國立交通大學

生醫工程研究所

碩士論文

PLS-based EMD 在臉部辨識應用之開發

Development of PLS-based EMD for face recognition

研 究 生: 蘇聖元

指導教授: 蕭子健 博士

中 華 民 國 102 年 1 月

PLS-based EMD 在臉部辨識應用之開發
Development of PLS-based EMD for face recognition

研 究 生：蘇聖元　　Student：Sheng-Yuan Su

指導教授：蕭子健　　Advisor：Tzu-Chien Hsiao

國 立 交 通 大 學

生 醫 工 程 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Biomedical Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

January 2013

Hsinchu, Taiwan, Republic of China

中華民國 102 年 1 月

PLS-based EMD 在臉部辨識應用之開發

研究生: 蘇聖元　　指導教授: 蕭子健

國立交通大學

生醫工程研究所

摘要

人臉辨識為圖形識別發展的一個重要應用，在過去的三十年來，研究學者致力於提升臉部辨識系統的可靠度，許多人臉辨識的演算法也應運而生。在過去的文獻曾提及，某些辨識的演算法易受到臉部影像間不同光線變化條件所影響，而產生誤判人臉的情況。例如以特徵為基礎的技術，此種技術藉由統計方法來計算特徵向量，進而擷取到能區分不同臉部影像差異性最大的特徵。例如: Eigenface 或 Tri-PLS 等方法。此種方法為對臉部影像作整體考量，也因此當臉部影像間存在不同光線條件變化時，會使辨識率大受影響。

針對此種問題，過去有許多演算法被提出，其中以 EMD 為基礎的技術在近年來漸漸受到重視。研究學者使用不同的內插法來實作 BEMD, 並應用所提出的方法來對臉部影像作前處理以達到可靠的辨識結果。 本論文嘗試以 Tri-PLS 對臉部影像資料庫建樣板，並以此樣板來取代典型 BEMD 方法的內插過程，再進一步檢視此種方法的有效性。

評估演算法辨識率的部份使用了 Yale 及 PIE 人臉資料庫做測試，並與 Eigenface 及 Tri-PLS 等方法做比較。結果顯示所提出的方法 PLS-EMD 在辨識率上優於此兩種方法，尤其在使用 PIE database 作評估時，Eigenface 及 Tri-PLS 的 leave-one-out cross-validation 辨識率約只有 20%及 40%，而 PLS-EMD 的辨識率則可達到 90%，如此大的差距說明了 PLS-EMD 相較於一般的辨識方法，對影

像間存在不同光線條件變化具有可靠的辨識能力。另外再進一步與典型的 BEMD 方法作比較，結果顯示所提出的方法無論是以 repeated random sub-sampling validation 或 leave-one-out cross-validation 做評估，在辨識率上均優於典型的 BEMD 作法。而在計算效能上，若以 PIE database 作評估，BEMD 在特定停止條件下，平均每張影像計算時間約需 50 秒; 而 PLS-EMD 在此條件下僅需花費約 7 秒。由此可知其計算效能的差異。

Development of PLS-based EMD for face recognition

student: Sheng-Yuan Su Advisor: Tzu-Chien Hsiao

Institute of Biomedical Engineering
National Chiao Tung University

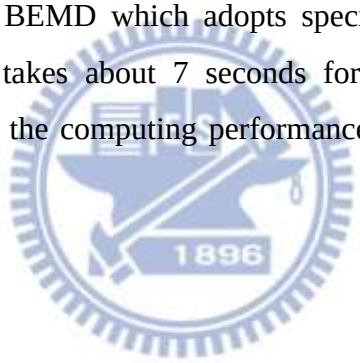
Abstract

Face recognition is an important application in the field of pattern recognition. Researchers devote themselves to enhance the reliability of the face recognition system in the past thirty years and many of the related algorithms emerge as the times require. As one of the major challenges in face recognition mentioned in literatures, some recognition algorithms will probably misclassify faces when different illumination conditions are present among face images. As one kind of feature extraction methods, feature-based methods extract features to distinguish different faces by solving eigenvalue problem; however this kind of method such as Eigenface and Tri-PLS suffers from the problem mentioned above. The feature-based methods take the whole regions of face image into consideration, and therefore the recognition rate will degrade significantly when the different illumination conditions are present in face images.

Some algorithms have been proposed to deal with this problem in the past. As one kind of solutions, EMD-based methods have received significantly attention in recent years. Researchers adopted different interpolation methods to implement BEMD, and then applied the proposed methods to preprocess face images in order to enhance the recognition accuracy. In this study, a different approach which replaces the interpolation process of BEMD by selecting templates is proposed. The templates of the face database are obtained by Tri-PLS.

Yale and PIE databases are applied to evaluate the recognition rate of the proposed method. The compared methods include Eigenface and Tri-PLS. Simulation results show that the recognition rate of the proposed method (i.e. PLS-EMD) is better than that of the two methods. The recognition rate of Eigenface and Tri-PLS evaluated by leave-one-out cross-validation are only about 20 and 40 percent respectively, and that of PLS-EMD is about 90 percent. The great disparity between the two recognition rates indicates that PLS-EMD has a reliable recognition ability to resist different illumination variations between images compared to other general methods.

Another comparison between BEMD and PLS-EMD is also given. The recognition results evaluated by sub-sampling validation and Leave-one-out cross-validation also shown that PLS-EMD has better recognition rate than that of BEMD. As for computing performance, the average computation time per image of PIE database computed by BEMD which adopts specific stop criteria is about 50 seconds; however, it only takes about 7 seconds for PLS-EMD under the same condition. It is obvious that the computing performance of PLS-EMD is better than that of BEMD.



誌謝

能完成這本碩士論文，首先我想感謝家人的支持，讓我有勇氣回來學校完成我一開始所選擇的路。在過去的兩年裡，先後到成功嶺、台中市政府觀光旅遊局服役，沉澱了許久，也逃避了好一陣子，最後因為家人的關係以及一念之間選擇了回來。我很慶幸能回來面對自己，從離開再到回來完成學業，走過這一遭，雖然沒有完美的結局，但我總算可以對過去的自己有所交待，甚至讓家人以我為榮，更讓我無所遺憾。

我想感謝我的指導教授蕭子健老師，由於他的殷勤指導及無限關愛，讓我能順利地完成學業，我從他身上學到了不僅是做研究上的知識，更學到了如何待人處事、如何關心他人，真的很感謝他過去一年來的照顧。此外，我也想感謝三位口試委員，分別是林志青教授、王智昱教授及陳錦龍教授。感謝林志青教授過去所傳授的圖形識別課程的相關知識，所學到的知識成為我往後研究的基石。也很感謝王智昱教授在百忙之中仍能前來出席口試會議以及感謝陳錦龍教授特地從花蓮過來擔任口試委員，非常感謝他們所給與的寶貴建議，讓我的論文更加完備。

最後要感謝的是 VBM 實驗室的成員，感謝你們精神上的支持讓我得以順利通過口試的挑戰! 這期間特別要感謝虹名、珮姍及岱凌的鼓勵，以及口試期間虹名及岱凌的幫忙。我永遠不會忘記去年年終 outside meeting 時和實驗室大家去尖石打漆彈及泡湯的情景，以及晚上我喝到掛、抽到頭獎最後被繼武等人扛上計程車的窘樣。謝謝你們讓我在這一年苦悶的研究生涯日子裡有個美好的回憶。

Table of Contents

List of Figures

List of Tables

I. Introduction

1.1 Background

As one of most successful and famous application of image analysis, face recognition has received significant attention in the past several years. It plays an important role in the areas of smart cards [1], information security [2], law enforcement [3], surveillance [4], etc. Literature has shown that the recognition accuracy will probably drop significantly in a natural or uncontrolled environment [1]. For the purpose of enhancing the recognition accuracy in real applications, the environmental conditions are usually well-controlled, i.e. neutral face expression, small variation of lighting condition, processed image background, and a certain range of head pose variation [1]. One of the major challenges of face recognition is indicated, that is the illumination variation problem [5]. It is believed that variations caused by lighting in face images are even larger than differences among distinct individuals [5]. Therefore, the variations of lighting conditions among different face images will lead to great influence on the recognition accuracy. Some potential methods which are introduced in the section 1.1.1 have been proposed to deal with this problem.

1.1.1 Approaches to illumination variation problem

Those approaches for solving illumination problem can be classified into three categories [6] as follows.

(a) Face and illumination modeling:

The main concept of this category is to build illumination or face model to deal with the illumination variation problem. The Illumination Cone method [7] which attempts to synthesize the face images under different poses and illumination conditions by reconstructing the shape and albedo of the face from a small number of training images. Lambertian surface [8] shows that the set of images under different lighting conditions can be simply characterized as a nine dimensional subspace in the space of all possible images. Three-dimensional morphable model [9] attempts to construct a general 3D human face model in order to fit different

illumination and pose conditions.

(b) Illumination invariant feature extraction:

Those methods in this category are mainly to extract the illumination insensitive facial features to perform further recognition. The representative methods include edge map, image intensity derivatives and Gabor-like filtering image [5]. Recently, quotient-image-based methods are reported to be a simple and efficient solution to illumination variations. Quotient Image (QI) [10] is defined as image ratio between a test image and linear combinations of three unknown independent illumination images. It only depends on the relative surface texture information and is illumination insensitive. Quotient Illumination Relighting (QIR) [11], Self-Quotient Image (SQI) [12] and Morphological Quotient Image (MQI) [13] are all derived from the idea of QI. As another recent work, Logarithmic Total Variation (LTV) [14] decomposes the image into two parts, one with small-scale features and the other one with large-scale features. Only small-scale features will be used for recognition.

(c) Preprocessing and normalization:

In this approach, the face images with different illumination conditions are preprocessed in order to obtain the normal lighting images. Further recognition will be performed based on the normalized images. The most commonly used methods include histogram equalization [15] and discrete cosine transform (DCT) based methods [16-19]. In the implementation of DCT, low-frequency discrete cosine transform coefficients are discarded to eliminate effects of illumination variations since illumination variations mainly lie in the low-frequency band. This method does not require multiple images to be trained.

1.1.2 EMD-based approaches

As one kind of illumination invariant feature extraction methods mentioned above, the Empirical mode decomposition (EMD) based methods have received significant attention in recent years. EMD was originally proposed by Huang *et al.* in 1998 [20] and was originally applied to analyze 1D signal. This method has been investigated how to apply to image analysis in recent years. Researchers have tried to apply EMD-based method to deal with the illumination problem. For example, Bhagavatula *et al.* [21] utilized EMD to decompose the face images in variant illumination conditions into different components and found that the illumination trend is presented in the last components. This observation gave them a thought for robust face recognition under variant illumination conditions by removing the components which contain the illumination effects. As a two-dimensional version of EMD, BEMD was proposed by Nunes *et al.* in 2003 [22] and applied to image analysis directly. Shao *et al*. [23] utilized BEMD to extract a series of normalization images from one subject, then canonical correlation analysis (CCA) is adopted to generate more discrimative features; Zhang *et al*. [24] proposed an improved BEMD to get bi-dimensional intrinsic mode functions (BIMF), then Riesz transform is subsequently applied to these obtained 2D analytic signals, i.e. BIMFs to get the corresponding monogenic signals. Finally, phase congruency (PC) [25] was calculated to get facial features under variant illumination conditions.

As mentioned above, EMD-based methods have received significant attention in recent years, especially in the face recognition field. Researchers have noticed the ability of EMD-based methods to deal with the illumination problem. Therefore, it is valuable to investigate the properties and the characteristics of EMD-based methods.

1.2 Motivation

As mentioned in above studies, the methods derived from EMD are different. What are the differences among these methods? How to apply them to solve the illumination problem? A further thought is that is it possible to propose a modified version of EMD-based methods to apply to the illumination problem?

1.3 Literature study

EMD was originally developed for analyzing one-dimensional non-stationary and nonlinear signals (Huang *et al*. 1998). Due to its adaptability, effectiveness and the fact that it can decompose signal to multi-scale components, this decomposition technique is widely used and has been investigated how to extend to 2D version to analyze two-dimensional (2D) data/images. The thinking of developing 2D EMD was originated from Linderhed [26]. He successfully used EMD to encode and decode the audio signal and gave a thought that image compression can be realized by a 2D version of EMD in the same way. This thought inspires the researchers to develop a two-dimensional EMD. Nunes *et al*. proposed a typical 2D EMD method in 2003, which is known as BEMD. He is the first one who proposed a 2D framework of EMD. The extrema detection method of BEMD adopts neighboring window or morphological operation, and the interpolation method for constructing 2D envelope (i.e. surface) adopts radial basis function (RBF). Researchers subsequently followed this framework to investigate how to improve BEMD mainly by proposing different interpolation methods. For instance, Damerval *et al.* [27] replaced the interpolation method RBF with Delaunay triangulation and fixed number of iterations to obtain BIMFs; Linderhed [28] used thin-plate spline which is an alternate choice of RBF as the interpolation method; Xu *et al.* [29] provided another interpolation approach by using a mesh fitting method based on finite elements; Bhuiyan *et al.* [30] used order statistic filter to obtain upper and lower surfaces rather than using interpolation method. Except the 2D implementation, 1D EMD has also been applied to decompose images to obtain BIMFs [31]. In this technique, each row and/or each column of the 2D data is processed by 1D EMD, which makes it a faster process. However, it has

been found that this 1D implementation results in poorer BIMF components compared to the standard 2D procedure since the former ignores the correlation among the rows and/or columns of a 2D image.

In recent years, EMD has been further developed to process the multidimensional data by multi-dimensional ensemble EMD method, designated as MEEMD [32]. The decomposition is based on ensemble empirical mode decomposition (EEMD) to slice an image in each and every dimension involved. This method bypasses major obstacles and difficulties in traditional BEMD, such as how to define the 2D extrema and the mode-mixing problem. Another advantage of MEEMD is that it can be applied to decompose spatially three or more dimensional fields without any barrier. Such extension would not be feasible for the traditional two-dimensional approaches. This method represents a new milestone in the progress of EMD.

The different decomposition methods mentioned above are given in table 1.

Table 1. Comparisons of different decomposition methods

Year	Authors	scheme	Envelope constructing method
1998	Huang *et al.*	1D	Cubic spline
2003	Nunes *et al.*	2D	RBF
2004	Liu *et al.*	1D	Based on EMD
2005	Damerval *et al.*	2D	Delaunay triangulation
2007	Xu *et al.*	2D	Finite element
2008	Bhuiyan *et al.*	2D	Order statistic filter
2009	Wu *et al.*	1D	Based on EEMD

1.4 Objective

The interpolation method of the typical BEMD is implemented in a self-organization scheme. It means the interpolating values only depend on the values of local extrema without any corresponding class information of this decomposed face image participating in the interpolation process. It gives us an idea that maybe we can construct the upper and lower surfaces by adding corresponding class information of the decomposed face image into the interpolation process. This different interpolation approach can be seen as a cross-correlation scheme. It has a cross relationship between the intrinsic values of this decomposed face image (**X**) and the corresponding class information (**Y**). Is it possible to implement the interpolation process by this different approach? Does this different scheme have better recognition accuracy or performance than the typical BEMD on the illumination variation issue?

1.5 Contribution

The contribution of this research is to propose and implement cross-correlation scheme of 2D-EMD which is different from traditional approach by self-organization scheme. Due to the cross-correlation property, PLS-EMD approach is possibly helpful to enhance recognition accuracy of face recognition under different illumination conditions.

1.6 Thesis organization

This thesis can be mainly divided into six chapters. The related works are given in chapter II and the proposed PLS-EMD method will be introduced in the end of this chapter. Chapter III displays the simulation results and compares some representative methods with the proposed method under different illumination conditions. The performance issue, the advanced comparisons and the validation of PLS-EMD will be discussed in chapter IV. Conclusion and future works are given in the last chapter of this thesis.

II. Material and methods

2.1 Empirical mode decomposition

Empirical mode decomposition (EMD) was originally developed for analysis of one-dimensional non-stationary and nonlinear signals (Huang *et al.* 1998). It decomposes a signal into a finite sum of intrinsic mode functions (IMF) that generally allow well-behaved Hilbert transforms. This makes it possible to construct a time–frequency representation, called the Hilbert spectrum, using instantaneous frequency. EMD combined with Hilbert spectrum, called the Hilbert–Huang Transform (HHT), has many advantages over the traditional time–frequency analysis techniques in the adaptability, the capability of decomposing nonlinear and unstable signals. In recent years, HHT has been applied with great success in various application areas.

2.1.1 Intrinsic mode function

Ideally, the IMFs of a signal obtained by EMD are expected to have the following properties [20, 30].

(i) In the whole data set of an IMF, the number of local extrema and the number of zero crossings must be equal or differ by at most one.

(ii) There should be only one mode of oscillation, that is, only one local maximum or local minimum, between two successive zero crossings.

(iii) At any point, the mean value of the upper and lower envelopes, defined by the local maxima and minima points, is zero or nearly zero.

(iv) The IMFs are locally orthogonal among each other.

2.1.2 Sifting process

The overall sifting process can be viewed in figure 1.

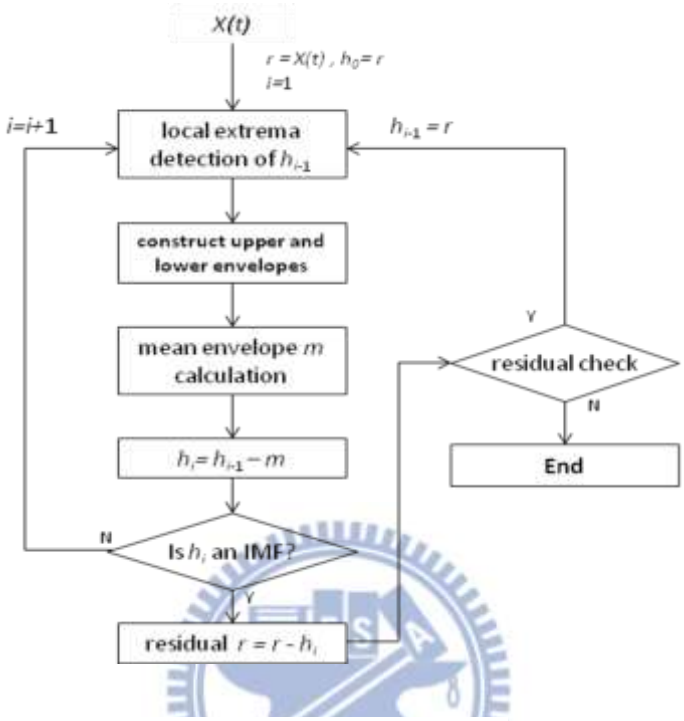


Figure 1. Sifting process of EMD

The sifting process can be illustrated as follow:

Define *X(t)* as an one-dimensional signal.

(i) Initialize: $r = X(t)$, $h_0 = r$, $i = 1$

(ii) Find the local extrema of h_{i-1} (local maxima and minima, respectively) by sliding window (local derivative is an alternate way to find local extrema [34]).

(iii) Interpolate all the local maxima and minima respectively to generate upper envelope *u(t)* and lower envelope *l(t)* by cubic spline.

(iv) Compute mean envelope *m* by the equation: $m = (u(t) + l(t)) / 2$.

(v) Calculate the difference between h_{i-1} and current mean envelope m.

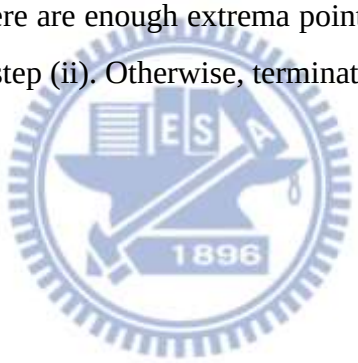
→ $h_i = h_{i-1} - m$

(vi) Compute standard deviation: $$\mathrm{SD} = \sum_{t=0}^{T} \left[\frac{|h_{i-1}(t) - h(t)|^2}{h_{i-1}^2(t)} \right]$$

*Note: SD has alternate definitions in other studies [35].

(vii) If SD is greater than the predefined threshold value, which is usually a small nearly zero value, then set iteration index $i = i+1$ and repeat the sifting process from step (ii).

Otherwise, if SD is less than the predefined threshold, then regard h_i as an IMF and set intermediate residual $r = r - h_i$. Finally, check r to insure enough extrema points. If there are enough extrema points, then set $h_{i-1} = r$ and repeat sifting process from step (ii). Otherwise, terminate the sifting process.



An example of sifting result can be viewed in figure 2.

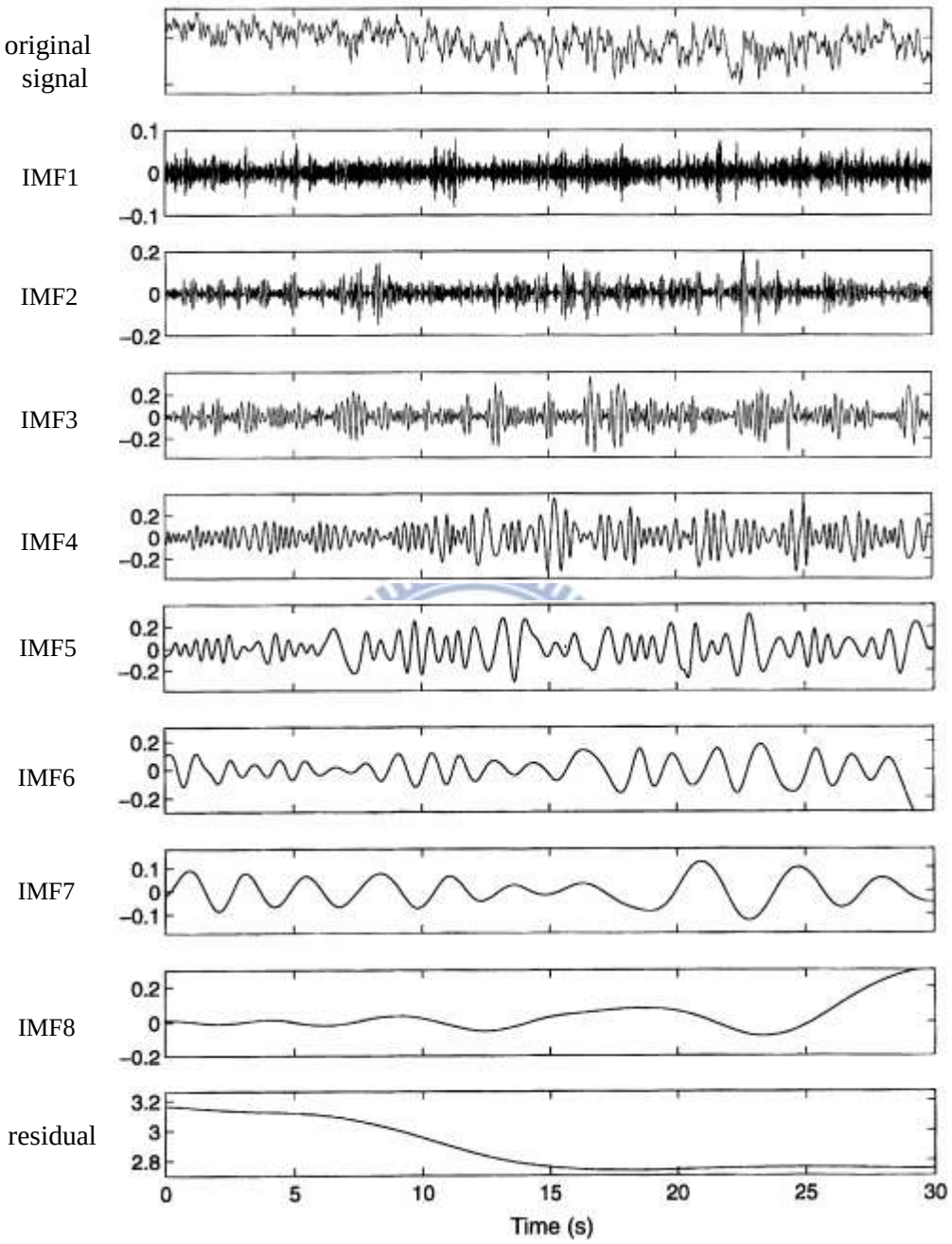


Figure 2. Example of sifting result (Revised from Huang *et al.* 1998, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis" [20])

From the figure, we can find that the IMF1 corresponds to the highest local frequency of oscillation of original signal; IMF2 corresponds to the second highest local frequency of oscillation and so on. The last component of the sifting result, i.e. residual corresponds to the trend of the data. By summing up all the components, we obtain

$$X(t) = \sum_{i=1}^{n} C_i + r \qquad (1)$$

That is, we can decompose a signal by EMD to find its different frequency components. This method is totally adaptively and can be applied to nonlinear and non-stationary signal compared with traditional Fourier transform.

2.1.3 Completeness and orthogonality

The property "completeness", which is established by equation (1), represents that an original data can be reconstructed by summing up all the IMFs and the residual. To further check the orthogonality of IMFs, an equation is given first:

$$X(t) = \sum_{i=1}^{n+1} C_i \ , \quad C_{n+1} = r \qquad (2)$$

in which residual is regarded as an extra component C_{n+1}. An overall index of orthogonality (IO) [20] is defined as:

$$\mathrm{IO} = \sum_{t=0}^{T} \left(\sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \frac{C_j(t) C_k(t)}{X^2(t)} \right) \qquad (3)$$

The orthogonality actually depends on the decomposition methods; however, a lower value of IO is preferred.

2.2 Bi-dimensional Empirical mode decomposition

Bi-dimensional Empirical mode decomposition (BEMD) is a two-dimensional extension of EMD and it has been applied in various real-world problems, e.g. medical image analysis, pattern analysis, and texture analysis. The typical BEMD was firstly proposed by Nunes in 2003 [22]. Researchers followed the framework of this typical BEMD to develop their own BEMD. They mainly focused on the modification or replacement of the interpolation method from the typical BEMD to enhance the performance or get better sifting results. A brief introduction of typical BEMD will be given from section 2.2.1 to section 2.2.4 and the interpolation method will be introduced in section 2.2.5.

2.2.1 Bi-dimensional intrinsic mode function

The definitions and the properties of BIMFs are slightly different from IMFs. It is sufficient for BIMFs to follow only the final two (iii and iv) properties mentioned in section 2.1.1.. In fact, due to the properties of an image and the BEMD process, it is impossible for a BIMF to satisfy the first two properties (i and ii), since the maxima and minima points are defined in a 2D scenario for an image. This viewpoint can be referred to this literature [30].

2.2.2 Sifting process

The sifting process of BEMD can be viewed in figure 3. It is obvious that the sifting process is nearly the same as EMD except extrema detection by neighboring window or morphological operators and surface construction by RBF.

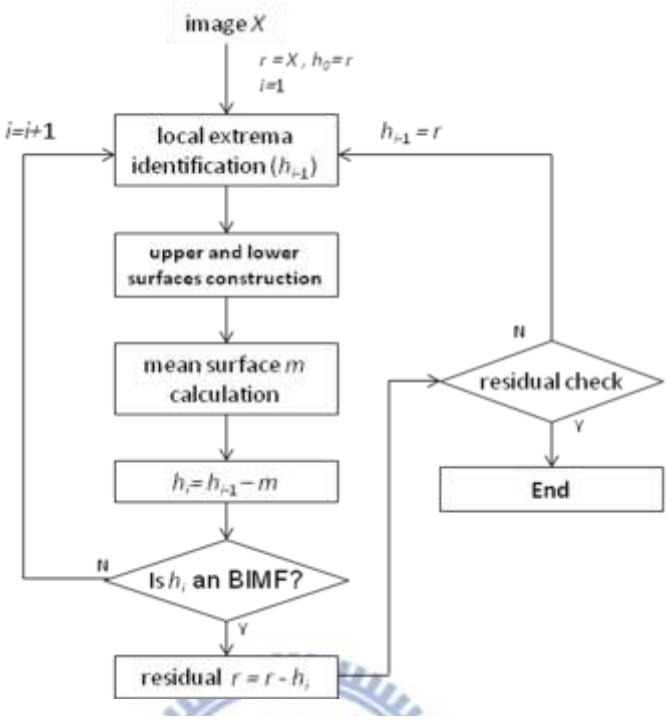


Figure 3. Sifting process of BEMD

The sifting process can be illustrated as follow:

Define X as an image data

(i) Initialize: $r = X$, $h_0 = r$, $i = 1$

(ii) Find the local extrema of h_{i-1} (local maxima and minima, respectively) by neighboring window.

(iii) Interpolate values between all the local maxima and minima respectively to generate upper surface u and lower envelope l by RBF.

(iv) Compute the mean surface m by the equation: $m = (u + l) / 2$.

(v) Calculate the difference between h_{i-1} and current mean surface m.

→ $h_i = h_{i-1} – m$

(vi) Compute standard deviation: $$\mathrm{SD} = \sum_{x=1}^{M}\sum_{y=1}^{N}\left[\frac{\left|h_{i-1}(x,y) - h_i(x,y)\right|^2}{\left|h_{i-1}(x,y)\right|^2}\right]$$

where (x, y) is the position of h and (M, N) is the total number of rows and columns of the decomposed image.

*Note: SD has alternate definitions in other studies [22, 34].

(vii) If SD is greater than a predefined threshold value, which is usually a small nearly zero value, then set iteration index $i = i + 1$ and repeat the sifting process from step (ii);

Otherwise, if SD is less than the predefined threshold, then regard h_i as an BIMF and set intermediate residual $r = r - h_i$. Finally, check r to ensure enough extrema points. If there are enough extrema points, then set $h_{i-1} = r$ and repeat sifting process from step (ii); Otherwise, terminate the sifting process.

It should be noted that different equations of SD affect the number of required sifting iterations to obtain a BIMF. As another stop criterion, fixed number of iterations is an alternate way to obtain BIMFs [27].

2.2.3 Completeness and orthogonality

The sifting results of BEMD also satisfy the completeness condition:

$$X = \sum_{i=1}^{n} C_i + r \qquad (4)$$

in which X is an image data, C_i is the corresponding i^{th} BIMF, n is the total number of BIMFs and r represents residual.

Equation (4) indicates that an original image can be reconstructed by summing up all the BIMFs and the residual. To further check the orthogonality of BIMFs, an equation is given first:

$$X(t) = \sum_{i=1}^{n+1} C_i \ , \ C_{n+1} = r \tag{5}$$

in which residual is regarded as an additional component C_{n+1}. The overall index of orthogonality (IO) for examining the sifting results is defined as:

$$\mathrm{IO} = \sum_{x=1}^{M} \sum_{y=1}^{N} \left(\sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \frac{C_j(x,y) C_k(x,y)}{X^2} \right) \tag{6}$$

A low value of IO is preferred for the local orthogonality among the components.

2.2.4 Extrema detection

In order to detect the distribution of extrema of an image, neighboring window is applied to compare each pixel location of an image with its neighbors (4, 6 or 8 connectivity). If its pixel value is strictly greater/lower than its neighbors, than it is a local maxima/minima point.

Another extrema detection method which is based on morphological operators is a useful operator provided by mathematical morphology [36]. This function finds regions of uniform pixel value whose neighboring regions all have smaller or larger values.

2.2.5 Surface interpolation

The typical BEMD adopts radial basis function (RBF) to perform surface construction. RBF-based interpolation methods are examples of global interpolation methods for scattered data points. They impose fewer restrictions on the geometry of the interpolation centers and are suited to problems where the interpolation centers do not form a regular grid as in the case of local maxima or minima maps of images or textures. However, a minimum number of interpolation centers with which an RBF interpolator can work may pose some limitations on its usefulness.

The scattered data interpolation can be implemented by the following formula:

$$f(x) = \sum_{j=1}^{N} w_j \phi(\| x - o_j \|) + p_m(x) \quad x \in R^d, w_j \in R \qquad (7)$$

Where,

x is the position of sifting components h

o_j are RBF centers,

N is total number of RBF centers,

$\| \ \|$ denotes Euclidean distance,

ϕ is a real-valued function called basis function,

w_j are coefficients of RBF, and

$p_m(x)$ is a low-degree polynomial, typically linear or quadratic.

Some examples of ϕ are given in Table 2.

Table 2. Choices of φ for various RBFs.

RBFs	Choices of ϕ
Linear	$\phi(r) = r$
Cubic splines	$\phi(r) = r^3$
Thin-plate splines	$\phi(r) = r^2 \log(r)$
Hardy's multiquadrics	$\phi(r) = \sqrt{r^2 + c^2}$ [a]
Inverse multiquadrics	$\phi(r) = \frac{1}{\sqrt{r^2+c^2}}$
Exponential splines	$\phi(r) = e^{-cr}$
Gaussian splines	$\phi(r) = e^{-cr^2}$
Compactly supported splines	$\phi(r) = (1-r)^m$ [b]

[a] c is a constant that governs the shape/spread of the basis function
[b] m is a function of spatial dimension

The sifting results, computation cost and the accuracy of different basis functions were studied and compared by Bhuiyan et al. [34]. It has been observed that RBF-cubic, RBF-thin-plate spline, RBF-inverse multiquadric, and RBF-Gaussian interpolators provide better results than the others.

2.3 Partial Least Square

Partial Least Square (PLS) is a method for prediction or feature extraction by finding latent vectors from modeling the relations between independent variables (X) and dependent variables (Y). Traditionally, this modeling of Y by means of X is done using multiple linear regression (MLR), which works well as long as the X-variables are fairly few and fairly uncorrelated, i.e., X has full rank. With modern measuring instrumentation, including spectrometers, chromatographs and sensor batteries, the X-variables tend to be many and also strongly correlated. Therefore, MLR shows its limitation under the situations mentioned above. PLS receives significant attention recently by the fact that it can analyze data with strongly collinear (correlated), noisy, and numerous X-variables, and also simultaneously model several response variables, Y.

Unlike traditional regression methods which directly find maximum covariance between X and Y, it firstly projects X ($n \times m$) and Y ($n \times d$) (where n denotes the sample number and m, d denote the feature dimension of X and Y, respectively) onto their own new spaces, and then the latent variable (T) is used to model their covariance structure.

Consider the general form of PLS algorithm to model the relation between the two blocks of variables:

$$\begin{aligned} X &= TP^{\mathrm{T}} + E \\ Y &= UQ^{\mathrm{T}} + F = TCQ^{\mathrm{T}} + F \end{aligned} \tag{8}$$

where X ($n \times m$) and Y ($n \times d$) are the matrices of zero-mean independent and dependent variables, respectively. T and U are ($n \times a$) matrices of the a extracted latent vectors (components, score vectors). P ($m \times a$) and Q ($d \times a$) represent matrices of loadings. E ($n \times m$) and F ($n \times d$) are the matrices of residuals. C ($p \times p$) is a diagonal matrix with regression coefficient.

2.3.1 PLS1 algorithm

PLS algorithm can be roughly divided into two categories: PLS1 (Y is one dimension) and PLS2 (Y is multi-dimension). The introduction will focus on PLS1 due to the fact that the dependent variables of different samples will only have one dimension in the subsequent application of this thesis.

PLS1 is a special case of PLS2, where Y is a vector in PLS1 and a matrix in PLS2. The algorithm is as follows (training phase):

1. set desired number of principal components $\rightarrow a = \ldots$
   set iteration number $i = 1$
   center X and y $\Rightarrow$ $X = X - \bar{X}$ $y = y - \bar{y}$
2. $w_i = \dfrac{X^T y}{\left\| X^T y \right\|}$
3. $t_i = X w_i$
4. $p_i = \dfrac{X^T t_i}{t_i^T t_i}$
5. $b_i = \dfrac{t_i^T y}{t_i^T t_i}$
6. $X = X - t_i p_i^T$, $y = y - t_i b_i$
7. if $i = a$ or $X \approx$ null matrix or $y \approx 0$
   $\Rightarrow$ terminate the algorithm.
   else
   $\Rightarrow i = i + 1$, repeat from step 2.

Interpretation of PLS1 algorithm:

Step2: we seek the direction in the space of X, which yields the maximum covariance between X and y. This direction is given by a unit vector w_i, and is such that large variations in x-values are accompanied by large variations in the corresponding y-values.

Step3: The score vector t_i is formed as a linear combination of the rows of X with weight w_i .

Step4: The regression coefficient p_i is obtained from simple linear regressions of the columns of X on t_i .

Step5: Regression coefficient b_i is calculated by ordinary linear regression of y on t_i.

Step6: $X = X - t_i p_i^T$, $y = y - t_i b_i$ represents the residuals after regressing X on t_i, and the residuals after regressing y on t_i , respectively.

After the first run through Steps 2-6, the procedure is repeated using the residuals X_2 and y_2. The algorithm then finds the best linear combination of the rows of X_2 for the purpose of predicting y_2, thus picking up any further structure in the connection between X and y not accounted for by t_1. This is repeated on and on, such that each run of the algorithm in principle reveals more and more information about the connection between X and y.

2.3.2 Prediction (testing) phase of PLS1

Now a new testing sample $x_{pred}(m \times 1)$ is about to be predicted, and do as follows :

1. Initialize iteration number $i = 1$, and $x = x_{pred} - x_{mean}$
2. $t_i = xw_i$ (w_i is obtained from training phase)
3. $x_{i+1} = x_i - t_i p_i^T$
4. stop if $i = a$; else repeat from step 2.

Now we have form a score vector set $t = [t_1\ t_2\ \ldots t_a]$, and the predicted y_{pred} is

$$y_{pred} = y_{mean} + tb \qquad (9)$$

in which $b = [b_1\ b_2\ \ldots\ b_a]$ is obtained from training phase

2.4 Tri-linear Partial Least Square

Tri-linear Partial Least Square (Tri-PLS) is an extension of PLS to analyze three-way data by a folding scheme, whereas PLS applies an unfolding scheme to analyze three-way data. Three-way data can occur when the variables are characterized by a matrix instead of a vector. When applying unfolding scheme, the three-way data (cube of the calibration matrices) are unfolded to give a matrix that

contains each sample in one row. This is done by concatenating all rows of matrix of one.

Applying Tri-PLS to analyze data is mostly when the data is organized as the form of three-way. For example, an image database is the form of three-way data, since an image corresponds to a matrix and all the images in the database can be seen as a cube of matrices, which is a three-way data. It is not necessary to use PLS to analyze three-way data, since PLS needs to unfold all the rows or columns of one sample to a vector for all the samples and combines all the vectors in one matrix, whereas Tri-PLS can handle the cube of matrices directly. Therefore, it is intuitive to use Tri-PLS to handle the three-way data and the advantages of Tri-PLS compared with PLS are mentioned in [37].

2.4.1 Tri-PLS1 algorithm

Tri-PLS algorithm can be divided into two categories as the same as PLS: Tri-PLS1 (*Y* is one dimension) and Tri-PLS2 (*Y* is multi-dimension). The introduction will focus on Tri-PLS1 due to the fact that *Y* will only be one column in the subsequent application of this thesis.

Tri-PLS1 is a special case of Tri-PLS2, where *Y* is a vector in Tri-PLS1 and a matrix in Tri-PLS2. The algorithm of Tri-PLS1 is given as follows (training phase):

set iteration number $i = 1$, and initialize $T =$ null matrix

center X and $y \Rightarrow \quad X = X - \bar{X} \quad y = y - \bar{y}$

1. Caluculate the matrix $Z = X^T y$
2. Determine w^J and w^K by SVD $\Rightarrow \quad (w^J, w^K) = \text{SVD}(Z)$
3. Caluculate the combined weight ve ctor using Kronecker product $\Rightarrow \quad w_i = w^k \otimes w^J$
4. Calculate latent vec tor $t \quad \Rightarrow \quad t = Xw_i$
5. add t into T as a column vec tor $\Rightarrow \quad T = [T\ t]$
5. Regress y on $T \quad \Rightarrow \quad b = (T^T T)^{-1} T^T y$
6. Calculate the residul of X and $y \quad \Rightarrow \quad X = X - tw_i^T \quad y = y - Tb$
7. $i = i + 1$ Continue from step 1 until proper description of y

2.4.2 Prediction (testing) phase of Tri-PLS1

Now a new testing sample X_{pred} is about to be predicted, and do as follows :

1. Initialize iteration number $i = 1$, and $X = X_{pred} - X_{mean}$
2. $t = XW$ ($W = [w_1\ w_2\ w_3\ \ldots\ w_a]$ is obtained from training phase)
3. $y_{pred} = tb$ (b is obtained from training phase)

2.5 PLS templates

In this study, a different approach which replaces the interpolation process of BEMD by selecting PLS templates is proposed. The templates are obtained by the calibration process of Tri-PLS on some well-known face databases.

2.5.1 Linkage between Tri-PLS and BEMD

The linkage between Tri-PLS and BEMD is shown in figure 4.

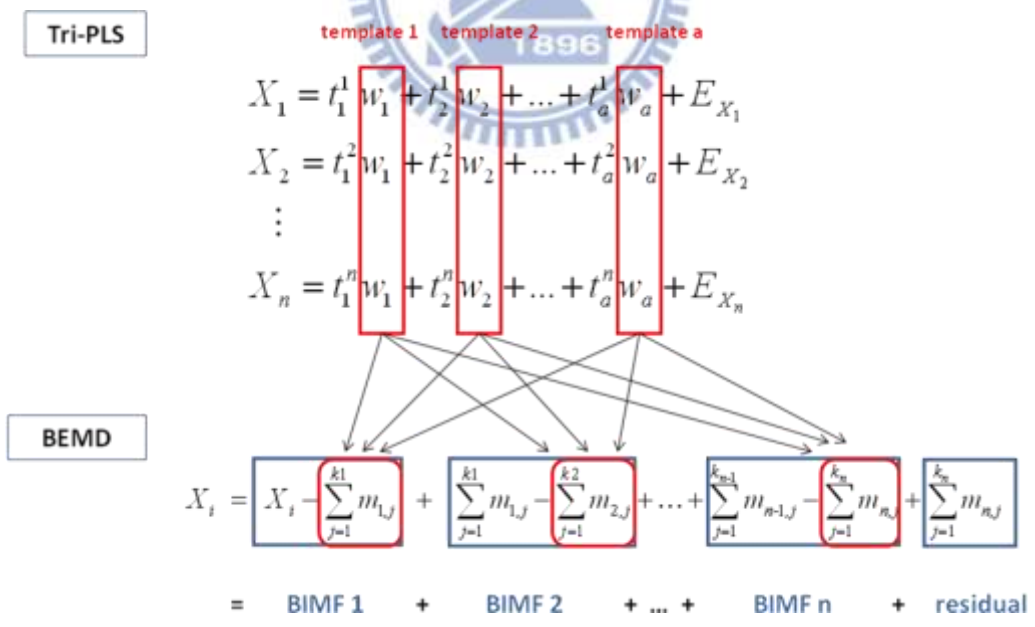


Figure 4. Linkage between Tri-PLS and BEMD

The lower part of figure 4 represents the decomposition result of one image in the database by the sifting process of BEMD. As introduced in section 2.2.3, the sifting result of BEMD follows the characteristic "completeness", i.e. an image can be represented as the summation of BIMFs and the residual. Each BIMF can be represented mathematically as the rectangle parts displayed in the lower part of figure 4. Traditionally, the sifting process of BEMD for obtaining each BIMF is to iteratively compute the mean surface *m* of the sifting component by some interpolation methods such as RBF, Delaunay triangulation and finite elements. Instead of constructing surfaces by the methods mentioned above, a new thought of surface construction is derived from Tri-PLS. The calibration results of Tri-PLS in the upper part of figure 4 is shown that each image X_i (i=1,2,…,n) in the database can be represented as the summation of partial regression quantity $t_a^i w_a$ (i=1,2,…,n; a is the number of latent vector) and the final regression error E_{xi}. It should be noticed that the size of the loading weights w_j (j=1,2,…,a) are identical to that of the images in the database and each score t_a^i (i=1,2,…,n) is a scalar. For example, if the images in the database are of the size 128x128, then all the w_j (j=1,2,…,a) are of the same size 128x128. We can find that the loading weights of each X_i are identical in first iteration (i.e. w_1) and are identical in second iteration (i.e. w_2), and so on. The loading weights w_j (j=1,2,…,a) can be seen as the common quantities of all images in the database described by the calibration process of Tri-PLS. In addition to the common quantities, the differences of the images can be discriminated by multiplying the scores t_k^i (i=1,2,…,n; k=1,2,…,a) by the common quantity w_k. We can therefore serve the common quantities w_j (j=1,2,…,a) as the templates of the selected face database and then apply these templates into the sifting process of BEMD to replace previous interpolation process.

How to select the templates in each sifting iteration of BEMD and how to adjust the templates in order to preserve the characteristic of the sifting process will be introduced in the following sections.

2.5.2 Surface construction by selecting PLS templates

A new approach of surface construction by selecting PLS templates is proposed. The modified sifting process, which is shown in figure 5, starts from detecting local extrema (local maxima and local minima, respectively) of the current sifting components h_i and then compare h_i with PLS templates w_1, w_2, ... , w_a. Two templates selected from these candidates will be served as the upper and the lower surfaces respectively (the detailed selecting process will be introduced in section 2.5.3). A further modification for the two selected templates will be performed in order to fit the local maxima and minima of h_i, respectively (the detailed adjusting process will be introduced in section 2.5.4). The mean surface is determined by calculating the mean of the two modified templates w_i and w_j. Once the mean surface is obtained, we can further obtain a new sifting component h_{i+1} by the equation $h_{i+1} = h_i - m$. Check whether h_{i+1} is BIMF or not and then repeat the sifting process just the same as that of BEMD.

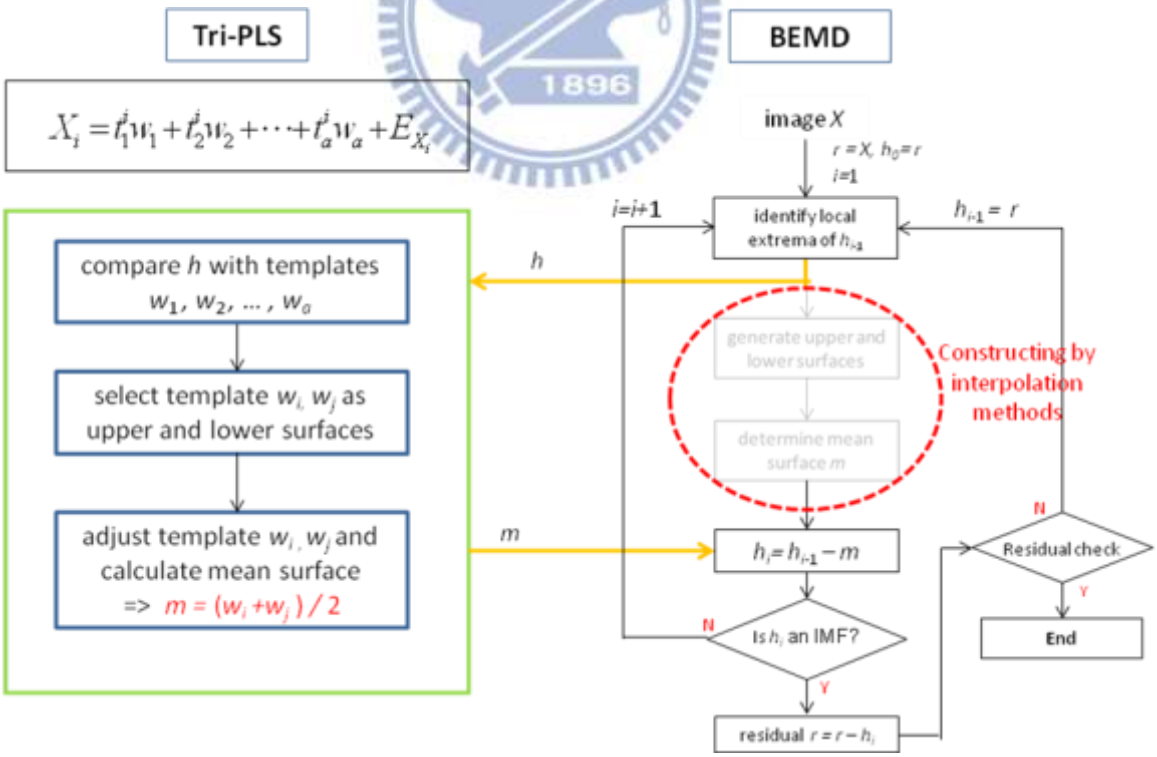


Figure 5. Surfaces construction by PLS templates

2.5.3 Selecting template from candidate

In order to select templates as upper and lower surfaces, the method proposed here is to decide which templates w_i has maximal correlation with the sifting component *h*. For instance, after determining the local maxima and minima of the sifting component *h*, the process compares local maxima and minima respectively with the intensity values of corresponding positions of all templates and then selects one template which has maximal correlation coefficient with *h*.

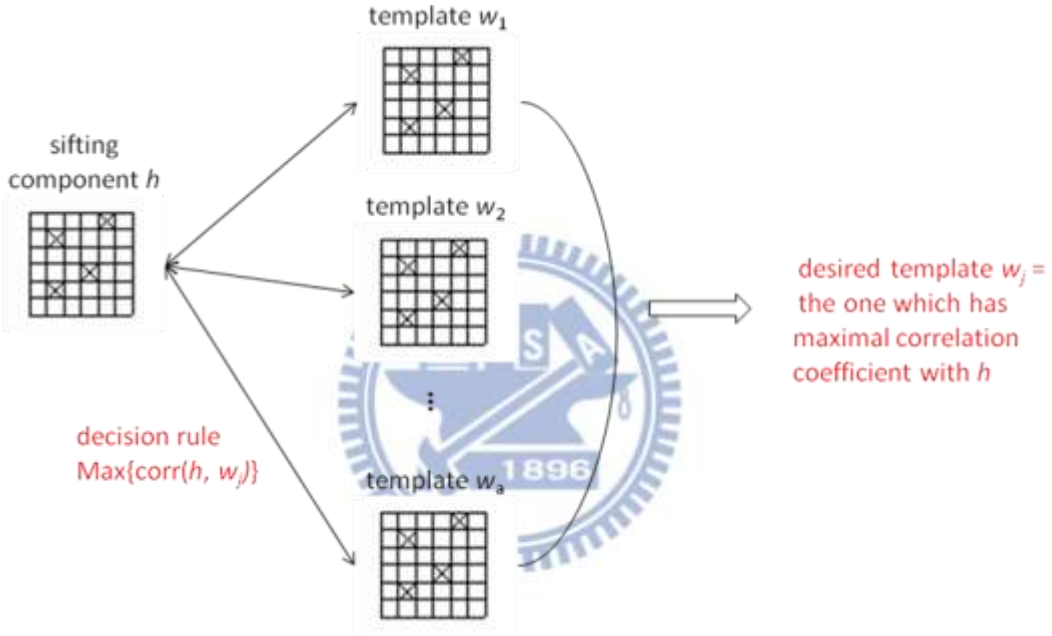


Figure 6. Template selecting

2.5.4 Adjusting the selected template

Selecting templates is the first step to construct the upper and lower surfaces; however, adequate modification is necessary for the selected template to be applied to the sifting process.

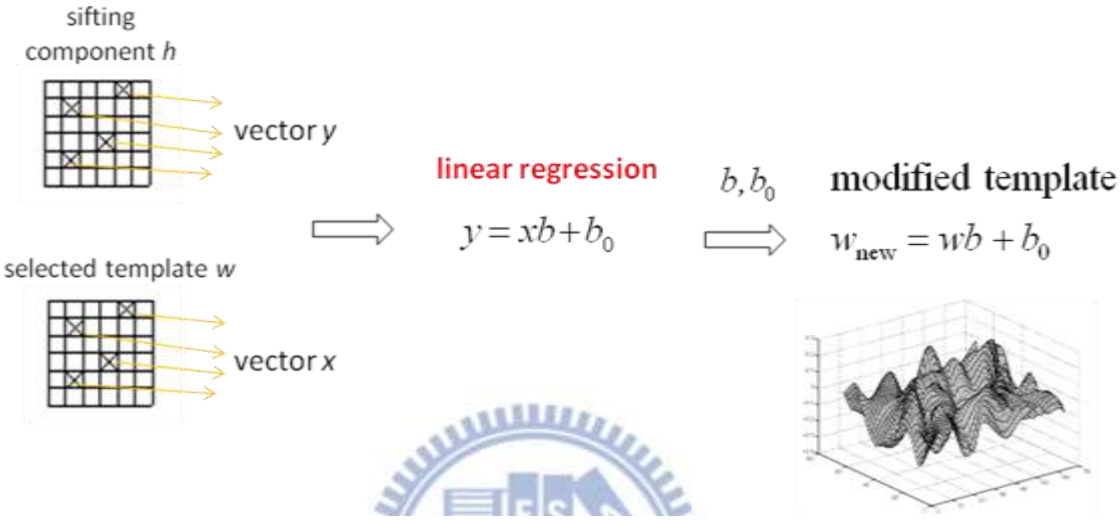


Figure 7. Adjusting templates by linear regression

For the purpose of preserving the sifting property of EMD-based methods, we have to consider the local extrema of the sifting component in each sifting iteration. Linear regression is adopted here to build a minimal sum of square error relation between the local extrema of the sifting component and the intensity values of corresponding positions of the selected template. The modification procedure is show in figure 7.

After applying linear regression, we obtain the regression coefficients b and error term b_0. The modified template can be obtained by multiplying b by the selected template and adding the error term b_0. The modified template is now regarded as the constructed surface.

III. Simulation result

3.1 Recognition procedure

In this section, the self-organization scheme method BEMD and the proposed cross-correlation scheme method, namely PLS-based EMD (PLS-EMD) are compared. This compared method BEMD is a typical type of BEMD which is proposed by Nunes in 2003. The adopted interpolation method is thin-plate spline which is a specious case of RBF. As shown in figure 8, a testing image is firstly preprocessed by BEMD or PLS-EMD, after the preprocessing stage, the features of this preprocessed image are extracted in the feature extraction stage and finally a classifier is applied to determine which class this testing image belongs to. The selected feature extraction method is PCA and the classifier is KNN. It should be noted that instead of using all the preprocessing results to extract features, the first mode (i.e. 1st BIMF) of BEMD and PLS-EMD are applied to extract features since the high spatial frequency component is believed more robust to illumination changes [38].

In order to examine whether the proposed method is an effective method to enhance recognition accuracy under different illumination conditions, two databases (i.e. Yale and PIE) are used in the simulation process. In addition to BEMD and PLS-EMD, two famous methods Eigenface and Tri-PLS are further taken into comparison. Eigenface and Tri-PLS both extract global features of face images and therefore they are believed more sensitive to illumination changes.

Figure 8. Recognition procedure of BEMD and PLS-EMD

3.2 Evaluation methods and databases

In the evaluation stage, repeated random sub-sampling validation and Leave-one-out cross-validation (LOOCV) were used to evaluate the performance of Eigenface, Tri-PLS, BEMD and PLS-EMD under different illumination conditions. The databases used in the simulation process are Yale database and PIE database.

3.2.1 Repeated random sub-sampling validation

In repeated random sub-sampling validation [39], a subject of training samples for each class is randomly selected, and the remaining samples are as testing samples. For each of the random sub-sampling procedure, it is repeated 30 times to take an average recognition rates for every database.

3.2.2 Leave-one-out cross-validation

Due to the fact that some sample might not be selected as testing sample in the above repeated random sub-sampling validation, further complete verification LOOCV is introduced. Considering a dataset with n samples, at each of the LOOCV procedure, one sample is selected as the testing sample, and the remaining n-1 samples are as training samples, thus n times iteration of LOOCV is executed. In LOOCV scheme, all the samples will be considered as testing samples and every sample will not be training and testing sample simultaneously in each LOOCV procedure

3.3 Yale database

Yale database is a well-known database in the field of face recognition. It is designed by the team of Yale University. The variations in Yale database are composed of illumination, expression and eyewear. There are 15 classes of different people with 11 images per person (i.e., 165 images in total). Each image is resized to 73x96 pixel in the simulation process. An example of Yale database is given in figure 9.

Figure 9. Example of Yale database

3.3.1 Preprocessing result of PLS-based EMD

The preprocessing results of PLS-EMD are given in figure 10. The 4 original images mainly differ from the illumination conditions. The light orientation of first image is from left side of the face and therefore the shadow is presented on the right side of background; the light orientation of second image is opposite to first image, therefore, the shadow is presented on the left side of background. An interesting observation is that the illumination effects of the 4 original images are presented in the residual respectively. That is the dark and bright regions in the original images' background are presented in the corresponding areas of the residual. A further observation is that each 1st BIMF is like a processed image by removing the illumination effects from original images. The illumination effects of the original

images can be seen as the low frequency part of spatial frequency. Therefore the illumination effects will be extracted in the residual due to the sifting property of EMD-based method.

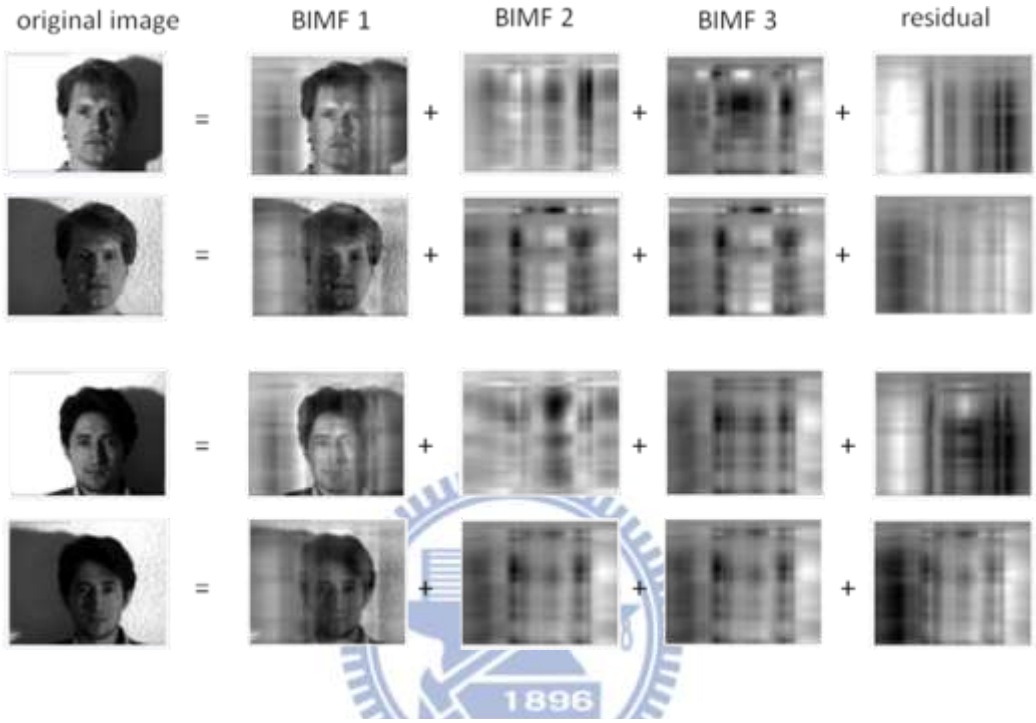


Figure 10. Preprocessing results of PLS-EMD – Yale database

3.3.2 Sub-sampling validation result of Yale database

In this experiment, the size of the training sample for each class is chosen to be 1, 3 and 5; in other words, there are 15, 45 and 75 training samples out of 165 samples. The recognition results of different methods are displayed in figure 11. The recognition rates of Eigenface and Tri-PLS are related to the number of selected principal components, therefore the horizontal axis is the number of selected principal components and the vertical axis is the corresponding recognition rate; however, the recognition rates of BEMD and PLS-EMD are independent to the principal components, thus the recognition rates of the two are just plotted on the graph by arrows. The number of templates of PLS-EMD applied in the simulation process is 30. The adopted standard deviation for both BEMD and PLS-EMD are defined as follows:

$$\mathrm{SD} = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left|h_{j+1}(x,y) - h_j(x,y)\right|^2}{\sum_{x=1}^{M}\sum_{y=1}^{N}\left|h_j(x,y)\right|^2}$$

And the threshold is 0.01 and 0.1 respectively.

The results show that BEMD and PLS-EMD outperform than Eigenface and Tri-PLS in the three cases of different number of training samples. The possible reasons of this observation might due to the fact that frequency-based methods are more robust to illumination changes, while Eigenface and Tri-PLS do not belong to this category. Eigenface and Tri-PLS take the whole region of an image into account and therefore the two methods are sensitive to illumination changes. Another observation is that the cross-correlation scheme PLS-EMD outperforms than the self-organization scheme BEMD in most cases except for the first case which the recognition rate of BEMD is 0.61 and that of PLS-EMD is 0.60.

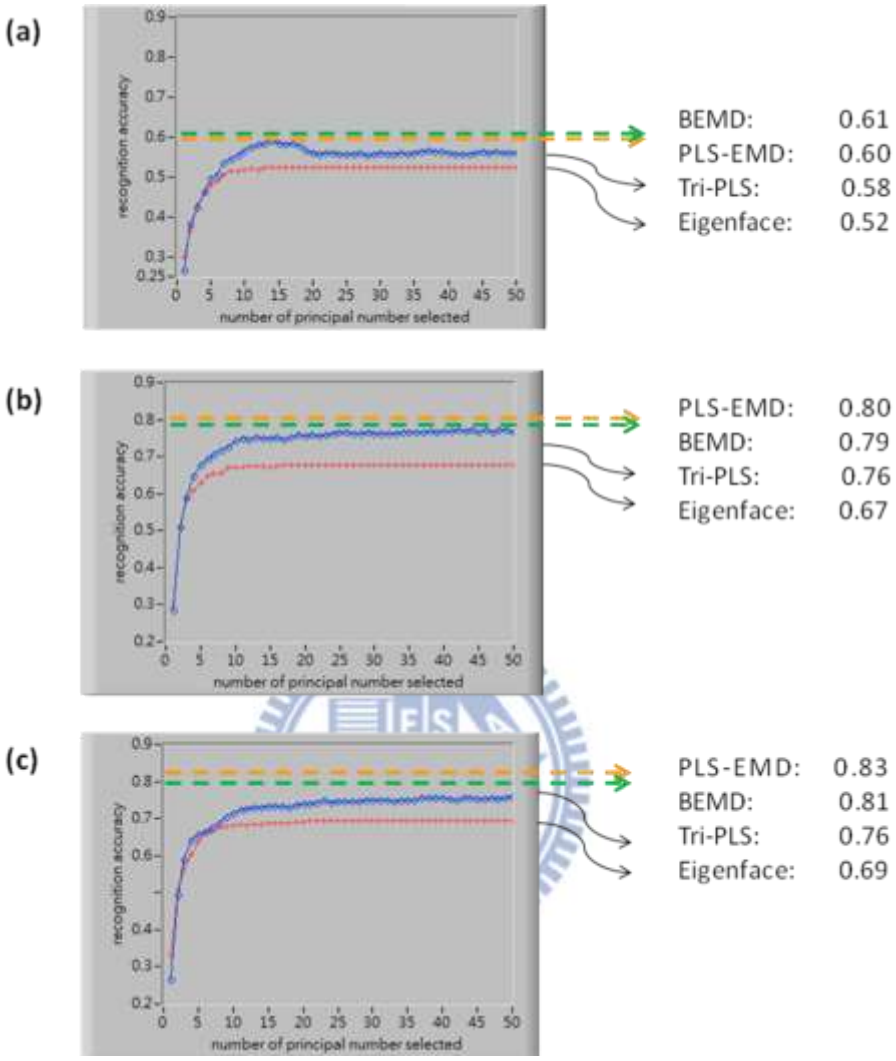


Figure 11. Results of sub-sampling validation – Yale database
(a) 1 training image and 10 testing images per class
(b) 3 training images and 8 testing images per class
(c) 5 training images and 6 testing images per class

3.3.3 LOOCV result of Yale database

In LOOCV procedure, the parameter setting of each method is identical to that of sub-sampling validation results mentioned above. The recognition accuracy of PLS-EMD outperform than the other three methods.

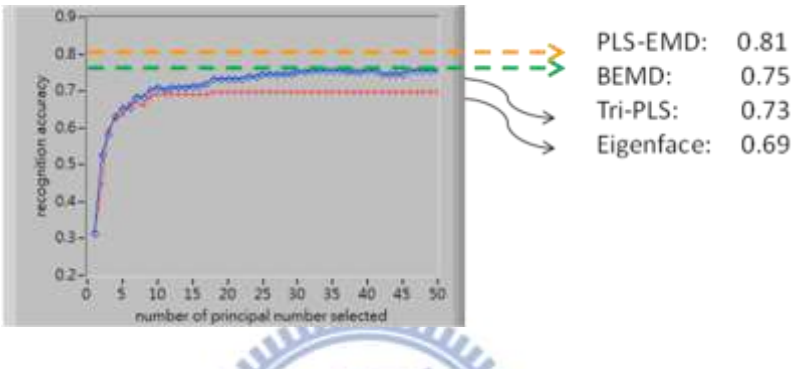


Figure 12. Results of LOOCV – Yale database

3.4 PIE database

PIE database is another well-known database in the field of face recognition. It is designed by the team of Carnegie Mellon University. PIE database contains 41368 images of 68 people, with each person under 13 different poses, 43 different illumination conditions, and 4 different expressions. In this experiment, we choose a subset with only the frontal view (set c27) under different illumination conditions with total images of 1428. All the selected images are cropped and resized to 32x32 pixel. An example of PIE database is given in figure 13.

Figure 13. Example of PIE database

3.4.1 Preprocessing result of PLS-based EMD

The preprocessing results of PLS-EMD are given in figure 14. The 4 original images mainly differ from the illumination conditions. The light orientation of first image is from right side of the face and therefore the shadow is presented on the left face; the light orientation of second image is opposite to first image, and therefore the shadow is presented on the right side of the face. An observation which is similar to that of the preprocessing results of Yale database is that the illumination effects of the 4 original images are presented in the residual respectively. That is the dark and bright regions of the face are presented in the corresponding areas of the residual. A further observation is that 1st BIMF is like a processed image by removing the illumination effects from original images. The illumination effects of the

original images can be seen as the low frequency part of spatial frequency. Therefore the illumination effects will be extracted in the residual due to the sifting property of EMD-based method.

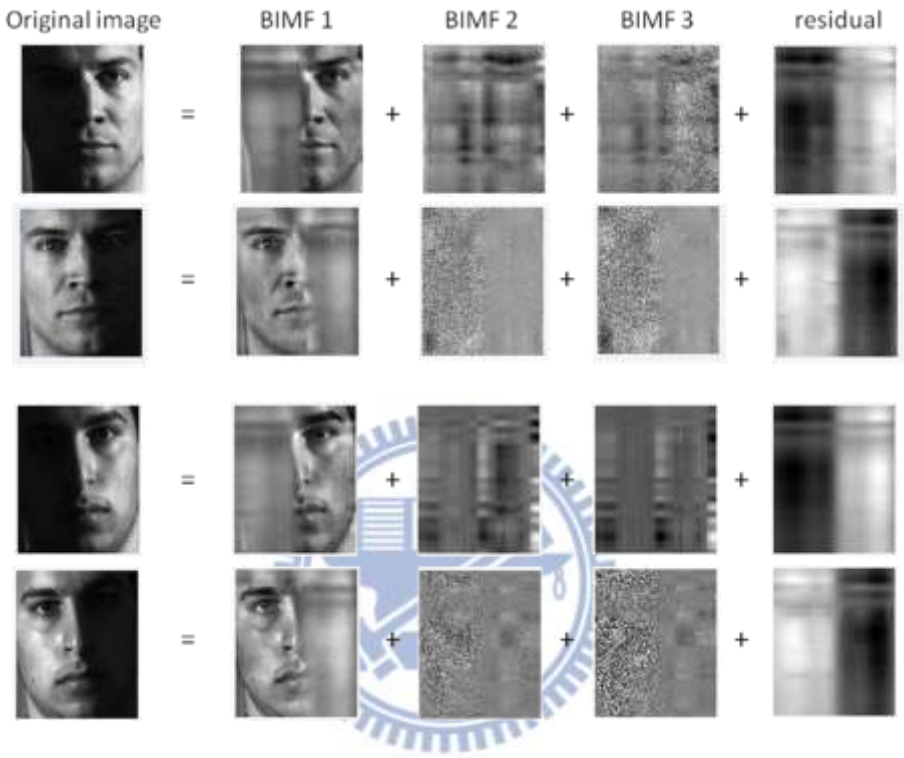


Figure 14. Preprocessing results of PLS-EMD – PIE database

3.4.2 Sub-sampling validation result of PIE database

In this experiment, the size of the training sample for each class is chosen to be 4, 8 and 12; in other words, there are 272, 544 and 816 training samples out of 1428 samples. The recognition results are compared in figure 15. The number of templates of PLS-EMD is still 30. The adopted SD type is identical to the simulation process of Yale database and the threshold of BEMD and PLS-EMD are 0.01 and 0.1 respectively.

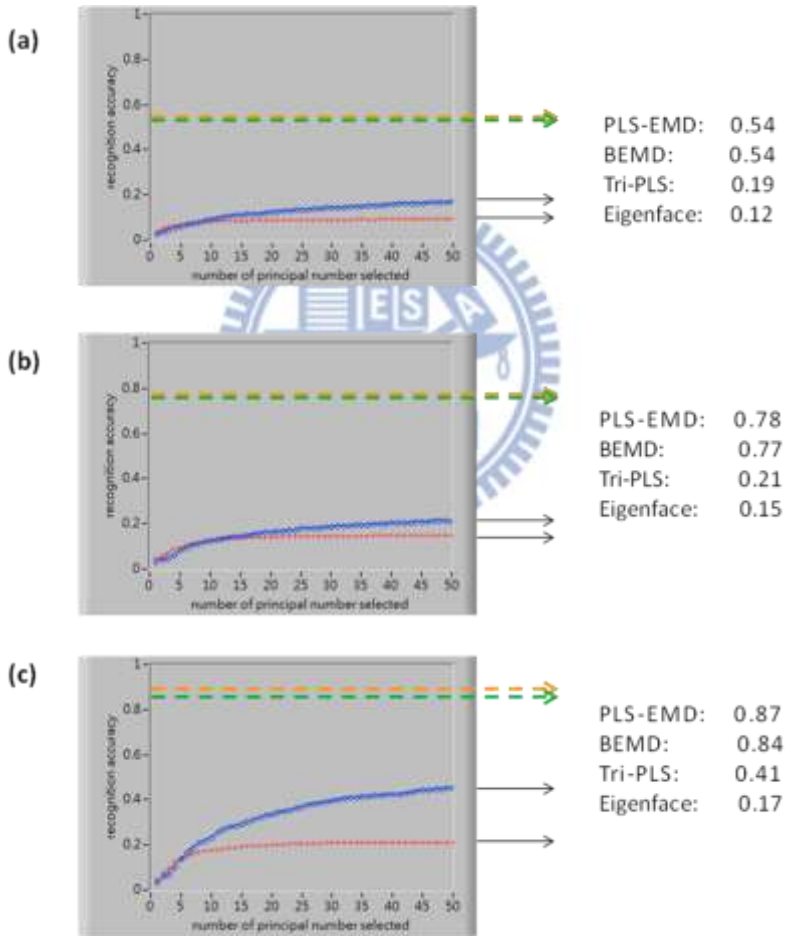


Figure 15. Results of sub-sampling validation – PIE database
(a) 4 training images and 17 testing images per class
(b) 8 training images and 13 testing images per class
(c) 12 training images and 9 testing images per class

It is obvious that BEMD and PLS-EMD outperform than Eigenface and Tri-PLS in all the cases of different training samples. There is a significant gap in the recognition rate between the two groups since PIE database has large illumination variations among different face images and Eigenface and Tri-PLS suffer from the severe illumination effects. The frequency-based methods (BEMD and PLS-EMD) are insensitive to large illumination variations and therefore are more robust to the illumination effects. Another observation is that the cross-correlation scheme PLS-EMD outperforms than the self-organization scheme BEMD in most cases except for the first case which the recognition rate of BEMD and PLS-EMD are equal.

3.4.3 LOOCV result of PIE database

In LOOCV procedure, the parameter settings of each method are identical to that of sub-sampling validation results mentioned above. The recognition accuracy of PLS-EMD outperform than the other three methods.

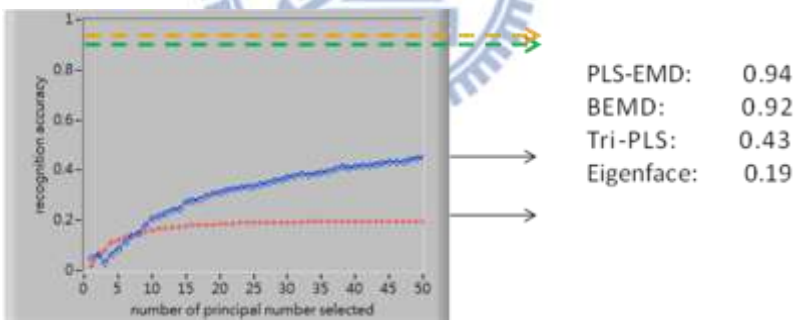


Figure 16. Results of LOOCV – PIE database

IV. Discussion

In this chapter, the performance issue of PLS-EMD and the validation of BIMFs obtained by PLS-EMD are discussed. The global mean of BIMF is regarded as the validation index to examine whether the obtained BIMFs satisfy the original characteristics. As for the performance issue, the number of iterations needed for BIMFs under different stop criteria and the computation time are considered as the performance index. The following discussion will focus on the comparison of EMD-based methods between self-organization scheme (typical BEMD) and cross-correlation scheme (PLS-EMD).

4.1 Computing performance of sifting process

Different types of SD will lead to different sifting results and thus the different number of iterations needed for BIMFs. To compare the performance of PLS-EMD with BEMD, two types of SD are adopted in the simulation process. The definitions are given as follows:

$$\mathrm{SD1} = \sum_{x=1}^{M}\sum_{y=1}^{N}\frac{\left|h_{j+1}(x,y)-h_j(x,y)\right|^2}{\left|h_j(x,y)\right|^2}$$

$$\mathrm{SD2} = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left|h_{j+1}(x,y)-h_j(x,y)\right|^2}{\sum_{x=1}^{M}\sum_{y=1}^{N}\left|h_j(x,y)\right|^2}$$

The corresponding simulation results under SD1 and SD2 are given in Table 3, 4 and Table 5, 6 respectively.

Table 3. Average number of iterations for BIMFs under SD1
(The threshold values of BEMD and PLS-EMD are 0.1 and 500 respectively.)

	1st BIMF		2nd BIMF		3rd BIMF	
	Yale	PIE	Yale	PIE	Yale	PIE
BEMD	27.58±15.73	57.31±20.53	20.83±7.10	62.78±60.32	15.63±8.09	35.11±16.14
PLS-EMD	2.99±1.33	3.34±1.22	9.57±11.67	16.06±14.97	12.50±12.86	20.82±16.83

Table 4. Average computation time per image (second / image) under SD1
(The threshold values of BEMD and PLS-EMD are 0.1 and 500 respectively.)

	Yale	PIE
BEMD	1.86±1.00	53.88±19.18
PLS-EMD	0.70±0.04	7.06±0.10

We can observe from Table 3 that the average number of iterations needed for BIMFs of PLS-EMD is fewer than that of BEMD. This result shows that the sifting process of PLS-EMD for searching BIMFs converges faster than that of BEMD. Here, only the comparison results of first three BIMFs are shown since the comparison results of other BIMFs are consistent with the observation of first three BIMFs. Therefore, only the first three BIMFs are examined in the following discussion. Another comparison in Table 4 is to compare the average computation time per image for both Yale and PIE databases. PLS-EMD takes less time than BEMD on both databases. It should be noticed that the threshold values of PLS-EMD and BEMD are 500 and 0.1 respectively. Literature [34] has shown that the value 0.1 is regarded as an appropriate threshold value for BEMD under SD1. The value 500 of PLS-EMD is empirically obtained from the simulation results. The results indicate that the threshold value of PLS-EMD for obtaining BIMFs is not necessary to set as small as that of BEMD, therefore the spending time of sifting process of PLS-EMD is less than that of BEMD.

Table 5. Average number of iterations for BIMFs under SD2
(The threshold values of BEMD and PLS-EMD are 0.01 and 0.1 respectively.)

	1st BIMF		2nd BIMF		3rd BIMF	
	Yale	PIE	Yale	PIE	Yale	PIE
BEMD	6.93±1.27	6.93±1.27	4.48±1.14	5.84±1.30	2.44±1.46	4.84±1.21
PLS-EMD	2.89±0.61	2.53±0.57	22.69±24.36	35.95±18.56	31.05±24.39	41.09±18.88

Table 6. Average computation time per image (second / image) under SD2
(The threshold values of BEMD and PLS-EMD are 0.01 and 0.1 respectively.)

	Yale	PIE
BEMD	0.29±0.07	4.35±1.06
PLS-EMD	0.14±0.10	1.84±0.09

The comparison of the average computation time per image under SD2 is given in Table 6. PLS-EMD also takes less time than BEMD on both databases. It should be noticed that the threshold values of PLS-EMD and BEMD are 0.1 and 0.01 respectively. Literature [34] has shown that the value 0.01 is an appropriate threshold value for BEMD under SD2. As for PLS-EMD, the value 0.1 is also empirically obtained from the simulation process. The threshold value of PLS-EMD for obtaining BIMFs is also not necessary to set as small as that of BEMD, therefore the spending time of sifting process of PLS-EMD is less than that of BEMD.

The average number of iterations needed for BIMFs under SD2 is compared in Table 5. The number of iterations needed for the first BIMF of PLS-EMD is less than that of BEMD on both Yale and PIE databases; however, the number of iterations needed for the second and the third BIMFs of PLS-EMD are more than that of BEMD. A detailed investigation of this observation will be given in section 4.2.1.

4.2 Evaluation of BIMFs

In addition to the performance issue, another issue what we concern about is whether BIMFs obtained by PLS-EMD satisfy the original characteristics. Two indices are utilized to evaluate BIMFs found by BEMD and PLS-EMD. One is the global mean of BIMF and the other is the index of the orthogonality (IO).

4.2.1 Global mean of BIMF

A 2D envelope can be seen as BIMF when it satisfies the condition that the mean value (local mean) of the 2D mean envelope (i.e. surface) is zero or nearly zero at any points. The term 'Global mean' indicates the mean value of the local mean of the sifting component (i.e. intermediate surface). In order to investigate the process of obtaining BIMFs, the value of SD and the global mean of the intermediate sifting components during sifting process will be checked. It should be noticed that the stop criteria here is adopted SD2 since the observation of more iterations of the second and the third BIMFs of PLS-EMD under SD2 than that of BEMD need to be investigated.

The variations of the value of SD of intermediate sifting components during sifting process are shown in figure 17(a). The horizontal axis denotes the number of iterations of the sifting process for finding each BIMF and the vertical axis denotes the corresponding value of SD of the intermediate sifting components. ' • ' stands for the sifting process for finding first BIMF, '×' for that of the second BIMF, and ' ◦ ' for that of the third BIMF. It is obvious that the value of SD of the intermediate sifting component drops nearly zero within 10 iterations during the sifting process of finding each BIMF. If the value of SD of the intermediate sifting component is smaller than a predefined threshold value which is usually a small value close to zero, then this intermediate sifting component is regarded as BIMF. Figure 17(b) shows that the global mean of the intermediate sifting component drops nearly zero within 10 iterations during the sifting process for each BIMF. We can find that the variations of the global mean are similar to that of the value of SD; that is to say, the first three BIMFs can be obtained within 10 iterations by checking the value of SD; similarly, if we adopt global mean as the stop criteria, then the first three BIMFs can also be

obtained within 10 iterations since the local mean of the sifting components is nearly zero at each point which ensures the obtained BIMFs can satisfy the original characteristic of BIMF.

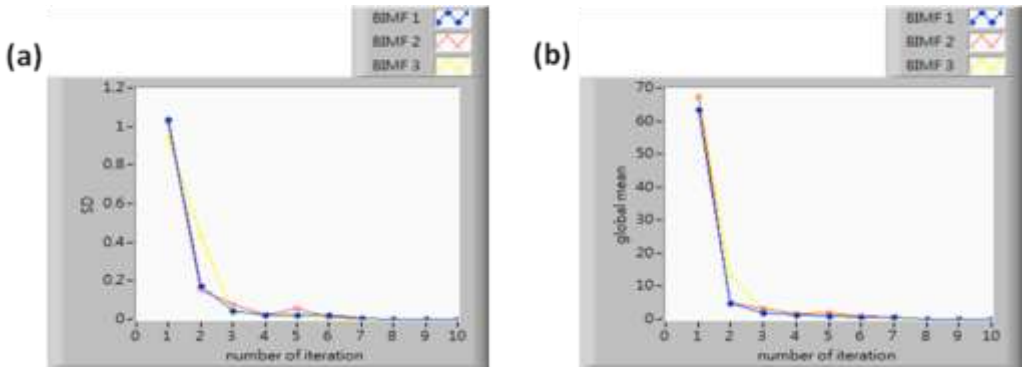


Figure 17. Variations of the value of SD and the global mean during sifting process
(a) Variations of the value of SD
(b) Variations of the global mean

Next, the sifting process of PLS-EMD is examined. Figure 18(a) shows that the first BIMF can be obtained within 10 iterations since the value of SD is approximately nearly zero and will be smaller than the predefined threshold value; however, the second and the third BIMFs are obtained over 50 iterations. This simulation result is consistent with the above comparison results in section 4.1 that the numbers of iterations needed for the second and the third BIMFs of PLS-EMD are more than that of BEMD. A further examination of global mean is given in figure 18(b). The global mean of the intermediate sifting component is approximately nearly zero within 10 iterations for finding each BIMF. It is reasonably to infer that BIMFs will be obtained within 10 iterations since the global mean is nearly zero; however, this finding is inconsistent with the results of figure 18(a). A possible inference is given that checking the value of SD to obtain BIMFs will lead to additional and unnecessary iterations for searching the second and the third BIMFs. Global mean is probably a better choice of the stop criteria of PLS-EMD.

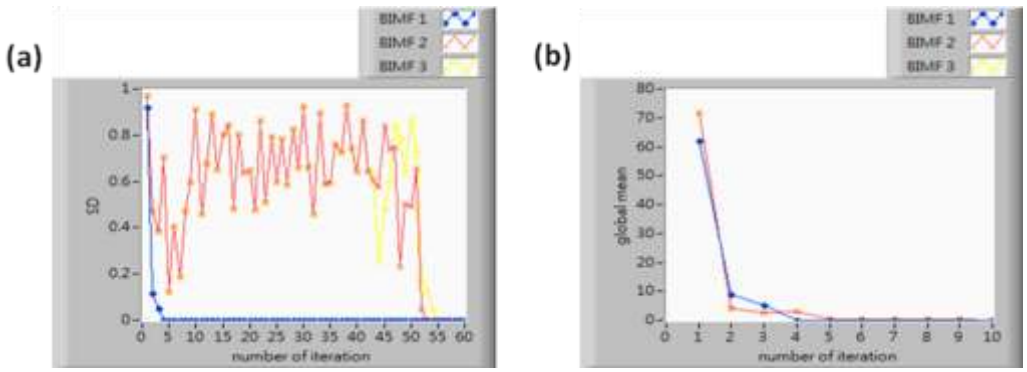


Figure 18. Variations of the value of SD and the global mean during sifting process

(a) Variations of the value of SD

(b) Variations of the global mean

4.2.2 Index of orthogonality

For the purpose of examining the orthogonality of components (BIMFs and residual) obtained by BEMD and PLS-EMD, the index of orthogonality (IO) of the components is computed and is given in Table 7. As introduced in section 2.2.3, a low value of IO is preferred for orthogonality among components. It is obvious that the value of IO of BEMD is smaller than that of PLS-EMD when both Yale and PIE databases are taken into account; however, the differences of the value are small. We can say that the decomposition result of BEMD is little better than that of PLS-EMD when IO is taken into account.

Another viewpoint is how significance of orthogonality it is to the subsequent applications? Just as the results displayed in chapter III, the recognition rate of PLS-EMD is better than that of BEMD, but the orthogonality among the components obtained by PLS-EMD is worse than that of BEMD. We apply the quantity measure (i.e. IO) here to compare the decomposition results of PLS-EMD and BEMD; however, the orthogonality among the components might not important at all for the recognition purpose since we only utilize 1st BIMF as the preprocessed result to extract features. Therefore a conclusion is given that whether IO is a means of comparing the EMD-based methods or not depends on the applications themselves.

Table 7. Comparison of index of orthogonality

	IO	
	Yale	PIE
BEMD	0.14±0.10	0.11±0.09
PLS-EMD	0.15±0.09	0.15±0.07

4.2.3 Illumination tendency in residual

Literatures [22, 27] have shown that the residual of the decomposition results of BEMD presents the trend of the decomposed image. The trend of a face image can be seen as the illumination effect or illumination tendency of this image. Three face images with different illumination conditions are given in figure 19(a). The residuals decomposed from figure 19(a) by BEMD and PLS-EMD are given in figure 19(b) and 19(c) respectively. We can find that PLS-EMD has the ability to find the illumination tendency in the residual just the same as that of BEMD. This observation is consistent with the fact that EMD-based methods have the ability to extract low frequency part of spatial frequency into the residual. Another observation from figure 19(c) is that the residual found by PLS-EMD has an effect of enhancing the contrast of the illumination tendency compared with the results of BEMD. However, there is no quantity measure here to evaluate which illumination tendency found by BEMD and PLS-EMD is better. We just can say PLS-EMD has an effect of enhancing the contrast of illumination tendency compared to BEMD.

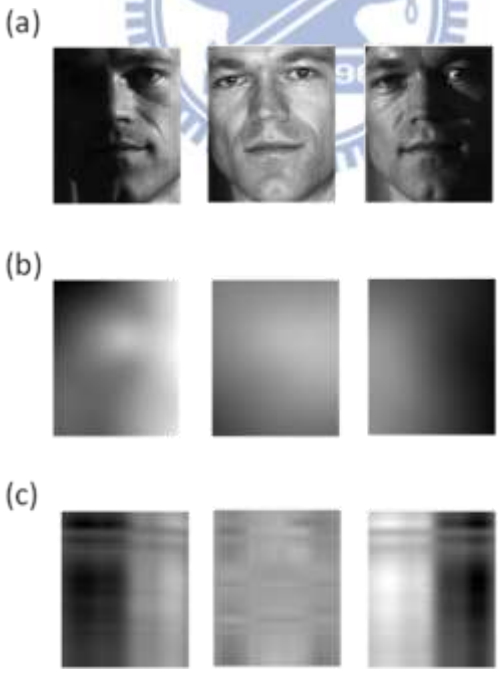


Figure 19. (a) Face images with different illumination conditions
(b) Residuals decomposed from figure 19(a) respectively by BEMD
(c) Residuals decomposed from figure 19(a) respectively by PLS-EMD

V. Conclusion

In this study, a new thought of surface construction by a different approach, i.e. cross-correlation scheme for implementing 2D-EMD is proposed. The proposed method is based on selecting templates obtained from the calibration process of Tri-PLS to replace RBF in typical BEMD. It is shown that the recognition rate of the proposed method PLS-EMD is better than that of BEMD under different illumination conditions and the performance of PLS-EMD is also better than that of BEMD.

VI. Future Work

The proposed method still has some implementation issues about templates selection and the subsequent adjustment. The sifting result of PLS-EMD reveal the mode mixing problem to some extent from the preprocessing results displayed above. Are there any alternate methods to select the candidate templates and adjust the selected template to get better recognition results? How to relieve the mode mixing problem? It is valuable to investigate this issue in the future.

References:

[1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–459, Dec. 2003.

[2] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.

[3] A.K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, Jan. 2004

[4] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, May. 1995.

[5] Y. Adini, Y. Moses and, S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.19, no. 7, pp.721–732, Jul.1997.

[6] Z.C. Lian, M.J. Er, and Y.C. Liang, "A novel efficient local illumination compensation method based on DCT in logarithm domain," *Pattern Recognition Letters*, vol. 33, no. 13, pp. 1725–1733, Oct. 2012.

[7] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 6, pp. 643–660 , Jun. 2001.

[8] R. Basri and D.W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.

[9] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 9, pp. 1063–1073. Sep. 2003.

[10] A. Shashua and T. Riklin-Raviv, "The quotient image: class-based re-rendering

and recognition with varying illuminations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 2, pp. 129–139, 2001.

[11] S. Shan, W. Gao, B. Cao, and D. Zhao, "Illumination normalization for robust face recognition against varying lighting conditions," in *Proc. IEEE Workshop on Analysis and Modelling of Faces and Gestures*, France, 2003.

[12] H.T. Wang, S.Z. Li, Y.S. Wang, and J.J. Zhang, "Self quotient image for face recognition," in *Proc. Internat. Conf. on Image Processing*, Singapore. 2004.

[13] X.G He, J. Tian, L.F. Wu, Y.Y. Zhang, and X. Yang, "Illumination normalization with morphological quotient image," *Journal of Software*, vol. 18, no. 9, pp. 2318–2325 , Sep. 2007

[14] T. Chen, W. Yin, X.S. Zhou, D. Comaniciu and T.S. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. Pattern Anal. Machine, Intell*, vol. 28, no. 9, pp. 1519–1524, Sep. 2006

[15] R.C. Gonzales and R.E. Woods, Digital Image Processing, 2nd Ed. Upper Saddle River, New Jersey: Prentice Hall, ISBN: 0-201-18075-8, 1992.

[16] W. Chen, M.J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Systems Man Cybernet. B Cybernet*, vol. 36, no. 2, pp. 458–466, 2006.

[17] V.P. Vishwakarma, S. Pandey and M.N. Gupta, "A novel approach for face recognition using DCT coefficients re-scaling for illumination normalization," in *Proc. 15th Internat. Conf. on Advanced Computing and Communications*, Guwahati, 2007.

[18] C.A. Perez and L.E. Castillo, "Genetic improvements in illumination compensation by the discrete cosine transform and local normalization for face recognition," in *Proc. SPIE –The International Society for Optical Engineering, International Symposium on Optomechatronic Technologies*, San Diego, 2008.

[19] M.V. Heydi, G.R. Edel, and C.M. Yadira, "A new combination of local appearance based methods for face recognition under varying lighting conditions," in *Proc. 13th Iberoamerican Congress on Pattern Recognition*, Havana, 2008.

[20] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu , H.H. Shih, Q.N. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, "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. 903–995, 1998.

[21] R. Bhagavatula and M. Savvides, "Analyzing facial images using empirical mode decomposition for illumination artifact removal and improved face recognition," *International Conference on Acoustics Speech and Signal Processing (ICASSP 2007)*, Honolulu, HI, Apr. 15-20, 2007, pp. 505–508.

[22] J.C. Nunes, Y. Bouaoune, E. Del´echelle, O. Niang, and Ph. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image and Vision Computing*, vol. 21, no. 12, pp. 1019–1026, 2003.

[23] M. Shao, Y. H. Wang, and X. Liang, "A BEMD based normalization method for face recognition under variable illuminations," *International Conference on Acoustics Speech and Signal Processing (ICASSP 2010)*, Dallas, TX, Mar. 14-19, 2010, pp. 1114－1117.

[24] D. Zhang, J. J. Pan, Y. Y. Tang, and C. Wang, "Illumination invariant face recognition based on the new phase features," *IEEE International Conference on Systems Man and Cybernetics (SMC 2010)*, Istanbul, Turkey, Oct. 10-13, 2010, pp. 3909－3914.

[25] M.C. Morrone, J.R. Ross, D.C.Burr, and R.A. Owens, "Mach bands are phase dependent," *Nature* 324, pp. 250–253, Nov. 1986.

[26] A. Linderhed, "2-D empirical mode decompositions in the spirit of image compression," in *Proc. SPIE Wavelet and Independent Component Analysis Applications IX*, vol. 4738, Orlando, USA, Apr. 2002, pp. 1–8

[27] C. Damerval, S. Meignen, and V. Perrier, "A fast algorithm for bidimensional EMD," *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 701–704, 2005.

[28] A. Linderhed, "Variable sampling of the empirical mode decomposition of two-dimensional signals," *Int. J. Wavelets Multresolution Inform. Process.*, vol. 3, no. 2005, pp. 435–452, Sep. 2005.

[29] Y. Xu, B. Liu, J. Liu, and S. Riemenschneider, "Two-dimensional empirical mode decomposition by finite elements," *Proceedings of the Royal Society A*,

vol. 462, no. 2074, pp. 3081–3096, 2006.

[30] S. M. A. Bhuiyan, R. R. Adhami, and J. F. Khan, "Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–18, 2008.

[31] Z. Yang, D. Qi, and L. Yang, "Signal period analysis based on Hilbert–Huang transform and its application to texture analysis," *Proc. 3rd Int. Conf. on Image and Graphics*, Hong Kong, China, Dec. 18–20, 2004, pp. 430–433.

[32] Z. Wu, N.E. Huang, and X. Chen, "The multi-dimensional ensemble empirical mode decomposition method," *Advances in Adaptive Data Analysis*, vol. 1, no. 3, pp. 339–372, 2009.

[33] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[34] S. M. A. Bhuiyan, N. O. Attoh-Okine, K. E. Barner, A. Y. Ayenu, and R. R. dhami, "Bidimensional empirical mode decomposition using various interpolation techniques," *Advances in Adaptive Data Analysis*, vol. 1, no. 2, pp. 309–338, 2009.

[35] G. Wang, X.Y. Chen, and F.L. Qiao, "On intrinsic mode function," *Advances in Adaptive Data Analysis*, vol. 2, no. 3, pp. 277–293, 2010.

[36] S. Beucher, "Geodesic reconstruction, saddle zones and hierarchical segmentation," *Image Analysis and Stereol*, vol. 20, no. 3, pp. 137–141, 2001.

[37] R. Bro, "Multiway calibration. Multilinear PLS," *Journal of Chemometrics*, vol. 10, no. 1, pp. 47–62, Dec. 1996.

[38] C. Nastar, "The image shape spectrum for image retrieval", *INRIA Technical Report RR-3206*, 1997.

[39] F. Dieterle. (2006, Aug. 14). *Theory – Fundamentals of the Multivariate Data Analysis* [online]. Available: http://www.frank-dieterle.de/phd/2_4_3.html