Extraction of Illumination-Invariant Features in Face Recognition by Empirical Mode Decomposition

Dan Zhang and Yuan Yan Tang

Department of Computer Science,
Hong Kong Baptist University,
Hong Kong SAR, China
{dzhang,yytang}@comp.hkbu.edu.hk

Abstract. Two Empirical Mode Decomposition (EMD) based face recognition schemes are proposed in this paper to address variant illumination problem. EMD is a data-driven analysis method for nonlinear and non-stationary signals. It decomposes signals into a set of Intrinsic Mode Functions (IMFs) that containing multiscale features. The features are representative and especially efficient in capturing high-frequency information. The advantages of EMD accord well with the requirements of face recognition under variant illuminations. Earlier studies show that only the low-frequency component is sensitive to illumination changes, it indicates that the corresponding high-frequency components are more robust to the illumination changes. Therefore, two face recognition schemes based on the IMFs are generated. One is using the high-frequency IMFs directly for classification. The other one is based on the synthesized face images fused by high-frequency IMFs. The experimental results on the PIE database verify the efficiency of the proposed methods.

Keywords: Empirical Mode Decomposition, Face recognition.

1 Introduction

Empirical mode decomposition (EMD) is an adaptive signal analysis method for nonlinear and non-stationary data. It has been originally introduced by Huang et al. in [1], which is the first step of Hilbert-Huang transform [2,3]. Essentially EMD decomposes a signal into a set of Intrinsic Mode Functions (IMFs), which recover the original input signal features in a multiscale sense from high frequency to low frequency. The IMFs heavily rely on the original signal rather than using predetermined filter or wavelet functions. Due to the non-parametric and data-driven advantages, EMD has been successfully applied in one dimensional signal analysis such as ocean waves, rogue water waves, sound analysis, and earthquake time records. Recently it has been more and more applied in image analysis. The first track can be retrospected to [4], in which S. Long et al. applied EMD to analyze digital slope images. The image data has been expressed in terms of an array of rows and columns, then the EMD is applied to these

arrays row by row. Linderhed et al. also adopted this row-by-row EMD method in the fusion of visible and infrared images [5,6]. The input images are vectorized in lexicographical order and EMD is performed on each channel vector separately. R. Bhagavatula et al. [7] addressed face recognition in the same way.

The EMD has been extended for two dimensional data in the literatures toward texture extraction [8] and image compression [10]. The two dimensional EMD (2DEMD) has a more wide application in various kinds of image analysis, such as image fusion [12], image compression [11], texture analysis [9], feature extraction [14], rainfall analysis [13], watermarking [15], and temperature variation [16]. Faces as two dimensional digital signal are generally processed by wavelets or other filters. However, these methods are not adaptive to the original faces and always affected by the predetermined functions. The facial features captured in this way sometimes are anamorphic, which always leads to low recognition rate. Though EMD based methods can capture the intrinsic feature of the iamges, few work done in face recognition.
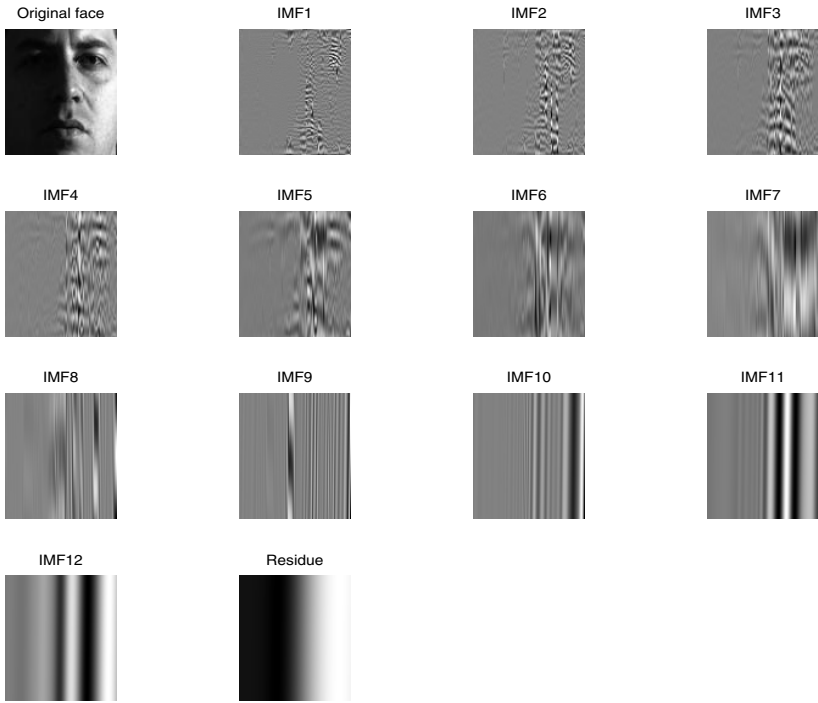
Face recognition refers to technologies for human authentication or verification based on human facial characteristics. Though this biometric technique has been improved a lot, however, many present facial biometric systems are still inefficient when identifying the same person under varying illumination. Earlier studies [17,18,19] show that only the low-frequency component is sensitive to illumination changes. It indicates that the corresponding high-frequency components are more robust to the illumination changes. Therefore, the high-frequency components alone without the low-frequency components are sufficient for classification. In terms of EMD, due to its adaptiveness it is able to capture more representative features and especially more singular information in high-frequency IMFs. In this paper, two available face recognition schemes based on EMD are generated. One is using the high-frequency IMFs directly for classification. The other one is based on the synthesized face images fused by high-frequency IMFs. The proposed two methods efficiently address the illumination variance problem. The experimental results on the PIE database verify their efficiency.

The paper is organized as follows. Section 2 presents an overview of the EMD and its details of application in image. Section 3 presents the flow of the proposed two face recognition schemes. The experiments are demonstrated in Section 4. Finally, conclusions and future work outlook are presented.

2 Empirical Mode Decomposition and Its Applications in Images

EMD is a signal analysis technique for adaptive representation of non-stationary signals as sum of a set of IMFs. It captures information about local trends in the signal by measuring oscillations, which can be quantized by a local high frequency or a local low frequency corresponding to finest detail and coarsest content. Here we briefly introduce the sifting process of EMD, four main steps are contained. S_1, S_2, S_3 and S_4 are abbreviations of Step 1, Step 2, Step 3 and Step 4 respectively. Given a signal x(t),
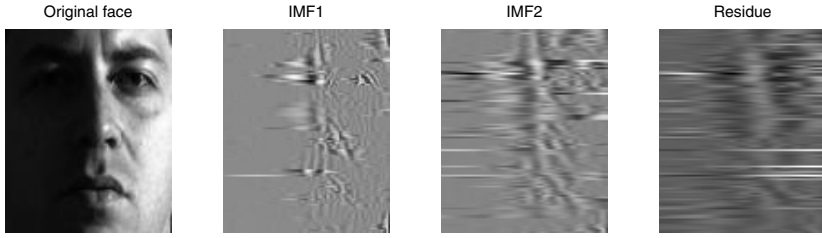
S_1. Identify all the local minima and maxima of the input signals $x(t)$;

S_2. Interpolate between all minima and maxima to yield two corresponding envelopes $E_{max}(t)$ and $E_{min}(t)$. Calculate the mean envelope $m(t) = (E_{max}(t) + E_{min}(t))/2$;

S_3. Compute the residue $h(t) = x(t) - m(t)$. If it is less than the threshold predefined then it becomes the first IMF, go to S_4. Otherwise, repeat S_1 and S_2 using the residue $h(t)$, until the latest residue meets the threshold and turns to be an IMF;

S_4. Input the residue r(t) to the loop from S_1 to S_3 to get the next remained IMFs until it can not be decomposed further.

Fig. 1. Twelve IMFs and the residue decomposed by 1DEMDa on a face image with right point illumination from the PIE database. The IMFs captures the information of the original face from the finest to the coarsest. It is clear in this figure that the right illumination effects focus in the residue.

There are three key points in the algorithm implementation.

- Judge whether the residue signal $h(t)$ is an IMF or not. Theoretically, the judgement should accord two conditions [1]. First, the number of extrema and the number of zero-crossing must be at most differ by one. Second, the mean envelops obtained by the maximum envelop and the minimum envelop must equal to zero. However, it is difficult to achieve these two conditions
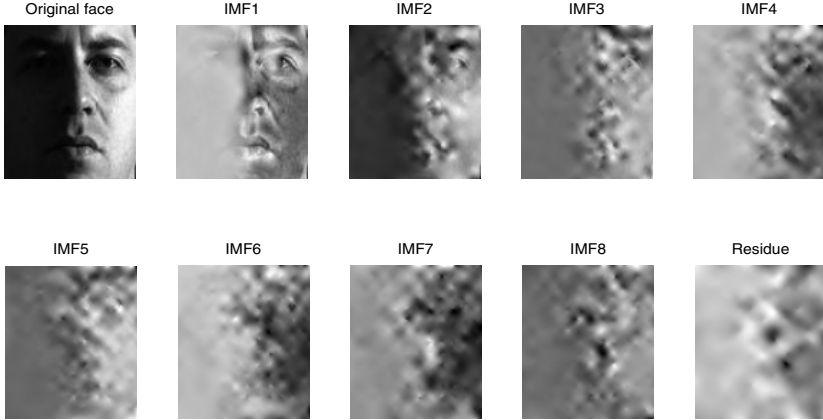
Fig. 2. Two IMFs and the residue decomposed by 1DEMDb on a face image with right point illumination from the PIE database

    strictly in the sense of realistic implementation. Researchers always define a small threshold instead. If $h(t)$ is small enough then it is an IMF.

- Choose an appropriate interpolation functions, such as linear interpolation, cubic spline interpolation and so on.
- Stop criterion. How to judge the residue can be decomposed further or not? Generally, the decomposition stops when the residue becomes so small that it is less than a predefined threshold.

    The EMD is originally proposed for one dimensional data. However, it has also been widely applied in image analysis. The idea is to express images as an array of one dimensional vectors. Generally two ways are adopted. One is to view the image as a long lengthened vector and then apply EMD on it. Another one is to apply on each row or column one by one as literatures [4,5,6,7] did. In order to discriminate these two methods, we name the former one as 1DEMDa and the latter one as 1DEMDb. Fig.1 and Fig.2 show all the IMFs and residues obtained by these two methods respectively. The IMFs captures the finest to the coarse information of the original face. The illumination effects are generally distributed in a few low level IMFs and the residues. Fig.1 shows an example. It is obvious that the right light effects mainly focus on the residue. By removing these low-frequency IMFs and residues, the remained high-frequency IMFs are less sensitive to the illumination effects, themselves alone are sufficient for classification. The extended 2DEMD has the similar sifting process as 1DEMD. Given the digital image $I = f(x, y)$, $x = 1, \cdots, M$, $y = 1, \cdots, N$, the process is summarized as follows.

S_1. Identify the extrema (both maxima and minima) of the image $I = f(x, y)$;

S_2. Generate the 2D envelopes by connecting maxima points (respectively, minima points) using surface interpolation. Determine the local mean m by averaging the two envelopes;

S_3. Subtract out the mean from the image to get an residue $h = I - m$, judge whether h is an IMF, if it is, go to S_4. Otherwise, repeat S_1 and S_2 using the residue h, until the latest residue turns to be an IMF;

| Original face | IMF1 | IMF2 | IMF3 | IMF4 |
| IMF5 | IMF6 | IMF7 | IMF8 | Residue |

Fig. 3. Eight IMFs and the residue decomposed by 2DEMD on a face image with right point illumination from the PIE database

S_4. Input the residue h to the loop from S_1 to S_3 to get the next remained IMFs until it can not be decomposed further.

In the implementation of two dimensional case, the Stop Criterion (SD) should be defined to judge whether residue h is an IMF and when to stop the loops:

$$SD = \sum_{k=0}^{m} \sum_{i=0}^{n} [\frac{|h_{i(j-1)}(k,l) - h_{ij}(k,l)|^2}{h_{i(j-1)}(k,l)^2}].$$

Different with 1DEMD, the extrema detection and the surface interpolation seem more complicated. Choosing appropriate extrema detection method and surface interpolation functions are more important, otherwise, poor decomposition results are got. More details can be seen in [13]. Fig.3 demonstrates the decomposition results by the 2DEMD. The IMFs are from finest to coarsest as well, and the illumination effects focus on the residue.

3   Face Recognition Schemes

As described in the former sections, two face recognition schemes based on EMD are generated. One is using the high-frequency IMFs directly as facial features. The other one is based on the synthesized face images fused by high-frequency IMFs. Fig.4 and Fig.5 show the flowcharts of two methods respectively. In our implementation, the following four items should be illustrated.

- Down-sampling is adopted twice before the EMD process. It aims to decrease the dimension of the faces.
- We choose the 1st IMFs to represent the facial features since the 1st IMF is the highest-frequency component.
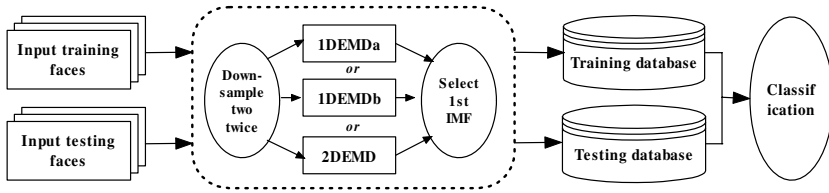
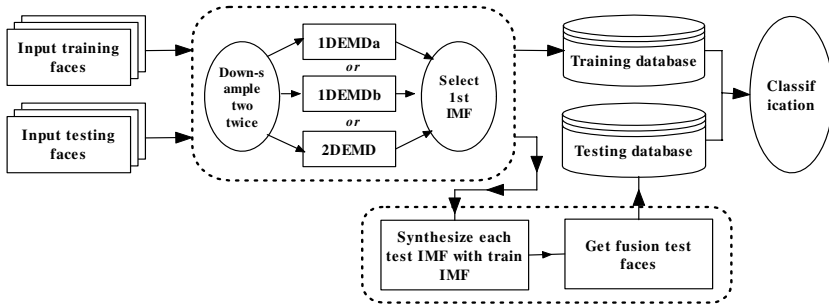**Fig. 4.** Flowchart of frequency face recognition scheme

Fig. 5. Flowchart of fusion based face recognition scheme

- Fusion the 1st IMFs of the train faces and test faces. In our experiment, we adopt only one neutral illumination face for training, and aim to verify the other four faces with serious illumination effects. As described in former sections, illumination effects always exist in low frequency IMFs, i.e., the 1st IMF contains minimum illumination effects. The synthesized faces of the two 1st IMFs enhance their mutual features while the serious illumination effects are weakened a lot in the test faces.

4   Experimental Results

We evaluated the proposed two schemes using the PIE face database, which is accessible at http://www.ri.cmu.edu/projects/project_418.html. This database contains 41368 images of 68 people, each person under 13 different poses, 43 different illumination conditions, and 4 different expressions. Here, we only focused on the images with varying illuminations. Fig.6 showed 21 face samples of one person. All the face images were normalized into 112×92 size.

In our experiments, we have established two sub-database for evaluating the two schemes separately. The first one is designed for the frequency method. As the rectangles denoted in Fig.6, the first 6 images were chosen from each subject orderly. Thereinto 3 face images were used for training and the remained 3 face images for testing. There are total 20 cases, here we only evaluated 5 cases and the results listed in Table.1 are average ones. Another sub-database contains

Fig. 6. Face image samples from the PIE database

5 faces per person. Each face is quite different from the other 4 in terms of illumination effects. We used the neutral one for training and the remained 4 for testing. Fig.7 showed an example.

Additionally, we have also compared our EMD based methods with the traditional wavelet based face recognition methods. In order to guarantee equitable comparisons, Daubechies wavelet 'db4' is adopted for wavelet decomposition, since 'db4' generally got the best performances as reported in [20,21]. Here we adopted Support Vector Machine (SVM) as classifier. Our experiments were implemented in a personal computer with Genuine Inete(R)T2300 CPU and 1.5G RAM and Matlab version 7.0 was used.

Table 1. Recognition rate versus different methods by scheme 1

Methods	1DEMDa	1DEMDb	2DEMD	'db4' wavelet
Correctness	99.51%	97.55%	93.63%	78.92%
	(203)	(199)	(191)	(161)

All the EMD based methods got much higher correctness than the wavelet based methods. Especially, 1DEMD applying on the lengthened image vector performed best, and the inferior one was 1DEMD applying row by row.

The next experiment we focused on testing the scheme 2. Fig. 7 demonstrated some face samples from the database and simultaneously showed the corresponding test fusion images. It is obvious that the illumination on test images have been removed a lot. Table. 2 showed the performances before fusion and after the fusion. The recognition rate are improved significantly comparing the poor correctness before fusion. All the EMD based methods performed nearly perfectly.

Fig. 7. Original testing face images and fusion face images by 1DEMDa

Table 2. Recognition rate versus different methods by scheme 2

Methods	1DEMDa	1DEMDb	2DEMD	'db4' wavelet
Correctness (before)	84.93% (231)	32.73% (89)	63.24% (172)	74.24% (202)
Correctness (after)	100% (272)	100% (272)	100% (272)	97.06% (264)

5   Conclusions

We have proposed two EMD based face recognition schemes in this paper. Both of the schemes have efficiently addressed the illumination effects problem. Compared with the traditional wavelet based methods, the proposed EMD based methods have much advantages in recognition rate. However, we also found that the EMD based methods have cost much time. Looking for the fast algorithm of EMD is one of our future work. Additionally, compared with 1DEMD, 2DEMD performed faster but poorer recognition rate. This is properly affected by the decomposition results which affected by the choices of surface interpolation methods. Finding the appropriate surface interpolation functions to improve the performances of 2DEMD is our remaining work in the coming days as well.

References

1. Huang, N.E., Shen, Z., Long, S.R., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A 454(1971), 903–995 (1998)
2. Huang, N.E., Shen, S.S.P.: Hilbert-Huang transform and its applications. Book in Interdisciplinary Mathematical Sciences, vol. 5 (2005)
3. Huang, N.E., Wu, M.L.C., Long, S.R., et al.: A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proceedings of the Royal Society A 459(2037), 2317–2345 (2003)
4. Long, S.R.: Applications of HHT in image analysis. In: Huang, N.E., Shen, S.S.P. (eds.) Hilbert-Huang Transform and Its Applications. World Scientific, River Edge (2005)
5. Hariharan, H., Gribok, A., Abidi, B., Abidi, M.: Multi-modal Face Image Fusion using Empirical Mode Decomposition. In: The Biometrics Consortium Conference, Crystal City, VA (2005)
6. Hariharan, H., Koschan, A., Abidi, B., Gribok, A., Abidi, M.A.: Fusion of visible and infrared images using empirical mode decomposition to improve face recognition. In: IEEE International Conference on Image Processing, ICIP 2006, Atlanta, GA, pp. 2049–2052 (2006)
7. Bhagavatula, R., Savvides, M., Acoustics, M.: Analyzing Facial Images using Empirical Mode Decomposition for Illumination Artifact Removal and Improved Face Recognition. In: IEEE International Conference on Speech and Signal Processing, 2007 (ICASSP 2007), April 15-20, vol. 1, pp. 505–508 (2007)
8. Nunes, J.C., Bouaoune, Y., Delechelle, E., Niang, O., Bunel, P.: Image analysis by bidimensional empirical mode decomposition. Image and Vision Computing 21(12), 1019–1026 (2003)
9. Nunes, J.C., Guyot, S., Deléchelle, E.: Texture analysis based on local analysis of the Bidimensional Empirical Mode Decomposition. Machine Vision and Applications 16(3), 932–8092 (2005)
10. Linderhed, A.: 2-D empirical mode decompositions in the spirit of image compression. In: Wavelet and Independent Component Analysis Applications IX, Orlando, Fla, USA. Proceedings of SPIE, vol. 4738, pp. 1–8 (2002)
11. Linderhed, A.: Compression by image empirical mode decomposition. In: IEEE International Conference on Image Processing (ICIP 2005), vol. 1, pp. 553–556 (2005)
12. Hariharan, H., Gribok, A., Abidi, M., Koschan, A.: Image Fusion and Enhancement via Empirical Mode Decomposition. Journal of Pattern Recognition Research 1(1), 16–32 (2006)
13. Sinclair, S., Pegram, G.G.S.: Empirical Mode Decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting. Hydrol. Earth Syst. Sci. Discuss. 2, 289–318 (2005)
14. Wan, J., Ren, L., Zhao, C.: Image Feature Extraction Based on the Two-Dimensional Empirical Mode Decomposition. In: 2008 Congress on Image and Signal Processing, vol. 1, pp. 627–631 (2008)
15. Taghia, J., Doostari, M.A., Taghia, J.: An Image Watermarking Method Based on Bidimensional Empirical Mode Decomposition. In: 2008 Congress on Image and Signal Processing, vol. 5, pp. 674–678 (2008)
16. Fauchereau, N., Sinclair, S., Pegram, G.: 2-D Empirical Mode Decomposition on the sphere, application to the spatial scales of surface temperature variations. Hydrol. Earth Syst. Sci. Discuss. 5, 405–435 (2008)

17. Nastar, C.: The image shape spectrum for image retrieval. Technical report, INRIA, No. 3206 (1997)
18. Nastar, C., Moghaddam, B., Pentland, A.: Flexible images: matching and recognition using learned deformations. Computer Vision and Image Understanding 65(2), 179–191 (1997)
19. Zhang, Z.B., Ma, S.L., Wu, D.Y.: The application of neural network and wavelet in human face illumination compensation. In: Wang, J., Liao, X.-F., Yi, Z. (eds.) ISNN 2005. LNCS, vol. 3497, pp. 828–835. Springer, Heidelberg (2005)
20. Feng, G.C., Yuen, P.C., Dai, D.Q.: Human face recognition using PCA on wavelet subband. Journal of Electronic Imaging 9(2), 226–233 (2000)
21. Ekenel, H.K., Sanker, B.: Multiresolution face recognition. Image and Vision Computing 23, 469–477 (2005)