Multiscale Fusion of Visible and Thermal IR Images for Illumination-Invariant Face Recognition

SEONG G. KONG

Imaging, Robotics, and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering, The University of Tennessee, Knoxville, TN 37996-2100

JINGU HEO

Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213-3890

FAYSAL BOUGHORBEL, YUE ZHENG, BESMA R. ABIDI, ANDREAS KOSCHAN, MINGZHONG YI, AND MONGI A. ABIDI

Imaging, Robotics, and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering, The University of Tennessee, Knoxville, TN 37996-2100

Abstract. This paper describes a new software-based registration and fusion of visible and thermal infrared (IR) image data for face recognition in challenging operating environments that involve illumination variations. The combined use of visible and thermal IR imaging sensors offers a viable means for improving the performance of face recognition techniques based on a single imaging modality. Despite successes in indoor access control applications, imaging in the visible spectrum demonstrates difficulties in recognizing the faces in varying illumination conditions. Thermal IR sensors measure energy radiations from the object, which is less sensitive to illumination changes, and are even operable in darkness. However, thermal images do not provide high-resolution data. Data fusion of visible and thermal images can produce face images robust to illumination variations. However, thermal face images with eyeglasses may fail to provide useful information around the eyes since glass blocks a large portion of thermal energy. In this paper, eyeglass regions are detected using an ellipse fitting method, and replaced with eye template patterns to preserve the details useful for face recognition in the fused image. Software registration of images replaces a special-purpose imaging sensor assembly and produces co-registered image pairs at a reasonable cost for large-scale deployment. Face recognition techniques using visible, thermal IR, and data-fused visible-thermal images are compared using a commercial face recognition software (FaceIt[®]) and two visible-thermal face image databases (the NIST/Equinox and the UTK-IRIS databases). The proposed multiscale data-fusion technique improved the recognition accuracy under a wide range of illumination changes. Experimental results showed that the eyeglass replacement increased the number of correct first match subjects by 85% (NIST/Equinox) and 67% (UTK-IRIS).

Keywords: face recognition, visible-thermal image fusion, multisensor image registration, thermal infrared imaging, eyeglass replacement, personal identification, security

1. Introduction

Machine recognition of human faces from still and video images is an active research area due to increasing demands for security in commercial and law enforcement applications (Chellappa et al., 1995). Face recognition addresses the problem of verifying or identifying a given person in the scene by comparing an input face image with images stored in the database. Face recognition offers a non-intrusive and perhaps the most natural way of personal identification. Several biometric authentication methods are also available (Jain et al., 2004) based on other physiological characteristics such as fingerprint, retina or iris patterns, hand geometry, and voice. Multimodal biometric authentication methods are also extensively explored (Snelick et al., 2005). However, such identification systems mostly rely on the cooperation of the participants, which cannot be readily expected in security applications. Authentication with face recognition is intuitive and does not need to interrupt user activities. The analysis of frontal or profile face images is often effective for personal identity recognition tasks without the participants' cooperation or knowledge.

Despite a significant level of maturity and several practical successes, face recognition is still a highly challenging task in pattern recognition and computer vision (Craw et al., 1999; Burton et al., 1999). Face recognition based only on the visible spectrum has shown difficulties in performing consistently in uncontrolled operating conditions (Phillips et al., 2000; Bone and Blackburn, 2002). The accuracy of face recognition degrades quickly when the lighting is dim or when the face is not uniformly illuminated (Adini et al., 1997). Since the face is essentially a three-dimensional (3D) object, lighting sources from different directions may significantly change visual appearances. This is one of the primary reasons why current face recognition technology is often constrained to indoor access control applications where illumination is well controlled. Light reflected from human faces also varies depending on the skin color of people from different ethnic groups (Störring et al., 2001). This variability, coupled with dynamic lighting conditions, may cause great difficulties in recognizing the face in applications such as outdoor surveillance tasks. The use of an artificial illuminator may reduce light variability, but will distract the eyes of the people in the scene and reveal the presence of a surveillance system. Different pose angles are also responsible for the changes in visual appearances.

Shadows, glint, makeup, and disguise also cause great errors in extracting reliable features for face recognition.

Face recognition using different imaging modalities, particularly infrared (IR) imaging sensors, has become an area of growing interest (Yoshitomi et al., 1997; Prokoski, 2000). The use of thermal IR images can improve the performance of face recognition in uncontrolled illumination conditions (Wolff et al., 2001). Electromagnetic spectral bands below the visible spectrum such as X-rays and ultraviolet radiation are known to be harmful to the human body, and therefore cannot be employed in face recognition applications. Thermal IR spectrum comprising mid-wave IR (3–5 μm) and long-wave IR (8–12 μm) bands has been suggested as an alternative source of information for face detection and recognition. Thermal IR sensors measure the emitted heat energy, not reflected, from the object. IR energy can be viewed in any light conditions and is less subject to scattering and absorption by smoke or dust than visible light. Hence thermal imaging has great advantages in face recognition in low illumination conditions or even in darkness. Thermal IR spectrum also reveals anatomical information of a subject that is useful in detecting disguised faces (Pavlidis and Symosek, 2000). However, thermal imaging needs to solve several challenging problems. Thermal signatures are subject to change according to different body temperatures caused by physical exercise or ambient temperatures. Eyeglasses may result in loss of useful information around the eyes in thermal face images since glass is nearly opaque to thermal radiation.

This paper presents a data fusion technique of visible and thermal IR images for robust face recognition regardless of illumination conditions. Data fusion finds a sensor-level combination of multi-modal images to produce an integrated image enhanced in terms of information content for pattern recognition and classification. The combination of multi-modal images significantly improves the performance of face recognition in uncontrolled operating environments. Recent advances in face recognition technology using visible and infrared spectra are reviewed in Kong et al. (2005). Visible and thermal IR sensors capture complementary information of reflectance and radiation from the face. A pair of visible and thermal IR face images must be spatially co-registered before any useful information is extracted from the fused images. A software approach to image registration is desirable since camera systems that can produce co-registered visible and

thermal IR image pairs may not be readily available in practice or very expensive. In this paper, visible and thermal images are spatially aligned using a new Gaussian criterion to measure spatial proximity and visual similarity (Boughorbel et al., 2004a, b). A pair of co-registered visible and thermal IR images is fused in the wavelet transform domain to produce images insensitive to illumination variation. For subjects with eyeglasses, the eyeglass regions are approximated using ellipse fitting and replaced by thermal eye templates to retain details of the eyes useful for face recognition. Experimental results show that the visible-thermal image fusion technique significantly improves the recognition accuracy over individual sensing modalities. Performance of the fusion-based face recognition technique was evaluated using FaceIt$^{®}$, a commercial face recognition software developed by Identix, Inc. The evaluation used two independent databases of visible and thermal IR face images developed by the National Institute of Standards and Technology and Equinox Corporation (NIST/Equinox database) (http://www.equinoxsensors.com/ products/HID.html) and by the Imaging, Robotics, and Intelligent Systems laboratory at the University of Tennessee, Knoxville (UTK-IRIS database) (http://www.cse.ohio-state.edu/otcbvs-bench/).

2. Face Recognition Techniques

2.1. Visible Face Recognition

Face recognition algorithms can be classified into two broad approaches according to feature extraction schemes for face representation: feature-based and appearance-based methods (Brunelli and Poggio, 1993). Feature-based face recognition techniques compute a set of geometrical features on the face such as the eyes, the nose, and the mouth. A number of early face recognition algorithms are based on feature-based methods (Kanade, 1973; Cox et al., 1996; Manjunath et al., 1992). Properties and geometric relations such as the areas, distances, and angles between the feature points are often used as descriptors for face recognition. The performance of feature-based face recognition methods depends on the accuracy of the feature locating algorithms used.

Appearance-based methods find the global properties of the face pattern. The face is recognized as a whole, not by certain fiducial points obtained from different regions of the face. Many face recognition

algorithms in this category project an image into a subspace with linear transforms and find the similarity of reference images to an input image. Several leading commercial face recognition products use the face representation methods based on the Karhunen-Loeve expansion, such as the eigenface (Turk and Pentland, 1991) and the local feature analysis (LFA) (Penev, 1998, 1999). The LFA represents face images in terms of locally correlated features derived statistically from a representative ensemble of faces. Local representations offer robustness against variability due to the changes in localized regions of the objects. A selection (or sparsification) step produces a minimally correlated and topographically indexed subset of features that define the subspace of interest. The features used in the LFA methods are less sensitive to illumination changes, easier for estimating the rotations, and have less computational burden than the eigenface method. FaceIt$^{®}$, a commercial face recognition software package highly ranked in the face recognition vendor test (FRVT) (Blackburn et al., 2001; Phillips et al., 2003), uses the LFA as a face representation framework.

2.2. Thermal Face Recognition

Advantages of thermal IR imaging in face recognition include the invariance to illumination changes. Face recognition using the visible spectrum may not work properly in low lighting conditions. The use of the thermal IR spectrum for face recognition reported improvements in the recognition accuracies under a wide variety of illumination conditions (Wolff et al., 2001; Socolinsky et al., 2003). Thermal IR imagery is less sensitive to the variations in face appearance caused by illumination changes due to the fact that thermal IR sensor measures the heat energy radiation, not the reflectance, from the object. While the visible spectrum provides features that depend only on surface reflectance, thermal IR images produce features that uncover thermal characteristics of the face. Thermal face recognition utilizes such anatomical information of the human face as features unique to each individual that can be measured at a distance using passive IR sensors.

Similar recognition techniques proposed for visible face recognition have been applied in thermal face recognition as well. Appearance-based face recognition algorithms applied to thermal IR imaging consistently performed better than when applied to visible imagery. Initial research approaches to thermal face recognition extracts and matches thermal contours for

identification using techniques that include elemental shape matching and eigenface methods. Automated face recognition using elemental shapes in real time has reported 96% accuracy for cooperative access control applications (Prokoski, 2000). A non-cooperative, non-real-time faces-in-the-crowd version of thermal face recognition reported 98% accuracy with no false positives for more than 100 people represented in a database of 500 images.

2.3. Fusion of Visible and Thermal Face Recognition

Fusion techniques exploit synergistic integration of the information obtained from different data sources or from multiple pattern classifiers to improve the overall classification accuracy (Hall and Llinas, 2001). Data fusion combines the data from multiple sources to produce a more informative form than the original. The registration is intrinsic when the various wavelength bands come from the same sensor, but is more complicated when several different sensors are involved. Numerous attempts have been made in face recognition based on the fusion of different types of data. Face recognition algorithms based on the fusion of visible and thermal IR images demonstrated higher performance than individual image types (Chen et al., 2003; Heo et al., 2003; Singh et al., 2004). Multi-biometrics demonstrated improvement in face recognition by providing multiple evidences of the same identity (Brunelli and Falavigna, 1995; Ross and Jain, 2003). Biometric systems that integrate face and speech signals (Ben-Yacoub et al., 1999), the face and fingerprint information (Hong and Jain, 1998), and the face and the ear images (Chang et al., 2003) improved the accuracy in

personal identification. Fusion of local and global face features can also increase face recognition accuracy as well (Fang et al., 2002). Face recognition based on fusion-of-experts algorithms demonstrated better performance than when applied to either visible or thermal IR imagery alone (Wilder et al., 1996; Selinger and Socolinsky, 2001). Multiple neural network classifiers have been used to classify gender, ethnic origin, and pose of human faces (Gutta et al., 2000; Ho et al., 1994). Methods of decision fusion include majority voting, ranked list combination, AND fusion, OR fusion using the ranks, or the scores generated from the classifiers (Dasarathy, 1994). Most fusion schemes in biometrics use this type of classification fusion after reordering rank, score normalization or applying different weights from the results of each classifier. Score normalization and finding optimal weights are key issues in classifier fusion.

This paper presents a multiscale data fusion method of visible and thermal IR images of the face for illumination-invariant face recognition. Face recognition techniques based on visible spectrum (Vi), thermal IR (Th), and the fusion of visible-thermal data (Df) are compared. Figure 1 shows a schematic diagram of face recognition scenarios discussed in this paper. As a first step, visible and thermal IR image pairs are spatially co-registered for data fusion. The proposed registration method uses directional derivative maps instead of image intensities. The transformation modeled by a projective warp was recovered using a gradient-based optimization technique. The multiscale data fusion involves discrete wavelet decomposition and eyeglass detection and replacement with eye templates in thermal images. The discrete wavelet transform (DWT) finds a multi-scale representation of an image in terms of the
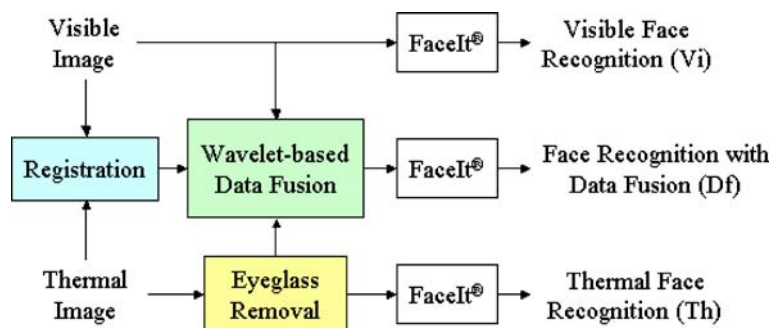

Figure 1. Block diagram for the face recognition schemes with visible images (Vi), thermal IR images (Th), and data fusion (Df) of visible and thermal images.

approximation and the details components. In the data fusion procedure, the approximation and the details co-efficients of visible and thermal images are combined independently. When a subject is wearing eyeglasses, thermal IR imaging will result in dark eyeglass regions in the face. Then a direct intensity combination of visible and thermal images may not produce useful information. An ellipse fitting method finds the location and shape of the eyeglasses and replaces the eyeglass regions with an average thermal template of the eyes to reduce the eyeglass effect. A face recognition software tool FaceIt$^{®}$ is used as a face recognition module to find ranked matches of the faces in the database.

3. Multiscale Fusion of Visible and Thermal Face Images

The NIST/Equinox database contains image pairs co-registered using hardware setting. The image pairs in the UTK-IRIS database are not originally registered and therefore spatial alignment is required before fusion. A new software-based image registration technique based on the Gaussian criterion is used to align each pair of visible and IR images in the UTK-IRIS database.

3.1. Visible and Thermal Face Image Registration

Image data fusion requires pixel-wise registration of the image pairs to be fused. Bringing multiple images of the same scene into spatial correspondence is an essential step in typical image fusion procedures. A special-purpose imaging sensor assembly that produces co-registered image pairs in the visible and thermal IR domains may not be practical in many face recognition scenarios due to elevated cost and low availability. Software registration of images enables the use of off-the-shelf visible and IR cameras for large-scale deployment at a reasonable cost. While the accuracy of the pixel-to-pixel alignment can be reduced compared to the hardware registration, the salient image features can be matched sufficiently well in a software registration process. In general, registration methods vary depending on the similarity measures between the images. Popular similarity measures used in image registration include cross-correlation, correlation ratio, and mutual information (Zitova and Flusser, 2003). Most area-based registration criteria generally assume global statistical dependence of the images to be aligned. This condition

may not be satisfied in the case of multi-sensor imagery such as visible and thermal IR images. For images containing prominent features, or sharp changes in intensity in certain regions, feature maps such as edge or frequency maps enhance registration accuracies. To align visible and IR images, Irani and Anandan (1998) designed a criterion that employed directional energy maps computed from the visible and IR images, which were used to reduce the visual differences between the two modalities and to highlight the common features. The method assumes local statistical dependence between the two modalities and the criterion was obtained by summing local cross-correlation measures of small image patches. This method was applied to the alignment of man-made structures such as airport runways and buildings.

In previous work, we introduced a Gaussian fields technique for the registration of 3D point-sets in the context of scene modeling (Boughorbel et al., 2004a, b). This method relies on a differentiable criterion and employs a standard gradient-based optimization scheme. The inputs to the Gaussian fields algorithm are the coordinates of the points that can be augmented by adding attributes such as color and local shape descriptors. In this paper, the original framework is applied to the two-dimensional (2D) registration of visible and thermal IR images. The criterion can be customized and applied to the task of binarized edge-maps matching, which was addressed previously by Huttenlocher et al. (1993) through the use of the Hausdorff distance. The registration criterion maximizes the overlap between the salient structures that are present in both visible and thermal images. As a point-sets registration technique, our method is also closely related to the popular class of methods known as the iterative closest point (ICP) techniques (Besl and McKay, 1992; Dalley and Flynn, 2002), although ICP was mostly employed for rigid registration. In the Gaussian fields framework, a smooth point-set and a shape-matching criterion substitutes for conventional non-differentiable registration algorithms such as ICP (Fitzgibbon, 2003) and Hausdorff distances (Charpiat et al., 2003). Another advantage of the Gaussian fields approach is its computational efficiency that leads to fast linear implementation (Elgammal et al., 2003).

A point X of either datasets to be registered can be augmented with an associated 1D attribute vector $S(X)$ whose components can be local shape descriptors as well as intensity or color information. Let $S(X)$ and $S(Y)$ denote the attribute vectors of X and Y in the

binarized feature-map computed from thermal IR and visible images. The binary feature-maps can be described as point-sets $M = \{X, S(X)\}$ of N_M points and $D = \{Y, S(Y)\}$ containing N_D points respectively of the visible and IR data. The basic registration function used in this paper measures the spatial proximity and local feature similarity of the two points X and Y in terms of a Gaussian function F defined by:

$$F(X, Y) = \exp\left(-\frac{d^2(X, Y)}{\sigma^2} - [S(X) - S(Y)]^T\right.$$
$$\left. \times \Sigma^{-1}[S(X) - S(Y)]\right) \qquad (1)$$

where $d(X, Y)$ denotes the Euclidean distance between points X and Y. The Gaussian function can be interpreted as a force field decaying with Euclidean distance between the image points and the Mahalanobis distance between the attribute vectors. The parameter σ controls the decay with Euclidean distance, while a diagonal matrix Σ with small components penalizes the difference in the attributes. The matrix Σ is used to de-correlate the feature descriptors and is obtained by computing the statistical dependence of the different descriptors in an approach similar to that proposed by Sharp et al. (2002). The attribute vectors correspond to the edge maps extracted from visible and thermal IR face images. Commonly used local feature descriptors such as curvature are not useful in visible-thermal image registration due to wide intensity variations between the image pair and large noise in thermal data. The criterion that measures the registration of the two point sets is defined as the integration over all the Gaussian forces exerted by one point set over the other:

$$E(T) = \sum_{\substack{X \in M \\ Y \in D}} \exp\left(-\frac{d^2(X, T[Y])}{\sigma^2} - [S(X)\right.$$
$$\left. - S(T[Y])]^T \Sigma^{-1}[S(X) - S(T[Y])]\right) \quad (2)$$

where $T[\cdot]$ denotes an affine transformation for the registration of the two point-sets.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = T\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

Moments that are invariant to affine transformations (Zitova and Flusser, 2003) are used as visual attributes. For a small value of the parameter σ, the criterion will

consist of Boolean operators that 'count' the number of overlapping points between the two binary images for a given transformation. Equation (2) can be derived from a simple combinatorial function that determines point-to-point overlap by analytic mollification (Murio, 1993), or by the Gaussian expression and relaxation. Our registration method uses a continuously differentiable function that maximizes both overlap and local shape similarity between the feature maps. To avoid spurious results (mainly due to local maxima in the criterion), a regularizing term is introduced to restrict the 6-parameter affine deformation $T[\cdot]$. The resulting registration criterion E_{reg} can be defined as:

$$E_{\text{reg}}(T) = E(T) + \lambda \sum_X \|Y - T[Y]\|^2 \qquad (4)$$

where λ denotes a Lagrange multiplier associated with the constraint. The parameters of the affine transformation are computed by minimizing the criterion function in Eq. (4) using a standard quasi-Newton algorithm (Press et al., 1992).

The image registration method was tested on several visible and thermal IR image pairs from the UTK-IRIS database using edge maps extracted by the Canny edge detectors. Figure 2 shows an example of the visible-thermal image registration method used in this paper. Figure 2(a) and (b) are visible and thermal images of the same person. The edge maps in Fig. 2(c) show the unregistered position of the image pair. Figure 2(d) and (e) show the image pair after registration using the proposed method. The algorithm emphasizes the importance of prominent face components such as the eyes, mouth, and eyebrows features marked by a rectangular bounding-box and minimizes the effect of the occluding contours. The occluding contours are labeled and eliminated using a simple search heuristic that uses knowledge about the shape of the face and starts the process by scanning horizontally and vertically the edge maps until encountering the outer edges. The hair region is also removed using a projection histogram technique.

The registration criterion consists of a mixture of N_D Gaussian functions evaluated at N_M points then summed together. Direct evaluation requires the computational amount of $O(N_M \times N_D)$. A numerical method called the Fast Gauss Transform (FGT) reduces the computational complexity of the Gaussian mixture evaluation to $O(N_M + N_D)$ in color modeling and tracking applications (Elgammal et al., 2003). This method belongs to a class of fast evaluation algorithms
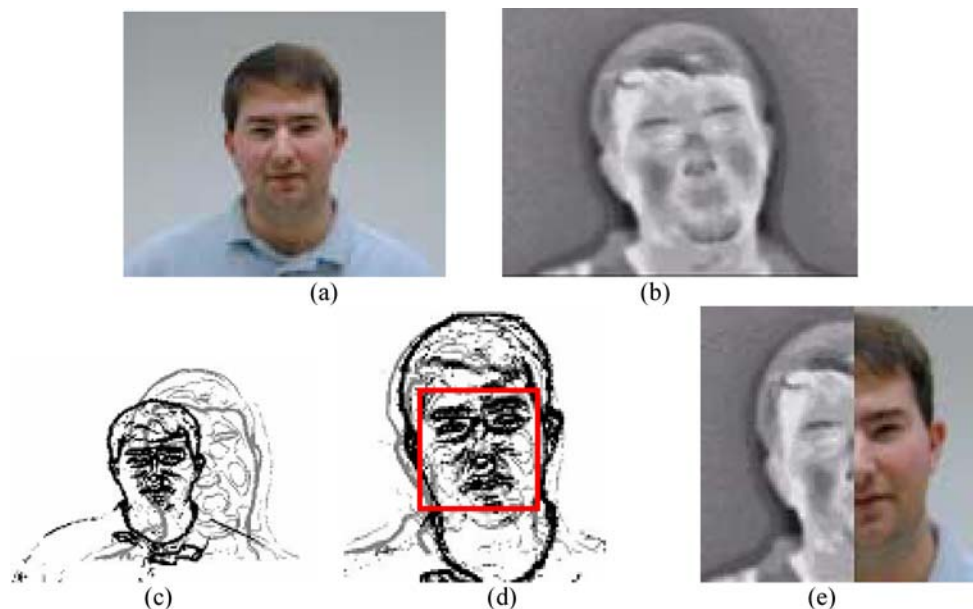
Figure 2. Registration of visible and thermal face images using the Gaussian criterion. (a) Visible image, (b) Thermal image, (c) Edge maps before registration, (d) Edge maps after registration with square region of registration, and (d) Aligned image pair after registration. Note that the emphasis here is to align prominent features of the face, i.e., eyes, nose, mouth, and not occluding contours.

known as "fast multipole" methods (Greengard, 1988) and was applied to the computation of potential fields. The sources and targets of potential fields are clustered using suitable data structures, and the sums are replaced by smaller summations that are equivalent to a given level of precision. Implementation details and analysis can be found in Elgammal et al. (2003), and Greengard and Strain (1991). The main problem with the original FGT is the exponential increase of the complexity with the number of dimensions. Yang et al. (2003) improved the original FGT to address this limitation in that a data-clustering scheme along with a more intelligent multivariate Taylor expansion allows for further computational gains even in high dimensions. The computational efficiency of the criterion evaluation is shown in Fig. 3 for different size point-sets. The gains in computational costs increase dramatically with the size of the point-set.

3.2. Eyeglass Detection in Thermal Images Using Ellipse Fitting

When eyeglasses are present in face images, direct intensity-based fusion of visible and thermal images will fail to produce the fused images useful for robust face recognition. Detection and replacement of eyeglass regions with eye templates in thermal face images can improve the quality of visible-thermal image fusion. Eyeglass regions can be segmented out easily from thermal face images due to large intensity differences with other parts of the face. Figure 4(a) and (b) represent result of the thermal images averaged
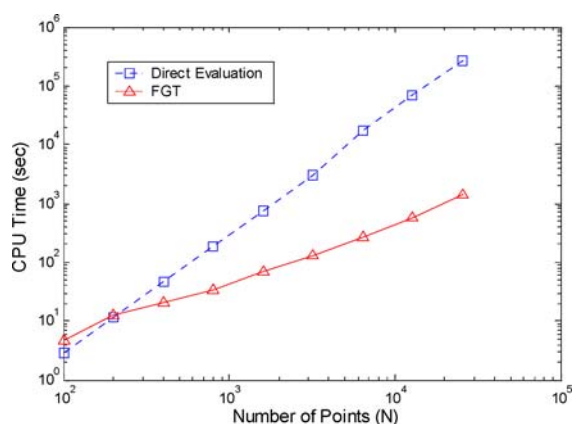

Figure 3. Computational efficiency of the Fast Gauss Transform. The CPU time (Pentium IV, 2.3 GHz) required for the evaluation of the Gaussian criterion for a given number of points ($N = N_M = N_D$).
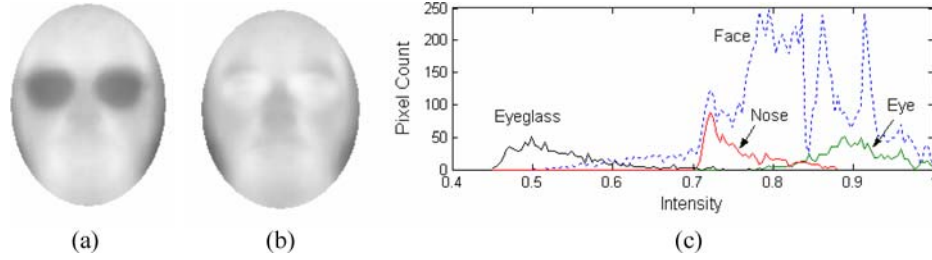
Figure 4.   Thermal images (a) with eyeglasses and (b) with no eyeglasses averaged over 60 face samples, (c) Histograms of facial components in terms of relative intensity (temperature).

over 60 face samples from the NIST/Equinox database (http://www.equinoxsensors.com/products/HID.html) for the cases with and without eyeglasses. Figure 4(c) shows the histograms of various facial components such as eyeglasses, the eyes, the nose, and the facial region in the averaged thermal face images. The sizes of the face components used in the histogram are 35×25 pixels for the eye and glasses, 24×35 pixels for the nose, and 150×120 for the face in images of size 240×320. The eyeglass region is usually associated with the lowest temperature of the face in thermal images. The eyes are relatively warmer areas, while the eyebrows, the mouth, and the cheeks often all have similar temperature ranges.

The eyeglass regions in thermal face images can be represented by ellipses. A thermal image, binarized using a threshold, provides the data points for ellipse fitting. Morphological operations reduce the noisy patterns and small blobs in the edge images. After morphological filtering for noise reduction, the data points in the binarized image are connected using the Freeman chain coding with 8-connectivity (Freeman and Davis, 1977). Small-sized connected components are ignored to reduce the computational amount of ellipse fitting. Non-elliptical shapes such as the connected components close to a line or non-closed contours are also removed.

A non-iterative ellipse-fitting algorithm is applied to each of the connected components to produce an ellipse. An equation for an ellipse can be denoted as a product of a coefficient vector A and an independent variable (data points) X:

$$F(A, X) = AX = ax^2 + bxy + cy^2 + dx + ey + f = 0 \qquad (5)$$

where $A = [a\,b\,c\,d\,e\,f]$ and $X = [x^2\,xy\,y^2\,x\,y\,1]^T$.

Ellipse fitting minimizes the algebraic distance $E = |F(A, X)|^2$ to a given set of data points in terms of least squares. If a quadratic constraint is set on the parameters, this minimization problem can be solved by a generalized eigenvalue analysis (Bookstein, 1979) described as

$$D^T DA = SA = \lambda CA \qquad (6)$$

where $D = [X_1 X_2 \ldots X_n]^T$ is called a design matrix of n data points, $S = D^T D$ refers to a scatter matrix, and C is a constant matrix. Fitzgibbon et al. (Fitzgibbon et al., 1999) proposed a direct least-square conic fitting algorithm with the constraint $A^T CA = 1$, where C is a 6×6 matrix with $C(1,3) = -2$, $C(2,2) = 1$, and $C(3,1) = -2$, and zeros, elsewhere. This can be interpreted as applying constraints using $b^2 - 4ac = 1$ with unique solution A, which minimizes the algebraic distance E. Only one negative eigenvector can exist and becomes the solution for the ellipse. Direct least squares method is used as a non-iterative ellipse-fitting algorithm. It yields the best least square ellipse fitting approach and is robust to noise. After finding the solutions for fitting ellipses, it is convenient to convert into the standard form of an ellipse to use its parameters effectively. Figure 5 shows an ellipse with the parameters used for eyeglass detection in thermal face images. Ci denotes the center of the ith ellipse. $2\alpha_i$, and $2\beta_i$ are the
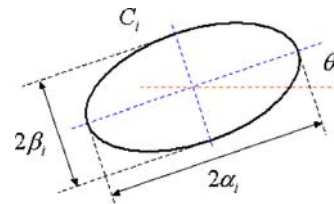

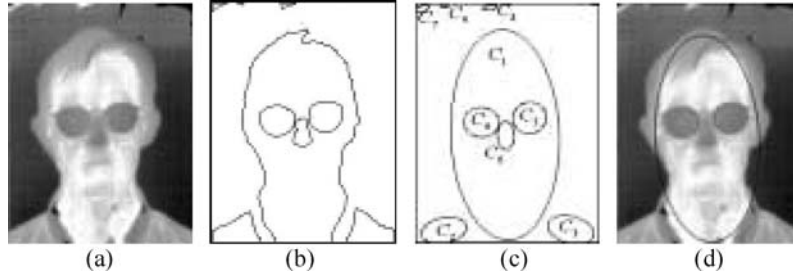*Figure 5.*   Parameters of an ellipse C_i.

Figure 6. Eyeglass detection example using ellipse fitting. (a) Original thermal image, (b) Connected components of the binarized image, (c) Ellipse fitting results, and (d) Eyeglass regions detected using the ellipse fitting method.

lengths of the major and the minor axes respectively, and θ_i indicates the orientation angle of the ellipse in the range of $-\pi/2 < \theta_i < \pi/2$.

Ellipse fitting can also be used for face detection since the biggest ellipse will always be associated with the face. Similarities of all the ellipses within the face region, or inside the biggest ellipse, are tested for possible eyeglass candidates. Among all the ellipses available, two ellipsoidal shapes of similar shape, size, and symmetrically located are considered as candidate eyeglasses in thermal images. In this paper, we used the similarity of ith and jth ellipsoids as a measure to determine eyeglasses. The similarity between ellipses C_i and C_j is defined as:

$$S_{ij} = \frac{\alpha_i \beta_i}{\alpha_j \beta_j} \left(\frac{1}{1 + |\theta_{ij}|} \right) \left(\frac{1}{1 + |\theta_i + \theta_j|} \right) \quad (7)$$

where θ_{ij} represents the angle of the line segment that connects the centers of the two ellipses C_i and C_j. We assume that $\alpha_j \beta_j > \alpha_i \beta_i$ so that the similarity measure S_{ij} is less than 1. Two ellipses with the highest similarity measures of $S_{ij} > 0.7$ are considered as eyeglasses. Figure 6 illustrates an example of detecting eyeglasses in thermal images using the ellipse fitting. Complete ellipses are generated from each connected component. The biggest ellipse (C_1) corresponds to the face. Ellipses outside the face region (C_2, C_3, C_7, C_8, and C_9) are not considered for similarity checking. For the three ellipses inside the face region, the similarities are calculated as $S_{45} = 0.96$, $S_{46} = 0.38$, and $S_{56} = 0.40$. As a result, the two ellipses C_4 and C_5 with the highest similarity are identified as eyeglasses.

3.3. Fusion of Visible and Thermal Face Images

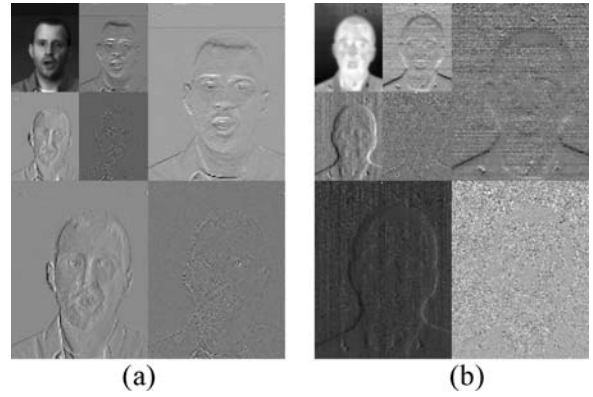This paper applies data fusion of visible and thermal IR image pairs in the discrete wavelet transform

Figure 7. 2D level-2 discrete wavelet decomposition of visible and thermal images. (a) Visible image, (b) Thermal image.

domain. Assuming the two images are co-registered and of the same sizes, visible and thermal IR images are transformed into a set of approximation and details components at a predefined analysis level using the DWT. The wavelet coefficients of the fused image are obtained by computing a weighted average of the wavelet coefficients of visible and thermal IR images. The visible and thermal images are decomposed into approximation coefficients $W_\varphi(m, n)$ and detail coefficients $W_\psi(m, n)$. Figure 7 shows an example of level-2 DWT decompositions of visible and thermal images using the Haar wavelets (Daubechies, 1992).

The data fusion used in this paper consists of a weighted combination of corresponding DWT coefficients of visible and thermal images.

$$W_\varphi(m, n) = \alpha_1 W_\varphi^{\text{visible}}(m, n) + \beta_1 W_\varphi^{\text{thermal}}(m, n) \quad (8)$$

$$W_\psi(i, j) = \alpha_2 W_\psi^{\text{visible}}(i, j) + \beta_2 W_\psi^{\text{thermal}}(i, j) \quad (9)$$
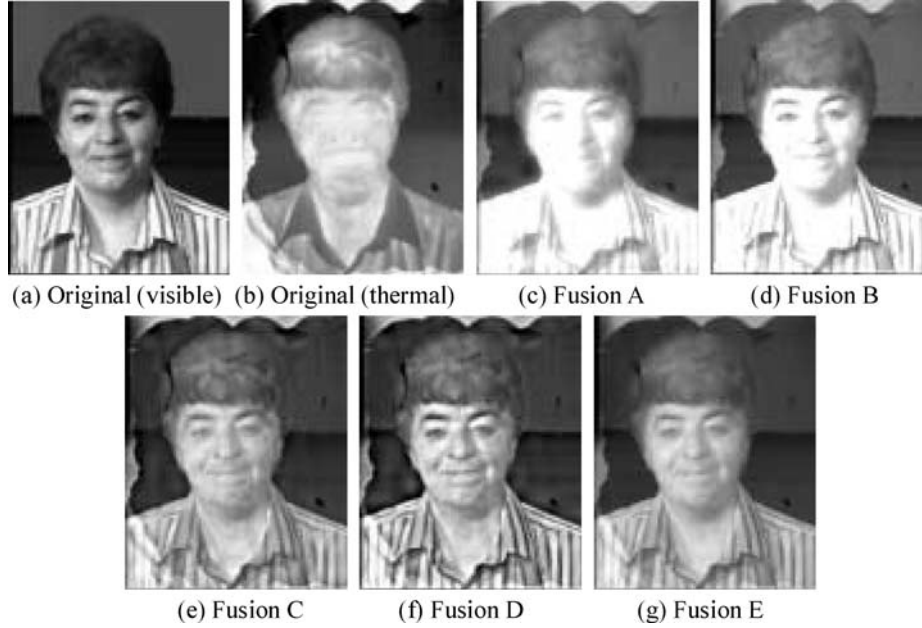
(a) Original (visible) (b) Original (thermal)   (c) Fusion A   (d) Fusion B

(e) Fusion C   (f) Fusion D   (g) Fusion E

Figure 8.   Sample fused images in side lighting condition.

The coefficients α_1 and β_1 represent the weight factors of each approximation coefficient, while α_2 and β_2 are used to combine the detail coefficients. Generally, the weight factors can be determined according to the information content of the visible and thermal images. When a subject is measured in side lighting or darkness conditions, the visible image may not contain useful information for face recognition. Then the weights will be chosen so that thermal information is emphasized. When the variation in ambient temperature becomes too big, thermal components may not provide reliable information. Hence the weights will need to be adjusted to reflect more visible information in the fusion. Frequency contents of visible and thermal images will also be considered in the fusion. Fused images are reconstructed by the inverse DWT of the approximation and the detail coefficients.

$$F_w(x, y) = IDWT[W_\varphi(m, n), W_\psi(i, j)] \quad (10)$$

Several weight settings are tested to validate the proposed multiscale visible-thermal image fusion technique. This test involves a gallery set and two probe sets selected from the NIST/Equinox database. The gallery set contains 62 frontal face images of 62 individuals with frontal lighting and neutral expressions. The probe sets have 216 visible and thermal face image pairs with varying expressions drawn from challenging illumination conditions: side lighting and low illumination. Fusion A uses more low-frequency (LF) visible information in the fusion ($\alpha_1 = 1$, $\beta_1 = 0.5$, $\alpha_2 = \beta_2 = 0.5$). Fusion B emphasizes both the low- and high-frequency visible components ($\alpha_1 = 1$, $\beta_1 = 0.5$, $\alpha_2 = 1$, $\beta_2 = 0.5$). In Fusion C, more high-frequency (HF) thermal information is used ($\alpha_1 = \beta_1 = 0.5$, $\alpha_2 = 0.5$, $\beta_2 = 1$) while Fusion D puts emphasis on HF visible and thermal information ($\alpha_1 = \beta_1 = 0.5$, $\alpha_2 = \beta_2 = 1$). Fusion E combines all the components with equal weights ($\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0.5$). The other cases with emphases on thermal information ($\alpha_1 = 0.5$, $\beta_1 = 1$, $\alpha_2 = 0.5$, $\beta_2 = 1$), LF thermal ($\alpha_1 = 0.5$, $\beta_1 = 1$, $\alpha_2 = 0.5$, $\beta_2 = 0.5$), and LF visible/thermal ($\alpha_1 = \beta_1 = 1$, $\alpha_2 = \beta_2 = 0.5$) produce severely faded fused images with little visual details. The weight setting with emphasis on HF visible ($\alpha_1 = \beta_1 = 0.5$, $\alpha_2 = 1$, $\beta_2 = 0.5$) showed fused images very similar to Fusion E. Figures 8 and 9 show sample probe images in side lighting and low-illumination conditions. Fusion D, which emphasizes the high frequencies in both the thermal and the visible modalities, generates most visually enhanced fused images among the weight settings examined in this paper.

Table 1 shows the first match and cumulative first-10 match percentage accuracies of the fused images

Table 1. First match and cumulative first-10 match percentage accuracies of the multiscale visible/thermal fusion with different weight settings in side lighting and low-illumination conditions.

Fusion	Weights ($\alpha_1, \beta_1, \alpha_2, \beta_2$)	Emphasis	Side lighting	Low illumination
A	1.0, 0.5, 0.5, 0.5	LF Visible	86.1 (95.8)	84.7 (94.9)
B	1.0, 0.5, 1.0, 0.5	Visible	90.2 (98.1)	89.8 (97.2)
C	0.5, 0.5, 0.5, 1.0	HF Thermal	87.5 (91.2)	88.9 (91.7)
D	0.5, 0.5, 1.0, 1.0	HF Visible/Thermal	**93.5** (97.2)	**94.9** (96.7)
E	0.5, 0.5, 0.5, 0.5	Equal	90.2 (96.2)	94.0 (95.8)

with different weight settings for the probe sets in side lighting and low-illumination conditions. Cumulative first-10 match accuracies are listed inside the parentheses. The weight setting with emphasis on HF visible/thermal information demonstrates highest first match accuracy for both of the probe sets among the five weight settings chosen.

Figure 10 visually compares the multiscale image fusion of Fusion D ($\alpha_1 = \beta_1 = 0.5$, $\alpha_2 = \beta_2 = 1$) with a simple direct intensity-average of a visible image $V(x,$

(a) Original (visible) (b) Original (thermal) (c) Fusion A (d) Fusion B

(e) Fusion C (f) Fusion D (g) Fusion E
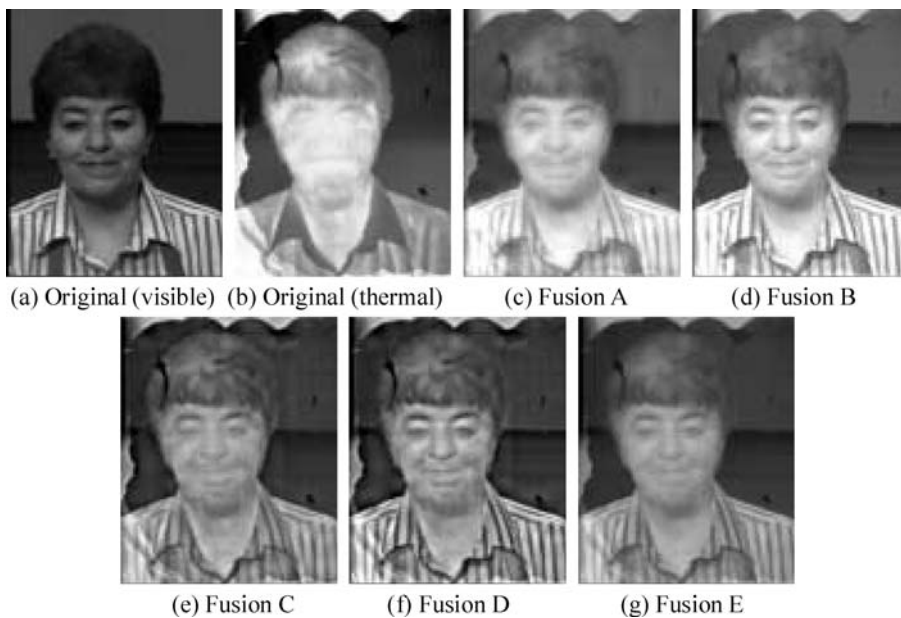
Figure 9. Sample fused images in low-illumination condition.

(a) $V(x,y)$ (b) $T(x,y)$ (c) $F_a(x,y)$ (d) $F_w(x,y)$
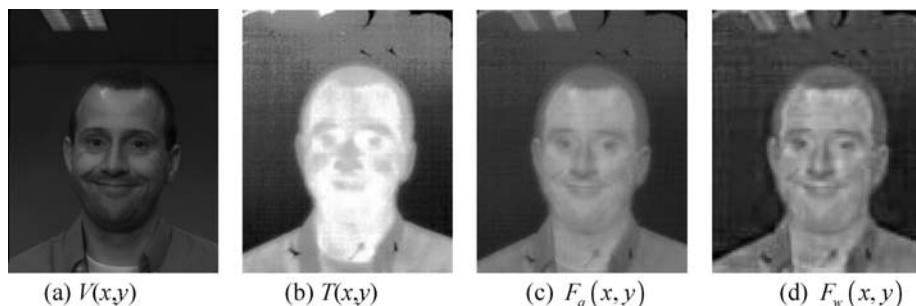
Figure 10. Data fusion of visible and thermal images with no eyeglasses. (a) Visible image, (b) Thermal image, (c) Intensity-average image fusion, and (d) multiscale image fusion.
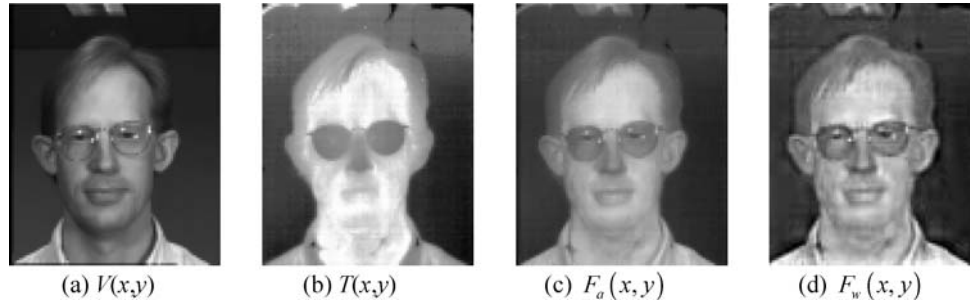
Figure 11.　Data fusion of visible and thermal images with eyeglasses. (a) Visible image, (b) Thermal image, (c) Intensity-average fusion, and (d) Multiscale image fusion.

y) and a thermal image $T(x, y)$ in the spatial domain:

$$F_a(x, y) = 0.5\, V(x, y) + 0.5\, T(x, y) \qquad (11)$$

High-frequency details are significantly emphasized in the images of multiscale fusion, which attribute to the improvement of face recognition accuracy as well as visual quality.

Data fusion based on fixed weights will encounter difficulties especially when the subject is wearing eyeglasses. Eyeglass regions appear dark in thermal images due to the fact that glass blocks a large portion of thermal energy from the face.Figure 11 compares the data fusion methods of visible and thermal images of a person with eyeglasses using the same parameters as in the experiment of Fig. 10.

Detected eyeglass regions need to be replaced with an eye template in the thermal images to enhance the visual quality around the eyes in data-fused images. The eye templates are obtained from the average of all thermal face images with no glasses. A geometrical transformation of the eye template is performed to fit the templates to the eyeglass regions detected by ellipse fitting. Left and right eye templates are superimposed on the eyeglass regions respectively after rotating and resizing the original ellipse pattern obtained from the eyeglass detection procedure discussed in Section 3.2.Figure 12 shows the procedure of eyeglass detection and replacement with thermal eye template patterns.

Figure 13 shows a multiscale data fusion result with eyeglass detection and replacement by the eye template of thermal images for image pairs from the NIST/Equinox and the UTK-IRIS databases. Eyeglass replacement substantially enhances the visual quality of data-fused images.
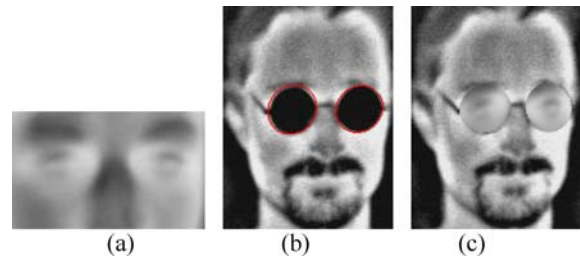

Figure 12.　An eyeglass detection and replacement example. (a) Average thermal images used to compute the eye template, (b) Detected eyeglasses, and (c) Eyeglasses replaced by the eye template.

4.　Performance Evaluation of Face Recognition with Data Fusion

The performance of face recognition with fusion of visible and thermal images is evaluated using the NIST/Equinox and UTK-IRIS databases. Visible-thermal image pairs in the UTK-IRIS database are software registered using the image registration algorithm described in Section 3.1 before the fusion step. The fusion of visible and thermal images is conducted in DWT domain to produce visible-thermal fused images. Face recognition accuracies are compared with single modality face recognition methods in terms of first match and first 10 match rates.

4.1.　NIST/Equinox Database

The NIST/Equinox database (http://www.equinox-sensors.com/products/HID.html) consists of 1,622 pairs of visible and calibrated thermal IR face images from 90 individuals. The visible and thermal IR image pairs are co-registered within 1/3 pixel with a hardware setting have a spatial resolution of 320 × 240

Table 2. The NIST/Equinox database of visible and thermal IR face images.

Dataset	Visible (Thermal)	Eyeglass	Lighting	Expression
Gallery	90 (90)	Off	Frontal	Neutral
Probe 1	283 (283)	Off	Frontal	Various
Probe 2	370 (370)	Off	Left	Various
Probe 3	365 (365)	Off	Right	Various
Probe 4	177 (177)	On	Frontal	Various
Probe 5	172 (172)	On	Left	Various
Probe 6	165 (165)	On	Right	Various

pixels, and a grayscale resolution of 8 bits (visible) and 12 bits (IR). Table 2 describes the NIST/Equinox database used in the experiments. "Gallery" denotes the images in the database with known identity while "probe" indicates the images presented to the system

for identification. One image for each face taken with frontal lighting conditions is used for the gallery. Probe images are divided according to the conditions such as the presence of eyeglasses, lighting directions, and facial expressions. Figure 14 shows two pairs of visible and thermal IR image samples with and without eyeglasses from the NIST/Equinox database.

Figure 15 shows the face recognition results using data-fused visible and thermal images (Df) in the DWT domain. The images on the first column are probe images. Face recognition results corresponding to the five best matches are displayed in descending order of matching score S_D in the range [0, 10] generated by FaceIt$^{®}$. Multiscale data fusion with eyeglass replacement obtained matching scores higher than 9.0 for the first match for probe images with and without eyeglasses.

Figure 16 compares the performances of face recognition techniques in terms of the first 10 best matches for data with glasses (probes 4, 5, 6) and without glasses

Figure 13. DWT-based image fusion with eyeglass replacement for image samples from (a) NIST/Equinox database and (b) UTK-IRIS database. From left to right are: Visible image, Thermal image, Fused image without eyeglass replacement, and Fused image after eyeglass replacement.
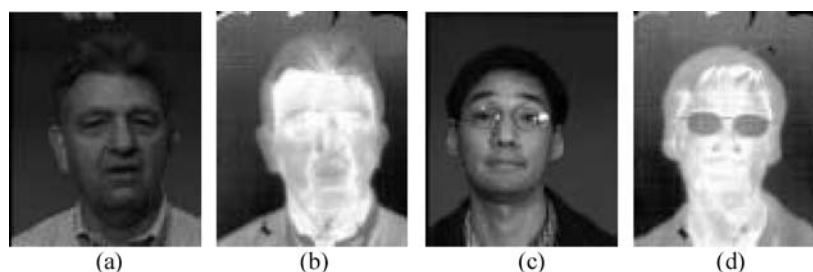

Figure 14. Visible and thermal image pairs from the NIST/Equinox database. (a, b) Image pair with no eyeglasses, (c, d) Image pair with eyeglasses.

Figure 15. First five matches of face recognition using data fusion (Df). (a) Probe with no eyeglasses, (b) Probe with eyeglasses.

Figure 16. Performance comparison of face recognition methods using the NIST/Equinox database. (a) With no eyeglasses, (b) With eyeglasses, (c) Eyeglasses replaced with the eye template, and (d) First match accuracies before and after eyeglass replacement for visible (Vi), thermal (Th), and data fusion (Df).
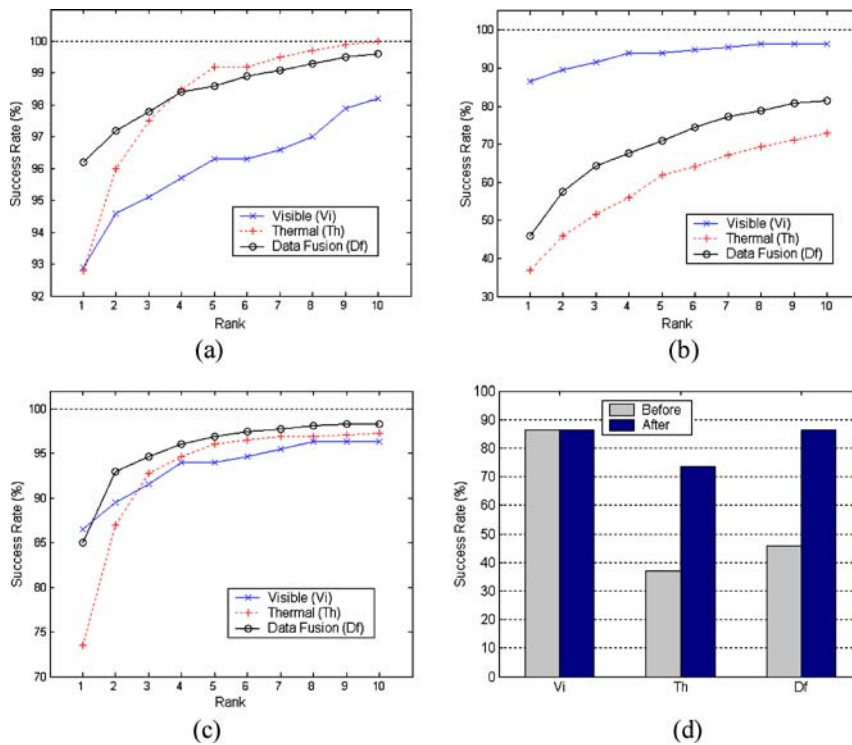
(probes 1, 2, 3). A plot of probabilities of correct match versus the number of best matching scores called a cumulative match characteristic (CMC) curve is used for the performance evaluation. The CMC curve is typically used for the analysis of face identification performance. Figure 16(a) shows that face recognition based on individual spectra and data fusion all produce higher than 90% accuracies when the subjects are not wearing glasses. However, face recognition with thermal images and data fusion show higher recognition accuracies over visible images. In Fig. 16(b), face recognition using thermal images shows unsatisfactory
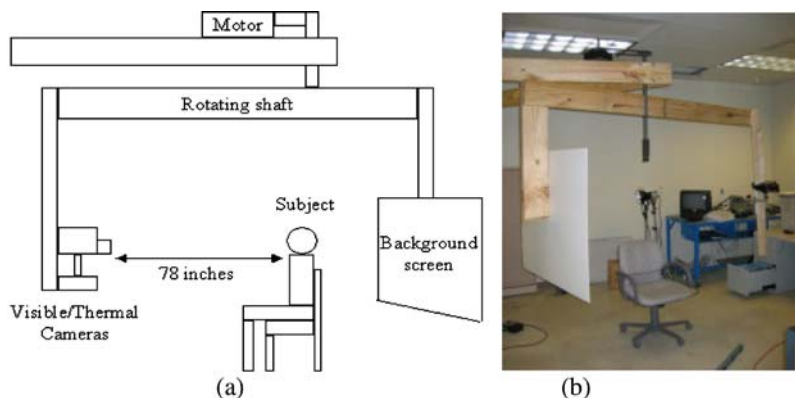
Figure 17. The UTK-IRIS rotary imaging platform. (a) A schematic diagram, (b) Imaging setup.

results with lower than 40% accuracies for the first match since glass blocks a large portion of thermal radiation energy. The presence of eyeglasses slightly lowers the performance of visible face recognition, but not as much as in thermal face recognition. In this case, a direct fusion of visible and thermal images using intensity average also results in low recognition accuracy. Figure 16(c) demonstrates that eyeglass replacement greatly improves the face recognition performance based on thermal images and the data fused images. After eyeglass replacement, the performance of thermal face recognition increases. Figure 16(d) shows first match rates before and after eyeglass replacement. Eyeglass replacement significantly improved the first match rates for face recognition with both thermal images and data fusion.

4.2. UTK-IRIS Database

The UTK-IRIS database (http://www.cse.ohio-state.edu/otcbvs-bench/) contains 4,350 pairs of visible and thermal IR face images of 32 individuals,

Table 3. The UTK-IRIS database of visible and thermal IR face images.

Dataset	Visible (Thermal)	Eyeglass	Lighting	Expression
Gallery	29 (29)	Off	Frontal	Neutral
Probe 1	41 (41)	Off	Various	Various
Probe 2	30 (30)	On	Various	Various

with 132–250 images per person from diverse ethnic groups. The imaging setup has visible and thermal IR cameras located on a 32-inch long shaft rotating around the subject at the center (see Fig. 17).The cameras scan the face in 180 degrees at a constant speed in a clockwise circular movement from one side to the other. At each rotational scanning, the system produces 11 images under challenging environments that involves variations in illumination, facial expressions, pose angles, and appearances including disguises. Images are taken under normal room lights on and off to test face recognition in darkness. Two additional front spotlights created four different illumination
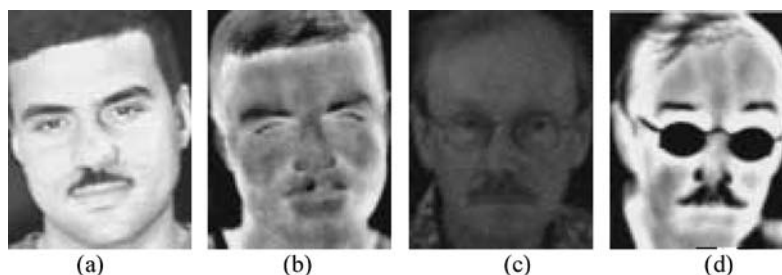

Figure 18. Visible and thermal image pairs from the UTK-IRIS database. (a; b) Image pair with no eyeglasses, (c; d) Image pair with eyeglasses.

Figure 19. First five matches of face recognition with data fusion (Df). (a) Probe with no eyeglasses, (b) Probe with eyeglasses.
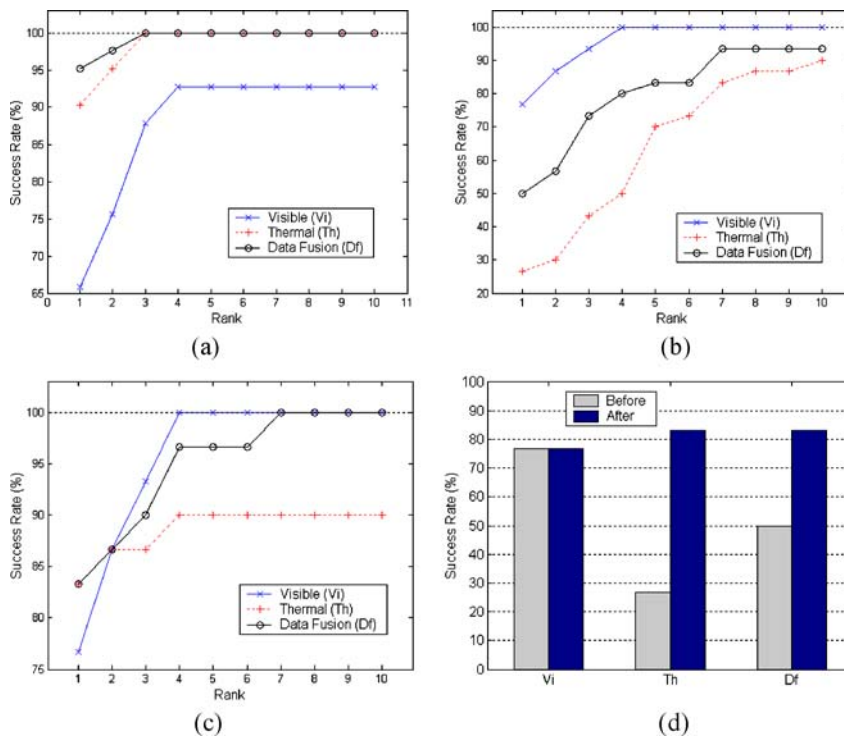

Figure 20. Performance comparison of face recognition methods using the UTK-IRIS database. (a) With no eyeglasses, (b) With eyeglasses, (c) Eyeglasses replaced with the eye template, and (d) First match accuracies before and after eyeglass replacement for visible (Vi), thermal (Th), and data fusion (Df).

conditions under normal room lights: Both on, Left on, Right on, and Both off. Three facial expressions are tested: talking, laughing, and angry. People who wear eyeglasses are asked to take pictures with and without eyeglasses. Raytheon Pro, an uncooled type IR camera, in the spectrum of 7–12 μm is used to take thermal IR images. Visible and thermal image pairs are captured at 30 frames per second of the size 320 × 240. Table 3 summarizes a subset (100) of frontal image pairs from the UTK-IRIS database used for performance evaluation in this paper. Figure 18 shows image examples from the UTK-IRIS databases.

Table 4. First match success rates in percentage before and after eyeglass replacement. First ten match rates are shown in parentheses.

	NIST/Equinox			UTK-IRIS		
	Vi	Th	Df	Vi	Th	Df
With glasses	86.5 (96.3)	37.0 (72.8)	45.9 (81.5)	76.7 (100)	26.7 (90.0)	50.0 (93.3)
Glasses replaced	86.5 (96.3)	73.5 (97.3)	85.0 (98.3)	76.7 (100)	83.3 (90.0)	83.3 (100)

Figure 19 shows a typical example of first five matches of face recognition using fused visible-thermal images from the UTK-IRIS database in the order of matching score.

Figure 20 reveals that face recognition using the UTK-IRIS database and a software registration approach produces consistent results as obtained using the NIST/Equinox database, which uses a hardware registration in Fig. 16. Figure 20(a) shows face recognition results with probe images without glasses (Probe 1). When individuals are not wearing eyeglasses, thermal images and data fusion show better performance over visible images. In Fig. 20(b) and (c), faces with glasses (Probe 2) were used to test recognition accuracies. Figure 20(d) shows first match success rates of face recognition using thermal and data-fused images before and after glass replacement. After glass replacement, accuracies are significantly improved for face recognition with both thermal IR and data-fused images.

Table 4 summarizes the values of first match and first ten match success rates in percentage before and after eyeglass replacement shown in Figs. 16(d) and 20(d). The data fusion achieved 85% (NIST/Equinox) and 83.3% (UTK-IRIS) first match success rates and 98.3% (NIST/Equinox) and 100% (UTK-IRIS) first 10 match success rates. The eyeglass replacement increased the number of correct first match subjects by 85% from 236 to 437 (NIST/Equinox) and 67% from 15 to 25 (UTK-IRIS).

5. Conclusion

This paper presents the registration fusion of visible and thermal IR image data for robust face recognition in challenging operating environments that involve illumination variations. The combined use of visible and thermal IR image sensors offers a viable means for improving the performance of face recognition techniques based on single imaging modalities. Recognizing the faces with visible images demonstrates difficulty in low illumination conditions. Thermal IR sensors measure energy radiation from the object, which is less sensitive to illumination changes and operable in darkness. Multiscale data fusion of visible and thermal images produces combined face images less sensitive to illumination variations and hence improves face recognition accuracies. When a subject is wearing eyeglasses, however, thermal face images may lose the information around the eyes useful for recognition since glass blocks a large portion of thermal energy. In this paper, eyeglass regions are detected by ellipse fitting of thermal images, and are replaced with eye template patterns to preserve the details in the fused image. Face recognition techniques using visible, thermal IR, and data-fused visible-thermal images are compared using a commercial face recognition software package (FaceIt$^{®}$) and two visible-thermal face image databases (the NIST/Equinox and the UTK-IRIS databases).

The proposed multiscale data-fusion technique greatly improves the recognition accuracy under a wide range of illumination changes, compared with the cases of single modality and intensity average fusion. Extensive experimental results show that thermal face recognition performed better than visible face recognition under various illumination and facial expression conditions in case of no eyeglasses. When eyeglasses are present in the image, thermal face recognition and a direct data fusion without eyeglass replacement produced lower performance. Eyeglass replacement significantly improved the recognition accuracy of the data fusion. Multiscale fusion approaches with eyeglass replacement consistently outperformed signal modality face recognition. Experimental results showed that the data fusion achieved 85% (NIST/Equinox) and 83.3% (UTK-IRIS) first match success rates and 98.3% (NIST/Equinox) and 100% (UTK-IRIS) first 10 match success rates. The eyeglass replacement increased the number of correct first match subjects by 85% (NIST/Equinox) and 67% (UTK-IRIS).

Acknowledgment

References

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

Ben-Yacoub, S., Abdeljaoued, Y., and Mayoraz, E. 1999. Fusion of face and speech data for person identity verification. *IEEE Trans. on Neural Networks*, 10(5):1065–1074.

Besl, P.J. and McKay, N.D. 1992. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(2):239–256.

Blackburn, D.M., Bone, J.M., and Phillips, P.J. 2001. Face recognition vendor test 2000. Evaluation Report, National Institute of Standards and Technology, pp. 1–70.

Bone, M. and Blackburn, D. 2002. Face recognition at a chokepoint: Scenario evaluation results. Evaluation Report, Department of Defense.

Bookstein, F.L. 1979. Fitting conic sections to scattered data. *Computer Graphics and Image Processing*, 9(1):56–71.

Boughorbel, F., Koschan, A., Abidi, B., and Abidi, M. 2004a. Gaussian fields: A new criterion for 3D rigid registration. *Pattern Recognition*, 37(7):1567–1571.

Boughorbel, F., Koschan, A., Abidi, B., and Abidi, M. 2004b. Gaussian energy functions for registration without correspondences. In *Proc. of 17th Int'l Conf. on Pattern Recognition*, pp. 24–27 Cambridge, UK.

Brunelli R. and Poggio, T. 1993. Face recognition: Features versus templates. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(10):1042–1052.

Brunelli, R. and Falavigna D. 1995. Personal identification using multiple cues. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(10):955–966.

Burton, A.M., Bruce, V., and Hancock, P.J.B. 1999. From pixels to people: A model of familiar face recognition. *Cognitive Science*, 23(1):1–31.

Chang, K., Bowyer, K.W., Sarkar, S., and Victor, B. 2003. Comparison and combination of ear and face image in appearance-based biometrics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1160–1165.

Chellappa, R., Wilson, C.L., and Sirohey, S. 1995. Human and machine recognition of faces: A survey. In *Proceedings of the IEEE*, Vol. 83, no. 5, pp. 705–740.

Cox, I.J., Ghosn, J., and Yianilos, P.N. 1996. Feature-Based Face Recognition Using Mixture-Distance. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 209–216.

Charpiat, G., Faugeras, O., and Keriven, R. 2003. Shape Metrics, Warping and Statistics. In *Proc. Int'l Conf. on Image Processing*, Barcelona, Vol. 2, pp. 627–630.

Chen, X., Flynn, P. and Bowyer, K. 2003. Visible-light and infrared face recognition. In *Proc. of Workshop on Multimodal User Authentication*, pp. 48–55.

Craw, I., Costen, N., Kato, T., and Akamatsu, S. 1999. How should we represent faces for automatic recognition?. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(8):725–736.

Dalley, G. and Flynn, P. 2002. Pair-wise range image registration: A study in outlier classification. *Computer Vision and Image Understanding*, 87(1–3):104–115.

Dasarathy, B.V. 1994. *Decision Fusion*, Washington. DC: IEEE Computer Society Press.

Daubechies, I. 1992. Ten lectures on wavelets. *CBMS-NSF Lecture Notes*, No. 61, Society for Industrial and Applied Mathematics.

Elgammal, A., Duraiswami, R., and Davis, L. 2003. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(11):1499–1504.

Fang, Y., Tan, T., and Wang, Y. 2002. Fusion of global and local features for face verification. In *Proc. of Int'l Conf. on Pattern Recognition*, pp. 382–385.

Fitzgibbon, A.W. 2003. Robust registration of 2D and 3D Point Sets. *Image and Vision Computing*, 21(13):1145–1153.

Fitzgibbon, A., Pilu, M., and Fisher, R.B. 1999. Direct least square fitting of ellipses. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(5):476–480.

Freeman, H. and Davis, L.S. 1977. A corner-finding algorithm for chain-coded curves. *IEEE Trans. on Computer*, 26(3):297–303.

Greengard, L. 1988. *The Rapid Evaluation of Potential Fields in Particle Systems*. Cambridge, MA: MIT Press.

Greengard, L. and Strain, J. 1991. The fast gauss transform. *SIAM Journal on Scientific Computing*, 12:79–94.

Gutta, S., Huang, J.R.J., Jonathon, P., and Wechsler, H. 2000. Mixture of experts for classification of gender, ethnic origin, and pose of human faces. *IEEE Trans. Neural Networks*, 11(4):948–960.

Hall, D.L. and Llinas, J. 2001. *Handbook of Multisensor Data Fusion*. CRC Press.

Heo, J., Abidi, B., Kong, S.G., and Abidi, M. 2003. Performance comparison of visual and thermal signatures for face recognition. *Biometric Consortium Conference*, Arlington, VA.

Ho, T.K., Hull, J.J., and Srihari, S.N. 1994. Decision combination in multiple classifier systems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(1):66–75.

Hong, L. and Jain, A. 1998. Integrating faces and fingerprints for personal identification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(12):1295–1307.

http://www.cse.ohio-state.edu/otcbvs-bench/

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

Huttenlocher, D.P., Klanderman, G.A., and Rucklidge, W.J. 1993. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(9):850–863.

Irani, M. and Anandan, P. 1998. Robust multi-sensor image alignment. In *Proc. of the 6th Int'l Conf. on Computer Vision*, pp. 959–965.

Jain, A.K., Ross, A., and Prabhakar, S. 2004. An introduction to biometric recognition. *IEEE Trans. Circuits and Systems for Video Technology*, 14(1):4–20.

Kanade, T. 1973. Picture processing by computer complex and recognition of human faces. *Technical Report*, Kyoto University.

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

Manjunath, B.S., Chellappa, R., and von der Malsburg, C. 1992. A feature based approach to face recognition. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 373–378.

Murio, D.A. 1993. *The Mollification Method and the Numerical Solution of Ill-Posed Problems*. New York, NY: John Wiley & Sons.

Pavlidis, I. and Symosek, P. 2000. The imaging issue in an automatic face/disguise detection system. In *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pp. 15–24.

Penev, P.S. 1998. Local feature analysis: A statistical theory for information representation and transmission. Ph.D. Thesis, The Rockefeller University.

Penev, P.S. 1999. Dimensionality reduction by sparsification in a local-features representation of human faces. Technical Report, The Rockefeller University.

Phillips, P.J., Grother, P., Micheals, R.J., Blackburn, D.M., Tabassi, E., and Bone, M. 2003. Face recognition vendor test 2002. Evaluation Report, National Institute of Standards and Technology, pp. 1–56.

Phillips, P.J., Moon, H., Rizvi, S.A., and Rauss, P.J. 2000. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(10):1090–1104.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. Second edition. Cambridge University Press.

Prokoski, F. 2000. History, current status, and future of infrared identification. In *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pp. 5–14.

Ross, A. and Jain, A. 2003. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125.

Selinger, A. and Socolinsky, D.A. 2001. Appearance-based facial recognition using visible and thermal imagery: A comparative study. Technical Report 02-01, Equinox Corporation.

Sharp, G.C., Lee, S.W., and Wehe, D.K. 2002. ICP Registration using invariant features. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24(1):90–102.

Singh, S., Gyaourova, A., Bebis, G., and Pavlidis, I. 2004. Infrared and visible image fusion for face recognition. In *Proc. SPIE Defense and Security Symposium (Biometric Technology for Human Identification)*, pp. 585–596.

Socolinsky, D.A., Selinger, A., and Neuheisel, J.D. 2003. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, 91(1–2):72–114.

Snelick, R., Uludag, U., Mink, A., Indovina, M., and Jain, A. 2005. Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(3):450–455.

Störring, M., Andersen, H.J., and Granum, E. 2001. Physics-based modelling of human skin colour under mixed illuminants. *Journal of Robotics and Autonomous Systems*, 35(3/4):131–142.

Turk, M. and Pentland, A. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):72–86.

Wilder, J., Phillips, P.J., Jiang, C. and Wiener, S. 1996. Comparison of visible and infrared imagery for face recognition. In *Proc. Int. Conf. Automatic Face and Gesture Recognition*, pp. 182–187.

Wolff, L.B., Socolinsky, D.A., and Eveland, C.K. 2001. Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery. In *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum*.

Yang, C., Duraiswami, R., Gumerov, N.A., and Davis, L. 2003. Improved fast Gauss transform and efficient kernel density estimation. In *Proc. 9th Int'l Conf. on Computer Vision*, Nice, France, pp. 464–471.

Yoshitomi, Y., Miyaura, T., Tomita, S., and Kimura, S. 1997. Face identification using thermal image processing. In *Proc. IEEE Int'l Workshop on Robot and Human Communication*, pp. 374–379.

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