås, E., and Low, B. K., 2001).

Face recognition is a generally utilized biometric system with promising arrangements and application fields. The sort of utilizing securing can order its methodology. Three dimensions (3D) face recognition is known for attaining high recognition rate and for being exceptionally secure, particularly against attacks, then again, obliges the utilization of very convoluted and costly gadgets. Face recognition additionally obliges a huge processing exertion and, by and large, the cooperation of the client. Two dimensions (2D) face recognition methodologies give a minimal effort and exceptionally usable recognition framework. That is because of the required equipments such as: (The basic optical camera), and the exerted effort of processing is low and does not need cooperation from client. The attacks of face pictures are one of the primary issues in the field of 2D face recognition arrangements, with a final goal to counteract attacks, many scientists submitted distinctive methods to differentiate between real pictures and the unreal ones (assault cases). These methodologies are generally alluded to as online face detection (Kahm, O., and Damer, N., 2012).

1.2 Problem statement

Automatic Teller Machines (ATM) fields have become an important technology, which needs a specific security; because the customer depends on the ATM to conveniently meet his banking needs. In the same time, there are many methods of proliferation of ATM frauds. So it is a must to provide the technology that works to minimize frauds problems, propose solve them.

Today, user identification systems play an important role in security field, but there are still some deficiencies, because of the many factors that continue to pose a significant challenge when applying the detection and recognition of people, including techniques trick normal camera by identifying the users that their images are available. Thus this field needs to focus on vulnerabilities to protect ATM.

The Kinect device presented different methods to solve classical problems in computer vision, because it has many features such as: low cost, reliability, speed of the measurement, ability to provide RGB (Red, Green and Blue) data and depth data. Therefore, it can solve problems in the normal camera, and provide security to the ATM machine.

In this research, the problem is represented in how to recognize the user authenticates or not, and how to eliminate cheating. Therefore, the Kinect features investigated: protection of the problem and reduced the danger. In this research, numerous inquiries were founded as follows:

- ✓ How will be selecting the identification for the user?
- ✓ What is its verification degree in perceive of the user?
- ✓ What is the utilizing environment in applying idea?
- ✓ What are investigated Kinect features in this research?
- ✓ How will the Kinect's features be utilizing in this research?

1.3 Significance of the study

The significance of this study conceals in specifying a proposed technique to fill gaps, from which techniques of face detection and recognition to be depended at different environment through proposing a model submitting solution to increase trust in results of these techniques.

1.4 Objectives

The main objectives of this thesis are:

1- Finding an acceptable method that prevents cheating face detection and recognition techniques.

2- Designing and implementing a system that provides security to the proposed environment.

3- Using the recognition technique of user based on results of Kinect camera at the dangerous environment, such as: Automated Teller Machine environment.

4- Obtaining the best results to specify the identity and reduce the rate of mistakes.

5- The results will be discussed through an experiment on the different user's images.

1.5 Motivation

Our motivation leans, in turn, on techniques of face detection and recognition are still worried because of plenty of gaps that do not qualify them to be used in applications of dangerous environments. Benefits of these techniques extend into in control and security applications, because of low costs and easiness of obtaining information to specify the identity of the human being, being deals with the face. So this thesis proposes a model solution of the classical problems for these techniques.

1.6 Contribution

The face detection and recognition are the methods of verifying, or identifying a face from its image; these methods assisted various systems to provide appropriate security.

Therefore, the proposed system presented an approach to the identification of the human face. The goal of this research is to discuss the following points:

- Design a system provides security through the user identification techniques by using low-cost (Kinect) camera.
- Avoid the error rate of face recognition algorithm by using motion property in Kinect.
- Using RGB data and depth data altogether from Kinect camera to distinguish face between photograph and original face, and so it is preventing cheating operations and human recognition.
- Applying systems of static biometrics and dynamic biometrics in the proposed system.
- The Kinect device had a role in decision-making of cashing operations through identifying the user, whether he is (authorized or not).

- Using the proposed identification system to provide security at an automated teller machine environment.

1.7 Thesis organization

This thesis will be organized as follows:

Chapter One: Introduction.

Chapter Tow: Detailed information for literature reviews for basic concepts of face detection and face recognition.

Chapter Three: Illustrates the scientific techniques that are used in research work.

Chapter Four: The main phases of the proposed model have been discussed. In addition to that, the methods of each phase and the methods and techniques used in the design have been discussed in details.

Chapter Five: Conclusion and Recommendations.

Chapter Two

Theoretical Background

And

Literature Review

2.1 Introduction

Face detection and recognition is one of the most active areas in computer science, so there are a lot of efforts and researchers for more than 30 years submitted by those interested in image processing field, whether they are researchers, computer engineers or developers. Due to achieve great results in the face detection and recognition fields, through Kinect device, that represents one of the latest devices that are used in an unlimited number of applications to specify the identity of the human (Ghiass, R. S., Arandjelović, O., Bendada, A., &Maldague, X. 2014).

There are many research areas for this device that include: hand-gesture recognition, human activity recognition, body biometrics estimation, 3D surface reconstruction, and healthcare applications (Han, J., Shao, L., Xu, D., &Shotton, J. 2013). So, it is impossible to survey all related work to it. This chapter attempted to focus on the main research methodologies related to the face detection and recognition through the Kinect device, to achieve the major aim of the proposed work.

2.1.1 Face detection

In general, detection is a type of image segmentation technique, leads to simplifying the image processing; the amount of data that are of most concern minimized in the image (Mohammed, T. S., et al, 2013).

Face is considered as one best way to identify the people, because it gives so much information about the person and personality such as: the distance or angle between mouth, nose or the left and right eye. These details called structural relationships, which are useful to constrain the facial feature detection process (Tin, H. H. K., 2012). So, the face detection is a necessary first-step in face recognition systems, with the purpose of localizing and extracting the face region from the background. The goal of this step can simply be described by determining that there are any faces in the image or not. The face detection processing has to detect the faces of people, both males and females, from different races, depending on their appearance (facial, hair, glasses/no glasses), orientation and background. Human face has a complex structure and image may have complex background, so to discover human face in a given photo has specific steps to be done beginning with the dimension reduction for data of face image then analyze the image to locate a face and focus face structure to be identified, as shown in figure 2.1, (Maraqa, R, 2014).

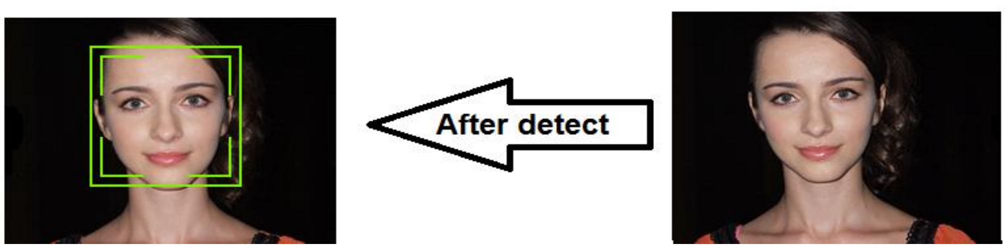

Figure 2.1: Detect Image

There are many methods to classify face detection, but mainly divided into four categories according to the structural relationships (Maraqa, R, 2014):

- The first category is called Knowledge-based that studies the patterns shape of faces that give information of the face structure component and the relation between these components.
- The second category is called feature invariant or feature-based approaches, finding facial features that are invariant to pose, lighting or rotation. These structural

components are tested in different conditions of illumination and position factors that affect detection and recognition results.

- The third category is called template matching which is spare layouts for facial feature structure, calculate the correlation between a test image and pre-selected facial formats, may utilize one format or more for better results, where one format is possibly considered insufficient, which needs high memory assets.
- The fourth category is called appearance-based methods using the teaching system through training sets of face images to help it find the facial feature structure.

After literary narration of face detection methods, we can say each method has its own methodology, restrictions and determinants in searching for the face features that may be unsuitable for other ones that make the comparison among different algorithms limited, where no standards of testing the algorithm for evaluating their results are available (Maraqa, R, 2014).

In 2001, Viola and Jones propose Object Detection utilizing Haar feature-based cascade classifiers in their paper (Viola, P., & Jones, M. 2001). It is considered as machine learning based method, where a cascade function is trained from a lot of positive images (images of faces) and negative images (images without faces). It is then used to detect objects in other images (testing images) (Cobos-May, C., Uc-Cetina, V., Brito-Loeza, C., & Martin-Gonzalez, A. 2015). Although it is more than 15 years since this method was proposed, it is still being used for training robust systems which require fast patterns detection such as (detecting human faces in noisy images) (Rezaei, M., Nafchi, H. Z., & Morales, S. 2014). It has been applied not only on face detection, but also to detect or recognize many other kinds of objects, such as; (hand gestures, text within images, and even vehicle detection). We have used Haar algorithm for face detection, which needs minimum computation time while achieving high detection accuracy (Cobos-May et al, A. 2015).

2.1.2 Face Recognition

Biometric technologies that currently offer greater accuracy, such as fingerprint and iris, require much greater explicit cooperation from the user. A biometric framework is basically a pattern recognition system that works by securing biometric data from an individual, extracting a feature set from the acquired data and comparing this feature set against the layout set in the database. Biometric recognition is the automated recognition of individuals based on their behavioral and biological characteristic, Biometric recognition has been applied to the identification of people (Bowyer, K. W., Chang, K., & Flynn, P. 2006).

In the computer vision field, the process of human face recognition system consists of three components: first component is face detection and face verification which means to detect the locations, sizes of face image in the arbitrary digital pictures or dynamic video picture of scene and extracting the face region from the background. The second component is feature extraction which transfers the original human face image data into a reduced representation set of features, which can represent the main information in this image. The third component is face recognition, which determines the right person in static picture or dynamic video picture, and then search for matching against a collection of images already registered into the human face image database (Ma, X., Liu, C., & Zhao, L. 2015).

Eigenface which represents Principal Component Analysis (PCA) considered as one of the best and most effective approaches used in face recognition systems. It has become the mainstream criterion to test the performance of various human face recognition systems. The advantage of this approach is simplicity, speed and insensitivity to small or gradual changes on the face. However, the traditional algorithm of Eigenface has some deficiencies in technology systems, because of the many factors, including techniques (dealing with lighting, changing facial expressions, facial change mode), posing as a hindrance in the practical application. Also, it can be cheated by simply showing a photo of a registered user. Some problems are limited to files that can be used to recognize the face when the human face images must be vertical frontal views, this is due to use the simple hardware needed for acquisition (e.g. a simple optical camera). It is known most researches endeavors in face identification field had concentrated on the Two-dimensional image to provide a low cost (Hg, R. I., Jasek, P., Rofidal, C., Nasrollahi, K., Moeslund, T. B., &Tranchet, G. 2012) (Slavković, Marijeta, and Dubravka Jevtić 2012).

In spite of that recognition of the face with 3D images improve averages of recognition but they require using highly complicated devices and with high price. Availability of algorithm with high accuracy (Eigenface), and depth information unentrusted depth device (Kinect sensor), lead to solve problems of high prices of the sensor devices and problems of face recognition in 2D images from which algorithm of recognizing the face suffered. Kinect devices are extremely popular and machine through operations of face recognition, the matter that makes them play an important role in different fields, such as security, carrying out the low, documentation banking service (Hg, R. I., et al 2012).

2.1.3 Kinect Camera

The face recognition of components suffers from general object recognition challenges, in addition to suffering from distortions or covariates specific to faces such as expression, accessories, and high inter-class similarity of human faces. The key recognition system is how to efficiently combine heterogeneous features, such as RGB images, and depth images. 2D images are providing information about a face, while the depth cameras have implicit advantages of handling illumination changes and identity protection. It is known Kinect sensor has the possibility of providing depth information and RGB images. This makes these devices quite suitable for many applications including 2D and 3D facial analysis systems (Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. 2013).

The Kinect technology depends on projecting Infrared (IR) patterns on the scene by utilizing IR pattern projector and estimating the depth map; the IR projector casts an IR speckle dot pattern into the 3D scene, while the IR camera catches the reflected IR dots. It also contains a Motorized Tilt (is a pivot for sensor adjustment). The sensor can be tilted up to 27◦ either up or down. Therefore, Kinect device is an instance of a structured light depth sensor. The IR projector into the 3D scene is invisible to the RGB camera, but can be viewed by the IR camera where RGB camera provides three basic colors the video and can offer images at 640×480 pixels with 8-bit per channel, The depth camera operates at 30 Hz, as shown in figure 2.2, (Han, J., Shao, L., Xu, D., &Shotton, J. 2013).

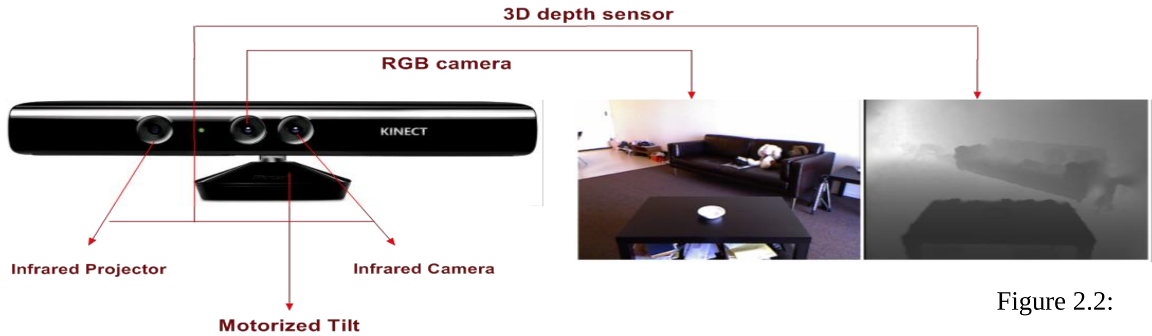

Figure 2.2: Hardware configuration of Kinect. In addition, Depth information and RGB-D information which obtained by using an infrared camera and RGB-D camera (Han, J., et al 2013).

The Kinect camera utilized the depth information to find the closest object camera, which is a legitimate assumption for most facial analysis systems, as it is typically the case that a user stands before the camera in these systems. After reducing the search space of the input image, the depth information is used another time to find some face candidate regions. The data in depth images may contain areas that are considered undefined depth, where the raw depth information are extremely precarious and many pixels in the image may have no depth due to multiple reflections, transparent objects or scattering in certain surfaces (such as human tissue and hair). These areas are usually small, but they can seriously effect on the face detection result. Therefore, it should be first checked if the user's face has any such holes, in this case, it has to recover prior using. A set of filters has been used for this purpose, as shown in figure 2.3, the results of using such filters (Hg, R. I., et al 2012).

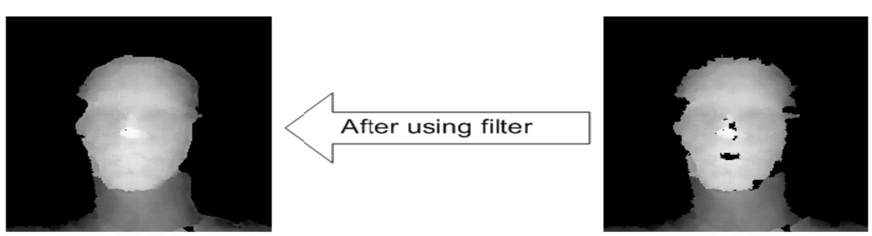

Figure 2.3: filling up the holes (Hg, R. I., et al 2012)

Kinect sensor was initially used as an input device by Microsoft for the Xbox game console. With a 3D human motion capturing algorithm, it enables interactions between users and a game without the need to touch a controller. The advent of this kind of sensors has spurred the developers, engineers, and researchers to work in the field of objects tracking and gesturing recognition. Human gesture recognition is defined as a non-vocal communication, used instead of or in combination with a verbal communication, intended to express meaning. It may be combined with the hands, arms or body, and also can be a movement of the head, face and eyes. Human gesture can be called several names: human pattern, human posture, human pose, and human behavior (Kumar, M. S., &Babu, R. V. 2012) (Han, J., Shao, L., Xu, D., &Shotton, J. 2013).

Gesture recognition is an extremely important element for building advanced Human–computer interaction (HCI) systems. The ability to automatically understand, not only hand gesture specifically defined by communication protocols, but also hand and body natural movements, will open to a completely new set of applications. There are many researches in the gesture recognition area since it is an extensive application in virtual reality, sign language recognition, and computer games. But these researches still are not satisfactory for real-life applications, because the quality of captured images is sensitive to some circumstances such as lighting and cluttered backgrounds. The optical sensor is unable to more gesture recognition, one effective way to achieve it is using other sensors to capture the motion, but those sensors are usually more expensive than optical sensors. There is a way, not a very popular in the real life, that is gesture recognition by using Kinect sensor can also gesture detection and recognition, and thus it provides a valid base for gesture recognition

(Ren, Z., Yuan, J., Meng, J., & Zhang, Z. 2013) (Pedersoli, F., Adami, N., Benini, S., &Leonardi, R. 2012).

This study focuses on use of the Kinect camera, because it is the latest consumer market camera which is enabling users to control applied system with body motion and gestures without extra input devices.

2.2 Literature Review

In the recent time, human face recognition has been based on computer vision and video-based techniques, in which the execution of recognition depends fundamentally on light conditions, shadow, camera angles, and head position. So, the system performance may suffer in these circumstances using a 2D camera (Maraqa, R, 2014). There are many researchers presented in this area, every one of them employed different methods for solving the mentioned problems:

Thomas S. Huang (1998) presented the progress towards face recognition using video sequences, defining a visual learning method, and its application to face detection in a complex background, and accurate facial feature detection/tracking. While, a fast algorithm for 2D-template coordinating is proposed as well as its application to face recognition. At last, the authors report an automatic, real-time face recognition model that utilizes video sequences continuously to: that uses video sequences continuously to: (i) detect the presence of a face in the input video, (ii) accurately locate/track the outer eye corners, (iii) match the input face with the database, and (iv) learn faces by adding entries to the database based on the confidence level of the matching and tracking procedures (Colmenarez, A. J., and Huang, T. S., 1998).

Kakadiaris, et al (2006) presented algorithmic solutions to most of the challenges faced by 3D facial recognition systems. By utilizing a deformable model we map the 3D geometry information onto a 2D regular grid, thus combining the descriptiveness of 3D data with the computational efficiency of 2D data. In the same time, scalability in both time and space is realizing by changing 3D facial scans into compact wavelet metadata. This paper displays results on the largest known, and now publicly-available, Face Recognition Grand Challenge 3D facial database consisting of various thousand scans. This method has produced the highest accuracy on this dataset (Kakadiaris, I. A., Passalis, G., Toderici, G., Murtuza, M. N., &Theoharis, T. 2006).

ABAS et al (2010) focused mainly on the employment of an efficient and low-level technique for facial identifications within the thermal spectrum. They had demonstrated a novel strategy in infrared imagery examination. This is finished by initially applying seeded region growing technique for background filtering, followed by a 3-valued threshold decomposition for thermal region decomposition, and introducing the Center Of Gravity (CoG) points alongside with its respective thermal values representation as feature to be extracted from each decomposed thermal region. This approach was tested on terrific facial IR database that is publicly available, where it achieved 81.46% in correct identification rate. Utilizing an inexpensive minimum distance measurement strategy between strings of attributes acquired from test and enrolled images do arrangements. Empirical results obtained shows that with the non-complex and low computational cost of this proposed approach, the performance surpasses the Eigenfaces approach and it could also achieve higher performance rate if more registered images with multiple angular poses were to be considered, plus more intensive analysis were to be conducted. It too suggests more complicated analysis and decomposition approach is valuable to be studied. Since the proposed methodology is independent of facial features; such as the length of eyes, nose, and mouth. The application of this finding is substantial for security vehicle-robots' vision system, since possible images are acquired without prior knowledge of the situations' surround lightings and individuals' pose (Abas, K. H., and Ono, O. 2010).

Oprisescu, S., et al (2012) this paper presents an automatic algorithm for static hand gesture recognition relying on both depth and intensity information provided by a time-of-flight (ToF) camera. The compound depth and intensity information facilitates the segmentation process, even in the presence of a cluttered background (2 misses out of 450 images). Gesture classification is based on a decision tree using structural descriptions of partitioned contour segments. Classification was tested on 9 different gestures. The final mean recognition rate is satisfactory, of about 93.3%. Recently we have tested the algorithm also with a Kinect camera, and the results are a bit worse because of the Kinect’s depth image instability and poor depth resolution (Oprisescu, S., Rasche, C., & Su, B. 2012) .

Preis et al (2012) explain a method for gait recognition based on Microsoft Kinect, with an integrated depth sensor allowing for skeleton detection and tracking in real time. The authors estimate the number of body features synchronically with step length and speed, their relevance for user identification, and display the results of an empirical evaluation of the model, where this research was ready to accomplish a success rate of more than 90% with nine test persons. The authors utilize 13 biometric features like height, the length of limbs, and step length which are calculated from the skeleton frames produce by Kinect. Based on test data from different users, the three basic classifiers Naive Bayes, 1R, and C4.5 were trained and evaluated concerning the success rate of their classification.

Based on the features utilizing in the decision tree C4.5, the authors attain out that just four features, namely height, length of legs, length of torso, and length of the left upper arm, were enough to correctly identify a user in 91% of all sample handle the complete video from the specific experiment and the Naive Bayes classifier. Classification based solely on step length and speed still yielded 55.2% success rate practice either Naive Bayes or the decision tree. The authors consider that the results from gait recognition with Kinect are assuring and showing that reliable discrimination of individuals in a small set of users is possible. However, a larger trial setup should present more insight into the modification of body parameters. Particularly in application scenarios with many people, tracking and classifying the trajectory of certain limbs like hand and feet could add to the accuracy of the model. Furthermore, a mixture of different identification systems like facial recognition could add to the dependability of such a system (Preis, J., Kessel, M., Werner, M., &Linnhoff-Popien, C. 2012)

Zhang et al (2012) introduced a framework for recognizing activities of daily living to facilitate the independence of older adults living in the community, reduce risks, and enhance the quality of life at home by using RGB-D (red, green, blue, depth) cameras. The researcher's contributions include three aspects. First, to detect abnormal activities which are dangerous for elderly people, we recognize 5 activities related to fall including standing, fall from standing, fall from sitting, sit on chair, and sit on floor. Second, to recognize finer activities of daily living, the researchers propose a discriminative representation of structure-motion features based on skeleton joints. Third, to continually track same person when there are multiple people appear in the same camera view, the researchers further develop a binary classification based person identification method by combining appearance and depth information. The proposed framework is evaluated on a dataset we collected under different lighting conditions for fall detection and a benchmark dataset for daily living activity recognition. Experiments demonstrate that our framework is effective and robust to recognize activities related to falling event and finer activities of daily living. Our RGB-D camera-based framework can handle lighting changes and pose variations, as well as provide a good solution for privacy protection. This method correctly labels most classes of activities and achieves an average accuracy rate of 98.1%. For the activity of "cooking (stirring)," some of the videos are mislabeled as "cooking (chopping)", because these two types of activities are very similar and very difficult to distinguish only from the skeleton features they have extracted. The average accuracy of this method for "New Person" test is 81.79% (Zhang, C., &Tian, Y. 2012).

Patsadu et al (2012) proposed human gesture recognition using Kinect camera. In particular, the researchers are involved in a specific stream of the vector of twenty body joint positions, which are characteristic of the human body captured by Kinect camera. The recognized gesture patterns of the study are stood, sit down, and lie down. In the experiment, the camera is located on the plane of the body in two distance settings (2-3 meters). Data mining classification methods, which include BPNN (back-propagation neural network), SVM (support vector machines), decision tree, and naïve Bayes, are investigated for gesture recognition. Test results have confirmed that the back propagation neural network system outperforms other classification methods and can perform recognition with 100% accuracy. Furthermore, the average accuracy of all classification approaches apply in this study is 93.72%, which confirms the high potential of adopting the Kinect camera in human body recognition applications (Patsadu, O., Nukoolkit, C., &Watanapa, B. 2012).

Tayal, Y., et al (2012) suggest a skin based segmentation algorithm for face detection in color images with detection of many faces and skin regions. Skin color has confirmed to be a helpful and robust cue for face detection, localization and tracking. Though there are some cases of false positives, the overall performance of the proposed algorithm is quite satisfactory. The training images on which the algorithm is tested are natural images taken under uncontrolled conditions. The efficiency of the face detection was found to be 73.68% (Tayal, Y., Lamba, R., &Padhee, S. 2012).

Ong, W, et al (2013) the researchers demonstrate the use of K-means clustering and simple template models to achieve human activity detection and recognition in an unsupervised manner. The features used are extracted from the skeleton data obtained from an inexpensive RGBD (RGB-Depth) sensor. The clustering achieved an average of 80.4% precision and 83.8% recall. The recognition performance achieved an average precision of 81.4% and recall of 83.3%. These results suggest the potential of performing unsupervised human activity recognition using just the skeleton data from an inexpensive RGBD camera (Ong, W. H., Koseki, T., & Palafox, L. 2013).

Liao et al (2013) submit a general face recognition method that does not need face alignment by eye coordination or some other fiducial points. They explain an alignment-free face representation approach based on Multi-Key point Descriptors (MKD), where the descriptor size of a face is selected by the actual content of the image. In this case, a large dictionary of gallery descriptors can sparsely describe any probe face image, holistic or partial. A new key point descriptor called Gabor Ternary Pattern (GTP) is also produced for robust and discriminative face recognition. Trial results are summarized in four public domain face databases (FRGCv2.0, AR, LFW, and PubFig) under the open-set identification and verification scenarios, Comparisons with two leading commercial face recognition SDKs (software development kit) (PittPatt (Pittsburgh Pattern Recognition) and FaceVACS) and two baseline algorithms Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Local Binary Patterns (LBP), (PCA+LDA and LBP). In case a partial face cannot be detected, the approach proposed can still provide a matching score given a manually cropped face region. Given the general framework of MKD-SRC (Sparse Representation-based Classification), it would be useful to apply MKD-SRC to other image classification areas, such as object categorization (Liao, S., Jain, A. K., and Li, S. Z. 2013).

Li, B. Y. et al (2013) discuss an algorithm that uses a low-resolution 3D sensor for robust face recognition under challenging conditions. A preprocessing algorithm is introduced which exploits the facial symmetry at the 3D point cloud level to acquire a canonical frontal view, shape and texture, of the faces irrespective of their initial pose. This algorithm also fills holes and smoothes the noisy depth data provided by the low-resolution sensor. The canonical depth map and texture of a query face are then sparse approximated from separate dictionaries learned from training data. The texture is adapted from the RGB to Discriminate Color Space before meager coding and the reconstruction errors from the two sparse coding steps are added for individual identities in the lexicon. The query face has assigned the identity of the smallest reconstruction error. results are performed applying a publicly available database containing over 5000 facial images (RGB-D) with varying poses, expressions, illumination and disguise, acquired utilizing the Kinect sensor. Recognition rates are 96.7% for the RGB-D data and 88.7% for the noisy depth data alone. The results paper justify the feasibility of low-resolution 3D sensors for robust face recognition (Li, B. Y., Mian, A. S., Liu, W., & Krishna, A. 2013).

Goswami et al (2013) introduce a novel algorithm that uses the depth information along with RGB images obtained from Kinect, to raise the recognition performance. The suggested algorithm calculates a descriptor based on the entropy of RGB-D faces along with the saliency feature obtained from a 2D face. The probe RGB-D descriptor is accepted as input to a random decision forest classifier to establish the identity. Also, this paper presents a novel RGB-D face database consisting of RGB-D images pertaining 106 subjects captured exclusively utilizing Kinect is prepared. Illustrated the identification accuracy along with standard deviation of face recognition algorithms on the IIIT-D RGB-D face database is 91.6 ± 1.2 and the EURECOM Kinect Face Dataset is 98.1 ± 1.1. The results, performed on Kinect face databases, indicate that the RGB-D information obtained by Kinect can be adopted to achieve improved face recognition performance compared to existing 2D and 3D approaches (Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. 2013).

Ghiass et al (2014) present a comprehensive and timely review of the literature on this subject. There key commitments are:

(i) A synopsis of the inherent properties of infrared imaging that makes this methodology promising in the context of face recognition,

(ii) A deliberate audit of the most influential approaches, with a focus on emerging common trends and additionally key differences between alternative methodologies,

(iii) A description of the fundamental databases of infrared facial images accessible to the scientist,

(iv) And lastly exchange of the most promising avenues for future examination. Additionally, it was reviewed that a range of data sets currently available to researchers. Considering the results distributed to date, in the conclusion of these authors two particularly promising ideas stand out: the first thought is the improvement of identity descriptors in light of constant physiological features such as vascular networks, and the second thought is utilization of methods for multi-modal fusion of complementary data types, most notably those based on visible and infrared images. Both are still in their initial stages, with a potential for huge further improvement (Ghiass, R. S., Arandjelović, O., Bendada, A., &Maldague, X. 2014).

Meng, R., et al (2014). Submit a real-time face recognition system based on CUDA that effectively completed the face detection and recognition tasks. The proposed face recognition system has a high acceleration performance, because that the researchers not only did parallel optimization to detection part, but also did it to recognition part in this system. In the face detection phase with Viola-Jones cascade classifier, the authors executed and developed novel parallel methodologies of image integral, calculation scan window processing and the amplification and correction of classifiers. In the face recognition phase, the authors examine the parallelizing of the algorithm and parallelized some part of the testing phase. Through the optimization of the two major part of face recognition model, the suggested model makes a big difference. The results demonstrate that, in comparison with traditional CPU program, submitted program implements a high speed. This increase in speed will enable real-time face recognition applications on occasions that need more real-time applications (Meng, R., Shengbing, Z., Yi, L., &Meng, Z. 2014).

Wei Lee. H, et al (2015) used Kinect skeleton information to conduct body height measurements with the purpose of improving character recognition performance. If the search scope settings for the database are set to within ±5 cm of the height, the value of height measurement reduces the amount of time spent making comparisons in the database. Simultaneously utilizing height and facial features to conduct character recognition can similarly reduce the frequency of mistaken recognition. Accepting the Kinect for measurement and tracking would allow the use of depth functions to obtain the actual height of the character on the screen, increasing the precision of the search. The approach suggests in this paper can be used in public safety, student attendance registration, commercial VIP recognition and many others (Lee, H. W., et al. 2015).

Solomon & Wang (2015) proposed a system that detects driver fatigue and distraction have been developed using non-invasive machine vision concepts to monitor observable driver behavior. Moreover, cellular telephone detection was also considered. Solomon & Wang also explore how driver error can be reduced through inexpensive technology (a Windows developer Toolkit Kinect) which observes driver behavior and reacts when certain unwanted patterns in behavior have been recognized. The experiments have been done to validate the efficiency of this new system. The detection and control system monitors pitch angle, roll angle and yaw angle. If any of the angles exceeds a predefined value for a certain period of time, a warning is triggered, to remind the driver to be attentive. The angles were also used to determine if the driver was nodding off or whether he has fallen asleep. If a driver uses a cellular telephone while driving, the control system uses the X, Y and Z coordinates of the driver's wrist joint and compares this coordination and vector to the driver's head position. The warning is triggered if the head position and wrist position are in close proximity of each other, but only (Solomon, C., & Wang, Z. 2015).

Summary

There are plenty of researchers submitted in the filed of detection and recognition through the face and walking. Some of these researches used Kinect camera, to make the primary objective of each research available.

The proposed model works to enlarge the range of the camera; Kinect outside the limits of games, and used in operations of human identification through face detection and recognition by using Eigenface algorithm. This model is different from other researches mentioned below.

- Description: application of the apparatus, Kinect to face detection and recognition in the real life being, to prove the characteristic traits outputs of Kinect compared with classical cameras in the context of problems in the classical vision, such as (deception operation across a photographic form).
- Characteristic traits of the proposed method:
  - The model will use the Kinect camera (low cost device) in the banking environment. This surely will belittle the average of deceptive activities on

Automatic Teller Machine in order the card-owner only can reach his banking account

- Avoid the error rate in the algorithm of recognition the face (Used to authorize the user) through using the motion trait layed as a stage in the proposed method to make available the greatest possible amount of security.
- In the proposed model, the Kinect device from had a role in finance exchanging decision-making from the banking accounts by excusing the allowed person only to enter.
- The proposed model submits a new method to differentiate between the real human face and human face available in a photographic form and this operation leads to get rid of deception operations done by technicalities of face detection and recognition.

Usually, the image of the face, acquired leans on the visual spectrum may affect the security performance of systems of detecting the face, because they provide a limited amount of information. And so, there will be fears from using them to make security available in many environments. Among the different approaches, that had been proposed, in an attempt to overcome these restrictions through picturing by the infrared. The proposed model will make a detection method of the depth of the human available by using RGB and depth information obtained from Kinect camera. This method makes the technicalities of face detection and recognition more trustful in the different environments. The follow chapter will illustrate the scientific technicalities used in solving these problems.

Chapter Three

Methodology

3.1Introduction

In the last years, face recognition had attracted the important research benefits, because of its large usages in applications of control and applications of presence and presence and departure. The classical techniques attempted to achieve recognition of faces by using 2D images. But this type of images provides a limited amount of information about the face. And this creates a challenge in recognizing the face. The researchers considered 2D incorporation with depth information is one of the best solutions to provide what is enough of information to face detection and recognition (Hg, R. I., et al 2012).

In spite of merging 2D images with depth information led to improvement in averages of recognition compared with using 2D images only, but high prices of specialized sensor devices costs limits the application of these curriculums in the practical application. Appearance of new sensor technology, such as (Kinect sensor) it is characterized with low cost, high speed in picking up and small size. All these characteristics make Kinect sensor more attracting to applications of face recognition (Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. 2013). The researcher seeks the possibility of using RGB and depth information in face detection and recognition using Kinect. The purpose of the proposed model is finding a trustful method to solve problems, from which the techniques of detection and recognition by using the data submitted by Kinect. Therefore, this chapter offers a description of the proposed model that provides security to protect Automatic Teller Machines. Figure 3.1 illustrates the three procedures and the main plans at each procedure.

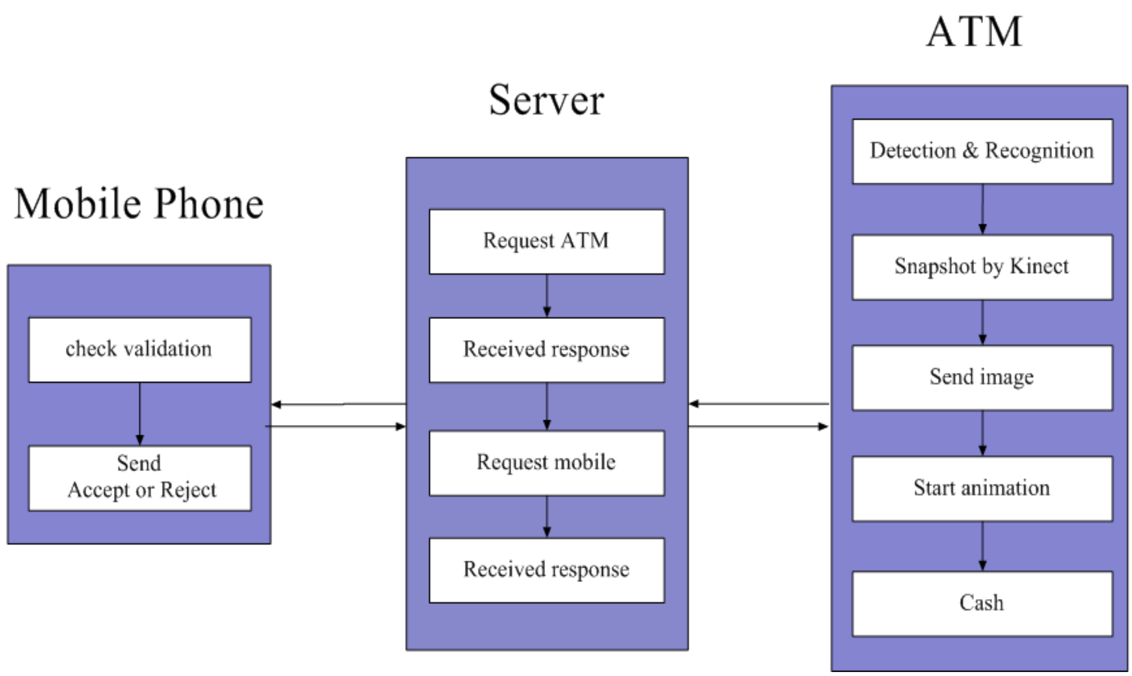

Figure 3.1: Proposed Model

3.2 Proposed work

The proposed model aim is identifying a client through face recognition process. A clear and reliable methodology was determined by using data obtained from the Kinect camera.

3.3 Methodology

The proposed model provides security to the Automatic Teller Machine through the techniques of face detection and recognition, that is by using depth and RGB data, obtained from the Kinect Camera. This Model provides security through making sure if the user is allowed to enter or not by recognizing the face as illustrated in the figure 3.6 and 3.7

Algorithm (Methodology algorithm)

This algorithm is implemented in several steps to reduce Kinect problems at operating outside gaming field. All these steps attempt to find a suitable solution to client identification problem as declared below.

Step1: Enter the required PIN code

//The user enters his/her code that is recorded it during the registration system

Step2: Check if the PIN Code is true then Go To step3

//Comparison between entered PIN code with stored PIN code

Else Go To step 6

Step3: Applying Haar and Eigenfaces algorithms to identify the user of ATM machine

//Applying Haar algorithm to determine the user's facial area and Eigenfaces algorithm to recognize user's facial

Step4: Check if the user is authorized then Go To step 5

//Ensured reliability of the user depending on the Haar and Eigenfaces algorithms

Else Go To step9

Step5: Check if depth image is true then Go To step7

//Ensured images taken by Kinect camera is whether of the user or the virtual image, as shown details in figure 3.3

Else Go To step12

Step6: Check if the pin counter is less than (3) then Go To step1

//Allowing the user to experience the PIN Code up to 3 times

Else Go To step16

Step7: Start animation

//The user performs its own animation.

// Clarification of the animation of this side is done as follows:

//The employed movement is lifting the right hand at the level of the head in order the complete hand-fist be higher than the fore-head of the user, and then raising the left hand at the same level of the right hand. The purpose of this movement is to fix the points at a certain level, through which the camera can accept the movement, or reject it.

//The differentiation method is done by specifying the location the selected points to a number of frames (30 frames per one second). So as the person is accepted and rejected through these points, and these points are represented with the right hand, the head and the left hand.

Step8: Check if animation is true then Go To step15

// Assurance of the animation to the user through Kinect camera

Else Go To Step 16

Step9: Taking a Snapshot to user

// Taking image of the user

Step10: Send the Snapshot to phone

//Sent the snapshot to the account owner mobile for authenticated user

Step11: Check if response timer is less than (30) then Go To step12

// Give a specific time duration of 30 seconds to receive the answer from the account owner

Else Go To step 16

Step12: Check if response phone is accepted then Go To step 13

// Depending on the account owner response is the approval of the user reliability or not

Else Go To step16

Step13: Start animation

//The user performs its own animation (more details in step 7)

Step14: Check if animation is true then Go To step15

//Assurance of the animation to the user through Kinect camera

Else Go To Step 16

Step15: Process

//Implementing specific processes for cashing

Step16: Finish

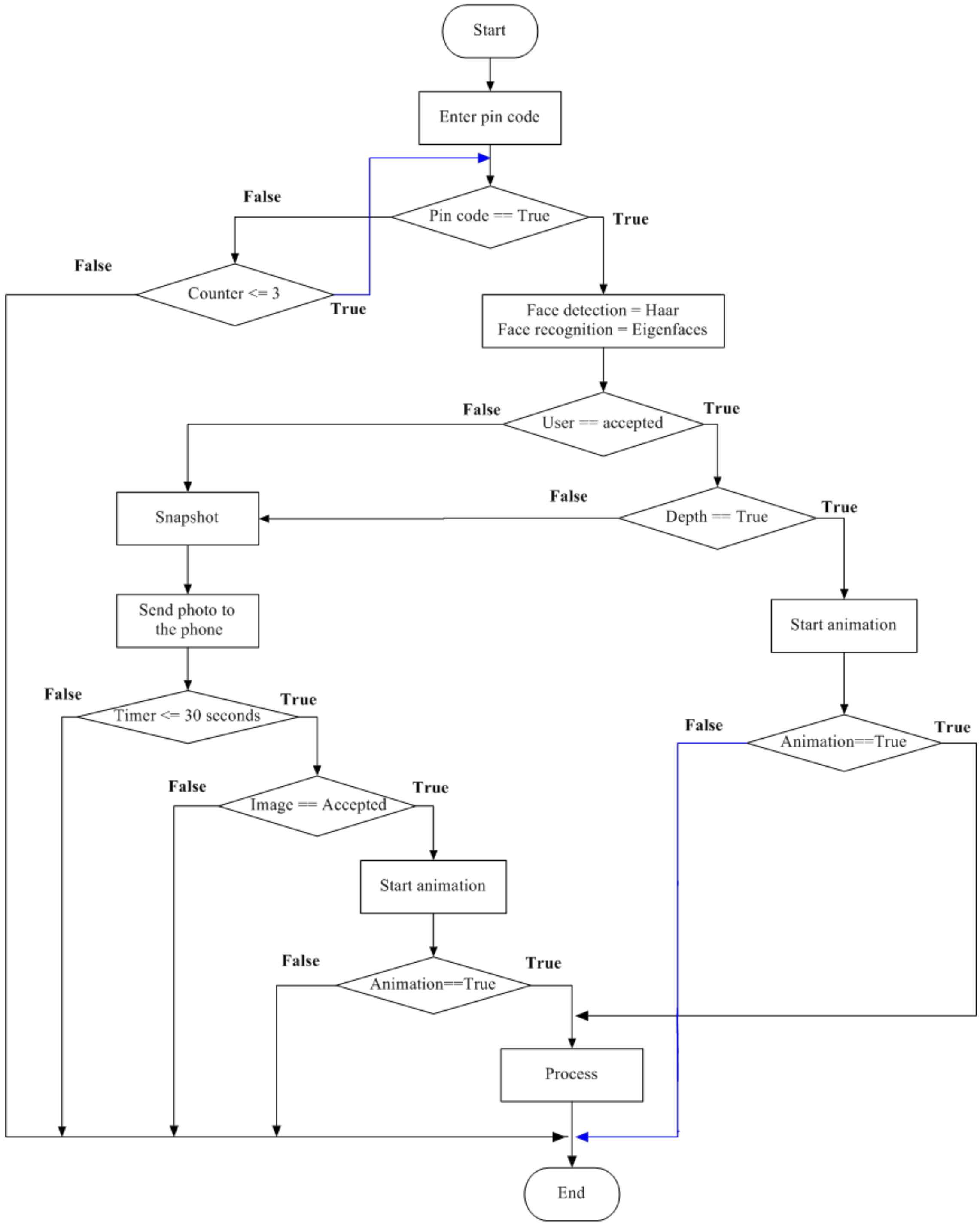

Figure 3.2: Methodology of the proposed approach

3.3.1 Implementation of depth detect:-

In this thesis, the researcher implements model in two procedures: the first procedure called RGB frame procedure which applies standard Haar and Eigenface algorithms to detect and recognize human face, the second procedure called Depth frame procedure which gets the minimum and maximum reliable depth for the current frame.

These procedures produced results for mapping between locations on the depth image and their corresponding locations on the RGB image by using handy utility, named CoordinateMapper(opensource:https://msdn.microsoft.com/enus/library/windowspreview.kinect.coordinatemapper.aspx). After that, the researcher determine seven points on the human face depending on Haar algorithm results by shifting some pixels from the image boundary of the nearby ribs.

Matching operation among each two points (matching between the depths of top points, matching between the depths of middle points, and matching between the depths of bottom points), all these depths points are almost similar because of the facial symmetry, then matching the depths of each point against the maximum point. Thus, when there are similarities between the depths of each point and the maximum point that leads to refuse trick a photograph. The following flowcharts declared the main steps of Depth Detection algorithm as shown in figure 3.7.

Algorithm (Depth Detection)

This algorithm represents Depth Detection approach; it is differentiating between deception operation as photographic form, and correct operation of entry (true user) through several steps as declared below:-

Step1: Take Snapshot to user.

// Taking Kinect Snapshot of the user.

Step 2: Applying Haar and Eigenfaces algorithms to detect and recognize human face.

//Applying the algorithms to determine the user's facial area on RGB frame

Step 3: Calculate the minimum and maximum reliable depth for the current frame (on the user face).

//Calculated the minimum distance (maximum value) and maximum distance (minimum value) of the user image

Step4: Building mapping between locations on the depth image and RGB image.

// Build mapping using CoordinateMapper tool

Step5: Determined seven points on the human face

//Identified points depending on the distance from the boundary 20 pixel upon rectangle of Haar algorithm

Step 6: Checking if the top points is similarity then Go to step 7

// Check if differences in depth between two top points are similarity and it does not exceed the rate (0-2) millimeters

Else Go To step11

Step 7: Checking if the bottom points is similarity then Go to step 8

// Check if differences in depth between two bottom points are similarity and it does not exceed the rate (0-2) millimeters

Else Go To step11

Step 8: Checking if the middle point is similarity then Go to step 9

//Check if differences in depth between two middle points are similarity and it does not exceed the rate (0-2) millimeters

Else Go To step11

Step 9: Checking if the Maximum point is not equal the other points then Go to step 10

// Check if differences in depth between the Maximum point with other points (The other 7 points on the human face) are not equal and it is greater than (10) millimeters.

Else Go To step11

Step 10: The user is authorized then Go to step 12

//The user is reliable to complete the requested operation, go to animation process as shown in figure 3.2

Step 11: The user is fake

// Depth information not detected, it is considered cheating process through displaying a photo of a registered user, the camera reports for the owner of account about cheating process as shown in figure 3.2

Step 12: Finish

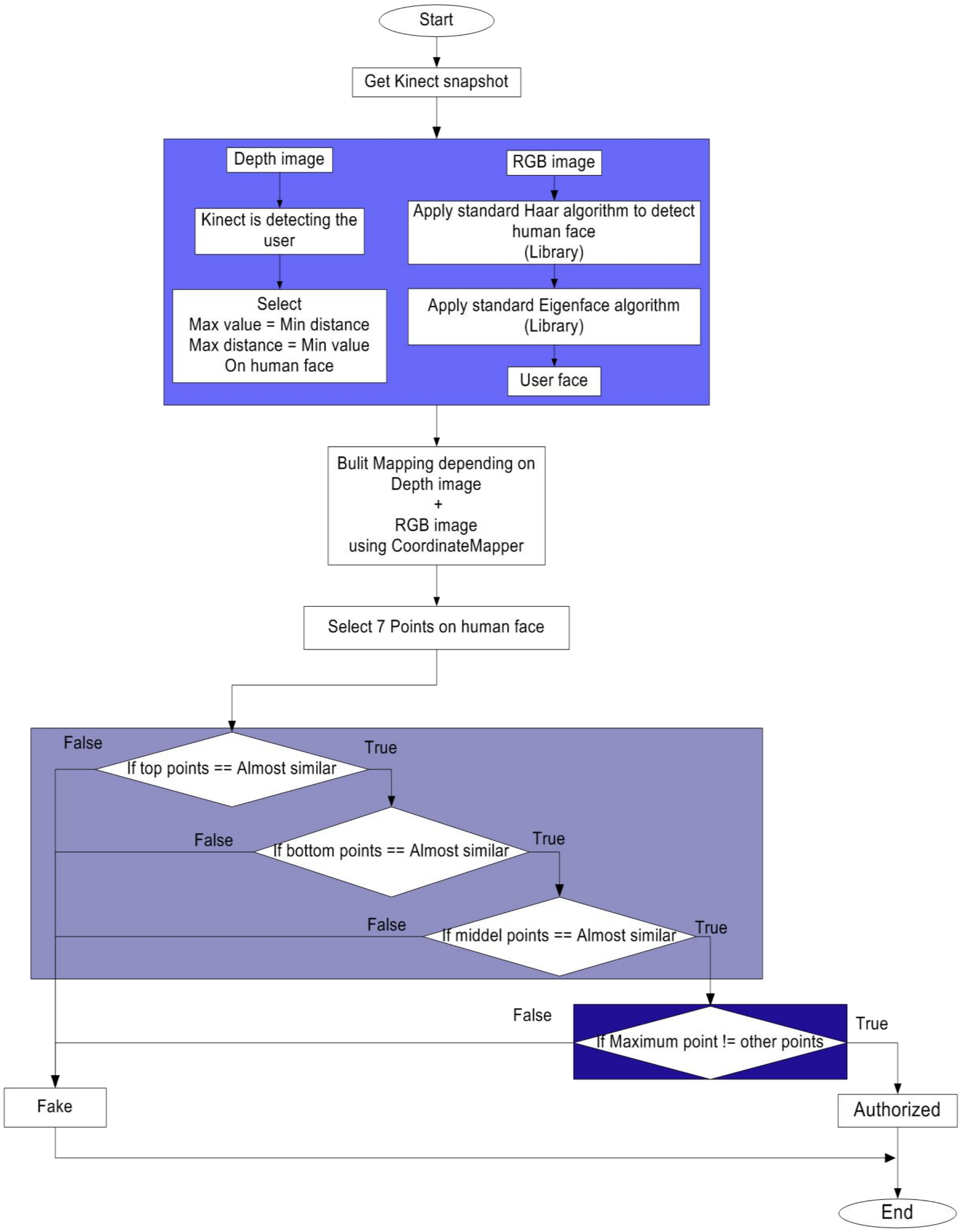

Figure 3.3: Depth detection

Chapter Four

Evaluation

And

Experimental results

4.1 Introduction

This chapter studies the results of the model obtained by applying Kinect device for face detection and recognition in the real-world, shown by using Haar with Eigenface algorithms. The objective is proving Kinect traits of outputs compared to classical cameras in the context of classical vision problems, such as (fraud by photograph). The model given for used techniques (face detection, face recognition, depth mapping and tracking).

4.1.1 Haar cascade classifiers

One of the most famous successful algorithms is the Viola & Jones approach, called Haar Classifiers, to rapidly detect an object, including human faces, using AdaBoost (Adaptive Boosting) classifier cascades that are based on Haar-like features and not pixels (Wilson, P. I., & Fernandez, J. 2006). The proposed model utilized the Haar classifier to detect human face as a first step for face recognition.

4.1.2 Eigenfaces

A face recognition system was implemented by using an Eigenfaces approach which represents the Principal Component Analysis (PCA). PCA method is a small set of significant features used to describe the variation among face images. Advantage of this approach is its simplicity, speed and insensitivity to small or gradual changes on the face, but it can be cheated by simply showing a photo of a registered user (Slavković, M., & Jevtić, D. 2012). The proposed model is approaching, able to utilize depth information from Kinect sensor for addressing the challenging cheat problem.

4.1.3 Depth information

Depth map representation is an important issue in many applications. In spite of information in the depth map produced by the Kinect sensor looks insignificant for feature extraction, it can be used to extract meaningful information from an image (Li, B. Y., Mian, A. S., Liu, W., & Krishna, A. 2013). Where is applying the model depend on the result of Haar algorithm by determining seven points to distinguish between human face as photographic and real human face. So the model detects and recognizes human uses RGB and depth information.

4.1.4 Animation

Difference in the animation from one person to another assists in differentiation among persons, where this difference by depending on time of animation, and technique of the animation. We can get benefit from this principle, considered as on basis of Kinect camera work bases in our research, in addition to the employed technicalities. The principle of animation in Kinect camera is represented in a group of points as an ordinate to joints of the human body. If these points were adjacent to each other, Kinect camera cannot recognize the movement powerfully.

4.1.5 Simulation Database (database file)

The purpose of establishing the database is to prove the operation of investigation, database file includes 25 RGB images taken from 25 persons upon whom experiments were done, and achieving recognition operation on persons and through which specifying the rate of validity of the system and its failure could be done, and the rate of the possibility of its acceptance in the proposed environment.

4.2 The setup of the system

In this section all system will be run, the ATM simulation, virtual server and android application as shown below.

Figure 4.1 shows cheating process through displaying a photo of a registered user. The camera refuses cheated operations, because depth information not detected. The camera takes a snapshot of the user and sends it to the server. The server sends a snapshot to the mobile.

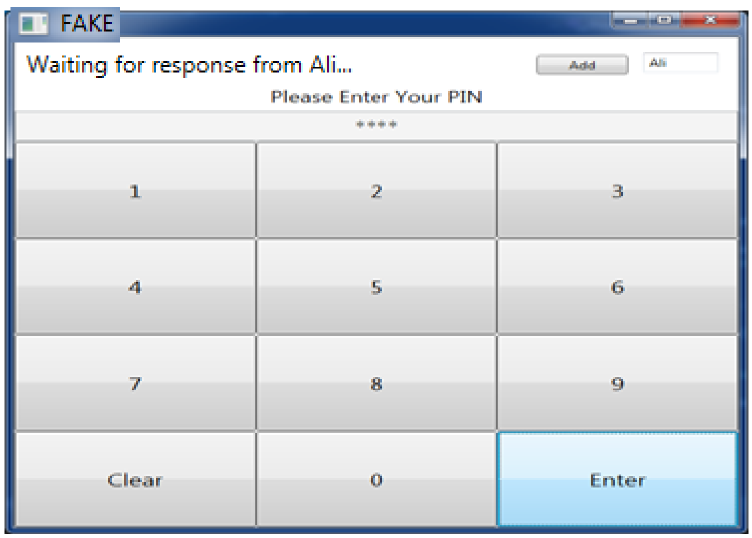

Figure 4.1: Cheated operations by showing a photo of a registered user

Figure 4.2 the researcher in this figure selected seven points on the human face to determine depth human face to prevent cheating operations by the photographic images.

Figure 4.2: Seven points on human face

Figure 4.3 shows the server which opens a channel between Kinect camera and the phone to receive the requests and responses between them.

Figure 4.3: Kinect server

Figure 4.4 shows android application works on owner's mobile. It is waiting for 30 seconds to answer (Accept or Reject). If owner didn't respond the user of ATM will receive a message (owner didn’t response, please try again later!).

Figure 4.4: Android application

4.3 Experimental results

The experimental results show that our method can recognize the persons by clear and reliable methodology, where the system will be able to identify the user who poses a threat, or a certain risk in order to reach the highest degree of protection in various fields. The experiments are applied as following:

1) The distance should be (500mm - 1500mm) for the Kinect camera.

2) Head is fixed angle (without any rotate) of the Kinect camera (0º).

Kinect device was placed on a table approximately in a height of the head of the user. Differences in heights among users were reduced by adjusting of the Kinect's height. The distance of the user from Kinect camera was measured using Manual engineering tools and we placed small laser at the same angle of Kinect that shines through the lens which splits it into a vertical line, this line placed on the middle of human face (to place human face in fixed angle).

4.3.1 Test of using photographic images for each individual

This section described the experiments performed to evaluate our method. The main steps of proposed algorithm have been tested using the real persons and the results are shown in table 4.3.

Experiments execute through simulation database, which contains RGB images obtained by using Kinect camera. Results have shown, Averages of detecting depth information are 93% rejected, and 7% accepted. The false acceptance rate is caused by effect of Haar algorithm performance and the implications or overlap in the IR, that sometimes occur as a result foldable image to defraud the system, where Kinect camera in most detection tests of surfaces was able to distinguish between deception operation as photographic form, and correct operation of entry (true user). This proves through this thesis, where these tests have been carried out, through 100 attempts to penetrate 10 persons' accounts from different ages at photographic forms at rate of 10 attempts for each user.

All attempts were completed at different conditions, from part of (remoteness of photos from camera and status of face in the photo and difference of enlightenment). Results of this experiment justify the feasibility of using depth information in the Kinect camera at different fields. The following tables are describing the experiments to select best points (which achieved highest rate in less time) to determine depth human face and detect deception by a photographic photo to record the highest protection rates in the shortest time for ATM security. As shown in figure 4.5, figure 4.6, and figure 4.7.

Table 4.1: Results of applying a detection programs on 3 points

Depth Detection by 3 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	1	9	6.9 MS
Talal	10	0	10	6.9 MS
Nada	10	3	7	6.9 MS
Hassan	10	2	8	6.9 MS
Mahmood	10	1	9	6.9 MS
Yousef	10	4	6	6.9 MS
Aseel	10	1	9	6.9 MS
Said	10	0	10	6.9 MS
Mohamed	10	3	7	6.9 MS
Café officer	10	0	10	6.9 MS
Rate	100	14%	86%	6.9 MS

Table 4.2: Results of applying a detection programs on 5 points

Depth Detection by 5 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	0	10	11.5 MS
Talal	10	2	8	11.5 MS
Nada	10	0	10	11.5 MS
Hassan	10	1	9	11.5 MS
Mahmood	10	3	7	11.5 MS
Yousef	10	0	10	11.5 MS
Aseel	10	1	9	11.5 MS
Said	10	0	10	11.5 MS
Mohamed	10	2	8	11.5 MS
Café officer	10	1	9	11.5 MS
Rate	100	10%	90%	11.5 MS

Table 4.3: Results of applying a detection programs on 7 points

Depth Detection by 7 Points (Our approach)				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	0	10	16.1 MS
Talal	10	2	8	16.1 MS
Nada	10	0	10	16.1 MS
Hassan	10	0	10	16.1 MS
Mahmood	10	1	9	16.1 MS
Yousef	10	0	10	16.1 MS
Aseel	10	1	9	16.1 MS
Said	10	0	10	16.1 MS
Mohamed	10	2	8	16.1 MS
Café officer	10	1	9	16.1 MS
Rate	100	7%	93%	16.1 MS

Table 4.4: Results of applying a detection programs on 9 points

Depth Detection by 9 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	1	9	19.7 MS
Talal	10	1	9	19.7 MS
Nada	10	0	10	19.7 MS
Hassan	10	2	8	19.7 MS
Mahmood	10	1	9	19.7 MS
Yousef	10	0	10	19.7 MS
Aseel	10	1	9	19.7 MS
Said	10	0	10	19.7 MS
Mohamed	10	0	10	19.7 MS
Café officer	10	1	9	19.7 MS
Rate	100	7%	93%	19.7 MS

Table 4.5: Results of applying a detection programs on 12 points

Depth Detection by 12 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	1	9	23.2 MS
Talal	10	1	9	23.2 MS
Nada	10	0	10	23.2 MS
Hassan	10	2	8	23.2 MS
Mahmood	10	1	9	23.2 MS
Yousef	10	0	10	23.2 MS
Aseel	10	1	9	23.2 MS
Said	10	0	10	23.2 MS
Mohamed	10	0	10	23.2 MS
Café officer	10	1	9	23.2 MS
Rate	100	7%	93%	23.2 MS

Table 4.6: Results of applying a detection programs on 15 points

Depth Detection by 15 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	0	10	26.2 MS
Talal	10	1	9	26.2 MS
Nada	10	1	9	26.2 MS
Hassan	10	0	10	26.2 MS
Mahmood	10	1	9	26.2 MS
Yousef	10	0	10	26.2 MS
Aseel	10	2	8	26.2 MS
Said	10	0	10	26.2 MS
Mohamed	10	1	9	26.2 MS
Café officer	10	1	9	26.2 MS
Rate	100	7%	93%	26.2 MS

Table 4.7: Results of applying a detection programs on 20 points

Depth Detection by 20 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	2	8	32.2 MS
Talal	10	0	10	32.2 MS
Nada	10	1	9	32.2 MS
Hassan	10	0	10	32.2 MS
Mahmood	10	1	9	32.2 MS
Yousef	10	0	10	32.2 MS
Aseel	10	0	10	32.2 MS
Said	10	1	9	32.2 MS
Mohamed	10	1	9	32.2 MS
Café officer	10	0	10	32.2 MS
Rate	100	6%	94%	32.2 MS

Table 4.8: Results of applying a detection programs on 25 points

Depth Detection by 25 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	1	9	37.8 MS
Talal	10	2	8	37.8 MS
Nada	10	1	9	37.8 MS
Hassan	10	0	10	37.8 MS
Mahmood	10	0	10	37.8 MS
Yousef	10	1	9	37.8 MS
Aseel	10	0	10	37.8 MS
Said	10	1	9	37.8 MS
Mohamed	10	0	10	37.8 MS
Café officer	10	0	10	37.8 MS
Rate	100	6%	94%	37.8 MS

Table 4.9: Results of applying a detection programs on 30 points

Depth Detection by 30 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	1	9	41.7 MS
Talal	10	1	9	41.7 MS
Nada	10	0	10	41.7 MS
Hassan	10	1	9	41.7 MS
Mahmood	10	0	10	41.7 MS
Yousef	10	1	9	41.7 MS
Aseel	10	0	10	41.7 MS
Said	10	1	9	41.7 MS
Mohamed	10	0	10	41.7 MS
Café officer	10	0	10	41.7 MS
Rate	100	5%	95%	41.7 MS

Table 4.10: Results of applying a detection programs on 35 points

Depth Detection by 35 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	0	10	45.5 MS
Talal	10	0	10	45.5 MS
Nada	10	0	10	45.5 MS
Hassan	10	1	9	45.5 MS
Mahmood	10	0	10	45.5 MS
Yousef	10	1	9	45.5 MS
Aseel	10	0	10	45.5 MS
Said	10	1	9	45.5 MS
Mohamed	10	0	10	45.5 MS
Café officer	10	2	8	45.5 MS
Rate	100	5%	95%	45.5 MS

Table 4.11: Results of applying a detection programs on 40 points

Depth Detection by 40 Points				
Name	**Try**	**False acceptance**	**True Rejection**	**Time**
Hashim	10	0	10	49.4 MS
Talal	10	1	9	49.4 MS
Nada	10	0	10	49.4 MS
Hassan	10	1	9	49.4 MS
Mahmood	10	0	10	49.4 MS
Yousef	10	1	9	49.4 MS
Aseel	10	0	10	49.4 MS
Said	10	1	9	49.4 MS
Mohamed	10	1	9	49.4 MS
Café officer	10	0	10	49.4 MS
Rate	100	5%	95%	49.4 MS

Table 4.12: Total results of applying a detection program

Total Table		
Points	**Rate**	**Time**
3	86%	6.9 MS
5	90%	11.5 MS
7	93%	16.1 MS
9	93%	19.7 MS
12	93%	23.2 MS
15	93%	26.2 MS
20	94%	32.2 MS
25	94%	37.8 MS
30	95%	41.7 MS
35	95%	45.5 MS
40	95%	49.4 MS

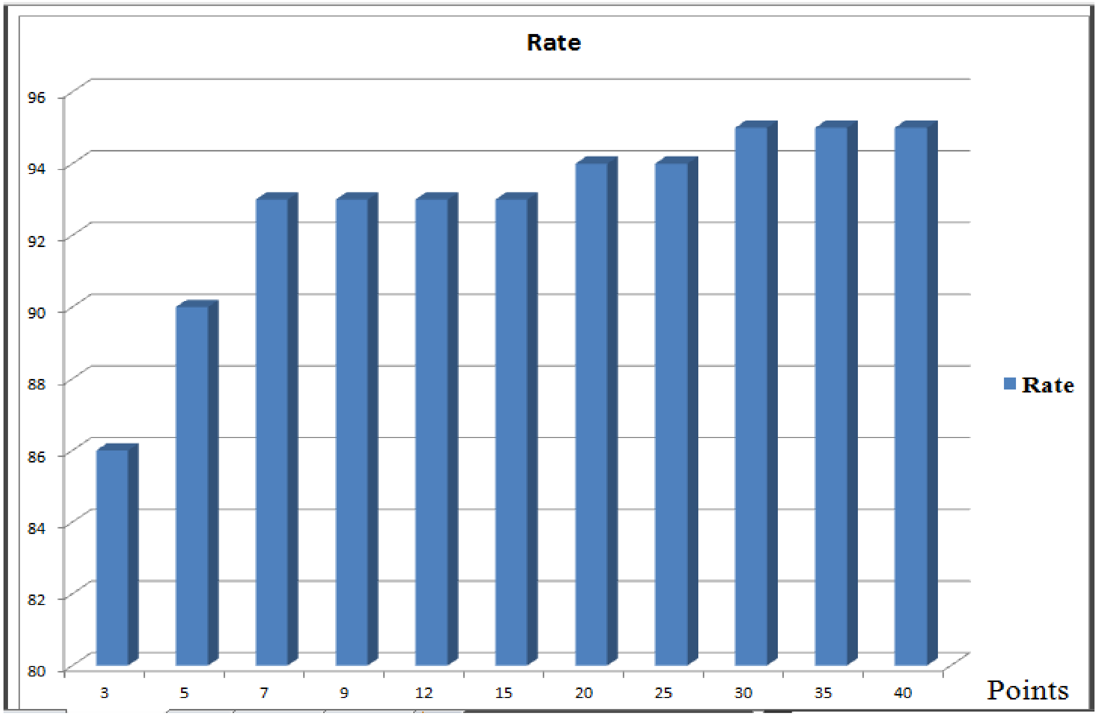

Figure 4.5: Impact of increasing points on the rate of depth detection

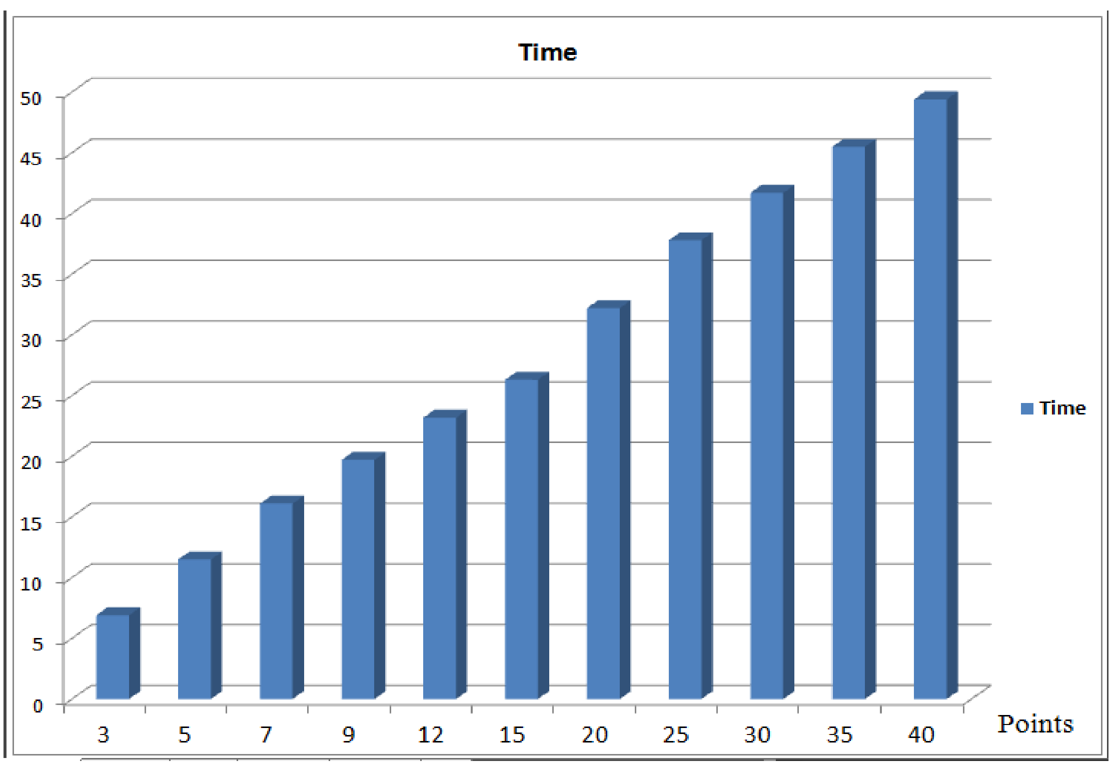

Figure 4.6: Impact of increasing points on the time of depth detection

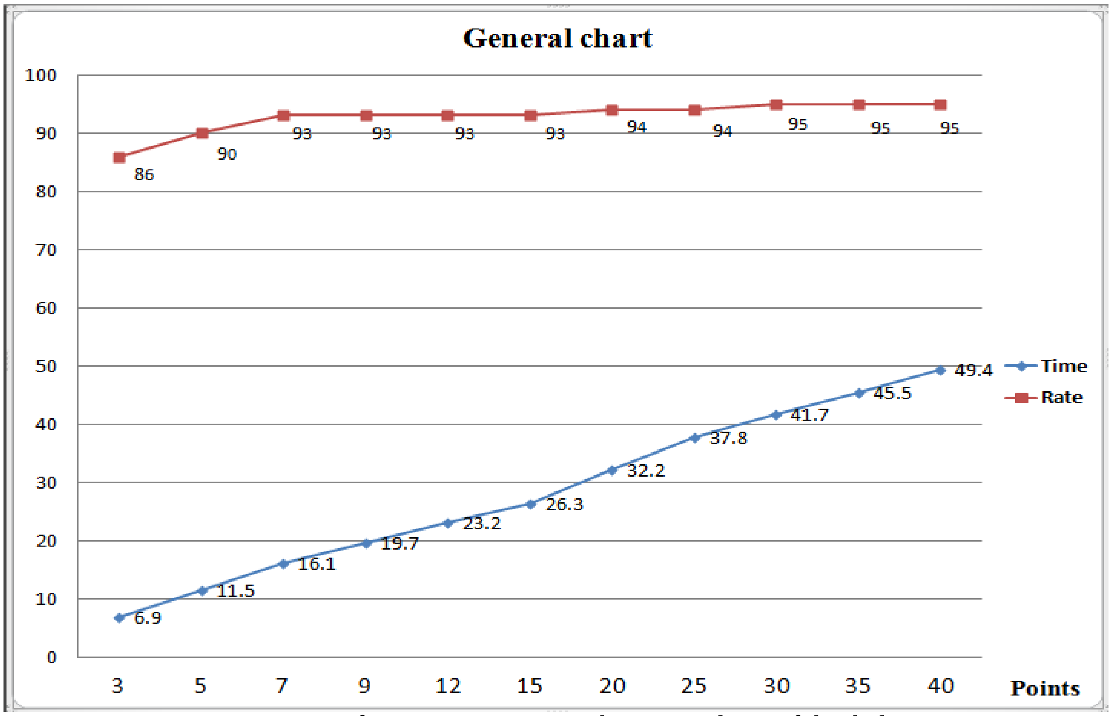

Figure 4.7: Impact of increasing points on the time and rate of depth detection

4.3.2 Testing on Standard method

These experiments had carried out on RGB camera to evaluate the results of our approach. These experiments are done in the same circumstances and algorithms in face detection and recognition by Haar, Eigenface, and using simulation database. Where contains the RGB images with the same testing dataset. Averages of detect deception at photographic form are 0%, in the RGB camera, making it easily susceptible to penetration. As shown in table 4.13.

Table 4.13: Results of deception RGB camera

RGB Results			
Name	**Try**	**False acceptance**	**True Rejection**
Hashim	10	10	0
Talal	10	10	0
Nada	10	10	0
Hassan	10	10	0
Mahmood	10	10	0
Yousef	10	10	0
Aseel	10	10	0
Said	10	10	0
Mohamed	10	10	0
Café officer	10	10	0
Rate	100	100%	0%

4.3.3 Comparison with Standard method

Following table presents the comparison between the results of proposed model and the results of RGB camera.

Table 4.14: Comparison between proposed approach and RGB camera

Name	Try	False acceptance	True Rejection
Results using depth detection approach	100	7%	93%
Results using RGB camera	100	100%	0%

4.3.4 Test using real persons for intelligent system access

This section is presented the experimental results of different aspects of the system, these results are tested and presented in terms of (face detection, face recognition, depth detection and animation).

Results in the table (4.15) are showing that proposed model is able to identify authentication user for using system. The system tested 10 persons from different ages and cultures, 10 systematic attempts for each entering, 10 attacking attempts against the system at different conditions. The security rate by preventing entering an unauthorized user was 98% and this rate assures the ability to use Kinect camera at different environments. The results are given in the following sections.

Table 4.15: Results using real persons

User Identification					
Owner			**Attacker**		
Name	**True Acceptance**	**False Rejection**	**Name**	**False Acceptance**	**True Rejection**
Mohamed	9	1	Mohamed	0	10
Ali	8	2	Ali	0	10
Aseel	9	1	Aseel	0	10
Mohamed co.	10	0	Mohamed co.	1	9
Omer co.	8	2	Omer co.	0	10
Akmar	8	2	Akmar	0	10
Shmose	10	0	Shmose	0	10
Yosef	10	0	Yosef	0	10
Said	9	1	Said	1	9
Hassan	10	0	Hassan	0	10
Rate	91%	9%	**Rate**	2%	98%

Previously results are showing 9% false rejection that caused by many factors and may affect the performance of Haar and Eigenface algorithms, such as (illumination changes, pose and expression changes), all these may pose as a hindrance in the practical application. This rate don't prevent account owner from entering to the system but may increase complexity of using this system, that caused when Kinect camera will send account's owner photo to the phone application and waiting to response.

4.4 Method Analytical Results

The method was implemented used C#, as three stages program. First stage to detect human face used Haar algorithm. Second stage is to recognize human face used Eigenfaces algorithm, third stage is to detect depth information to human face, and last stage detect human animation. The proposed model is implemented to handle ATM machine through the Kinect camera, and focused on the suitable security method for ATM machine, by dealing with the face detection and recognition. Good results are shown previously in this chapter. The results show ability of Kinect camera to handle problems of human identification in the different fields. The following chapter shows the conclusion of this thesis work, and demonstrates some views of future work remarks, that could be done in researches related to the field of this thesis. The main contribution of this work is the use of Kinect device for human identification in the real world. This system enhanced the reliability of previous systems, such as (using Kinect camera for human activity recognition, and for face detection and recognition) in a context other than gaming. The proposed system presented an approach to the human identification for increased reliability Kinect camera in the security field, through using many approaches in one system, while in the previously mentioned systems used one approach only.

4.5 Tools

➢ **Laptop**

Windows7, Ram 4 GB, HDD 500 GB

➢ **Visual Studio 2010**

✓ **Virtual Server**

The researcher creates a virtual server by using web service on the grounds that will be accommodated recognizing and perceiving

✓ **ATM system**

It's the process of creating a hypothetical ATM system on computer, so that it can be studied to see how the system works, in order to identify and understand the factors which control the system, and predict the future behavior of the system without having to experiment on a real system.

➢ **Android Studio**

✓ Mobile application

✓ The researcher used android studio to create mobile application.

➢ **Mobile device by using Android operating system version (4.2.1)**

✓ Android Jelly Bean is the term given to three main releases of the Android mobile operating system developed by Google, spanning versions between 4.1 and 4.3.1. Android 4.1 "Jelly Bean" added extra features to the user interface such as: (lock screen widgets, quick settings, and screensavers, one-finger gestures to expand/collapse notifications).

➢ **Kinect for windows (Version 1)**

✓ It provides color and depth images.

➢ **Data**

✓ Handling different cultures.

Chapter Five

Conclusion

And

Recommendations

5.1 Conclusion

Face detection and recognition in the field of biometrics issues, used widely in many applications such as: (surveillance systems, attendance and leave applications) but not used in determining a person's identity to provide security in hazardous environments, because there are many gaps, such as the possibility of deception using photo of the person registered, as well as there is a high probability suspicion among people.

The proposed model presented approach uses RGB images and depth information obtained through Kinect camera to improve the performance of detection and recognition techniques. Where a system was organized to provide security to the most dangerous environments (ATM environment), where people are approved by detection and facial recognition techniques. Also the influence of suspected cases of over-mover advantage provided by Kinect camera had got rid of, as well as a proposal depth approach uses pictures to get rid of fraud in photos and advantages of this approach: -

1. The absence of an excessive load on the database because it does not contain pictures of the depth, therefore the comparison takes place among points within the depths of memory.

2. In this model does not need training all algorithms to recognition of faces on the proposed database, for it contains RGB images only. They are axioms of facial recognition technology.

3. The proposed approach does not affect the processing time of any algorithm, because it is completely separated.

4. Using any RGB camera to collect image in the database.

We have implemented 1400 experiments, in thesis and based it on the obtained results. And the physical limitations of the Kinect device could be assessed. Considering depth information precision, the achieved results look promising.

5.2 Future Work

This study forms the base for several future researches. The following points are suggested, in order to get more benefits from using Kinect device and improve the recognition performance.

1. Convert IR projector in Kinect device to the Thermal infrared (IR) which is based on heat emission for addressing the cheating problem through image printed by 3D printer.
2. Design security system for the channels between server, ATM simulation and Android application.
3. Using cloud rather than server.
4. Design voice recognition system using Kinect capabilities.
5. Continue the research on perform face detection and recognition by using Kinect.

References

Abas, K. H., & Ono, O. (2010). Multilayer Infrared-Based Face Identification System for Security Vehicle-Robot Vision. International Journal of Simulation--Systems, Science & Technology, 11(5), 17-23.

Ahmad, F., Najam, A., & Ahmed, Z. (2013). Image-based face detection and recognition:" state of the art". arXiv preprint arXiv:1302.6379.

Arora, K. (2012). Real Time Application of Face Recognition Concept. International Journal of soft computing and Engineering (IJSCE), ISSN, 2(5), 191-196.

Bowyer, K. W., Chang, K., & Flynn, P. (2006). A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition. Computer vision and image understanding, 101(1), 1-15.

Cobos-May, C., Uc-Cetina, V., Brito-Loeza, C., & Martin-Gonzalez, A. (2015).A Convex Set Based Algorithm to Automatically Generate Haar-Like Features. Computer Science and Applications, 2(2), 64-70.

Colmenarez, A. J., & Huang, T. S. (1998). Face detection and recognition. In Face Recognition. Springer Berlin Heidelberg. 174-185

De Marsico, M., Nappi, M., Riccio, D., &Dugelay, J. (2012). Moving face spoofing detection via 3D projective invariants. In Biometrics (ICB), 2012 5th IAPR International Conference on, 73-78

Ghiass, R. S., Arandjelović, O., Bendada, A., &Maldague, X. (2014). Infrared face recognition: A comprehensive review of methodologies and databases. Pattern Recognition, 47(9), 2807-2824.

Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. (2013). On RGB-D face recognition using Kinect. In Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on, 42-46.

Han, J., Shao, L., Xu, D., &Shotton, J. (2013). Enhanced computer vision with microsoftkinect sensor: A review**.** IEEE Transactions on Cybernetics**,** 43(5), 1318-1334.

Hg, R. I., Jasek, P., Rofidal, C., Nasrollahi, K., Moeslund, T. B., &Tranchet, G. (2012).An RGB-D Database Using Microsoft's Kinect for Windows for Face Detection**.** In Signal Image Technology and Internet Based Systems (SITIS), Eighth International Conference on, 42-46

Hjelmås, E., & Low, B. K. (2001). Face detection: A survey. Computer vision and image understanding, 8 (3), 236-274.

Hossny, M., Filippidis, D., Abdelrahman, W., Zhou, H., Fielding, M., Mullins, J.,&Nahavandi, S. (2012). Low cost multimodal facial recognition via Kinectsensors**.** In Proceedings of the land warfare conference (LWC): potent land force for a joint maritime strategy. Commonwealth of Australia, 77-86.

Kahm, O., &Damer, N. (2012). 2D face liveness detection: An overview. In Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the IEEE, 171-182

Kakadiaris, I. A., Passalis, G., Toderici, G., Murtuza, M. N., &Theoharis, T. (2006).3D Face Recognition. In BMVC 869-878.

Kumar, M. S., &Babu, R. V. (2012). Human gait recognition using depth camera: a covariance based approach. In Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing,(20), 1-6.

Lee, H. W., Liu, C. H., Chu, K. T., Mai, Y. C., Hsieh, P. C., Hsu, K. C., & Tseng, H. C. (2015). Kinect Who's Coming-Applying Kinect to Human Body Height Measurement to Improve Character Recognition Performance. Smart Science, 3(2), 117-121.

Li, B. Y., Mian, A. S., Liu, W., & Krishna, A. (2013). Using Kinect for face recognition under varying poses, expressions, illumination and disguise. In Applications of Computer Vision (WACV), 2013 IEEE Workshop on, 186-192.

Liao, S., Jain, A. K., & Li, S. Z. (2013). Partial face recognition: Alignment-free approach. Pattern Analysis and Machine Intelligence, IEEE Transactions on**,** 35(5), 1-14.

Ma, X., Liu, C., & Zhao, L. (2015). Face Fusion Recognition Based on Bit-Plane Image. The Open Automation and Control Systems Journal, 7(1), 835-841.

Maraqa, M. R. (2014). A new approach for real time Face Recognition system using modified Eigenfaces with Edge Tracking**.** MasterThesis, Middle East University, Amman.

Meng, R., Shengbing, Z., Yi, L., &Meng, Z. (2014). CUDA-Based real-time face recognition system. In Digital Information and Communication Technology and it's Applications (DICTAP), 2014 Fourth International Conference on 237-241, IEEE.

Mishra, A. K., Skubic, M., & Abbott, C. (2015).Development and Preliminary Validation of an Interactive Remote Physical Therapy System. In 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, 1-4.

Mohammed, T. S., Al-Azzo, W. F., & Al Mashani, K. M. (2013,).Image processing development and implementation: a software simulation using Matlab®. In Proceedings of the 6th International Conference on Information Technology.

Ong, W. H., Koseki, T., & Palafox, L. (2013). An Unsupervised Approach for Human Activity Detection and Recognition. International Journal of Simulation--Systems, Science & Technology, 14(5), 42-49.

Oprisescu, S., Rasche, C., & Su, B. (2012). Automatic static hand gesture recognition using Tof cameras. In Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European 2748-2751. IEEE.

Patsadu, O., Nukoolkit, C., &Watanapa, B. (2012).Human gesture recognition using Kinect camera. In Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference on. 28-32.IEEE

Pedersoli, F., Adami, N., Benini, S., &Leonardi, R. (2012).XKin-: eXtendable hand pose and gesture recognition library for kinect. In Proceedings of the 20th ACM international conference on Multimedia.1465-1468

Preis, J., Kessel, M., Werner, M., &Linnhoff-Popien, C. (2012). Gait recognition with Kinect. In 1st International Workshop on Kinect in Pervasive Computing.

Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013).Robust part-based hand gesture recognition using kinect sensor. Multimedia, IEEE Transactionson**,** 15(5), 1110-1120.

Sadeghi, M., Maghooli, K., &Moein, M. S. (2014). Using Artificial Immunity Network for Face Verification. International Arab Journal of Information Technology,11(4), 100-106.

Slavković, Marijeta, and DubravkaJevtić (2012). Face recognition using Eigenface approach. Serbian Journal of electrical engineering 9(1), 121-130.

Solomon, C., & Wang, Z. (2015).Driver Attention and Behavior Detection with Kinect. Journal of Image and Graphics, 3(2), 84-89.

Tayal, Y., Lamba, R., &Padhee, S. (2012). Automatic face detection using color based segmentation. International Journal of Scientific and Research Publications, 2(6), 1-7.

Tin, H. H. K. (2012).Robust Algorithm for Face Detection in Color Images. International Journal of Modern Education and Computer Science (IJMECS), 4 (2), 31-37.

Viola, P., & Jones, M. (2001). Robust real-time object detection. International Journal of Computer Vision, 4, 51-52.

Wilson, P. I., & Fernandez, J. (2006). Facial feature detection using Haar classifiers. Journal of Computing Sciences in Colleges, 21(4), 127-133.

Xia, L., Chen, C. C., &Aggarwal, J. K. (2011).Human detection using depth information by Kinect. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on IEEE**.**15-22.

Zhang, C., &Tian, Y. (2012). RGB-D camera-based daily living activity recognition. Journal of Computer Vision and Image Processing, 2(4), 1-7.

Rezaei, M., Nafchi, H. Z., & Morales, S. (2014). Global Haar-like features: A new extension of classic Haar features for efficient face detection in noisy images. In Image and Video Technology. Springer Berlin Heidelberg, 302-313.

Appendix (A)

- **Software code**

Main Function where five points are selected to detect surface part is done:

```
private void SensorDepthFrameReady(object sender, DepthImageFrameReadyEventArgs
e)
    {
        using (DepthImageFrame depthFrame = e.OpenDepthImageFrame())
        {
            if (depthFrame != null)
            {
                // Copy the pixel data from the image to a temporary array
                depthFrame.CopyDepthImagePixelDataTo(this.depthPixels);
                // Get the min and max reliable depth for the current frame
                int minDepth = depthFrame.MinDepth;
                int maxDepth = depthFrame.MaxDepth;
                // Convert the depth to RGB
                int colorPixelIndex = 0;
                for (int i = 0; i < this.depthPixels.Length; ++i)
```

```
                {
                    // Get the depth for this pixel
                    short depth = depthPixels[i].Depth;
// We're preserving detail, although the intensity will "wrap."
// Values outside the reliable depth range are mapped to 0 (black).
// Note: Using conditionals in this loop could degrade performance.
// Consider using a lookup table instead when writing production code.
// See the KinectDepthViewer class used by the KinectExplorer sample
// for a lookup table example.
                    byte intensity = (byte)(depth >= minDepth && depth <= maxDepth ? depth :
0);
                    // Write out blue byte
                    this.colorPixels[colorPixelIndex++] = intensity;
                    // Write out green byte
                    this.colorPixels[colorPixelIndex++] = intensity;
                    // Write out red byte
                    this.colorPixels[colorPixelIndex++] = intensity;
```

```
                // We're outputting BGR, the last byte in the 32 bits is unused so skip it
                // If we were outputting BGRA, we would write alpha here.
                ++colorPixelIndex;
            }
            // Write the pixel data into our bitmap
            this.colorBitmapDepth.WritePixels(
                new Int32Rect(0, 0, this.colorBitmapDepth.PixelWidth,
this.colorBitmapDepth.PixelHeight),
                this.colorPixels,
                this.colorBitmapDepth.PixelWidth * sizeof(int),
                0);
            lock (this)
            {
                foreach (var item in faces)
                {
                    for (int i = 0; i < 5; i++)
                    {
```

```
                    WriteableBitmapExtensions.DrawRectangle(colorBitmapDepth, item.X
+ i, item.Y + i, item.Right + i, item.Bottom + i, Colors.Red);
                    }
                }
            }
        }
    }
}
/// <summary>
/// Event handler for Kinect sensor's ColorFrameReady event
/// </summary>
/// <param name="sender">object sending the event</param>
/// <param name="e">event arguments</param>
//  private void SensorColorFrameReady(object sender,
ColorImageFrameReadyEventArgs e)
List<System.Drawing.Rectangle> faces = new List<System.Drawing.Rectangle>();
DepthImageFrame DF;
```

```
        private string ProcessColorFrame(bool AddFace, bool RecognizeFace, bool GetFace,
ref Bitmap face)
        {
            face = null;
            //using (ColorImageFrame colorFrame = e.OpenColorImageFrame())
            System.Threading.Thread t = new System.Threading.Thread(new
System.Threading.ThreadStart(() =>
           {
               if (DF != null)
               {
                   DF.Dispose();
               }
               DF = sensor.DepthStream.OpenNextFrame(500);
           }));
            t.IsBackground = true;
            t.Start();
            using (ColorImageFrame colorFrame = sensor.ColorStream.OpenNextFrame(250))
            {
```

```
if (colorFrame != null)
{
    // Copy the pixel data from the image to a temporary array
    colorFrame.CopyPixelDataTo(this.colorPixelsFace);
    // Write the pixel data into our bitmap
    this.colorBitmapFace.WritePixels(
        new Int32Rect(0, 0, this.colorBitmapFace.PixelWidth,
this.colorBitmapFace.PixelHeight),
        this.colorPixelsFace,
        this.colorBitmapFace.PixelWidth * colorFrame.BytesPerPixel,
        0);
    lock (this)
    {
        if (AddFace)
        {
            faceRecognition.AddFace(colorBitmapFace, textBox.Text);
        }
```

```
else
{
    if (RecognizeFace)
    {
        System.Drawing.Rectangle frect = new System.Drawing.Rectangle();
        string name = faceRecognition.RecognizeFace(colorBitmapFace, ref
face, ref frect);
        // if (face == null)
        {
            face = ConvertWriteableBitmapToBitmapImage(colorBitmapFace);
        }
        if (!string.IsNullOrEmpty(name))
        {
            Title = "Name:" + name;
            if (DF != null)
            {
                WriteableBitmapExtensions.DrawRectangle(colorBitmapFace,
(int)frect.X, (int)frect.Y, (int)frect.Right, (int)frect.Bottom, Colors.Red);
```

```
DepthImagePixel[] px = new
DepthImagePixel[DF.PixelDataLength];
DepthImagePoint[] dp = new
DepthImagePoint[DF.PixelDataLength];
DF.CopyDepthImagePixelDataTo(px);
sensor.CoordinateMapper.MapColorFrameToDepthFrame(colorFrame.Format,
DF.Format, px, dp);
Dictionary<FacePoint, DepthImagePoint> Points = new
Dictionary<FacePoint, DepthImagePoint>();
bool ttop = false;
bool tbottom = false;
bool tleft = false;
bool tright = false;
bool tcenter = false;
bool tLeftcenter = false;
bool tRightcenter = false;
bool known = false;
foreach (var p in dp)
```

```
{
    if (p.Depth == 0)
        continue;
    //   WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 1, (int)p.Y - 1, (int)p.X + 1, (int)p.Y + 1, Colors.Blue);
    known = true;
    int shift = 30;
    if (!ttop && (Math.Abs(frect.X + shift - p.X) < 5 &&
Math.Abs(frect.Bottom - shift - p.Y) < 5))
    {
        ttop = true;
        Points.Add(FacePoint.BottomLeft, p);
        WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
    }
    if (!tbottom && (Math.Abs(frect.X + shift - p.X) < 5 &&
Math.Abs(frect.Y + shift - p.Y) < 5))
    {
```

```
            tbottom = true;
            Points.Add(FacePoint.TopLeft, p);
            WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
        }
        if (!tleft && (Math.Abs(frect.Right - shift - p.X) < 5 &&
Math.Abs(frect.Bottom - shift - p.Y) < 5))
        {
            tleft = true;
            Points.Add(FacePoint.BottomRight, p);
            WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
        }
        if (!tright && (Math.Abs(frect.Right - shift - p.X) < 5 &&
Math.Abs(frect.Y + shift - p.Y) < 5))
        {
            tright = true;
            Points.Add(FacePoint.TopRight, p);
```

```
                    WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
                }


                if (!tcenter && (Math.Abs((frect.Width / 2 + frect.X) - p.X) < 5
&& Math.Abs((frect.Height / 2 + frect.Y) - p.Y) < 5))
                {
                    tcenter = true;
                    Points.Add(FacePoint.Center, p);
                    WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
                }
                if (!tLeftcenter && (Math.Abs((frect.Width / 2 + frect.X -20) -
p.X) < 5 && Math.Abs((frect.Height / 2 + frect.Y) - p.Y) < 5))
                {
                    tLeftcenter = true;
                    Points.Add(FacePoint.CenterLeft, p);
```

```
                    WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
                }
                if (!tRightcenter && (Math.Abs((frect.Width / 2 + frect.X + 20)
- p.X) < 5 && Math.Abs((frect.Height / 2 + frect.Y) - p.Y) < 5))
                {
                    tRightcenter = true;
                    Points.Add(FacePoint.CenterRight, p);
                    WriteableBitmapExtensions.FillEllipse(colorBitmapFace,
(int)p.X - 5, (int)p.Y - 5, (int)p.X + 5, (int)p.Y + 5, Colors.Red);
                }
            }
            if (known)
            {
                Title = "Depth detected";
                bool equal = false;
                int fakeR = 10;
```

```
if (Points.ContainsKey(FacePoint.BottomLeft) &&
Points.ContainsKey(FacePoint.BottomRight))
{
  if
(Math.Abs(Math.Abs(Points[FacePoint.BottomLeft].Depth) -
Math.Abs(Points[FacePoint.BottomRight].Depth)) < fakeR)
  {
    equal = true;
    if (Points.ContainsKey(FacePoint.TopLeft) &&
Points.ContainsKey(FacePoint.TopRight))
    {
      if
(Math.Abs(Math.Abs(Points[FacePoint.TopLeft].Depth) -
Math.Abs(Points[FacePoint.TopRight].Depth)) < fakeR)
      {
        equal = true;
        if (Points.ContainsKey(FacePoint.TopLeft) &&
Points.ContainsKey(FacePoint.Center))
        {
```

```
if
(Math.Abs(Math.Abs(Points[FacePoint.TopLeft].Depth) -
Math.Abs(Points[FacePoint.Center].Depth)) < fakeR)
{
    equal = true;
    Title = "FAKE";
    name = "Fake";
}
else
{
    if (Points.ContainsKey(FacePoint.BottomLeft)
&& Points.ContainsKey(FacePoint.Center))
    {
        if
(Math.Abs(Math.Abs(Points[FacePoint.BottomLeft].Depth) -
Math.Abs(Points[FacePoint.Center].Depth)) < fakeR)
        {
            equal = true;
```

```
    Title = "FAKE";
    name = "Fake";
}
else
{
    if
(Points.ContainsKey(FacePoint.CenterLeft) && Points.ContainsKey(FacePoint.Center))
    {
        if
(Math.Abs(Math.Abs(Points[FacePoint.CenterLeft].Depth) -
Math.Abs(Points[FacePoint.Center].Depth)) < fakeR)
        {
            equal = true;
            Title = "FAKE";
            name = "Fake";
        }
        else
        {
```

```
if
(Points.ContainsKey(FacePoint.CenterRight) && Points.ContainsKey(FacePoint.Center))
{
    if
(Math.Abs(Math.Abs(Points[FacePoint.CenterRight].Depth) -
Math.Abs(Points[FacePoint.Center].Depth)) < fakeR)
    {
        equal = true;
        Title = "FAKE";
        name = "Fake";
    }
    else
    {
        Title = "Welcome " + name;
    }
}
}
}
```

```
                                }
                            }
                        }
                    }
                }
            }
        }
    }
}
else
{
    Title = "Unknown";
    name = "Not detected";
}
}
}
else
```

```
    {
        Title = "Not detected";
        name = "Not detected";
    }
    return name;
}
else
{
    if (GetFace)
    {
        faces = faceRecognition.GetFaceRect(colorBitmapFace);
        foreach (var item in faces)
        {
            for (int i = 0; i < 5; i++)
            {
                WriteableBitmapExtensions.DrawRectangle(colorBitmapFace,
item.X + i, item.Y + i, item.Right + i, item.Bottom + i, Colors.Red);
```

```
                                if (Head != null)
                                {
                                    WriteableBitmapExtensions.DrawEllipse(colorBitmapFace,
(int)Head.X - 20, (int)Head.Y - 20, (int)Head.X + 20, (int)Head.Y + 20, Colors.Red);
                                }
                            }
                        }
                    }
                }
            }
        }
    }
    }
    return "";
}
enum FacePoint
{
```

```
    TopLeft,

    TopRight,

    Center,

    BottomLeft,

    BottomRight,

    CenterLeft,

    CenterRight

}

public Bitmap ConvertWriteableBitmapToBitmapImage(WriteableBitmap wbm)

{

    Bitmap bmImage = null;

    using (MemoryStream stream = new MemoryStream())

    {

        PngBitmapEncoder encoder = new PngBitmapEncoder();

        encoder.Frames.Add(BitmapFrame.Create(wbm));

        encoder.Save(stream);

        bmImage = new Bitmap(Bitmap.FromStream(stream));
```

```
        }
        return bmImage;
    }
    private void Window_Closing(object sender,
System.ComponentModel.CancelEventArgs e)
    {
        if (null != this.sensor)
        {
            this.sensor.Stop();
        }
    }
    private void Grid_Loaded(object sender, RoutedEventArgs e)
    {
        SelectSensor();
    }
    private void button_Click(object sender, RoutedEventArgs e)
    {
```

```
            Bitmap f = null;

            ProcessColorFrame(false, false, true, ref f);

        }

        bool fwaitingforresponse = false;

        KinectServer.KinectServerSoapClient KServer = new
KinectServer.KinectServerSoapClient();

        private void btnEnter_Click(object sender, RoutedEventArgs e)

        {

            if (fwaitingforresponse)

                label.Content = "Waiting for response...please wait";

            else

            {

                if (lblPIN.Password == "0258")

                {

                    label.Content = "PIN accepted, requesting permission from Ali...";

                    Bitmap f = null;

                    var name = ProcessColorFrame(false, true, false, ref f);
```

```
        if (string.IsNullOrEmpty(name) || name != "Ali")
        {
            fwaitingforresponse = true;
            //call web service to authorize
            //  f = new Bitmap(f, new System.Drawing.Size(320, 240));
            String g = Guid.NewGuid().ToString();
            KServer.Authorize(g, "Ali", ImageToB64(f));
            System.Threading.Thread t = new System.Threading.Thread(new
System.Threading.ParameterizedThreadStart(GetResponse));
            t.IsBackground = true;
            t.Start(g);
        }
        else
        {
            label.Content = "Welcome " + name;
        }
    }
```

```
            else
            {
                label.Content = "Invalid PIN code";
            }
        }
    }
    void GetResponse(object g)
    {
        int i = 0;
        while (i++ <= 30)
        {
            var res = KServer.GetAuthorizationResponse(Convert.ToString(g));
            Application.Current.Dispatcher.Invoke(new Action(() => { label.Content =
"Waiting for response from Ali..."; }));
            if (res == 1)
            {
                Application.Current.Dispatcher.Invoke(new Action(() => { label.Content =
"Please do the secret move"; }));
```

```
            state = SecretState.None;

            fwaitingforresponse = false;

            return;

        }

        else

        {

            if (res == 0)

            {

                fwaitingforresponse = false;

                state = SecretState.NotAuthorized;

                Application.Current.Dispatcher.Invoke(new Action(() => { label.Content =
"Ali rejected your request."; }));

                return;

            }

        }

        System.Threading.Thread.Sleep(1000);

    }
```

```
            state = SecretState.NotAuthorized;

            fwaitingforresponse = false;

            Application.Current.Dispatcher.Invoke(new Action(() => { label.Content = "Ali
didn't respond, please try again later!"; }));

        }

        private void button10_Click(object sender, RoutedEventArgs e)

        {

            Bitmap f = null;

            ProcessColorFrame(true, false, false, ref f);

        }

        Bitmap B64ToImage(string B64S)

        {

            byte[] bytes = Convert.FromBase64String(B64S);

            Bitmap image;

            using (MemoryStream ms = new MemoryStream(bytes))

            {

                image = (Bitmap)System.Drawing.Image.FromStream(ms);
```

```
    }
    return image;
}
string ImageToB64(Bitmap img)
{
    string base64String = "";
    using (MemoryStream ms = new MemoryStream())
    {
        // Convert Image to byte[]
        img.Save(ms, System.Drawing.Imaging.ImageFormat.Jpeg);
        // Convert byte[] to Base64 String
        base64String = Convert.ToBase64String(ms.ToArray());
    }
    return base64String;
}
private void button1_Click(object sender, RoutedEventArgs e)
{
```

```
            lblPIN.Password = lblPIN.Password + ((Button)sender).Content.ToString();
        }
        private void btnDelete_Click(object sender, RoutedEventArgs e)
        {
            lblPIN.Password = "";
        }
    }
}
```