# Fusion of Infrared and Visible Images for Face Recognition

Aglika Gyaourova[1], George Bebis[1], and Ioannis Pavlidis[2]

[1] Computer Vision Laboratory, University of Nevada, Reno
{aglika,bebis}@cs.unr.edu
[2] Visual Computing Laboratory, University of Houston
pavlidis@cs.uh.edu

**Abstract.** A number of studies have demonstrated that infrared (IR) imagery offers a promising alternative to visible imagery due to it's insensitive to variations in face appearance caused by illumination changes. IR, however, has other limitations including that it is opaque to glass. The emphasis in this study is on examining the sensitivity of IR imagery to facial occlusion caused by eyeglasses. Our experiments indicate that IR-based recognition performance degrades seriously when eyeglasses are present in the probe image but not in the gallery image and vice versa. To address this serious limitation of IR, we propose fusing the two modalities, exploiting the fact that visible-based recognition is less sensitive to the presence or absence of eyeglasses. Our fusion scheme is pixel-based, operates in the wavelet domain, and employs genetic algorithms (GAs) to decide how to combine IR with visible information. Although our fusion approach was not able to fully discount illumination effects present in the visible images, our experimental results show substantial improvements recognition performance overall, and it deserves further consideration.

## 1 Introduction

Considerable progress has been made in face recognition research over the last decade [1] especially with the development of powerful models of face appearance (e.g., eigenspaces [2]). Despite the variety of approaches and tools studied, however, face recognition has shown to perform satisfactorily in controlled environments but it is not accurate or robust enough to be deployed in uncontrolled environments. Several factors affect face recognition performance including pose variation, facial expression changes, face occlusion, and most importantly, illumination changes.

Previous studies have demonstrated that IR imagery offers a promising alternative to visible imagery for handling variations in face appearance due to illumination changes more successfully. In particular, IR imagery is nearly invariant to changes in ambient illumination [3], and provides a capability for identification under all lighting conditions including total darkness [4]. Thus, while visible-based algorithms opt for pure algorithmic solutions into inherent phenomenology problems, IR-based algorithms have the potential to offer simpler

and more robust solutions, improving performance in uncontrolled environments and deliberate attempts to obscure identity [5].

Despite its advantages, IR imagery has other limitations including that it is opaque to glass. Objects made of glass act as a temperature screen, completely hiding the face parts located behind them. In this study, we examine the sensitivity of IR imagery to facial occlusion due to eyeglasses. To address this serious limitation of IR, we propose fusing IR with visible information in the wavelet domain using GAs [6]. To demonstrate the results of our fusion strategy, we performed extensive recognition experiments using the popular method of eigenfaces [2], although any other recognition method could have been used. Our results show overall substantial improvements in recognition performance using IR and visible imagery fusion than either modality alone.

## 2   Review of Face Recognition in the Infrared Spectrum

An overview of identification in the IR spectrum can be found in [7]. Below, we review several studies comparing the performance of visible and IR based face recognition. The effectiveness of visible versus IR was compared using several recognition algorithms in [8]. Using a database of 101 subjects without glasses, varying facial expression, and allowing minor lighting changes, they concluded that there are no significant performance differences between visible and IR recognition across all the algorithms tested. They also concluded that fusing visible and IR decision metrics represents a viable approach for enhancing face recognition performance. In [9,10], several different face recognition algorithms were tested under various lighting conditions and facial expressions. Using radiometrically calibrated thermal imagery, they reported superior performance for IR-based recognition than visible-based recognition. In [11], the effect of lighting, facial expression, and passage of time between the gallery and probe images were examined. Although IR-based recognition outperformed visible-based recognition assuming lighting and facial expression changes, their experiments demonstrated that IR-based recognition degrades when there is substantial passage of time between the gallery and probe images. Using fusion strategies at the decision level based on ranking and scoring, they were able to develop schemes that outperformed either modality alone. IR has also been used recently in face detection [12]. This approach employs multi-band feature extraction and capitalizes on the unique reflectance characteristics of the human skin in the near-IR spectrum.

## 3   Fusion of Infrared and Visible Imagery

Despite its robustness to illumination changes, IR imagery has several drawbacks. First, it is sensitive to temperature changes in the surrounding environment. Currents of cold or warm air could influence the performance of systems using IR imagery. As a result, IR images should be captured in a controlled environment. Second, it is sensitive to variations in the heat patterns of the face.

Factors that could contribute to these variations include facial expressions (e.g. open mouth), physical conditions (e.g. lack of sleep), and psychological conditions (e.g. fear, stress, excitement). Finally, IR is opaque to glass. As a result, a large part of the face might be occluded (e.g. by wearing eyeglasses).

In contrast to IR imagery, visible imagery is more robust to the above factors. This suggests that effective algorithms to fuse information from both spectra have the potential to improve the state of the art in face recognition. In the past, fusion of visible and IR images has been successfully used for visualization purposes [13].

In this study, we consider the influence of eyeglasses to IR-based face recognition. Our experiments demonstrate that eyeglasses pose a serious problem to recognition performance in the IR spectrum. To remedy this problem, we propose fusing IR with visible imagery. Visible imagery can suffer from highlights on the glasses under certain illumination conditions, but the problems are considerably less severe than with IR. Since IR and visible imagery capture intrinsically different characteristics of the observed faces, intuitively, a better face description could be found by utilizing the complimentary information present in the two spectra.

## 3.1   Fusion 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 with different spatial extend to be fused at the resolution at which they are most salient. In this way, important features appearing at lower resolutions can be preserved in the fusion process.

Multiple resolution features have been used in several face recognition systems in the past (e.g. [14]). The advantages of using different frequencies is that high frequencies are relatively independent of global changes in the illumination, while the low frequencies take into account the spatial relationships among the pixels and are less sensitive to noise and small changes, such as facial expression.

The slow heat transfer through the human body causes natural low resolution of IR images of human face. Thus, we decided to implement our fusion strategy in the wavelet domain, taking into consideration the benefits of multi-resolution representations and the differences in resolution between the IR and visible-light images. Our fusion strategy is thus different from fusion strategies implemented at the decision level, reported earlier in the literature (i.e., [8,11]).

## 3.2   Method Overview

The proposed method contains two major steps: (a) fusion of IR and visible images and (b) recognition based on the fused images. Fusion is performed by combining the coefficients of Haar wavelet [15] decompositions of a pair of IR and visible images having equal size. The fusion strategy is found during a training phase using GAs [6]. The coefficients selected from each spectrum are put together and the fused face image is reconstructed using the inverse wavelet transform. To demonstrate the effectiveness of the fusion solutions found by GAs,

we perform recognition using the popular method of eigenfaces [2] although any other recognition technique could have been used.

# 4   Mathematical Tools and Background Information

## 4.1   Eigenfaces

The eigenface approach uses Principal Components Analysis (PCA), a classical multivariate statistics method, to linearly project face images in a low-dimensional space. This space is spanned by the principal components (i.e., eigenvectors corresponding to the largest eigenvalues) of the distribution of the training images. After a face image has been projected in the eigenspace, a feature vector containing the coefficients of the projection is used to represent the face image. Representing each image $I(x, y)$ as a $N \times N$ vector $\Gamma_i$, first the average face $\Psi$ is computed: $\Psi = \frac{1}{R} \sum_{i=1}^{R} \Gamma_i$ where $R$ is the number of faces in the training set. Next, the difference $\Phi$ of each face from the average face is computed: $\Phi_i = \Gamma_i - \Psi$. Then the covariance matrix is estimated by: $C = \frac{1}{R} \sum_{i=1}^{R} \Phi_i \Phi_i^T = AA^T$, where, $A = [\Phi_1 \Phi_2 \ldots \Phi_R]$. The eigenspace can then be defined by computing the eigenvectors $\mu_i$ of $C$. Usually, we need to keep a smaller number of eigenvectors $R_k$ corresponding to the largest eigenvalues. Each image $\Gamma$ is transformed by first subtracting the mean image ($\Phi = \Gamma - \Psi$), and then projecting in the eigenspace $w_i = \mu_i^T \Gamma$.

## 4.2   Wavelet Transform (WT)

Wavelets are a type of multi-resolution function approximation that allow for the hierarchical decomposition of a signal or image. In particular, they decomposes a given signal onto a family of functions with finite support. This family of functions is constructed by the translations and dilations of a single function called *mother wavelet*. The finite support of the *mother wavelet* gives exact time localization while the scaling allows extraction of different frequency components. The discrete wavelet transform (DWT) is defined in terms of discrete dilations and translations of the *mother wavelet* function: $\psi_{jk}(t) = 2^{-j/2} \psi(2^{-jt} - k)$ , where the scaling factor $j$ and the translation factor $k$ are integers: $j, k \in \mathbf{Z}$. The wavelet decomposition of a function $f(t) \in L^2(R)$ is given by: $f(t) = \sum_j \sum_k h_{j,k} \psi_{jk}(t)$ , where the coefficients $h_{j,k}$ are the inner products of $f(t)$ and $\psi_{jk}(t)$.

## 4.3   Genetic Algorithms (GAs)

GAs are a class of randomized, parallel search optimization procedures inspired by the mechanisms of natural selection, the process of evolution [6]. They were designed to efficiently search large, non-linear, poorly-understood search spaces. In the past, GAs have been used in target recognition [16], object recognition [17], face detection/verification [18,19], and feature selection [20,21].

GAs operate iteratively on a population of structures, each of which represents a candidate solution to the problem, encoded as a string of symbols (i.e.,

chromosome). A randomly generated set of such strings forms the initial population from which the GA starts its search. Three basic genetic operators guide this search: selection, crossover and mutation. Evaluation of each string is based on a fitness function which is problem-dependent. The fitness function determines which of the candidate solutions are better. Selection probabilistically filters out poor solutions and keeps high performance solutions for further investigation. Mutation is a very low probability operator that plays the role of restoring lost genetic material. Crossover in contrast is applied with high probability. It is a randomized yet structured operator that allows information exchange between the stings.

## 5   Evolutionary IR and Visible Image Fusion

Our fusion strategy operates in the wavelet domain. The goal is to find an appropriate way to combine the wavelet coefficients from the IR and visible images. The key question is which wavelet coefficients to choose and how to combine them. Obviously, using un-weighted averages is not appropriate since it assumes that the two spectra are equally important and, even further, that they have the same resolution which is not true. Several experiments for fusing the wavelet coefficients of two images have been reported in [22]. Perhaps, the most intuitive approach is picking the coefficients with maximum absolute value [23]. The higher the absolute value of a coefficient is, the higher is the probability that it encodes salient image features. Our experiments using this approach showed poor performance.

In this paper, we propose using GAs to fuse the wavelet coefficients from the two spectra. Our decision to use GAs for fusion was based on several factors. First, the search space for the image fusion task at hand is very large. In the past, GAs have demonstrated good performance when searching large solution spaces. Much work in the genetic and evolutionary computing communities has led to growing understanding of why they work well and plenty of empirical evidence to support this claim [24,25]. Second, the problem at hand appears to have many suboptimal solutions. Although, GAs cannot guarantee finding a global optimum, they have shown to be successful in finding good local optima. Third, they suitable for parallelization and linear speedups are the norm, not the exception [26]. Finally, we have applied GAs in the past for feature selection, a problem very much related to fusion, with good success [20,21].

**Encoding:** In our encoding scheme, the chromosome is a bit string whose length is determined by the number of wavelet coefficients in the image decomposition. Each bit in the chromosome is associated with a wavelet coefficient at a specific location. The value of a bit in this array determines whether the corresponding wavelet coefficient is selected from the IR (e.g., 0) or from the visible spectrum (e.g., 1).

**Fitness Evaluation:** Each individual in a generation represents a possible way to fuse IR with visible images. To evaluate its effectiveness, we perform the fusion based on the information encoded by this individual and apply the

eigenface approach. Recognition accuracy is computed using a validation dataset (see Section 7) and is used to provide a measure of fitness.

**Initial Population:** In general, the initial population is generated randomly, (e.g., each bit in an individual is set by flipping a coin). In this way, however, we will end up with a population where each individual contains the same number of 1's and 0's on average. To explore subsets of different numbers of wavelet coefficients chosen from each domain, the number of 1's for each individual is generated randomly. Then, the 1's are randomly scattered in the chromosome.

**Selection:** Our selection strategy was cross generational. Assuming a population of size N, the offspring double the size of the population and we select the best N individuals from the combined parent-offspring population

**Crossover:** In general, we do not know how different wavelet coefficients depend on each other. If dependent coefficients are far apart in the chromosome, it is more probable that traditional 1-point crossover, will destroy the schemata. To avoid this problem, uniform crossover is used here. The crossover probability used in our experiments was 0.96.

**Mutation:** Mutation is a very low probability operator which flips the values of randomly chosen bit. The mutation probability used here was 0.02.

## 6  Face Dataset

In our experiments, we used the face database collected by Equinox Corporation under DARPA's HumanID program [27]. Specifically, we used the long-wave infrared (LWIR) (i.e., $8\mu$-$12\mu$) and the corresponding visible spectrum images from this database. The data was collected during a two-day period. Each pair of LWIR and visible light images was taken simultaneously and co-registered with 1/3 pixel accuracy (see Fig. 1). The LWIR images were radiometrically calibrated and stored as grayscale images with 12 bits per pixels. The visible images are also grayscale images represented with 8 bits per pixel. The size of the images in the database is 320×240 pixels.

The database contains frontal faces under the following scenarios: (1) three different light direction - frontal and lateral (right and left); (2) three facial expression - "frown", "surprise" and "smile"; (3) vocals pronunciation expressions - subjects were asked to pronounce several vocals from which three representative frames are chosen; and (4) presence of glasses - for subjects wearing glasses, all of the above scenarios were repeated with and without glasses. Both IR and visible face images were preprocessed prior to experimentation by following a procedure similar to that described in [9,10]. The goal of preprocessing was to align and scale the faces, remove background, and account for some illumination variations (see Fig. 1).

## 7  Experimental Procedure

In this study, we attempted to test the effect on recognition performance of each factor available in the Equinox database. In addition, we have performed

**Fig. 1.** Examples of visible and IR image pairs and preprocessed images

experiments focusing on the effect of eyeglasses. For comparison purposes, we have attempted to evaluate our fusion strategy using a similar experimental protocol to that given in [9,10]. Our evaluation methodology employs a training set (i.e., used to compute the eigenfaces), a gallery set (i.e., set of persons enrolled in the system), a validation set (i.e., used in the fitness evaluation of the GA), and a test set (i.e., probe image set containing the images to be identified). Our training set contains 200 images, randomly chosen from the entire Equinox database.

For recognition, we used the Euclidean distance and the first 100 principal components as in [9,10]. Recognition performance was measured by finding the percentage of the images in the test set, for which the top match is an image of the same person from the gallery. To mitigate for the relatively small number of images in the database, the average error was recorded using a three-fold cross-validation procedure. In particular, we split each dataset used for testing randomly three times by keeping only 75% of the images for testing purposes and the rest 25% for validation purposes. To account for performance variations due to random GA initialization, we averaged the results over three different GA runs for each test, choosing a different random seed each time. Thus, we performed a total of 9 runs for each gallery/test set experiment.

### 7.1   Facial Expression Tests

The test sets for the facial expression experiments include the images containing the three expression frames and three vocal pronunciation frames. There are 90 subjects with a total of 1266 pairs of images for the expression frames and 1299 for the vocal frames. Some of the subjects in these tests sets wear glasses while others not. Following the terminology in [9,10] we have created the following test sets: EA (expression frames, all illuminations), EL (expression frames, lateral illuminations), EF (expression frames, frontal illumination), VA (vocal frames, all illumination), VL (vocal frames, lateral illumination), VF (vocal frames, frontal illumination). The inclusion relations among these sets are as follows: EA = EL ∪ EF, VA = VL ∪ VF, and VA ∩ EA = ∅.

### 7.2   Eyeglasses Tests

Measuring the effect of eyeglasses is done by using the expression frames. There are 43 subjects wearing glasses in the EA set making a total of 822 images. Following the terminology in [9,10] we created the following test sets: EG (expression

frames with glasses, all illuminations), EnG (expression frames without glasses, all illuminations), EFG (expression frames with glasses, frontal illumination), ELG (expression frames with glasses, lateral illumination), EFnG (expression frames without glasses, frontal illumination), ELnG (expression frames without glasses, lateral illumination). The inclusion relations among these sets are as follows: EG = ELG ∪ EFG, EnG = ELnG ∪ EFnG and EG ∩ EnG = ∅.

# 8  Experimental Results

## 8.1  Eyeglasses

The results shown in Table 1 illustrate that IR-based recognition is robust to illumination changes but performs poorly when glasses are present in the gallery set but not in the test set and vice versa. Considerable improvements in recognition performance have been achieved in this case by fusing IR with visible images. The improvement was even greater when, in addition to eyeglasses, the test and the gallery set contained images taken under different illuminations. For example, in the EFG/ELnG test case the fusion approach improved recognition performance by 46% compared to recognition using visible-light images and by 82% compared to recognition using LWIR images.

Recognition using LWIR images outperformed recognition using fused images when the only difference between the images in the test and gallery sets was the direction of illumination. This is accounted to the inability of our fusion scheme to fully discount illumination effects contributed by the visible-light images. Recognition performance using visible-light images was always worse than using fused images.

## 8.2  Facial Expression

The facial expression tests had varying success as shown in Table 2. In general, fusion led to improved recognition compared to recognition using visible-light images. In several cases, however, the accuracy using LWIR images was higher than using fused images. These were cases again where the illumination directions between the gallery and the test sets were different. This result is consistent with that of the eyeglasses tests and was caused by the inability of our fusion scheme to fully discount the illumination effects in the visible images. Note that we did not performed experiments when the intersection between gallery and test sets is not empty.

# 9  Discussion

The presence/absence of eyeglasses proved to be a big obstacle for IR-based recognition. To better understand this, let's take a closer look of the results shown in Table 1. The horizontal and vertical double lines through the center of the table divide the table into four quadrants (i.e., $I$ to $IV$, starting from

**Table 1.** Averages and standard deviations for the eyeglasses experiments. The columns represent the gallery set and the rows represent the test set. The first entry in each cell shows the performance measured from the visible-light images, the second entry is from the LWIR images, and the third entry is from the fused images. The bottom entry shows the minimum and maximum recognition performances from the three cross-validation runs achieved when using the fused images. Test scenarios for which the test and the gallery sets had common subsets were not performed.

| | EG | ELG | EFG | EnG | ELnG | EFnG |
|---|---|---|---|---|---|---|
| EG | ✕ | ✕ | ✕ | (84.8, 1.4)(15.1, 1.0)(92.5, 1.3) 91.0 − 93.6 | (84.8, 1.4)(13.1, 1.0)(88.9, 1.4) 88.0 − 90.5 | (64.3, 1.7)(21.7, 1.0)(82.1, 3.1) 79.2 − 85.4 |
| ELG | ✕ | ✕ | (71.4, 1.0)(99.6, 0.3)(93.2, 3.0) 90.2 − 94.2 | (85.8, 0.7)(16.2, 0.3)(92.3, 2.0) 90.2 − 94.2 | (85.8, 0.7)(14.2, 0.4)(92.7, 0.6) 92.0 − 93.0 | (56.0, 1.2)(22.4, 0.4)(83.9, 1.3) 82.5 − 84.8 |
| EFG | ✕ | (78.3, 1.1)(100, 0)(97.9, 0.6) 97.4 − 98.5 | ✕ | (83.7, 3.9)(14.5, 0.6)(91.7, 1.3) 90.9 − 93.1 | (50.7, 1.7)(13.0, 0)(77.1, 3.1) 74.3 − 80.4 | (82.6, 4.0)(22.1, 0.6)(92.2, 2.0) 90.2 − 94.2 |
| EnG | (79.4, 1.1)(2.6, 0.2)(84.9, 0.8) 84.3 − 85.8 | (72.0, 2.2)(2.4, 0.2)(81.6, 0.3) 81.2 − 81.8 | (60.2, 2.0)(17.4, 0.5)(98.0, 0.1) 97.9 − 98.1 | ✕ | ✕ | ✕ |
| ELnG | (82.0, 1.7)(2.7, 0.5)(84.2, 1.9) 82.1 − 85.7 | (82.0, 1.7)(2.5, 0.3)(84.1, 1.4) 82.5 − 85.2 | (52.6, 0.8)(17.7, 0.6)(96.7, 1.4) 95.1 − 97.8 | ✕ | ✕ | (73.0, 1.9)(98.4, 0.6)(96.9, 0.6) 96.2 − 97.5 |
| EFnG | (78.6, 1.6)(2.1, 0)(85.6, 3.2) 83.2 − 89.2 | (56.8, 1.1)(2.1, 0)(80.7, 2.9) 78.3 − 83.9 | (78.6, 1.6)(18.6, 1.2)(87.0, 2.1) 85.2 − 89.3 | ✕ | (71.9, 1.2)(100, 0)(97.8, 1.0) 97.2 − 99.0 | ✕ |
| number of images | 398 | 260 | 138 | 424 | 281 | 143 |

**Table 2.** Recognition results for the facial expression tests.

| | VA | EA | VF | EF | VL | EL |
|---|---|---|---|---|---|---|
| VA | ✕ | (65.0, 1.0)(92.3, 0.7)(93.6, 0.7) 92.8 − 94.2 | ✕ | ✕ | ✕ | ✕ |
| EA | (65.2, 1.2)(93.7, 0.4)(93.1, 0.4) 92.7 − 93.4 | ✕ | ✕ | ✕ | ✕ | ✕ |
| VF | ✕ | ✕ | ✕ | (65.3, 1.8)(90.5, 0.4)(93.3, 0.9) 92.5 − 94.2 | (62.4, 0.4)(99.3, 0)(98.1, 0.5) 97.7 − 98.6 | (36.4, 0.7)(87.9, 0.6)(91.4, 1.4) 89.8 − 92.3 |
| EF | ✕ | ✕ | (59.2, 0.9)(91.3, 0.7)(91.2, 1.3) 90.3 − 92.8 | ✕ | (30.4, 0.7)(91.2, 0.4)(90.0, 1.4) 88.6 − 91.5 | (68.5, 1.1)(99.5, 0.4)(98.2, 0.4) 97.7 − 98.4 |
| VL | ✕ | ✕ | (60.5, 0.7)(98.4, 0.2)(97.6, 0.8) 97.0 − 98.4 | (28.0, 1.2)(90.4, 0.1)(86.9, 0.8) 85.9 − 87.4 | ✕ | (66.4, 0.4)(91.3, 0.4)(94.5, 1.5) 92.9 − 95.7 |
| EL | ✕ | ✕ | (31.6, 1.1)(85.8, 0.7)(84.4, 0.3) 84.2 − 84.7 | (63.0, 0.9)(98.6, 0.2)(96.4, 0.7) 95.8 − 97.2 | (68.9, 1.0)(91.7, 0.2)(93.7, 0.4) 93.3 − 94.1 | ✕ |
| number of images | 1299 | 1266 | 435 | 429 | 864 | 837 |
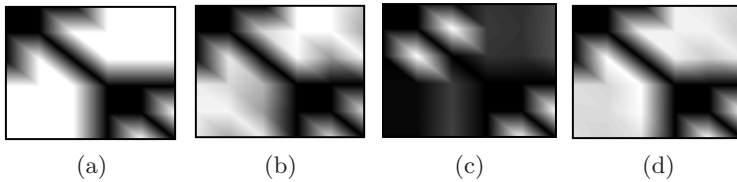
**Fig. 2.** The average performance values from Table 1, visualized as a grayscale image. See text for details. (a) ideal case (b) visible images (c) IR images, (d) fused images.

the upper-right corner and moving counterclockwise). Each quadrant represents a set of experiments testing some specific difference between the gallery and the test sets: (1) Experiments in quadrant $I$ evaluate the effect of eyeglasses being present in the probe but not in the gallery; (2) Experiments in quadrant $III$ evaluate the effect of eyeglasses being present in the gallery but not in the probe; (3) Experiments along the off-diagonals within each of these two quadrants represent tests where the illumination conditions between the gallery and probe sets are the same; (4) Experiments in quadrants $II$ and $IV$ evaluate the effect illumination changes only.

To illustrate the performance of our fusion approach, we have interpolated the results from Table 1 and used a simple visualization scheme to remove small differences and emphasize major trends in recognition performance (see Fig. 2). Our visualization scheme assigns a grayscale value to each average from Table 1) with black implying 0% recognition and white 100% recognition. The empty cells from Table 1 are also shown in black.

By observing Fig. 2, several interesting conclusions can be made. As expected, face recognition success based on IR images (see Fig. 2.(b))is not influenced by lighting conditions. This is supported by the prevailing white color in quadrants $II$ and $IV$ (case (3)) and by the high recognition rates in quadrants $II$ and $IV$ (case (4)). However, IR yielded very low success when eyeglasses were present in the gallery but not in the probe and vice-versa (cases (1) and (2)). The success of visible-based face recognition was relatively insensitive to subjects' wearing glasses (see Fig. 2.(c)). This follows from the relatively uniform color in quadrants $I$ and $III$ (cases (1) and (2)). Lighting conditions had big influence on the success of face recognition in the visible domain. There are distinguishable bright lines along the main diagonals in quadrants $I$ and $II$ (case (3)). The success of face recognition based on fused images was similar in all four quadrants of the image (see Fig. 2.(d)). This implies that we were able to achieve relative insensitivity to both eyeglasses and variable illumination.

The image fusion approach led to higher recognition performance compared to recognition in the visible spectrum but was not able to completely compensate for the effects of illumination direction in the visible images. We have noticed that in all the cases where LWIR performed better than fusion, the illumination direction in the gallery set was different from that in the test set (assuming no difference in glasses). The presence of illumination effects in the fused images

|     |     |     |     |
|:---:|:---:|:---:|:---:|
| (a) | (b) | (c) | (d) |
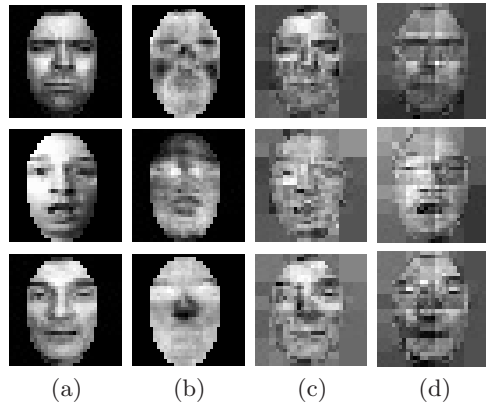
**Fig. 3.** The original (a) visible and (b) IR images followed by two different fused image results – (c) trained on a data set with lateral illumination and without glasses and (d) trained on a data set with glasses.
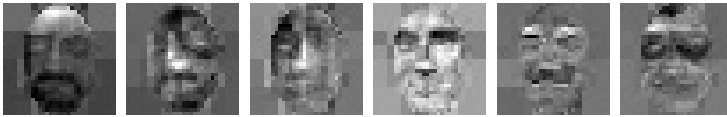


**Fig. 4.** The first few eigenfaces of a fused image data set. The second and third eigenfaces show clear influence of the right and left lateral illumination.

can be visually confirmed by observing the reconstructed fused images shown in Fig. 3, and their first eigenfaces shown in i.e., Fig. 4. Fused images had higher resolution compared to LWIR images, however, they were also affected by illumination effects present in the visible images. Obviously, the first eigenfaces of the fused images still encode the effects of illumination direction, present in the visible images. More effective fusion schemes (e.g., weighted averages of wavelet coefficients) and more powerful fitness functions (i.e., add extra terms to control the number of coefficients selected from different bands of each spectrum) might help to overcome these problems and improve fusion overall.

Also, further consideration should be given to the existence of many optimal solutions found by the GA. Although optimal in the training phase, these solutions showed different recognition performances when used for testing. In investigating these solutions, we were not able to distinguish any pattern in the content of the chromosomes that might have revealed why some chromosomes were better than others. On the average, half of the coefficients were selected from the visible spectrum and the other half from the IR spectrum. The use of larger validation sets and more selective fitness functions might help to address these issues more effectively.

## 10   Conclusions and Future Work

We presented a fusion method for combining IR and visible light images for the purposes of face recognition. The algorithm aims at improved and robust recognition performance across variable lighting, facial expression, and presences of eyeglasses. Future work includes addressing the issues mentioned in the previous section, considering fitness approximation schemes [28] to reduce the computational requirements of fitness evaluation, and investigating the effect of environmental (e.g., temperature changes), physical (e.g., lack of sleep) and physiological conditions (e.g., fear, stress) to IR performance.

## References

1. Zhao, W., C.R.P.P., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Surveys **35** (2003) 399–458
2. Turk, M., Pentland, A.: Eigenfaces for recognition. CogNeuro **3** (1991) 71–96
3. Wolff L., S.D., , C., E.: Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methodsand Applications, Hawaii (2001)
4. Jain, A., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society. Kluwer Academic Publishers (1999)
5. Pavlidis, I., Symosek, P.: The imaging issue in an automatic face/disguise detection system. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications. (2000) 15–24
6. Goldberg, D.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley (1989)
7. Prokoski, F.: History, current status, and future of infrared identification. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum, Hilton Head (2000)
8. Wilder, J., Phillips, J., Jiang, C., Wiener, S.: Comparison of visible and infra-red imagery for face recognition. In: 2nd International Conference on Automatic Face and Gesture Recognition, Killington (1996) 182–187
9. Selinger, A., Socolinsky, D.: Appearance-based facial recognition using visible and thermal imagery: A comparative study. Technical report, Equinox Corporation n.02–01 (2002)
10. Socolinsky, D., A., S.: Comparative study of face recognition performance with visible and thermal infrared imagery. In: International Conference on Pattern Recognition. (2002) 217–222
11. Chen, X., Flynn, P., Bowyer, K.: Pca-based face recognition in infrared imagery: Baseline and comparative studies. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, France (2003)
12. Dowdall, J., Pavlidis, I., Bebis, G.: Face detection in the near-ir spectrum. Image and Vision Computing **21** (2001) 565–578
13. Scheunders, P.: Local mapping for multispectral image visualization. Image and Vision Computing **19** (2001) 971–978
14. Manjunath, B., Chellappa, R., von der Malsburg, C.: A feature based approach to face recognition. In: Computer Vision and Pattern Recognition. (1992) 373–378
15. Chui, C.: An introduction to wavelets. Academic Press (1992)
16. Katz, A., Thrift, P.: Generating image filters for target recognition by genetic learning. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994)

17. Bebis, G., Louis, S., Varol, Y., Yfantis, A.: Genetic object recognition using combinations of views. IEEE Transactions on Evolutionary Computing **6** (2002) 132–146
18. Swets, D., Punch, B.: Genetic algorithms for object localization in a complex scene. In: IEEE International Conference on Image Processing. (1995) 595–598
19. Bebis, G., Uthiram, S., Georgiopoulos, M.: Face detection and verification using genetic search. International Journal of Artificial Intelligence Tools 9 (2000) 225–246
20. Sun, Z., Yuan, X., Bebis, G., Louis, S.: Genetic feature subset selection for gender classification: A comparison study. In: IEEE Workshop on Applications of Computer Vision. (2002)
21. Sun, Z., Bebis, G., Miller, R.: Boosting object detection using feature selection. In: IEEE International Conference on Advanced Video and Signal Based Surveillance. (2003)
22. Chipman, L., Orr, T.: Wavelets and image fusion. In: IEEE International Conference on Image Processing. Volume 3. (1995) 248–251
23. Li, H.and Manjunath, B., Mitra, S.: Multisensor image fusion using the wavelet transform. In: IEEE International Conference on Image Processing. Volume 1., Austin, Texas (1994) 51–55
24. Fogel, D.: Evolutionary Computation, Toward a New Philosophy of Machine Intelligence. IEEE Press (1992)
25. Koza, J.: Genetic Programming. MIT Press (1993)
26. Cantu-Paz, E.: Efficient and Accurate Parallel Genetic Algorithms and Evolutionary Computation. Kluwer Academic Publishers (2000)
27. Available at: http://www.equinoxsensors.com/products/HID.html.
28. Jin, Y., Sendhoff, B.: Fitness approximation in evolutionary computation – a survey. In: Genetic and Evolutionary Computation Conference. (2002)