Multi-Spectral Hybrid Invariant Moments Fusion Technique for Face Identification

Shaymaa Hamandi, Abdul Monem Rahma, and Rehab Hassan
Computer Science Department, University of Technology, Iraq

Abstract: For reliable face identification, the fusion process of multi-spectral vision features produces robust classification systems. This paper exploits the power of thermal facial image invariant moments features fused with the visible facial image invariant moments features to propose a new multi-spectral hybrid invariant moment fusion system for face identification. And employs Feed-forward neural network to train the moments' features and make decisions. The evaluation system uses databases of visible thermal pairs face images CARL and UTK-IRIS databases and gives an accuracy reaches 99%.

Keywords: Feed-forward neural network, affine moments, face recognition, invariant moments, shape descriptor, zernike moments.

Received February 21, 2021; accepted March 7, 2021
https://doi.org/10.34028/iajit/18/3A/3

1. Introduction

Face identification is one of the common and strong human biometric authentication methods among other authentication techniques such as fingerprinting, iris scanning, and voice recognition [8, 27].

Most of the current face identification systems use the visible spectrum, which corresponds to the wavelength band from 0.38 μm to 0.78 μm. The use of this spectrum suffers from some limitations, such as light changing, facial expression and pose variations. In the last years, many researches in biommetry have focused their interest on the Infrared (IR) spectrum, for the advantages that offers. It is more reliable and robust in many aspects, compared to the visible spectrum [2].

The IR spectrum gives the easiest and more robust solution to improve the identification performance. But IR image has some limitations like it is sensitive to temperature variations of the surrounding environment and changes in the heat patterns of the face. In contrast to thermal image visual image is more robust to the above factors but very sensitive to illumination variations. This implies that efficient algorithms to fuse data from both visual and thermal spectra to enhance face identification performance [3, 14].

Robust facial feature extraction is an effective and important process for face recognition and identification system. The facial features should be invariant to scaling, translation, illumination and rotation, several feature extraction techniques may be used to increase the recognition accuracy. Recently, the most utilized descriptors are the moment invariants which is "The first choice descriptors" for making an estimate of the implementation about different descriptor types [17].

This paper is a revised and expanded version of a conference paper [16] which investigated three different moment invariants techniques for robust facial features extraction and then determined how each one of these moments was affected by whether the face image was thermal or on a greyscale with the proposal of a hybrid technique that dealt with the robust descriptors of each method. This hybrid technology has improved the results and gave robust facial features for face identification. In this paper, the robust hybrid descriptors proposed in the previous study for both thermal and visible imaging will be fused to produce a new reliable and more robust descriptor.

The rest of this paper is organized as follows: section 2 reviews the related works, section 3 describes Thermal and Visible Grey Imaging, section 4 illustrates the details of the face shape moments calculation, the robust feature selection is illustrated in section 5. section 6 describes the adopted feed-forward neural network technique, the proposed work and algorithms are described in section 7, experimental results and discussions presented in section 8. Finally, section 9 summarizes the paper's conclusion and future work.

2. Related Works

There are many fusion methods for face identification, a lot of studies have been published to enhance the performance of the regular existent face biometric methods are discussed here and summarized in Table 1.

Viswanathan et al. [32] have developed an algorithm that was based on the invariant moment features of face and fingerprint which was fused using a variation. In this algorithm the fingerprint and face were segmented and the invariant moment features were obtained. The invariants were fused into a single recognition value by using a coefficient variance.
Experimental tests were conducted on the EURECOM VIS face database, which fully integrates the features extracted from visible light and IR channels. Through experiments on the KAIST dataset, it was proved that the MSFFN model was superior to a number of state-of-the-art multispectral pedestrian detectors methods in accuracy and speed but only for small input images.

In this paper, we propose a new multi-spectral fusion system for face identification that uses the proposed hybrid feature selection that discussed in the previous conference paper [16] by combining only the robust moment invariants descriptors.

### 3. Thermal and Visible Grey Imaging

Although visual images have been traditionally used in automatic facial recognition systems, researchers have recently been interested in using other parts of the electromagnetic spectrum as is shown in Figure 1 such as long-wavelength IR because of their special properties which depending on temperature and emission properties, emits different ranges of IR energy.

![Electromagnetic spectrum map](Image 308x325 to 564x494)

Figure 1. Electromagnetic spectrum map.

Thus, thermal image acquisition is possible even in completely dark conditions, where the formation of the visual image is impossible [2]. Thus, in night security surveillance systems, using thermal images is practical, where traditional face identification systems have failed [25]. Table 2 lists the face recognition problems that can be solved or cannot using visible or thermal images collected and organized from these researches [5, 19, 21, 24, 29] which presented in the previous study [16].

The fourth column is added to this study to present the face recognition problems which solved by the visible-thermal fusion process.

<table>
<thead>
<tr>
<th>Authors and Year</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viswanathan et al. [32] 2010</td>
<td>invariant moment features of face and fingerprint which was fused using a variation</td>
</tr>
<tr>
<td>Jasiński and Forczmański [19] 2016</td>
<td>created a dataset of face images captured in both thermal and visual visions. Many image fusion methods have been performed to merge visual and thermal features for use in face identification algorithms</td>
</tr>
<tr>
<td>Pal and Singha [27] 2017</td>
<td>fusion depended on the highest assortment of DB4 wavelet coefficients of visible and thermal images</td>
</tr>
<tr>
<td>Dong et al. [7] 2019</td>
<td>Visible-thermal fused face identification method utilizing non-linear hashing</td>
</tr>
<tr>
<td>Benamara et al. [2] 2019</td>
<td>IR and visible spectra with an optimization based on a feature selection algorithm and the PSO-SVM</td>
</tr>
<tr>
<td>Song et al. [30] 2020</td>
<td>multispectral feature fusion network (MSFFN) for pedestrian detection</td>
</tr>
</tbody>
</table>

Jasiński and Forczmański [19] have created a dataset of face images captured in both thermal and visual visions. Many image fusion methods have been performed to merge visual and thermal features for use in face identification algorithms.

Pal and Singha [27] have presented fusion depended on the highest assortment of wavelet coefficients of visible and thermal images. The dimensions of the fused images are minimized and the fewer dimension fused images were classified with the aid of the support vector machine method. The test outcome of fusion gave almost 2% and 6% higher accuracy than the visible and thermal images respectively for IRIS database.

Dong et al. [7] have proposed a visible-thermal fused face identification method utilizing non-linear hashing. To obtain features from both visible and thermal face images, a deep neural network design pre-trained by visual images, the accuracy of various scenarios has experimented on the EURECOM VIS-TH face database.

Benamara et al. [2] have proposed a multispectral framework that used both IR and visible spectra with an optimization based on a feature selection algorithm and the support vector machine. Experimental tests were conducted on IRIS OTCBVS Thermal/Visible and CSIST Lab2 databases.

Song et al. [30] have presented a multispectral feature fusion network (MSFFN) for pedestrian detection, which fully integrates the features extracted
Table 2. Face recognition problems which solved by the visible, thermal and fusion imaging.

<table>
<thead>
<tr>
<th>Face Recognition Problems</th>
<th>Solved using Grey Face Images Illumination</th>
<th>Solved using Thermal Face Images Illumination</th>
<th>Solved using Thermal - visible Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose Variation</td>
<td>√</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Variation in Expression</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>Different Scaling</td>
<td>√</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Disguises</td>
<td>×</td>
<td>(except the use of glasses)</td>
<td>√</td>
</tr>
<tr>
<td>Aging Problem</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Variation in Temperature</td>
<td>√</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>locating facial features and parts</td>
<td>√</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>Work well at dark</td>
<td>×</td>
<td>×</td>
<td>√</td>
</tr>
<tr>
<td>work through smoke and fog</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Differentiate twin brothers</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>preservation against spoofing efforts</td>
<td>×</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

4. Face Shape Moments Calculation

Moments described as scalar quantities that are used for characterizing functions and for computing their different features. Moments are very useful, since calculating them is algorithmically easy and clearly adjusted for any of the picture functions [5, 11, 12, 21].

Invariance to translations is done by object moving to a specific position before the computing of moments, then each moment has the translation invariance. Gaining “a central geometric moments” as in \( \mu_{pq} \) in (1) for a 2-D face image \( f(x,y) \) and \( (p+q) \) order moment [10].

\[
\mu_{pq} = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} (x-x_e)^p (y-y_e)^q \ast f(x,y)
\]  

(1)

Where \( x_e \) and \( y_e \) are the dimensions of the face centroid. Equation (1) represents the base of the following progresses in the moment descriptors types as illustrated in Figure 2 [15].

From each type of the moments descriptors this study picks one sub type for the experiments on thermal and visible face images, the chosen types are [15]:

4.1. Geometric Hu Moments Computation

Flusser et al. [11], has investigated the algebraic invariants theory, relying on normalized central moments Equation in (1) and derived his seven famous invariants to rotations around the origin and the seven equations are computed.

In spite of these seven characteristics were represented to be valuable, they’re invariant just to translation, scaling, and rotation [31].

4.2. Algebraic Affine Moment Computation

The first researcher who considered the invariants to affine transform was the well-known mathematician Celebi and Aslandogan [6], they depended on (1) and the affine moments’ equations are computed [1]:

4.3. Orthogonal Moments Calculation

The Zernike Moments (ZM) calculation has been derived by Teague [26] who has used ZM for the computation of the rotation invariants [18]. First, mapping the images to a disk of a unit before computing ZM. Then, the ZM will be computed.

5. Robust Feature Selection

The purpose of feature selection is to gain the best subset of features that gives the optimal achievements. Sequential Forward Feature Selection (SFFS) technique is utilized to choose the optimal feature combination which gives optimal accuracy. SFFS is done on the features which give the highest discrimination ranks to minimize the search time and guarantying only optimal features will be found in the latest feature subset and this process will minimize the number of used features [23]. SFFS calculation is starts from an unfilled list and continuously add highlights chose by some assessment work, that limits the Mean Square Error (MSE). At every emphasis, the component to be remembered for the list of capabilities is chosen among the leftover accessible highlights of the list of capabilities, which have not been added to the list of capabilities, the new broadened highlights set should create a base order mistake contrasted and the expansion of some other element.

6. Feed-Forward Neural Network Technique

Classification is the procedure for discriminating class data from other data sources in the feature space. A feed-forward neural network is an artificial neural
network wherein connections between the nodes do not form a cycle [20], in this network as shown in Figure 3, the data moves in one direction, only beginning with the input nodes, through the hidden nodes ending with the output nodes. There are no cycles in the network [28].

Figure 3. Feedforward neural network design.

7. The Proposed Work

Different types of features have different intensity of description. For improving that power, various kinds of features have been used together rather than only one kind. This study utilizes the selected hybrid moments from both the visible and the thermal face images and fuse them to improve the feature description power and benefits from the advantages of each type and overcome the drawback of them. The process of the proposed technique is shown in Figure 4.

The proposed method includes seven main stages:

- **Image Preprocessing Stage:**
  - Detect face area using Viola-Jones method and then crop the face area for both the visible face images and thermal face images.
  - Converting from RGB to a greyscale image.
  - Threshold the greyscale image to have either 0 or 1 value for each pixel in the image.

- **Moment Invariants Computation Stage:**
  - Hu, Affine, and Zernike moments equations are calculated for each visible and thermal face image.

- **Moment Invariants Evaluation Stage:**
  - Computing each moment feature intra-class (within the class) and inter-class (between-discrimination between classes).
  - The features are passed to the next stage of the robust moments selection stage.

- **Robust Moments Selection Stage:**
  - SFFS technique is used for choosing the robust combination of features that gives the best accuracy.

- **Fusion Stage:**
  - The robust features for both visible and thermal face images fused together.

- **Classification Stage:**
  - Feed-forward neural network as in Figure 7 is used as a classification model.
  - Feed-forward neural network is used to classify the resulted best feature subset into an appropriate class.
  - The network is trained with the moments’ features to classify the inputted feature vector into one of the basic faces class.

- **Face recognition evaluation Stage:**
  - The test images are used at this stage to examine the robustness of the proposed system.

8. Experimental Results and Discussion

The performance of each moment descriptor for thermal and grey images will be checked using two datasets. Iris thermal/visible face database [13] which contains 4228 pairs of 320×240-pixel images. The images are acquired under illumination, expressivity and pose variation as illustrated in Figure 5.

Figure 4. The basic stages of the proposed hybrid moments descriptor fusion.

Figure 5. Some samples of the IRIS database.
The second dataset is the CARL Database [9]. Each image is (160×120) pixel. CARL database consists of thermal, visible and near IR images of 41 subjects (32 males and 9 females) in various illumination conditions as is shown in Figure 6.

a) Visible image.  
b) Near IR image.  
c) Thermal image.

d) Visible image.  
e) Near IR image.  
f) Thermal image.

Figure 6. Some samples of the CARL dataset. (section(a) is the first subject's visible picture, (b) first subject's in NIR vision, and (c) in thermal vision) and (d, e, f are the visions for the second subject)

8.1. Moments Computation and Fusion

The conducted results of the proposed descriptor for each thermal and grey image are presented in this section besides all results of moments descriptors for thermal and grey faces from the CARL face dataset and IRIS face dataset which is used in the experiments. The datasets divided into three parts 70% of the data for the training part, 15% of the data for the testing part and 15% of the data for the validation part. Each dataset is entered through the first five stages of the proposed system to compute the robust moments invariants for each face image.

8.2. Neural Network Classifier

A feed-forward neural network classifier was applied to build a suitable neural network structure and update neural network weights.

The input to the neural network in this experiment was the hybrid robust features for both thermal and visible face images. The neural network was with one hidden layer containing 5 neurons and the output layer with ten neurons which represents the ten classes of the persons in the dataset as is depicted in Figure 7.

Figure 7. The structure of the Feed-forward neural network classifier.

Figure 8 plots error vs. epoch for the training, validation, and test performances of the training in the Feed-forward neural network classifier for CARL thermal-visible faces image fusions.

Figure 8. Error vs. epoch for the training, validation, and test performances for the Feed-forward neural network classifier when applied to the CARL thermal-visible faces image fusion.

Figure 9 shows the results for the Feed-forward neural network classifier when applied to the IRIS thermal-visible hybrid descriptor fusion using only the robust features.

Figure 9. Error vs. epoch for the training, validation, and test performances for the Feed-forward neural network classifier when applied to the IRIS thermal-visible hybrid descriptor fusion.

Our proposed work as shown in Figure 10 is to take only the robust moment descriptors from the above results for thermal and grey face images and create a new fused descriptor that takes all the robustness of them and gives a better result in recognizing the shapes. For thermal face images the features (Hu3, Hu4, Hu5, Aff3, Z2, Z3, Z7, Z9) are chosen as the robust features and for visible (Hu3, Hu4, Hu5, Hu7, Aff1, Aff5, Aff7,
Z1, Z4, Z5, Z8) these features fused and be the input to the Feed-forward neural network classifier.

The results of the Feed-forward neural network classifier when applied to the IRIS thermal faces dataset with thermal-visible hybrid moments fusion for five faces only is shown on Figure 11 which gives a very good accuracy and distinctively between faces.

The difference in recognition accuracy results when applying different descriptors on the CARL and IRIS datasets are shown in Figure 12.

When comparing the proposed system results with different current researches that worked on IRIS database, it shows that an improvement of 1% is achieved on the accuracy of face recognition using the proposed method for thermal face image as shown in Table 5.
Table 5. A Comparison between current research works on IRIS database and the proposed hybrid fusion system.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methodology</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal and Singha [27]</td>
<td>Haar/maximum selection of DB4 wavelet coefficients</td>
<td>96.49%</td>
</tr>
<tr>
<td>Bhowmik et al. [4]</td>
<td>Pixel fusion using 70% visual and 30% thermal images + IPCA + SVM</td>
<td>97.28%</td>
</tr>
<tr>
<td>Malikarjunu et al. [22]</td>
<td>Local Active Pixels Pattern</td>
<td>83%</td>
</tr>
<tr>
<td>Kong et al. [21]</td>
<td>Ellipse fitting method</td>
<td>83.3%</td>
</tr>
<tr>
<td>Visswanathan et al. [32]</td>
<td>SIFT+invariant moment</td>
<td>70%</td>
</tr>
<tr>
<td>The proposed</td>
<td>Fused hybrid moment invariants</td>
<td>97.2%</td>
</tr>
</tbody>
</table>

9. Conclusions and Future Work

In this paper, a new face recognition technique is proposed. A Thermal-visible fusion of hybrid face descriptors is presented to enhance the results for identifying face images. Three kinds of moment shape descriptors are calculated and analyzed (Geometric Hu moments, Algebraic Affine moments, and Orthogonal Zernike moments) for the face images. The recognition accuracy is not quite enough if using the features of the thermal face image only or features of visible face images but when from each spectral only the robust descriptors were chosen and fused together to produce the proposed fused hybrid descriptor, the results are enhanced to reach identification accuracy of 99% in the proposed system which produces a very robust descriptor.

In future work, the features of multi-spectral visions will be fused and processed by deep learning to enhance the accuracy results.

References


Shaymaa Hamandi awarded her M.Sc. from Al Nahrain University, Computer science department in 2006. She worked at Baghdad Governorate as a manager of information technology department until 2011, programmed a lot of database systems for financial, Stock, and human resources departments. Then worked as a programmer in a public governorate library. Started her PhD studying in 2018 and now she is in the research stage. Her research interests include image processing, pattern recognition and biometrics.
Research Gate: https://www.researchgate.net/profile/Shaymaa_Hamandi
EMAIL: 111797@student.uotechnology.edu.iq.

Abdul Monem Rahma has an extensive background in Cryptography and Information Security Image Processing, Pattern Recognition and Biometrics. He received his PhD in Computer Science in 1984, from the Loughborough University of Technology in the United Kingdom, and become a professor in Computer Science since 2008. His main work experience involves teaching at Iraqi universities and supervising postgraduate students; He also was Deputy Dean of the Department of Computer Science, University of Technology, Iraq from 2005 to 2013. From 2013 to 2015, he became the Dean of the department. Now He is Lecturer and the head of the Department of Computer Engineering Techniques, Imam Ja’afar Al-Sadiq University. Prof. Rahma published 180 Papers, 4 Books in Computer Science; supervised 36 PhD and 66 M.Sc. students.
ORCID 0000-0001-6323-9148
PHONE:+9647712890216
Research Gate: https://www.researchgate.net/profile/Abdul_Monem_Rahma
Google Scholar: https://scholar.google.com/citations?user=63NQ8lcAA AJ&hl=en&oi=ao
EMAIL:110003@uotechnology.edu.iq

Rehab Hassan Awarded her M.Sc. and Ph.D. degree from University of Technology, computer science department in 1995 and 2005, respectively. She taught at the University of Technology, Computer Science department.
She published a lot of papers in the field of computer science and supervised Ph.D. and M.Sc. students. Her research interests include computer graphics image processing, and computer security.
Research Gate: https://www.researchgate.net/profile/Rehab_Hassan
EMAIL: 110019@uotechnology.edu.iq.