**Face Recognition with**

**Visible and Thermal IR Images**

A Thesis Proposal

Submitted

to the Temple University Graduate Board

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science in Engineering

By

Lei Guan

December, 2009

Thesis Advisor: Dr. Seong G. Kong

Committee Member: Dr. Joseph Picone

Committee Member: Dr. Li Bai

# ABSTACT

This research proposes to improve the performance of a face recognition system using visible and thermal infrared (IR) images. Four new ideas are proposed in this research: eyeglasses which block energy emitted from face will cause partially occluded face in thermal IR images. This will degrade the performance of a face recognition system. (1) To deal with this problem, eyeglasses are detected using support vector machine classifier and replaced by information predicted from corresponding visible image using a neural network. (2) A new software based registration method called Edge-based Mutual Information is proposed to register visible and thermal IR images. (3) After registration, a new method based on particle swarm optimization and wavelet transform is proposed to fuse images in data level. (4) A new face recognition algorithm which uses predictable information is also proposed. To demonstrate the viability and advantages of face recognition with visible and thermal IR images, a real time face recognition system which takes visible, thermal IR, fused images as input, and uses different classification methods for recognition will be built.

Contents

[ABSTACT I](#_Toc250320948)

[i. INTRODUCTION 1](#_Toc250320949)

[1. Fundamental Issues in Face Recognition 2](#_Toc250320950)

[2. Face Recognition in Visible Spectrum 5](#_Toc250320951)

[3. Face Recognition in Thermal IR Spectrum 8](#_Toc250320952)

[4. Face Recognition with FUSION of Visible and Thermal IR images 9](#_Toc250320953)

[ii. IMAGE REGISTRATION 12](#_Toc250320954)

[1. Advantages of Edge-based Mutual Information (EMI) Image Registration 12](#_Toc250320955)

[2. EMI Image Registration Algorithm 14](#_Toc250320956)

[iii. EYEGLASSES REPLACEMENT IN THERMAL IR IMAGES 25](#_Toc250320957)

[1. Eyeglasses Detection in Thermal IR Images 25](#_Toc250320958)

[2. Eyeglasses Replacement in Thermal IR Images 28](#_Toc250320959)

[iv. Face recogntion system 35](#_Toc250320960)

[1. Structure of Face Recognition System 35](#_Toc250320961)

[2. Image Registration and Fusion 36](#_Toc250320962)

[3. Face Detection and normalization 36](#_Toc250320963)

[4. Feature Extraction 37](#_Toc250320964)

[5. Face Recognition Algorithm 37](#_Toc250320965)

[v. FUTURE WORK 43](#_Toc250320966)

[1. Fusion of Visible and Thermal IR Images 43](#_Toc250320967)

[2. Face Recognition Algorithm WITH PREDICTABLE INFORMATION 44](#_Toc250320968)

# INTRODUCTION

As one of the most successful applications of image analysis and understanding, face recognition has been a rapid growing research area for many years. Face recognition addresses the problem of identifying or verifying a person by comparing his face with face images stored in the database. General procedures for face recognition including face detection, face normalization, feature extraction and recognition. Face detection can segment face from complicated background. Face normalization is used to normalize the face to ensure the input face and faces stored in the database are of the same size and position. Feature extraction is to represent the normalized face as low-dimensional vectors with discriminant power. Face recognition includes both identification and verification. Face identification is the process of providing a ranked listing of candidates whose faces best match with the input face. For face verification, a person presents an identity claim and his face to the system, and then the system either accepts or rejects his claim based on the result of comparing his faces with those stored in the databases.

Although significant progress has been made in the area of face recognition during past few years, face recognition is still a challenging task. Face recognition in visible spectrum is influenced by variations in illumination, pose, facial expression, viewpoint, disguise and etc. Thermal infrared (IR) band, which consists of Mid-wave infrared (MWIR, 3-5), Long-wave infrared (LWIR, 8-14), can capture the emitted energy from an object, thus is more robust to illumination variation. The use of thermal IR imagery has great advantages over visible images for face recognition in variant illumination conditions. However, thermal IR imagery is sensitive to changes in body and ambient temperature and the existence of the glass. On the contrary, visible imagery is more robust to these factors. Considering the complementary information provided by visible and thermal IR images, fusion of them provides a viable way for face recognition. In this chapter, current status of face recognition will be introduced. Section 1 will introduce general issues of face recognition. Section 2 and 3 will discuss the face recognition technique in visible and thermal IR spectrum respectively. Section 4 will talk about fusion of visible and thermal IR imagery for face recognition.

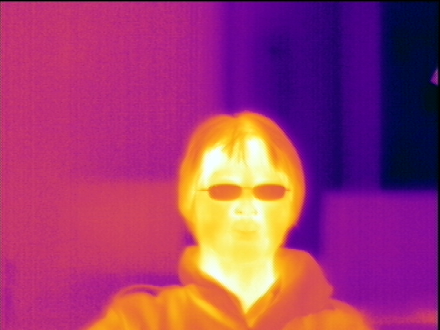
## Fundamental Issues in Face Recognition

Human beings have used biometric characteristics such as face, ears, voice, gait, iris, fingerprint and etc. for recognizing each other for a long time. Face recognition is one of the most successful applications in biometric areas. Unlike fingerprint analysis, iris scan, which depend on the cooperation of participants, face recognition is often effective without the participant’s cooperation or knowledge. Advantages and disadvantages of different biometrics are described in (Jain, 2004). Application areas of face recognition include entertainment (video game, virtual reality), human computer interface, smart cards (driver’s license, passports, voter registration), information security (personal device and desktop logon, internet access, database security), law enforcement and surveillance (CCTV control, shoplifting, suspect tracking and investigation) (Zhao, Chellappa, Phillips, & Rosenfeld, 2003).

Face recognition in uncontrolled environment is a challenging task. The performance of face recognition in visible spectrum is sensitive to variations of facial expression (Tian, Kanade, & Cohn, 2001), pose (Ben-Arie & Nandy, 1998), illumination condition (Adini, Moses, & Ullman, 1997), view point and disguise (Pavlidis & Symosek, 2000). The performance of visible face recognition system in outdoor environment is especially sensitive to the illumination variation because face is a 3D object. Light from different directions will cast different shadows on the face. This will cause variance on face images. The use of an artificial illumination source can reduce such variance. However, it may distract the person or expose the presence of a surveillance system.

Thermal IR imagery, which captures the energy emitted from an object, is robust to illumination variation and thus offers a promising alternative to visible imagery for face recognition (Friedrich & Yeshurun, 2009). Thermal IR band, which consists of Mid-wave infrared (MWIR, 3-5), Long-wave infrared (LWIR, 8-14), is associated with thermal radiation emitted by objects. Human face can emit thermal radiation in both bands, which can be captured by thermal IR camera and produce a 2D image. Thermal radiation from human face is an intrinsic property, which is influenced by vein pattern beneath the skin. Therefore unlike visible images which only provide the reflective information, thermal IR images can provide the anatomical information which is less sensitive to outside variation. However, thermal IR imagery has its limitations: (1) Thermal IR imagery is influenced by the presence of glass. Glass blocks the energy emitted from an object, which will result in the loss of information. This will bring difficulties for recognizing a person wearing eyeglasses or behind a glass (like sitting in a car). shows a pair of visible and thermal IR images of a person wearing eyeglasses. (2) Thermal IR imagery is influenced by the variation of body and ambient temperature. Figure 2 illustrates the influence of body and room temperature on thermal IR images. As we can see, after exposed in outside or after physical exercise, the variation of body temperature will change thermal characteristics of face. However, in visible images, no big difference can be observed.

As new algorithms and more systems are built, the evaluation of existing systems becomes very important. A summary of current popular evaluation protocols can be seen in (Zhao, et al., 2003). Although the evaluation of face recognition should be dependent on its application area, the robustness of a face recognition system can still be evaluated

(a) (b)

Figure 1: Influence of eye glasses on thermal IR images

(a) (b) (c)

Figure 2: Changes in body and ambient temperature: (a) Room temperature. (b) Outside temperature: lower than the room temperature. (c) After 10 minutes physical exercise

in terms of three tasks: identification, verification and watch list (Grother, Micheals, & Phillips, 2003). In identification, the face recognition system provides a ranked listing of candidates that best match the input face from the unknown person. The performance is often described by the Cumulative Match Characteristic (CMC), which measures the rate at which the input image will be classified at rank *n* or better. In verification, a person presents his face to a face recognition system and claims his identity. Then the system either accepts or rejects his claim by comparing the input face with those stored in the database. In the process, two types of errors can occur in the process: (1) false accept, in which the system falsely accepts the claim by an imposter; (2) false reject, in which the system falsely rejects the claim by the person who actually belongs to the system. The Receiver Operation Characteristic (ROC) is often used to quantify the performance of a system’s verification performance. The watch list task can be seen as a combination of both verification and identification. The face recognition system will first determine if an individual is on the watch list. If the individual is on the watch list, then the system identifies the individual. The measure for correctly detecting and identifying an individual on the watch list is called the detection rate and identification rate. For real applications of a face recognition system, the false recognition is very critical. Due to the imperfection of face recognition system, innocent people can be detained and embarrassed.

## Face Recognition in Visible Spectrum

As suggested by research in psychology, human beings recognize a human face from both holistic and local features. Therefore the face recognition algorithm is divided into the following categories here: 1) Holistic approach 2) Feature-based approach.

1. **Holistic Approach**

These approaches use the whole face region as the input to the face recognition system. Among them, Principal Component Analysis (PCA), which provides a compact representation of face, has been widely used in face representation and recognition. A vector can be expressed as linear combinations of orthogonal basis:. The orthogonal basis can be obtained by solving the following problem: , where is the covariance matrix for input vector . Advantages of PCA includes: (1) it offers a compact representation. (2) It is less sensitive to noise. The first successful application of PCA for face recognition can be seen from (Turk & Pentland, 1991). In their system, every face in the database is projected onto the eigenfaces to obtain a vector of weights. When a probe face comes, it is also represented as a vector of weights by projecting it onto eigenfaces. Then the recognition of face is done by searching for the image in the database whose weights are closest to those of probe face.

As suggested by (Belhumeur, Hespanha, & Kriegman, 1997), eigenfaces method only reconstructs face from a low dimensional basis which ignores class-specific information. Therefore this reconstruction may not be optimal from a discriminant point of view. They proposed to use Fisher’s Linear Discriminant (FLD) for face recognition. The idea is to maximize the determinant of between-class scatter matrix to the determinant of within-class scatter matrix. The within-class and between-class scatter matrix is computed as follows:





Where is the number of classes,  is the mean image of class , is the mean image of all sample images,  is an image in class and is the number of samples in class .Then a matrix  is chosen as the matrix whose orthonormal columns maximize the ratio of the determinant of the between-class scatter matrix of projected samples to the determinant of the within-class scatter matrix:



In face recognition problem, because the number of training samples is less than the number of pixels, it is therefore is always singular. In order to overcome the complication of a singular , the author (Belhumeur, et al., 1997) proposes a method called Fisherface which use PCA to reduce the dimension of the feature space and then uses the standard FLD defined by equation (2) to further reduce the dimension:







Where is the total scatter matrix of N training images (), and  is the mean image of all training samples:



Besides PCA and FLD, other popular matching methods also show good results. These methods include: Independent Component Analysis (ICA) uses higher-order statistics information for face representation and recognition based on a set of basis vectors that possess maximum statistical independence (Bartlett, Movellan, & Sejnowski, 2002). Support Vector Machines (SVMs) try to reduce the error of misclassification by finding the optimal separating hyper-plane that maximizes the margin of separation (Heisele, Ho, & Poggio, 2001). Local Binary Patterns (LBP) use LBP histograms in small regions on the face for face representation and recognition (Ahonen, Hadid, & Pietikäinen, 2004). Active Appearance Model tries to reconstruct a new face for face recognition by using an integrated statistical model that combines shape and texture of a face (Cootes, Edwards, & Taylor, 2001; Edwards, Cootes, & Taylor, 1998).

1. **Feature-based Approach**

Feature-based approaches try to extract local features such as eyes, nose, mouth and use them for classification. One of successful algorithms among them is Elastic Bunch Graph Matching (Wiskott, Fellous, Krüger, & von der Malsburg, 1997). This algorithm uses the idea that all human faces share a similar topological structure. Faces can be represented by labeled graphs, with nodes positioned at fiducial points (eyes, noses and etc.). Each node is described by several complex Gabor wavelet coefficients at different scales and rotations based on fixed wavelet bases. These wavelet coefficients are robust to illumination change, translation, distortion, rotation, and scaling. Recognition is based on these labeled graphs and the distance between two nodes.

## Face Recognition in Thermal IR Spectrum

Face recognition in visible spectrum has difficulties in performing consistently under uncontrolled environments. It is especially sensitive to illumination variations. Thermal IR images, which capture the energy emitted from an object, are almost robust to illumination variation. Objects emit IR energy according to their temperature. Since human temperature is quite uniform (varying from 35.5 to 37.5 °C), it provides consistent thermal IR signature. Thermal patterns of faces are derived primarily from superficial blood vessels under the skin, which are unique to each individual (Kong, Heo, Abidi, Paik, & Abidi, 2005). It is shown that even similar twins have different thermal patterns. Therefore face recognition in thermal IR spectrum tries to utilize these characteristics. Eigenfaces approaches have been successfully applied to face recognition in thermal IR spectrum, and perform much better than applied in visible spectrum under different illumination conditions (D. A. Socolinsky, Wolff, Neuheisel, & Eveland, 2001). Besides eigenfaces approach, Yoshitomi, et al., also proposed to indentify thermal face using Neural Networks. They combine the output from 3 neural networks using histograms, mosaic images and shape factors of an image (Yoshitomi, Miyaura, Tomita, & Kimura, 1997) .

It is also argued that face recognition with thermal IR images has time lapse problem. Under relatively controlled illumination condition, the performance of a face recognition system using visible images outperforms that using thermal IR images when the probe image is acquired at a substantial time lapse from gallery image (Chen, Flynn, & Bowyer, 2005) . In this paper, the author used PCA-based recognition algorithm on a database which was collected in 10 weeks, with 1 acquisition session per week. The results show that in the same-session recognition, neither modality is clearly significantly better than the other. However, if acquisitions of gallery and probe images are in different sessions, PCA-based recognition using visible-light images performed better than PCA-based recognition using thermal IR images.

To improve the performance of face recognition using thermal IR images over time, the author in (Buddharaju, Pavlidis, Tsiamyrtzis, & Bazakos, 2007) tries to use the innate characteristics under skin for face recognition. In that paper, a Bayesian framework is used to segment the face from the background. Then the superficial blood vessel network under skin is extracted by image morphology. The branching points of the skeletonized vascular network (referred as thermal minutia points) are used as features for classification. Experiments show obvious improvements on the time-gap database of University of Notre Dame.

## Face Recognition with FUSION of Visible and Thermal IR images

Information fusion has been widely used in biometrics because the result generated from one sensor may not be reliable. Concerning the same object, different sensors can provide different information. Information fusion tries to improve the performance of a system by integrating complementary information from different sources (Ross & Jain, 2003). As shown before, for face recognition, visible and thermal IR images have their own advantages and disadvantages: visible images are sensitive to illumination variations while thermal IR images are almost robust to illumination variations. Thermal IR images are easily influenced by variations of body and ambient temperature as well as the presence of glass (eyeglasses) while visible images are more robust in these aspects. Therefore fusion of visible and thermal IR images provides a way to improve the overall performance of a face recognition system.

Information fusion can be mainly divided into three categories: 1) data level fusion, 2) feature level fusion, and 3) decision level fusion (Llinas & Hall, 1998). Data fusion combines raw data from different sources to produce new data which is more informative than original input. Feature fusion tries to combine various features in feature space, like texture or shape information, eigenfaces and etc. Decision fusion combines the output from different classifiers. Traditional methods include majority voting, AND, OR fusion, weighted average and etc.

In literature, it has been shown that face recognition using fusion of visible and thermal IR images provides better results than face recognition using single modality. In (Diego A. Socolinsky, Selinger, & Neuheisel), a simple adaptive weighting scheme is used to fuse the score generated by two classifiers. In (Bebis, Gyaourova, Singh, & Pavlidis, 2006), Genetic Algorithm is used to fuse visible and thermal IR images for face recognition both in pixel level in wavelet domain and in eigenspace domain. In (S. G. Kong, et al., 2007), eyeglasses are first replaced by an average template in thermal IR images to deal with the occlusion problem and then a weighted average technique is used to fuse visible and thermal IR images in wavelet domain. One disadvantage with this method is that it did not provide an empirical way to choose the weight for wavelet coefficients of visible and thermal IR images. This problem will be addressed in the proposed fusion method.

The remainder of this paper is arranged as follows: Chapter II will propose a new method called edge-based mutual information (EMI) to register visible and thermal IR images, a preliminary step for fusing visible and thermal IR images. Chapter III will introduce a novel eyeglasses replacing method in thermal IR image to alleviate the influence of occlusion to the face recognition system. Chapter IV will give an overview of the face recognition system being built. Chapter V will give an outline of future research work.

# IMAGE REGISTRATION

A preliminary step before fusing visible and thermal IR images is to register them. In this chapter, a new method for registering visible and thermal IR images called Edge-based Mutual Information (EMI) is proposed. When complicated background is present, it is very difficult to find a transformation that can register background and foreground at the same time. This is because when objects are of different depths to a camera, they will appear differently on two images captured at different locations due to different parallaxes. The proposed method can successfully register faces in front of complicated background, where traditional method like Mutual Information (MI) is more likely to fail. Besides, it also can register images under lateral illumination conditions, partially occluded conditions. Image registration in these tough conditions has never been addressed before to our knowledge.

## Advantages of Edge-based Mutual Information (EMI) Image Registration

A preliminary step before fusing visible and thermal IR images is to register two images (S. Kong, et al., 2007; D. A. Socolinsky & Selinger, 2002). Image registration is the process of overlaying two or more images of the same scene taken at different time, from different viewpoints, and/or by different sensors (Zitov & Flusser, 2003). Although there exists a system that can produce co-registered visible and thermal IR images (D. A. Socolinsky & Selinger, 2002), this kind of system may not be readily available or too expensive. Therefore, a general approach to register visible and thermal IR images is needed.

To register visible and thermal IR face images, Wang has used the vanishing-point approach to determine the 3D pose of the head, then registered visible and thermal IR face images with the known relationship between visible and thermal IR cameras (Jian-Gang, Sung, & Venkateswarlu, 2004). Kong and Heo have designed a Gaussian criterion and used the edge map of visible and thermal IR images for registration (S. Kong, et al., 2007). However, these methods only apply on registering faces present in clean background. When complicated background is present, feature extraction may not be reliable thus the registration result will be influenced.

To register visible and thermal IR images, a novel method called Edge-based Mutual Information (EMI), is proposed here. Characteristics that distinguish EMI from other approaches are: (1) EMI emphasizes the role of face in the registration process, which can handle registering faces in complicated background. If complicated background is present, foreground and background will lie on different planes because of their different distances to cameras. Due to different parallaxes, the same scene will appear quite different even on images captured by cameras side by side. As shown in Figure 3, a face in front of a switch is captured by a pair of visible and thermal IR cameras. Cameras are located side by side as shown in . In thermal IR image, we can see switch is next to the face while in visible image it is of a distance from the face. Therefore it is very hard to find a global transformation that can register both foreground (face) and background (switch) at the same time. Previous approaches try to register objects with clean background, which only consider objects at the same distance. Therefore, those methods are more likely to fail to register faces in complicated background. Our approach, EMI can recognize face from complicated background and assign more weight to it in the registration process, thus ensure the accuracy of face registration. (2) EMI is also robust to illumination variation and partial occlusion to some extent.

(a) (b)

Figure 3: Different Parallaxes (a) Image captured by visible camera (b) Image captured by thermal IR camera

****

**Figure 4: A pair of visible and thermal IR cameras**

## EMI Image Registration Algorithm

1. **Overview of EMI Image Registration Algorithm**

EMI image registration algorithm can be seen from . First, face is detected in visible spectrum using face detector introduced by Viola and Jones (Viola & Jones, 2001). Then canny edge detector(Canny, 1987) is used to extract edges in the face region of visible image as well as in the whole thermal image. Pixels that are on the edge will be assigned with more weight in the registration process. This can emphasize the face part in the registration process and also include spatial information into MI, which makes the similarity measure more robust and address the problem of different parallaxes. After that, both visible and thermal IR images as well as their edge maps are rectified. Image rectification simplifies the optimization process, thus avoids the problem of local maximum in searching for the transformation parameters. After image rectification, the difference between two images is only on the image’s horizontal axis. Therefore the transformation parameter only needs to be searched in 1 dimension (the horizontal axis of the image). During shifting, EMI objective function is calculated, the shifted value with the maximum EMI value will be used as the transformation parameter that can register visible and thermal IR images.

****

Figure 5: Algorithm for Edge-Based Mutual Information Registration

1. **Face Detection**

Face Detection is used to segment the face from complicated background, which enables accurate registration of face from visible and thermal IR images. To detect face, we use the method proposed by Viola and Jones (Viola & Jones, 2001), which is well known for its high detection rate and fast speed. The method uses Haar-like features and Adaboost algorithm to select features and train the classifier. One example of face detection in lateral illumination condition can be seen from .

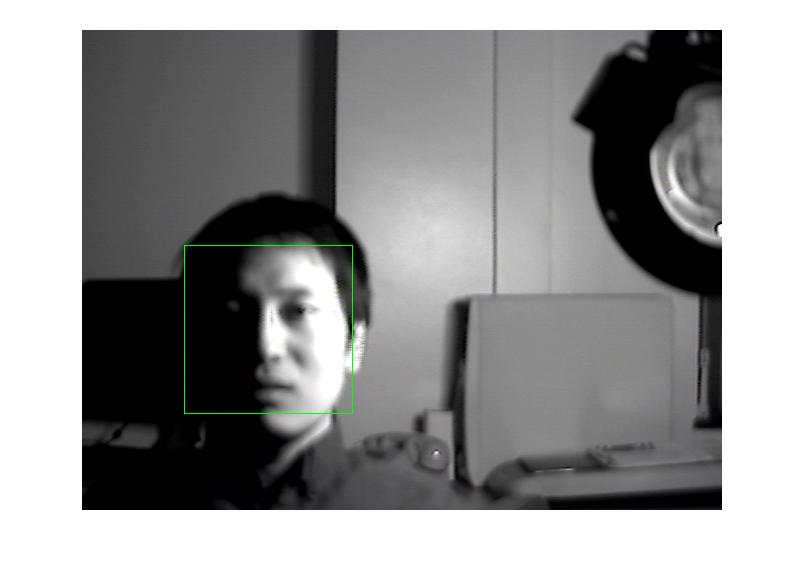


Figure 6: Detected face in lateral illumination

1. **Image Rectification**

Given a pair of images, image rectification is a transformation of image plane that makes the conjugate epipolar lines of two images become parallel to the horizontal image axis, enabling disparity matching to be a one dimensional searching (Hartley, 2000). Before two images can be rectified, it is necessary to get the intrinsic and extrinsic calibration parameters of visible and thermal IR cameras. The intrinsic parameter is a matrix that consists of a camera’s horizontal and vertical focal length and , coordinate of principle point , skew coefficient , as shown in Equation . The extrinsic parameter is another matrix , where  is a rotation matrix, and  is a  translation vector.  and  can be related using the coordinate of the optical center of the camera C, as shown in Equation .







To obtain these parameters, standard technique like camera calibration can be used. After getting camera 1’s extrinsic and intrinsic parameters  and camera 2’s extrinsic and intrinsic parameters, image rectification can be performed. Take images captured by camera 1 as an example. As shown in Figure 8, a spatial point is projected onto the camera 1’s image plane as point  using a projection matrix . Image rectification will find a transformation matrix  that transforms the original projection matrix  to a new projection matrix . projects the spatial point M to another point , which lies on the plane that is parallel to the base line . The original transformation matrix  can be obtained from intrinsic and extrinsic parameters of a camera as shown in Equation . After image rectification, the new X axis of the camera is parallel to the baseline . Therefore the new transformation matrix  can be expressed as Equation . ,, are vectors that represent the *X,Y,Z* axes of the camera after rectification., which represents that new X axis parallel to the baseline . ,means that the new *Y* axis is orthogonal to  and .is an arbitrary unit vector, that can be chosen as the *Z* unit vector of original coordinate for convenience. Finally,  is orthogonal to and, can be expressed as .





Because and are associated with the same line , we can express  as a set of spatial pointsthat satisfies equation and .





Hence,

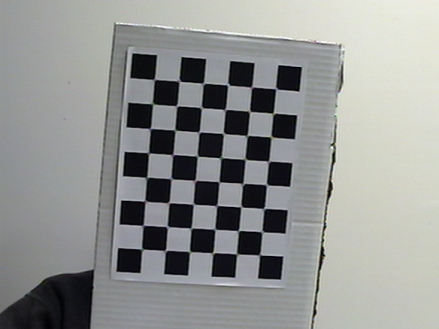
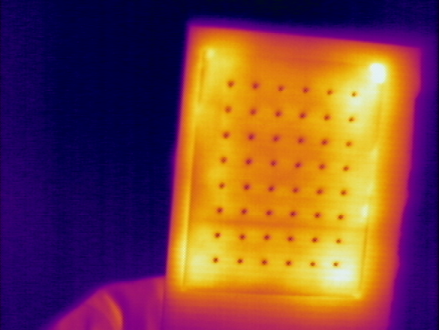


Then the rectifying transformation matrix for images taken from camera 1 is:



The same method can be applied for finding rectifying transformation matrix for images taken from camera 2.

Both Camera calibration and image rectification are implemented using Camera Calibration Toolbox ("http://www.vision.caltech.edu/bouguetj/calib\_doc/,"). For camera calibration, the toolbox uses control points on input images where a calibration board is visible. For getting visible images, it is simply placing a checkerboard pattern in front of a visible camera, and then taking some photos. For thermal IR images, 48 () nails are placed on the corner of the checkerboard’s grid. Then the board is heated with hair dryer so that 48 control points will become visible in thermal IR images. Examples of checkerboard and rectified images can be seen from and respectively.

(a) (b)

Figure 7: A pair of visible and thermal IR images of checkerboard for camera calibration (a) Visible image (b) Thermal IR image of heated checkerboard. 8x6 points are nails, used as control points

Figure 8: Image rectification

(a) (b)

Figure 9: A pair of rectified visible and thermal IR images (a) Visible image (b) Thermal IR image

1. **Edge-based Mutual Information (EMI)**

In information theory, mutual information (MI) measures the degree of dependence between two random variables (Vajda, 1989). Assume that two random variablesandwith events and, have marginal probability and joint probability.MI, measures the degree of dependence between  and by calculating the distance between the joint probabilityand, the case when andare statistically independent, using Kullback-Leibler measure (Vajda, 1989):



If we consider two images as random variable and, their intensities as events  and, then MI can be considered as a measure of dependence between two images. The marginal probability and joint probability of both images can be estimated using histogram and joint histogram of two images. When two images of the same object are spatially aligned, the dependence between them is the maximum, and their MI is also the maximum. Therefore, by maximizing MI, a transformation can be found to register two images.

One problem that influences the robustness and accuracy of MI is that it only assumes statistical information while ignoring spatial information (Keller & Averbuch, 2006; Kim, Lee, & Ra, 2008; Pluim, Maintz, & Viergever, 2000; Russakoff, Tomasi, Rohlfing, & Jr, 2004). EMI is proposed to alleviate this problem by including edge information. Another advantage of EMI is that it can register face in front of complicated background accurately by assigning more weight to face in the registration process.

The inspiration of EMI is that: different pixels, even of the same intensity, have different roles in registering two images. For example, pixels on the face may have the same intensity with those in the background. Because we are more interested in face, pixels on the face should play a more important role in the process of registration comparing with those in the background. However, when we use the histogram and joint histogram to estimate the marginal probability and joint probability of an image, they are treated equally. Our method tries to overcome this problem using edge information. Although the difference between visible and thermal IR images is large, edge information in both images has more information in common, because it captures the shapes of face and features like nose, mouth and etc. When two images are spatially aligned, the amount of edges that are overlapped is also at the maximum. Based on this, an edge coefficient is incorporated in MI.

To obtain the EMI objective function, edge is detected in thermal IR image and in the face part of visible image with Canny Detector (Canny, 1987). Letand denote the overlapping areas of visible and thermal IR images; denote the edge points and  denote non-edge points inand. Then the joint probability is calculated by constructing four joint histograms: 1) joint histogram of 2) joint histogram of , 3) joint histogram of , 4) joint histogram of,. By dividing four joint histograms with the total number of pixels in the overlapping area, the following four joint probabilities can be obtained:, , , . The marginal probability of imageand image:,,,can be also estimated from histograms of two images. After we get the joint and marginal probability, EMI can be obtained using the following equations:





 and  can be either 1 or 0. When, it means that the two corresponding points   
are both edge points. Then more emphasis will be placed on this combination by assigning a bigger coefficient s to , which is called the edge coefficient. In our experiment, we choose s as 100. When either  or , it means that either a point of oris not edge point, then they will be treated equally by assigning 1 to .

1. **Experiments and Results**

To evaluate the proposed algorithm, 15 pairs of visible and thermal IR images are collected and used for registration. To collect the images, a pair of visible and thermal IR cameras is placed on a platform which is supported by a tripod, as can be seen from . The visible camera used in the experiment is *Samsung SHC-735* camera and thermal IR camera is *ThermoVisionTM A40M* . The vertical distance of the platform to the ground is 115. The horizontal distance between centers of two cameras is 9. The distance of the object to cameras is about 150.

The collected 15 pairs of visible and thermal IR images include: 5 pairs of images with partially occluded face, 5 pairs of images taken under lateral illumination condition, 5 pairs of images taken with complicated background. Partially occluded face is obtained by placing a glass in front of face. Because glass can block the energy emitted from the face, the part of face behind the glass will be occluded in thermal IR images. MI and EMI are both used as objective function to register these images and their performances are evaluated using target registration error (TRE), which denotes the average distance between the positions of pixels determined by the estimated parameters and their ground truth (Maurer, Maciunas, & Fitzpatrick, 1998). Ground truth is obtained by manually registering visible and thermal IR images.

Results of image registration using partially occluded face, under lateral illumination condition and with complicated background can be seen from . Comparing with MI, using EMI as objective function has better performance. Examples of image registration under these situations can be seen from , , and .

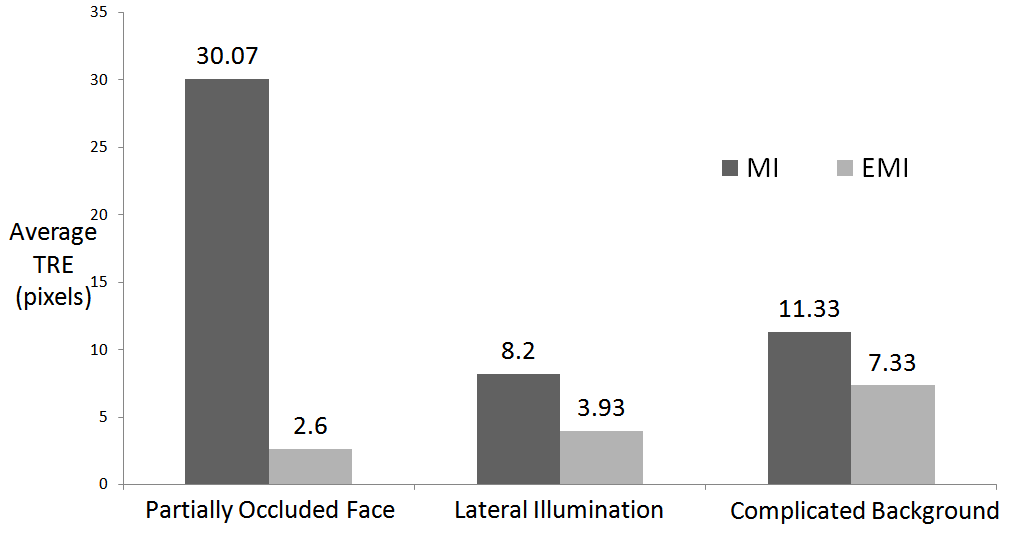
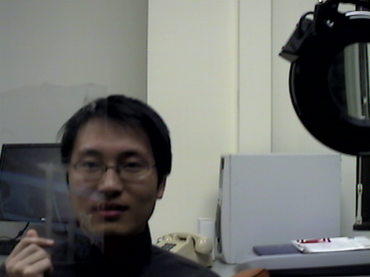


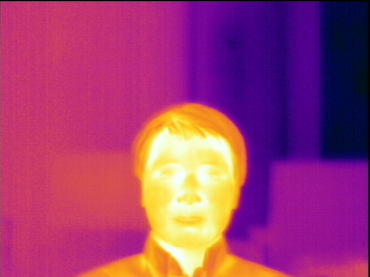
Figure 10: Registration of visible and thermal IR images using MI or EMI as objective function

   (a) (b) (c)



(d) (e)

Figure 11: Image registration with partially occluded face (a) Visible Image (b) Thermal IR image (c) Before registration (d) Registration using MI as objective function (e) Registration using EMI as objective function

   (a) (b) (c)

 ****

(d) (e)

Figure 12: Image registration under lateral illumination (a) Visible Image (b) Thermal IR image (c) Before registration (d) Registration using MI as objective function (e) Registration using EMI as objective function

   (a) (b) (c)

 ****

(d) (e)

Figure 13: Image registration with complicated background (a) Visible Image (b) Thermal IR image (c) Before registration (d) Registration using MI as objective function (e) Registration using EMI as objective function

# EYEGLASSES REPLACEMENT IN THERMAL IR IMAGES

One disadvantage of using thermal IR images for face recognition is the occlusion problem caused by eyeglasses. Eyeglasses block the energy emitted from face, which result in loss of information around eyes. The lost information will influence the performance of a face recognition system which uses thermal IR images (Bebis, et al., 2006; S. Kong, et al., 2007). In order to overcome this, it is desirable to compensate the lost information around eyes.

In this chapter, a novel approach to address this eyeglass problem is proposed. First, an input face image is classified as either with or without eyeglasses using a Support Vector Machine (SVM) classifier. If the input face is with eyeglasses, the eyeglasses can be segmented by thresholding concerning the big temperature difference between eyeglasses and face. A neural network model is used to predict the lost information in thermal IR images from the information around eyes in visible images. Eyeglasses in thermal IR images will be replaced by the predicted information.

## Eyeglasses Detection in Thermal IR Images

A support vector machine (SVM) classifier is trained to determine whether a person is wearing eyeglasses or not. If the person is wearing eyeglasses, thresholding is applied on the face to segment the eyeglasses from face. This algorithm assumes that input face is normalized, which is necessary in many face recognition algorithms (Zhao, et al., 2003). A face is usually normalized using the location of eyes, mouth, nose and etc. It may be argued that it is extremely hard to normalize a face in a thermal IR image because it is difficult to locate key facial features like eyes, nose, and mouth. Since visible and thermal IR images can be registered, we can use location of facial features in visible images to normalize the face on thermal IR images.

The diagram of the proposed algorithm can be seen from . Input faces are normalized to 32x32 pixels and then further converted into feature vectors. Then we use these feature vectors to train a SVM classifier that can classify a probe face image falling into two categories: 1) with eyeglasses, 2) without eyeglasses. SVM is a supervised learning method widely used in pattern recognition. An tutorial on SVM can be seen from (Burges, 1998). If the probe image is classified as with eyeglasses, thresholding is applied on the image to locate eyeglasses.

The proposed algorithm is tested on the NIST/Equinox Database ("http://www.equinoxsensors.com/products/HID.html,"). The NIST/Equinox Database consists of 1643 pairs of registered visible and thermal IR images. Visible and thermal IR images are registered within 1/3 pixels. In this experiment, we use 100 images, 50 images with eyeglasses and 50 images without eyeglasses, to train a SVM classifier with RBF Kernel. SVM is implemented using LIBSVM (Lin). Some sample images from this training set can be seen from . After training, we use the remaining 1543 thermal IR images to test the accuracy of thermal eyeglasses detection. The result can be seen from .



Figure 14: Diagram of eyeglasses detection algorithm

t5.bmp t1.bmp t2.bmp t3.bmp

(a)

tt21.bmp tt41.bmp tt1.bmp tt11.bmp

(b)

Figure 15: Training samples for eyeglasses detection: (a) Examples of training images with eyeglasses (b) Examples of training images without eyeglasses

**­**

Table 1: Performance of Eyeglasses Detection

|  |  |  |
| --- | --- | --- |
| Descriptions | Matched/Total Images | Accuracy |
| True Positive  (Eyeglasses->Eyeglasses) | **484/487** | **99.38%** |
| True Negative  (No Eyeglasses-> No Eyeglasses) | **1054/1056** | **99.81%** |
| False Positive  (No Eyeglasses->Eyeglasses) | 2/1056 | 0.19% |
| False Negative  Eyeglasses->No Eyeglasses | 3/487 | 0.62% |

## Eyeglasses Replacement in Thermal IR Images

After eyeglasses are detected in thermal IR images, eyeglasses region can be replaced by information predicted from visible images. The idea is like this: since visible and thermal IR images are sensing the same person, some relationship might exist between visible eye regions and thermal IR eye regions. If this relationship can be described by some model, then this model can be used to predict the lost information in thermal IR images from visible images. The relationship between visible and thermal IR images is highly complex and hard to be described by any closed-form mathematical model. A non-linear neural network model is proposed to describe this relationship. Neural networks can be trained to learn any nonlinear input-output relationship from a set of training data (Gurney, 1997). In the training process, the internal weights of a neural network are adjusted so as to minimize the estimation error over a set of examples. Here a multi-layer feed-forward neural network is trained to learn the relationship between eyes in thermal IR images and those in visible images.

The diagram of proposed algorithm can be seen from . Eye regions are first extracted from both visible training images with eyeglasses and thermal IR training images without eyeglasses based on the visible eye coordinates. Then Principal Component Analysis (PCA) is performed on both extracted eye regions for dimension reduction. The purpose of PCA operation is to represent each eye region as a low dimension vector and to obtain eigenvectors. The low dimension vector and eigenvectors can be used to reconstruct the original image. Details of using PCA for dimension reduction and reconstruction are described as follows:

**Dimension Reduction using PCA:**

* obtain images *I*1, *I*2, ..., *IM* (training faces)
* represent every image *Ii* as a vector *i*
* compute the average image vector 



* subtract the mean face from each vector to get the normalized face



* compute the covariance matrix C



* compute the eigenvectors *ui* of *AAT*
* keep only *K* eigenvectors (corresponding to the *K* largest eigenvalues)
* project each training face vector on the eigenspace :



* each normalized training face i is represented in this basis by a vector:



**Image Reconstruction:**

* Using eigenvectors *uj*, the weight *wj* and the average image to approximate the original image :





Figure 16: Diagram of eyeglasses replacement algorithm

After each image has been represented by a low dimension vector, these vectors will be used as input to train a neural network. A neural network which describes the relationship between visible and thermal IR eye regions can be obtained after training. When a pair of probe visible and thermal IR images comes, visible eye region is first extracted and represented by a vector using PCA. This vector is fed into the trained neural network. The trained neural network will predict a feature vector that describes corresponding thermal eye regions. Then thermal eye regions will be reconstructed using the predicted vector and calculated eigenvectors from thermal IR training images. The reconstructed eye region will replace the eyeglasses in thermal IR images.

The proposed algorithm is tested on NIST/Equinox Database, as can be seen from Table 2. “Gallery” denotes images in the database with known identity. One image for each person taken with frontal lighting condition and neutral facial expression is used for the gallery. “Probe” indicates images presented to the system for identification. Probe images are divided according to conditions such as the presence of eyeglasses, lighting conditions, and facial expression. “Training” refers to images used for training the SVM classifier for eyeglasses detection, the neural network for predicting thermal IR eye regions and eigenfaces for face recognition. Training images contain 30 people which neither exist in “Probe” nor “Gallery”. Although this makes the performance worse than otherwise, it is useful to eliminate any bias that might be introduced in the face recognition algorithm and make the experimental result more objective. Examples of training samples can be seen from Figure 17.In the experiment, an SVM classifier with RBF Kernel is trained to classify an input face as with or without eyeglasses. A feed-forward neural network which has two hidden layers, with 6 nodes in each layer, using log-sigmoid function as its transfer function, is trained to transform visible eyeglasses region to thermal eye regions. When a new face comes, it is first fed into the trained SVM classifier. If it is classified as face with eyeglasses, eyeglasses regions are converted to a low dimension vector (dimension 10) using PCA and further fed into the neural network to predict a new vector. Then thermal eye regions are reconstructed from the predicted new vector and are used to replace eyeglasses on the input face. Examples of faces after replacing eyeglasses can be seen from Figure 18.

In the literature, there are not many methods existing to deal with eyeglasses problem in thermal IR images directly. The only method we found is the one which replaces eyeglasses with an average thermal eye template (S. Kong, et al., 2007) . This method uses ellipse fitting technique to detect eyeglasses region in thermal IR images and then replace eyeglasses with the average eye template. We try to compare our algorithm with their algorithm based on the performance of face identification using eigenfaces-based approach (Turk & Pentland, 1991). The eigenfaces-based approach is chosen because of its great impact in face recognition community. It should be noted that the proposed algorithm can also be evaluated through other face recognition algorithms. Figure 19 shows the performance of face identification in terms of top 10 matching using probe sets with eyeglasses (probe 1, 2, 5, 6). As can be seen, after replacing eyeglasses, the performance of face identification is much better than before replacing eyeglasses. Also, our eyeglasses replacing algorithm performs better than the method of replacing with average template.

Table 2: NIIST/Equinox database of visible and thermal IR images

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Visible (Thermal) | Eyeglasses | Lighting | Expression |
| Gallery | 58 (58) | Off | Frontal | Neutral |
| Probe 1 | 27 (27) | On | Frontal | Neutral |
| Probe 2 | 85 (85) | On | Frontal | Various |
| Probe 3 | 36 (36) | Off | Frontal | Neutral |
| Probe 4 | 165 (165) | Off | Frontal | Various |
| Probe 5 | 57 (57) | On | Lateral | Neutral |
| Probe 6 | 139 (139) | On | Lateral | Various |
| Probe 7 | 121 (121) | Off | Lateral | Neutral |
| Probe 8 | 349 (349) | Off | Lateral | Various |
| Training | 582 (582) | On+Off | Frontal+Lateral | Neutral+Various |

ttt1.bmpttt6.bmpttt11.bmpttt16.bmp

(a)

vvv1.bmpvvv6.bmpvvv11.bmpvvv16.bmp

(b)

Figure 17: Examples of training images (a) Thermal IR images (b) Visible images

ttt43.bmpttt10.bmpttt22.bmpttt30.bmp

(a)

vvv43.bmpvvv10.bmpvvv22.bmpvvv30.bmp

(b)

vvv43.bmpvvv10.bmpvvv22.bmpvvv30.bmp

(c)

Figure 18: Examples of images before and after eyeglasses replacement: (a) Images before replacing eyeglasses (b) Images after replacing eyeglasses (c) thermal IR Images of the corresponding people without eyeglasses

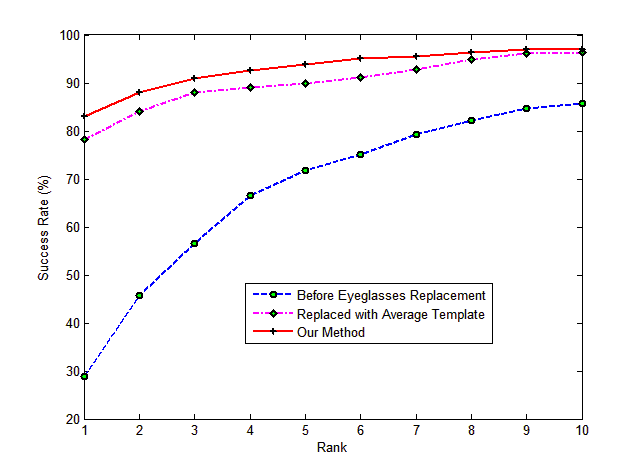


Figure 19: Performance of face identification before and after eyeglasses replacement

# Face recogntion system

In this chapter, a real-time face recognition system which uses visible, thermal IR images for face recognition is introduced. This system contains 5 face classification methods: (1) Eigenfaces; (2) Local Binary Pattern (LBP); (3) Support Vector Machines (SVM); (4) Minimum Average Correlation Energy (MACE) Filter; (5) our own classification algorithm (will be developed soon). Also, two databases are used in this system: a public database NIST/Equinox Database and our own database IPR Database.

## Structure of Face Recognition System



Figure 20: Structure of face recognition system

The whole system consists of a PC, a visible camera (*Samsung SHC-735*) and a thermal IR camera (*ThermoVisionTM A40M* ). The visible camera and thermal IR camera can communicate with the PC through a 4-channel frame grabber (*ADLINK RTV-24*). The frame grabber can be programmed to capture images from visible and thermal IR images at the same time, enabling further registration and fusion of visible and thermal IR images. The system can capture images at a maximum frame rate of 30 frame per second, with the resolution of 320×240.

The whole system is designed to select one from four sources as input: (1) image captured by visible camera; (2) image captured by thermal IR camera; (3) fused image from visible and thermal IR camera (4) image file stored in the hard disk. Then face will be detected and normalized on the selected image. After that, discriminant features are extracted from face and used for face recognition.

## Image Registration and Fusion

If fused image from visible and thermal IR cameras are selected as input, the system will capture one frame from two cameras at the same time. Then the proposed registration method Edge-based Mutual Information will be used to register visible and thermal IR images. After image registration, fusion is performed on captured images. Right now, only fusion based on weighted average of two images (Equation ) has been implemented.



## Face Detection and normalization

Face is detected from the selected image using the method proposed by Viola and Jones (Viola & Jones, 2001). This method is popular for its fast speed and high accuracy. The detected face will be further normalized to 100×100 pixels and used for feature extraction and face recognition.

## Feature Extraction

This will be introduced with face recognition algorithm together.

## Face Recognition Algorithm

Right now, four face recognition algorithms have been implemented: (1) Eigenfaces; (2) LBP; (3) Mace Filter; (4) SVM.

1. ***Eigenfaces***

The eigenfaces method is based on (Turk & Pentland, 1991). Detailed algorithm for face recognition is:

1. ***Training***

* obtain face images *I*1, *I*2, ..., *IM* (training faces)
* represent every image *Ii* as a vector *i*
* compute the average face vector  using



* subtract the mean face from each vector to get the normalized face



* compute the covariance matrix C



* compute the eigenvectors *ui* of *AAT*
* keep only *K* eigenvectors (corresponding to the *K* largest eigenvalues)
* project each training face vector on the eigenspace :



* each normalized training face i is represented in this basis by a vector:



1. ***Testing***

* normalize an unknown face image  :



* project on the eigenspace:



* represent as a vector



* compute the distance between unknown face and training faces using *Mahalanobis distance*:



* Find the face in the database with minimum distance as the matched person.

1. ***Local Binary Pattern (LBP)***

This algorithm is based on (Ahonen, et al., 2004). LBP is a non-parametric 3×3 kernel which summarizes the local spatial structure of an image. The operator labels pixels of an image by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number (). An extension to LBP is called uniform LBP (Ojala, Pietikainen, & Maenpaa, 2002). A Local Binary Pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. For example, 00000000, 00011110 and 10000011 are uniform patterns.

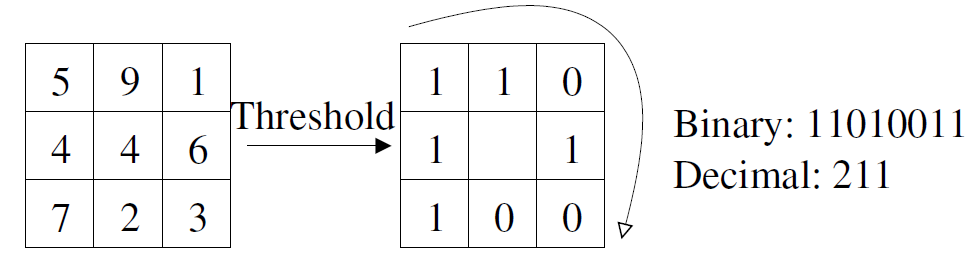


Figure 21: LBP Operator

The algorithm for face recognition can be described as:

* An individual face image is divided into 25 (5x5) small non-overlapping blocks (or regions) of same size.
* Uniform LBP operator is performed in each region.
* Calculate the histogram of the LBP codes in each region.
* Concatenate histograms of each block into a long single histogram. (Figure 22)
* Measure the distance between concatenated histograms of two images using Chi square dissimilarity measure: (*P* is the probe Image, G is Gallery Image)



* Find the face in the database with minimum distance as the matched person.

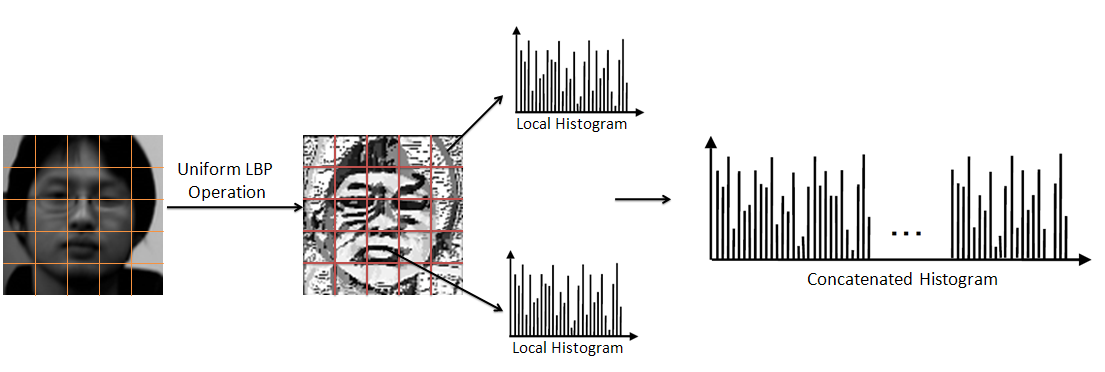


Figure 22: Uniform LBP for face recognition

1. ***MACE Filter***

The MACE filter method is based on (M. Savvides, 2002). The overall structure of this method can be seen from Figure 23. The key step for this method is how to design the filter. Details for designing the filter and used it for face recognition are:

1. ***Training: Get an individual filter for each class***

* *N* images in each class *I*1, *I*2, ..., *IN*, each image has *d* pixels.
* Perform 2D-FFT on each image *Ij*and convert it into a vector 1-D column (*d*×1) vector *Vj.*
* Arrange *N* vectors (corresponding to *N* images) into a matrix *X* (*d*×*N*)
* Specify a column vector *c* (*N*×1) with correlation peak values of the training images.
* Specify another matrix *D* (*d*×*d*) containing along its diagonal the average Fourier spectrum of training images.
* Calculate the filter (*d*×*1*) using equation



* Reorder into the 2-D array (same size as training image).

1. ***Testing:***

* Giving a test image *Itest*
* Correlation with MACE Filter using equation ,and get result *R*



* Calculate the *PSR* from *R* using equation . The way to measure peak and mean value can be seen from Figure 24



* Find the face in the database with highest PSR as the matched person.



Figure 23: Mace Filter for face recognition

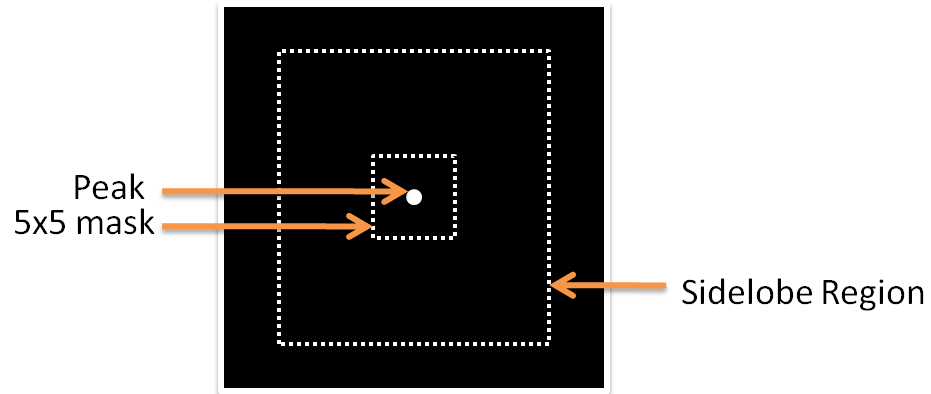


Figure 24: Calculate PSR from R: Peak (bright point) is located; then the standard deviation and mean are calculated over the sidelobe region (the rectangle area excluded the 5×5 mask region)

1. ***SVM***

The algorithm of SVM for face recognition is based on (Heisele, et al., 2001). An SVM classifier can separate two classes by maximizing their margin. In our application, the whole face is used for classification. To equip SVM with multi-class classification ability, one-vs.-all strategy is used: if there are *N* classes, then *N* SVM classifiers are trained. Each SVM classifier is trained to separate one class from the remaining classes. For testing, a new face is fed into all trained SVM classifiers. Then the SVM classifier with the highest output score assigns the label to the new face. Both training and testing can be implemented using LIBSVM (Lin).

# FUTURE WORK

## Fusion of Visible and Thermal IR Images

Fusion of visible and thermal IR images will be evaluated on decision level and data level. Decision fusion can be accomplished by combining the confidence rates. For a pair of visible and thermal IR probe images, two confidence rates will be generated by classifiers. Confidence rates tell how two images (gallery, probe) are similar. Several ways of combing these two confidence rates will be evaluated like the weighted average of two scores, choosing the maxim score.

On the data level, fusion will be performed in wavelet domain. The main reason to choose wavelet domain is that: (1) it can represent an image at multiple resolutions. Pixel by pixel fusion does not preserve the spatial information in the image. In contrast, fusion at multiple resolution levels allows features to be fused at the resolution at which they are most salient. In this way, important features appearing at different resolutions can be preserved in the fusion process. (2) It can represent an image at different frequencies. The advantage of using different frequencies is that high frequencies are relatively independent of global changes in illumination, while the low frequencies take into account the spatial relationships among pixels and are less sensitive to noises and small changes. Therefore a more rich description of images can be achieved and more salient features can be preserved in the fusion level. (3) The resolution of thermal IR images is lower than visible images. Fusion in wavelet domain can represent images in a multiple resolution scheme, thus can better deal with differences in resolution between visible and thermal IR images.

Fusion in wavelet domain is performed by combining coefficients of wavelet decompositions of a pair of visible and thermal IR images. Then the problem becomes how to determine the weight for each coefficient. An optimization method called particle swarm optimization (PSO) will be used to learn the weights of each coefficient by maximizing an objective function: the rank one face identification rate. Particle swarm optimization is a population based stochastic optimization technique, inspired by social behavior of bird flocking or fish schooling (Kennedy & Eberhart, 1995; Poli, Kennedy, & Blackwell, 2007). Advantages of PSO include they are a global optimization method and is much faster than other global optimization technique like Genetic Algorithm (GA) (Srinivas & Patnaik, 1994). After combining the weighted coefficients of visible and thermal IR images, a fused image will be reconstructed and fed into face recognition algorithm. The performance of fusion method will be evaluated by comparing the face identification rate using fused images with using visible images and thermal IR images. We expect that after fusion, the performance of face recognition can be greatly improved.

## Face Recognition Algorithm WITH PREDICTABLE INFORMATION

A new face recognition algorithm which uses predictable information will be investigated. It has already been shown that eye information in thermal IR image is predictable from visible image. The predicted information can be used to improve the performance of face recognition (Chapter III). Inspired by this, we believe that it is also possible to predict visible (thermal IR) face from corresponding thermal IR (visible) face and the predicted information can be used with the original information for recognition. To realize this idea, we need to find out how to predict the face and how to extract information from predicted face and original face. A new feature extraction is proposed to extract discriminant feature vector as shown in Figure 25 and Figure 26. Given thermal IR and visible training images, Principal Component Analysis (PCA) is performed on them for extracting low dimension feature vectors. Details of using PCA for dimension reduction are described in Chapter III. After each image has been represented by a low dimension vector, these vectors will be used as input to train a neural network. If we want to predict thermal IR (visible) images from visible (Thermal IR) images, visible (thermal IR) feature vectors will be used as training input and thermal IR (visible) feature vectors will be used as training target for the neural network training. After training, a neural network which predicts thermal IR (visible) information from visible (Thermal IR) information can be obtained. For a new input image, PCA is performed on it to extract a low dimension vector. The trained neural network takes this vector as input and predicts another vector as output. A new image can be reconstructed from the predicted vector (Details for reconstruction referring to Chapter III). Then the reconstructed image are concatenated with original image to form a larger image (if original image is of size *M*×*N* then the concatenated image is of size 2*M*×*N*). Finally, a discriminant low dimension feature vector can be obtained by using PCA on the larger image.

After the feature vector has been extracted, face recognition is simply calculating the distance (Euclidian distance for instance) between the extracted vector from probe image and that from gallery image. The gallery image which has the minimum distance to the probe image is the matched image.

The new face recognition algorithm will be compared with popular face recognition algorithms like Eigenfaces, LDA using NIST/Equinox Database. The performance will be measured by the face identification rate. We expect that its performance will outperform these popular algorithms.



Figure 25: Training for Feature Extraction



Figure 26: Feature Extraction

**References:**

Adini, Y., Moses, Y., & Ullman, S. (1997). Face recognition: the problem of compensating for changes in illumination direction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19*(7), 721-732.

Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face Recognition with Local Binary Patterns (pp. 469-481).

Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *Neural Networks, IEEE Transactions on, 13*(6), 1450-1464.

Bebis, G., Gyaourova, A., Singh, S., & Pavlidis, I. (2006). Face recognition by fusing thermal infrared and visible imagery. *Image and Vision Computing, 24*(7), 727-742.

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19*(7), 711-720.

Ben-Arie, J., & Nandy, D. (1998). A volumetric/iconic frequency domain representation for objects with application for pose invariant face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 20*(5), 449-457.

Buddharaju, P., Pavlidis, I. T., Tsiamyrtzis, P., & Bazakos, M. (2007). Physiology-Based Face Recognition in the Thermal Infrared Spectrum. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 29*(4), 613-626.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. [10.1023/A:1009715923555]. *Data Mining and Knowledge Discovery, 2*(2), 121-167.

Canny, J. F. (1987). A computational approach to edge detection *Readings in computer vision: issues, problems, principles, and paradigms* (pp. 184-203): Morgan Kaufmann Publishers Inc.

Chen, X., Flynn, P. J., & Bowyer, K. W. (2005). IR and visible light face recognition. [doi: DOI: 10.1016/j.cviu.2005.03.001]. *Computer Vision and Image Understanding, 99*(3), 332-358.

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23*(6), 681-685.

Edwards, G. J., Cootes, T. F., & Taylor, C. J. (1998). *Face Recognition Using Active Appearance Models*. Paper presented at the Proceedings of the 5th European Conference on Computer Vision-Volume II - Volume II.

Friedrich, G., & Yeshurun, Y. (2009). Seeing People in the Dark: Face Recognition in Infrared Images (pp. 348-359).

Grother, P., Micheals, R., & Phillips, P. (2003). Face Recognition Vendor Test 2002 Performance Metrics (pp. 1057-1057).

Gurney, K. N. (1997). *An introduction to neural networks*: CRC.

Hartley, R. a. A. Z. (2000). *Multiple View Geometry in Computer Vision*: Cambridge University Press.

Heisele, B., Ho, P., & Poggio, T. (2001). *Face recognition with support vector machines: global versus component-based approach.* Paper presented at the Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on.

<http://www.equinoxsensors.com/products/HID.html>.

<http://www.vision.caltech.edu/bouguetj/calib_doc/>.

Jain, A. K. a. R., A. and Prabhakar, S. (2004). An introduction to biometric recognition. *Circuits and Systems for Video Technology, IEEE Transactions on, 14*, 4-20.

Jian-Gang, W., Sung, E., & Venkateswarlu, R. (2004). *Registration of infrared and visible-spectrum imagery for face recognition.* Paper presented at the Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on.

Keller, Y., & Averbuch, A. (2006). Multisensor image registration via implicit similarity. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28*(5), 794-801.

Kennedy, J., & Eberhart, R. (1995). *Particle swarm optimization.* Paper presented at the Neural Networks, 1995. Proceedings., IEEE International Conference on.

Kim, Y. S., Lee, J. H., & Ra, J. B. (2008). Multi-sensor image registration based on intensity and edge orientation information. *Pattern Recognition, 41*(11), 3356-3365.

Kong, S., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B., Koschan, A., et al. (2007). Multiscale Fusion of Visible and Thermal IR Images for Illumination-Invariant Face Recognition. *International Journal of Computer Vision, 71*(2), 215-233.

Kong, S. G., Heo, J., Abidi, B. R., Paik, J., & Abidi, M. A. (2005). Recent advances in visual and infrared face recognition--a review. *Computer Vision and Image Understanding, 97*(1), 103-135.

Kong, S. G., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B. R., Koschan, A., et al. (2007). Multiscale Fusion of Visible and Thermal IR Images for Illumination-Invariant Face Recognition. *Int. J. Comput. Vision, 71*(2), 215-233.

Lin, C.-C. C. a. C.-J. LIBSVM: a library for support vector machines. 2001

Llinas, J., & Hall, D. L. (1998). *An introduction to multi-sensor data fusion.* Paper presented at the Circuits and Systems, 1998. ISCAS '98. Proceedings of the 1998 IEEE International Symposium on.

M. Savvides, B. V. K. V. K. a. P. K. (2002). *Face verification using correlation filters*. Paper presented at the Proc. Of the Third IEEE Automatic Identification Advanced Technologies.

Maurer, C. R., Jr., Maciunas, R. J., & Fitzpatrick, J. M. (1998). Registration of head CT images to physical space using a weighted combination of points and surfaces [image-guided surgery]. *Medical Imaging, IEEE Transactions on, 17*(5), 753-761.

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24*(7), 971-987.

Pavlidis, I., & Symosek, P. (2000). *The imaging issue in an automatic face/disguise detection system.* Paper presented at the Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000. Proceedings. IEEE Workshop on.

Pluim, J. P. W., Maintz, J. B. A., & Viergever, M. A. (2000). Image registration by maximization of combined mutual information and gradient information. *Medical Imaging, IEEE Transactions on, 19*(8), 809-814.

Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. [10.1007/s11721-007-0002-0]. *Swarm Intelligence, 1*(1), 33-57.

Ross, A., & Jain, A. (2003). Information fusion in biometrics. [doi: DOI: 10.1016/S0167-8655(03)00079-5]. *Pattern Recognition Letters, 24*(13), 2115-2125.

Russakoff, D., Tomasi, C., Rohlfing, T., & Jr, C. (2004). Image Similarity Using Mutual Information of Regions *Computer Vision - ECCV 2004* (pp. 596-607).

Socolinsky, D. A., & Selinger, A. (2002). *A comparative analysis of face recognition performance with visible and thermal infrared imagery.* Paper presented at the Pattern Recognition, 2002. Proceedings. 16th International Conference on.

Socolinsky, D. A., Selinger, A., & Neuheisel, J. D. Face recognition with visible and thermal infrared imagery. [doi: DOI: 10.1016/S1077-3142(03)00075-4]. *Computer Vision and Image Understanding, 91*(1-2), 72-114.

Socolinsky, D. A., Wolff, L. B., Neuheisel, J. D., & Eveland, C. K. (2001). *Illumination invariant face recognition using thermal infrared imagery.* Paper presented at the Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on.

Srinivas, M., & Patnaik, L. M. (1994). Genetic algorithms: a survey. *Computer, 27*(6), 17-26.

Tian, Y. I., Kanade, T., & Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23*(2), 97-115.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *J. Cognitive Neuroscience, 3*(1), 71-86.

Vajda, I. (1989). *Theory of Statistical Inference & Information*: Kluwer Academic Publishers.

Viola, P., & Jones, M. (2001). *Robust real-time face detection.* Paper presented at the Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on.

Wiskott, L., Fellous, J.-M., Krüger, N., & von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching (pp. 456-463).

Yoshitomi, Y., Miyaura, T., Tomita, S., & Kimura, S. (1997). *Face identification using thermal image processing.* Paper presented at the Robot and Human Communication, 1997. RO-MAN '97. Proceedings., 6th IEEE International Workshop on.

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Comput. Surv., 35*(4), 399-458.

Zitov, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing, 21*(11), 977-1000.