Fusion of Thermal and Visual Images for efficient Face Recognition using Gabor Filter

Jahanzeb Ahmad, Usman Ali, Rashid Jalal Qureshi,
1Dept of Electrical Engineering, COMSATS Institute of Information Technology, Abbottabad, Pakistan
2Laboratoire d'Informatique, Université François-Rabelais de Tours, France
1jahanz1982@yahoo.com, 1usman_gikian_abbott@hotmail.com, 2rashid-jalal.qureshi@etu.univ-tours.fr,

Abstract

Face recognition from visual images is a difficult task due to illumination problem and in thermal imaging the main problem is of glasses. The solution of both these problem is data fusion of thermal and visual images. This paper presents the implementation of Human face recognition system using data fusion of visual and thermal images. Gabor filtering technique which extracts facial features is used in the proposed face recognition. To our knowledge, this is the first visual and thermal data fusion recognition system which utilizes Gabor filter. Paper also discusses the performance improvement of face recognition along with issues of memory requirements.

1. Introduction

Human face recognition is a challenging task and its domain of applications is very vast, covering different areas like security systems, defense applications, and intelligent machines. It involves different image processing issues like face detection, recognition and feature extraction.

Face recognition using visual images is gaining acceptance as a superior biometric [1]. Images taken from visual band are formed due to reflectance. Therefore they depend on the external light source, which sometimes might be absent e.g. night time or when there are heavy clouds. Imagination is also difficult because it depend on the intensity of light and angle of incident of light.

Some of the commonly used visual face recognition techniques are Principal Component Analysis (PCA) [2], Independent Component Analysis (ICA) [3], Linear Discriminate Analysis (LDA) [4], and Elastic Bunch Graph Matching (EBGM) [5], and Gabor Filtering technique [6].

Recently, face recognition on thermal infrared spectrum [7, 8]. It has gained popularity because thermal image is formed due to emission not reflection. They do not depend on the external light source and intensity of light [9, 10] and are also less dependent on angle of incident of light. Different methodologies of face recognition have been performed on visual and thermal images [11, 12, 13], and compared.

A face recognition method in the infrared spectrum is presented in [14], where Gabor filtering is used as a spectral analysis tool. They have used Equinox facial database [15] and have provided comparison with eigenfaces method for infrared images. For the detailed review on current advances in visual and thermal face recognition refer to [16].

In [17] many multisensor data fusion architectures are presented to create a color night vision capability for different purposes. Fusion of visual and thermal signatures for robust face recognition is explained in [18] has outperformed the individual visual and thermal face recognizers. Both data and decision fusion is discussed in [15], where facelt recognition is used for testing the fusion of Equinox face database images.

In this paper we present the data fusion architecture for the fusion of thermal and visual face images to achieve better biometric face recognition system, which utilizes Gabor filter. Paper is organized as follow: In section 2 Data fusion is discussed. Sections 3 and 4 discuss Feature point calculation and Feature point selection respectively. Feature vector generation, Similarity calculation in proposed architecture is presented in subsequent sections. Finally, the results and conclusions are discussed.

2. Data Fusion Architecture

Data fusion technique provides an advantage of combining the information of both the sources (visual and thermal) to produce new more informational image is obtained for recognition, which produces more accurate results. The Equinox facial database [15], the most extensive infrared facial database that is publicly available at the moment, was used for testing. The Equinox database has a good mix of subject
images with accessories (e.g., glasses) as well as expressions of happiness, anger, and surprise, which account for pose variation. Figure 1 shows some examples from this database.

The visual and thermal image is combined by using equation as proposed in [18]:

\[ F(x, y) = F_w T(x, y) + (F_w - 1)V(x, y) \]

Where \( F(x,y) \) is the fused image \( T(x,y) \) and \( V(x,y) \) are the thermal and visual image respectively and \( F_w \) is the weight of fusion and its value is from 0 to 1. Fused image \( F(x,y) \) is passed to Face recognition system as input. Few fused images are shown in Figure 2.

3. Feature point calculation

Physiological studies found simple cells in human visual cortex that are selectively tuned to orientation as well as to special frequency, it was suggested that the response of a simple cell could be approximated by 2-D Gabor filters [19]. One of the most successful recognition methods is based on graph matching of coefficients, which
have disadvantages due to their matching complexity.

Escobar and Javier [20] proposed a model in which they manually located the feature points and then calculated the Gabor jet manually which describes the behavior of the image around that point.

As for automatic face recognition we cannot locate feature points manually so to detect points automatically we have calculated the Gabor filter response at every point to see the behavior of the image around that point. A filter response at any point can be calculated by convolving the filter kernel with the image at that point. For point \((X, Y)\), filter response denoted as \(R\) is defined as:

\[
R = \sum_{x=-X}^{X-1} \sum_{y=-Y}^{Y-1} I(x+X, y+Y) f(x, y, \theta, \lambda)
\]

\[
f(x, y, \theta, \lambda, \sigma X, \sigma Y) = e^{-0.5 \frac{(R1^2 + R2^2)}{\sigma X^2 + \sigma Y^2}}
\]

Where \(\sigma X\) and \(\sigma Y\) are the standard deviation of the Gaussian envelop along the x and y dimensions respectively, \(\lambda\), \(\theta\) and \(n\) are the wavelength, orientation and no of orientations respectively. \(I(x, y)\) denotes NxM image.

When we apply all Gabor filters at multiple frequencies and orientations at a specific point we thus get filter response for that point. We have chosen four orientations and a constant wavelength because feature points are relatively insensitive to the Gabor kernel wavelength, while vary significantly across different orientations [5]. We have constant \(\lambda = 2\times1.414\) and \(\sigma X = \sigma Y = \lambda/2\).

**4. Feature Point selection**

Mostly eyes, nose, mouth and corner of lips are taken as feature points. However, in our implementation we do not fix the feature points because of varying facial characteristics of different faces such as dimples, moles etc. Human mind also uses these characteristics for face recognition. We chose the point in a particular window of size \(S \times T\) around which the behavior or response of Gabor filter kernel is maximum, as feature point.

\[
S = \frac{N}{\sqrt{W}} \quad \text{and} \quad T = \frac{M}{\sqrt{W}}
\]

Where \(N\) = no of columns and \(M\) = no of rows and \(W\) is the no of windows. Feature point located at any point can be evaluated as

\[
R_f(x_o, y_o) = \max_{(x,y) \in C} (R_f(x, y))
\]

Where \(R_f\) is the response of the image to the \(j^\text{th}\) Gabor filter and \(C\) is any window. Window size is one of the important constraints of our implemented model. It should be small enough to capture all important facial feature points, but it should be large enough so that no redundancy occurs. Feature responses are obtained by applying above method on all windows.

**5. Feature vector generation**

Feature vectors are generated at feature points as discussed in previous sections. \(p^{th}\) feature vector of \(j^{th}\) reference face is defined as:

\[
v_{i,p} = [x_p, y_p, R_f(x_p, y_p)]
\]

Where \(j = \text{no of responses}\). Feature vector contains response with location information.

**6. Similarity calculation**

The degree of similarity is calculated between input image and all the images from the database. Similarity between features of input image and any image from database is calculated using:

\[
S(p, j) = \frac{\sum_{i} |v_{i,p}(l)| |v_{i,j}(l)|}{\sqrt{\sum_{i} |v_{i,p}(l)|^2 \sum_{i} |v_{i,j}(l)|^2}}
\]
Where $S(p, j)$ represents the similarity of $j^{th}$ feature vector of input face ($v_i, j$) to $p^{th}$ feature vector of $i^{th}$ reference face, $(v_i, p)$, where $l$ the no of vector elements.

We chose the greatest similarity value (i.e. nearest to one) of a feature vector of input image with all the feature vectors of any image from the database, as it determines the highest degree of similarity between two feature vectors.

$$D(T, I) = \min \{ |S(i, 1) - 1|, |S(i, 2) - 1|, \ldots, |S(i, z) - 1| \}$$

Where $D(T, I)$ is the difference of $i^{th}$ feature vector of input image $T$ with all the feature vectors of image $I$ of the respective degree from the database where $z = n \times W$.

Finally, we calculate the overall difference $D(T, I)$ between an input image and an image from data base by using following equation:

$$D(T, I) = \frac{1}{z} \sum_{i=1}^{z} D(T, I)$$

7. Architecture of system

The proposed architecture consists of four main processing modules a) Data fusion b) feature value calculation c) Feature vector selection d) Similarity calculation. Figure 3 shows the block diagram of architecture of the system, there are few pre-processing and storage modules as well.

![Figure 3](image)

8. Results

Experiments are performed on 24 candidates from Equinox facial database [14] of images. Three images (thermal, visual, and fused) of each candidate with frontal illumination are used. In our experimentations we have value of $F_w$ to 0.5. Accuracy results achieved for two different resolution images provided in Table 1.

<table>
<thead>
<tr>
<th>Input type</th>
<th>Accuracy of Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>21x30</td>
<td>38%</td>
</tr>
<tr>
<td>37x49</td>
<td>38.6%</td>
</tr>
<tr>
<td>Thermal</td>
<td>63%</td>
</tr>
<tr>
<td>Visual</td>
<td>74.38%</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>96%</td>
</tr>
<tr>
<td>97.4%</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Performance of Gabor filter

The above table shows that the accuracy of recognition system over fused images is greater then the accuracy of system over each visual and thermal image individually.
9. Conclusion

As Recognition algorithms are well developed for visual images, in this paper a well developed Gabor filter is employed on the fused data of visual and thermal images for face recognition. Because only test visual and thermal images are stored in memory, so memory requirement is very less. Only feature vectors achieved by applying Gabor filter on the fused data are stored in database, hence further reducing the memory requirement. Our system design has almost the same computations as the individual computational cost of visual or thermal face recognition systems, but at the same time the efficiency of designed fusion architecture is more then the individual visual and thermal face recognition.

10. Acknowledgement

We are thankful to COMSATS Institute of IT, HEC and PTCL Pakistan, for providing us a research environment, development facilities and travel grant support to attend conferences.

11. References