UNIVERSITI PUTRA MALAYSIA

REAL TIME TRACKING AND FACE RECOGNITION USING WEB CAMERA

PUNITHAVATHI THIRUNAVAKKARASU.

REAL TIME TRACKING AND FACE RECOGNITION USING WEB CAMERA

By

PUNITHAVATHI THIRUNAVAKKARASU

Thesis Submitted to School of Graduate Studies, Universiti Putra Malaysia, in Partial Fulfillment of the Requirements for the degree of Master of Science

August 2005

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in partial fulfilment of the requirement for the degree of Master of Science

REAL TIME TRACKING AND FACE RECOGNITION USING WEB CAMERA

By

PUNITHAVATHI THIRUNAVAKKARASU

August 2005

Chairman : Associate Professor Abdul Rahman bin Ramli, PhD

Faculty : Engineering

Much interest has been shown in the field of biometric surveillance over the past decade. Face Recognition is a biometric recognition system that has gained much attention due to its low intrusiveness and easy availability of input data. To humans, face recognition is a natural ability that is an easy task. However, computerized face recognition is often complex and inaccurate. Several good techniques such as template matching, graph matching and eigenfaces have been developed by researchers to accomplish this task to varying degrees of success.

In this dissertation, the eigenface approach is combined with neural networks to perform face recognition. Face images are first projected into a feature space where eigenvectors are extracted. The neural network performs identification and is used to train the computer to recognize faces.

A number of very good approaches to face recognition are already available. Most of them work well in constrained environments. Here the development of a real time face

recognition system that should work well in an unconstrained environment is studied. A tracking system is developed to work together with the face recognition algorithm. A method using pixel difference is used to detect movements in the camera's view. A pan-tilt system, using stepper motors is used to enable horizontal and vertical movements.

The face recognition algorithm is found to be working well with a recognition rate of around 95%. Eigenface method combined with neural networks displays good performance in terms of accuracy and the ability for learning and generalization. The tracking system works well for objects traveling speeds below 5m/s and at distances from between 0.5m to 2m from the camera. Several improvements are suggested to improve the tracking system performance. An overview of some leading tracking and face recognition systems and scope of future work in this area is discussed.

Abstrak tesis yang di kemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi sebahagian syarat keperluan ijazah Master Sains

PENJEJAKAN DALAM MASA NYATA DAN PENGECAMAN MUKA MENGGUNAKAN KAMERA WEB

Oleh

PUNITHAVATHI THIRUNAVAKKARASU

Ogos 2005

Pengerusi : Professor Madya Abdul Rahman bin Ramli, PhD

Fakulti : Kejuruteraan

Minat serta kajian adalah meningkat di dalam bidang pengawasan biometrik dalam dekad ini. Perbandingan diantara semua teknik biometrik menunjukkan bahawa pengecaman muka makin diberi perhatian. Ini adalah kerana ianya kurang memerlukan interaksi diantara sistem dan subjek yang dicam serta mudah mendapat kemasukan data. Pengecaman muka menggunakan komputer pada asasnya rumit dan kurang tepat. Namun beberapa teknik yang baik seperti kepadananan templat, kepadanan graf dan "eigenface" telah dihasilkan untuk melakukan tugas ini dan ianya menunjukkan peningkatan dalam kejayaan yang berbeza-beza.

Di dalam tesis ini, kaedah "eigenface" digabungkan dengan kaedah rangkaian neural untuk melakukan pengecaman muka. Imej muka pada mulanya ditampilkan ke ruang ciri di mana "eigenvectors" dihasilkan. Rangkaian neural kemudiannya melakukan pengenalpastian dan seterusnya melatih komputer untuk melaksanakan pengecaman muka.

Beberapa teknik pengecaman muka yang baik sudahpun tersedia ada. Kebanyakan teknik-teknik ini berkesan di dalam persekitaran yang terkawal. Di sini, kajian dilakukan untuk menghasilkan sistem pengecaman muka yang lebih berkesan dalam situasi masa nyata. Sebuah sistem penjejak dibina untuk digabung bersama sistem pengecaman muka. Kaedah perbezaan piksel digunakan untuk mengesan pergerakkan dalam imej kamera.

Algoritma pengecaman muka ini berkesan dengan mencapai kadar pengecaman dalam lingkungan 95%. Kaedah "eigenface" digabung dengan kaedah rangkaian neural menampilkan peningkatan dalam kebolehan dan ketepatan serta berkebolehan untuk menjalankan pembelajaran dan pengitlakan. Sistem penjejakan menunjukkan keputusan yang memuaskan apabila menjejak objek yang bergerak dengan kelajuan yang tidak melebihi 5m/s dan pada jarak diantara 0.5m hingga 2m daripada kamera. Beberapa cadangan untuk meningkatkan pencapaian sistem ini dibincangkan. Beberapa teknik menjejak dan pengecaman muka yang terkenal diperkenalkan sebagai skop kajian lanjutan di dalam bidang ini.

ACKNOWLEDGEMENTS

Firstly, the author would like to take this opportunity to express her sincere gratitude to her research supervisor, Professor Madya Dr Abdul Rahman Ramli for his valuable advice and guidance throughout the duration of this project.

The author would also like to thank the members of her supervisory committee Encik Syed Abdul Rahman Al-Haddad and Puan Roslizah Ali for their support and assistance during this period. The author is also indebted to the respective lecturers, staff and fellow students of the engineering faculty for all assistance and encouragement given.

Finally, the author would like to express her gratitude to her husband, Sukumar Nair, for his support, assistance and encouragement during this period.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two Dimensions
3D	Three Dimensions
AFIS	Accurate Fingerprint Identification System
ATM	Automated Teller Machine
CCTV	Close Circuit Television
DNA	Deoxyribonucleic Acid
EER	Equal Error Rate
FAR	False Acceptance Rate
FRR	False Rejection Rate
IC	Integrated Circuits
JPEG	Joint Photography Expert Group
LED	Light Emitting Diode
MLP	Multilayer Perceptron
NN	Neural Network
PC	Personal Computer
PCB	Printed Circuit Board
PTU	Pan-Tilt Unit
RGB	Red, Blue Green (used in three dimension colour representation)

CHAPTER 1

INTRODUCTION

Over the past century, computers have evolved from being high tech word processors and memory devices to an amazing device that is capable of performing a vast variety of tasks. In many fields, computers outperform humans when it comes to speed and accuracy. Computers have become a major tool in performing many intelligent tasks in the fields of engineering, medical diagnostics, security and many more. The use of computers has made many tedious and complicated tasks relatively high speed and easy, giving high accuracy and fidelity. With the availability of the Internet, computers are becoming increasingly popular as many tasks can now be performed with ease.

Among the many fields of computer usage, one that is gaining much attention over the last two decades is remote monitoring and surveillance [1]. Computers are now being used in places where it is difficult or unsafe to place a human being.

1.1 Remote Surveillance

Remote surveillance and security monitoring devices are becoming increasingly popular in the world today. Remote surveillance is widely used in the monitoring of restricted areas, hazardous environments, unfriendly territory, and many others. With the recent threats to security in countries throughout the world, it is also desirable to have remote monitoring devices in places such as airports, government buildings, prominent buildings

and so on. It has also become necessary to have remote surveillance combined with face identification or recognition to increase security in high-risk areas.

Face recognition is the identification of a person from an image of their face. It is a pattern recognition task performed specifically on faces. There are two different modes of operation for a face recognition system, which are authentication and identification [2]. In the authentication mode, the system accepts or rejects the claimed identity of the individual. In the identification mode, the system compares the face image to a database of known people and returns the most likely identity or identities. It is advantageous if the system has the capability of learning to recognize unknown faces.

1.2 Problem Statement

In view of the high rate of crime, terrorism and fraud in the world today, it is becoming increasingly important to have remote monitoring systems that work well with other security devices. As security threats and frauds become increasingly rampant, it is necessary to have systems that allows monitoring and recognition of unauthorized people being in an area, using an equipment or making a transaction such as an Automated Teller Machine (ATM) withdrawal.

Identity fraud starts when an individual uses multiple identification documents such as driver's licenses, passports, visas, national Identification Card (IC), etc., under assumed identities. This is possible because in most countries, documents such as birth certificates

are very easy to fake. Databases may contain facial photographs and thus in principle, have the information required to prevent duplication. However, in practice, it is impossible for a human to search over millions of photos to find those duplicates. Fortunately, computers are able to do this function. Using Face Recognition, millions of images can be checked for possible matches quickly, automatically and with phenomenal accuracy. The software returns to the investigator any matches exceeding a confidence threshold and rank these matches in terms of diminishing resemblance or likelihood of a match.

With the emergence of biometric recognition systems, it is possible to achieve the above by combining a password, for example, with a face recognition system or a fingerprint recognition system. The International Biometric Group [3] predicts that the biometrics industry will experience tremendous growth by the year 2005, especially in the area of face recognition.

Although face recognition systems are popular and highly accurate, many do not perform well in real time environment. Factors such as scale, rotation, pose and lighting prove to be limiting factors in a real time face recognition system performance [3]. Preprocessing is normally necessary in order to obtain results that are satisfactory. A tracking module can be incorporated into the face recognition system in order to obtain the best view of a person's face in a real time environment where the subject is in motion.

1.3 Scope

In this thesis, a surveillance system that incorporates face tracking and recognition using a web camera is explored. A real-time face recognition system should be able to work in real world situations such as environments with bad lighting, crowds, differing scales and rotations. A good face recognition system should be able to work reliably with the constraints mentioned above. However it is difficult to develop a face recognition that is capable of adapting to all the constraints faced in the real world environment. Therefore this thesis will be limited to the development of a real time face tracking and recognition system under limited constraints such as lighting and pose, for 20 people in a database.

1.4 Objective

The objectives of this thesis are as follows:

i. To design a tracking system, which incorporates both software and hardware. The hardware uses two stepper motors to control horizontal and vertical motion.

ii. To develop an algorithm that performs facial recognition. A neural network approach will be used in this algorithm and the software will be tested using a small database of 20 people.

iii. To implement the motion tracking and face recognition system so that it is capable of performing real time surveillance and recognition of persons in an area.

1.5 Thesis Organization

This thesis consists of five chapters. Chapter One introduces the reader to face recognition and remote surveillance. Chapter Two presents the literature review on biometrics, face tracking and recognition. It gives an overview on the work previously done by others in this field. The methodology of this research is presented in Chapter Three, which includes the design and development of the system and steps involved in setting up the hardware and developing the software. The results of the tests carried out using the algorithm is discussed in Chapter Four. The results are compared and analyzed with other similar systems. Finally, the conclusion and future work for this thesis is discussed in Chapter Five, where drawbacks and further improvements to the system is presented.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction to Object Tracking

Real-time Object Tracking refers to the use of special purpose computer vision hardware and software to follow up on the motion of objects detected in dynamic scenes at rates which are high enough to be used in surveillance systems or decision-making, in real-world situations.

Many applications have been developed for monitoring public areas such as offices, shopping malls or traffic highways. In order to control normal activities in these areas, tracking of pedestrians and vehicles play the key role in video surveillance systems.

2.1.1 Tracking Techniques

Tracking techniques [4] can be classified into the following categories:

a. Tracking based on a moving object region. This method identifies and tracks a blob token or a bounding box, which are calculated for connected components of moving objects in 2 dimensions (2D) space. The method relies on properties of these blobs such as size, color, shape, velocity, or centroid. A benefit of this method is that it is time efficient, and it works

well for small numbers of moving objects. Its shortcoming is that, problems of occlusion cannot be solved properly in 'dense' situations. Grouped regions will form a combined blob and cause tracking errors. Azarbayejani [5] presents a method for 2D and 3D blob tracking to track motion in cluttered scenes. Non-linear modeling and a combination of iterative and recursive estimation methods are used to perform tracking.

b. Tracking based on an active contour of a moving object. The contour of a moving object is represented by a *snake*, which is updated dynamically. It relies on the boundary curves of the moving object. For example, it is efficient to track pedestrians by selecting the contour of a human's head. This method can improve the time complexity of a system, but its drawback is that it cannot solve the problem of partial occlusion, and if two moving objects are partially overlapping or occluded during the initialization period, this will cause tracking errors. Coue et. al. [6] proposed a stochastic algorithm for tracking of objects. This method uses factored sampling, which was previously applied to interpretations of static images, in which the distribution of possible interpretations is represented by a randomly generated set of representatives. It combines factored sampling with learning of dynamical models to propagate an entire probability distribution for object position and shape over time. This improves the mentioned drawback of contour tracking in the case of partial occlusions, but increases the computational complexity.

c. Tracking based on a moving object model. Normally model based tracking refers to a 3 dimension (3D) model of a moving object. This method defines a parametric 3D geometry of a moving object. High accuracy is obtained only for a small number of moving objects. Fablet and Black [7] solved the partial occlusion problem by considering 3D models. The definition of parameterized vehicle models make it possible to exploit the a-priori knowledge about the shape of typical objects in traffic scenes.

d. Tracking based on selected features of moving objects. Feature based tracking is to select common features of moving objects and tracking these features continuously. Corners can be selected as features for vehicle tracking. Even if partial occlusion occurs, a fraction of these features is still visible, so it may overcome the partial occlusion problem. The difficult part is how to identify those features, which belong to the same object during a tracking procedure (feature clustering). Several papers have been published on this aspect. Horn [8] extracts corners as selected features using the Harris corner detector. These corners then initialize new tracks in each of the corner trackers. Each tracker tracks any current corner to the next image and passes its position to each of the classifiers at the next level. The classifiers use each corner position and several other attributes to determine if the tracker has tracked correctly.

2.1.2 Other Approaches to Object Tracking

Other approaches to object tracking includes works by Fishler and Bolles [9] who presented a tracking method based on wavelet analysis. A wavelet-based neural network (NN) is used for recognizing a vehicle in extracted moving regions. The wavelet transform is adopted to decompose an image and a particular frequency band is selected for input into the NN for vehicle recognition. Vehicles are tracked by using position coordinates and wavelet feature differences for identifying correspondences between vehicle regions.

Zisserman et. al. [10] employed a second order motion model for each object to estimate its location in subsequent frames, and a 'cardboard model' is used for a person's head and hands. Kalman models and Kalman filters are very important tools and often used for tracking moving objects. Kalman filters are typically used to make predictions for the following frame and to locate the position or to identify related parameters of the moving object. Hayden and Sparr [11] implemented an online method for initializing and maintaining sets of Kalman filters. At each frame, they have an available pool of Kalman models and a new available pool of connected components that they could explain.

Yang and Chern [12] used an extended Kalman filter for trajectory prediction. It provides an estimate of each object's position and velocity. But, as pointed out by Coue et. al. [6], Kalman filters are only of limited use, because they are based on unimodal