Illumination Invariant Face Recognition based on the New Phase Local Features

Dan Zhang

Abstract

Hilbert-Huang transform (HHT) is a novel signal processing method which can efficiently handle non-stationary and nonlinear signals. It contains two key parts: Empirical Mode Decomposition (EMD) and Hilbert transform. EMD decomposes signals into a complete series of Intrinsic Mode Functions (IMFs), which capture the intrinsic frequency components of original signals. Hilbert transform is adopted on the IMFs to get the analytical local features. Recently, bidimensional version has been studied for advanced image processing. EMD has been extended to bidimensional EMD (BEMD), then the corresponding monogenic signals are studied. Phase information is an important local feature of signals in frequency domain because it is robust to contrast, brightness, noise, shading in the image. Phase congruency (PC) is a quantity that is invariant to changes in image illumination. In this paper, we firstly proposed a new BEMD method based on the improved evaluation of local mean, then the Riesz transform is applied to get the corresponding monogenic signals. Finally, PC calculated by the new phase information has been adopted as facial features to classify different faces under variant illumination conditions. The experimental results demonstrate the efficiency of our approach.

1 Introduction

Face recognition under varying illumination is always a challenging problem. Illumination changes could be larger than the differences between individuals. Many methods have been proposed to solve this problem such as illumination cone, quotient image, model based method and so on. There are one kind of methods that based on the frequency domain. Generally, researchers focused on two ways. One directly adopts high-frequency components as facial features due to the results that high-frequency components are more robust to the illumination changes while only the low-frequency component is sensitive to illumination changes[32, 33, 34]. Another one is based on the local phase information[28, 29].

Phase is an important local feature of signals in frequency domain. Earlier studies show that phase information alone is sufficient to completely reconstruct a signal within a scale factor[27]. It has been demonstrated both in experiments and in theory that edge features occur at places where phases at all frequencies are in congruency or within a small variation with each other. Therefore, the quantity named Phase Congruency (PC) is exploited as a frequency-based approach for feature detection. The studies show that PC is a dimensionless quantity that is invariant to changes in image brightness or contrast[31, 30]. PC can be calculated in Fourier domain, wavelet domain, and monogenic domain as well. This paper adopts monogenic filters to calculate PC.

We obtain the monogenic filters using the Hilbert-Huang transform (HHT) framework[1, 3, 2]. HHT is a novel signal processing method which can efficiently handle non-stationary and nonlinear signals. It contains two key parts: Empirical Mode Decomposition (EMD) and Hilbert transform. EMD decomposes signals into a complete series of Intrinsic Mode Functions (IMFs), which capture the intrinsic components of original signals. Hilbert transform is adopted on the IMFs to get the analytical local features. Recently, bidimensional version has been studied for advanced image processing[5, 7]. EMD has been extended to bidimensional EMD (BEMD) and the corresponding frequency components bidimensional IMFs (BIMFs) are generated, then the corresponding monogenic signals are studied[5, 17, 19, 15, 16, 18, 21]. In this paper, for convenience we also call the BEMD and monogenic signals process as HHT framework. The obtained monogenic phase is used for the PC calculation.

The motivation of this paper comes from three aspect advantages:

1. Earlier studies[32, 33, 34] and our studies show that high-frequency components are comparatively more robust to the illumination changes, while the low-frequency component is sensitive to them. Generally, high-frequency component only is enough for illumination invariant facial feature extraction.

2. HHT theory provides us another efficient method to decompose signals into different frequency IMFs com-

ponents. Because of the data-driven property and adaptiveness of the sifting process, it is able to capture more representative features and especially more singular information in high-frequency IMFs. It is reasonable to infer that the high-frequency components obtained by HHT framework may have more discriminate ability.

3. Phase information was found to be crucial to feature perception. Phase congruency is a dimensionless quantity that is invariant to changes in illumination.

Therefore, it is believable that the combination of high-frequency BIMFs and PC will further enhance the discriminate ability under variant illumination conditions.

In this paper, we firstly proposed a new BEMD method based on the improved evaluation of local mean, then the Riesz transform is applied to get the corresponding monogenic signals. Finally, PC based on the new phase information has been adopted as facial features to classify different faces under variant illumination conditions. The experimental results demonstrate the efficiency of our approach.

This paper is organized as follows: Section 2 presents an introduction to PC theory and introduce the calculation of PC using monogenic filters. In Section 3, we introduce the details of the modified 2DEMD and then show the monogenic signals of BIMFs. The simulation results are demonstrated in Section 4. Finally, the conclusions are given.

2 Phase Congruency

2.1 Phase Congruency Theory

Phase Congruency (PC) was first defined by Morrone and Owens [26] in terms of the Fourier series expansion of a 1D signal

$$PC(x) = \max_{\bar{\varphi}(x)\in[0,2\pi]} \frac{\sum_n A_n cos(\varphi_n(x) - \bar{\varphi}(x))}{\sum_n A_n}, \quad (1)$$

where A_n represents the amplitude of the nth Fourier component, and $\varphi_n(x)$ represents the local phase at position x. By Taylor expansion of $cos(\varphi_n(x) - \bar{\varphi}(x))$, we found that finding where PC is a maximum is transferred to finding where the weighted variance of local phase angles relative to the weighted average local phase $\bar{\varphi}(x)$ is minimum. However, this definition does not offer satisfactory local features and it is sensitive to noise. P. Kovesi [30] constructed a more sensitive measure of PC

$$PC_2(x) = \frac{\sum_n W(x)\lfloor A_n \Delta\phi_n(x) - T\rfloor}{\sum_n A_n + \varepsilon}, \quad (2)$$

where $\Delta\phi_n(x)$ is the phase deviation defined as

$$\Delta\phi_n(x) = cos(\varphi_n(x) - \bar{\varphi}(x)) - |sin(\varphi_n(x) - \bar{\varphi}(x))| \quad (3)$$

ε is a small constant to avoid division by zero and T is the estimated noise influence.

P. Kovesi also extended PC to two dimensions. The framework was based on wavelets. Since the response vectors of wavelet transform can also form the basis of localized representation of the signal. The need frequency information can be obtained via wavelet filter banks. Given signal s, even-symmetric (cosine) and odd-symmetric (sine) wavelets B_n^e, B_n^o at scale n, the response vector is

$$[e_n(x), o_n(x)] = [s(x) * B_n^e, s(x) * B_n^o]. \quad (4)$$

Then the amplitude is given by $A_n = \sqrt{e_n^2 + o_n^2}$, and the phase is given by $\varphi_n = atan2(e_n, o_n)$. They can be used to calculate PC as the same way as Fourier components used.

In terms of 2D image, one can use the above 1D analysis over m orientations.

$$PC_2(x) = \frac{\sum_m \sum_n W_m(x)\lfloor A_{n,m}\Delta\phi_{n,m}(x) - T_m\rfloor}{\sum_m \sum_n A_{n,m} + \varepsilon}, \quad (5)$$

Generally, researchers adopted the log Gabor filters.

2.2 Calculate PC Using Monogenic Filters

Actually, PC can also be calculated by monogenic filters. The monogenic signal was first proposed by Felsberg [22, 23] based on the Riesz transform. As a two-dimensional generalization of the 1D analytic signal, it similarly preserves the equivariance property of signal decomposition based on local amplitude and local phase.

The Riesz transformed signal in the frequency domain:

$$F_R(\mathbf{v}) = \frac{i\mathbf{v}}{\mathbf{v}} F(\mathbf{v}) = H_2(\mathbf{v})F(\mathbf{v}), \quad (6)$$

where $F(\mathbf{v})$ is the Fourier transform, and the transfer function of Riesz transform H_2 of the Riesz transform is a generalization of the Hilbert transform. **x** is a point in 2D space, then the multiplication in the Fourier domain above corresponds to the convolution in spatial domain is

$$f_R(\mathbf{x}) = -\frac{\mathbf{x}}{2\pi|\mathbf{x}|^3} * f(\mathbf{x}) = h_2(\mathbf{x}) * f(\mathbf{x}). \quad (7)$$

The monogenic signal $f_M(\mathbf{x})$ is composed by the the original signal and corresponding Riesz transformed signal.

$$f_M(\mathbf{x}) = f(\mathbf{x}) - (i,j)f_R(\mathbf{x}). \quad (8)$$

The amplitude of the monogenic signal $f_M(\mathbf{x})$ can be expressed as

$$|f_M(\mathbf{x})| = \sqrt{f_M(\mathbf{x})\overline{f_M(\mathbf{x})}} = \sqrt{f^2(\mathbf{x}) + |f_R(\mathbf{x})|^2}, \quad (9)$$

It is known that the polar representation of the complex $z = x + iy$ is $(r, \varphi) = (\sqrt{z\overline{z}}, \arg(z))$. Where $\overline{z}$ is the conjugate of z, $\arg(z)$ is the phase of the complex $\arg(z) = atan2(y,x) = sign(y)atan(|y|/x)$, $sign(y)$ represents the direction of rotation. The phase of the 2D analytical signal is

$$atan3(\mathbf{x}) = \frac{\mathbf{x}_D}{|\mathbf{x}_D|} atan(\frac{|\mathbf{x}_D|}{< (0,0,1)^T, \mathbf{x} >}), \qquad (10)$$

where $\mathbf{x}_D = (0,0,1)^T \times \mathbf{x}$ is the direction of the rotation vector. The expression of monogenic phase is

$$\varphi(\mathbf{x}) = atan3(f_M(\mathbf{x})) = \arg(f_M(\mathbf{x})). \qquad (11)$$

Thus using monogenic filters, the quantity PC can be calculated.

3 Monogenic Signals Obtained From BEMD

3.1 Hilbert-Huang Transform Framework

Hilbert-Huang transform was first proposed by N.E.Huang in 1998. It contains two key parts: Empirical Mode Decomposition (EMD) and Hilbert transform. EMD decomposes signals into a complete series of Intrinsic Mode Functions (IMFs), which capture the intrinsic components of original signals. Hilbert transform is adopted on the IMFs to get the analytical features. Since this method is local, data-driven, it is capable of handling nonlinear and non-stationary signals.

EMD 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 review the sifting process of EMD. Four main steps are contained, S1, S2, S3 and S4 are abbreviation for Step 1 to Step 4. Given a signal x(t),

- S1. Identify all the local minima and maxima of the input signals $x(t)$;
- S2. Interpolate between all minima and maxima to yield two corresponding envelopes $E_{max}(t)$ and $E_{min}(t)$. Calculate the mean envelope $m(t) = (E_{max}(t) + E_{min}(t))/2$;
- S3. 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 Step 4. Otherwise, repeat Step 1 and Step 2 using the residue $h(t)$, until the latest residue meets the threshold and turns to be an IMF;
- S4. Input the residue r(t) to the loop from Step 1 to Step 3 to get the next remained IMFs until it can not be decomposed further.

The analytical signal provides a way to compute the 1D signal's local amplitude and phase, which is obtained by the Hilbert transform on a real signal. The Hilbert transform $f_H(x)$ of a real 1D signal f is given by:

$$f_H(x) = f(x) * \frac{1}{\pi x},$$

where $*$ is convolution. $f_H(x)$ is the imaginary part of the signal. The analytical signal can be written as

$$f_A = f(x) + i f_H(x) = a(t)e^{i\theta(t)},$$

in which, $a(t)$ is the amplitude, $\theta(t)$ is the phase.

3.2 The improved BEMD

For advanced image processing, the EMD has been extended to two dimensions, i.e., bidimensional EMD (BEMD)[17, 19, 15, 16, 18, 21]. However, one counters a lot of challenges such as inaccuracy of surface interpolation, high computational complexity and so forth. Here we propose an alternative algorithm for EMD. Instead of using the envelopes generated by splines we use a low pass filter to generate a "moving average" to replace the mean of the envelopes. The essence of the sifting algorithm remains.

The moving average is the most common filter in digital signal processing. It operates by averaging a number of points from the input signal to produce each point in the output signal, it is written:

$$y[i] = \frac{1}{M}\sum_{j=0}^{M-1} x[i+j],$$

where $x[]$ is the input signal, $y[]$ is the output signal, and M is the number of points used in the moving average. It is actually a convolution using a simple filter $[a_i]_{i=1}^{M}, a_i = \frac{1}{M}$, and $[A_{i,j}]_{i=1,j=1}^{M,N}, A_{i,j} = \frac{1}{M \times N}$ for the 2-dimensional case.

Detection of local extrema means finding the local maxima and minima points from the given data. No matter for 1D signal or 2D array, neighboring window method is employed to find local maxima and local minima points. The data point/pixel is considered as a local maximum (minimum) if its value is strictly higher (lower) than all of its neighbors.

We illustrated 1-dimensional case and 2-dimensional case separately.

- 1-dimensional case:
  For each extrema map, the distance between the two neighborhood local maxima (minima, extrema, zero-crossing) has been calculated called as adjacent maxima (minima, extrema, zero-crossing) distance vector Adj_max ($Adj\_min, Adj\_ext, Adj\_zer$). Four types of window size:

– Window-size I: $\max(Adj\_max)$;
– Window-size II: $\max(Adj\_min)$;
– Window-size III: $\max(Adj\_zer)$;
– Window-size IV: $\max(Adj\_ext)$.

- 2-dimensional case:
  The window size for average filters is determined based on the maxima and minima maps obtained from a source image. For each local maximum (minimum) point, the Euclidean distance to the nearest local maximum (minimum) point is calculated, denoted as adjacent maxima (minimum) distance array Adj_max (Adj_min).

  – Window-size I: $\max(Adj\_max)$;
  – Window-size II: $\max(Adj\_min)$;

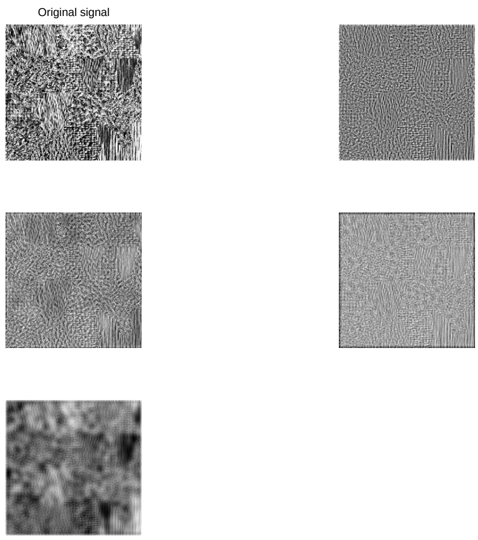

Figure 1. The BIMFs obtained by the improved 2DEMD.

3.3 Monogenic Features of BIMFs

Similar with HHT, Riesz transform was applied on the bidimensional IMFs (BIMFs) got by the improved BEMD. Fig.2 and Fig.3 demonstrated the local features obtained by convolving the BIMFs with monogenic filters.

4 Experimental Results

We first briefly describe the algorithm of our feature extraction method.

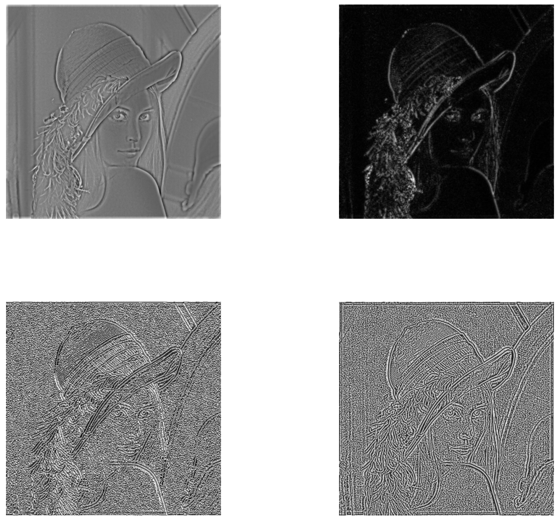

Figure 2. left-up: 1st IMF, right-up: amplitude, left-down: phase orientation, right-down: phase angle.

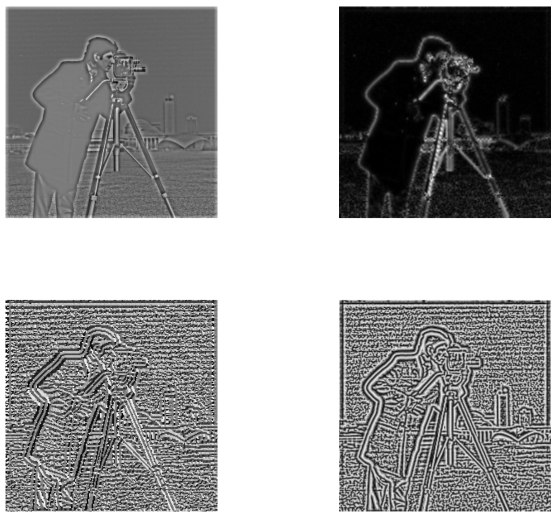

Figure 3. left-up: 1st BIMF, right-up: amplitude, left-down: phase orientation, right-down: phase angle.

step1. Input face images;

step2. Decompose the face images by applying the former BEMD.

step3. Apply Riese transform on the first three BIMFs obtained to get monogenic phase.

step4. Calculate the PC, the weighted mean PC is our facial features which will be used for classification.

We evaluated the proposed method 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.4 showed 21 face samples of one person. All the face images were normalized into 112×92 size.

Figure 4. Face image samples from the PIE database

In our experiments, we have established a sub-database for evaluating the scheme. As the rectangles denoted in Fig.4, the first 6 images were chosen from each subject orderly. Thereinto 2 face images were used for training and the remained 4 face images for testing. There are total 5 cases, here we only evaluated 5 cases and the results listed in Table.1 are average ones.

Additionally, we have also compared the proposed method with other phase based face recognition methods. 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 R2009b was used.

Table 1. Recognition Rate Comparison.

Methods	[4]	[28]	[29]	our method
Correctness	68.41	81.25%	91.83%	94.63%
	186	(221)	(249)	(257)

5 Conclusions

In this paper, we use the phase congruency quantity based on the BIMFs to address the illumination face recognition problem. We firstly proposed a new BEMD method based on the improved evaluation of local mean, then apply the Riesz transform to get the corresponding monogenic signals. Based on the new phase local information obtained, PC is calculated. We combine the PC on different BIMFs and use the weighted mean as the facial features input to the classification process. Compared with other phase based face recognition method, our proposed method shows its efficiency.

References

[1] N. E. Huang, Z. Shen, S. R. Long, et al.. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A, vol. 454, no. 1971, pp. 903C995, 1998.*

[2] Hilbert-Huang Transform and its applications. *Book in Interdisciplinary Mathematical Sciences, Vol. 5, edited by N. E. Huang, and Samuel S P Shen, 2005.*

[3] N. E. Huang, M. L. C. Wu, S. R. Long, et al.. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society A, vol. 459, no. 2037, pp. 2317C2345, 2003.*

[4] Bhagavatula, R., Marios Savvides, and M. Acoustics. Analyzing Facial Images using Empirical Mode Decomposition for Illumination Artifact Removal and Improved Face Recognition. *IEEE International Conference on Speech and Signal Processing, 2007 (ICASSP 2007). Vol. 1, Issue , 15-20 April 2007 pp. I 505-508.*

[5] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and Ph. Bunel. Image analysis by bidimensional empirical mode decomposition. *Image and Vision Computing Volume 21, Issue 12, Pages 1019-1026, November 2003.*

[6] Nunes J. C., Guyot S., and Deléchelle E. Texture analysis based on local analysis of the Bidimensional Empirical Mode Decomposition. In *Machine Vision and Applications 16, 3, pp. 0932-8092, 2005.*

[7] A. Linderhed. 2-D empirical mode decompositions in the spirit of image compression. *in Wavelet and Independent Component Analysis Applications IX, vol. 4738 of Proceedings of SPIE, pp. 1C8, Orlando, Fla, USA, April 2002.*

[8] A. Linderhed. Compression by image empirical mode decomposition. *IEEE International Conference on Image Processing, 2005 (ICIP 2005), Vol. 1, pp. I 553-6, 2005.*

[9] H. Hariharan, A. Gribok, M. Abidi, and A. Koschan. Image Fusion and Enhancement via Empirical Mode Decomposition. *Journal of Pattern Recognition Research, Vol. 1, No. 1, pp. 16-32, January 2006.*

[10] Sinclair, S. and 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.*

[11] Jian Wan, Longtao Ren, and Chunhui Zhao. Image Feature Extraction Based on the Two-Dimensional Empirical Mode Decomposition. *2008 Congress on Image and Signal Processing, Vol. 1, pp. 627-631, 2008.*

[12] Jalil Taghia, Mohammad Ali Doostari and Jalal Taghia. An Image Watermarking Method Based on Bidimensional Empirical Mode Decomposition. *2008 Congress on Image and Signal Processing, Vol. 5, pp. 674-678, 2008.*

[13] Fauchereau, N., Sinclair, S., and 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.*

[14] C. Damerval, S. Meignen, and V. Perrier. A fast algorithm for bidimensional EMD. *IEEE Signal Processing Letters, vol. 12, no. 10, pp. 701C704, 2005.*

[15] Sharif M. A. Bhuiyan, Reza R. Adhami, and Jesmin F. Khan. Fast and Adaptive Bidimensional Empirical Mode Decomposition Using Order-Statistics Filter Based Envelope Estimation. *EURASIP Journal on Advances in Signal Processing, vol. 2008, Article ID 728356, 18 pages, 2008.*

[16] Sharif M. A. Bhuiyan, Reza R. Adhami, and Jesmin F. Khan. A novel approach of fast and adaptive bidimensional empirical mode decomposition. *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008), pp. 1313-1316, 2008.*

[17] Sherman Riemenschneider, Bao Liu, Yuesheng Xu and Norden E. Huang. B-spline based empirical mode decomposition. *Hilbert-Huang Transform and Its Applications, Book chapter 2, 2005.*

[18] Louis Yu Lu. Fast intrinsic mode decomposition of time series data with sawtooth transform. *Technical report, Nov 2007.*

[19] Yong-Ping Huang, Xue-Yao Li and Ru-Bo Zhang. A research on local mean in empirical mode decomposition. *Proceedings of the 7th international conference on Computational Science, Part III: ICCS 2007, Lecture Notes In Computer Science, Vol. 4489, pp. 125-128.*

[20] Luan Lin, Yang Wang, and Haomin Zhou. A new approach to empirical mode decomposition. *Preprint.*

[21] Lixin Shen. Local mean and empirical mode decomposition. *Report on The Second International Conference on the Advances of Hilbert-Huang Transform and Its Applications, Guangzhou, Dec 2008. Preprint.*

[22] Felsberg M and Sommer. G. The monogenic signal. *IEEE Trans Signal Process 49, 2001.*

[23] Felsberg M and Sommer. G. The Monogenic Scale-Space: A Unifying Approach to Phase-Based Image Processing in Scale-Space. *2004.*

[24] Felsberg M and Sommer. G. Structure multivector for local analysis of images. *Workshop on theoretical foundations of computer vision, SpringerVerlag, pp 93-104, 2001.*

[25] Felsberg M and Sommer. G. Low Level Image Processing with the Structure Multivector. *2002.*

[26] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters, 6: 303 - 313, 1987.*

[27] Alan V. Oppenheim, and Jae S. Lim. The importance of phase in signals. *Processings of The IEEE, Vol. 69, No. 5, May 1981.*

[28] Jingu Heo, Marios Savvides, and B.V.K Vijayakumar. Illumination tolerant face recognition using phase-only support vector machines in frequency domain. *Processings of ICAPR2005, pp. 66-73, 2005.*

[29] Zhikai Huang, Dehui Liu, Weizhong Zhang and Lingying Hou. Multi-view face database recognition using phase congruency and svm classifier. *Processings of 2008 International Conference on Computer and Electrical Engineering, pp. 219-222, 2005.*

[30] Peter Kovesi. Image features from phase congruency. *MIT Press, Quarterly Journal, Vol. 1, No. 3, 1999.*

[31] Michael Felsberg and Gerald Sommer. A new extension of linear signal processing for estimating local properties and detecting features. *In 22. DAGM Symposium Mustererkennung, pp 195-202, 2000.*

[32] C Nastar. The image shape spectrum for image retrieval. *Technical report, INRIA, No. 3206, June 1997.*

[33] C Nastar, B Moghaddam and A Pentland. Flexible images: matching and recognition using learned deformations. *Computer Vision and Image Understanding, Vol. 65, No. 2, pp. 179- 191, 1997.*

[34] Zhang, Z. B., Ma, S. L. and Wu, D. Y. The application of neural network and wavelet in human face illumination compensation. *Proc. Advances in Neural Networks, pp. 828C 835, 2005.*

[35] Feng, G. C., Yuen, P. C. and Dai, D. Q. Human face recognition using PCA on wavelet subband. *Journal of Electronic Imaging, Vol. 9, No. 2, pp. 226C233, 2000.*

[36] Ekenel, H. K. and Sanker, B. Multiresolution face recognition. *Image and Vision Computing, Vol. 23, pp. 469C477, 2005.*

[37] Xinge You, Dan Zhang, Qiuhui Chen, Patrick Wang, and Yuan Yan Tang. Face representation by using nontensor product wavelets. *18th International Conference on Pattern Recognition, pp.503–506, Aug 2006.*