



Combined and Weighted Features for Robust Multispectral Face Recognition

Nadir Kamel Benamara¹(✉), Ehlem Zigh², Tarik Boudghene Stambouli¹,
and Mokhtar Keche¹

¹ Laboratory Signals and Images (LSI), University of Sciences
and Technology of Oran Mohamed Boudiaf (USTO-MB), Oran, Algeria
nadirkamel.benamara@univ-usto.dz

² Laboratory of Research in Applied ICT (LARATIC), National Institute
of Telecommunications and ICT of Oran (INTTIC), Oran, Algeria

Abstract. Face recognition has been very popular in recent years, for its advantages such as acceptance by the wide public and the price of cameras, which became more accessible. The majority of the current facial biometric systems use the visible spectrum, which suffers from some limitations, such as sensitivity to light changing, pose and facial expressions. The infrared spectrum is more relevant to facial biometric, for its advantages such as robustness to illumination change. In this paper, we propose two multispectral face recognition approaches that use both the visible and infrared spectra. We tested the new approaches with Uniform Local Binary Pattern (uLBP) as a local descriptor and Zernike Moments as a global descriptor on IRIS Thermal/Visible and CSIST Lab 2 databases. The experimental results clearly demonstrate the effectiveness of our multispectral face recognition system compared to a system that uses a single spectrum.

Keywords: Multispectral face recognition · Infrared
Uniform Local Binary Pattern · Zernike moments · Feature fusion

1 Introduction

Face recognition is a routine task for humans to identify persons in their everyday lives. Since 1960, automatic face recognition become an area that interest more and more researches in computer vision and biometric technologies [1]. Their applications are useful mainly for security purposes, such as access control, authentication systems and crime investigations. Compared to the other biometric modalities (fingerprint, iris and palmprint), face recognition has advantages of the ease of capturing subject samples without interacting with the person to identify and it is accepted by the wide public.

Face recognition, using visible spectrum (from $0.38\ \mu\text{m}$ to $0.78\ \mu\text{m}$), has been a major interest of researchers in recent years, with the evolution of cameras, which have become cheaper and more sophisticated. Several algorithms have been developed in this field, like eigenfaces [2], ficherfaces [3] and elastic bunch graph matching [4], but the visible images are vulnerable to light changing, poses, facial expressions and also disguise and faking. To overcome these limitations, multispectral face recognition has grown in interest for the advantages that offers, such as more discriminative features than those given by other spectra.

Infrared spectrum is divided into 2 parts, the active infrared, with Near Infra-Red (NIR) ($0.74\ \mu\text{m}$ – $1\ \mu\text{m}$) and short-wave IR ($1\ \mu\text{m}$ – $3\ \mu\text{m}$), and the passive infrared or thermal infrared, with middle-wave IR ($3\ \mu\text{m}$ – $5\ \mu\text{m}$) and the long-wave IR ($8\ \mu\text{m}$ – $12\ \mu\text{m}$). The infrared does not suffer from the limitations of the visible spectrum, mainly the light changing, However, these spectra face other limitations and challenges, due to face expressions and outdoor applications for NIR, dark face images caused by skin moisture absorption of infrared wavelength above $1.45\ \mu\text{m}$ in SWIR, glasses opacity and body metabolism (fever, sporting activity) that change the thermal image for MWIR and LWIR (thermal infrared).

Based on the advantages and disadvantages of each spectrum, it is possible to design a multispectral facial recognition biometric system with a better recognition rate than a system with a single spectrum modality.

To this effect, we propose, in this article, a face recognition method that combines the infrared and visible spectrum. The features from each spectrum are merged to obtain a more discriminative information.

This paper is organized as follows: In Sect. 2, we present some related work. Then, in Sect. 3, the two proposed multispectral face recognition approaches are described. In Sect. 4, we talk about the used benchmark databases for our experiments. Following, in Sect. 5, we present and discuss the obtained results. Finally, we draw some conclusions in Sect. 6.

2 Related Works

Many research works have been carried out, in order to ensure robust solutions for face recognition problems, by combining visible and infrared spectra. We can cite for example Kong et al. [5] who proposed a multiscale fusion of visible and thermal face images. They detected and replaced the glasses opacity in thermal infrared with an eye template, to improve the recognition performance. Buddharaju et al. [6] proposed a multispectral system based on score fusion between results of eigenspace matching on visible faces and those obtained by the physiology-based face recognition method in thermal infrared spectrum, introduced by [7–9]. Bhowmik et al. [10] presented the effect of infrared spectrum on the enhancement of recognition rate when it is fused with visible spectrum. Hermosilla et al. [11], proposed a multispectral face recognition system based on fusion of visible and thermal descriptors using a genetic algorithm. More recently, Guo et al. [12] proposed a deep network with an adaptive score fusion strategy for visible - near infrared face recognition.

3 Proposed Approaches

The proposed multispectral face recognition system comprises four stages: data base preprocessing, features extraction, features fusion and finally classification. Our contribution in this paper is related to the third stage, in which two features vectors are merged into a single characteristic vector containing the visible and thermal infrared information.

Two fusion approaches, named the features weighted average and the combined feature vectors, were tested. The flowcharts of these approaches are presented in Figs. 1 and 2, respectively.

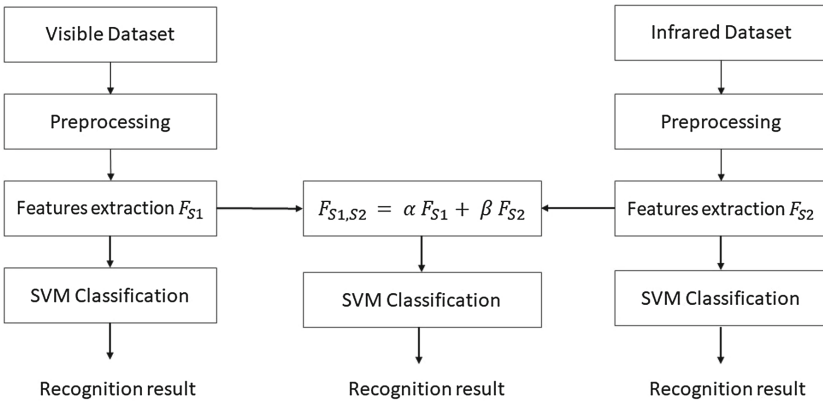


Fig. 1. Flowchart of the features weighted average approach

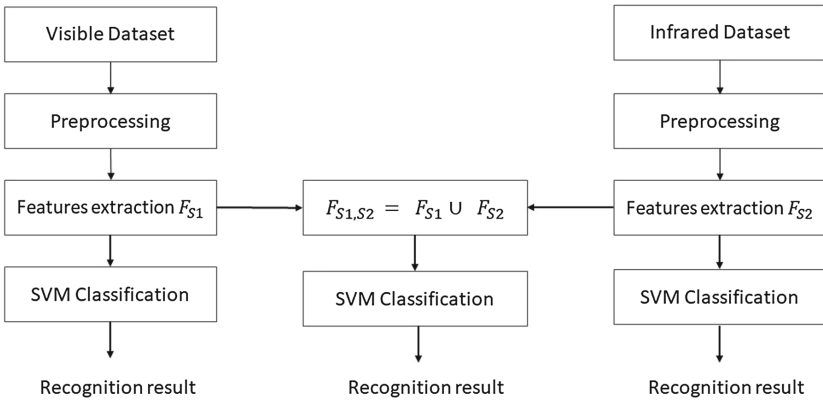


Fig. 2. Flowchart of the combined feature vectors approach

3.1 Database Preprocessing

The databases images are generally taken with a background, which it is useless for facial recognition. Therefore, the images must be cropped to keep only need the face of the person. Also, since the color information is not used, the cropped images should be converted from the color space to the grayscale one, as shown below in Fig. 3.

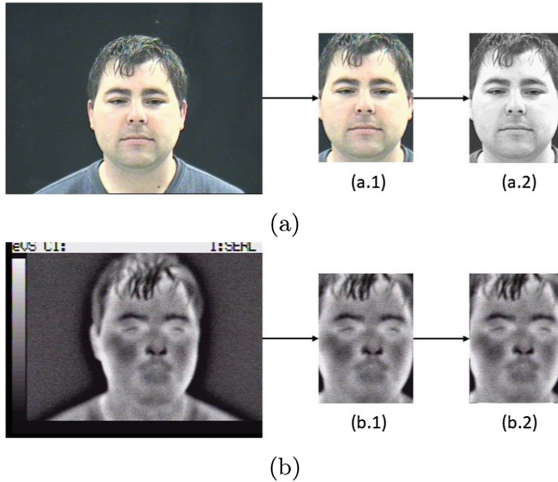


Fig. 3. Face preprocessing applied on an IRIS Thermal/Visible image. (a) Visible image (a.1) Visible cut image (a.2) Visible grayscale cut image. (b) Thermal image (b.1) Thermal cut image (b.2) Thermal grayscale cut image

3.2 Features Extraction

We have used two different descriptors: the uniform Local Binary Pattern (uLBP), as a local descriptor, and Zernike Moments (ZMs), as a global descriptor.

The Uniform Local Binary Pattern Feature. Local Binary Pattern was used for the first time in [13], for texture analysis. It proved its effectiveness in many image analysis applications, like biomedical, motion and biometric applications. Applied in face recognition for first time by Ahonen [14], the advantages of LBP are: the invariance to monotonic gray level changes or in otherwise illumination changing, the powerfulness for textural descriptions and the computational efficiency.

This local descriptor consists of the distribution (histogram) of the pixels, according to their neighborhoods. The neighborhood of a central pixel $P_0(x_c, y_c)$

is characterized by the pair (P, R) , where P is the set of points (pixels) located around P_0 , inside the circle of radius R . The coordinates of point P_i are given by:

$$P_i = (x_c + R \cos(\frac{2\pi i}{P}), y_c - R \sin(\frac{2\pi i}{P})) \tag{1}$$

As illustrated in Fig. 4, the value of the LBP pixel, for a neighborhood of $P = 8$ pixels, is calculated by the following formula:

$$I_{LBP}(P_0) = \sum_{i=1}^8 P'_i * 2^{i-1} \tag{2}$$

where

$$P'_i = \begin{cases} 1 & \text{for } I(P_i) \geq I(P_0) \\ 0 & \text{for } I(P_i) < I(P_0) \end{cases} \tag{3}$$

where $I(P)$ denotes the intensity, *i.e.* the gray level of pixel P .

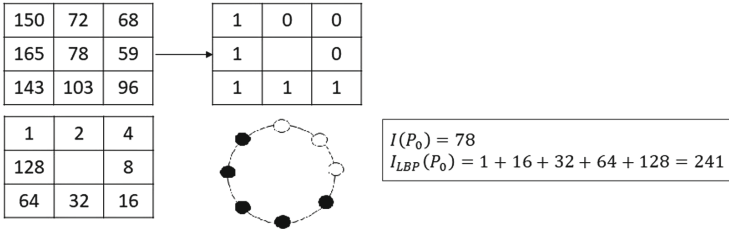


Fig. 4. LBP Calculation ($P = 8, R = 1$), for a circular binary pattern representation.

Uniform Local Binary Pattern was proposed in [15], where the smaller non-uniformity measure is described as the less likely pattern that undergoes unwanted changes, as rotation, the non-uniformity measure represents the number of transition in the circular bitwise LBP representation, for example: 00000100 and 11111000 have non-uniformity measure of 2, 0 and 255 (00000000 and 11111111) have a measure of 0. Other patterns have at least a non uniformity of 4. In [15], they selected nine uniform pattern that have non uniformity measure of at most 2 which are: 00000000, 00000001, 00000011, 00000111, 00001111, 00011111, 00111111, 01111111 and 11111111, these patterns and their circular rotated versions correspond to a subset of 58 patterns, from the original 256 patterns LBP set. The remaining patterns are accumulated in the 59th bin.

The resulting histogram has 59 bins could be expressed as a feature vector that describes the face image.

The Zernike Moments Feature. Zernike moments are a set of orthogonal polynomials that describes the whole image. Defined as a global descriptor,

they have very interesting properties [16,17], like orthogonality, which means less information redundancy, rotation invariance and high accuracy for detailed shapes. Zernike moments are calculated by:

$$Z_{n,m} = \frac{n+1}{\pi} \sum_{r \leq 1} \sum_{\theta \leq 2\pi} I(r, \theta) \cdot [V_{n,m}(r, \theta)]^* \tag{4}$$

where $I(r, \theta)$ is the representation of image in polar coordinates, $V_{n,m}(r, \theta)$ represents an orthogonal radial basis function, on which the image is projected, defined by:

$$V_{n,m}(r, \theta) = R_{n,m}(r) \cdot e^{jm\theta} \tag{5}$$

$R_{n,m}(r)$ is equal to:

$$R_{n,m}(r) = \begin{cases} \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s (n-s)!}{s! [\frac{1}{2}(n+m)-s]! [\frac{1}{2}(n-m)-s]!} \cdot r^{n-2s} & \text{for } (n-m) \text{ even} \\ 0 & \text{for } (n-m) \text{ odd} \end{cases} \tag{6}$$

and $n = 0, 1, 2, 3, \dots$, $m = 0, 1, 2, 3, \dots$ and $m \leq n$.

3.3 Features Fusion

Features fusion is the third stage of our multispectral face recognition system. The aim of this stage is to get a data that contains a combined information from different spectrums. Two different approaches are proposed, namely: the features weighted average and the combined feature vectors.

The Features Weighted Average Approach. In this approach, a new feature vector is obtained by a linear combination of the feature vectors, obtained from two different spectrums, the visible and the infrared ones:

$$F_{S1,S2} = \alpha F_{S1} + \beta F_{S2} \tag{7}$$

where F is a feature vector, $S1$ and $S2$ are two different spectra and α, β are the weights of each spectrum, with $\alpha + \beta = 1$.

The Combined Feature Vectors Approach. In the combined feature vector approach, the features vectors, from the visible and infrared spectrums, are concatenated to form a new feature vector:

$$F_{S1,S2} = F_{S1} \cup F_{S2} \tag{8}$$

3.4 Classification

The role of a classification algorithm is to assign a class to each input feature vector.

We have chosen, as a classifier, the well-known powerful one-versus-all multiclass Support Vector Machine (SVM) classifier, with a linear kernel.

4 The Used Face Databases

Two databases have been used to evaluate the performance of the proposed multispectral face recognition methods, the CSIST database and IRIS Thermal/Visible face database (Fig. 5).

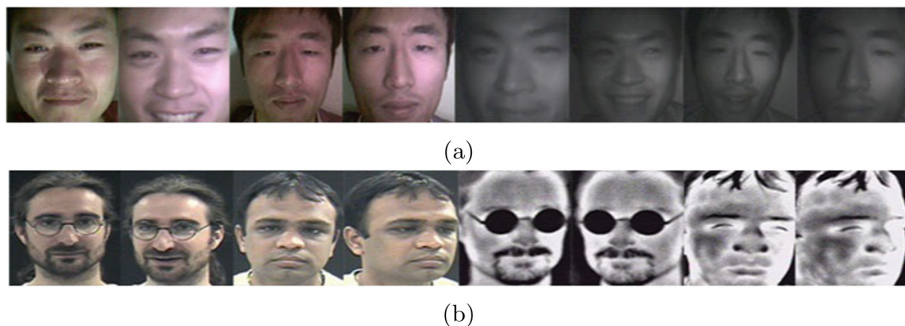


Fig. 5. Database samples: (a) CSIST Lab 2 (b) IRIS database

4.1 CSIST Database

CSIST Database was built by Harbin Institute of Technology Shenzhen Graduate School [18]. It contains two subsets of two different spectrum: Near Infrared & Visible spectrum. The first subset named Lab1 contains 1000 face images at a resolution of 100×80 , 500 images for each spectrum, of 50 different subjects (10 images for each subject), The second subset, Lab2, contains 2000 face images at a resolution of 200×200 , 1000 images for each spectrum, of 50 different subjects (20 images for each subject).

For our experiments, we have chosen the second subset Lab2 because the first subset, Lab1, does not contain an illumination change in its face images. The step of preprocessing will not be performed on CSIST Database, because face images in this data set are already cropped by its founder.

4.2 IRIS Thermal/Visible Face Dataset

IRIS Thermal/visible face dataset [19] is a public database comprising 4228 pairs 320×240 pixel visible and thermal face images of 30 individuals, with different poses, variable illumination conditions and different facial expressions.

In our experiment, we have selected a sub-set of 954 images, 477 images from each spectrum: front view, 2 different poses (right and left orientation), with presence and absence of light, and 3 different facial expressions, with different right and left orientations too.

5 Results and Discussion

To evaluate our proposed approaches, we have used two different descriptors: Uniform Local Binary Pattern, as a local descriptor, and Zernike Moments, as a global descriptor. The evaluation was performed on two different datasets, IRIS Thermal/Visible Database and CSIST Lab 2, as described above.

In classification stage, 70% of total samples were dedicated for learning and the remaining 30% for testing. All experiments were carried out with the same computer configuration which is Intel i3 5010u 2.10 Ghz with 4 GB RAM.

5.1 Results Obtained with the Features Weighted Average Approach

In order to get the features weighted average, we calculate features (uLBP or ZMs) from each spectrum and fuse them, according to Eq. (7), for different values of α . The classification results are presented in Figs. 6a and b, for the uLBP and ZMs, respectively and for both the IRIS and the CSIST Lab 2 datasets.

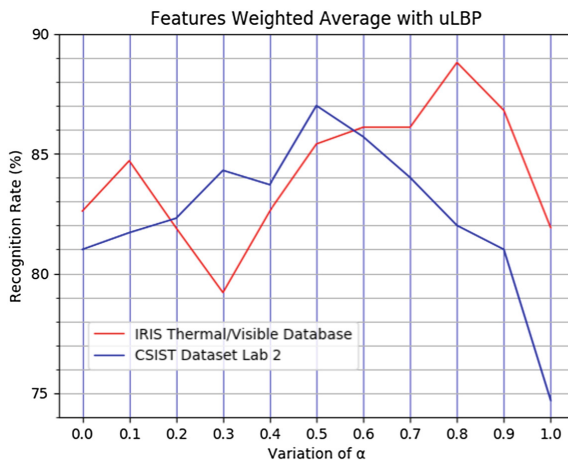
Results Obtained with the uLBP Local Features. For the IRIS Dataset, Fig. 6a clearly shows that the features weighted average method provides a recognition rate that is significantly higher than those given by using the single spectrums. The highest 88.8% recognition rate was obtained with $\alpha = 0.8$. This rate is to be compared to the 82.6 % and 81.9% rates, obtained using the single thermal and visible spectra, respectively. We notice that the lowest performance was obtained with $\alpha = 0.3$, *i.e.* a heavy weighting in favor of thermal data. This is because third of the used subset contains face images with glasses, which rises the glasses opacity problem that the thermal spectrum suffers from.

The results obtained with CSIST Lab 2 Dataset confirm that combining the visible spectrum and the NIR spectrum features, using the features weighted average approach, gives better results than using these features separately. The highest recognition rate of 87%, obtained with $\alpha = 0.5$, is to be compared to the rates 81% and 74.7%, obtained, respectively, with the near infrared and the visible spectrums features.

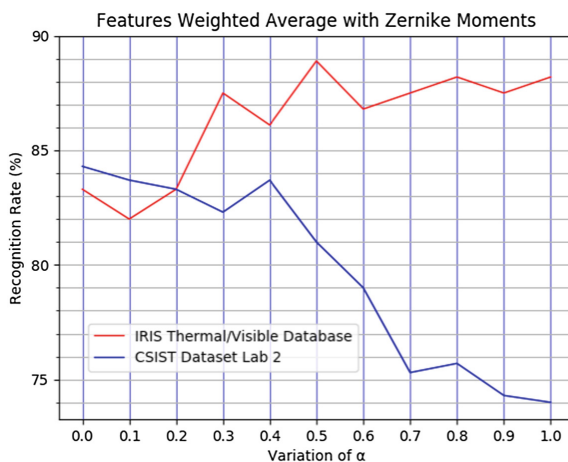
Results Obtained with the ZMs Global Descriptor. For the second evaluation of the features weighted average approach, the Zernike moments with a polynomial degree of order 10 were used, as a global descriptor. The classification results are presented in Fig. 6b.

With the IRIS Dataset, using the features weighted average approach, gives, for $\alpha = 0.5$, a 88.9% recognition rate that is better than the 83.3% rate, obtained with the thermal spectrum features, and slightly better than the 88.2% rate, obtained using the visible spectrum features only.

With the CSIST Lab 2 Dataset, the highest recognition rate, obtained by combining the features extracted from the visible and the near infrared spectra, using the features weighted average approach, with $\alpha = 0.4$, is 83.7%. It is higher



(a)



(b)

Fig. 6. Features weighted average results: (a) uLBP 2 (b) ZMs $n = 10$

than the 74% rate, obtained with the visible spectrum features, and slightly lower than the 84.3% rate, obtained by using the near infrared features alone.

The optimal value of α depends essentially on the conditions in which the database has been captured, for that reason we notice the difference between IRIS database and CSIST Lab2 dataset regarding the optimal weighting α and β .

5.2 Results obtained with the Combined Feature Vectors Approach

To evaluate the combined feature vectors approach, the obtained features from the visible and infrared spectra, using either the uLBP or the ZMs descriptors, were gathered in a unique features vector that was input to the SVM for classification.

Results Obtained with the uLBP Local Features. The results obtained by using the uLBP, for features extraction, are presented in Table 1, for both the IRIS and CSIST Lab2 databases. It can be seen from this table that for both of these databases, compared to using the features from the visible and infrared spectra separately, merging them improves significantly the performance.

Table 1. Combined feature vector using uLBP results

Database	Visible only	Infrared only	Combined feature vector
IRIS Thermal/Visible	81,9%	82.6%	91,7%
CSIST Lab2	74,7%	81%	94%

Results Obtained with the ZMs Global Features. For the second evaluation of combined feature vector, Zernike Moments of order 10 were used. The classification results are shown in Table 2. These results confirm those obtained with the uLBP local features.

Table 2. Combined feature vector using Zernike moments results (polynomial degree $n = 10$)

Database	Visible only	Infrared only	Combined feature vector
IRIS Thermal/Visible	88.2%	83.3%	97.2%
CSIST Lab2	74%	84.3%	91%

5.3 Discussion

We have noticed that the combined feature vectors proposed approach gives better results than the features weighted average approach for both databases, in terms of recognition rate, however it has a longer training phase as shown in Table 3. The reason of this result is that the first features fusion approach combines the totality of the visible and invisible vectors, whereas the second one takes just a percentage from each spectrum.

Concerning the near infrared-visible database, the local descriptor has a better recognition rate than the global one for both features fusion approaches. This can be explained by the robustness of the local descriptors, in comparison to the

Table 3. The computational learning time for the two proposed approaches

Approach	IRIS Thermal/Visible		CSIST Lab2	
	uLBP	Zernike moments	uLBP	Zernike moments
Features weighted average	31.3 ms	15.6 ms	78.1 ms	62.5 ms
Combined feature vectors	62.5 ms	31.3 ms	125 ms	109.4 ms

global descriptors, to variations, like illumination changes, which is the case for the CSIST Lab 2 database that contains an important light changing in face images.

Regarding the IRIS thermal/visible database, the global descriptor has a better performance than the local one, for both proposed approaches. This result is due to the rotation invariance property of the Zernike moments, which performs well for the slightly rotated faces, present in this database.

To summarize, our proposed approaches perform well in presence of illumination changes and slight variations of poses (rotated faces), which could be ideal for building a robust multispectral face recognition system.

6 Conclusion

In this paper, a features weighted average and a combined feature vectors approaches were proposed and applied for multispectral face recognition. The features from both the visible and the infrared spectra were extracted by using either the local uLBP descriptor or the global Zernike moments descriptor. These features were combined, by using one of the above mentioned approaches, and then input to a SVM classifier. The results obtained by fusing the features from the visible and invisible spectra were compared to those obtained by using the features from these spectra singly. The comparison shows that fusing the features from the visible and infrared spectra improves the performance. It also shows that the fusion by the combined feature vectors approach is better than the fusion by features weighted average approach.

For future work, a more elaborated fusion method and an other SVM kernels will be applied for a further improvement of classification rate.

References

1. Li, S.Z., Jain, A.K. (eds.): Handbook of Face Recognition, 2nd edn. Springer, London (2011). <https://doi.org/10.1007/978-0-85729-932-1>
2. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
3. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997)

4. Wiskott, L., Fellous, J.M., Kuiger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 775–779 (1997)
5. Kong, S.G., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B.R., Koschan, A., Yi, M., Abidi, M.A.: Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. *Int. J. Comput. Vis.* **71**(2), 215–233 (2007)
6. Buddharaju, P., Pavlidis, I.: Multispectral face recognition: fusion of visual imagery with physiological information. In: Hammoud, R.I., Abidi, B.R., Abidi, M.A. (eds.) *Face Biometrics for Personal Identification*, pp. 91–108. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-49346-4_7
7. Buddharaju, P., Pavlidis, I., Tsiamyrtzis, P.: Physiology-based face recognition, pp. 354–359. *IEEE* (2005)
8. Buddharaju, P., Pavlidis, I., Tsiamyrtzis, P.: Pose-invariant physiological face recognition in the thermal infrared spectrum, pp. 53–53. *IEEE* (2006)
9. Buddharaju, P., Pavlidis, I.T., Tsiamyrtzis, P., Bazakos, M.: Physiology-based face recognition in the thermal infrared spectrum. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 613–626 (2007)
10. Kanti Bhowmik, M., Saha, K., Majumder, S., Majumder, G., Saha, A., Nath, A., Bhattacharjee, D., Basu, D.K., Nasipuri, M.: Thermal infrared face recognition a biometric identification technique for robust security system. In: Corcoran, P., (ed.) *Reviews, Refinements and New Ideas in Face Recognition*. InTech, July 2011
11. Hermosilla, G., Gallardo, F., Farias, G., Martin, C.: Fusion of visible and thermal descriptors using genetic algorithms for face recognition systems. *Sensors* **15**(8), 17944–17962 (2015)
12. Guo, K., Wu, S., Xu, Y.: Face recognition using both visible light image and near-infrared image and a deep network. *CAAI Trans. Intell. Technol.* **2**(1), 39–47 (2017)
13. Ojala, T., Pietikinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* **29**(1), 51–59 (1996)
14. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J. (eds.) *ECCV 2004*. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24670-1_36
15. Topi, M., Timo, O., Matti, P., Maricor, S.: Robust texture classification by subsets of local binary patterns, vol. 3, pp. 935–938. *IEEE Computer Society* (2000)
16. Teh, C.H., Chin, R.: On image analysis by the methods of moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(4), 496–513 (1988)
17. Singh, C., Mittal, N., Walia, E.: Face recognition using Zernike and complex Zernike moment features. *Pattern Recogn. Image Anal.* **21**(1), 71–81 (2011)
18. Xu, Y.: Bimodal biometrics based on a representation and recognition approach. *Opt. Eng.* **50**(3), 037202 (2011)
19. OTCBVS: IRIS Thermal/Visible Face Database. <http://vcipl-okstate.org/pbvs/bench/>. Accessed 17 Feb 2017