

RESEARCH

Open Access

Adaptive polar transform and fusion for human face image processing and evaluation

Debotosh Bhattacharjee

Correspondence: debotosh@ieee.org
Department of Computer Science
and Engineering, Jadavpur
University, Kolkata 700032, India

Abstract

Human face processing and evaluation is a problem due to variations in orientation, size, illumination, expression, and disguise. The goal of this work is threefold. First, we aim to show that the variant of polar transformation can be used to register face images against changes in pose and size. Second, implementation of fusion of thermal and visual face images in the wavelet domain to handle illumination and disguise and third, principal component analysis is applied in order to tackle changes due to expressions up to a particular extent of degrees. Finally, a multilayer perceptron has been used to classify the face image. Several techniques have been implemented here to depict an idea about improvement of results. Methods started from the simplest design, without registration; only combination of PCA and MLP as a method for dimensionality reduction and classification respectively to the range of adaptive polar registration, fusion in wavelet transform domain and final classification using MLP. A consistent increase in recognition performance has been observed. Experiments were conducted on two separate databases and results yielded are very much satisfactory for adaptive polar registration along with fusion of thermal and visual images in the wavelet domain.

Introduction

Due to enhancement in accuracy, face recognition has gradually increased its recognition as a biometric trait for identification and authentication. Biometric security system is in active research areas for more than last four decades but till date a tractable, robust, and low cost solution is yet to be produced. Complications and difficulties arise in designing such a system lie heavily due to the requirement of unconstrained face recognition. Some constrained face recognition systems may be created specific to applications like monitoring daily attendance, recording frequency of visits for known personalities, or identity checking at non-critical official dealings, with ease, but to maintain security and surveillance e.g. to counter terrorism if some suspect is to be restricted to the protected area like an airport, unconstrained face recognition becomes a necessity. Many techniques have already been developed to tackle sources of different complications like, changes in illumination level and direction [1,2], variation in pose [3,4], changes in expression [5]; changes in skin colour [6], disguises due to cosmetics [7], glasses [8,9], skin colour [6], beard, moustaches [7] etc. These complications are multifaceted and it becomes deterrent to achieve better

recognition performance when two or more complicacies are bundled together, some of them are given below:

- i. Different illumination levels
- ii. Direction of illumination may vary
- iii. Distance of camera from the face may vary and, as a result, effective facial area is bound to be different. In this case if it is possible to crop the facial area still scaling of cropped images is necessary
- iv. Rotation of head about one or more axes in 3-D space leads to significant differences in face images
- v. Presence or absence of beard and/or moustache
- vi. Presence or absence of spectacles and/or coloured-glasses, adornments
- vii. Deliberate change in colour of skin and/or hair to appear before the designed system in disguise.
- viii. Change in expressions and
- ix. Others

In order to tackle all these detrimental factors let us take them into account one by one.

To handle different pose and sizes all the images of each and every person are to be registered against a frontal and neutral face design of that person. Illumination and disguise related problems are dealt with fusion of visual and its corresponding infrared or thermal face images. Finally, Principal Component Analysis (PCA) is used to manage the results of beard, moustache, glasses, and expressions. However, with all these tools in action, classification job is not necessarily a trivial one because all those tools are capable of handling these ill effects to some extent whereas they are not safe enough to abolish those unwanted circumstances. Therefore, like other complex pattern classification tasks, here also, an efficient classifier is needed for final acceptance. Here, one multilayer perceptron (MLP) has been chosen because of its simplicity.

Face recognition is an active research area and researchers are still trying different modalities to increase accuracy for recognition in an unconstrained environment. In this category one, inclusion is thermal infrared imagery [9], and another very recent addition is a using of fusion procedure [9] over different types of images.

Infrared (IR) or thermal images are considered a viable alternative to manage changes in illumination level and in detecting disguised faces [10-12]. Thermal or IR cameras capture images based on the heat patterns emitted from an object. Heat patterns emitted by an object depend on the body temperature and characteristics of the constituent material of the object. Since the blood vessel pattern, muscle, tissue etc. are different for each person, the radiated pattern of heat should be unique for each and every individual. The use of thermal face images has great advantages over visual images. Face recognition systems based on visual images results very poor under different illumination conditions, colours, disguises, and typical conditions like identical twins. However, these situations can be handled by IR images very easily but it has many drawbacks. Firstly, IR imagery depends on the temperature, and if there is a large difference in body temperature then the heat patterns generated by human body would certainly differ. Heat pattern produced by a person sitting in an air-conditioned room and the same person when sitting under strong sunlight must be different. As a matter of fact, IR images should be captured

in a controlled environment i.e. with minimum variation in external temperature. Second, IR images are also sensitive to variations in the internal temperature of the face. Factors that could contribute to these variations include diseases (e.g. cold, fever etc.), facial expressions (e.g. open mouth, closed eyes), physical conditions (e.g. lack of sleeps), and psychological conditions (e.g. fear, stress, excitement). Finally, IR is opaque to spectacles and coloured glasses. This may introduce partial occlusion to the face images.

It is apparent from the above discussion; neither thermal images nor visual images are capable of tackling complications of face recognition, and therefore a combination of both the imaging modalities namely fusion of images have come up. Fusion is actually a natural mechanism built in man and other mammals. It serves as perceiving the real world by the simultaneous use of various sensing modalities [13]. The principal motivation for the fusion approach is to exploit the benefits of two or more modalities in one hand and simultaneously suppressing disadvantages of those on the other.

According to the method and data sources, image fusion techniques can be grouped into [14] following categories:

- (i) Multiview fusion: images from the same source and taken at the same time but from different viewpoints.
- (ii) Multimodal fusion: images coming from various sources like visual, thermal, X-ray etc.
- (iii) Multitemporal fusion: images taken at different times
- (iv) Multifocus fusion: images taken with different focal lengths.

For all these different categories, the fusion comprises of two primary steps: (i) image registration, which is nothing but spatial alignment of input images, and (ii) combining those aligned images.

Due to the rotation, tilting, and panning of head, it is difficult to match face images efficiently and the situation worsens when images are taken at different time intervals. Also, due to differences in distance from the camera to the source face, there may be a difference in dimension which is not so easy to ignore. Any unknown image should be registered against the probe images (neutral and frontal face image) stored in the images database. Many works have been made in this area in the past twenty years [15]. Recently Zokai and Wolberg [16] proposed an innovative design by using Log-Polar transform (LPT). LPT [16,17], is well known tool for image processing for its rotation and scale invariant properties. Scale and rotation in Cartesian coordinate appears as a translation in the log-polar domain. These invariant properties provide a significant benefit in registering images. Log-polar transformation utilizes the feature of applying greater weights to pixels at the centre of the interpolation region and logarithmically decreasing weights to pixels away from the centre.

In this paper, two different face recognition procedures have been performed, and their respective results are compared with each other. The general steps for each process are image registration, fusion of images in the wavelet domain, dimensionality reduction using PCA, and finally, classification of projected images using a multilayer perceptron neural network (MLPNN).

This paper deals with recognition of human face images in semi-uncontrolled environment. Here, semi-uncontrolled term has been used because uncontrolled environment may be any environment like wild where pose variation may be around 90 degrees, almost invisible face under overexposed or underexposed condition, very low resolution

etc. Main contribution of this paper is to handle face recognition problem in moderate conditions of pose, illumination, disguise, and occlusions termed as semi-uncontrolled environment. Some techniques in literature exists where authors targeted one or two sources of complications but in the system, developed here, a comprehensive technique was developed, which is capable to handle combination of different ailments generally present in most of the face recognition systems.

The organization of the rest of this paper is as follows. In Overview of the present system section, the overview of the system is discussed. In Image registration section, details of image registration have been discussed. Fusion of visual and thermal images section describes the fusion of visual and thermal face images in detail. Principal component analysis section and Multilayer perceptron neural network section describe PCA and MLP respectively in brief. In Experimental results and discussions section, experimental results and discussions are given. Finally, Conclusion section concludes this work.

Overview of the present system

In the following sections, we present the techniques which form the elements of our system, shown in Figure 1 and which also describe our motivation for using them. Briefly, we explore the use of polar transform to register images against rotation and scale invariant face images, fusion of face images is made to achieve illumination invariance, principal component analysis to incorporate expression changes to some degree and ultimately, a multilayer perceptron is used for classification of images.

Image registration

Image registration is a means of finding correspondence between two images depicting common visual information, like, images of the same object taken at different conditions like changes in illumination levels and illumination direction, changes in the environment; many geometrical position e.g. orientations about X-, Y-, and Z-axes; with a difference in a time interval; and considered by several sensors, by which there may be a difference in resolution, intensity etc. In general, image registration works in four steps.

- (i) Feature detection: Salient and distinctive features like corners, intersection and end points of lines, regions, edges, closed contours etc. are automatically detected. These features are described by critical data structures and termed as control points to establish a correspondence between images brought under registration.
- (ii) Feature matching: Once control points are computed; correspondences between them are found through finding similarity between particular features computed.
- (iii) Transform model estimation: Depending on corresponding control points, estimation of transformation is done. This estimation has to find the possible transformation functions and their respective parameters, so that once applied to transformed image; it becomes closest to the original one.
- (iv) Image transformation: Final job is to transform the image as per transformation model discussed above. Once transformation is completed it may require resampling of images with appropriate brightness interpolation for pixels represented by non-integer coordinates.

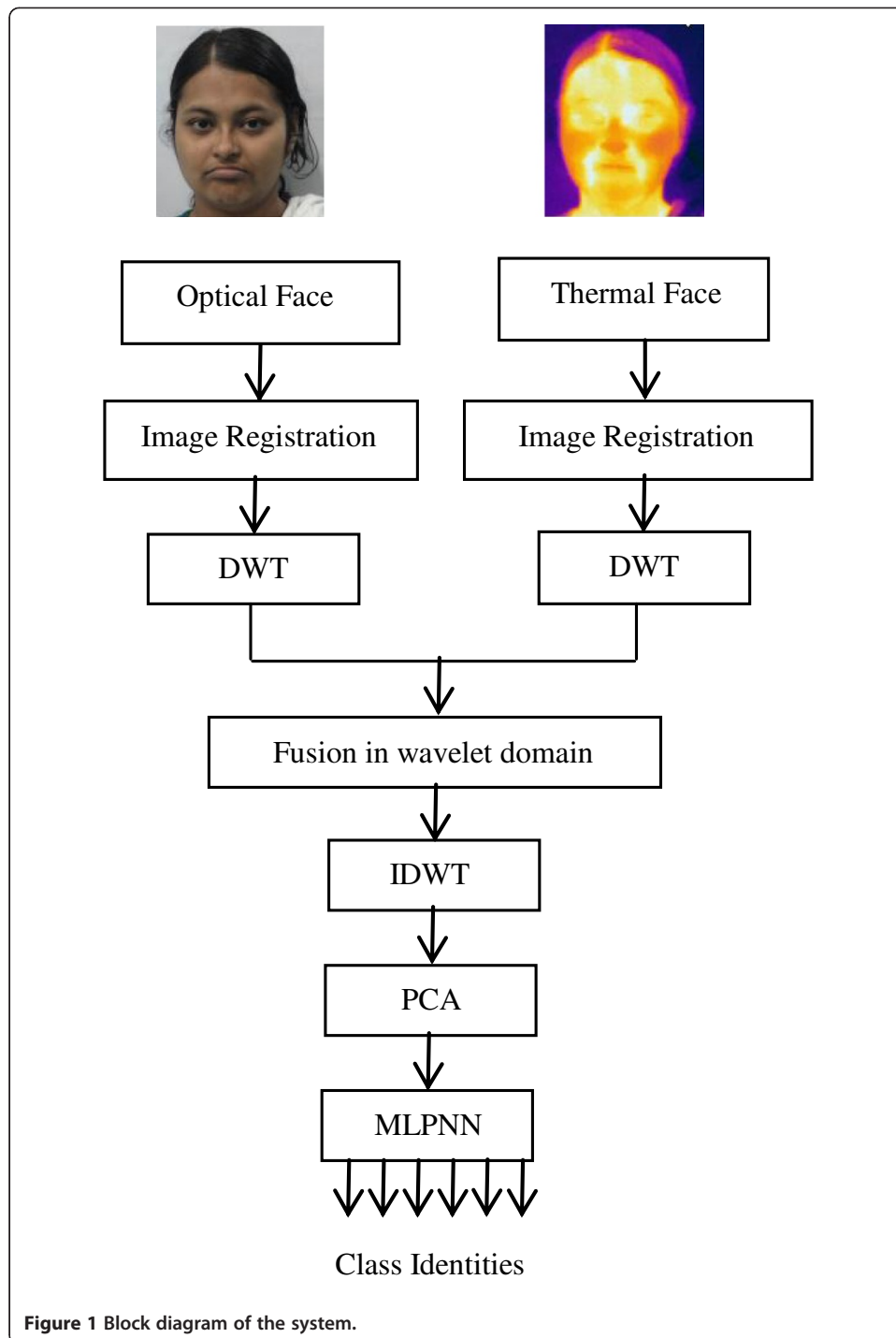


Image registration plays an important role in the fusion of images captured with different cameras with different features like viewing angle, resolution and focus.

Log polar transform

Log-Polar transform (LPT) is a process that converts an image described in Cartesian coordinate $f(x, y)$ into the log-polar coordinate system $s(r, \theta)$. Image represented in

log-polar coordinates are known for rotation and scale invariant properties. The transformation can be given as

$$r = \log_{10} \sqrt{(x-x_c)^2 + (y-y_c)^2}$$
$$\theta = \tan^{-1} \frac{y-y_c}{x-x_c}$$

where, (x_c, y_c) is the centre pixel of the transformation in the Cartesian coordinates.

Figure 2 shows an image in 2(a) and its corresponding log-polar transformed image in 2(d). Moreover, 2(b) and 2(c) are rotated images of the image in 2(a) and 2(e), 2(f) are log-polar transformed images of 2(b) and 2(c) respectively.

It is evident from Figure 2 is that the rotation in Cartesian coordinate system is mere shifting in log-polar coordinate system. Since, in case of shifting, detection is easier, and retranslation to its earlier state back can be done afterwards, log-polar transformation has been used in image registration very efficiently. The benefit of log-polar coordinate system over the Cartesian coordinate representation is that any rotation and (or) scale in the Cartesian coordinates are (is) represented as shifts in the angular and log-radius directions, respectively. Since log-polar transform uses non-uniform sampling, it becomes very efficient in object recognition. If the central message content becomes the centre of the transform then for that part more samples are taken, whereas for background etc., which are in the distant locations from the centre could influence very limited due to consideration of the lesser number of samples for those.

Adaptive polar transform

In some cases, the benefit of non-uniform sampling becomes disadvantageous. For example, in case of human face recognition, generally cropped images are considered and therefore the importance of all the parts should be given uniformly. In log polar transformation, the

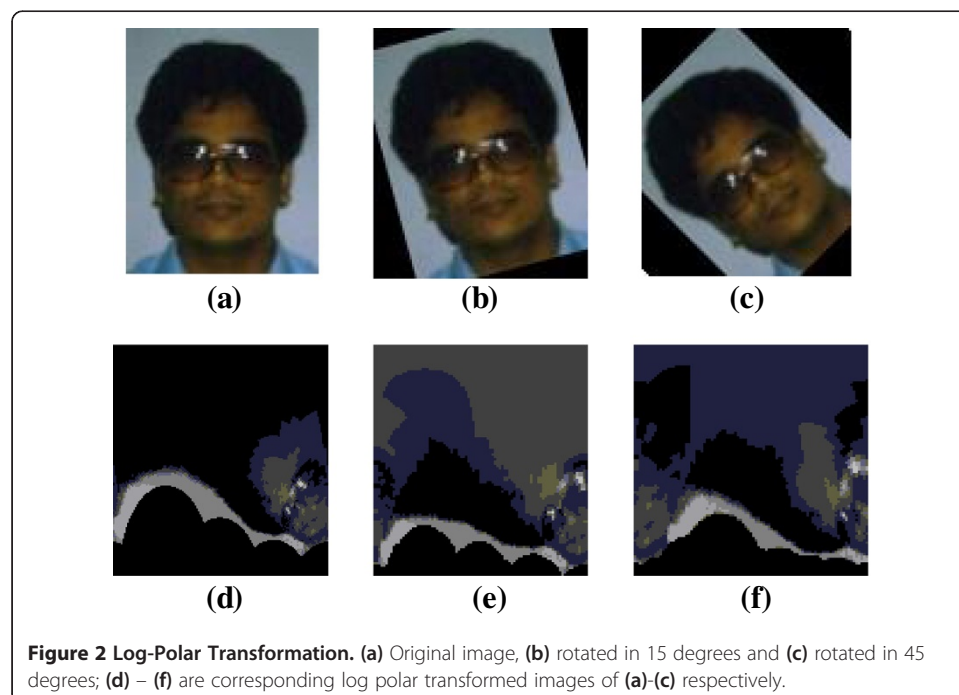


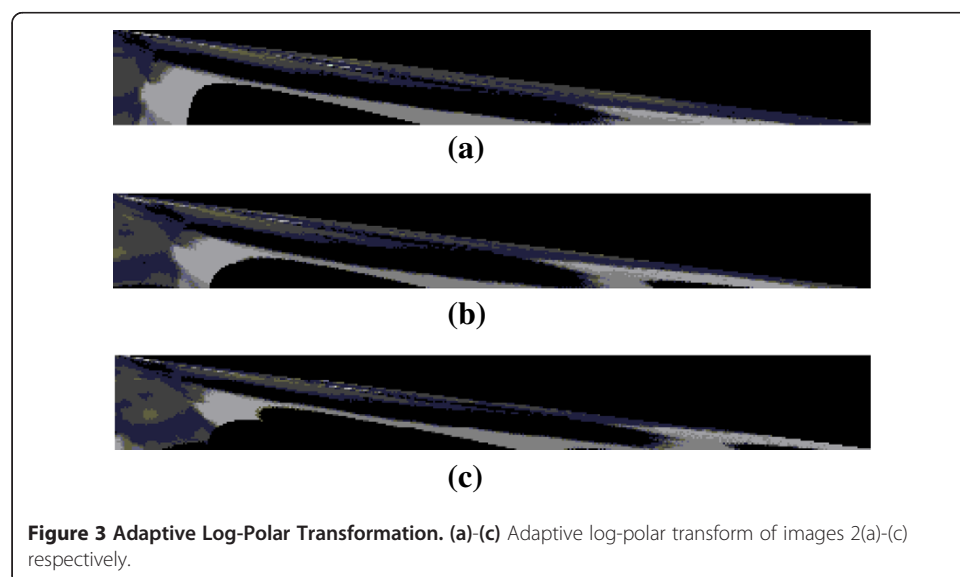
Figure 2 Log-Polar Transformation. (a) Original image, (b) rotated in 15 degrees and (c) rotated in 45 degrees; (d) – (f) are corresponding log polar transformed images of (a)-(c) respectively.

pixels which were away from centre point may be missed, and as a result, the part of pixels in the periphery in recognition becomes negligible and therefore the accuracy should degrade. To get rid of this disadvantage Matungka et al. [18] proposed a method called adaptive polar transformation. In adaptive polar transformation number of samples near centre and those at the periphery differ. It increases directly proportionally to the increase in radius.

In order to achieve consistent sampling over the circumference (C_i) at every sample of radius (r_i), the i^{th} circular sample from centre through radius, number of samples in an angular direction (θ_i) should be adaptive i.e. θ_i should increase along with the increase in C_i . The circumference C_i at radius r_i is known to be $C_i = 2 \times \pi \times r_i$. For an adaptive polar transform of an image, $f(x, y)$, of size $2R_{\max} \times 2R_{\max}$ in Cartesian coordinate system into polar domain (with uniform samples over the circumference), $f_{\text{apt}}(x, y)$, is given as

```
for  $i = 1$  to  $R_{\max}$   
begin  
 $\theta = 2 \times \pi \times i$   
for  $j = 1$  to  $\theta$   
begin  
 $f_{\text{apt}}(x, y) = (f(R_{\max} + i \times \cos(2\pi j/\theta), R_{\max} + i \times \sin(2\pi j/\theta))$   
end for  
end for
```

Adaptive polar transforms of Figure 2(a)-(c) are given in Figure 3(a)-(c) respectively. Registration of these images is made in the adaptive polar transformed domain using phase correlation because it is very easy to find the number of change by phase correlation and registration can be made by realignment opposite to shifting in transformed domain.



Phase correlation

Shifting theorem states that Fourier Transform of a function $f(x+\alpha)$ i.e. shifting of $f(x)$ by an amount α can be obtained by multiplying Fourier Transform of $f(x)$ by $e^{-j\alpha x}$. This may be extended to 2-D (two-dimension) i.e. if $f(x, y)$ is to be shifted by (α, β) then the shifted function would be

$$f_1(x, y) = f(x + \alpha, y + \beta)$$

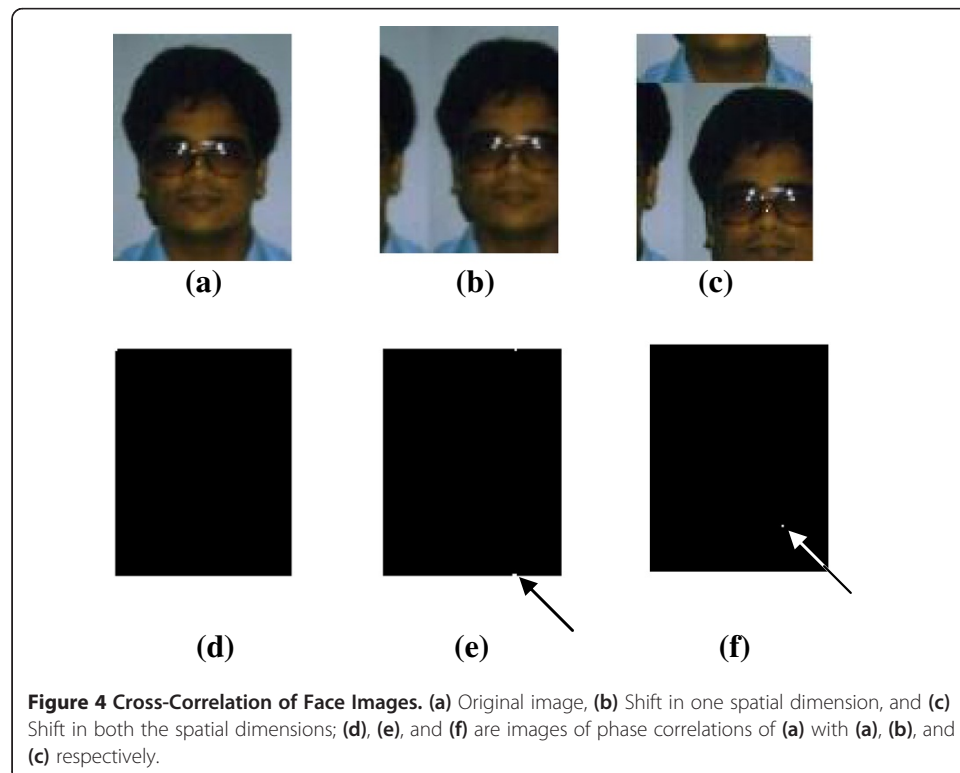
and its corresponding Fourier Transform is

$$F_1(u, v) = e^{-j(\alpha u + \beta v)} F(u, v)$$

where $F(u, v)$ is the Fourier transformed image of $f(x, y)$. That means translation between two images in the spatial domain can be described as a phase difference in the frequency domain.

Cross correlation can be computed by the normalized multiplication in the frequency domain between the first image and the complex conjugate of the second image. To represent this phase difference as translation in the spatial domain, we apply the 2D inverse Fourier transform to it. The peak value of this inverse Fourier transform indicates the translation between the two images.

Figure 4 shows the use of cross-correlation in finding shifting of images. Figure 4(a), (b), and (c) show original image, change in the horizontal direction, and change in both horizontal and vertical direction respectively. Phase correlations of the original image given in Figure 4(a) with all these three images are shown in Figure 4(d)-(f) respectively. Figure 4(d) shows complete black representing no translation. Figure 4(e) shows translation in the horizontal direction and the number of translation is pointed by black arrow. Figure 4(f) shows shifting in both vertical and horizontal direction, and the number



of translation is pointed by white arrow. Once the number of shifting is known then the images can be translated back to its original state. Since a rotation and scaling in Cartesian coordinate is represented in Polar coordinate a shifting in horizontal and vertical directions then retranslation of images in polar coordinate would produce scale and rotation invariant images.

Fusion of visual and thermal images

The fusion of images can be described as a means by which different images acquired through different sensors or modalities are combined to produce a new image with necessary and complementary information from the input images and capable of resolving inconsistencies or ambiguities encountered by classifiers used for classification of face images.

Human face recognition based on visual images has shown poor performance under uncontrolled working conditions. Among others, handling of larger variations in illumination level and disguises are very cumbersome. Whereas thermal face images are invariant towards variations in illumination levels as because those images are representations of the map for blood vessels. Other kind of visual impairments like disguise etc. can also be managed by thermal imaging. But in thermal images apart from blood vessels all other parts in the image are alike. As a matter of fact, two images of different persons may appear to be same by considering this similarity. Therefore to take the advantages of both the visual and thermal images combination of both the image types have been performed and the name of this blend comes under the field of image fusion.

Image fusion techniques

In image fusion, two or more images are combined, and a finally a single fused image is obtained which retains the necessary features from each of the original images. This image fusion is a lossy process because none of the original image can be reconstructed from a fused image. As discussed in Introduction, there could be many fusion situations. This existing design is a multi-modal fusion, where images from two different sources namely visual and thermal are fused. In [19] pixel level fusion method has been applied. In that work, it has been assumed that each face is represented by a pair of images, one in the IR spectrum, and another is in the visible spectrum. For a particular imaging condition, both the thermal and visual images have been combined a priori to get corresponding fused images. During combining, 70% of a pixel in the visual image is added to 30% of the corresponding pixel in thermal images subject to a maximum. The best result reported is 95.71% of recognition accuracy on IRIS face database [20], but that idea didn't consider the entire IRIS face database. When it comes to the full database the performance degrades drastically, and the present method is an extension to that effort to acquire better recognition result for the whole face database. The purpose of such drastic reduction is quite logical from the fusion point of view. Since, in [19] pixel level fusion was considered, that should not work for the images with a subtle change in position of pixels due to expression, shift, tilt, or even noise. To overcome this category of problems, fusion of the wavelet coefficients is considered.

Wavelet-based fusion

Wavelet transform is a tool for time-frequency localization. It performs such localization by separating given signal into different frequency components and then each section is

analysed with a resolution matched to its scale. Since face images are two dimensional discrete signals, 2-D discrete wavelet transforms are used in this situation. Several researchers have already applied wavelet transforms for fusing images in time-frequency domain [21-23]. Results of most of those investigations reveal that wavelet-based fusion algorithms outperform other image fusion algorithms.

In general, discrete wavelet transform decomposes a discrete signal into two subsignals of half of its length with the first one as a running average and another one as a difference or fluctuation. Considering all the discrete wavelet transforms Haar wavelet is the simplest transform. It has benefits like fast, requires less memory space, reversible etc., but it lacks capturing of all the high frequency components. This is because Haar wavelet captures through a window of size two starting from odd index and if there is a huge difference in components with even index in comparison with that of odd index then that the large difference is not taken into account during computation of high frequency components. To get rid of such demerit, in this work, Daubechies wavelet transform have been used. Keeping all other calculation same as Haar wavelet transform Daubechies wavelet uses overlapping windows and therefore all the changes are reflected in the high frequency coefficients. In this work, Daubechies 4-tap wavelet has been considered with the filter coefficients shown in Table 1. Here; H0 and H1 are input decomposition filters and G0 and G1 are output decomposition filters.

The fusion of images using wavelets [24,25] follows a standard procedure and is performed at the decomposition level. The input images are decomposed by a discrete wavelet transform and the wavelet coefficients are selected using a set of fusion rules and an inverse discrete wavelet transform is performed to reconstruct the fused image. Wavelet fusion methods differ mostly in the fusion rule used for selection of wavelet coefficients. Numerous fusion rules for combining the wavelet coefficients can be explored: (i) simple average, (ii) weighted average, (iii) Maximum, (iv) Maximum for high frequency components and minimum for low frequency components, (v) Maximum for high frequency components and average for low frequency components, and (vi) Maximum for high frequency components and the weighted average for low frequency components. It has been recognized that these combination schemes are having little effect on the overall performance of the fusion process. In this work, fusion rule “Maximum for high frequency components and the weighted average for low frequency components” has been chosen. The main reason for this choice is that the highest values capture the salient information (e.g. edges) in the images and a weighted average of low frequency coefficients to avoid the effects of noises.

Principal component analysis

Principal Component Analysis (PCA) is a well established tool for dimensionality reduction. A human face with thousands of pixels, if represented immediately, needs a

Table 1 Daubechies 4-tap wavelet coefficients

H0	H1	G0	G1
0.4830	0.1294	-0.1294	0.4830
0.8365	0.2241	0.2241	-0.8365
0.2241	-0.8365	0.8365	0.2241
-0.1294	0.4830	0.4830	0.1294

huge dimensional space for statistical analysis and subsequent classification. This will require extremely large memory space and processing capabilities for a large database which is common in real life applications. This problem is not as significant as storage space and processing requirement may be manageable, but the concept of “curse of dimensionality” may apply here. For a trainable classifier, it confirms that the required number of training samples grows exponentially with the increase in dimensionality. The fundamental objective of this step is dimensionality reduction by representing face space in lower dimensional space by generating features that are optimally uncorrelated. In the year 1987, Sirovich and Kirby used PCA to reduce dimensions for features vector of faces [26]. In the year 1991, Turk and Pentland used the same PCA again for dimensionality reduction in solving complex pattern recognition problem like problem of face recognition [27]. As reported in [28-30], PCA not only reduces the size of face images but also handles variations in face images due to changes in facial expressions.

Multilayer perceptron neural network

Artificial neural network (ANN) has already established its capability for classification of models due to its adaptivity, generalization strength, robustness, and fault tolerance. A Multilayer Perceptron (MLP) is a supervised neural network that has been used as a classifier in different pattern classification tasks successfully. This learning algorithm applied to multilayer feed forward networks here is consisting of processing elements with continuous differentiable activation functions. Such networks associated with the back propagation learning algorithm are also called backpropagation networks [31-35]. In this work, a multilayer neural network which incorporates backpropagation learning with momentum is used. This is a homogeneous system with tansigmoid transfer function in all the processing elements or neurons. The system consists of three layers. In the first layer, the number of neurons is equal to the size of the feature vectors achieved after dimensionality reduction and in output layer number of neurons is only one; a binary classifier. To be more specific for each course, there is a system to classify members of this class from others. Weights for the system are initialised with pseudorandom numbers. Weights associated with all the interconnections incident on node k are initialized with the pseudorandom numbers generated in the range $[-3/\sqrt{n_k}, 3/\sqrt{n_k}]$, where n_k is the number of interconnections to the k^{th} node. Different parameters used to train the system are goal, learning rate (lr) and momentum constant (mc), epochs with values 10^{-6} , 0.06, 0.8, and 6,00,000 respectively.

Experimental results and discussions

For the performance analysis, several experiments were conducted. For this investigation, two face image databases were used. The first one is the IRIS database [20], and the second one is a face database created at our own laboratory and its proposed name is UGC-JU face database [36]. In all the experiments, images are normalized and cropped into corresponding images of size with width \times Height as 40×50 . Original images are normalized after registration of images with respect to frontal and neutral face images of each person. In order to obtain higher recognition accuracy, all the images are registered using variant of polar transforms and phase correlation. Corresponding registered visual and thermal face images are fused in the wavelet domain and then fused images are

transformed back into the spatial domain. In this work, a multilayer perceptron neural network has been applied for classification of fused images. To evaluate the performance of the classification, the classifier used here is considered as binary classifier i.e. for each class, there is one classifier with elements of that class as positive exemplars and elements from other classes taken as negative exemplars. During the testing, four different results are possible: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). False positive is considered when the outcome is incorrectly classified and marked as “yes” (or “positive”), when actually it is “no” (or “negative”). A false negative is marked when the outcome is incorrectly classified as negative but in real it is positive. Therefore, true positives and true negatives are correct classifications. Based on those four results other two statistical measures have been derived: Sensitivity and Specificity. Sensitivity is the measure that finds how good the test is in detecting real positives. It measures the proportion of actual positives, which are correctly identified and it is given as

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$

Whereas, specificity, is the measure that finds how well the test is at detecting real negatives. It measures the proportion of actual negatives, which are correctly identified and described as

$$\text{Specificity} = \text{TN}/(\text{FP} + \text{TN})$$

Another measure called accuracy is also calculated, which describes a ratio between number of samples correctly classified and total number of samples used in classification, given as

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

There are other measures exist, which can be computed from Specificity or Sensitivity and for that reason those are not computed here; those include:

$$\text{False positive rate} = \text{FP}/(\text{FP} + \text{TN}) = 1 - \text{Specificity}$$

$$\text{False negative rate} = \text{FN}/(\text{TP} + \text{FN}) = 1 - \text{Sensitivity}$$

Sensitivity, Specificity, and Accuracy measures computed for both the databases using different techniques implemented here are discussed in following sections.

IRIS thermal/visible database

This database [20] was initially formed in response to the IEEE Int'l Workshop on “Object Tracking and Classification in and Beyond the Visible Spectrum”. The benchmark is used for educational and research purposes only and is available for all researchers in the international computer vision communities. The description relating to the database is given below:

Sensor Details: Thermal - Raytheon Palm-IR-Pro

Visible - Panasonic WV-CP234

Data Details: Total size of 1.83 GB

Image size: 320 × 240 pixels (visible and thermal), Total 4228 pairs of thermal and visible images with 176–250 images/person, 11 images per rotation (poses for each expression and each illumination) 30 persons - Expression, pose, and illumination.

Expression: ex1, ex2, ex3 - surprised, laughing, angry (varying poses)

Illumination: Lon (left light on), Ron (right light on), 2on (both lights on), dark (dark room), off (left and right lights off), varying poses.

Some visual and corresponding thermal images of a person are shown in Figure 5.

To measure the sensitivity, specificity, and accuracy, the IRIS database for visual and thermal face images were considered. There are several methods implemented for this experiment. Without fusion, visual and thermal face images are considered separately, and six different methods were investigated. They are based on dimensionality reduction using PCA and classification using Multilayer Perceptrons (MLP). Only difference is in registration; first process is without any registration at all; second one is registration with polar transform, and the third one is with adaptive polar transform for both thermal and visual counts to six different designs. When fusion is taken into consideration, unregistered, registered with polar transform, and registered with adaptive polar transform along with pixel fusion and wavelet fusion we got total six different designs. For the calculation of results, 3-fold cross-validation [37] was considered, and the results are shown in Table 2. All the images are partitioned into three subsets. Out of those three groups one is used for training and rest two are used for testing. From the results, in obtaining better recognition performance we can infer the followings:

- i) Registration of face images produces better in comparison to an unregistered one.
- ii) Adaptive Polar Registration produces better results that polar one.
- iii) Fusion of visual and thermal face images outperforms the techniques which consider those individually.
- iv) Wavelet based fusion produces better results that pixel fusion.

UGC-JU face database

This face database has been created at Department of Computer Science and Engineering, Jadavpur University, India under a project with a grant from University Grants Commission (UGC), India. Infrared images are captured using an FLIR 7 camera, and visual images are captured using Sony DSC-W350 digital camera at our own laboratory. Images are captured in pairs, one thermal and one visual, under different constraints. Every person sitting on a chair at a distance of about 2 feet from both the cameras. Each person has 34 different templates: different expressions with eye movements and emotions without changing head orientation, different views about x-axis, y-axis, and z-axis etc. for 20 persons.

All the experiments conducted on IRIS face database have also been conducted on UGC-JU face database [36], results obtained is shown in Table 3. Trend of results for



Figure 5 Some visual images (first line) of a person with corresponding thermal images (second line) from IRIS face database. (All are in grayscale).

Table 2 Experimental results on IRIS face database

Sl. no.	Method		Sensitivity	Specificity	Accuracy
1	Unregistered + No Fusion + PCA + MLP	Visual images	90.42	81.60	88.73
		Thermal images	87.04	84.62	85.38
2	Unregistered + Pixel Fusion + PCA + MLP		91.93	86.32	90.86
3	Unregistered + Wavelet Fusion + PCA + MLP		94.21	85.30	92.50
4	Log Polar Registered + No fusion + PCA + MLP	Visual images	90.18	80.80	88.39
		Thermal images	89.24	78.70	87.22
5	Adaptive Log Polar Registered + No fusion + PCA + MLP	Visual images	91.84	84.70	90.47
		Thermal images	90.99	82.80	89.42
6	Log Polar Registered + Pixel Fusion + PCA + MLP		92.53	86.70	91.41
7	Adaptive Log Polar Registered + Pixel Fusion + PCA + MLP		94.77	88.10	93.50
8	Log Polar Registered + Wavelet Fusion + PCA + MLP		95.84	92.80	95.26
9	Adaptive Log Polar Registered + Wavelet Fusion + PCA + MLP		98.46	97.90	98.36

different techniques obtained in this database is also similar as in the case of IRIS face database and therefore the inferences made in IRIS thermal/visible database section. Some visual and corresponding thermal images of a person of UGC-JU database are shown in Figure 6.

Comparative study

Discussion won't be complete without comparison of results of the present process is not made with other existing recent designs. Comparison is given in Table 4, where all the methods are not using the same database. Here, we have tried to compare our method with other through the same database and also result due to our database UGC-JU has also been depicted.

In Table 4, a complete overview with regard to acceptance rate of different fusion techniques has been presented. The IRIS thermal visible face dataset (which is one of the OTCBVS benchmark dataset) has been used for all the experiments conducted

Table 3 Experimental results on UGC-JU face database

Sl. no.	Method		Sensitivity	Specificity	Accuracy
1	Unregistered + No Fusion + PCA + MLP	Visual images	81.91	60.00	77.53
		Thermal images	76.32	54.87	72.06
2	Unregistered + Pixel Fusion + PCA + MLP		86.32	73.53	83.76
3	Unregistered + Wavelet Fusion + PCA + MLP		87.21	77.06	85.18
4	Log Polar Registered + No fusion + PCA + MLP	Visual images	85.44	72.35	82.82
		Thermal images	82.79	64.12	79.06
5	Adaptive Log Polar Registered + No fusion + PCA + MLP	Visual images	86.62	75.88	84.47
		Thermal images	84.26	68.82	81.18
6	Log Polar Registered + Pixel Fusion + PCA + MLP		88.38	80.00	86.71
7	Adaptive Log Polar Registered + Pixel Fusion + PCA + MLP		92.06	82.94	90.24
8	Log Polar Registered + Wavelet Fusion + PCA + MLP		92.94	87.65	91.88
9	Adaptive Log Polar Registered + Wavelet Fusion + PCA + MLP		96.77	91.77	95.77



Figure 6 Some thermal images (first line) of a person with corresponding visual images (second line) from UGC-JU face database.

by M. K. Bhowmik et al. [19,39-45]. In [40], a pixel level fusion of visual and thermal image has been used, and 97.05% acceptance rate is achieved. In [45], fused images have been classified using radial basis function and multilayer perceptron and the listed images using RBF shown better accuracy than MLP, which is 96.0%. In [19], log polar transform of fused images has been analysed over MLP, and the acceptance rate is 93.81% which is much lesser than other techniques [40,42,45]. In [41], an optimum level fusion of visual and thermal images has been introduced, and 93% acceptance rate is achieved. In [42], a new dimension reduction technique, candid covariance-free

Table 4 A comparative study between different fusion methodologies

Sl. no.	Author	Techniques	Database used	Reported performance
1.	M. K. Bhowmik et al. [19]	log-polar transformed + PCA	OTCBVS (IRIS)	93.81%
2.	Mohammad Hanif et al. [38]	Gabor Filter Technique	Equinox	DWT - 90.31% OIF - 95.84%
3.	M. K. Bhowmik et al. [39]	Daubechies wavelet transform + PCA + ICA	OTCBVS (IRIS)	PCA - 91.13% ICA I - 94.44% ICA II - 89.72%
4.	M. K. Bhowmik et al. [40]	Pixel fusion + RBF	OTCBVS (IRIS)	97.05%
5.	M. K. Bhowmik et al. [41]	Optimum + Eigenspace projection + Multilayer Perceptron	OTCBVS (IRIS)	93%
6.	D. Bhattacharjee et al. [35]	Eigenspace projection + Multilayer Perceptron + Backpropagation learning	OTCBVS (IRIS)	95.07%
7.	M. K. Bhowmik et al. [42]	Pixel fusion + CCI PCA + SVM	OTCBVS (IRIS)	97.28%
8.	M. K. Bhowmik et al. [43]	Wavelet transformation + multiresolution analysis + MLP + RBF	OTCBVS (IRIS)	Feature level - 87.28% Decision level - 94.95%
9.	M. K. Bhowmik et al. [44]	Daubechies wavelet co-efficient + ICA + MLP	OTCBVS (IRIS)	91.5%
10.	M. K. Bhowmik et al. [45]	Eigenspace projection + MLP + RBF	OTCBVS (IRIS)	RBF - 96% MLP - 95.07%
11.	D. R. Kisku et al. [46]	Dempster-Shafer decision theory + SIFT features	ORL, IITK	ORL - 98.93% IITK - 96.29%
12.	R. Singh et al. [47]	Match score fusion + 2v GSVM + Dezert Smarandache theory + SVM	Equinox	DSm match score fusion - 98.08% SVM - 95.05% DST - 96.51% 2v GSVM - 94.98%
13.	Present method	Adaptive Log Polar Registered + Wavelet Fusion + PCA + MLP	OTCBVS(IRIS) UGC-JU	98.36% 95.77%

incremental principal component analysis (CCIPCA) is applied over the fused images, and finally those lower spaced fused images has been classified using different SVM kernel and 97.28% acceptance rate is achieved, which is the maximum of all fusion methodologies, those performed by M. K. Bhowmik et al. Since, Equinox database is not available so no comparison could have been completed. Results are compared for IRIS database only. All other methods, mentioned here, worked with IRIS database did not consider complete database for the evaluation of results, whereas the present method considered the entire database and has shown a very convincing result for acceptance of this method for development of the face identification system.

Conclusion

We present a fully automatic human face recognition system which efficiently registers face images using adaptive polar transformation. This technique also uses a wavelet based fusion method that combines visual and thermal faces images in a robust manner. Alternative methods were also implemented with lower computational cost, but recognition performance achieved is much lower than this. Different sources of complicacies when present in face images it becomes difficult to solve this recognition problem in efficient manner. This method has shown a very good performance in recognising almost all the possible complicacies to a large extent.

There are possibilities for further improvements of the existing design. Registration may be done based on some particular landmarks and considering symmetry of the face image to handle images captured across poses. Support vector machine (SVM) may improve classification result instead of a simple MLP. There may be fusion of different region classifiers used at various distinct region of the face images.

Competing interests

Author acknowledges the incessant and encouraging support of his wife Ms. Rituparna Bhattacharjee, which helped him in pursuing this research.

Received: 21 June 2013 Accepted: 17 March 2014

Published online: 06 June 2014

References

1. Liu Z, Liu C (2008) A hybrid color and frequency features method for face recognition. *IEEE Trans Image Process* 17(10):1975–1980. doi:10.1109/TIP.2008.2002837
2. Basri R, Jacobs DW (2003) Lambertian reflectance and linear subspaces. *IEEE Trans Pattern Anal Machine Intel* 25(2):218–233. doi:10.1109/TPAMI.2003.1177153
3. Hsieh CK, Chen YC (2007) Kernel-based pose invariant face recognition. In: *Proc. IEEE Int. Conf. Multimedia and Expo.*, pp 987–990. doi:10.1109/ICME.2007.4284818
4. Ashraf AB, Lucey S, Chen T (2008) Learning patch correspondences for improved viewpoint invariant face recognition. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '08)*, pp 1–8. doi:10.1109/CVPR.2008.4587754
5. Gizatdinova Y, Surakka V (2006) Feature-based detection of facial landmarks from neutral and expressive facial images. *IEEE Trans Pattern Anal Machine Intel* 28(1):135–139. doi:10.1109/TPAMI.2006.10
6. Kawato S, Ohya J (2000) Two-step approach for real-time eye tracking with a new filtering technique. In: *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics*, pp 1366–1371
7. Pavlidis I, 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 (CVBVS 2000)*, p 15. doi:10.1109/CVBVS.2000.855246
8. Park JS, Oh YH, Ahn SC, Lee SW (2005) Glasses removal from facial image using recursive error compensation. *IEEE Trans Pattern Anal Machine Intel* 27(5):805–811. doi:10.1109/TPAMI.2005.103
9. Heo J, Kong SG, Abidi BR, Abidi MA (2004) Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In: *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, p 122. doi:10.1109/CVPR.2004.35
10. Buddharaju P, Pavlidis I, Kakadiaris I (2004) Face recognition in the thermal infrared spectrum. In: *Proceedings of the IEEE Workshop on Computer Vision and Pattern Recognition Workshop (CVPRW '04)*
11. Pavlidis I, Buddharaju P, Manohar C, Tsiamyrtzis P (2006) Biometrics: face recognition in thermal infrared. In: *Biomedical Engineering Handbook, 3rd Edition*, CRC Press, pp 1–15
12. Dowdall JB, Pavlidis I, Bebis G (2003) Face detection in the near-IR spectrum. *Image Vis Comput* 21(7):565–578

13. Yin Z, Malcolm AA (2000) Thermal and visual image processing and fusion. In: SIMTech Technical Report (AT/00/016/MVS). Machine Vision & Sensors Group, Automation Technology Division, Singapore
14. Flusser J, Sroubek F, Zitova B (2007) Image Fusion: Principles, Methods, and Applications. In: Tutorial EUSIPCO
15. Zitova B, Flusser J (2003) Image registration methods: a survey. *Image Vis Comput* 21(11):977–1000. doi:10.1016/S0262-8856(03)00137-9
16. Zokai S, Wolberg G (2005) Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. *IEEE Trans Image Process* 14(10):1422–1434
17. Matungka R, Zheng YF, Ewing RL (2008) 2D invariant object recognition using log-polar transform. In: Proc. World Congress on Intelligent Control and Automation (WCICA '08), pp 223–228
18. Matungka R, Zheng YF, Ewing RL (2009) Image registration using adaptive polar transform. *IEEE Trans Image Process* 18(10):2340–2354
19. Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M (2011) Polar fusion technique analysis for evaluating the performances of image fusion of thermal and visual images for human face recognition. In: IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM), 11–15 April 2011, Paris, pp 62–69
20. OTCBVS (2013) OTCBVS Benchmark Dataset Collection., <http://www.cse.ohio-state.edu/otcbvs-bench/>. Accessed 20 November 2013
21. Nikolov SG, Canga EF, Lewis JJ, Loza A, Canagarajah CN, Bull DR (2004) Fusion of Visible and Infrared Image Sequences Using Wavelets. In: Technical report. The University of Bristol, Bristol, UK, p 202
22. Lewis JJ, O'Callaghan RJ, Nikolov SG, Bull DR, Canagarajah N (2007) Pixel- and region-based image fusion with complex wavelets. *Informa Fusion* 8:119–130
23. Zaveri T, Zaveri M (2011) A novel region based multimodality image fusion method. *J Pattern Recogn Res* 2(2011):140–153
24. Nunez J, Otazu X, Fors O, Prades A, Pala V, Arniol R (1999) Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans Geosci Remote Sensing* 37(3):1205–1211
25. Singh S, Gyaourva A, Bebis G, Pavlidis I (2006) Face recognition by fusing thermal infrared and visible imagery. *Image Vis Comput* 24:727–742
26. Kirby M, Sirovich L (1990) Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Trans Pattern Anal Machine Intel* 12(1):103–108
27. Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3(1):71–86
28. Lekshmi VP, Sasikumar M, Naveen S (2008) Analysis of Facial Expressions from Video Images using PCA. In: Proceedings of the World Congress on Engineering 2008 Vol I. WCE 2008, London, UK
29. Sun W, Ruan Q (2006) Two-Dimension PCA for Facial Expression Recognition. In: ICSP2006
30. Lin D (2006) Facial expression classification using PCA and hierarchical radial basis function network. *J Inf Sci Eng* 22:1033–1046, 1033
31. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2008) Classification of polar-thermal eigenfaces using multilayer perceptron for human face recognition. In: Proc. 3rd IEEE Conf. Industrial and Information Systems (ICIIS '08), p 118
32. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2009) Human face recognition using line features. In: Proc. National Seminar on Recent Advances on Information Technology (RAIT '09), p 385
33. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2008) Classification of log-polar-visual eigenfaces using multilayer perceptron. In: Proc. 2nd Int. Conf. on Soft computing (ICSC '08), pp 107–123
34. Bhattacharjee D, Basu DK, Nasipuri M, Kundu M (2010) Human face recognition using fuzzy multilayer perceptron. *Soft Comput* 14:559–570
35. Bhattacharjee D, Bhowmik MK, Nasipuri M, Basu DK, Kundu M (2009) Classification of fused face images using multilayer perceptron neural network. In: Proc. Int. Conf. on Rough sets, Fuzzy sets and Soft Computing, pp 289–300
36. Seal A, Bhattacharjee D, Nasipuri M, Basu DK (2013) UGC-JU face database and its benchmarking using linear regression classifier. In: *Multimedia Tools and Applications Journal of Springer*. doi:10.1007/s11042-013-1754-8
37. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc 14th Int Joint Conf Artif Intel* 2:1137–1143
38. Hanif M, Ali U (2006) Optimized visual and thermal image fusion for efficient face recognition. In: 9th International Conference on Information Fusion, Florence, pp 1–6
39. Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M (2011) Independent Component Analysis (ICA) of fused Wavelet Coefficients of thermal and visual images for human face recognition. In: SPIE Defense, Security, and Sensing 2011 (track of Independent Component Analysis, Wavelets, Neural Networks, Biosystems and Nanoengineering, Conference 8058). Published by SPIE and SPIE Digital Library, in Orlando World Center Marriott Resort & Convention Center, Orlando, Florida, USA
40. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2009) Classification of fused images using Radial basis function Neural Network for Human Face Recognition. In: Proc. of The World congress on Nature and Biologically Inspired Computing. NaBiC-09, Coimbatore, India
41. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2010) Optimum Fusion of Visual and Thermal Face Images for Recognition. Proc of 6th Int. Conf on Information Assurance and Security (IAS 2010), Atlanta, USA, pp 311–316
42. Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M (2012) Multisensor Fusion of Visual and Thermal Images for Human Face Identification using Different SVM Kernels. In: IEEE Conference on Long Island Systems, Applications and Technology (LISAT 2012). Farmingdale, New York
43. Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M (2012) Human face Recognition Using Multisource Fusion. In: Track of Multi Sensor, Multisource Information Fusion: Architecture Algorithms, and Applications 2012 (DS223) SPIE Defense, Security and Sensing 2012, Baltimore Convention Center. Published by SPIE and SPIE Digital Library, Baltimore, Maryland, USA
44. Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M (2012) Eye region based fusion technique of thermal and optical images for human face recognition in dark. *Optical Eng J SPIE* 51(7):2012
45. Bhowmik MK, Bhattacharjee D, Nasipuri M, Basu DK, Kundu M (2009) Image Pixel Fusion for Human Face Recognition. In: Published in International Journal of Recent Trends in Engineering [ISSN 1797–9617]. Academy Publishers, Finland, pp 258–262, Vol. 2, No. 2

46. Kisku DR, Tistarelli M, Sing JK, Gupta P (2010) Face Recognition by Fusion of Local and Global Matching Scores using DS Theory: An Evaluation with Uni-classifier and Multi-classifier Paradigm. In: IEEE Computer Vision and Pattern Recognition Workshop on Biometrics
47. Singh R, Vatsa M, Noore A (2008) Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition. *Pattern Recognition Journal of Elsevier Science Inc, New York, USA*, Vol. 41 Issue 3

doi:10.1186/s13673-014-0004-z

Cite this article as: Bhattacharjee: Adaptive polar transform and fusion for human face image processing and evaluation. *Human-centric Computing and Information Sciences* 2014 **4**:4.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
