Gabor wavelet transform and its application

Wei-lun Chao R98942073

Abstract

This term project report introduces the well-know Gabor wavelet transform and its applications. Wavelet transform could extract both the time (spatial) and frequency information from a given signal, and the tunable kernel size allows it to perform multi-resolution analysis. Among kinds of wavelet transforms, the Gabor wavelet transform has some impressive mathematical and biological properties and has been used frequently on researches of image processing.

1. Introduction

From the last three lectures of the "time-frequency analysis and wavelet transform" course [3], we have learned that the wavelet transform could perform multi-resolution time-frequency analysis. The tunable kernel size results in different time-frequency resolution pair and the size is related to the analytical frequency. For example, smaller kernel size (in time domain) has higher resolution in time domain but lower resolution in frequency domain, and is used for higher frequency analysis; while bigger kernel size has higher resolution in frequency domain but lower resolution in time domain, and is used for lower frequency analysis. This great property makes wavelet transform suitable for applications such as image compression, edge detection, filter design, and some kinds of image object recognition, etc.

Among various wavelet bases, Gabor functions provide the optimal resolution in both the time (spatial) and frequency domains, and the Gabor wavelet transform seems to be the optimal basis to extract local features for several reasons [2]:

- *Biological motivation:* The simple cells of the visual cortex of mammalian brains are best modeled as a family of self-similar 2D Gabor wavelets.
- *Mathematical and empirical motivation:* Gabor wavelet transform has both the multi-resolution and multi-orientation properties and are optimal for measuring local spatial frequencies. Besides, it has been found to yield distortion tolerance space for pattern recognition tasks.

Based on these advantages of Gabor wavelet transform, it has been used in many image analysis applications, and this report focus it's applications on face recognition, texture classification, facial expression classification, and some other excellent

researches. The report is organized as follows: In section 2, the definition of 1D Gabor functions and 2D Gabor wavelets are introduced, and important mathematical properties and derivations are included. Applications on face recognition are presented in section 3, followed by texture features and classification in section 4. In section 5, we talk about other applications, and the conclusion of this term project report is given in section 6.

2. Fundamentals of Gabor wavelet transform

The Fourier transform has been the most commonly used tool for analyzing frequency properties of a given signal, while after transformation, the information about time is lost and it's hard to tell where a certain frequency occurs. To solve this problem, we can use kinds of time-frequency analysis techniques learned from the course [3] to represent a 1-D signal in time and frequency simultaneously. There is always uncertainty between the time and the frequency resolution of the window function used in this analysis since it is well know that when the time duration get larger, the bandwidth becomes smaller. Several ways have been proposed to find the uncertainty bound, and the most common one is the multiple of the standard deviations on time and frequency domain:

$$\sigma_t{}^2 = \frac{\int t^2|x(t)|^2dt}{\int |x(t)|^2dt}, \sigma_f{}^2 = \frac{\int f^2|X(f)|^2df}{\int |X(f)|^2df} \tag{1}$$

$$\sigma_t \times \sigma_f \geq \frac{1}{4\pi} \tag{2}$$

Among all kinds of window functions, the Gabor function is proved to achieve the lower bound and performs the best analytical resolution in the joint domain [4]. This function is a Gaussian modulated by a sinusoidal signal and shown below:

$$\varphi(t) = exp(-\alpha^2 t^2)exp(j2\pi f_0 t) \tag{3}$$

$$\Phi(f) = \sqrt{\frac{\pi}{\alpha^2}}exp(-\pi^2/a^2(f-f_o)^2) \tag{4}$$

where α determines the sharpness and f_0 the is modulated center frequency of $\varphi(t)$, and $\Phi(f)$ is its Fourier transform. Fig.1 shows the example of $\varphi(t)$ with three different f_0 but the same α and their time-frequency analysis by Gabor transform. These three distributions have the same area but don't meet the multi-resolution requirement: the window size should depend on the center frequency. To achieve this requirement, we substitute a with f_0/γ, where γ is an self-defined constant, and make the time duration of $\varphi(t)$ dependent on the central frequency f_0. The generalized $\varphi(t)$ with normalization of the maximum response in frequency

domain is now defined as:

$$\varphi(t) = \frac{|f_0|}{\gamma\sqrt{\pi}} \exp(-\frac{f_0^2}{\gamma^2} t^2) exp(j2\pi f_0 t) \quad (5)$$

Fig.2 shows the example of this new-defined $\varphi(t)$ with three different f_0 but the same α and their time-frequency analysis by Gabor transform.

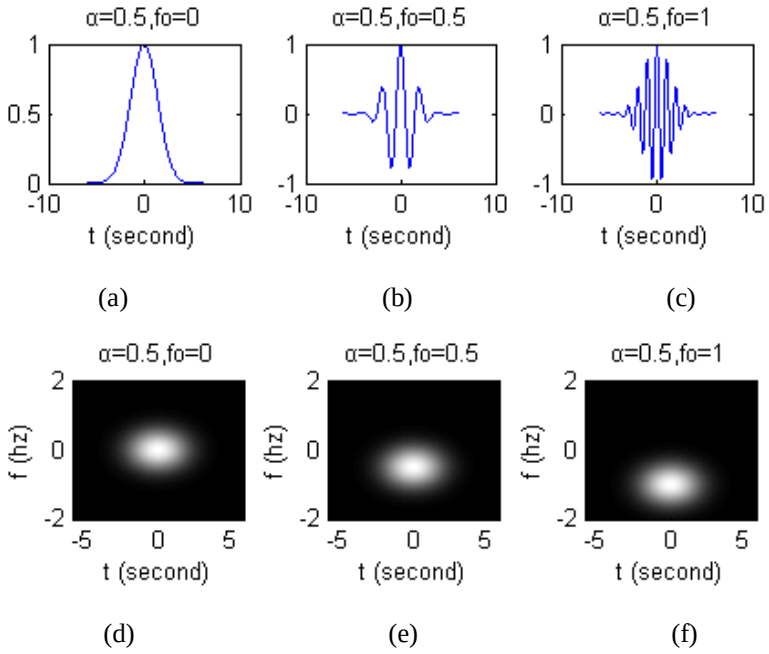


Figure 1: Example of $\varphi(t)$ with three different $f_0 = 0, 0.5, \text{and } 1$ but the same $\alpha = 0.5$ and their time-frequency analysis by Gabor transform, where (a)-(c) show the real part of $\varphi(t)$and (d)-(f) show the magnitude of the Gabor transform of $\varphi(t)$.

This 1-D Gabor function could be extended into 2-D form and also achieve the lower bound of uncertainty principle [5]. This 2-D Gabor function is defined as [2]:

$$\varphi(x,y) = \frac{f^2}{\pi\gamma\eta} exp(-(\frac{f^2}{\gamma^2} x_r^{\,2} + \frac{f^2}{\eta^2} y_r^{\,2})) exp(j2\pi f x_r) \quad (6)$$

$$x_r = x\cos\theta + y\sin\theta\,, y_r = -x\sin\theta + y\cos\theta$$

where f is the frequency of the modulating sinusoidal plane wave and θ is the orientation of the major axis of the elliptical Gaussian. The 2-D Fourier transform of $\varphi(x,y)$ is shown below:

$$\Phi(u,v) = exp(-\pi^2(\frac{\gamma^2}{f^2}(u_r - f)^2 + \frac{\eta^2}{f^2} v_r)) \quad (7)$$

$$u_r = u\cos\theta + v\sin\theta\,, v_r = -u\sin\theta + v\cos\theta$$

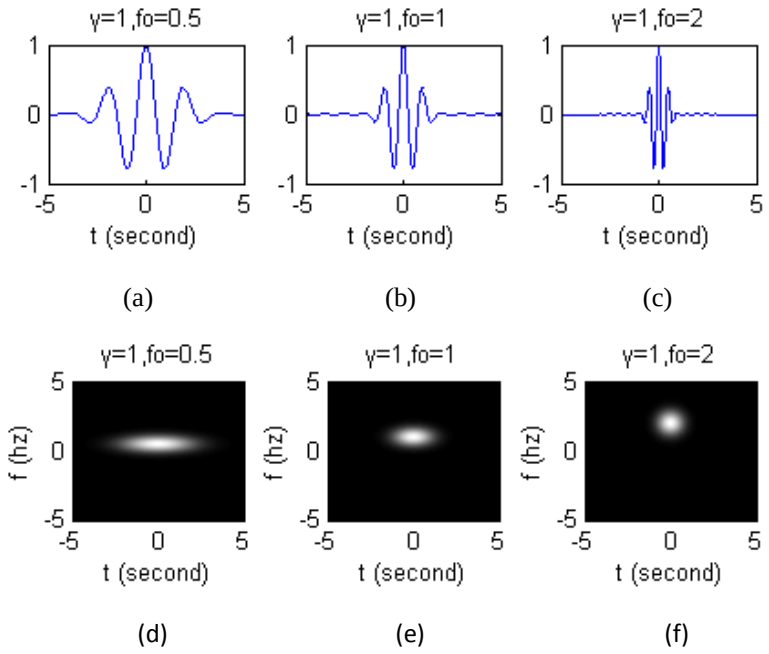


Figure 2: Example of the generalized $\varphi(t)$ with three different $f_0 = 0.5, 1, \text{and } 2$ but the same $\gamma = 1$ and their time-frequency analysis by Gabor transform, where (a)-(c) show the real part of $\varphi(t)$and (d)-(f) show the magnitude of the Gabor transform of $\varphi(t)$.

From the course [3], we know that a set of 1-D wavelets is defined as:

$$\psi(t, a, b) = \psi((t - a)/b) \tag{8}$$

where $\psi(t)$ is the mother wavelet and a and b determines the temporal shifting and scaling of this function. This definition could be further extended into 2-D wavelet transform as:

$$\psi_\theta(b_x, b_y, x, y, x_0, y_0) = \frac{1}{\sqrt{b_x b_y}} \psi_\theta\left(\frac{x - x_0}{b_x} + \frac{y - y_0}{b_y}\right) \tag{9}$$

where $\psi_\theta(x, y)$ is the 2-D mother wavelet, with b_x and b_y the scaling parameters, x_0 and y_0 the spatial shifting, and θ the orientation parameter. The 2-D Gabor function defined in Eq.6 meets this form and could be seen as a set of self-similar Gabor wavelets. The spatial shifting terms are missing in Eq.6 while these could be compensated by the convolution operation between this equation and the input image. To make $\varphi(x, y)$ as a set of continuous wavelets, we should make sure that $\varphi(x, y)$ obeys the five constraints [3]: compact support, real, even symmetric or odd symmetric, vanishing moments, and admissibility criterion. The former three constraints are achieved by setting a magnitude threshold to $\varphi(x, y)$ and separate it into real and imaginary part, while the later two need the DC-free modification to make this wavelet transform reversible and with at least one vanishing moment. This

modified Gabor wavelet is defined as:

$$\varphi(x,y) = \frac{f^2}{\pi\gamma\eta} exp\left(-\left(\frac{f^2}{\gamma^2}{x_r}^2 + \frac{f^2}{\eta^2}{y_r}^2\right)\right)(exp(j2\pi f x_r) - K) \quad (10)$$

where K is an offset parameter dependent on γ and η.

In practical cases, the Gabor wavelet is used as the discrete wavelet transform with either continuous or discrete input signal, while there is an intrinsic disadvantage of the Gabor wavelets which makes this discrete case beyond the discrete wavelet constraints: the 1-D and 2-D Gabor wavelets do not have orthonormal bases. If a set of wavelets has orthonormal bases, the inverse transform could be easily reconstructed by a linear superposition, and we say this wavelet transform provides a complete representation. The nonorthonormal wavelets could provide a complete representation only when they form a frame [1]. The concepts of the frame is beyond the scope of this report because it's too theoretical, while in most of the applications, we don't really care about these nonorthonormal properties if the Gabor wavelets are used for ***"feature extractions"***. When extracting features for pattern recognition, retrieval, or computer vision purpose, the transformed coefficients are used for distance measure or compressed representation but not for reconstruction, so the orthogonal constraint could be omitted.

Besides the form defined in Eq.6, there are some other equivalent definitions of the Gabor wavelets, for example, in [2], there is another form defined as:

$$\varphi(\vec{z}) = \frac{1}{2\pi}\frac{\|\vec{k}\|^2\|\vec{z}\|^2}{2\sigma^2} exp(j\vec{k}\cdot\vec{z}) \quad (11)$$

where $\vec{k} = 2\pi f exp(j\theta)$, and the scaling functions for the two elliptical axes are the same as σ. For most of the applications, a family of $U \times V$ Gabor wavelets is usually required to perform the multi-resolution and multi-orientation analysis, which is defined as below [2]:

$$\{\varphi_{discrete\ (f_u,\theta_v,\gamma,\eta)}(x,y)\} \quad (12)$$
$$f_u = \frac{f_{max}}{\sqrt{2}^u}, \theta_v = \frac{v}{V}\pi, u = 0, \dots\dots U-1, v = 0, \dots\dots V-1$$

where f_u and θ_v defines the orientation and scale of the Gabor wavelets, f_{max} is the maximum central frequency and $\sqrt{2}$ is the spacing factor between different central frequencies. Fig.3 shows a family of Gabor wavelets with 4 scales and 8 orientations, and Fig.4 shows the coverage of spatial frequency plane of these Gabor

wavelets.

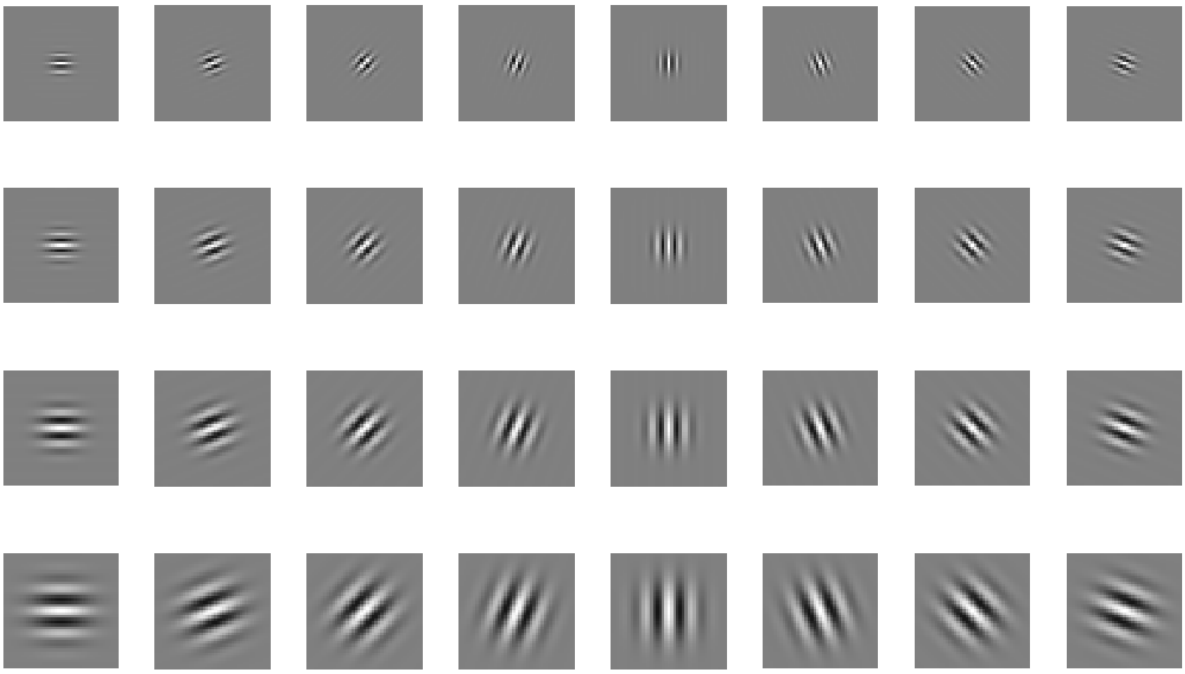


Figure 3: An example of the real part of Gabor wavelets with 4 scales and 8 orientations. The f_{max} is set as 0.35 with $\gamma = \eta = 1.2$.

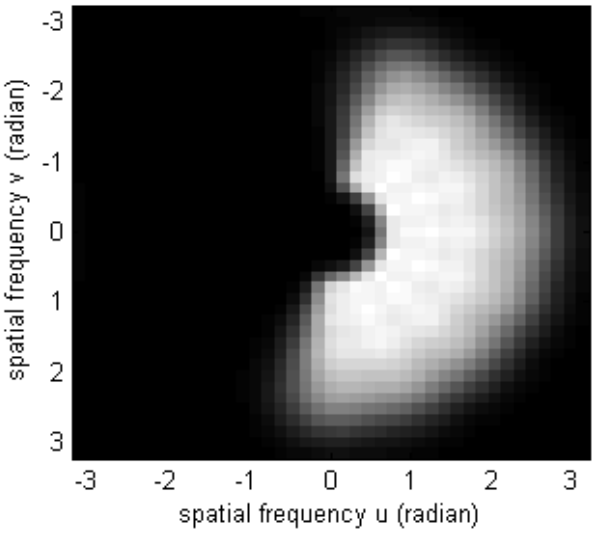


Figure 4: The coverage of the spatial frequency plane of the same Gabor wavelets family with 4 scales and 8 orientations. We can see that with well-chosen parameters. The whole frequency space could be covered.

3. Face recognition

Face recognition is one of the most important applications of Gabor wavelets. The face image is convolved with a set of Gabor wavelets and the resulting images are further processed for recognition purpose. The Gabor wavelets are usually called Gabor filters in the scope of applications. There have been a great amount of researches on face recognition recently, and various proposed approaches could be roughly classified into analytic and holistic approaches [2]:

- *Analytic approaches:* Some feature points are detected from the face, especially the important facial landmarks such as eyes, noses, and mouths. These detected points are called the fiducial points, and the local features extracted on these points, distance and angle between these points, and some quantitative measures from the face are used for face recognition. The main advantage of analytic approaches is to allow for a flexible deformation at the key feature points so that pose changes, different angles of view can be compensated for.
- *Holistic approaches:* In contrast to using information only from key feature points, holistic approaches extracts features from the whole face image. Normalization on face size and rotation is a really important pre-processing to make the recognition robust. The eigenface based on principal component analysis (PCA) [6] and the fisher face based on linear discriminant analysis (LDA) [7] are two of the most well-know holistic approaches.

In this section, I focus more on the analytic face recognition methods, especially based on the graph matching framework.

3.1 Elastic graph matching based methods

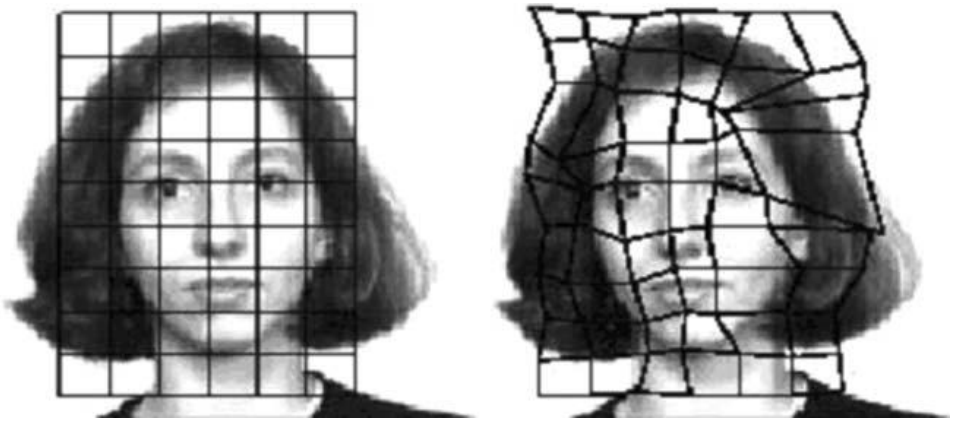
The application of Gabor wavelet for face recognition is pioneered by Lades et al.'s work [8]. In their work, the elastic graph matching framework is used for finding feature points, building the face model and performing distance measurement, while the Gabor wavelets are used to extract local features at these feature points, and a set of complex Gabor wavelet coefficients for each point is called a jet. Graph matching based methods normally requires two stages to build the graph g^I for a face image I and compute its similarity with a model graph g^M. During the first stage, g^M is shifted within the input image to find the optimal global offset of g^I while keeping its shape rigid. Then in the second stage, each vertex in g^I is shifted in a topological constraint to compensate the local distortions due to rotations in depth or expression variations. It is actually the deformation of the vertices that makes the graph matching procedure elastic. To achieve these two stages, a cost measure function $S(g^M, g^I)$ is

neccesarily to be defined and these two stages terminate when this function reaches the minimum value.

Lades et al.'s [8] used a simple rectangular graph to model faces in the database while each vertex is without the direct object meaning on faces. In the database building stage, the deformation process mentioned above is not included, and the rectangular graph is manually placed on each face and the features are extracted at individual vertices. When a new face *I* comes in, the distance between it and all the faces in the database are required to compute, which means if there are totally *N* face models in the database, we have to build *N* graphs for *I* based on each face model. This matching process is very computationally expensive especially for large database.Fig.5 shows an example of a model graph and a deformed graph based on it, and the cost function is defined as:

$$S(g^M, g^I) = \sum_n S_m({J_n}^I - {J_n}^M) - \lambda \sum_e (\Delta {\vec{x}_e}^I - \Delta {\vec{x}_e}^M)^2 \qquad (13)$$

where λ determines the relative importance of jet similarity and the topography term. $\Delta\vec{x}_e$ is the distance vector of the labeled edge *e* between two vertices, J_n is the set of jets at vertex *n*, and S_m is the distance measure function between two jets based on the magnitude of jets.



(a) (b)

Figure 5: The graphic models of face images. The model graph (a) is built to represent a face stored in the database, and features are directly extracted on vertices in the rectangular graph. When a new face comes in and we want to recognize this person, a deformed graph (b) is generated based on the two-stage process [8].

Wiskott et al. [9] proposed an improved elastic graph matching framework to deal with the computational-expensive problem above and enhance the performance. They employed object-adaptive graph to model faces in the database, which means the

vertices of a graph refer to special facial landmarks and enhance the distortion-tolerant ability (see Fig.6). The distance measure function here not only counts on the magnitude information, but also takes in the phase information from the feature jets. And the most important improvement is the used of face bunch graph (FBG), which is composed of several face models to cover a wide range of possible variations in the appearance of faces, such as differently shaped eyes, mouths, or noses, etc. A bunch is a set of jets taken from the same vertex (the same landmark) from different face models and Fig.7 shows the FBG structure. The cost function is redefined as:

$$S(B, g^I) = \frac{1}{N}\sum_n \max_m S_p(J_n{}^I - J_n{}^{B_m}) - \frac{\lambda}{E}\sum_e \frac{(\Delta\vec{x}_e{}^I - \Delta\vec{x}_e{}^B)^2}{(\Delta\vec{x}_e{}^B)^2} \qquad (14)$$

where B is the FBG representation, and N and E are the total amounts of vertices and edges in the FBG. B_m denotes the m^{th} model graph of B and S_p is the new-defined distance measure function which takes the phase of jets into account. To build the database, a FBG is first generated and models for individual faces are generated by the elastic graph matching procedure based on FBG. When a new face comes in, the same elastic graph matching procedure based on FBG is executed to generate a new face model, and this model could directly compare with the face models in the database without re-modeling. The FBG serves as the general representation of faces and reduce the computations for face modeling.

Besides these two symbolic examples using the elastic graph matching framework, a number of varied versions have been proposed in literature and readers could found a brief introduction in [2].

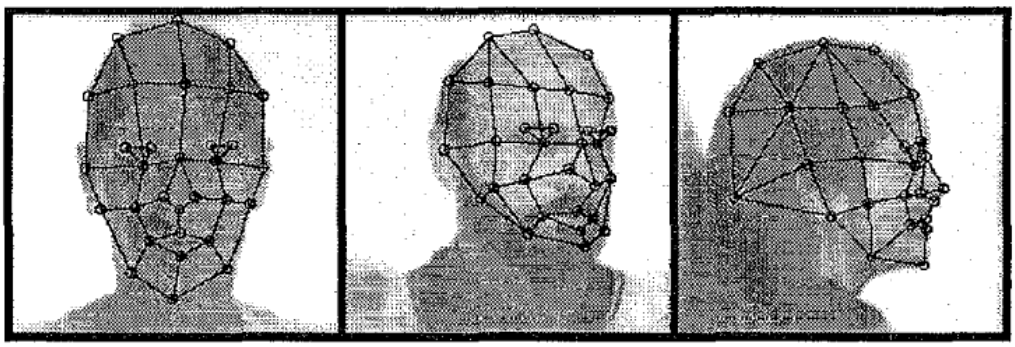


Figure 6: The object-adaptive grids for difference poses. Now the vertices are positioned automatically by elastic bunch graph matching and are located at special facial landscapes. One can see that, in general, the matching finds the fiducial points quite accurately, but still with some mispositioning [9].

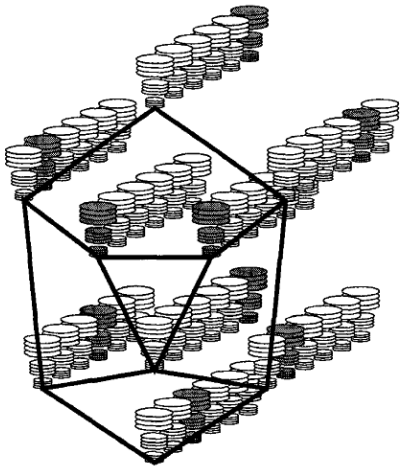


Figure 7: The face bunch graph serves as the general representation of faces. As shown in this figure, there are nine vertices in the graph and each of them contains a bunch of jets to cover the variations in the facial appearance. The edges are represented by the averaged distance vector calculated from all face models sued to build the FBG [9].

3.2Other face recognition methods based on Gabor wavelets

The elastic graph matching framework is one of the analytic face recognition approaches, and this framework is used mainly for feature points detection and face modeling. Other previous works used different feature point detection algorithm such as the method proposed by Manjunath et al. [10] (see Fig.8), or manually located these key points.

Figure 8: Feature locations marked for a pair of face images. The algorithm is based on a model of end-inhibition property, and makes use of local scale intersections between simple oriented features [10].

Besides the analytic approaches, Gabor wavelets can be used for holistic approaches, too. A face image is first convolved with a set of Gabor wavelets and the resulting images can be further combined with PCA or LDA to reduce the feature dimension and generate the salient representation. In Fig.9, I show an example of face images undergone convolution with the same 4×8 Gabor wavelet sets shown in Fig.3, and the original face image is shown in Fig.10.

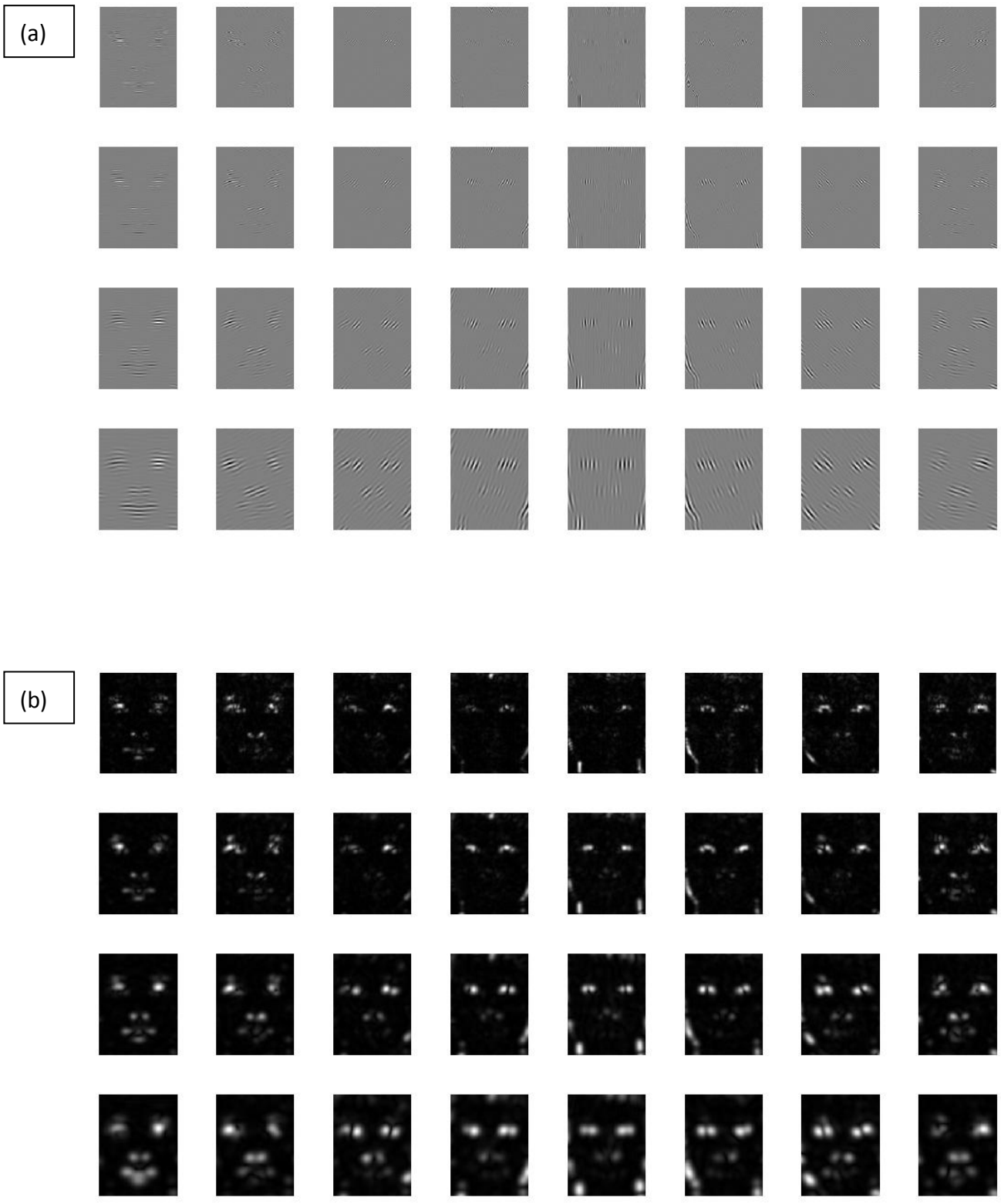


Figure 9: Gabor wavelet representation of a face image shown in Fig.10. (a) The real part of the resulting images, and (b) the magnitude.

Figure 10: The original face image used in Fig.9.

4. Texture features and classification

The Gabor wavelets could not only be used for feature extraction for face images, but also for other images such as textures, remote sensing images, and landscape images, etc. Manjunath and Ma [11] used Gabor wavelets for browsing and retrieval of image data, and the experimental result showed that Gabor wavelets achieve the best performance against other feature extraction methods such as conventional pyramid-structured wavelet transform (PWT) features, three-structured wavelet transform (TWT) features, and the multi-resolution simultaneous autoregressive model (MR-SAR) features. They also proposed an algorithm for adaptive filter selection which could reduce the feature dimension without seriously degrading the performance. Arvazhagan et al. [12] proposed a new approach for rotation invariant texture classification using Gabor wavelets where the features are found by calculating the mean and variance of the Gabor filtered images. An feature vector is generated for a image feature, for example, if a 4×9 Gabor wavelet set is used, then there will be 72 elements in this feature vector. The element order in their method is based on the dominant direction, where the first 9 elements are ordered as $20^o, 40^o, \ldots\ldots 160^o, 0^o$ of the same scale with 20^o detected as the dominant direction. Fig.11 shows an example for determining the dominant direction.

Here I'd like to show an example of using Gabor wavelet features for image retrieval. Fig.12 shows three images and there corresponding Gabor wavelet features (the mean magnitude is used as the feature), where two of them are pictures of Taipei 101 and the other one is a sky picture. Set the Taipei 101 image with blue sky background as the query image and the other two as images in the database, if we use color information for computing image similarity, the top retrieved image will be the sky not the Taipei 101 image with sunset glow. Using the Gabor wavelets, we can see

from Fig.12 that the features between the two images are more similar and the expected retrieval result could be achieved.

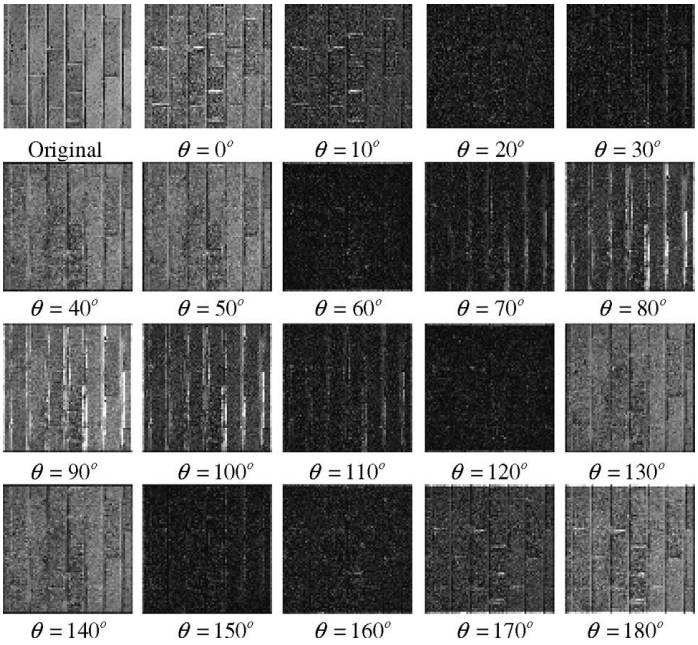


Figure 11: Gabor transform result obtained for 19 orientations (steps: 10^o) at a certain scale. The image with $\theta = 0^o$ is defined as the dominant orientation because it has the largest magnitude, then the feature order is set as $0^o, 10^o, \ldots\ldots 170^o, 180^o$ [12].

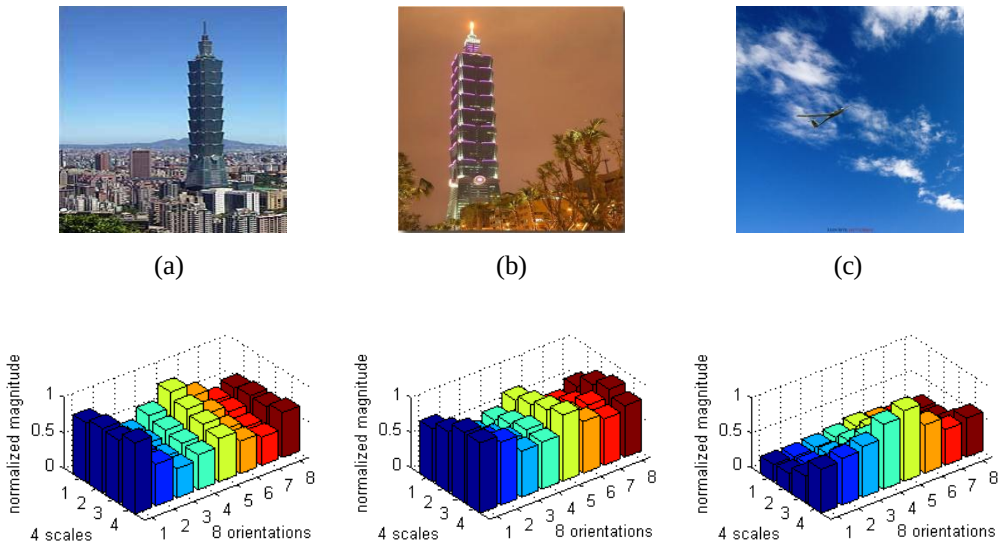


Figure 12: The features extracted from three images based on a 4×8 Gabor wavelet set. The 32 mean magnitude features are normalized by the maximum magnitude in each set of feature vector. Image (a) is used as the query and (b), (c) as database images, and it is shown that the query feature of (d) is more similar to (e) than to (f), which gives the correct retrieval result.

5. Other applications

There are many other application of Gabor wavelets, such as the facial expression classification [13,14], Gabor networks for face reconstruction [15], fingerprint recognition [16], facial landmark location [17], and iris recognition [18], etc. In this section, I'll briefly introduce the former two applications.

5.1 Facial expression classification

This technique is proposed by Lyons et al. [13,14],where the filtered image is used for feature generation and a two-layer perceptron (neural network framework) is trained for classification purpose. It is shown that the Gabor-wavelet-based method outperforms the geometry-based method and they can be combined to achieve the better performance. Fig.13 gives an example of their work.

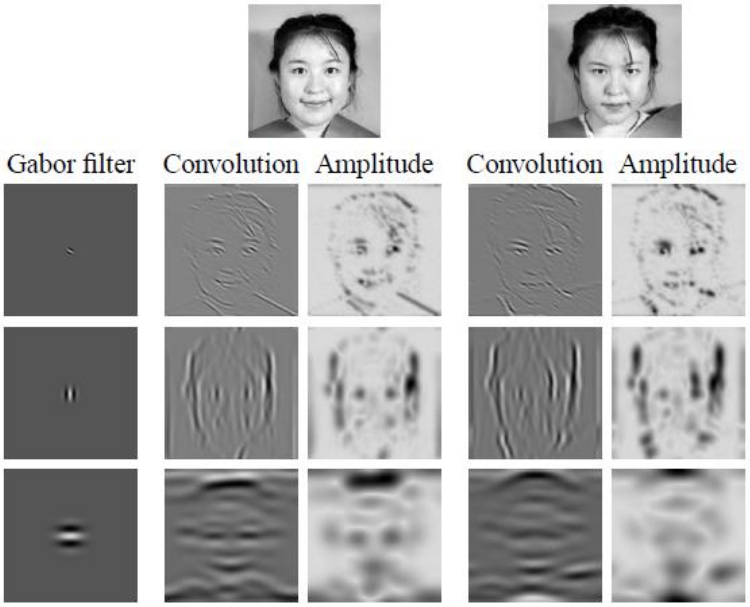


Figure 13: Example of three Gabor filter responses to two facial expression images [14].

5.2 Gabor networks for face reconstruction

Among most of the works based on Gabor wavelets for feature extraction, the characteristics and compression ability of wavelets have not been fully explored due to the nonorthogonality. Krueger [15] proposed the use of Gabor wavelet networks (GWN) for object representation and face processing, and the face image could be reconstructed by a set of Gabor wavelet coefficients with a certain quality. The quality of the reconstruction undoubtedly depends on the number of wavelets used. Fig. 14 shows the image reconstruction result based on different number of coefficients.

Figure 14: The original image and the reconstructed images with 16, 53, 116, and 216 Gabor wavelets from left to right [15].

Lee [1] has done a great work on the mathematical derivations and proofs of 2-D Gabor wavelets. In his study, the image could be perfectly reconstructed in a numerical stable way if the set of Gabor wavelets forms a frame.

6. Conclusion

This term project report introduces the well-known Gabor wavelet transform and its application. The multi-resolution and multi-orientation properties of the Gabor wavelet transform makes it a popular method for feature extraction even if the intrinsic nonorthogonality exists. Among all the works based on Gabor wavelet, face recognition and texture representation are the most noticeable applications, and other research used the Gabor wavelets mainly for feature extraction. Several Matlab implementations are presented in this report and show both the theoretical and application aspects of Gabor wavelets. There seems no further necessity to modify the formula of Gabor wavelets while the feature representation and more possible applications remain spaces for future works.

7. Reference

[1] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(10), 1996

[2] L. Shen and L. Bai, "A review of Gabor wavelets for face recognition," *Patt. Anal. Appl.* 9: 273-292, 2006

[3] J. J. Ding, "Time-Frequency Analysis and Wavelet Transform-Course website," 2009, [online], Available: http://djj.ee.ntu.edu.tw/TFW.htm. [Accessed Jan. 20, 2010].

[4] D. Gabor, "Theory of communications," in *J. IEE*, 93: 429-457

[5] J. G. Daugman, "Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Optical Soc. Amer.*, 2(7): 1160-1169, 1985

[6] M. Turk and A. Pentland, "Eigenfaces for recognition," *J.Cognitive. Neuroscience*, 3(1): 71-86, 1991

[7] P, Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection, " *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7): 711-720, 1997

[8] M. Lades, J. C. Vorbriiggen, J. Buhmann, J. Lange,C. von der Malsburg, R. P. Wiirtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. on Computers,* 42(3): 300-311, 1993

[9] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Analxsis and Machine Intelligence,* 19(7): 775-779, 1997

[10] B. S. Manjunath, R. Chellappa, and C. von der Malsburg, "A feature based approach to face ecognition," *Proc. IEEE Conf. CVPR'92*: 373-378, 1992

[11] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Analysis and Machine Intelligence,* 18(8): 837-842, 1996

[12] S. Arivazhagan, L. Ganesan, and S. P. Priyal, "Texture classification using Gabor wavelets based rotation invariant features," *Pattern recognition letters,* 27(16): 1976-1982, 2006

[13] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," *Proc. Int'l Conf. Automatic Face and Gesture Recognition,* 200-205, 1998

[14] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron," *Proc. Int'l Conf. Automatic Face and Gesture Recognition,* 454-459, 1998

[15] V. Kruger and G. Sommer, "Gabor wavelet networks for efficient head pose estimation," *Image and Vision Computing,* 20(9-10), 665-672, 2002.

[16] C. J. Lee and S. D. Wang, "Fingerprint feature extraction using Gabor filters," *Electronic letters,* 35(4), 288-290, 1999

[17] F. Smeraldi and J. Bigun, "Facial feature detection by saccadic exploration of the Gabor decomposition", *Proc. Int'l Conf. Image Processing,* 163-167

[18] C. Sanchez-Avila and R. Sanchez-Reillo, "Two different approaches for iris recognition using Gabor filters and multiscale zero-crossing representation," *Pattern recognition letters,* 38(2): 231-2440, 2005