Face identification by using fusing Photographic and Thermal Images

[1]M. Parisa Beham, [2]M.R.H. Prasanna, [2]SM.Mansoor Roomi and [1]H. Jebina
[1]Vickram College of Engineering, Tamilnadu, India.
[2]Thiagarajar College of Engineering, Tamilnadu, India.
*Corresponding Author:E-Mail parisa@vickramce.org

ABSTRACT

Face recognition is a biometric security system which is mainly used for authentication and surveillance applications. Face recognition system must be automatic and it must be robust. It is affected by variations like illumination variations, pose variations, expression variations and occlusions. Face recognition using visual images can't efficiently solve all these variations. Face recognition using thermal images can solve few of these variations but they are very sensitive to the temperature variations. By fusing different source images such as visual and thermal images the efficient system can be designed. In this system, the acquired visual and thermal images are first shaped using the affine transformation. The registered images are fused using Discrete Wavelet Transform and the features obtained using Principal Component Analysis method. The SVM classifier is trained using these obtained features and it is observed that 94% of recognition accuracy is achieved using the proposed system.

KEY WORDS: Affine Transformation, Discrete Wavelet Transform, Image Fusion and Principal, Component Analysis.

1. INTRODUCTION

In the recent decades because of the increased interest in biometric security systems in general, face recognition has recently gained a lot of importance. Since 1970 the face detection system has been keep on developed. As the normal system can't track the need of users, many techniques are proposed during 1990's based on the face recognition and it leads to improve the face detection application much. Face recognition is one of the biometric being used for authentication and security applications. It has a wide range of applications and it is very difficult to implement the system which can be able to recognize the faces under any constraints. The face detection aim is to detect faces under any constraints. In the face recognition system face is the primary focus and it plays a major role in conveying the identity and emotion.

If the face direction is only frontal then the face detection is very simple and recognition can be easily done. But in real case the face direction can't be always frontal and the head may be rotated in any angle which leads to pose variations. It is not mandatory that the environment at which the images are captured will always be bright and the illumination level will be same. These different illumination conditions will leads to illumination variations. Similarly there will be faces with different expressions and occluded by different objects which leads to expression variations and occlusions. Face recognition using visual images can effectively recognize the face when there is less percentage of variations in face. However it can't handle the variations like occlusions, illumination and pose variations. To solve these variations thermal images have played a major role. The thermal images are insensitive to the illumination variations, disguises and expression variations. But the thermal images are very sensitive to the temperature changes in the ambience and the visual images are insensitive to these temperature changes. Although there have been many methods developed to recognize the face using either thermal or visual images but not equals human ability to identify the face despite many variations. The human can identify face under any constraints throughout the life time. Hence the methods have been proposed to design a system based on the fusion of visual and thermal images. The image fusion is nothing but combining the best information available from both the images. There are many fusion techniques available such as simple fusion technique, pyramid based fusion and wavelet based fusion. Each fusion technique has some advantages and disadvantages and the choice of the fusion technique plays a major role in achieving the good accuracy. Apart from fusion technique the choice of image registration technique and classifier also plays a major role in achieving good recognition accuracy. The algorithm should be designed in such a way that it should be able to recognize the face under any constraint and the system should be invariant to scaling, translation, rotation, illumination, expression variations and occlusions.

Related works: In the recent years, series of researches have been conducted on the fusion of visual and thermal images for efficient face recognition. In (Tarek and Nabil, 2014), a thermo-visual fusion schemes for background subtraction is presented. They take advantage of both sensors to provide a more complete assessment of the foreground. To ensure an automatic 24/7 persistent operation, an efficient image classification algorithm is introduced. This work gives the basis for complicated video surveillance tasks such as the continuous tracking of moving objects where it is not dealt pose variations and disguise problems. The paper (Debotosh, 2014), represents human face recognition system which is fully automatic and efficiently changes the images using transformation techniques. This proposed method combines images in a robust manner. This method has shown its best performance in identifying almost all the situations to a large extent. Occlusions problems were not dealt in that proposed algorithm. In the paper (Bhowmik, 2012), face identification is performed using different types of SVM kernels. The

paper (Bhowmik, 2010), represents a comparative study of two different methods, based on fusion of polar transformed visual and thermal images with fusion of visual and thermal images with log-polar transformation to handle pose variations, partial occlusions etc. Both the methods shows good accuracy but occlusion problems were not completely dealt. In the paper (Bhowmik, 2010), a weighted sum of pixels from each images have been used for fusion and solved the illumination problem. Expression changes using pixel based fusion scheme was not discussed. In the paper (Ahmad, 2006), visual and thermal images are fused and the feature vectors are extracted by using Gabor filter. A design's computational cost is same as the previous existing systems and at the same time efficiency is also higher than the existing systems. Problems of illumination and pose variations were not discussed. In the paper (Besma, 2004), an eye region is given the at most importance and that region is replaced with an eye template. In the presence of eyeglasses the recognition rate is poor and after replacing with an eye template, recognition rate is improved.

The existing literature has revealed that by using computer vision systems the fast and efficient recognition system had been achieved. To perform automatic identification of the faces a number of complicacies had to be overcome. Thus inspired by the success of fusion techniques, in this work, recognition task has been accomplished by fusing thermal and photographic images.

2. PROPOSED METHODOLOGY

The proposed work considers both visual and thermal images as the input images. Since the captures images are affected by translation, scaling and rotation variations, the images have to be registered in such a way that both the coordinates system must get satisfied. Once the images are registered the fusion is performed using a Discrete Wavelet Transform (DWT). Thus the fused image is obtained and the dominant feature is extracted using Principal Component Analysis (PCA) method. These features are trained by the SVM classifier to accomplish the classification task.

Database description: Lack of availability of thermal face images motivated us to create a new face dataset consist both thermal and photographic images. To demonstrate the efficacy of the proposed algorithm, internal dataset using 8 different persons is developed and stored. The face images of the 8 subjects are captured at different illuminations, facial expressions, pose variations and occlusions. Thus each subject has 14 samples. For the 8 subjects both thermal and photographic images have been captured. The images were taken at a distance of 150 cm from the image acquisition point.

Image Acquisition: In the prototype of the proposed work, the visual and thermal images of faces are captured. The images are acquired at the controlled environment (Senthilarasi, 2014). In a standard capture condition, images were taken on a same background (Senthilarasi, 2014). Each face image of a person is captured in different variations using (Flexible Readout and Integration Sensor) FRIS camera.

Process Flow:

Preprocessing: The objective of the proposed work is to recognize the face under any constraints (Senthilarasi, 2014). The acquired images are converted into gray scale for further processing (Senthilarasi, 2014), using Eq.1.

$$t(a,b) = 0.299R(a,b) + 0.587G(a,b) + 0.114B(a,b) \qquad (1)$$

Where $t(a, b)$ is the intensity at position (a, b) of the gray image and $R(a, b)$, $G(a, b)$ and $B(a, b)$ represents the channels of a colour image (Senthilarasi, 2014).

Image Registration: When the images are taken at different illumination level or at different orientations or at comibnation of many different conditions the images have to be matched for further processing. Image registration is a process of matching the images acquired at different conditions by using the common features depicted by both the images. Procedure of Image registration is described as follows:

Feature detection and Matching: In this step the important parts of face from both the images are identified and they are computed as features. These features are the control points and correspondences between the images are found using similarity measure between particular features computed.

Transform model estimation and image transformation: With the use of control points, estimation of transformation is done by using affine transformation technique. It is one of the most widely used registration methods in registering two images. A geometric transformation refers to a combination of translation, scaling and rotation with a general form of

$$X = \text{RS}(u + t) = \text{A}u + b \qquad (2)$$
$$u = \beta(x - b) = \beta x + c \qquad (3)$$

Where $A = \text{RS}$, $b = \text{RSt}$, $c = t$ and $\beta = \text{inverse}(A)$, these equations represent the combinations of translation, scaling and rotation coefficients respectively.

An affine transformation corrects some global distortions in the images to be registered. A linear combinhion is a function that preserves all linear combinations. An affine transformation is a function that preserves all affine combinations. Using this transformation, an object matrix is estimated. Finally an image warping is performed to transform the image that becomes closest to the desired image.

Image Fusion: Image fusion is a process of combining two or more images and the resultant image contains the best features from all the source images (Bhowmik, 2010). The proposed multimodal fusion fused the images obtained by using two different sources namely visual and thermal. There are several methods of fusion is available in the literature. Here the wavelet based fusion approach is used. Wavelet is a powerful signal analysis tool which is widely used in extracting the feature points, compressing the image size and noise removing applications. The main advantage of using wavelet transform is that it is well suited to manage different resolution of images and allows the image decomposition in desired coefficients, while preserving the information of images. In this work, the facial image is decomposed into high and low frequency components using DWT decomposition. In DWT decomposition (Bhowmik, 2010), the filters are specially designed so that successive layers include details which are absent at the preceding levels. The DWT decomposition uses a series of special low pass and high-pass filters and sub sampling operation. The outputs from 2D-DWT image having size equal to half the size of the original image. Thus from the photographic and thermal images, obtained coefficients are HHp, HLp, LHp, LLp and HHt, HLt, LHt and LLt respectively. These coefficients are nothing but an approximation details, horizontal details, vertical details and diagonal details. To fuse the coefficients maximum value of coefficients from both images are obtained and these coefficients are given to the inverse 2D-DWT transform to obtain the fused image.

Feature Extraction: In this proposed step, the task is to describe the regions based on the selected representation. Principal Component Analysis (Abate, 2007), is a tool which is used here to extract the features from the fused image. A human face image have huge number of pixels and processing such huge amount of pixels in a short span of time is a very complicated task. The memory requirement is also very high to store such huge amount of pixels. The main aim of this step is to reduce the dimensionality by using features of face images that are optimally uncorrelated. PCA reduces the size of face images and manages some of the complicacies in face images. From the fused photographic and thermal images, PCA is applied and thus obtained features are trained by the SVM classifier for the recognition process.

Classification of Faces: The classification of faces is performed using the Support vector Machine Classifier (Senthilarasi, 2014). The objective of the proposed work (Senthilarasi, 2014), is to recognize the human face. The SVM is a theoretically superior machine learning methodology (Senthilarasi, 2014) and SVM method was derived from optimal hyper plane in the linearly separable case (Senthilarasi, 2014). Consider the following classification problem, the samples to be taken as training are represented as *pj; qj; j=1,2,...M* and *qj ε -1; +1*, where *M* is the count of samples to be used as training *, qj = +1* for class *t1* and *qj = -1* for class *t2=2*. Suppose the two classes are linearly separable (Senthilarasi, 2014), it is possible to find at least one hyper plane defined by a vector *t* with a bias *t0* (Senthilarasi, 2014), which can separate the classes with high recognition accuracy.

$$f(x) = t \cdot p + t_0 \qquad (4)$$

To find such a hyper plane, *t* and *t0* should be estimated in a way that

$$q_j(t \cdot p + t_0) = +1 \text{ for } q_j = +1 (\text{Class} t_1)$$
$$q_j(t \cdot p + t_0) = -1 \text{ for } q_j = -1 (\text{Class} t_2) \qquad (5)$$

These two equations are combined as:

$$q_j(t \cdot x_j + t_0) - 1 \geq 0 \qquad (6)$$

Many hyper planes could be fitted to separate the two classes but there is only one optimal hyper plane that is expected to generalize better than other hyper planes (Senthilarasi, 2014).

3. RESULTS AND DISCUSSION

The internally created face database consists of 8 subjects with 14 samples each for photographic as well as thermal images (totally 224 images). For experimentation, 168 images for training and 56 images for testing were considered. Fig. 1(a) and (b) shows the sample visual images and their respective thermal images.

Figure.1. (a) Sample visual images   (b) Sample thermal images

From the input images it is observed that the images were affected by geometric variations. In order to fuse these images both the images must have the same size and should be geometrically corrected. Hence an affine transformation is performed and image warping was done on the visual image to get matched to the coordinates of the thermal image. It is well known that the visual image is a high resolution image and thermal image is a low resolution one. In order to avoid the quality loss and degradation effect, the visual image is warped to get matched with the thermal image. After transformation the images are converted into grayscale and resized to 256×256.

Figure.2 (a) Registered Visual Images   (b) Fused Images

The images are registered and the DWT based fusion is performed to fuse the two images.

PCA is performed on the fused images to extract the discriminant features. These features are trained by the SVM classifier. Here the proposed face recognition is considered to be a multi-class classification problem. There are two basic methods for face recognition with SVMs (Bhowmik, 2012): one- against-one and one-against-all (Bhowmik, 2012). The one-against-all method was used in this proposed classification and it is between each class and all the rest classes (Bhowmik, 2012). Both positive samples and negative samples were taken as input samples to train a SVM classifier to get a corresponding support vectors and optimal hyper plane (Bhowmik, 2012). The SVM was labeled as SVM_1. In turn we can get the SVM for every individual.

When a test sample is in turn inputted to SVM classifier, there would be several cases (Bhowmik, 2012):

a) If the sample is decided to be positive by SVM_i and to be negative by others SVMs at the same time, then the sample is classified as class I (Bhowmik, 2012).

b) If the sample is decided to be negative by several SVMs synchronously and to be positive by other SVMs, then the classification is false (Bhowmik, 2012).

c) If the sample is decided to be negative by all SVMs synchronously, then the sample is decided not belong to the database (Bhowmik, 2012).
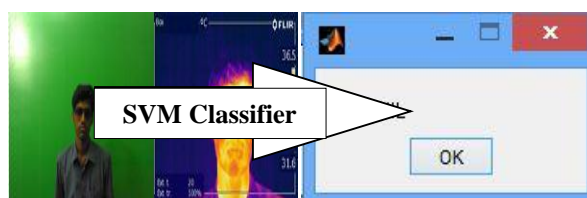
Figure.3. Recognition Result

Thus the classification was performed using SVM classifier as shown in Fig.3 and 94% of recognition accuracy is achieved.

4. CONCLUSION

Human face recognition is a challenging task and it has more practical domain of applications. In this work, a novel face recognition method is implemented which is robust to occlusions and illumination, rotation and pose variations. Here the DWT fusion technique is applied to fuse the visual and thermal image of a person. From the fused image, PCA feature is obtained which is used for recognizing the faces. For the geometrical transformations affine transformation is used and thus achieved better recognition accuracy with very less computational cost. The database consists of both visual and thermal images have been internally created and are used for the evaluation of the proposed algorithm. Considerable deviations suggest needs for next movement in the process. The study was made with a small size of samples and for better improvement the sample size must be increased. The limitation of the proposed system is its inability to work with images with more complicacies. Robust image registration methods and efficient feature selection methods need to be implemented. The future goal is to extend the current face recognition application to design an efficient face recognition system, robust to any type of complicacies in unconstrained environment.

REFERENCES

Abate AF, Nappi M, Riccio D, & Sabatino G, 2D and 3D face recognition: A survey, Pattern Recognition Letters, 28 (14), 2007, 1885-1906.

Ahmad J, Ali U, and Qureshi RJ, Fusion of Thermal and Visual images for efficient face recognition using Gabor Filter, International conference on computer systems and applications, IEEE, 2006.

Besma Abidi, Shafik Huq, and Mongi Abidi, Fusion of Visual, Thermal and Range as a solution to illumination and pose restrictions in face recognition, International Carnahan Conference on Security Technology, IEEE, 2004.

Bhowmik MK, Bhattacharjee D, Basu DK, and Nasipuri M, Multisensor fusion of Visual and Thermal images for human face identification using different SVM kernels, IEEE, 2012.

Bhowmik MK, Bhattacharjee D, Basu DK, and Nasipuri M, Polar fusion technique analysis for evaluating the performances of image fusion of Thermal and Visual images for face recognition, Sixth International Conference on Information Assurance and Security, IEEE, 2010.

Bhowmik MK, Bhattacharjee D, Basu DK, Nasipuri M, and Kundu, Optimum fusion of Visual and Thermal face images for recognition, Sixth International Conference on Information Assurance and Security, IEEE, 2010.

Debotosh Bhattacharjee, Adaptive Polar transform and fusion for human face image processing and evaluation, Human centric computing and information science, Springer, 2014.

Senthilarasi M, Md.Mansoor Roomi S, and Prasanna M.R.H, Shape based approach for detecting musa species in fruit industry, Sixth International Conference on Advanced Computing (ICOAC), IEEE, 2014.

Tarek Mouats and Nabil Aouf, Fusion of Thermal and Visual images for Day/ Night moving objects detection, Sensor signal processing for defence, IEEE, 2014.