Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 381–388
© 2006 Birkhäuser Verlag Basel/Switzerland

Transfer Colors from CVHD to MRI Based on Wavelets Transform

Xiaolin Tian, Xueke Li, Yankui Sun and Zesheng Tang

Abstract. A new algorithm based on wavelet transform to transfer colors from image of Chinese Virtual Human Data (CVHD) to Magnetic Resonance Images (MRI) has been proposed and implemented. The algorithm firstly extracts the primary components from both CVHD and MR images by wavelet-based multi-resolution analysis, then mapping colors from CVHD to MRI between these primary components with the similar characters. Finally these colors in MRI will be transferred to all pixels of the MRI according to their characters. Several experiment results have been reported, which have confirmed the effectiveness of this new color-transferring scheme.

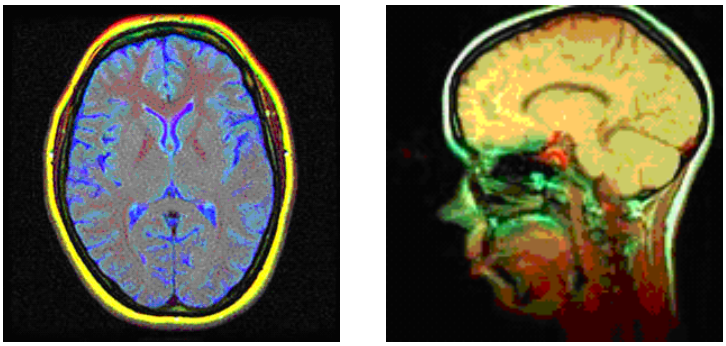
Keywords. Medical Image Processing, CVHD, Color MRI, Wavelet Transform, Multi-resolution Analysis.

1. Introduction

Magnetic resonance imaging (MRI) is a imaging method primarily used to demonstrate alterations of tissues in living organisms. Current MRI technology displays images as gray tone images. Because human eyes are more sensitive to different colors than to different gray tones, many scientists are working hard to have color MRI. A color MRI would significantly improve both the visualization and the parcellation of structures visualized by MRI, which enable more effective diagnosis.

The most popular and simplest way to have a color MRI is to assign individual red(R), green(G), and blue(B) channels to each pixel to separately encode three co-registered images with different contrasts into a single composite pseudo-color image[2][3][4]. This basic idea has several varieties but all varieties belongs to either a pseudo-color methodology or a semi-natural color composite MRIs[5][6][7], shown below in figure 1.

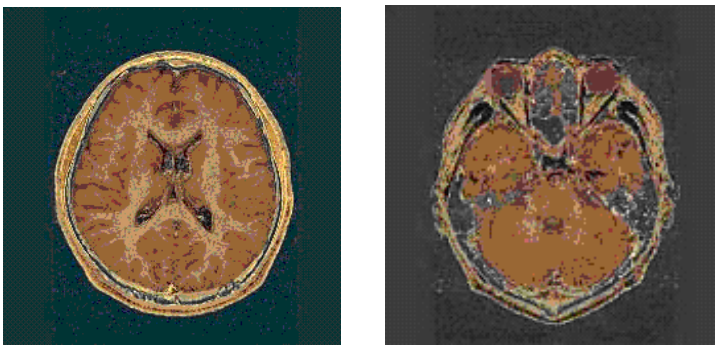
The work was supported by both Science and Technology Development Fund of Macao SAR (No. 001/2005/A) and the National Natural Science Foundation of China (No. 30470487).



(a) (b)

FIGURE 1. Pseudo color images from [2](left) and [4](right).

Another way of creating a colored MRI is to calculate the natural colors using other digital images of the same object. An algorithm has been proposed and implemented to transfer colors from the full-color cross-sections of the visible human datasets to the MRIs[8][9].In this approach, the MRI are transformed to components by independent component analysis(ICA)[10], which enhances physical characteristics of the tissue. The mapping functions for generating color values from independent components are obtained using the radial basis function(RBF)[11] network by training the network with sample data chosen from a visible female data set. Two transferred color MR images from [8] have been showed in figure 2.



(a) (b)

FIGURE 2. Natural color MR images from [8].

In this paper, a new algorithm based on wavelet transform has been proposed to transfer natural colors from images from the Chinese Virtual Human Data(CVHD)[12] to MRI. Instead of using complex artificial neural networks(RBF),

this new approach uses wavelet transforms to obtain multi-resolution images for both CVHD and MR images. The primary components of both the CVHD and the MR images will be extracted by the clustering of the lower resolution images, and then the transferring of colors from an image of CVHD to several MR images whose primary components have similar characteristics. Finally these colors in the primary components of the MR images will be transferred to all pixels in the MR images based on their characters. Several experiment results have been reported in the paper, which have confirmed the effectiveness of this new color-transferring scheme.

2. Major Steps of the New Algorithm

There are three major parts of the new algorithm: pre-processing, kernel processing, and post-processing, as described step by step below:

I. Preprocessing:

1. Select two corresponding images of the same human body part from the CVHD and the MR image database. Remove the background color[1], and then convert the RGB image from CVHD to a HSV image.
2. Normalize both the CVHD and MR images from step 1 to the same size and similar gray level distribution using gray level histogram transformation.
3. Register the two images of CVHD and MR so that more accurate corresponding spatial locations can be achieved.

II. Kernel Processing:

4. Calculate a series of different resolution images for the results of step 3 through a multilevel 2-D wavelet decomposition and reconstruction.
5. Calculate the primary components from the lower resolution images by clustering.
6. Calculate the statistical features for each primary component obtained from step 5. Setup the corresponding pairs of the primary components between the lower resolution images of CVHD and MR according to their spatial locations.
7. Calculate the same statistical features for both the original(highest) resolution images of CVHD and MR. Copy colors from the original(highest) resolution image of CVHD to the original(highest) resolution MR image pixel by pixel if their features belong to the pair of corresponding primary components.
8. Assign a color to every pixel in the MR image by recursively extending a color to the closest neighbor distance-wise in the feature space.

III. Post-processing:

9. Increase the color contrast of the colored MR image from step 8 through histogram equalization.
10. Smooth/denoise the resulting colored MR image of step 9.

3. Algorithm Implemented

In our research, the algorithm was implemented using Matlab[13]. Symlets Wavelets, which are improved Daubechies Wavelets, have been used to construct multiresolution images. For the colored images of CVHD, wavelet transform was applied to three images–the H, S, and V components of the CVHD images. A total of six levels of different resolution images for each original image have been reconstructed from wavelet decomposition coefficients. Four different color groups have been decided for the representation of the eyes, skin, nervous gray, and nervous white, which means that at least four primary components are needed in both images. The real color of each group and its varieties are calculated from these four primary components. Primary components are obtained from, in our case, the lowest resolution image of the V component of the HSV CVHD image in two steps: first, apply the nearest neighbor clustering, and then pass the component attributes to the other components of the CVHD image. The four primary components of the MR image are also calculated from the lower resolution image, which will significantly reduce the clustering and feature calculation time. The features used in this algorithm are the color values and gray level values combined with the locational and spatial info. The relationships of the corresponding primary components between the CVHD and the MR images, which are color-transferring pairs, are set up according to the similarities of their spatial locations only. The color transferring work is done in two steps: First, the color is transferred from the original CVHD image to the original MR image if both pixels belong to the color transferring component pair; Secondly, colors in the MR image are transferred according to the neighboring color that has the most similar features as the one to be colored. The second step is implemented recursively until all pixels of the MR image have been assigned a color.

4. Experiment Results

The images of the head part of a body have been chosen since we only have corresponding head parts of both the CVHD and the MR images in our testing database. Figure 3 shows the original image of CVHD used in the test with their lower resolution images through the wavelet transform. The original CVHD image is the coloring source image, and the four primary components are obtained from the lowest resolution image. Figure 4 shows the original MR image used in our test with their lower resolution images by the same wavelet transform. The original MR image is the coloring target image and the four primary components are also obtained from its lowest resolution image.

Figure 5 shows several resulting colored MR images calculated using the new algorithm. The first row of figure 5 shows the original MRIs. The second row of figure 5 shows the color MRIs after the first coloring step of the new algorithm. The third row of figure 5 shows the color MRIs after the second coloring step. The fourth row of figure 5 shows the color MRIs after post-processing.

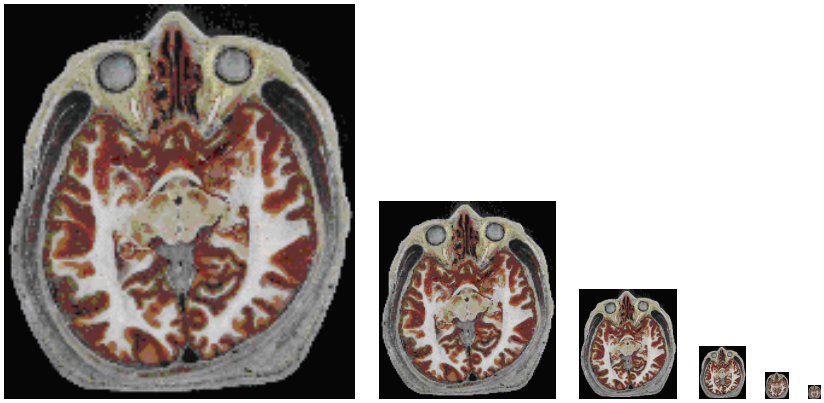


FIGURE 3. The original image of CVHD and its lower resolution images through the wavelet transform.

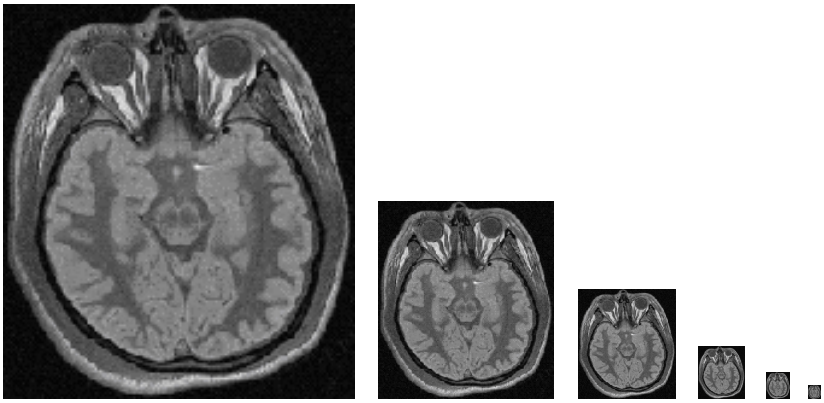


FIGURE 4. The original MR image and its lower resolution images through the wavelet transform

5. Discussions and Conclusions

In this paper, a new algorithm to obtain color MR Images has been proposed. The algorithm does not need the calculations of the independent components of the images and does not need the training mapping function by artificial neural network as proposed in [8]. This new algorithm simply calculates the primary components from the lower resolution images obtained through the wavelet transform, and the set up color mapping functions between corresponding primary components of the lower resolution images. The testing results have shown that color MR images can be obtained in this manner with more nature colors compared with the results

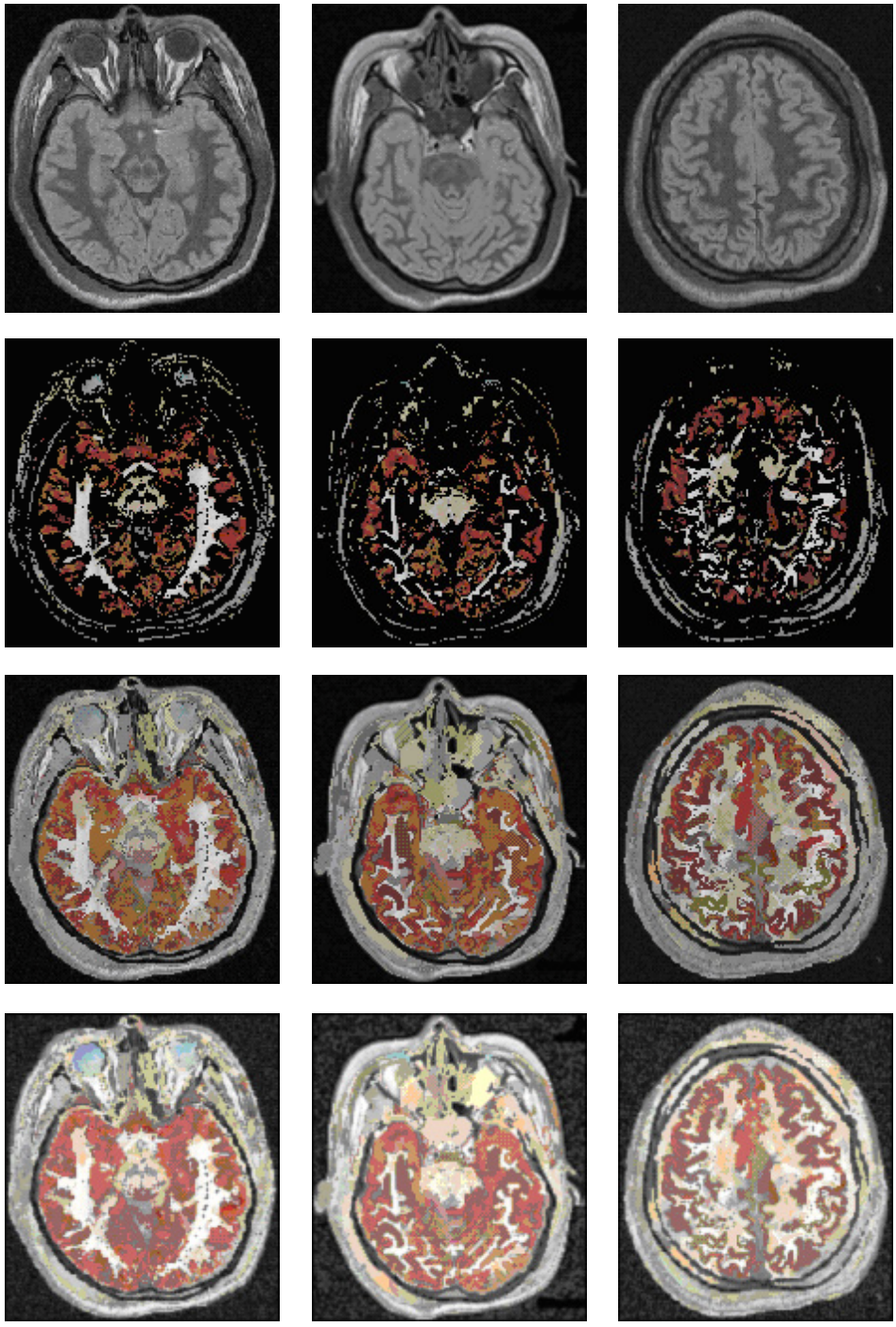


FIGURE 5. The original MR images with their coloring results at various steps

from [8], showed in the figure 2. In other words, this new algorithm can create color MRI more accurate.

All works reported in this paper are just beginning results. There are many issues that can be improved through further research to achieve better coloring results. For instance, instead of using one color source image, using different corresponding color CVHD images for a single MR images will achieve better results, but it also needs more time to calculate the different primary components and such. There is still a long way to go to have a real MRI coloring systems.

Acknowledgement

The MR images and the full color cross-section images of the Chinese Visible Human Project used in our research and shown in this paper are provided by Prof. Zhang Shaoxiang, who is the leader of the *Chinese Virtual Human* research team in the Third Military Medical University in Chongqing, China[12].

We would like to express our sincere appreciation for their help and support.

References

[1] Xiaolin Tian, Yan Zhao, Chengjun Tao, Zesheng Tang, *Background Removal for Color Images Based on Color Components Differences*, Proceedings of the Sixth IASTED International Conference on Signal and Image Processing, August 23-25, Hawaii, USA, 2004.

[2] Kenneth L. Weiss, Qian Dong, William J. Weadock, Robert C. Welsh, and Gaurang V. Shah, *Multiparametric Color-encoded Brain MR Imaging in Talairach Space*, Radiographics. 2002;22:e3-e3.

[3] K.L. Weiss, H. Pan, J. Storrs, W. Strub, J.L. Weiss, L. Jia, and O.P. Eldevik, *Clinical Brain MR Imaging Prescriptions in Talairach Space: Technologist- and Computer-Driven Methods*, AJNR Am. J. Neuroradiol., May 1, 2003; 24(5): 922 - 929.

[4] W.E. Phillips, H.K. Brown, J. Bouza, R.E. Figueroa, *Neuroradiologic MR applications with multiparametric color composite display*, Magn Reson Imaging. 1996;14(1):59-72.

[5] Michael D. Abràmoff, Ad P. G. Van Gils, Gerard H. Jansen and Maarten P. Mourits, *MRI Dynamic Color Mapping: A New Quantitative Technique for Imaging Soft Tissue Motion in the Orbit. Investigative Ophthalmology and Visual Science*, 2000; 41:3256-3260

[6] Bradley M. Hemminger, *Isoluminance: a color technique for visualizing multivariable medical image data*, Proceedings SPIE Vol 1896, pp.325-335.

[7] J. Ward, V. Magnotta, N.C. Andreasen, W. Ooteman, P. Nopoulos, and R. Pierson, *Color Enhancement of Multispectral MR Images: Improving the Visualization of Subcortical Structures*, Journal of Computer Assisted Tomography. 25(6):942-949, November/December 2001.

[8] Shigeru Muraki, Toshiharu Nakai and Yasuyo Kita, *Basic Research for Coloring Multichannel MRI Data*, Proceedings of the 11th IEEE Visualization 2000 Conference, p.33.

[9] Shigeru Muraki, Toshiharu Nakai and Yasuyo Kita, *An Attempt to Develop Color Magnetic Resonance Imaging (MRI) by using Visible Human Data*, The Third Visible Human Project Conference Proceedings 2000.

[10] S. Muraki and T. Nakai, *Independent Component Analysis of multichannel MRI Data*, Proc. ISMRM 2000, 582.

[11] C.M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, 1995.

[12] S.X. Zhang, P. A. Heng et. al., *Creation of the Chinese visible human data set*, The Anatomical Record (Part B: New Anatomist), Vol.275B, 2003, pp.190-195.

[13] Hu Chang Hua et. al., *System Analysis and design based on MATLAB—Wavelet Analysis*, Xidian University Press.

Xiaolin Tian
Faculty of Information Technology
Macau University of Science and Technology,
Taipa, Macao
e-mail: xltian@must.edu.mo

Xueke Li
Faculty of Information Technology
Macau University of Science and Technology,
Taipa, Macao
e-mail: kenry_lee@yahoo.com.cn

Yankui Sun
Department of Computer Science and Technology
Tsinghua University,
Beijing, China
e-mail: syk@tsinghua.edu.cn

Zesheng Tang
Faculty of Information Technology
Macau University of Science and Technology,
Taipa, Macao
e-mail: ztang@must.edu.mo

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 389–396
© 2006 Birkhäuser Verlag Basel/Switzerland

Medical Image Fusion by Multi-resolution Analysis of Wavelets Transform

Xueke Li, Xiaolin Tian, Yankui Sun and Zesheng Tang

Abstract. A novel algorithm for the multimodalities medical images fusion based on wavelet transform has been proposed and implemented. The auto-adaptive weighted coefficients have been calculated recursively to maximize the mutual information between the source image and the result image. Adopting multi-resolution analysis of wavelet transform, we achieved the MRI and CT image fusion. In addition, the new algorithm has been extended to MRI and color image fusion. The experiment results demonstrate that the new algorithm with wavelet transform have better fusion results compared with other mutual information fusion schemes without wavelet transform.

Keywords. Medical image fusion, Mutual information, Multi-resolution analysis, Wavelet transform.

1. Introduction

Medical image fusion is a process of combining the useful features from each of two or more different modalities medical images into a single medical image. In recent years, with the increase of clinic application demands, the research of multi-modality medical image fusion has attracted more and more attention. Just as the radiotherapy plan often benefits from the complementary information in images of different modalities, dose calculation is based on the computed tomography (CT) data, while tumor outlining is often better performed in the corresponding magnetic resonance (MR) scan. For medical diagnosis, CT provides the best information on denser tissue with less distortion, MRI provides better information on soft tissue with more distortion, PET provides better information on blood flow and flood activity with low space resolution, and color image data provides the best integrated information on natural body structure in general. With more available multimodality medical images in clinical applications, the idea of combining

The work was supported by both the Macao Science and Technology Development Fund (No. 001/2005/A) and the National Natural Science Foundation of China (No. 30470487).

images from different modalities has become highly important and medical image fusion has been marked as a new and promising research field.

Several image fusion approaches based on wavelet transforms have already been proposed [1][2][3][4]. Image fusion algorithms based on multi-scale wavelet transforms are also very popular in recent researches [5][6][7]. Existing algorithms for image data fusion are not quite satisfactory for object detection. In order to improve the resolution of the target image and to suppress the detection noise of each sensor, the images are decomposed with wavelet transform, and thus the wavelet coefficients and the approximation coefficients at different scales are obtained. We took the coefficients with large absolute values between the multi-resolution images as the important wavelet coefficients and computed the weighted mean value of the approximation coefficientsThe fusion image can then be obtained using the inverse wavelet transform for the wavelet coefficients and the weighted approximation coefficients.

However, the majority of fusion algorithms employing the wavelet transform have not attempted to acclimatize themselves to the wavelet coefficients based on the mutual information between the multimodality medical images. Specifically, no reports about the medical images fusion between MRI and color images have been found. In our research, the multi-resolution wavelet analysis were implemented for the multimodality medical images, and then in the wavelet transform region, the mutual information for the different weighted coefficient are compared. When the mutual information is maximal, we achieve the best fusion result by taking the corresponding wavelet coefficient through the inverse wavelet transform.

2. New Image Fusion Algorithm

2.1. The Scheme of Fusion

In this research, the image fusion scheme involves six steps recursively: Decompose the input images, which already have their backgrounds removed[8]; Select a fusion strategy; Fuse image F; Calculate the mutual information; Find the maximum of the mutual information; And output the fused image as shown in figure 2.1.

Step 1: Decompose the image A and the image B by forward wavelet transformation. In this case, each image is decomposed into the same level using a periodic discrete wavelet transform. The wavelet transform decomposes each image into low and high-frequency sub-layer images. At a given level, a series of parameter sets of the forward wavelet transformation can be obtained. When the input images are the CT and MRI, the MRI is described as image A and the CT image is described as image B. When the input images are the MRI and color image, the MRI is still described as image A and the color image is described as image B.

Step 2: Select both of the wavelet weighted approximation coefficient a and b. a and b are initialized to 0.5. In addition, a and b are kept so that $a+b=1, 0<a, b<1$.

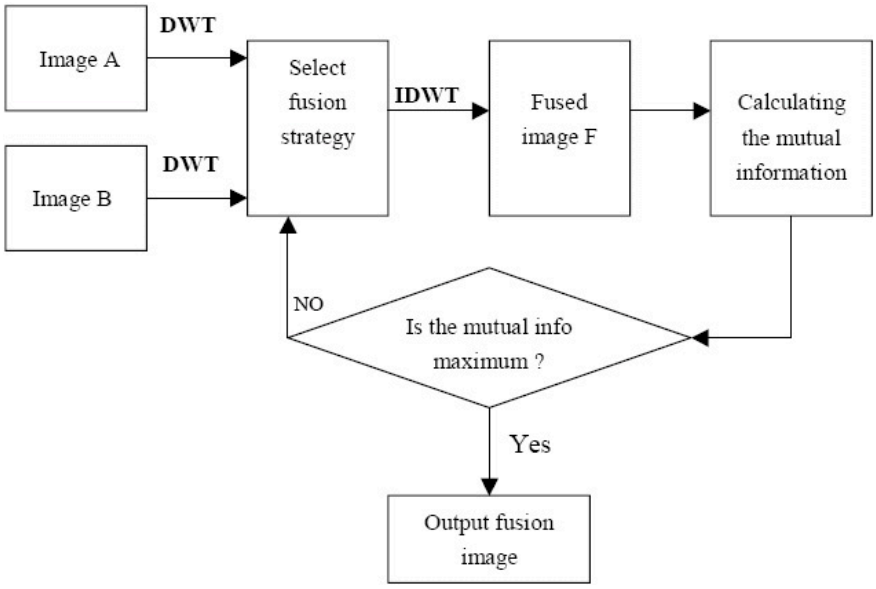


FIGURE 1. Image fusion scheme in this research

Step 3: Obtain the fusion image F by using the inverse wavelet transform incorporating with the selected coefficients as shown in expressions 2.1.

$$C_F(u,v) = aC_A(u,v) + bC_B(u,v) \tag{2.1}$$

In the above expressions, the parameter u is the decomposition vector and the parameter v is the corresponding bookkeeping matrix. $C_A(u,v), C_B(u,v)$ and $C_F(u,v)$ express respectively the wavelet transform coefficient of image A, image B and the fusion image F.

Step 4: Calculate the mutual information between the MRI image and the fusion image F.

Step 5: Compare the value of the latest mutual information with the value of the previous mutual information. Update the values of a and b for the new fused image F. Repeat steps 3 through 5 until a maximum for the mutual information is achieved.

Step 6: Output the fusion image F at the maximal mutual information as the experiment result.

2.2. Evaluation of the Fusion Result

In this research, we propose to use the mutual information(MI) to describe the fusion results. It is known that MI is a basic concept from information theory, measuring the statistical dependence between two random variables or the amount

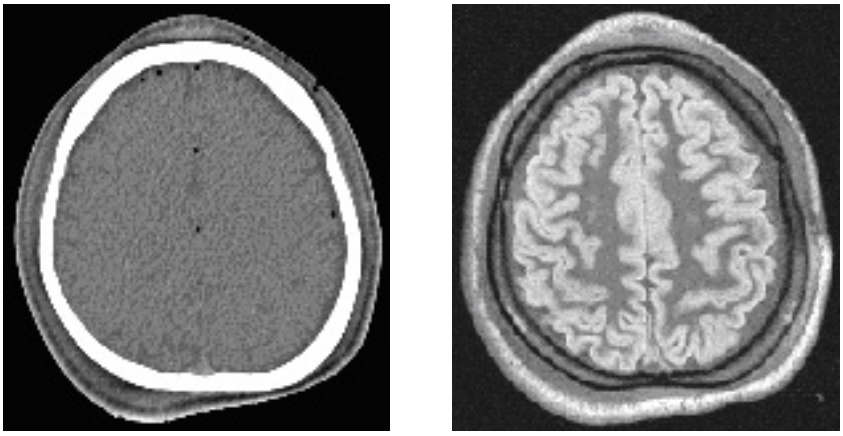
of information that one variable contains about the other. The mutual information or entropy is widely used in image compression and a large MI of the two medical image pixel pairs shows that the images have been well fused[12][13].

We can calculate the MI according to the following equation[12]:

$$MI_{AB}(x,y) = \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} P_{AB}(x,y)\log(P_{AB}(x,y)/P_A(x)P_B(y))$$

3. Experiment Results

In this section, we report the results of the experiments on fusing MRI and CT and fusing MRI and color images through the proposed image fusion scheme. Figure 2 shows the original CT and MRI images to be fused. Each image is resized to 146 pixels by 164 pixels with 256 gray levels.



(a) CT image (b) MR image

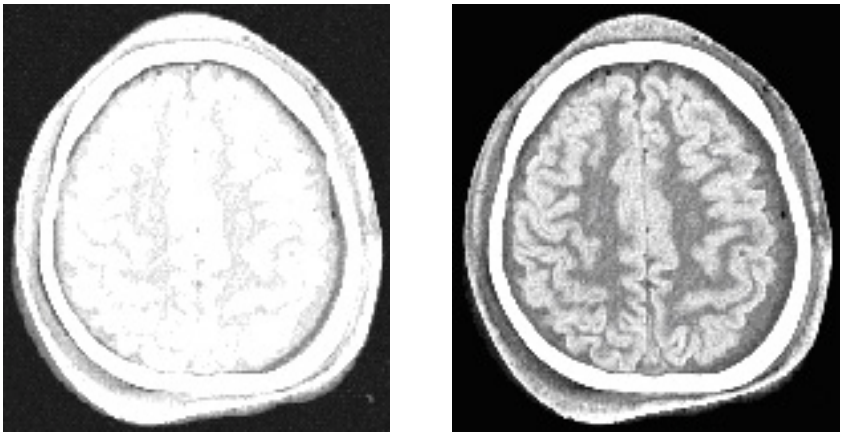
FIGURE 2. The original CT and MR images

Figure 3(a) shows the fusion result by the mutual information maximum without wavelet transform [13], Figure 3(b) shows the fusion result by the proposed algorithm with the wavelet transform.

Figure 4 shows the original MRI and Color images to be fused. Each image is resized to 256 pixels by 256 pixels with 256 gray levels.

Figure 5(a) shows the fusion result by a MI algorithm without wavelet transform [13]. Figure 5(b) shows the fusion result by our method.

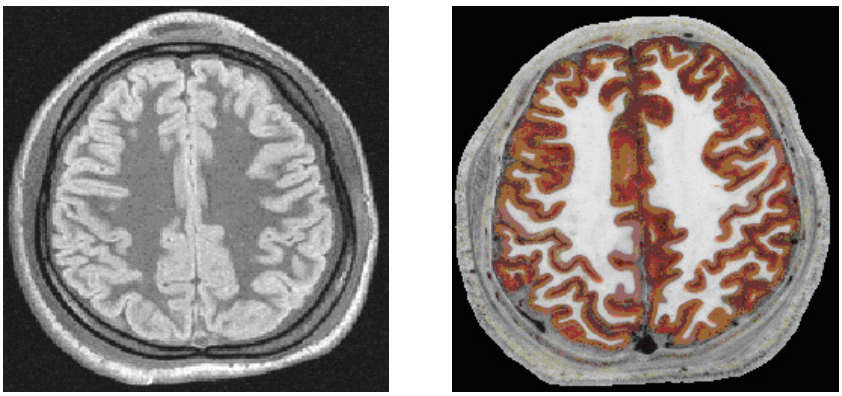
According to expression 2.1, the fusion performance results can are shown in Table 1. From the data in table 1, the advantage of the proposed fusion scheme has been confirmed.



(a) The fused image by original Max MI algorithm[12]

(b) The fused image by the algorithm of this paper

FIGURE 3. The fused image results of two different algorithms



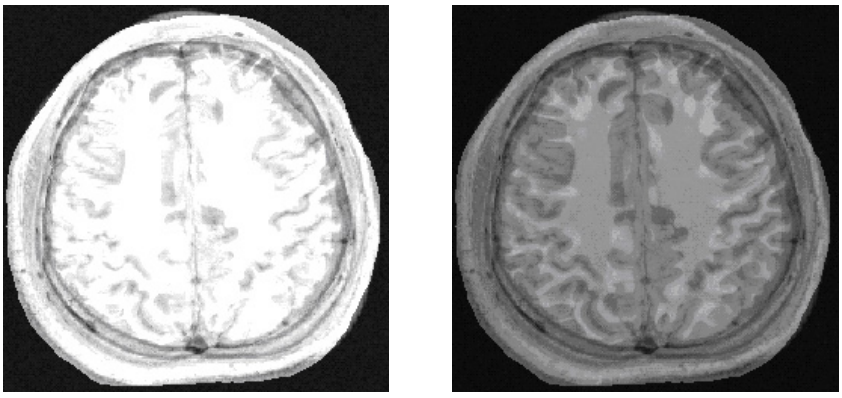
(a) MR image

(b) Color image

FIGURE 4. The original multi-modalities medical images

TABLE 1. The fusion performance assessing results

MRI and Fusion result image	Fig2(b) and 3(a)	2(b) and 3(b)	4(a) and 5(a)	4(a) and 5(b)
Mutual information MI_{AB}(x, y)	3.2667	3.5746	4.3292	4.4067



(a) The fused image by original algorithm of the Max. MI between the MRI and Color images

(b) The fused image by the algorithm of this paper

FIGURE 5. The fused image results of two different algorithms

4. Conclusions

A new image fusion algorithm based on wavelet transform to reach the maximum mutual information has been described and applied to the fusion of multimodality medical images. Our image fusion algorithm shows better results compared to the algorithm of maximal mutual information without wavelet transform [12],[13];

This algorithm can also be used for the fusion of MRI and color images.

Acknowledgement

The CT/MR images and the full color cross-section images of the Chinese Visible Human Project used in our research and shown in this paper are provided by Prof. Zhang Shaoxiang, who is the leader of the *Chinese Virtual Human* research team in the Third Military Medical University in Chongqing, China[11].

We would like to express our sincere appreciation for their help and support.

References

[1] Z. Zhang and R. S. Blum, *A categorization and study of multiscale-decomposition-based image fusion schemes*, Proceedings of the IEEE, vol. 87(8), pp. 1315–1328, August 1999.

[2] L. J. Chipman, T. M. Orr, and L. N. Graham, *Wavelets and image fusion*, in Proc. SPIE, vol. 2569, pp. 208–219, 1995.

[3] H. Li, B. S. Manjunath and S. K. Mitra, *Multisensor image fusion using the wavelet transform*, Graphical Models and Image Processing, vol. 57, pp. 235–245, May 1995.

[4] Paul Hill, Nishan Canagarajah and Dave Bull, *Image Fusion using complex wavelets*, BMVC, pp.478-496, 2002.

[5] Zhao Ruizhen, Xu Long and Song Guoxiang, *Muliscale image data fusion with wavelet transform*, Journal of computer-aided design & computer graphics, 361-364,vol 4(14), April 2002.

[6] Yan Hao, Hou Zengguang and Tan Min, *Multi-scale Image Data Fusion Based on Wavelet Transform*, computer engineering and applications, vol. 134, pp.131-134, 2003.

[7] Guihong Qu, Dali Zhang and Pingfan Yan, *Medical image fusion by wavelet transform modulus maxima*, OSA, pp.184-190, 2001.

[8] Xiaolin Tian, Yan Zhao, Chengjun Tao, Zesheng Tang, *Background Removal for Color Images Based on Color Components Differences*, Proceedings of the Sixth IASTED International Conference on Signal and Image Processing, August 23-25, Hawaii, USA, 2004.

[9] L. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*, Proceedings of the IEEE, vol. 77, pp.257–285, February 1989.

[10] Guixi Chang and Wanhai Yang,*A Wavelet-Decomposition-Based Image fusion Scheme and ITS Performance Evaluation*, ACTA Automatica Sinica, vol 6(28), November 2002

[11] S. X. Zhang, P. A. Heng and et al., *Creation of the Chinese visible human data set*, The Anatomical Record (Part B: New Anatomist), Vol.275B, 2003, pp.190-195.

[12] Josien P. W. Pluim, J. B. Antoine Maintz and Max A. Viergever, *Mutual information based registration of medical images: a survey*, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. XX, NO. Y, MONTH 2003.

[13] Zhou yongxin and Luo Shuqian, *Medical Image Registration Based on Mutual Information of Feature Points*, Journal Of Computer-Aided Design & Computer Graphics, vol.14.July 2002.

Xueke Li
Faculty of Information Technology
Macao University of Science and Technology, Macao
Taipa, Macau
e-mail: kenry_lee@yahoo.com.cn

Xiaolin Tian
Faculty of Information Technology
Macao University of Science and Technology, Macao
Taipa, Macau
e-mail: xltian@must.edu.mo

Yankui Sun
Department of Computer Science and Technology
Tsinghua University,
Beijing, China
e-mail: syk@tsinghua.edu.cn

Zesheng Tang
Faculty of Information Technology
Macao University of Science and Technology, Macao
Taipa, Macau
e-mail: ztang@must.edu.mo

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 397–405
© 2006 Birkhäuser Verlag Basel/Switzerland

Salient Building Detection from a Single Nature Image via Wavelet Decomposition

Yanyun Qu, Cuihua Li, Nanning Zheng,
Zejian Yuan and Congying Ye

Abstract. We describes how wavelet decomposition can be used in detecting salient building from a single nature image. Firstly, we use Haar wavelet decomposition to obtain the enhanced image which is the sum of the square of LH sub-image and HL sub-image. Secondly, we separate the candidates of building from the background based on projection profile. Finally, we discriminate the building by Principle Component Analysis(PCA) in RGB color space. The proposed approach has been tested on many real-world images with promising results.

Mathematics Subject Classification (2000). Primary 99Z99; Secondary 00A00.
Keywords. Salient building detection, wavelet decomposition, PCA.

1. Introduction

Man-made structure detection is important in many applications such as automatic navigation, virtual reality, image understanding etc. But there are some difficulties in automatic detection of building. Firstly, there are lots of clutters in the natural scene image, and the edge detection is very noisy. This make it complicate to separate the building from background. Secondly, unlike the aerial views of buildings, the realistic view is unconstrained which makes it difficult to use predefined models or model-specific properties in detection. Finally, it is difficult to represent building. Building detection is an active research field recently, which can be classified as detection from aerial images and from ground view images. For structure detection of buildings from aerial images, the state-of-art approaches mainly use hypothesis of the roof presence, the low-level image information such as edges, lines, and junctions, and as well as use other information

The work is supported by the Chinese National Nature Science Foundation No.**60205001** and the Foundation of Creative Research Group Programming No.**60021302**.

such as illumination, shadow, height, and DEM[1][2][3][4][5][6][7] . There are also some works about building detection from a building facade image. In [8] a perception grouping based method is used to detect building for image retrieval, and in [9] a model of multiscale random field is proposed to detect building. The disadvantages of these methods are the computing complexity of either line detection or line classification.

We propose a novel approach to detect the salient building which is distinct because of its height and other features. Firstly Haar wavelet decomposition is used to obtain four sub-images: LL sub-image, LH sub-image, HL sub-image, HH sub-image. We use LH sub-image and HL sub-image to enhance the image. Secondly, we propose a projection profile based approach to extract building candidates. Finally, the salient building is discriminated by applying Principle Component Analysis(PCA) in RGB color space.

The organization of this paper is as follows. In section 2 we briefly review two dimension multi-resolution analysis. In section 3 we proposed the wavelet-based segmentation method. Principle component analysis(PCA) in RGB color space is used to discriminate the building in section 4, and conclusion is given in the last section.

2. Review of 2-D Multi-resolution Analysis

Wavelet transformation is popular in 1990s and is widely used in signal processing, image processing and analysis, image compression and coding etc. $L^2(R^2)$ can be constructed by the separable wavelet orthogonal basis which are the product of the 1-D scaling function and the 1-D wavelet function. Let $\{V_j\}_{j\in \mathrm{Z}}$ be the sequence scaling subspace in $L^2(R)$, and $\{V_j^2\}_{j\in \mathrm{Z}}$ which are in space $L^2(R^2)$ can be constructed by the tensor product space $V_j^2 = V_j \otimes V_j$. Let $\{W_j^2\}_{j\in \mathrm{Z}}$ be the wavelet sequence subspace, so $W_j^2 = (V_j^2)^{\perp}$ and $V_j^2 = V_{j+1}^2 \oplus W_{j+1}^2$. The following theorem gives the fundamental theory of constructing the orthonormal bases of $L^2(R^2)$.

Theorem 2.1. [10]*Let ϕ denote the scaling function, ψ denote the corresponding the wavelet function in $L^2(R)$. Define the following wavelet function:*

$$\psi^1(x) = \phi(x_1)\psi(x_2). \tag{2.1}$$

$$\psi^2(x) = \psi(x_1)\phi(x_2) \tag{2.2}$$

$$\psi^3(x) = \psi(x_1)\psi(x_2) \tag{2.3}$$

Let

$$\psi_{j,n}^k(x) = \frac{1}{2^j}\psi^k\left(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j}\right)$$

where $1 \le k \le 3$. The function system $\{\psi_{j,n}^1, \psi_{j,n}^2, \psi_{j,n}^3\}_{n\in \mathrm{Z}^2}$ are the orthonormal basis in W_j^2, and $\{\psi_{j,n}^1, \psi_{j,n}^2, \psi_{j,n}^3\}_{(j,n)\in \mathrm{Z}^3}$ are the orthonormal basis in $L^2(R^2)$.

The three wavelet functions extract the image details at different scale and direction. From equ(2.1), (2.2), (2.3), we obtain that

$$\hat{\psi}^1(w_1, w_2) = \hat{\phi}(w_1)\hat{\psi}(w_2) \tag{2.4}$$

$$\hat{\psi}^2(w_1, w_2) = \hat{\psi}(w_1)\hat{\phi}(w_2) \tag{2.5}$$

$$\hat{\psi}^3(w_1, w_2) = \hat{\psi}(w_1)\hat{\psi}(w_2) \tag{2.6}$$

The magnitude $\hat{\psi}^1(w_1, w_2)$ emphasizes the details which has low frequency in horizontal direction and high frequency in vertical direction, the magnitude $\hat{\psi}^2(w_1, w_2)$ emphasizes the details which has high frequency in horizontal direction and low frequency in vertical direction, the magnitude $\hat{\psi}^3(w_1, w_2)$ emphasizes the details which has high frequency in both directions.

3. Extraction of Building Candidates via Wavelet Decomposition

There are many methods to segment regions of interest which can be classified as edge-based method and region-based method. However the performance is not generally because of the variance of image noise, illumination and building texture. To improve the performance of building detection, we propose a novel approach which is based on wavelet decomposition.

The wavelet decomposition can break down an image into four sub-images:

LL sub-image: Both horizontal and vertical directions have low frequency.

LH sub-image: The horizontal direction has low frequency, and the vertical one has high-frequency.

HL sub-image: The horizontal direction has high frequency, and the vertical one has low-frequency.

HH sub-image: Both horizontal and vertical directions have high frequency.

The building is detected from a ground view image. Since the wall and windows have regular shape in building facade, we separate building candidates by using an image projection profile based method. As mentioned in [11], LH-image is the result that the image is convolved by a lower-pass filter along the horizontal direction and by a higher-pass filter along the vertical direction. So the horizontal detail is enhanced while the vertical detail is smoothed. The result of the HL sub-image is opposite to that of the LH one. The vertical detail is enhanced while the horizontal detail is smoothed. This operation on the image is just like edge detection.By denoting the LH sub-image by $H_2(x, y)$ and the HL sub-image by $V_2(x, y)$, we construct an enhanced image as

$$B(x, y) = H_2(x, y)^2 + V_2(x, y)^2 \tag{3.1}$$

Although many orthogonal wavelets meet with our needs, we use the simplest orthogonal wavelet–Haar wavelet which is discontinuous and resemble the step function. It was defined as

$$h(t) = \begin{cases} 1, 0 \le t < \frac{1}{2} \\ -1, \frac{1}{2} \le t \le 1 \\ 0, others \end{cases} \tag{3.2}$$

and

$$h_{j,k}(t) = 2^{-\frac{1}{2}} h(2^{-j}t - k), j, k \in z \tag{3.3}$$

It is easy to obtain from equ(3.3)

$$h_{j,k}(t) = \begin{cases} 2^{-\frac{j}{2}} & 2^j k \le t < (2k+1)2^{j-1} \\ -2^{-\frac{j}{2}} & (2k+1)2^{j-1} \le t < (k+1)2^j \\ 0 & others \end{cases} \tag{3.4}$$

From equ(3.2) (3.3) (3.4), it's obvious that Haar wavelet functions act as the gradient-based edge detection, while the wavelets do not have the property. Furthermore Haar wavelet makes the computation easier. We compare the performance of Haar wavelet and other orthogonal wavelets on the synthesized image which contains a white rectangle in the black background. Due to the space limitation, we only show the result of **db4** wavelet and Haar wavelet in Figure 1. It is clear that the horizontals and verticals in the enhanced image are shifted when we use **db4** wavelete decomposition, while Haar wavelet does not cause the problem. Testing results demonstrate that Haar wavelet is more suitable to building detection. We use one level Haar wavelet decomposition, and select HL sub-image and LH sub-image to compute the enhanced image in order to improve the computational efficiency. An example using Haar wavelet decomposition for enhanced image is shown in Figure 2.

In the following, we explain the extraction of the building candidate based on projection profile. The image edges within the building of the same column or row are accumulated at the same location of the projection profile histograms. There exist high frequency peaks in the horizontal and vertical histograms where the building located. The building candidates can be separated based on the vertical projection profile and horizontal projection profile. Because there exist image textures other than buildings in the image, we should choose appropriate threshold values to select those peaks that corresponds to the horizontal and vertical positions of buildings in the image. The projection profile based approach is as follows.

Step 1. Denoise the image B(x,y) to make the background flatter.

Step 2. Compute the amplitude projections of $B(x, y) \in R^{m \times n}$ along its row, that is

$$X_r(x) = \sum_{y=1}^{m} B(x, y), 1 \le x \le n$$

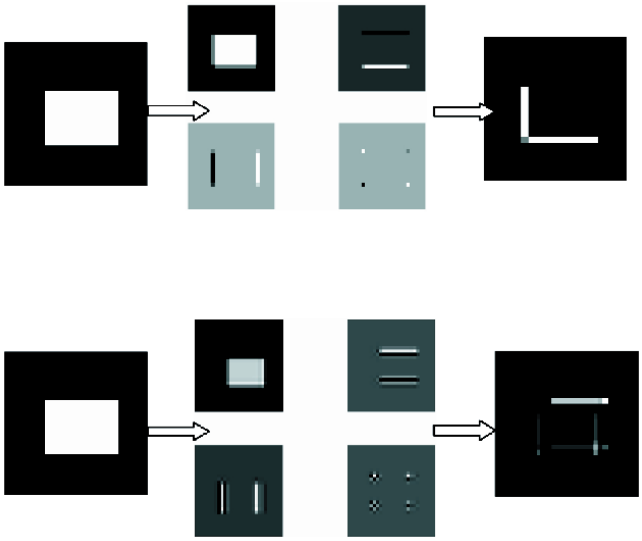


FIGURE 1. The comparison between Haar wavelet decomposition and **db4** wavelet decomposition. The first row is the result after Haar wavelet decomposition and the second row is the result after **db4** wavelet decomposition. In each row, the left is original image, the middle is the sub-images, and the right is the enhanced image.

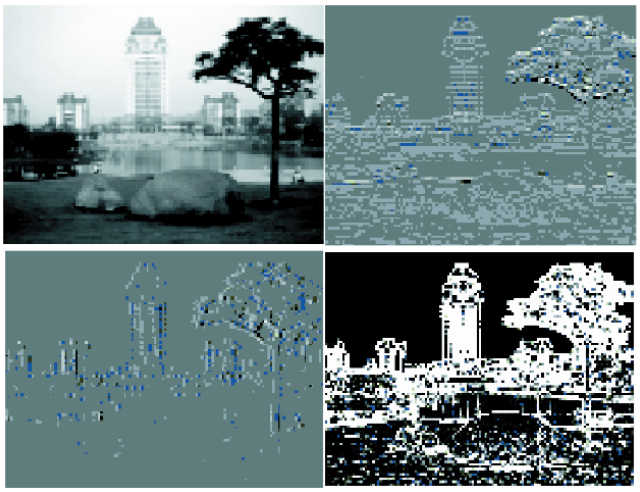


FIGURE 2. The sub-images and enhanced image of a building image. From left to right and up to down, they are the LL sub-image, LH sub-image, HL sub-image, and the enhanced image.

Step 3. Find the first non-zero value in the array $X_r(x)$, and obtain the corresponding index which is the upper bound of buildings, that is

$$upper = \min_x\{x\,|X_r(x) > 0\,\}$$

Delete the part above the upper bound, and obtain the image $Bcut(x, y)$.

Step 4. Find the non-zero value in every column from top to down in $Bcut(x, y)$, and obtain the corresponding index array, that is

$$S(y) = \min_x\{x\,|Bcut(x, j) > 0\ , j = y\}, 1 \le y \le m.$$

Step 5. Find the maximum and minimum in array $S(y)$, and compute the mean as the lower bound, that is

$$lower = (\max\{S(y)\} + \min\{S(y)\})/2.$$

Delete the part below the lower, and denote the image as $Btop(x, y)$

Step 6. Compute the amplitude projections of $Btop(x, y)$along its column, that is

$$X_c(y) = \sum_{x=1}^{n-upper-lower} Btop(x, y), 1 \le y \le m,$$

and cluster the indices whose corresponding values are non-zero, to obtain the left bound $Bleft(i)$ and the right bound $Bright(i)$, the length of the array $Bleft(i)$ is the number of building candidates.

Step 7. Find the region $region(i)$in the original image corresponding region in image B(x,y).

Remark: From step 1 to step 5, the operation is to separate the high buildings from the sky and the low building group. From step 6 to step 7, the operation is to separate salient building candidates from each other. Figure 3 illustrates the result of the approach.

4. Building Discrimination by PCA in RGB Color Space

From section 3, we obtain building candidates which may be building or not.We need to discriminate buildings in the next step. A hypothesis is generated that the desired region can be classified into two classes one of which is building, the other of which is tree, because we believe that the hypothesis can cover a significant fraction of the building image. In order to make the discrimination problem simple, we consider the negative sample, i.e. trees. We compute the principle color of the tree color samples and compare the color of building candidates with it. Let $\{A_i\}, 1 \le i \le m$, be the color vector generated from the desired region, where m is the number of the points which correspond to separated region, and the mean vector is denoted by μ . The new vectors are given by

$$\bar{A}_i = A_i - \mu, i = 1, 2 \cdots m \tag{4.1}$$

FIGURE 3. The segmentation result. The building candidates are indicated by rectangle.

The covariance matrix is given by

$$Cov = A^T A$$

where $A = [\bar{A}_1^T, \bar{A}_2^T, \cdots \bar{A}_m^T]^T$.

Compute the eigenvector corresponding to the largest eigenvalue of the matrix Cov, denoting it by the principle color in desired region, i.e.,

$$Cov * V = \lambda_{\max} V \tag{4.2}$$

The principle color of tree sample was denoted by V_{mod}. The color similarity between the model and the desired region is measured by the city block distance, and the desired region is classified as the building if the distance is larger than a threshold σ i.e.,

$$flag(i) = \begin{cases} 0 & \| V - V_{\text{mod}} \| < \sigma \\ 1 & others \end{cases}$$

where $flag(i)$ denote the class label of the desired region.

Experiments are performed on real-world images using matlab 6.5 on Intel P4 2.4G platform. The average time taken in processing an 372×496 image is no more than 2s, which is faster than that in [9]. Figure 4 shows some examples of building detection from a few of real ground view images.

5. Conclusion

This paper focuses on detecting salient building from a ground view building image. Wavelet decomposition and projection profile based approach are proposed to separate the building from the background. PCA is implemented in the RGB color space in order to discriminate the building. Real-world images are used to

FIGURE 4. Building detection example. The detected building is indicated by red rectangle

evaluate the proposed method, and the experimental results demonstrate that the proposed method is feasible and effect.

References

[1] D.M. McKeown,*Toward Automatic Cartographic Feature Extraction,Mapping and Spatial Modelling for Navigation.* L.F. Pau, ed., NATO ASI series **65** (1990), 149-180.

[2] R.B.Irvin and D.M. McKeown, *Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery.* IEEE Trans. Systems, Man, and Cybernetics**19** (1989), 1564-1575.

[3] J.C. McGlone and J.A. Shufelt, *Projective and Object Space Geometry for Monocular Building Extraction.* Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2000, 54-61.

[4] J.A. Shufelt, *Exploiting Photogrammetric Methods for Building Extraction in Aerial Images.* Int'l Archives of Photogrammetry and Remote Sensing**XXXI, B6/S** (1996), 74-79.

[5] J.A. Shufelt, *Projective Geometry and Photometry for Object Detection and Delineation.* PhD thesis, Computer Science Dept., Carnegie Mellon Univ., available as Technical Report CMU-CS-96-164, 29 July 1996.

[6] C. Lin, R. Nevatia, *Building Detection and Description from a Single Intensity Image.* Computer Vision and Image Understanding **72** (1998),101-121.

[7] H. Kauppinen, T. Seppanen, M. Pietikainen, *An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classi1cation.* IEEE Trans. Pattern Anal. Mach. Intell **17(2)** (1995), 201-207.

[8] A. Iqbal and J.K.Aggarwal, *Applying perceptual grouping to content-based image retrieval:Builing images.* In Proc. IEEE Int. Conf. CVPR **1** (1999), 42-48.

[9] S. Kumar and M. Hebert, *Man-made Structure Detection in Natural Images using a Causal Multiscale Random Field.* In Proc. IEEE Int. Conf. on CVPR**1** (1999), 119-126.

[10] S.Mallat, *A Wavelet Tour of Signal Processing.* 2rd Edition, Acadamic Press, 1999.

[11] Y.Y.Tang, H.Ma,J.Liu,B. Li, and D. Xi, *Multiresolution Analysis in Extraction of Refoerence Lines from Documents with Gray Level Background.* IEEE Trans. Pattern Anal. Mach. Intell **19(8)** (1997), 921-926.

Yanyun Qu
Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University
Xi'an,710049
Department of Computer Science, Xiamen University
P.R.China
e-mail: yyqu@xmu.edu.cn

Cuihua Li
Department of Computer Science, Xiamen University
P.R.China
e-mail: chli@xmu.edu.cn

Nanning Zheng
Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University
P.R.China
e-mail: nnzheng@mail.xjtu.edu.cn

Zejian Yuan
Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University
P.R.China
e-mail: zjyuan@aiar.xjtu.edu.cn

Congying Ye
Department of Computer Science, Xiamen University
P.R.China

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 407–417
© 2006 Birkhäuser Verlag Basel/Switzerland

SAR Images Despeckling via Bayesian Fuzzy Shrinkage Based on Stationary Wavelet Transform

Yan Wu, Xia Wang and Guisheng Liao

Abstract. An efficient despeckling method is proposed based on stationary wavelet transform (SWT) for synthetic aperture radar (SAR) images. The statistical model of wavelet coefficients is analyzed and its performance is modeled with a mixture density of two zero-mean Gaussian distributions. A fuzzy shrinkage factor is derived based on the minimum mean error (MMSE) criteria with bayesian estimation. In this case, the ideas of region division and fuzzy shrinkage are adopted according to the interscale dependencies among wavelet coefficients. The noise-free wavelet coefficients are estimated accurately. Experimental results show that our method outperforms the refined Lee filterwavelet soft thresholding shrinkage and SWT shrinkage algorithms in terms of smoothing effects and edges preservation.

Keywords. SAR image despeckling, fuzzy shrinkage factor, MMSE, regions division, Bayesian estimation, SWT.

1. Introduction

The synthetic aperture radar (SAR) generates images that are severely degraded by speckle, a type of multiplicative noise. Due to its granularity in an image, speckle noise makes it very difficult to visually and automatically interpret SAR data. Therefore, reduction of the speckle noise is important in most detection and recognition systems where speckle is present.
In the past ten years, many algorithms have been developed to suppress speckle noise in order to facilitate postprocessing tasks. The earliest method is multilook

This work was supported by the National Defence Foundation (51431020204DZ0101) and the Postdoctoral Science Foundation of China (J63104020156).

processing, but this method sacrifices the spatial resolution. Later, the spatial filters which are based on the minimum mean square error (MMSE) criteria are traditionally used, such as the Lee-filter[1], the Kuan-filter[2,3] and the Frost-filter[4], but an insufficient existing in these techniques is that the region with edges or strong textures are blurred after filtering. Then some statistical adaptive filters for speckle reduction appeared, they are capable of adapting the size or the shape of the local window according to the underlying structural features. Examples of such filters are the refined-Lee filter[5] and Epos filter[6], etc. They outperform the former filters. Recently, there has been considerable interest in using the wavelet transform as a powerful tool for recovering noisy data. As an outcome of the wavelet theory, despeckling in the domain where SAR images are defined by detail coefficients of the noisy image, either hard or soft, are firstly proposed by Donoho[7]. These methods involve a preprocessing step consisting of a logarithmic transform to convert the multiplicative noise into the additive noise. The key point in these thresholding techniques is the selection of an appropriate threshlold. If its value is too small, the recovery image will remain noisy. On the other hand, if its value is too large, artifacts like pseudo-Gibbs oscillations may appear near discontinuities in the reconstructed signal. Another approach involves the stationary wavelet transform (SWT). F.Argenti et al. applies a local linear MMSE estimator in the stationary wavelet transform domain[8]. The method avoids using the logarithmic transform, but it does not utilize statistical assumptions about the wavelet coefficients. This is a drawback in cases where good statistical models exist. Furthermore, the filter performs the same in wavelet domain.

In this paper, we propose an efficient SWT despeckling method based on bayesian estimation using MMSE criteria. We avoid using the log-transform. Firstly, we analyze the statistical properties of the wavelet coefficients, and model the coefficients with a mixture density of two zero-mean Gaussian distributions. Then we derive a bayesian wavelet shrinkage factor based on MMSE criteria. Finally, according to the interscale dependencies of the wavelet coefficients, we adopt fuzzy shrinkage to the factor and divide different regions among the coefficients. Different methods are used for different regions. The algorithm achieves good performance.

2. The Statistical Model in SWT Domain

2.1. SWT

SWT[9] is a special version of the DWT that has preserved translation invariance. This means that a translation of the original signal does not cause a translation of the corresponding wavelet coefficients. Instead of subsampling, the SWT utilizes recursively dilated filters in order to halve the bandwidth from one level to another. At scale 2^j the filters are dilated by inserting 2^{j-1} zeros between the filter coefficients of the prototype filters. This is performed in order to reduce the bandwidth by a factor of two from one level to another.

$$h_k^j = \begin{cases} h_{k/2^j} & \text{if } m \in Z \\ 0 & \text{else} \end{cases} \tag{2.1}$$

$$g_k^j = \begin{cases} g_{k/2^j} & \text{if } m \in Z \\ 0 & \text{else} \end{cases} \tag{2.2}$$

where $\{h_k\}$ and $\{g_k\}$ are low-pass and high-pass filters, respectively.
For images, image rows and columns are then filtered separately. Filtering equations to obtain the level $j+1$ from the level j are the following (where (a,b) is for the pixel position):

$$A_{X,j+1}(a,b) = \sum_{k,m} h_k^j h_m^j A_{X,j}(a+k, b+m) \tag{2.3}$$

$$W_{X,j+1}^h(a,b) = \sum_{k,m} g_k^j h_m^j A_{X,j}(a+k, b+m) \tag{2.4}$$

$$W_{X,j+1}^v(a,b) = \sum_{k,m} h_k^j g_m^j A_{X,j}(a+k, b+m) \tag{2.5}$$

$$W_{X,j+1}^d(a,b) = \sum_{k,m} g_k^j g_m^j A_{X,j}(a+k, b+m) \tag{2.6}$$

where $A_{X,j}$ is the approximation of the original image at the scale 2^j, giving the low-frequency content. Images details are contained in three high-frequency images $W_{X,j}^h$, $W_{X,j}^v$, $W_{X,j}^d$ corresponding to horizontal, vertical and diagonal detail orientations, respectively. Since SWT does not include downsampling operations, it is a redundant transform.

2.2. The Statistical Model in SWT Domains

Wavelet coefficients of SAR images typically exhibit strong non-Gaussian statistics. The correct choice of priors for wavelet coefficients is certainly a very important factor. Several different priors have been considered for the wavelet coefficients. In [10], wavelet coefficients are modeled as a generalized Gaussian distribution (GGD), which matches well the histograms of typical SAR images. However, in the bayesian estimation process, there usually does not exist a closed-form solution for the estimate of noise-free wavelet coefficients when the signal prior is described by the GGD. Among alternative methods, a mixture density of two zero-mean Gaussian distributions has been proposed due to its relatively simple form and high accuracy in modeling the distribution of wavelet coefficients[11]. One mixture component is corresponding to significant coefficients (representing "homogeneity"), and the other is corresponding to significant coefficients (representing "heterogeneity").
For SAR imagery, we have a multiplicative speckle model: $X = YF$, where X is the noisy observation, Y is the noise-free signal and F is normalized speckle random variable with unit mean. Assuming that W_X represents the noisy wavelet coefficients, its mixture probability density function (pdf) is given by

$$p_{W_X}(w_x) = \sum_{k=0,1} p(s=k) p(w_x | s=k) \tag{2.7}$$

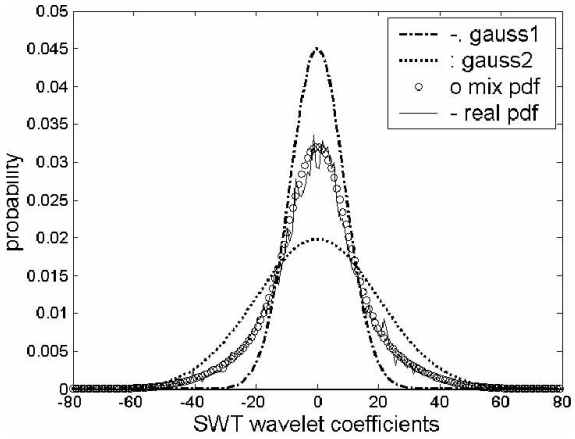


FIGURE 1. Histogram of SWT wavelet coefficients and the Mixture-Gaussian model.

where $p(w_x|s=k)$ is a zero-mean Gaussian distribution and $S=0$ or 1 represents each Gaussian component in the mixture distribution.
We demonstrate one typical example in Figs. 1 to show the performance of the mixture-Gaussian model in matching the distribution of SWT wavelet coefficients of a real SAR image. In the figure, the mixture Gaussian distribution is labeled as "mix pdf" and two pure Gaussian components are labeled as "gauss1" and "gauss2", respectively. As the figure shows, the mixture-Gaussian model provides a fairly accurate approximation to the distribution of wavelet coefficients of real SAR image.

3. Bayesian Fuzzy Wavelet Estimate

3.1. Bayesian Wavelet Estimation

The wavelet decomposition operation can be written as

$$W_X = W[X] = W[YF] = W[Y] + W[Y(F-1)] = W_Y + W_N \tag{3.1}$$

where W_N is an additive signal-dependent noise. Shrinkage estimate of the noise-free wavelet coefficient is $\hat{W}_Y = \eta W_X$, where η is the shrinkage factor. Based on the MMSE criteria, the optimal shrinkage factor η' is obtained by minimizing the mean square error between $\hat{W}_Y$ and W_Y

$$\eta' = \arg\min_{\eta} E[(\hat{W}_Y - W_Y)^2] \tag{3.2}$$

which has the MMSE solution in the form of

$$\eta' = \frac{E[W_X^2] - E[W_X W_N]}{E[W_X^2]} \tag{3.3}$$

To calculate the η, it is necessary to estimate all unknown parameters in Eq.(10). Based on a mixture-Gaussian model of wavelet coefficients, the MMSE estimate of noise-free wavelet coefficients $\hat{W}_Y$ is

$$\hat{W}_Y = \sum_{k=0,1} p(s=k|w_x)\frac{\sigma^2_{W_X} - \sigma^2_{W_N}}{\sigma^2_{W_X}} W_X \tag{3.4}$$

where $\sigma^2_{W_X}$ is the variance of the noisy wavelet coefficients W_X in the corresponding state S and $p(s=k|w_x)$ is calculated based on the Bayes rule

$$p(s=k|w_x) = \frac{p(w_x|s=k)p(s=k)}{p(W_X)} \tag{3.5}$$

The unknown probabilities in Eq.(12) are calculated with the expectation-maximization(EM) algorithm[11].

Proof. Since the speckle-noise random variable F is usually normalized as: $E[F] = 1$, $E[F-1] = 0$. Thus, due to the high-pass nature of wavelet functions, we have $E[W_N] = 0$. Furthermore, because of the zero-mean mixture-Gaussian distribution model of W_X, thus $E[W_X] = 0$. Therefore, we obtain

$$E[W_N^2] = \sigma^2_{W_N}; \quad E[W_X^2] = \sigma^2_{W_X} \tag{3.6}$$

Since W_Y and W_N are statistically independent, thus

$$E[W_X W_N] = E[(W_Y + W_N)W_N] = E[W_Y W_N] + E[W_N^2] = \sigma^2_{W_N}; \tag{3.7}$$

Combining Eq.s(10), (13), (14) with Eq.(7), we can obtain Eq.(11).

□

The variance of noise in wavelet domain $\sigma^2_{W_N}$ equals[12]

$$\sigma^2_{W_N} = \frac{\Psi_j \mu_X^2 + \sigma^2_{W_X}}{1 + C_F^2} C_F^2 \tag{3.8}$$

where $\mu_X = E[X]$, the normalized standard deviation of noise C_F equals $\sqrt{(4/\pi - 1)/L}$ for amplitude images and $\sqrt{1/L}$ for intensity images ($L \geq 1$), and parameter Ψ_j is defined as $\Psi_j = \left(\sum_k (h_k)^2\right)^2 \sum_m \left(g_m\right)^2)^{2(j-1)}$, where h and g are the high-pass and low-pass filters at the decomposition level j, respectively.

3.2. Fuzzy Shrinkage Factor

Since the shrinkage factor based on the MMSE criteria is the minimum mean value and this factor from each wavelet coefficient obtained by SWT is invariable, we adopt fuzzy transform to modify it according to the variety of the wavelet coefficients. This factor is redefined by

$$\eta_{f,j} = \frac{\eta'}{1 + exp(\alpha(c - \sigma_j^2))} \tag{3.9}$$

where $c = |W_{X,j}||\hat{W}_{Y,j+1}|$; For each image detail $W_{X,j} = W_{Y,j} + W_{N,j}$, $\sigma_j = Median(|W_{X,j} - Median(W_X, j)|)/0.6745$, σ_j^2 is its threshold value[13], and σ_j

is the noise standard deviation at scale 2^j. Since noise mostly exists in subimage HH, σ_j is calculated by the standard deviation of the wavelet coefficients in HH; $\alpha < 0$ and $\eta_{f,j} = \eta'$ when α =-∞. α determines the convergence speed of the function. To suppress the noise, we hope that the absolute value of α is smaller when noise variance is bigger, i.e., the convergence speed is slower. Thus we define the expression of α by

$$\eta = (\frac{W_X - W_{X,min}}{W_{X,max} - W_{X,min}} - 1) \times 1000\,. \tag{3.10}$$

Since we use the intensity image, we set the scope of α as [0,−1000]. $W_{X,min}$ is the minimization of W_X in the corresponding decomposition level and $W_{X,max}$ is the maximization. Thus,

$$\hat{W}_Y = \sum_{k=0,1} p(s = k|W_X)\eta_{f,j}W_X\,. \tag{3.11}$$

3.3. Regions Division

As the wavelet transform deepened its application in image coding and denoising, researchers proposed some more complicated and more accurate models that exploit interscale dependencies and intrascale dependencies among wavelet coefficients. Hidden markov random field tree model[14] can capture the interscale dependencies, but it costs too much time to calculate the parameters in the model. To take spatial dependence into account, we adopt the ideas of region division and fuzzy transform. The method is easy and feasible, furthermore, it costs little in computation.

As the scale increases, the large wavelet coefficients of the signal almost keep invariable, but the wavelet coefficients of the noise tend to decrease greatly. In [14], for the necessity to distinguish signal from noise ahead, the author indicated that the pixel was considered as noise if $c < \sigma_j^2$ and the pixel was considered as signal if $c > \sigma_j^2$. As this dividing method is a little coarse, we further redefine the criteria of region division. We choose proper n_1 and n_2 $(0 < n_1 < 1 < n_2 < \infty)$ so that we can obtain the following cases: the probability which denote that the pixel is in homogenous neighborhood (i.e. the point is noise) is close to 1 when $c < n_1\sigma_j^2$; the probability which denote that the pixel is an edge point (i.e. the point is signal) is close to 1 when $c > n_2\sigma_j^2$; and the pixel is considered to be in the regions between homogenous neighborhood or edge point when $n_1\sigma_j^2 \le c \le n_2\sigma_j^2$. Finally, choosing $n_1 = 1/2$ and $n_2 = 5/3$ in the paper according to our many experiments. For each image detail $W_{X,j}(a,b) = W_{Y,j}(a,b) + W_{N,j}(a,b)$, the estimate of noise-free $\hat{W}_Y$ is

$$\hat{W}_{Y,j} = \begin{cases} 0 & c < n_1\sigma_j^2 \\ \sum_{k=0,1} p(s = k|W_X)\eta_{f,j}W_{X,j} & n_1\sigma_j^2 \le c \le n_2\sigma_j^2 \\ W_{X,j} & c > n_2\sigma_j^2 \end{cases} \tag{3.12}$$

The steps of the proposed algorithm are as follows:

1. Decompose the original SAR image by SWT and the number of decomposition level is 3.
2. Estimate parameters of the mixture-Gaussian model with EM algorithm, and then calculate the value of $\sigma^2_{W_N}$.
3. Shrink three detail subimages at the decomposition level 3 according to Eq.(11), and obtain the estimate of $\hat{W}_{Y,3}$.
4. At decomposition level 2 and 1, the despeckling algorithm is applied separately to detail subimages according to Eq.(19).
5. Apply an inverse SWT to the despeckled detail images.

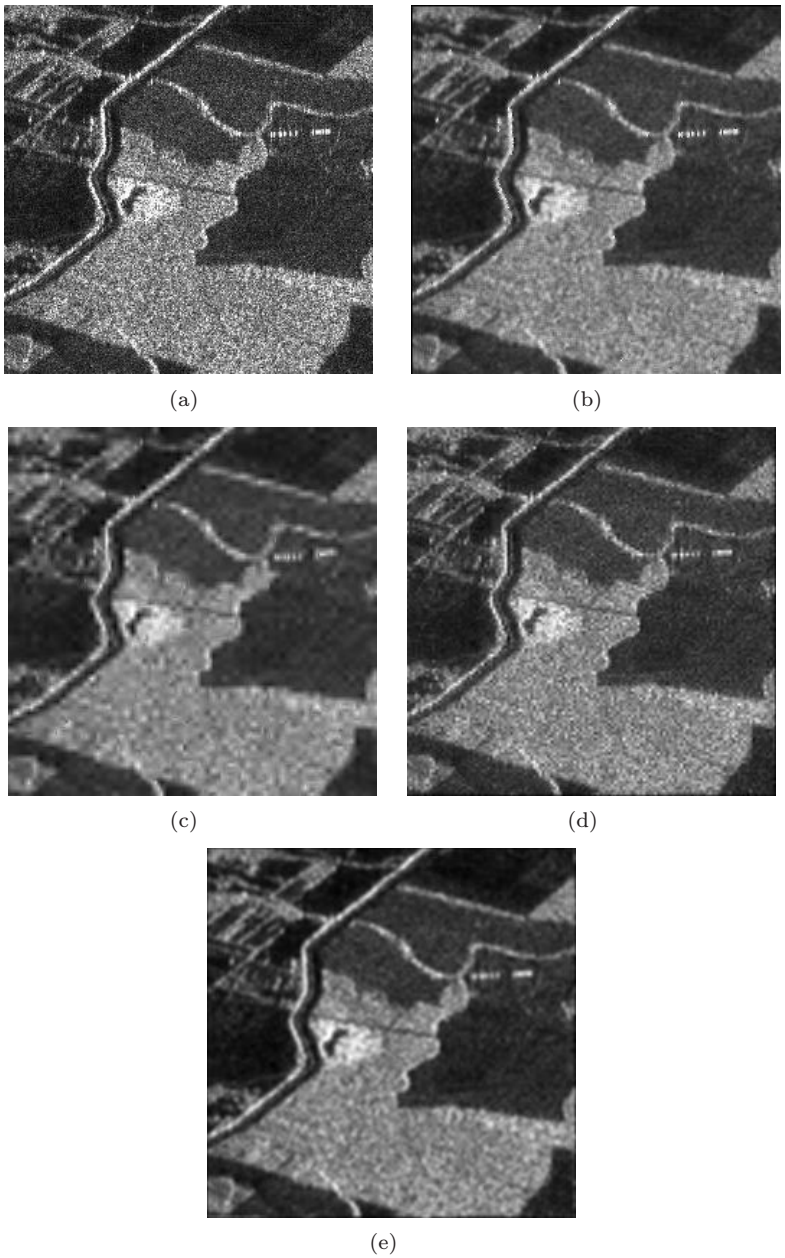
4. Experimental Results and Discussion

In our experiments, we are quantifying the algorithm performance in terms of smoothing effects and edge preservation. We use equivalent number of looks (ENL) to measure the smoothing effects of despeckling methods. It is defined as $ENL = \mu^2/\sigma^2$, where μ and σ are the mean and the variance of intensity value over a homogenous region. Furthermore, as a measure of edge preservation, we adopt the Pratt's figure of merit [16], which is defined by

$$FOM = \frac{1}{max(N_A, N_I)} \sum_{i=1}^{N_A} \frac{1}{1 + d_i^2 \alpha} \tag{4.1}$$

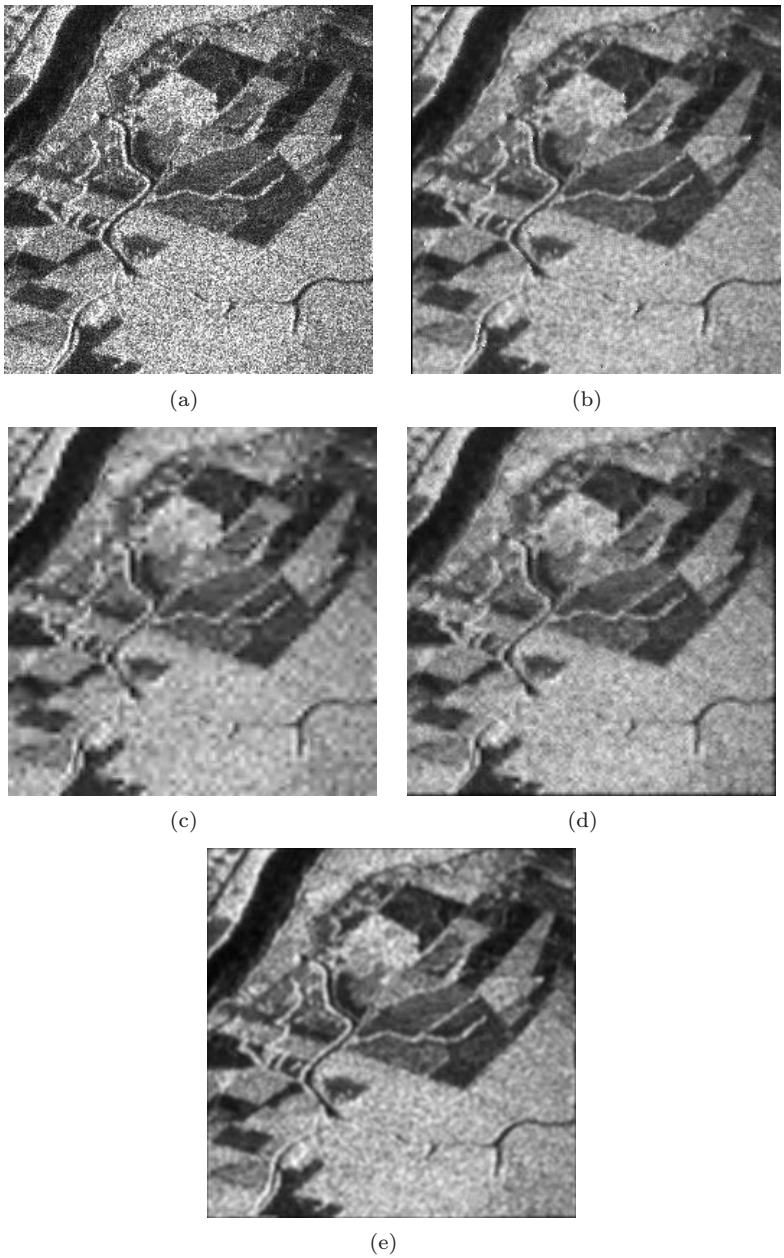
where N_A and N_I are the numbers of original and filtered edge pixels, respectively, d_i is the Euclidean distance between the original edge pixel and the nearest filtered edge pixel, and α is a constant, which is typically set to $1/9$. FOM ranges between 0 and 1, with unity for ideal edge detection.

The proposed algorithm is tested on two 256×256 real SAR images. We find that the refined Lee filter possesses the best standard spatial trade off between noise reduction and feature preservation. Therefore we also include the refined Lee filter in the comparison. The other methods include the following: wavelet soft thresholding shrinkage algorithm and SWT shrinkage algorithm[8]. The original SAR images are shown in Fig. 2 (a) and Fig. 3 (a), respectively. The filtered images by the refined Lee filter are shown in Fig. 2 (b) and Fig. 3 (b), respectively. The filtered images by the wavelet soft thresholding algorithm are shown in Fig. 2(c) and Fig. 3 (c), respectively. The filtered images by SWT shrinkage algorithm[8] are shown in Fig. 2 (d) and Fig. 3 (d), respectively. The filtered images by the proposed algorithm are shown in Fig. 2(e) and Fig. 3 (e), respectively. These results are obtained by using bior2.2 mother wave-let. We observe that oscillations appear near the edges in Fig. 2 (c), and it illuminates that DWT may cause Gibbs effects. Fig. 2 (d) is over-smoothed and thus many features are blurred, the reason is that the algorithm does not model the wavelet coefficients or divided regions. However, it is observed that Gibbs effects presented in Fig. 2 (c) disappear from Fig. 2 (d). Thus it still outperforms soft thresholding algorithm. As it appears, the refined Lee filter outperforms two algorithms above in terms of smoothing effects and edge preservation. From Fig. 2 (e), it is observed that the proposed algorithm



(a) (b) (c) (d) (e)

FIGURE 2. Comparison of different despeckling methods (a)The original SAR image (b) The refined Lee filter (c) Soft thresholding shrinkage algorithm (d)SWT shrinkage algorithm (e)The proposed algorithm



(a) (b) (c) (d) (e)

FIGURE 3. Comparison of different despeckling methods (a)The original SAR image (b) The refined Lee filter (c) Soft thresholding shrinkage algorithm (d)SWT shrinkage algorithm (e)The proposed algorithm

TABLE 1. Comparisons of different despeckling methods in terms of ENL and FOM

	SAR image1		SAR image2	
	ENL	FOM	ENL	FOM
The original image	8.579		11.993	
Soft thresholding	36.195	0.6877	64.990	0.6587
The refined Lee	44.363	0.7761	71.928	0.7419
SWT shrinkage algorithm	39.037	0.6973	67.884	0.6858
The proposed algorithm	58.881	0.7943	85.493	0.7518

has a good despeckling performance and typically preserves minor edges, at the same time, it sharpens edges. It is because that the proposed algorithm adopts the ideas of region division and fuzzy shrinkage. The ENL and FOM values of applying these four despeckling algorithms to the two real SAR images are listed in Table 1, from which we note the following observations: (1) under all conditions, the proposed algorithm achieves remarkable improvement over the other three methods; (2) the SWT shrinkage algorithm [8] outperforms the wavelet soft thresholding algorithm. The evaluation fits the visual analysis above. The visual effect analysis in Fig. 3 and its evaluation consist with the cases above.

5. Conclusions

The proposed algorithm combines image multiscale analysis with bayesian shrinkage. The multiplicative model introduced in high-frequency images, permits to retain coefficients produced by significant structures presented in the image and suppress those produced by the speckle noise. We modeled wavelet coefficients with a mixture density of two zero-mean Gaussian distributions. Then a fuzzy shrinkage factor is derived based on the MMSE criteria with bayesian estimation. The ideas of region division and fuzzy shrinkage are adopted according to the interscale dependencies among wavelet coefficients. Finally the noise-free wavelet coefficients are estimated accurately. Experimental results show that our method outperforms the other three despeckling methods in terms of speckle reduction and edges preservation.

References

[1] J.S.Lee, *Digital image enhancement and noise filtering by use of local statistics.* IEEE Trans on Pattern Anal. Machine Intell, vol.2, no.2, pp.156–163,1980.

[2] D.T.Kuan, AA.Sawchuk, TC.Strand, *Adaptive noise smoothingfilter for images with signal-dependent noise.* IEEE Trans on Pattern Anal. Machine Intell, vol.7, no.2,pp.165–177, 1985.

[3] D.T.Kuan, AA.Sawchuk, TC.Strand, *Adaptive restoration of images with speckle.* IEEE Trans on Acoust.,Speech,Signal Processing,vol.35 no.3, pp. 373–383,1987.

[4] V.S.Frost, JA.Stiles, KS.Shanmugan, JC.Holtaman, *A model for radar images and its application to adaptive digital filtering of multiplicative noise.* IEEE Trans on Pattern Analysis and Machine Intelligence, vol.4, No.2, pp. 157–166,1982.

[5] Lopes.A, Touzi.R, Nezry.E, *Adaptive speckle filters and scene heterogeneity.* IEEE Trans Geosci. Remote Sensing, vol. 28, pp.992–1000, Nov. 1990.

[6] Hagg W. Sites M., *Efficient speckle filtering of SAR images.* Processing of the International Geoscience and Remote Sensing ymposium (IGARSS5),1994

[7] D.L.Donoho, *De-Noising by soft-thresholding.* IEEE Trans. Information Theory.vol.41,pp.613–627, 1995

[8] F.Argenti, *Speckle removal from SAR images in the undecimated wavelet domain.* IEEE Trans Geosci. Remote Sensing, vol.40, no.11, pp.2363–2374, 2002.

[9] G.P.Nason,BW.Silverman, *The stationary wavelet transform and some statistical applications.* Dept. Math., Univ.Bristol, U.K., Tech. Rep.,BS8 1TW, 1995

[10] H.Chipman,E.Kolaczyk,R.McCulloch, *Adaptive bayesian wavelet shrinkage.* J.Amer. Statist.Assoc., vol.92, pp. 1413–1421,1997.

[11] R.A.Redner,HF.Walker, *Mixture densities, maximum likelihood and the EM algorithm.* SIAM Rev., vol, 26, no.2,pp.195-239,1984.

[12] F.Samuel, *Multiscale MAP filtering of SAR images.* IEEE Trans Image processing,vol 10, no.1, pp.49–60,Jan, 2001.

[13] P.Thitima, *A simple SAR speckle reduction by wavelet thresholding.* http://www.gisdevelopment.net/aars/acrs/1998/ps1/ps1014pf.htm.

[14] Matthew S. Crouse. Robert D. Nowak. *Wavelet-Based Statistical Signal Processing Using Hidden Markov Models.*IEEE Trans Image processing, vol 46, no.4, pp. 886–902,April, 1998.

[15] Yongjian Yu.,Scott TA. *Speckle reducing anisotropic diffusion.* IEEE Trans Image processing.vol 11,no.11,pp.927–935,Nov,2002.

Yan Wu
National key lab. of radar signal processing, Xidian Univ.,
Xi'an, 710071, China
School of electronics engineering, Xidian Univ.
Xi'an, 710071, China
e-mail: ywu@mail.xidian.edu.cn

Xia Wang
School of electronics engineering, Xidian Univ.,
Xi'an, 710071, China

Guisheng Liao
National key lab. of radar signal processing, Xidian Univ.,
Xi'an, 710071, China

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 419–430
© 2006 Birkhäuser Verlag Basel/Switzerland

Super-Resolution Reconstruction Using Haar Wavelet Estimation

C.S. Tong and K.T. Leung

Abstract. High resolution image reconstruction refers to the reconstruction of a high resolution image from a set of shifted, blurred low resolution images. Many methods have been developed, and most of them are iterative methods. In this paper, we present a direct method to obtain the reconstruction. Our method takes advantages of the properties of Haar wavelet transform of the high resolution image and its relationship with the low resolution images. Thus the coefficients of the Haar wavelet transform of the high resolution image can be estimated from the low resolution images. Our method is very simple to implement and is very efficient. Experiments show that it is robust to boundary conditions and superior to the least - squares method especially in the low - noise case.

Keywords. High resolution image reconstruction, Haar wavelet transform.

1. Introduction

Nowadays, more and more people prefer to take photos using a digital camera instead of the traditional camera. They seek a high resolution digital camera, but there are not many people can afford to buy this kind of camera. Therefore, high resolution image reconstruction (HRIR) can help these people to get a high resolution image from a set of low resolution images. These low resolution images are under-sampled with subpixel displacement from the high resolution image. More importantly, such a technique enables us to make the best use of the existing limits of digital photography and have important applications in remote sensing.

From Zhang *et al.*[16], we know that there are many HRIR algorithms such as frequency domain method [15, 9, 8], the projection onto convex sets method [11, 12], the maximum a posteriori (MAP) method [13, 5, 6, 10, 4, 14], and inversion problem with regularization [7, 6]. Recently, researchers proposed using wavelet technique or preconditioned conjugate gradient (PCG) method to restore the high resolution image [3, 2]. These techniques are very powerful but they seem

to be sensitive to boundary conditions. Chan *et al.*[1, 2] discussed the importance of the boundary condition in the regularization problem. They mention that ringing effect will appear at the boundary of the image using zero boundary condition. Reflective boundary condition (Neumann boundary condition) gives about 1.0dB PSNR better result than zero boundary condition and periodic boundary condition for test images. If the assumption is not suitable for the image, the error propagates in each iteration and gives a poor result. In the mathematical model, the matrix representation for the shift, blur and down-sampling operator have to be changed for different types of boundary condition. There is a parameter that controls the model regularity. If this parameter is overestimated, the solution is still blurred. Moreover, the solution is ill - conditioned. This means that choosing an appropriate parameter directly affects the quality of solution. All the problems mentioned above thus make regularization a difficult but key issue in super-resolution reconstruction. In section 2, we will propose a simple and effective method for solving HRIR problem that avoid the regularization problem and the boundary condition problem.

2. High Resolution Image Reconstruction

In HRIR, we assume the low resolution images have subpixel displacements between each other. This special relationship provides enough information for restoring the high resolution image. Suppose we have q^2 $M \times M$ $I_{l_1 l_2}$ low resolution images where q is the magnification factor, l_1 and l_2 are positive integers smaller than and equal to $q-1$ and they represent the low resolution image I which has a l_1 vertical downward and l_2 horizontal rightward shift. For $q=2$, the 4 low resolution images are $I_{00}, I_{01}, I_{10}, I_{11}$. The pixels in the low resolution image are the average of $h_{l_1 l_2} \times w_{l_1 l_2}$ pixels in the high resolution image. $h_{l_1 l_2}$, $w_{l_1 l_2}$ and q can be different values for different low resolution images. For simplicity, they are set to a certain value. We will propose a general mathematical model which is suitable for different values of $h_{l_1 l_2}$, $w_{l_1 l_2}$ and q in future. The $qM \times qM$ high resolution image defined as f. Figure 1 illustrates the pixel relationship between low resolution image I_{00}, I_{11} and high resolution image f with $h_{00} = w_{00} = h_{11} = w_{11} = 2$. The relationship of the pixel between the low resolution image and high resolution image is given by

$$(I_{l_1 l_2})ij = \sum_{\substack{x \in X_{l_1} \\ y \in Y_{l_2}}} f(x,y)/(h_{l_1 l_2} w_{l_1 l_2}), \tag{2.1}$$

where i and j are integers with $1 \le i, j \le M$, $X_{l_1} = \{2i - 1 + k + l_1 | k = 0, \ldots, h_{l_1 l_2} - 1\}$ and $Y_{l_2} = \{2j - 1 + k + l_2 | k = 0, \ldots, w_{l_1 l_2} - 1\}$. In our experiment, we set the magnification factor q equal to 2. Thus, we assume all the low resolution images have the same magnification factor and the pixels in the low resolution image are the average of 2×2 pixels in the high resolution image.

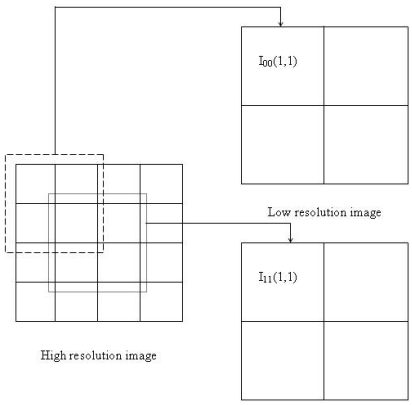


FIGURE 1. The relationship between the high resolution image and low resolution images

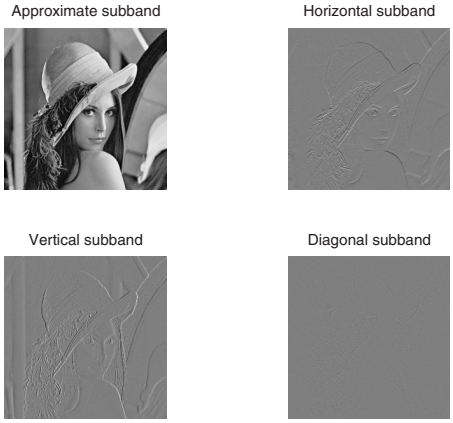


FIGURE 2. The Haar wavelet transform of Lena image

Haar wavelet is the well-known and simple example of an orthonormal wavelet. For simplicity, we use the 1 level wavelet decomposition. The $qM \times qM$ high resolution image after applying the Haar wavelet transform consists of 4 subbands: approximate subband (LL), horizontal subband (H), vertical subband (V) and diagonal subband (D). These subbands are the half size of the high resolution image. Approximate subband is down-sampling of the high resolution image. The pixel relationship between the approximate subband and the high resolution image is given by

$$(LL)ij = \sum_{\substack{x \in X_{LL} \\ y \in Y_{LL}}} f(x,y)/2, \tag{2.2}$$

where $X_{LL} = \{2i - 1 + k | k = 0, 1\}$ and $Y_{LL} = \{2j - 1 + k | k = 0, 1\}$. Horizontal subband contains the horizontal edges and details of the high resolution image. The horizontal detail coefficients are defined as:

$$(H)ij = \frac{1}{2}\left(\sum_{\substack{x \in X_{Ha} \\ y \in Y_{Ha}}} f(x,y) - \sum_{\substack{x \in X_{Hb} \\ y \in Y_{Hb}}} f(x,y) \right), \tag{2.3}$$

where $X_{Ha} = \{2i - 1\}$, $Y_{Ha} = \{2j - 1 + k | k = 0, 1\}$, $X_{Hb} = \{2i\}$ and $Y_{Hb} = \{2j - 1 + k | k = 0, 1\}$. The vertical and diagonal detail coefficients are defined similarly. Each subband have $M \times M$ coefficients. Figure 2 shows the image Lena after applying the Haar wavelet transformation. The top left image is the approximate subband which is the low resolution image of the original image. The top right image displays the horizontal edges, the bottom left image shows the vertical edges. The last image illustrates the diagonal edge. We found that there are connections between the set of low resolution images and these subbands. We use the pixel relationship between the low resolution images with the help of Taylor series to approximate those subbands. Since the coefficients of the horizontal, vertical and diagonal subbands are the pixels difference in the high resolution image and the pixels in the low resolution images are the averages of the pixels in the high resolution image, we take advantages of this dependence to interpolate those subbands. A good approximation of the subbands means that we can reconstruct the high resolution image well.

2.1. First Order Approximation

In last section we mentioned that we use the Taylor series to interpolate the wavelet coefficient subbands. Now we give more details on this algorithm. Recall equations (2.1) and (2.2), we can formulate the following relationship:

$$(I_{l_1 l_2})ij = \sum_{\substack{x \in X_{l_1} \\ y \in Y_{l_2}}} f(x,y)/(h_{l_1 l_2} w_{l_1 l_2}).$$

Putting $l_1 = l_2 = 0$ and $h_{00} = w_{00} = 2$. We obtain

$$(I_{00})ij = \sum_{\substack{x \in X_0 \\ y \in Y_0}} f(x,y)/4,$$

where $X_0 = \{2i - 1 + k | k = 0, 1\}$ and $Y_0 = \{2j - 1 + k | k = 0, 1\}$. Then the relationship between I_{00} and LL is

$$2(I_{00})ij = (LL)ij.$$

The approximate subband can be interpolated by low resolution image I_{00} exactly. To approximate the horizontal subband, we apply first order Taylor series expansion for the high resolution image:

$$f(x+\triangle x, y+\triangle y) \approx f(x,y) + \triangle x f_x(x,y) + \triangle y f_y(x,y), \tag{2.4}$$

where $f_x(x,y)$ and $f_y(x,y)$ are the first order derivatives. As an example, consider the coefficient of the horizontal subband $(H)ij$. From (2.3) we have

$$(H)ij = \frac{f(2i-1,2j-1) - f(2i,2j-1)}{2} + \frac{f(2i-1,2j) - f(2i,2j)}{2}. \tag{2.5}$$

Let $x = 2i-1$ and $y = 2j-1$, the coefficient of the horizontal subband become

$$(H)ij = \frac{f(x,y) - f(x+1,y)}{2} + \frac{f(x,y+1) - f(x+1,y+1)}{2}.$$

Using (2.4), we have

$$\begin{aligned} f(x,y) - f(x+1,y) &\approx f(x,y) - (f(x,y) + f_x(x,y)) \\ &= -f_x(x,y), \end{aligned}$$

and similarly when $\triangle x = 2$ we obtain

$$f(x,y) - f(x+2,y) \approx -2f_x(x,y)$$

which implies

$$f(x,y) - f(x+1,y) \approx \frac{f(x,y) - f(x+2,y)}{2}$$

or equivalently

$$f(2i-1,2j-1) - f(2i,2j-1) \approx \frac{f(2i-1,2j-1) - f(2i+1,2j-1)}{2}. \tag{2.6}$$

In the same way,

$$f(2i-1,2j) - f(2i,2j) \approx \frac{f(2i-1,2j) - f(2i+1,2j)}{2}. \tag{2.7}$$

Substituting (2.6) and (2.7) into (2.5), we have

$$\begin{aligned} (H)ij &\approx \frac{1}{2}\left(\frac{f(2i-1,2j-1) - f(2i+1,2j-1)}{2} + \frac{f(2i-1,2j) - f(2i+1,2j)}{2}\right) \\ &= \frac{f(2i-1,2j-1) - f(2i+1,2j-1)}{4} + \frac{f(2i-1,2j) - f(2i+1,2j)}{4} \\ &= (I_{00})ij - (I_{10})ij. \end{aligned}$$

We can obtain a similar relationship between the vertical subband and the low resolution images I_{00}, I_{01}. For the diagonal subband,

$$(D)ij = \frac{f(2i-1,2j-1) + f(2i,2j) - (f(2i,2j-1) + f(2i-1,2j))}{2}.$$

Putting $x = 2i-1$, $y = 2j-1$ and applying equation (2.4), we have

$$\begin{aligned} f(x,y) + f(x+1,y+1) &\approx f(x,y) + (f(x,y) + f_x(x,y) + f_y(x,y)), \\ f(x+1,y) + f(x,y+1) &\approx f(x,y) + (f(x,y) + f_x(x,y) + f_y(x,y)), \end{aligned}$$

which implies

$$f(2i-1,2j-1)+f(2i,2j-1)=f(2i,2j-1)+f(2i,2j),$$

and

$$(D)ij \approx 0.$$

Therefore, the first order approximation of the Haar wavelet subbands are given by:

$$\begin{aligned}(LL)ij &= 2(I_{00})ij,\\ (H)ij &\approx (I_{00})ij-(I_{10})ij,\\ (V)ij &\approx (I_{00})ij-(I_{01})ij,\\ (D)ij &\approx 0.\end{aligned}$$

The high resolution image can be obtained by taking inverse Haar wavelet transform on these subbands. This result shows our method no need to use I_{11}, hence we only need to use 3 low resolution images for reconstruction. The low resolution image I_{11} is difficult to obtain, since the camera hard to has a precise subpixel shift in both horizontal and vertical direction. The other low resolution images can be taken precisely. Therefore, our method need less data and easy to implement. Also, the first order approximation does not need boundary condition to provide extra information. This implies our method in first order approximation does not suffered any boundary condition problems.

2.2. Second Order Approximation

In section 2.1, we propose the first order approximation for high resolution image wavelet subbands. If the pixels in the high resolution image change linearly, the first order approximation gives a good reconstruction. But most of the changes are not linear; they may be occurred in quadratic or cubic form. When the changes are not linear, we use second or higher order Taylor series expansion to approximate the subbands. We are concentrate on interpolate the subbands except the approximate subband as the approximate subband can be interpolated by the low resolution image I_{00} exactly. The second order Taylor series is

$$\begin{aligned}f(x+\triangle x,y+\triangle y) \approx\ & f(x,y)+\triangle x f_x(x,y)+\triangle y f_y(x,y)+\frac{1}{2}((\triangle x)^2 f_{xx}(x,y)\\ & +2(\triangle x)(\triangle y)f_{xy}(x,y)+(\triangle y)^2 f_{yy}(x,y)).\end{aligned} \tag{2.8}$$

Recall the equation (2.5), we have

$$(H)ij=\frac{f(2i-1,2j-1)-f(2i,2j-1)}{2}+\frac{f(2i-1,2j)-f(2i,2j)}{2}.$$

Using the similar technique in section 2.1 we can show that

$$\begin{aligned}(H)ij &\approx \frac{f(2i-1,2j-1)-f(2i+1,2j-1)}{4}+\frac{f(2i-1,2j)-f(2i+1,2j)}{4}\\ &\quad +\frac{1}{4}(f_{xx}(2i-1,2j-1)+f_{xx}(2i-1,2j))\\ &= \frac{3}{2}((I_{00})ij-(I_{10})ij)-((I_{10})ij-(I_{00})i+1,j).\end{aligned}$$

The vertical subband can be interpolated alike. To interpolate the diagonal subband we define the $P_s(x,y)$ as,

$$P_s(x,y) = f(x,y)+f(x+s,y+s)-f(x+s,y)-f(x,y+s).$$

Then,

$$(D)ij = \frac{P_1(2i-1,2j-1)}{2}.$$

By using equation (2.8)

$$\begin{aligned}(D)ij &\approx \frac{P_2(2i-1,2j-1)}{8}\\ &= \frac{(I_{00})ij+(I_{11})ij-(I_{10})ij-(I_{01})ij}{2}.\end{aligned}$$

Then the second order approximation are:

$$\begin{aligned}(LL)ij &= 2(I_{00})ij,\\ (H)ij &\approx \frac{3}{2}((I_{00})ij-(I_{10})ij)-\frac{1}{2}((I_{10})ij-(I_{00})i+1,j),\\ (V)ij &\approx \frac{3}{2}((I_{00})ij-(I_{01})ij)-\frac{1}{2}((I_{01})ij-(I_{00})i,j+1),\\ (D)ij &\approx \frac{(I_{00})ij+(I_{11})ij-(I_{10})ij-(I_{01})ij}{2}.\end{aligned}$$

In general, the higher order approximation can give a better performance than the lower order approximation for low-noise image. But if the low resolution images are noisy, there could be noise amplification effects. Since the interpolation of the diagonal subband involved all the low resolution images, we believe that the second order approximation of the diagonal subband is more susceptible noise amplification with noisy low resolution images than first order approximation. Suppose all the images contain same amount of noise, the diagonal subband has the largest noise amplification. The other subbands would not have such large noise amplification, as they are only approximated by two noisy low resolution images. The noisy low resolution images enlarge the error of approximation, the reconstructed image also suffer from noise and gives a poor quality. Thus, we propose to use the first order approximation to interpolate the diagonal subband for higher order approximation. This idea provides two advantages. First, we use the first order approximation for the diagonal subband that means we only need 3 low resolution

images to obtain our reconstruction. Compare with other reconstruction methods, our method needs less data to reconstruct a high resolution image. Second, we reduce the noise amplification during the reconstruction. Since the first order approximation of the diagonal subband does not include any images, which is independent of noise. In the second order approximation, the approximated coefficients in horizontal and vertical subbands require boundary condition to provide extra information during reconstruction especially on the last row or column of the wavelet subband. Thus, we use 3 different boundary conditions (Periodic boundary condition, Reflective boundary condition and Zero boundary condition) in experiments which show our method in second order approximation is robust to the boundary condition. Also, numerical results show our method gives a better performance than the traditional Least-squares method.

3. Numerical Results

In this section, we illustrate the performance of the first order and second order approximation for high resolution image reconstruction. Our method is not a iterative method, it is direct and fast. All the results can be generated in about 1 to 2 seconds in matlab with Pentium-4 CPU 3.2GHz computer. To study the performance of the reconstruction method, we use the peak signal - to - noise ratio (PSNR) to compare the reconstructed image $\hat{f}$ with the original image f. The PSNR is defined as

$$\text{PSNR} = 10 \log_{10} \frac{255^2 NM}{\|f - \hat{f}\|_2^2}$$

where the size of image f is $N \times M$.

We use the Lena image of size 512×512 as the original image in our experiments. To obtain the set of low resolution images. We first add Gaussian white noise to the original image. The four 254×254 low resolution images can be obtained by down-sampling with factor of 2 with original boundary. Then we compare our method with Least-squares model (Tikhonov method) [2]: Sec:2. The results of the Least - squares model quote from Chan *et al.* [2]. Note that our imaging model is slightly different from Chan *et al.*'s, our reconstructed image differs from their reconstructed image by a half pixel shift in both horizontal and vertical direction so a direct comparison is not technically fair. However, the difference is small and we believe that the qualitative conclusions that we will draw from the results are still meaningful. In practice, the set of the low resolution images which have the precise subpixel shift in horizontal and vertical direction are difficult to obtain especially the low resolution image I_{11}. In our method, only the approximation of diagonal subband requires low resolution image I_{11}. Therefore, the second order approximation of the diagonal subband is replaced by the first order approximation. Now we only need 3 low resolution images for the reconstruction.

Table 1 shows the results of reconstructed Lena image with different noise level and boundary conditions. The second and third row shows the results of Lena

TABLE 1. The results for the first order approximation by using 3 low resolution images

SNR	Boundary Condition	Least-Squares model		1st Order Approximation
		PSNR	β^*	
40	Periodic	33.89	0.0197	35.82
	Reflective	35.89	0.0134	
	Zero	N/A	N/A	
30	Periodic	33.16	0.0216	34.54
	Reflective	34.51	0.0175	
	Zero	N/A	N/A	

image with noise level at SNR = 40dB and 30dB respectively. The results of zero boundary condition are not available, because Chan *et al.* claim that the reconstructed image would be suffered ringing effect when the zero boundary condition are used, and the reconstruction result is worst. In second row and third column, we compare the results of Least - square method among the boundary conditions. We can see that the results of periodic boundary condition has about 2.0dB lower than that of reflective boundary condition. For the first order approximation, the reconstruction do not require any boundary conditions. Thus the result has the unique value 35.82dB which is better than the Least - square method in periodic and zero boundary condition. The first order approximation's result is close to the Least - square method in reflective boundary condition. We only use 3 low resolution images then the result is comparable to the Least - square method by using 4 low resolution images with optimum parameter. For SNR = 30dB, our method works better than Least - square method among all boundary conditions and Least - square method still sensitive to the boundary conditions. There are about 1.5dB difference between periodic and reflective boundary condition. Therefore, first order approximation by using 3 low resolution image is comparable to the Least - square method by using 4 low resolution image with optimum parameter and it is robust to the boundary condition.

TABLE 2. The results for the second order approximation by using 3 low resolution images

SNR	Boundary Condition	Least-Squares model		2nd Order Approximation
		PSNR	β^*	
40	Periodic	33.89	0.0197	37.29
	Reflective	35.89	0.0134	37.27
	Zero	N/A	N/A	37.29
30	Periodic	33.16	0.0216	34.94
	Reflective	34.51	0.0175	34.93
	Zero	N/A	N/A	34.95

Table 2 illustrate the second order approximation results with different boundary conditions. The results of Least - square method are same as table 1. Compare table 1 and 2, we can find that there are about 1.4dB and 0.5dB improvement from second order to first order approximation at SNR = 40dB and 30dB respectively. In the column 4, we can see that the second order approximation obtain very close results among the boundary conditions in SNR = 40dB and 30dB. The difference between the boundary conditions are only 0.02dB. This shows our method is robust to boundary condition. Also, table 2 also shows the second order approximation is superior to Least - square method.

4. Conclusion and Further Work

From the experiment results, our method only uses three low resolution images which give a better performance than the traditional Least-square method by using 4 low resolution images with optimum parameter. One of the most important observation is that our method can work well with zero boundary condition thus avoiding the justify which boundary condition to use. Our method works extremely well in the low-noise case. In the noisy image case, it is only slightly better than the Least-squares model. However, we only need 3 low resolution images. We shall integrate wavelet denoising technique into our method to improve the results for noisy images. To sum up, our method can work with just 3 low resolution images and is robust to the boundary condition. Since the I_{11} image is very difficult to obtain accurately, our method do not need this image also give a good result. It is a direct, fast and easy to implement method. Our method can work with wider class of images as it can work with zero boundary condition. Also, our method has a far superior results in noise case. In the future, we will concentrate on remove noise with the reconstructed image and attempt other interpolate method to approximate the wavelet subbands.

Acknowledgment

This research was partially funded by RGC research grant HKBU2021/04P and a Faculty Research Grant from the Hong Kong Baptist University. The authors would also like to thank Prof. Tony Chan for his distinguished lecture series talk on ENO - Haar wavelet in HKBU on 15 April 2005 which inspired our approach, and to Prof. Michael Ng for his valuable comments and discussion.

References

[1] R. Chan, T. Chan, M. Ng, W. Tang, and C. Wong. Preconditioned iterative methods for high-resolution image reconstruction from multisensors. In Franklin Luk, editor, *Proceedings to the SPIE Symposium on Advanced Signal Processing: Algorithms, Architectures, and Implementations*, volume 3461, pages 348–357, 1998.

[2] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen. Wavelet algorithms for high-resolution image reconstruction. *SIAM Journal on Scientific Computing*, 24:1408–1432, 2003.

[3] R. H. Chan and M. K. Ng. Conjugate gradient methods for toeplitz systems. *SIAM Review*, 38(3):427–482, 1996.

[4] T. J. Connolly and R. G. Lane. Gradient methods for superresolution. In *Proceedings of the International Conference on Image Processing*, volume 1, pages 917–920, 26-29 Oct. 1997.

[5] R. C. Hardie, K. J. Barnard, and E. E. Armstrong. Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. on Image Processing*, 6(12):1621–1633, 1997.

[6] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson. High resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system. *Optical Engineering*, 37:247–260, 1998.

[7] M. C. Hong, M. G. Kang, and A. K. Katsaggelos. An iterative weighted regularized algorithm for improving the resolution of video sequences. *Proc. 1997 Int. Conf. Image Processing*, 2:474–477, 1997.

[8] E. Kaltenbacher and R. C. Hardie. High resolution infrared image reconstruction using multiple, low resolution, aliased frames. In OH Dayton, editor, *Proceedings of the IEEE National Aerospace Electronics Conference (NAECON)*, volume 2, pages 702–709, 1996.

[9] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38:1013–1027, 1990.

[10] N. Nguyen, P. Milanfar, and G. Golub. A computationally efficient superresolution image reconstruction algorithm. *IEEE Trans. on Image Processing*, 10(4):573–583, 2001.

[11] P. Oskoui-Fard and H. Stark. Tomographic image reconstruction using the theory of convex projections. *IEEE Transactions on Medical Imaging*, 7(1): 45–58, 1988.

[12] A. J. Patti and Y. Altunbasak. Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Trans. Image Processing*, 10(1):179–186, 2001.

[13] R. R. Schultz and R. L. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 6:996–1011, 1996.

[14] J. H. Shin, J. S. Yoon, J. K. Paik, and M. Abidi. Fast superresolution for image sequence using motion adaptive relaxation parameters. In *Proceedings of the 1999 International Conference on Image Processing*, volume 3, pages 676–680, 1999.

[15] R. Y. Tsai and T. S. Huang. Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing*, 1:317–339, 1984.

[16] D. Zhang, H. Li, and M. Du. Fast map-based multiframe super-resolution image reconstruction. *Image and Vision Computing*, 23:671–679, 2005.

C.S. Tong
Centre for Mathematical Imaging and Vision
Hong Kong Baptist University
Department of Mathematics
Kowloon Tong, HK
e-mail: cstong@hkbu.edu.hk

K.T. Leung
Centre for Mathematical Imaging and Vision
Hong Kong Baptist University
Department of Mathematics
Kowloon Tong, HK
e-mail: ktleung@math.hkbu.edu.hk

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 431–441
© 2006 Birkhäuser Verlag Basel/Switzerland

The Design of Hilbert Transform Pairs in Dual-Tree Complex Wavelet Transform

Fengxia Yan, Lizhi Cheng and Hongxia Wang

Abstract. An approach for designing biorthogonal dual-tree complex wavelet transform filters is proposed, where the two related wavelets form an approximate Hilbert transform pair. Different from the existing design techniques, the two wavelet filter banks obtained here are both of linear phase. By adjusting the parameters, wavelet filters with rational coefficients may be achieved. The designed examples show that the lengths of wavelet filters may be effectively shorted while efficient approximation to Hilbert transform pairs is still kept. The validity of the proposed design scheme is exhibited through an application to dual-tree complex wavelet for iris image enhancement.

Keywords. Hilbert transform pairs, biorthogonal wavelet, linear phase, dual-tree complex wavelet, iris image enhancement.

1. Introduction

The real Discrete Wavelet Transform (DWT) is a powerful tool for signal and image processing. It, however, has some disadvantages, including, (1) It is shift-sensitive because the input signal shift generates unpredictable changes in DWT coefficients; (2) It suffers from poor directionality because DWT coefficients reveal merely three spatial orientations; (3) It lacks of the phase information that accurately describes non-stationary signal behavior; that undermine its usage in many applications. Many researchers have proposed different techniques to overcome simultaneously some or all of these drawbacks [1-5]. In recent years, some researchers proposed to design two wavelet bases to work together in order to gain some specific effects.

In particular, Kingsbury has demonstrated that dramatic improvements can be achieved in wavelet-based signal processing by utilizing a pair of wavelet transforms where the wavelets form a Hilbert transform pair. Kingsbury calls a wavelet

This work was supported by the National Science Foundation of China (NSFC) under Grant 60573027.

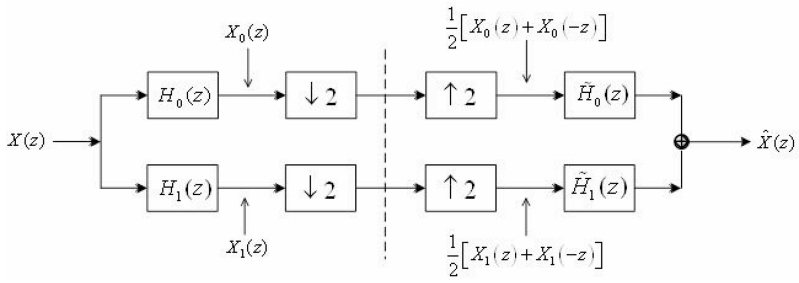


FIGURE 1. Biorthogonal wavelet filter banks

transform of this type a dual-tree complex wavelet transform (DT-CWT) [5-6], which is a form of discrete wavelet transform which generates complex coefficients by using a dual tree of wavelet filters to obtain their real and imaginary parts. The DT-CWT, which is based on (approximate) Hilbert pairs of wavelets, has many advantages such as shift-invariance, directional selectivity, perfect reconstruction, limited redundancy and offering phase information. The obtained phase information is associated with quasi-analytic wavelet decomposition and has been exploited for various image-processing applications [7-8]. Selesnick formulated the design problem in [9] for the minimal length scaling filters such that 1) the wavelets each has a specified number of vanishing moments, and 2) the half-sample delay approximation is flat at $\omega = 0$ with specified degree. However, his formulation leads to nonlinear design equations, and the filters would have to be obtained using Gröbner bases. In [10-11], he describes the other two design procedures based on spectral factorization. But the wavelet filters designed by these methods are of long length and nonlinear phase.

In this paper, we present a new algorithm for designing two biorthogonal wavelets forming a Hilbert transform pair approximately. Wavelet filter banks obtained here are all of linear phase, i.e., their coefficients are all symmetric. Results of the designed examples show that our algorithm can shorten the length of the filters efficiently while still keep the satisfied approximation to Hilbert transform pairs. Furthermore, we use the designed filter banks in dual-tree complex wavelet transform for iris image enhancement. Results of experiments show that our design method is effective for reducing computational complexity and improving performance.

2. The Design of Linear Phase Biorthogonal Wavelet Filters

2.1. Biorthogonal Wavelet Filter Banks

Fig. 1 is the filter scheme of one level biorthogonal wavelet decomposition and reconstruction. Multirate filter analysis shows that

$$\begin{aligned}\hat{X}(z) &= \frac{1}{2}[X_0(z)+X_0(-z)]\tilde{H}_0(z)+\frac{1}{2}[X_1(z)+X_1(-z)]\tilde{H}_1(z)\\ &= \frac{1}{2}[H_0(z)\tilde{H}_0(z)+H_1(z)\tilde{H}_1(z)]X(z)\\ &+ \frac{1}{2}[H_0(-z)\tilde{H}_0(z)+H_1(-z)\tilde{H}_1(z)]X(-z).\end{aligned} \tag{2.1}$$

The first perfect reconstruction (PR) condition requires aliasing cancellation and forces the above term in $X(-z)$ to be zero. Hence $H_0(-z)\tilde{H}_0(z)+H_1(-z)\tilde{H}_1(z)=0$, this can be achieved if

$$H_1(z)=z^{-k}\tilde{H}_0(-z),\tilde{H}_1(z)=z^kH_0(-z) \tag{2.2}$$

where k must be odd (usually $k=\pm 1$). The second PR condition is that the transfer function from $X(z)$ to $\hat{X}(z)$ should be unity; i.e. $H_0(z)\tilde{H}_0(z)+H_1(z)\tilde{H}_1(z)=2$. If we define a product filter $P(z)=H_0(z)\tilde{H}_0(z)$, then the PR condition becomes

$$P(z)+P(-z)=2. \tag{2.3}$$

Equivalently,

$$\sum_n h_0(n)\tilde{h}_0(n+2k)=\delta(k). \tag{2.4}$$

Besides, the low pass filers should satisfy the following normalization conditions

$$\sum_n h_0(n)=\sqrt{2}\,,\quad \sum_n \tilde{h}_0(n)=\sqrt{2}. \tag{2.5}$$

Considering the filter banks (FB) in Fig. 1 are not only PR FB but also wavelet systems, let N and $\tilde{N}$ denote the number of vanishing moments of the wavelet and its dual, respectively, we have the following equations

$$\begin{aligned}\sum_n (-1)^n n^k h_0(n) &= 0,\quad k=0,1,...,N-1\\ \sum_n (-1)^n n^k \tilde{h}_0(n) &= 0,\quad k=0,1,...,\tilde{N}-1.\end{aligned} \tag{2.6}$$

2.2. The Design of Linear Phase Biorthogonal Wavelet Filters with a Parameter

Assuming the biorthogonal wavelet filter banks are linear phase, and let L and $\tilde{L}$ be the length of $H_0(z)$ and $\tilde{H}_0(z)$, and N, $\tilde{N}$ be the vanishing moments. Then, one can represent $H_0(z)$ and $\tilde{H}_0(z)$ in the form of

$$H_0(z)=Q(z)(1+z^{-1})^N,\tilde{H}_0(z)=\tilde{Q}(z)(1+z^{-1})^{\tilde{N}}. \tag{2.7}$$

When felicitously select the length of the scale filters and the vanishing moments, we can obtain a group of biorthogonal PR filters with a parameter by solving the equations (2.4)-(2.6). As the technique in [12], in order to achieve biorthogonal wavelet filters, using equation (2.7), we employ Daubechies's Theorem [13, 14] to determine the interval which the parameter can be included. The following examples give the coefficient lists with a parameter and the intervals which the corresponding parameters can be included.

Example 1: with $L = 7,\ \tilde{L} = 9$ and $N = 2,\ \tilde{N} = 4$

$$\begin{aligned} h_0 &= \{-\frac{t}{16}, \frac{1-2t}{16}, \frac{t+4}{16}, \frac{3+2t}{8}, \frac{t+4}{16}, \frac{1-2t}{16}, -\frac{t}{16}\} \\ \tilde{h}_0 &= \{\frac{8t^3-6t^2+3t}{32(1+2t)}, \frac{-16t^3+20t^2-12t+3}{32(1+2t)}, \frac{-8t^3+6t^2+5t+20}{16(1+2t)}, \\ &\quad \frac{2t-3}{8(1+2t)}, \frac{16t^3-20t^2+28t+5}{32(1+2t)}, \frac{2t-3}{8(1+2t)}, \cdots\} \end{aligned}$$

The parameter $t \in [0.186, 1.129]$.

Example 2: with $L = 6,\ \tilde{L} = 10$ and $N = 3,\ \tilde{N} = 3$

$$\begin{aligned} g_0 &= \{s, \frac{1}{8}(1+8s), \frac{1}{8}(3-16s), \frac{1}{8}(3-16s), \frac{1}{8}(1+8s), s\} \\ \tilde{g}_0 &= \{\frac{s(3+48s+512s^2)}{8(-1+16s)}, \frac{(1+8s)(3+48s+512s^2)}{64(1-16s)}, \\ &\quad \frac{-1}{64}(9+112s+512s^2), \frac{7+288s+1280s^2+8192s^3)}{64(-1+16s)}, \\ &\quad \frac{45-304s-512s^2}{64(1-16s)}, \frac{7+288s+1280s^2+8192s^3)}{64(-1+16s)}, \cdots\} \end{aligned}$$

The parameter $s \in [-0.186, 0.31]$.

Example 3: with $L = 8,\ \tilde{L} = 8$ and $N = 5,\ \tilde{N} = 1$

$$\begin{aligned} g_0 &= \{s, \frac{96s+1}{32}, \frac{32s+5}{32}, \frac{5(16s-1)}{16}, \frac{5(16s-1)}{16}, \frac{32s+5}{32}, \frac{96s+1}{32}, s\} \\ \tilde{g}_0 &= \{\frac{2s(3+192s+8192s^2)}{64s-1}, \frac{(1+96s)(3+192s+8192s^2)}{16(1-64s)}, \\ &\quad \frac{15+736s+26624s^2+786432s^3}{16(1-64s)}, \\ &\quad \frac{5-40s+1536s+65536s^3}{4(1-64s)}, \frac{5-40s+1536s+65536s^3}{4(1-64s)}, \cdots\} \end{aligned}$$

The parameter $s \in [-0.336, 0.368]$.

Using this technique, we can obtain many other linear phase biorthogonal wavelet filters which include a parameter, such as biorthogonal wavelets filters with $L = 7,\ \tilde{L} = 13$ and $N = 4,\ \tilde{N} = 4$; $L = 11,\ \tilde{L} = 5$ and $N = 4,\ \tilde{N} = 2$; $L = 8,\ \tilde{L} = 4$ and $N = 3,\ \tilde{N} = 1$; etc.

3. The Hilbert Transform Pairs of Biorthogonal Wavelet Bases

3.1. The Hilbert Transform Pairs

Recall the definition of Hilbert transform, $\psi_g(t)$ is the Hilbert transform of $\psi_h(t)$, if

$$\hat{\psi}_g(\omega) = \begin{cases} -j\hat{\psi}_h(\omega), \omega > 0 \\ j\hat{\psi}_h(\omega), \omega < 0 \end{cases} \qquad i.e. \qquad \psi_g(t) = \mathcal{H}\{\psi_h(t)\}. \tag{3.1}$$

Let $\{H_0(z), H_1(z); \tilde{H}_0(z), \tilde{H}_1(z)\}$ and $\{G_0(z), G_1(z); \tilde{G}_0(z), \tilde{G}_1(z)\}$ are two biorthogonal wavelet filter banks. They both satisfy the above conditions (2.1)-(2.6), and the corresponding wavelets functions are $\{\psi_h(t), \tilde{\psi}_h(t)\}$ and $\{\psi_g(t), \tilde{\psi}_g(t)\}$ respectively. If $\psi_g(t) = \mathcal{H}\{\psi_h(t)\}$ and $\tilde{\psi}_g(t) = -\mathcal{H}\{\tilde{\psi}_h(t)\}$, then $\{\psi_h(t), \tilde{\psi}_h(t)\}$ and $\{\psi_g(t), \tilde{\psi}_g(t)\}$ are called as Hilbert transform pairs [15]. Furthermore, if both the scale filters $H_0(z)$ and $G_0(z)$ of the two biorthogonal wavelet filter banks are linear phase, and the corresponding wavelet functions $\psi_g(t)$ is the Hilbert transform of $\psi_h(t)$, then the lengths of $H_0(z)$ and $G_0(z)$ need to be odd and even separately. It fitly accords with the dual-tree complex wavelet transform introduced by Kingsbury. Since the wavelet bases form a Hilbert transform pair, the frequency response of the function $\psi_h(t) + j\psi_g(t)$ must have high attenuation in the region $-\infty < \omega < 0$, i.e. $\psi_h(\omega) + j\psi_g(\omega)$ approximates zero as $\omega < 0$. We consider finding suitable filters by maximizing

$$\frac{\max |O(\omega)|^2}{\int_{-\infty}^{0} |O(\omega)|^2 \, d\omega}, \tag{3.2}$$

where $O(\omega) \triangleq \psi_h(\omega) + j\psi_g(\omega)$.

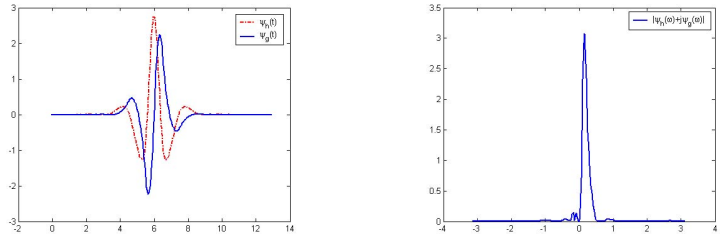
3.2. The Design of Hilbert Transform Pairs of Biorthogonal Wavelet

In section 2.2, we have obtained many groups of biorthogonal linear phase wavelet filters with a parameter. To form Hilbert transform pairs, we let the lengths of $H_0(z)$ and $G_0(z)$, which corresponding to the analysis filters of the two trees of the dual-tree complex wavelet, to be odd and even separately. The optimal parameter set $\alpha^* = (t^*, s^*)$ can be obtained from

$$\alpha^* = \arg\max_{\alpha} \frac{\max |O(\omega)|^2}{\int_{-\infty}^{0} |O(\omega)|^2 \, d\omega}, \tag{3.3}$$

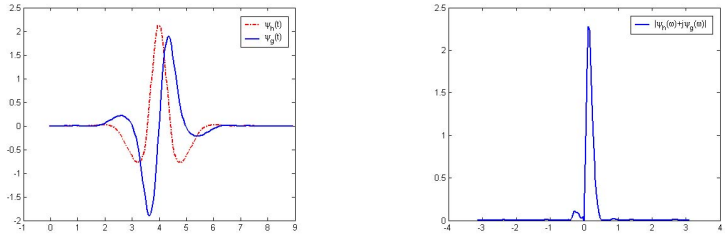
where $\alpha = (t, s)$ denote the parameter set of the biorthogonal wavelet filters to be designed. Once the parameters t and s are known, the scale and wavelet functions $H_0(z)$, $\tilde{H}_0(z)$, $G_0(z)$ and $\tilde{G}_0(z)$ can be easily computed. The wavelet functions are generated by the following equations ($z = e^{j\omega}$ and $\hat{H}(\omega) \hat{=} H(e^{j\omega})$):

$$\begin{cases} \hat{\varphi}_h(\omega) = \frac{1}{\sqrt{2}} \hat{H}_0\left(\frac{\omega}{2}\right) \hat{\varphi}_h\left(\frac{\omega}{2}\right) \\ \hat{\psi}_h(\omega) = \frac{1}{\sqrt{2}} \hat{H}_1\left(\frac{\omega}{2}\right) \hat{\varphi}_h\left(\frac{\omega}{2}\right) \end{cases}, \begin{cases} \hat{\tilde{\varphi}}_h(\omega) = \frac{1}{\sqrt{2}} \hat{\tilde{H}}_0\left(\frac{\omega}{2}\right) \hat{\tilde{\varphi}}_h\left(\frac{\omega}{2}\right) \\ \hat{\tilde{\psi}}_h(\omega) = \frac{1}{\sqrt{2}} \hat{\tilde{H}}_1\left(\frac{\omega}{2}\right) \hat{\tilde{\varphi}}_h\left(\frac{\omega}{2}\right) \end{cases}. \tag{3.4}$$



(a) (b)

FIGURE 2. (a) 13-7 and 10-6 biorthogonal wavelet functions (b)$|\psi_h(\omega) + j\psi_g(\omega)|$

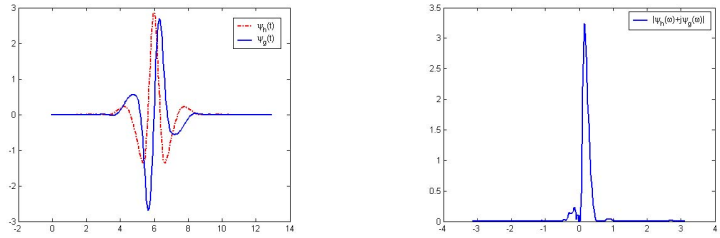


(a) (b)

FIGURE 3. (a) 9-7 and 8-8 biorthogonal wavelet functions (b)$|\psi_h(\omega) + j\psi_g(\omega)|$

According to the above design procedure, we construct three groups of Hilbert transform pairs. The corresponding optimal parameters are obtained by solving (3.3). The values of the optimal parameters (t^*, s^*) and the figures of the corresponding wavelet functions and $|\psi_h(\omega) + j\psi_g(\omega)|$ are given below.

Example 1: 13-7 tap and 10-6 tap linear phase biorthogonal wavelet filters which the corresponding wavelet functions constitute approximate Hilbert transform pairs. $t^* = -0.016$, $s^* = -0.008$ *Example 2:* 9-7 tap and 8-8 tap linear phase biorthogonal wavelet filters which the corresponding wavelet functions constitute approximate Hilbert transform pairs. $t^* = 0.244$, $s^* = -0.020$ *Example 3:* 13-7 tap and 8-8 tap linear phase biorthogonal wavelet filters which the corresponding wavelet functions constitute approximate Hilbert transform pairs. $t^* = -0.025$, $s^* = -0.025$ Especially, the parameters t^* and s^* can be approximated properly by some rational $t^*_{rational}$ and $s^*_{rational}$, so that the coefficients of the two filter banks are all rational numbers. For example, as to *Example 1* , we can choose $t^*_{rational} = -1/64$, $s^*_{rational} = -1/128$. Then the corresponding symmetric rational coefficients of the scaling filters are as follows:



(a) (b)

FIGURE 4. (a) 13-7 and 8-8 biorthogonal wavelet functions (b)$|\psi_h(\omega) + j\psi_g(\omega)|$

$$\begin{aligned}
h_0(n) &= \{-\frac{1}{64}\ \frac{1}{32}\ \frac{17}{64}\ \frac{7}{16}\ \frac{17}{64}\ \frac{1}{32}\ -\frac{1}{64}\}\\
3\times\tilde{h}_0(n) &= \{-\frac{1}{64}-\frac{1}{32}\ \frac{149}{1024}-\frac{105}{1024}-\frac{1267}{2048}\ \frac{443}{512}\ \frac{1259}{512}\ \frac{443}{512}\ \cdots\}\\
g_0(n) &= \{-\frac{1}{128}\ \frac{15}{128}\ \frac{25}{64}\ \frac{25}{64}\ \frac{15}{128}-\frac{1}{128}\}\\
9\times\tilde{g}_0(n) &= \{\frac{85}{4096}\ \frac{1275}{4096}-\frac{2349}{2048}-\frac{1235}{2048}\ \frac{1515}{256}\ \frac{1515}{256}-\frac{1235}{2048}\ \cdots\}. \qquad (3.5)
\end{aligned}$$

As to *Example 2* , we choose $t^*_{rational} = 1/8$, $s^*_{rational} = -1/64$, then the corresponding symmetric rational coefficients of the scaling filters are:

$$\begin{aligned}
h_0(n) &= \{0-\frac{1}{128}\ \frac{3}{64}\ \frac{33}{128}\ \frac{13}{32}\ \frac{33}{128}\ \frac{3}{64}-\frac{1}{128}\ 0\}\\
10\times\tilde{h}_0(n) &= \{\frac{19}{256}\ \frac{57}{128}-\frac{11}{4}\ \frac{263}{128}\ \frac{1325}{128}\ \frac{263}{128}\ \cdots\}\\
g_0(n) &= \{0-\frac{1}{64}-\frac{1}{64}\ \frac{9}{64}\ \frac{25}{64}\ \frac{9}{64}-\frac{1}{64}-\frac{1}{64}\}\\
\tilde{g}_0(n) &= \{\frac{1}{32}-\frac{1}{32}-\frac{7}{32}\ \frac{23}{32}\ \frac{23}{32}-\frac{7}{32}-\frac{1}{32}\ \frac{1}{32}\ 0\}. \qquad (3.6)
\end{aligned}$$

Since all the denominators of the coefficients have the form of 2^k, $k \in Z$, the operations needed for the wavelet transform using these two groups of filters are only additions and shifts. Hence the complex wavelet transform can be computed very quickly. The figures (Fig.2-Fig.4) show that the wavelet functions designed in these examples are quite smooth and short supported, and $|\psi_h(\omega) + j\psi_g(\omega)|$ approximate zero for $\omega < 0$ as expected in each example. Comparing these figures with the designed results in [11], we find that they perform as well as the Example 3 of [11] in approximating a Hilbert transform pair. But the lengths of filters in our examples are much shorter (the filters of Example 3 in [11] are of length 13 and 11) and the corresponding wavelet functions are smoother, furthermore, the coefficients are symmetric, while in [11] are all asymmetric. Since we focus on designing linear phase biorthogonal FIR wavelet filters to form approximate

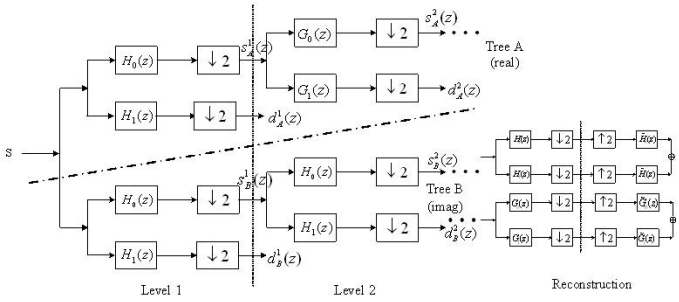


FIGURE 5. 1D dual-tree complex wavelet transform

Hilbert transform pairs, the corresponding complex wavelet transform provides a fast algorithm and a rapid, multiscale decomposition.

4. Application in Dual-Tree Complex Wavelet Transform for Iris Image Enhancement

Reliable automatic recognition of persons based on biometrics has been receiving extensive attention over the past decade. Recently, iris recognition is becoming an active topic in biometric due to its high reliability for personal identification [16-19]. In [19], the iris image preprocessing includes three steps: i) localization; ii) normalization; iii) enhancement. The iris image enhancement is an important step in the iris recognition system for the reasons in [19].

As far as image-processing applications are concerned, the main advantage as compared to the DWT is that the complex wavelet transform can give shift invariance and better directionality. As mentioned in section 3, the wavelet filter banks designed in this paper are of symmetric coefficients, shorter length and can generate smoother wavelets. These properties are quite advantageous to their application in complex wavelet transform for image processing, such as image enhancement. Fig.5 is the scheme of 1D dual-tree complex wavelet transform introduced by Kingsbury in [6], where $\{H_0(z), H_1(z); \tilde{H}_0(z), \tilde{H}_1(z)\}$ and $\{G_0(z), G_1(z); \tilde{G}_0(z), \tilde{G}_1(z)\}$ are two biorthogonal wavelet filter banks. When the two wavelets form a Hilbert transform pair, and two trees are sub-sampled differently, the whole transform is not only PR but also approximately shift invariant.

Using dual-tree complex wavelet transform of the linear phase biorthogonal wavelet pairs of *Example 2* (9-7 and 8-8 biorthogonal wavelet filters in equation (3.6)) in section 3.2, we consider enhancing the iris image by the enhancement algorithm introduced in [15]. Fig.6 represents the preprocessing procedures of the iris image. The methods of localization, normalization and estimation of local average intensity are as in [19]. Fig.6 (d) is subtracted from the normalized image to

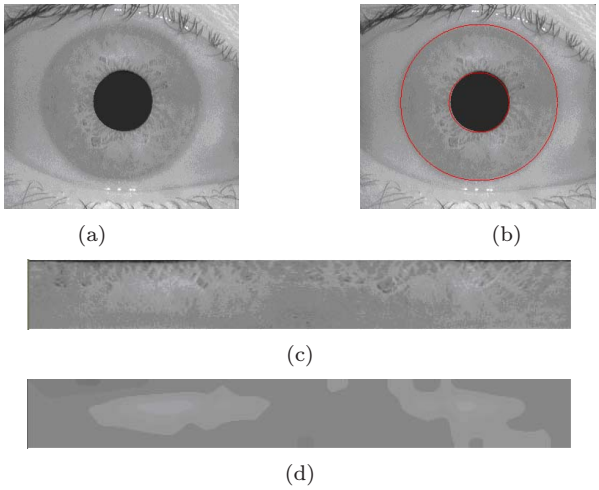
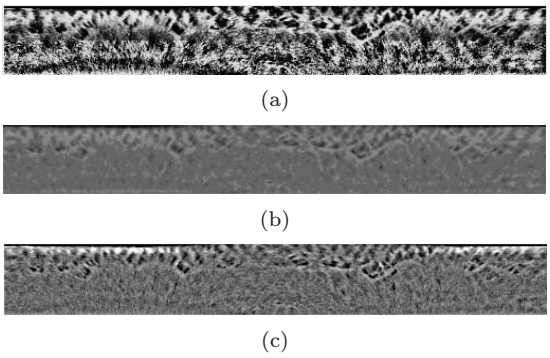


FIGURE 6. Iris image preprocessing: (a) original image; (b) localized image; (c) normalized image; (d) estimated local average intensity

compensate for a variety of lighting conditions. Fig.7 lists the enhanced results of the lighting corrected iris image: (a) is the enhanced image by the local histogram equalization [19]; (b) is the enhance image by DWT [20]; (c) is the enhanced image by our DT-CWT method. We can see that in the iris texture region, the enhanced iris image based on the dual-tree complex wavelet transform shows more clearly details of the randomly distributed and irregular small blocks, which constitute the most distinguishing characteristics of the iris; while in the non-texture region, it is smoother and introduces less distortional information. The improvement of the performance mainly due to the good properties of shift invariance and better directionality of the DT-CWT. Because the coefficients of all filters are symmetric and rational, the computational complexity is remarkably reduced. All these properties will benefit the subsequent processing in feature extraction and matching.

5. Conclusion

This paper presents the design of Hilbert transform pairs formed by two biorthogonal wavelet bases and their applications in DT-CWT for iris image enhancement. In terms of our new designing scheme, we can obtain wavelet filters which are all of linear phases and quite short lengths. Furthermore, the corresponding wavelet functions would be smoother than those designed by the existing algorithms. Especially, wavelet filters with rational coefficients have been obtained by adjusting the parameters, which can effectively reduce the computational complexity in the com-



(a)

(b)

(c)

FIGURE 7. (a) Enhanced image by local histogram equalization; (b) Enhanced image by DWT; (c) Enhanced image by our DT-CWT method

plex wavelet transform. The new designed Hilbert transform pairs of biorthogonal wavelet filters have been used in dual-tree complex wavelet transform to enhance the iris image. Experiment results show that the algorithm based on dual-tree complex wavelet do perform better than those based on real wavelet transform and local histogram equalization.

Acknowledgment

The authors are very grateful to professor Tao Qian and the anonymous reviewers who give us many valuable comments and advice.

References

[1] J. Liang and T. W. Parks, *A translation invariant wavelet representation algorithm with applications.* IEEE Trans. on SP, vol. 44, no. 2, pp. 225-232, Feb. 1996.

[2] E. J. Candes, *Ridglets: theory and applications.* PhD thesis, Stanford University, 1998.

[3] I. W. Selesnick and L. Sendur, *Iterated oversampled filter banks and wavelet frames.* In Wavelet Apllications VII, Proceedings of SPIE, 2000.

[4] M. Do and M. Vetterli, *Pyramidal directional filter banks and curvelets.* In IEEE Proc. Int. Conf. Image Processing, 2001.

[5] N. G. Kingsbury, *The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters.* In Proceedings of the Eighth IEEE DSP Workshop, Utah, August, 1998.

[6] N. G. Kingsbury, *Complex wavelets for shift invariant analysis and filtering of signals.* Applied and Computational Harmonic Analysis, vol. 10, no. 3, pp. 234-253, May 2001.

[7] J. K. Romberg, H. Choi, R. G. Baraniuk and N.G. Kingsbury, *Hidden Markov tree model for complex wavelet transform.* IEEE Transactions on Singal Processing, 2001.

[8] P. de Rivaz and N.G. Kingsbury, *Complex wavelet features for fast texture image retrieval.* In IEEE Proc. Int. Conf. Image Processing, 1999.

[9] I. W. Selesnick, *Hilbert transform pairs of wavelet bases.* IEEE Signal Processing Letters, vol. 8, no. 6, June 2001.

[10] I. W. Selesnick, *The design of Hilbert transform pairs of wavelet bases via the flat delay filter.* In Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), May 2001.

[11] I. W. Selesnick, *The design of approximate Hilbert transform pairs of wavelet bases.* IEEE Signal Processing, vol. 50, pp. 1144-1152, May 2002.

[12] L. Z. Cheng and D.L. Liang, *General construction of 9/7 wavelet filter and its application in image compression.* Optical Engineering, vol. 42, no.8, 2003.

[13] I. Daubechies, *Orthonormal bases of compactly supported wavelets.* Commun. Pure and Applied Maths., vol. 41, no. 2, pp. 909-980, 1988.

[14] A. Cohen, I. Daubechies and J.C. Feauveau, *Biorthogonal bases of compactly supported wavelets.* Commun. Pure and Applied Math, vol. 45, no. 1, pp. 485-560, 1992.

[15] H. X. Wang, *The research of new types of wavelets based on filter banks and their applications.* PhD thesis, National University of Defense Technology, 2004.

[16] J. Daugman, *High confidence visual recognition of persons by a test of statistical independence.* IEEE Trans. Pattern Aanly. Machine Intell., vol. 15, pp. 1148-1161, Nov. 1993.

[17] J. Daugman, *Demoulation by complex-valued wavelets for stochastic pattern recognition.* Int. J. Wavelets, Multi-Res. And Info. Processing, vol. 1, no. 1, pp. 1-17, 2003.

[18] J. Daugman, *How Iris Recognition Works.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, Jan. 2004.

[19] Li Ma, Tieniu Tan, Yunhong Wang and Dexin Zhang, *Efficient iris recognition by characterizing key local variations.* IEEE Transactions on Image Processing, vol. 13, no. 6, pp. 739-749, June 2004.

[20] P. Sakellaropoulos, L. Costaridou and G. Panayiotakis, *A wavelet-based spatially adaptive method for mammographic contrast enhancement.* Physics in Medicine Biology, vol. 48, no. 6, pp. 783-803, 2003.

Fengxia Yan
Department of Mathematics & System, School of Science, National University of Defense Technology, Changsha, 410073, Hunan, P.R. China
e-mail: xialang3@163.com

Lizhi Cheng
Department of Mathematics & System, School of Science, National University of Defense Technology, Changsha, 410073, Hunan, P.R. China
e-mail: lzcheng@hotmail.com

Hongxia Wang
Department of Mathematics & System, School of Science, National University of Defense Technology, Changsha, 410073, Hunan, P.R. China
e-mail: whx8292@hotmail.com

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 443–452
© 2006 Birkhäuser Verlag Basel/Switzerland

Supervised Learning Using Characteristic Generalized Gaussian Density and Its Application to Chinese Materia Medica Identification

S.K. Choy and C.S. Tong

Abstract. This paper presents the estimation of the characteristic generalized Gaussian density (CGGD) given a set of known GGD distributions based on some optimization techniques, and its application to the Chinese Materia Medica identification. The CGGD parameters are estimated by minimizing the distance between the CGGD distribution and known GGD distributions. Our experimental results show that the proposed signature based on the CGGD together with the use of Kullback-Leibler distance outperforms the traditional wavelet-based energy signature. The recognition rate for the proposed method is higher than the energy signature by at least 10% to around 60% - 70%. Nevertheless, the extraction of CGGD estimators still retains comparable level of computational complexity. In general, our proposed method is very competitive compared with many other existing Chinese Materia Medica classification methods.

Mathematics Subject Classification (2000). Primary 68T10; Secondary 65T60.

Keywords. supervised learning, generalized Gaussian density, Kullback-Leibler distance, similarity measurement, wavelets.

1. Introduction

Wavelets have recently emerged as an effective tool to analyze texture information as they provide a multi-resolution and orientation representation of an image by efficient transform. Thus, several well-known texture signatures such as generalized Gaussian density [1-2] signature, energy [3] signature, co-occurrence [4] signature, high order moments and wavelet-domain Hidden Markov Tree models [5-6] based on wavelet subbands are widely used in texture image retrieval.

For the image classification problem, a wide variety of texture feature extraction methods have been proposed based on different techniques. Some features are extracted directly from the gray-scale image and make use of its co-occurrence properties [7], whereas some are extracted by wavelet transform followed by applying statistical measure on the transformed coefficients [3-6].

For the GGD signature, it has been shown that the image retrieval performance is superior to many other existing techniques. However, the GGD signature is seldom used for supervised learning (image classification). The main reason is that a slight change in GGD parameters may lead to a substantial variation in its distribution and hence it is difficult to characterize a benchmark distribution from known GGD distributions. In this paper, we will show how to estimate the characteristic GGD given a series of known GGD distributions based on some optimization techniques. We will also study and compare the recognition performance of the CGGD with the energy signature on Chinese drugs identification.

The outline of this paper is as follows. In the next section, we review several wavelet-based features including the energy signature and the generalized Gaussian density signature. In section 3, we measure the discrepancy between GGD distributions based on Kullback-Leibler distance and define the CGGD as the distribution that minimizes the mean discrepancy to known GGD distributions. In section 4, we conduct an experiment and summarize the classification results of a feasibility study on 10 Chinese Materia Medica. Then, we conclude with some discussions in section 5.

2. Review of Wavelet Signatures

For N levels' discrete wavelet transform, the image is decomposed into $3N$ high-pass subbands and one low-pass subband. Note that all wavelet signatures discussed in this paper are derived from high-pass subbands and thus the low-pass subband will be neglected. The standard order of the $3N$ subbands (hereafter high-pass subband is referred to as subband) is labeled as follows

$$\{S_1, S_2, \ldots, S_{3N}\} = \{NHL, NLH, NHH, \ldots, 1HL, 1LH, 1HH\}, \tag{1}$$

where NHL, NLH and NHH are subbands with horizontal, vertical and diagonal detail coefficients at level N respectively. In this section, we shall only report the wavelet-based energy signature and the GGD signature to characterize each subband.

2.1. Energy Signature

Let S_k be the k^{th} subband of size $X \times Y$and $h_{i,j}$ be its $(i,j)^{th}$ coefficient, then the root-mean-square (RMS) energy of S_k is defined as

$$E_{S_k} = \left(\frac{1}{XY} \sum_{j=1}^{Y} \sum_{i=1}^{X} h_{i,j}^2 \right)^{1/2}. \tag{2}$$

The RMS energy signature $\{E_{S_k}\}_{k=1,2,...,3N}$ represents the distribution of the energy of the subbands. An alternative measure of the energy signature is the mean of the absolute value of the subband and is given by

$$AE_{S_k} = \frac{1}{XY}\sum_{j=1}^{Y}\sum_{i=1}^{X}|h_{i,j}|. \tag{3}$$

In practice, the combination of E_{S_k} and AE_{S_k} are often used to enhance the retrieval rate or as a feature vector for classification. Here, we refer to $\{E_{S_k}\}_{k=1,2,...,3N} \cup \{AE_{S_k}\}_{k=1,2,...,3N}$ as the Extended Energy (EE) signature.

2.2. Modelling of Wavelet Coefficients by Generalized Gaussian Density

Experiments show that for a variety of images, the distribution of the wavelet coefficients of all subbands can be modeled by the GGD, which is defined as

$$p(h;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)}e^{-(|h|/\alpha)^{\beta}}, \tag{4}$$

where $\Gamma(z) = \int_0^{\infty} e^{-t}t^{z-1}dt$, $z > 0$ is the Gamma function. Here, α models the width of the *pdf* (standard deviation) whereas β corresponds to the shape of the GGD. Note that $\beta = 2$ yields the Gaussian density function and $\beta = 1$ yields the Laplacian density function. The GGD parameters α and β can be estimated by Maximum Likelihood Estimation (MLE) [1] or Moment Estimation (ME) [2]. The GGD signature of an image is then referred to as $\{\alpha_{S_k}, \beta_{S_k}\}_{k=1,2,...,3N}$.

3. Similarity Measurement

In pattern recognition, classification of images can be treated as measuring the discrepancy between feature vectors (i.e. input test feature vector and characteristic feature vector). For the EE signature, the simplest measure is to use norm-based metric. For the GGD signature, it is noted that a slight change in GGD parameters may lead to a substantial deviation in its distribution. Thus, it is more reasonable to measure the similarity (or equivalently, discrepancy) between the actual GGD distributions rather than their parameters. Here, we shall adopt the use of Kullback-Leibler distance (KLD) to measure the discrepancy of two distributions p and q and it is given by

$$KLD(p,q) = \sum_i p_i log\frac{p_i}{q_i}, \tag{5}$$

where $log(p_i/q_i)$ is defined to be zero for $p_i = 0$ or $q_i = 0$. Substitute Eq. (4) into Eq. (5) and after some manipulations. We have the following form for the Kullback-Leibler distance between two GGD distributions [1]

$$D\{p_1(\bullet,\alpha_1,\beta_1),p_2(\bullet,\alpha_2,\beta_2)\}=log\left(\frac{\beta_1\alpha_2\Gamma(1/\beta_2)}{\beta_2\alpha_1\Gamma(1/\beta_1)}\right)+\left(\frac{\alpha_1}{\alpha_2}\right)^{\beta_2}\frac{\Gamma((\beta_2+1)/\beta_1)}{\Gamma(1/\beta_1)}-\frac{1}{\beta_2}. \tag{6}$$

Hence, two GGDs are close to each other when $D(p_1,p_2)$ is small and vice versa and the discrepancy between two wavelet subbands can be calculated accurately using only the GGD parameters.

3.1. Estimation of Characteristic GGD

In addition to computing the discrepancy between two GGDs, it is also desirable to estimate the characteristic GGD distribution $p(\bullet,\hat{\alpha},\hat{\beta})$ given L known GGD distributions $\{p(\bullet,\alpha_i,\beta_i)\}_{i=1,2,\ldots,L}$ in which the CGGD best fits the known GGD distributions. Let $\theta_i=(\alpha_i,\beta_i)$ and $J(\theta_i,\theta)=\sum_{i=1}^{L}D\{p(\bullet,\theta_i),p(\bullet,\theta)\}$. Mathematically, it computes

$$\hat{\theta}=arg\min_{\theta}J(\theta_i,\theta), \tag{7}$$

which yields

$$\hat{\alpha}=\left(\frac{\hat{\beta}}{L}\sum_{i=1}^{L}w_i\right)^{1/\hat{\beta}}, \tag{8}$$

$$1+\frac{1}{\hat{\beta}}\left[log\left(\frac{\hat{\beta}\sum_{i=1}^{L}w_i}{L}\right)+\Psi(\frac{1}{\hat{\beta}})\right]-\frac{\sum_{i=1}^{L}\left[w_i log\alpha_i+\frac{w_i}{\beta_i}\Psi(\frac{\hat{\beta}+1}{\beta_i})\right]}{\sum_{i=1}^{L}w_i}=0, \tag{9}$$

where $w_i=\frac{\alpha_i^{\hat{\beta}}\Gamma((\hat{\beta}+1)/\beta_i)}{\Gamma(1/\beta_i)}$, and $\Psi(z)=\frac{\Gamma'(z)}{\Gamma(z)}$ is a di-gamma function. The transcendental equation (9) can be solved numerically using Newton-Raphson or Brent's iterative algorithm. The solution $p(\bullet,\hat{\alpha},\hat{\beta})$ (see Fig. 1) is then the CGGD distribution generated by $\{p(\bullet,\alpha_i,\beta_i)\}_{i=1,2,\ldots,L}$. By Eq. (6), the total distance (TD) between a test image $Q_1=\{\alpha_1^{(k)},\beta_1^{(k)}\}_{k=1,2,\ldots,3N}$ and the CGGD $Q_2=\{\hat{\alpha}_2^{(k)},\hat{\beta}_2^{(k)}\}_{k=1,2,\ldots,3N}$ obtained from the training images is the sum of all the distance D between two corresponding subbands in the same level. i.e.

$$TD\{Q_1,Q_2\}=\sum_{k=1}^{3N}D(p_{Q_1}^{(k)},p_{Q_2}^{(k)}). \tag{10}$$

The CGGD provides an effective characterization of GGD distributions and thus the supervised learning based on CGGD parameters can be implemented very efficiently. In addition, the computational complexity in the similarity measurement and the storage of the signature is vastly reduced compared with using the whole set of wavelet coefficients, which may contain hundreds of parameters.

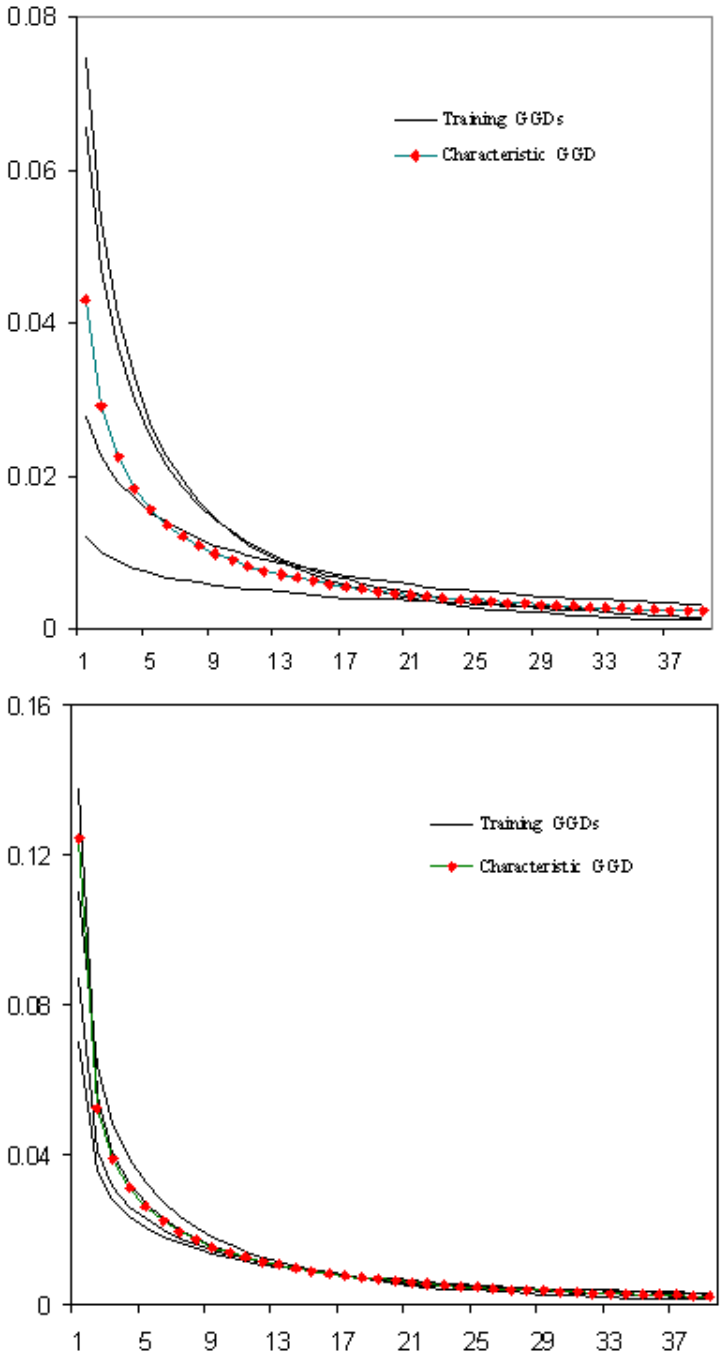


FIGURE 1. Two typical examples of training GGDs and CGGD with L=4.

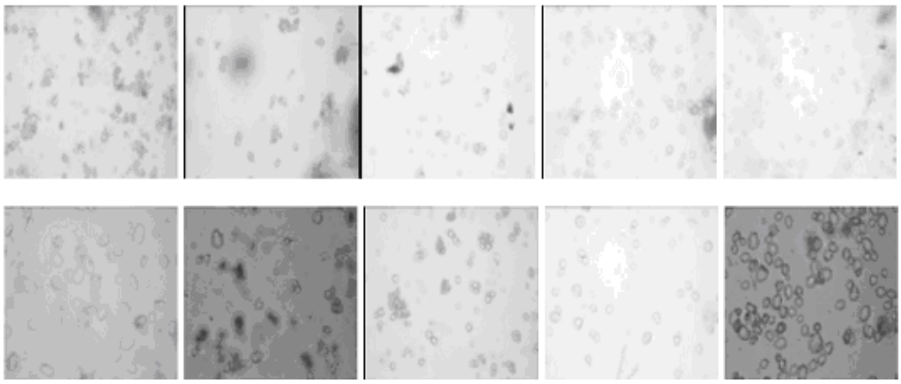


FIGURE 2. Ten different types of gray-scale starch grains images; from left to right, and top to bottom: BaiZhi, BanLanGen, ChuanWu, FenFangJi, GuangFangJi, ShanYao, TianHuaFen, TianNanXing, TuFuLing, ZheBeiMu.

4. Experimental Results and Discussion

4.1. Background of Chinese Materia Medica Identification

Chinese Materia Medica (CMM) is becoming very popular worldwide. But in order to secure its worldwide adoption, efficient yet effective methods for classification of the ingredients of CMM is of paramount importance. The traditional method relies on human senses and is quite subjective. It also requires very long training times and can only be taught in interactive environment. Since misclassification can have serious consequences, more accurate and objective technique is sought.

More recently, a number of research papers and reports in CMM identification using image processing techniques have already been published, and the methods of extracting information for herbs classification had been investigated. One of the classification techniques [8] was based on a powder-form starch grains image (see Fig. 2). A segmented sample of grains in the binarized image is chosen manually and randomly and each sample is further processed to obtain basic features such as area and perimeter for classification. This process, however, is time-consuming and subjective and therefore it is undesirable for herbs classification in real time application.

One of our motivations in this paper is to apply the CGGD signature with a suitable discrepancy measure on the starch grains image for classification and then to build an efficient, objective, and accurate classification system, thus rendering the drug identification process objective and automated, serving as a benchmark standard for CMM identification.

TABLE 1. Recognition rates of wavelet signatures using different measures

Signature+Discrepancy Measure	Recognition rate
EE+L_1	40.00%
CGGD+KLD	64.00%

4.2. Supervised Learning and Recognition Performance

Ten images have been taken for each of 10 different Chinese drugs (see Fig. 2): 5 sample images for each drug are used for training and the other 5 reserved for testing. Note that all signatures in this paper are derived by 2-levels' wavelet decomposition with coiflet 5 as a wavelet function. In the training phase, 10 characteristic EE signatures via Eq. (2) and Eq. (3) are obtained by considering the mean of the training feature vectors and the 10 sets of CGGD signatures $\{\hat{\alpha}_i^j, \hat{\beta}_i^j\}_{i=1,2,\ldots,10;j=1,2,\ldots,6}$ are also obtained via Eq. (8) and Eq. (9).

In the testing phase, we compare the EE signature of an input test sample with each of the 10 characteristic EE signatures using $L_p - norm$. For the GGD signature, we compare the test GGD signature with each of the 10 CGGD signatures by using Eq. (10). That is, if we denote Ω_i as a characteristic image signature of the i^{th} drug and Ω_t as a signature of the input test sample. Then the input test sample is assigned to the i^{th} drug if

$$\overline{d}[\Omega_i, \Omega_t] < \overline{d}[\Omega_j, \Omega_t], \ for\ i \neq j; j = 1, 2, \ldots, 10, \tag{11}$$

where $\overline{d}[\bullet]$ is some distance measures. In this paper, we shall only report the results based on two measures of discrepancy, namely, $L_1 - norm$ for the EE signature and the KLD distance for the CGGD. Table 1 shows the summary of the recognition rates for the two signatures using two measures of discrepancy. Note that using the GGD parameters obtained by optimization method leads to recognition rate of 64%. For the EE signature with $L_1 - norm$, the recognition rate is 40% which is considered unsatisfactory. This indicates that our proposed CGGD is more robust than the EE signature.

Instead of using 5 training samples, it is of interest to know the recognition rate using different number of training samples. Figure 3 plots the recognition rates vs. number of samples used in the training phase. Note that the classification performance for the CGGD+KLD is always better than the EE+L_1 and we can achieve 70% recognition using 7 training samples. In particular, it is observed that the classification rate for the CGGD+KLD increases monotonically with the number of training samples, while there is an unstable sharp change for the EE+L_1 (when no. of training samples = 5). This implies again our proposed CGGD is a robust feature and has a better recognition performance. The only drawback for the CGGD is that the computational complexity is higher than the EE signature as the transcendental Eq. (9) is a highly non-linear equation which needs to be solved numerically.

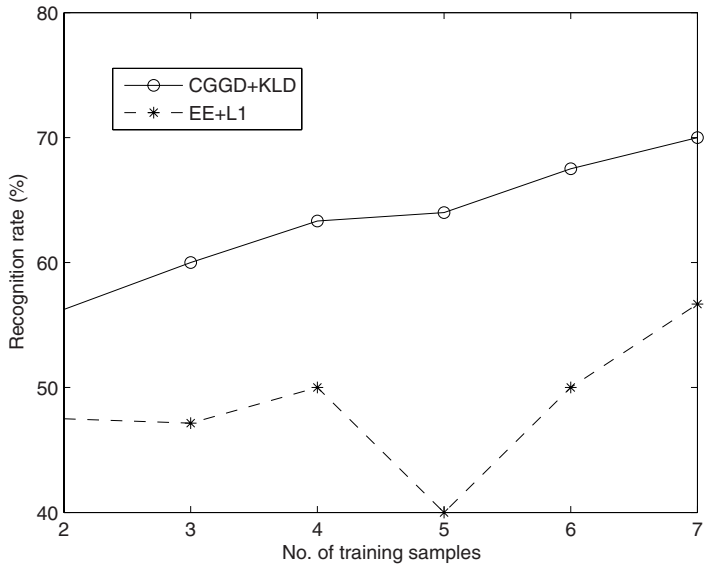


FIGURE 3. Recognition rates vs. no. of training samples

In addition to the top 1 recognition, it is often desirable to ensure that the correct drug should be the top M "closest" to the input test sample. This can be referred to as the cumulative percentage of the input test sample belonging to one of the top M best matches. In the 5-training samples case, we can achieve 80% top M recognition rate at $M = 2$ for the CGGD+KLD, which is considered satisfactory. If we use 7 training samples, the top M recognition rate is over 90% at $M = 2$. This suggests that we can apply the proposed feature in conjunction with other features (e.g. high order moments) to carry out a multi-stage recognition to further improve the classification performance.

5. Conclusions

We have introduced the estimation of the characteristic generalized Gaussian density and made use of it with the Kullback-Leibler distance to identify Chinese Materia Medica. The CGGD parameters are estimated by minimizing a mean distance between a CGGD distribution and a set of known GGD distributions. These CGGD parameters provide an effective characterization of GGD distributions and thus are suitable for supervised learning.

Our experiments show that the CGGD and KLD seem to be an effective signature and discrepancy measure in this small scale study. The recognition performance for the CGGD+KLD outperforms the EE+L_1 by at least 10% which is

considered satisfactory. Moreover, our proposed method is efficient and objective compared with many other existing CMM classification methods in which manual segmentation is necessary.

For a large scale study, we are currently extracting more independent features to further improve the recognition performance such as exploiting the feature from spatial statistics. We are also investigating the use of more sophisticated statistical techniques (e.g. using clustering analysis) and design a statistical multi-stage recognition system for classification.

Acknowledgment

This work is partially supported by Faculty Research Grant from the Hong Kong Baptist University.

References

[1] M. N. Do, M. Vetterli, Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance, *IEEE Trans. Image Processing*, **11**(2), pp. 146-158, 2002.

[2] K. Sharifi and A. Leon-Garcia, Estimation of Shape Parameter for Generalized Gaussian Distribution in Subband Decomposition of Video, *IEEE Trans. Circuits and Systems, Video Tech.*, **5**(1), pp. 52-56, 1995.

[3] M. Unser, Texture Classification and Segmentation Using Wavelet Frames, *IEEE Trans. Image Processing*, **4**(2), pp. 1549-1560, 1995.

[4] A. L. Amet, A. Ertuzun, A. Erqil, Texture Defect Detection Using Subband Domain Co-occurrence Matrices, *Proceedings of IEEE Southeast Symp. for Image Analysis and Interpretation*, pp. 976-979, April 6-7, 1998.

[5] G. L. Fan and X. G. Xia, On Context-Based Bayesian Image Segmentation: Joint Multi-context and Multiscale Approach and Wavelet-Domain Hidden Markov Models, *Proceedings of the* 35^{th} *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov 4-7, 2001.

[6] G. L. Fan and X. G. Xia, Maximum Likelihood Texture Analysis and Classification Using Wavelet-Domain Hidden Markov Models, *Proceedings of the* 34^{th} *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct 29-Nov 1, 2000.

[7] R. F. Walker, P. Jackway, and I. D. Longstaff, Improving Cross-Cooccurrence Matrix Feature Discrimination, *Proceedings of the* 3^{rd} *Conference on Digital Image Computing: Techniques and Application*, pp. 643-648, Dec 6-8, 1995.

[8] W. J. Zhang, Pharmacognostic study of Arctium Tomentosum Mill. and Microscopic Study of Chinese Materia Medica, Master of Medicine Thesis, Liaoning College of Traditional Chinese Medicine, 2001.

S.K. Choy
Centre for Mathematical Imaging and version
Department of Mathematics
Hong Kong Baptist Univeristy
Kowloong Tong
Hong Kong
e-mail: skchoy@math.hkbu.edu.hk

C.S. Tong
Centre for Mathematical Imaging and version
Department of Mathematics
Hong Kong Baptist Univeristy
Kowloong Tong
Hong Kong
e-mail: cstong@hkbu.edu.hk

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 453–462
© 2006 Birkhäuser Verlag Basel/Switzerland

A Novel Algorithm of Singular Points Detection for Fingerprint Images

Taizhe Tan and Jiwu Huang

Abstract. It is very important to effectively detect singularities (core and delta)for fingerprint matching, fingerprint classification and orientation flow modeling. In this paper, based on multilevel partitions in a fingerprint image, we present a new method of singularity detection to improve the accuracy and reliability of the singularities. Firstly, based on the information of the orientation field, with the Poincaré index method, we detect singularities which are estimated by different block sizes and various methods of orientation field estimation (smoothing or no smoothing). Secondly, based on the corresponding relationship between the singularities detected by multilevel block sizes and by different methods of orientation field estimation, we extract the singularities precisely and reliably. Finally, an experiment is done in the NJU-2000 fingerprint database that has 2500 fingerprints. The result shows that the method performs well and it is robust to poor quality images.

Keywords. fingerprint, singularity, multilevel partitions, orientation field smoothing.

1. Introduction

In recent years, the automatic fingerprint authentication technology has received more attention since fingerprint recognition based on distinctive personal traits has the potential to become an irreplaceable part of many identification systems [1, 2]. In a general way, in a fingerprint authentication system, it mainly includes fingerprint collection, fingerprint feature extraction, fingerprint matching and fingerprint classification and so on. Moreover, the fingerprint classification[3, 4, 5, 6, 7, 8] can be directly determined by the number, type and positions of the singular points. The singular points also represent the intrinsic points of reference within the fingerprint image and can therefore be used to register or approximately align two

This work was completed with the support of the Foundation for Doctor Start-up of Guangdong University of Technology(053017).

different images during matching. Thus, it is of great significance that the singularities are detected accurately and reliably.

Global patterns of ridges and furrows form special configurations in the central region of fingerprint, which is called pattern area. The pattern areas of loop or whorl types of fingerprints contain two types of singular points (delta and core). The delta provides the starting point for ridge tracing and ridge counting. The core can be used as reference point during fingerprint matching. The singular points are shown in Fig. 1.

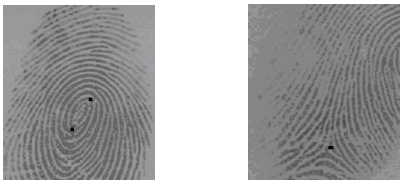


FIGURE 1. Examples of core point and delta point

There are some papers which have touched directly upon singularity detection of fingerprint images. Herny[9]defines the core as the upper point of the innermost ridge of the fingerprint. Srinivasan and Murthy[10]proposed an approach based on the examination of the local characteristics of the orientation image. They examined the dominant direction of the orientation in four quadrants around each point and used several rules to detect and classify the singular points. This method is very heuristic in its approach. An interesting implementation of the Poincaré method for locating singular points was proposed by Bazen and Gerez[11]. In this method, according to Green's theorem, a closed line integral over a vector field can be calculated as a surface integral over the rotation of this vector field; in practice, instead of summing angle differences along a closed path, the authors compute the "rotation" of the orientation image (through a further differentiation) and then perform a local integration (sum) in a small neighborhood of each element. Zhang et al.[12] suggested the corner detection method to find the region of singularities and gray level of ridges was tracked to get the position of singularities. However, estimating the location of the singular point itself is a difficult problem. Existing methods are based on computing the Poincaré index of the orientation image at each pixel position. These methods are sensitive to noise and generate many false singularities in noisy images. An elegant and practical method based on the Poincaré index being the main method for singularity detection was proposed by Kawagoe and Tojo[4]. Most of the approaches addressed in the literature for singularity detection operate on the fingerprint orientation image.

For images of better quality, the focus is how to fix the precise locations of singular points. For different partition sizes, the larger the partition size is, the more the position windage is. But the robust of the arithmetic is bad when the partition size is comparatively small. Whereas low quality images have serious noise pollution, the reliable orientation field estimation becomes difficult intrinsically, so

the precise locations of singular points are difficult to identify, meanwhile, there are many fake singularities detected in the mistaken directions. Thereby it is a very arduous task to effectively detect singularities from the fingerprint images. In this paper, in order to improve the precision and reliability of the detected singularities, we detect the singularities by partitioning the fingerprint image with multilevel sizes and using the relationship characters of locations of the estimated singularities based on the different methods for estimating the fingerprint orientation field, where the singularities detection is based on the Poincaré index.

This paper is organized as follows. The traditional Poincaré index theory and our method are introduced in section 2. Section 3 describes our detailed arithmetic. Experiments and discussions are given in the last section.

2. Singularity Detection Based on Poincaré Index

Poincaré index suggested by Kawagoe and Tojo[4] is used to detect singularities. Let o' be the orientation field of fingerprint image, the Poincaré index at pixel (i, j) of image which is enclosed by the digital curve can be computed as formula 2.1–2.4.

$$Poincaré(i,j) = \frac{1}{2\pi}\sum_{k=0}^{N_\Psi}\Delta(k) \tag{2.1}$$

$$\Delta(k) = \begin{cases} \delta(k), & \text{if } \mid \delta(k) \mid < \frac{\pi}{2}; \\ \pi + \delta(k), & \text{if } \delta(k) \leq -\frac{\pi}{2}; \\ \pi - \delta(k), & \text{otherwise.} \end{cases} \tag{2.2}$$

$$\delta(k) = o'(\Psi_x(i'), \Psi_y(i')) - o'(\Psi_x(i), \Psi_y(i)) \tag{2.3}$$

$$i' = (i+1) \mod N_\Psi \tag{2.4}$$

Where $\Psi_x(i)$ and$\Psi_y(i)$ are the x and y coordinates of the ith point of a closed digital curve with N_Ψ pixels respectively. The point(i, j) will be core point if the value of Poincaré index is 1/2, and delta point if the value of Poincaré index is -1/2.

K. Karu et al.[5] and J. G Chen et al.[8] both present the methods of singular points based on Poincaré index. In both methods, the measures of eliminating the noise pollution are taken. Although the method of smoothing time after time and iterative computing by K. Karu et al.[5] eliminates the noises partially, the model area information tends to be changed, besides the precision of the position of singularities is debased and /or the true singularities are lost, and the computation complexity is also increased. Moreover, using the different templates to check singularities, J. G Chen et al.[8] address the method of eliminating the noises in a way. But the singular points near fingerprint image margin are difficult to be detected since the sizes of the blocks and the template for Poincaré index computation are larger.

Both the methods above have their own merits and demerits respectively. In this paper, we present a method as follows: Firstly, fingerprint images are partitioned with multilevel block sizes and ridge directions of each block image are

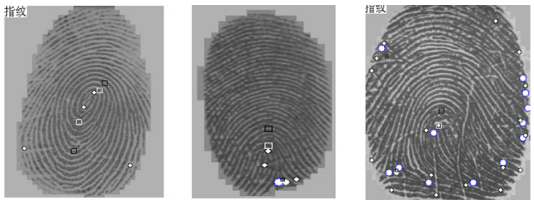


FIGURE 2. Singularities detected with our method, and corresponding relationship between these singularities' positions

estimated by smoothing and no smoothing processing respectively. By using the information of the estimated block directions, the Poincaré index of the centre point of each block is computed, and the types and locations of the singular points for different partitions are determined with the Poincaré index computed. Secondly, in order to ensure the locations of singularities are relatively precise, the centroid of the convergence area of the singularities detected in same way (block size) is regarded as the location of singular point of this partition way. There is a corresponding position relationship among the singularities detected by multilevel partitions and by the above-mentioned methods (shown in Fig. 2). For the sake of getting the precise locations of singularities, the locations of the singularities detected by the way of the small block size and no smoothing directions are regarded as the true locations of the final singularities. Finally, based on the corresponding relationship between the singularities detected by multilevel partitions and by different methods of orientation field estimation, we eliminate the false singularities, and retain the true, thereby the precision and reliability of the singularities are taken into account. The examples of singularities detected with our method and corresponding relationship between these singularities' positions are shown in Fig. 2. In the Figure, the large black square blocks denote the core points detected with 16×16 block image and smoothing the orientation field, the small black square blocks denote the delta points detected with 16×16 block image and no smoothing the orientation field, the large white square blocks denote the core points detected with 8×8 block image and smoothing the orientation field, the largest white ellipses denote the delta points detected with 8×8 block image and no smoothing the orientation field, the larger white ellipses denote the delta points detected with 8×8 block image and smoothing the orientation field, and the small white ellipses denote the core points detected with 8×8 block image and no smoothing the orientation field, where the central point of each plot corresponds with the location of the singular point detected in different ways.

3. Arithmetic Describing

Our arithmetic includes the following steps: (1) Image segmentation, (2) singularity detection based on multilevel partitions and different methods for direction

estimation, and (3) locating singularity's ultimate precise position based on the corresponding relationship between the singularities detected above.

3.1. Fingerprint Image Segmentation

Image area is segmented for the sake of getting the valid fingerprint area. In our method, the block-wise average grayscale and standard deviation are used to segment the image. The block is considered as the foreground if its grayscale mean and standard deviation satisfy some predefined standard, otherwise, the background.

3.2. Singularity Detection Based on Multilevel Partitions and Different Methods of Direction Estimation

3.2.1. Fingerprint Image Multilevel Partitions. The partition sizes are set with m levels, the size of each level block is $W_0 \times W_0, W_1 \times W_1, \ldots, W_{m-1} \times W_{m-1}, (W_i = W_0 \times (i+1), i = 0, 1, \ldots, m-1, where, W_0 = 8, m = 2)$.

3.2.2. Multilevel Partitions Based Singularity Detection. 1. **Fingerprint Image Orientation Estimation.** After each partition, a set of block images is gained. We use the method presented by Jain, Hong and Bolle[13] to estimate the orientation of each block image (u, v). But, here the orientation field is not processed with smoothing. 2. **Singularity Detection.** For a certain block size, a set of singularities is gotten by the Poincaré index. Assume that the sets of core and delta detected by the different block sizes are $\{CP_{i,j}\}$ and $\{DP_{i,j}\}$ respectively, where let i be the partition level and j be the jth point. To be more concrete, singularity detection includes the steps as follows:

Step 1. The Poincaré index of each block (u, v) is computed according to the orientation field, and based on the value of Poincaré index of each block, the core point and delta point are determined with the interested point within the template. Because the smaller the size of template is, the easier it is for the singularities on the image edge to be detected and the less the computation complexity is, the template size for computing the Poincaré index of core is 2×2, and for delta is 3×3. The interested points are all adopted as the center point of the template. It needs noting that only the delta points are detected in the second level of partition (16×16).

Step 2. Once the centroid coordinate of the convergence area of the singularities detected by Poincaré index in same block size is computed, hereby, CP_i and DP_i, the sets of the singularities which have the comparatively precise locations will be gained, where i denotes the partition level. The method for computing the centroid coordinate is k-means arithmeticz[14].

Step 3. The orientation of each block is smoothed with 2 dimension low-pass filter, and the size of filter is 5×5. Again, the value of Poincaré index of each block is computed by the smoothed orientation field and the new core and delta points are determined. The method is the same as before, but only the core points are detected in the second level of partition (16×16). Lastly, new sets of singularities CP'_0, DP'_0 and CP'_1 are gained.

3.3. Locating Singularity's Ultimate Precise Position

In this section, the singularity's ultimate location will be decided. Once the location of the singularity detected by the smallest block size and no smoothing orientation field is taken as the true location of the final singularity, and at the same time the character of corresponding relationship between the positions of singularities detected in our method in the same fingerprint image is considered, the fake singularity will be eliminated and the true be gotten. The approaches and rules of determining singularities are described as follows:

3.3.1. Locating Core Point.

1. Let each core point $CP'_{0,r}$ (r denotes the rth core) detected by the block size $W_0 \times W_0$ and smoothing the orientation field be the norm core point, and the cores of the set CP_0 are searched out within a certain distance according to the formula 3.1:

$$(CP_{0,j} \cdot x - CP'_{0,r} \cdot x)^2 + (CP_{0,j} \cdot y - CP'_{0,r} \cdot y)^2 < Threshold_0 \tag{3.1}$$

   If there is only one core $CP_{0,j}$, then the core $CP_{0,j}$ is determined as the true core point and is marked. And if there are cores CP'_1 which meet the formula 3.2 condition, then the core $CP_{0,j}$ is marked as the highest score and the core $CP'_{1,min}$ which is the nearest to the norm core $CP'_{0,r}$ is marked.

$$(CP'_{1,j} \cdot x - CP'_{0,r} \cdot x)^2 + (CP'_{1,j} \cdot y - CP'_{0,r} \cdot y)^2 < Threshold_1 \tag{3.2}$$

   Otherwise, the core $CP_{0,j}$ is marked as the second grade score.
2. If the conditions above are not met, then another judgment depends on whether there is a set of the core $CP_{0,j}$ in which the core $CP_{0,j}$ is very near to the norm core $CP'_{0,r}$ within a certain distance of the threshold value, where the distance threshold value is smaller. (1) If there is only one core $CP_{0,j}$ in the set, then $CP_{0,j}$ is determined as the true core point and is marked as the highest score. (2) If there is no $CP_{0,j}$ at all but there are some core points CP'_1 which meet with the formula 3.2, then the core $CP'_{1,min}$ which is the nearest to the norm core $CP'_{0,r}$ is marked, and the core $CP_{0,min}$ that is the nearest to the norm core $CP'_{0,r}$ and that is located between the core $CP'_{1,min}$ and the norm core $CP'_{0,r}$ is determined as the true core and is marked as the second grade score. (3) In cases other than the above two, the next step is performed.
3. First of all, It is important to tell whether there is a set of the core CP_0 that meet the condition that the form is the same as the formula 3.1, where the $Threshold_0$ in the formula 3.1 is changed to $Threshold_2 (Threshold_2 > Threshold_0)$. If there is no set of the core CP_0 that meet the condition above, then step 6 is performed. If there is this set, then the core CP_0 is marked as TCP_0, and whether there are the core CP'_1 which meet the condition has to be judged further according to formula 3.2. If it is valid, then the next step is performed. Otherwise, step 6 is carried out.
4. The core $CP'_{1,min}$ that is the nearest to the norm core $CP'_{0,r}$ is determined and marked from within the core CP'_1 which meet the formula 3.2. For every

core $TCP_{0,j}$, if it is located between the core $CP'_{1,min}$ and the norm core $CP'_{0,r}$, then let $TCP_{0,j}$ be $TCP'_{0,s}$ and again judge whether the distances between $TCP'_{0,s}$, $CP'_{0,r}$ and $CP'_{1,min}$ are nearer and whether the core $TCP'_{0,s}$ is located between the core $CP'_{1,min}$ and the norm core $CP'_{0,r}$. (1) If there is only one core $TCP'_{0,s}$ which meets the condition, then it is the true core point and marked as the highest score. (2) If there is more than one, then the true core points are determined by the same method as in the condition **(2)** of step **2**. (3) Otherwise, step 6 is carried out.

5. In the precondition of step **4**, if the core $TCP'_{0,s}$, $CP'_{0,r}$ and $CP'_{1,min}$ are largely in line, then the linearity measure between $TCP'_{0,s}$ and the line formed by $CP'_{0,r}$ and $CP'_{1,min}$ is computed according to the formulas 3.3 and 3.4, the minimum is found.

$$LineLevel = (a'^2 + b'^2 - c'^2)/2a'b' \tag{3.3}$$

$$Min_{LL} = \min\{Linelevel_0, LineLevel_1, LineLevel_2 \cdots\} \tag{3.4}$$

Where a', b' and c' denote the distance of $TCP'_{0,s}$ and $CP'_{0,r}$, the distance of $CP'_{0,r}$ and $CP'_{1,min}$ and the distance of $TCP'_{0,s}$ and $CP'_{1,min}$ respectively. (1) If Min_{LL} is smaller than a predefined threshold value, then the corresponding core $TCP'_{0,s}$ is regarded as the true core point and marked as the highest score. (2) If the condition above is not met, then judge once again whether there is a set of $TCP'_{0,s}$ that is within the predefined distance of the norm core $CP'_{0,r}$. If there is only one core $TCP'_{0,s}$ which meet the condition, then it is the true core point and marked as the highest score, or else, the core $TCP'_{0,s}$ nearest to the norm core $CP'_{0,r}$ is regarded as the true core and marked as the second grade score. (3) Otherwise, step **6** is carried out.

6. In other cases, the norm core $CP'_{0,r}$ is regarded as the true core point. And if there is a core $CP'_{1,min}$ that is within the predefined distance of the norm core and is the core nearest to the norm core, then the core $CP'_{1,min}$ is marked and the norm core $CP'_{0,r}$ is marked as the second grade score. Otherwise, as the third grade score.

3.3.2. Locating Delta Point. Step 1. Let each delta point $DP'_{0,r}$ (r denotes the rth core) detected by the block size $W_0 \times W_0$ and smoothing the orientation field be the norm delta point, and the delta set DP_0 are searched out within the predefined distance of the norm delta point. If there is only one delta $DP_{0,j}$, then the delta $DP_{0,j}$ is determined as the true delta point and is marked. And if there are some deltas DP'_1 which are within the predefined distance of the norm delta point, then the delta $DP_{0,j}$ is marked as the highest score and the delta $DP'_{1,min}$ which is the nearest to the norm delta $DP'_{0,r}$ is marked. Otherwise, the delta $DP_{0,j}$ is marked as the second grade score.

Step 2. If there is more than one delta like $DP_{0,j}$ which meets the conditions above, then the delta $DP_{0,min}$ that is nearest to the norm delta $DP'_{0,r}$ is marked as the true delta. And if there are deltas like DP_1 which are within the predefined

distance of the norm delta $DP'_{0,r}$, then the corresponding delta $DP_{0,min}$ is given the highest score, otherwise, given the second grade score.

Step 3. If there are only the delta DP_1 and no delta $DP_{0,j}$, which meet the conditions above, then the norm delta $DP'_{0,r}$ is marked as the delta and given the second grade score.

In all the steps above, the true deltas that have been detected are marked so as to avoid being involved in other delta detections.

3.3.3. Core Point Checking. Several cores detected above may be fake. So the detected cores ought to be checked according to the inherent structural characteristic of the fingerprint model area. The principles include: 1. the core with the highest score serves as the new norm and the core which is not reliable is checked; 2. If there are two or more than two cores detected, then they are checked by the concavo-convex characteristics of the model area which is defined by two cores; and 3. the distance between the two cores detected is smaller than that of the corresponding two norm cores CP'_0 and than that of the corresponding two cores CP'_1 which are detected on the large block size (16×16), and the directions of the vector radius of the curvature circle formed on the two true cores respectively are opposite generally. Last, the checked cores are sorted by their scores and placed in array of cores.

4. Experimental Results and Conclusions

A fingerprint singularity detection experiment based on the proposed algorithm is carried out on 500 typical fingerprint samples of NJU fingerprint database which contains 2500 images taken from 250 different fingers, 10 images per finger, and

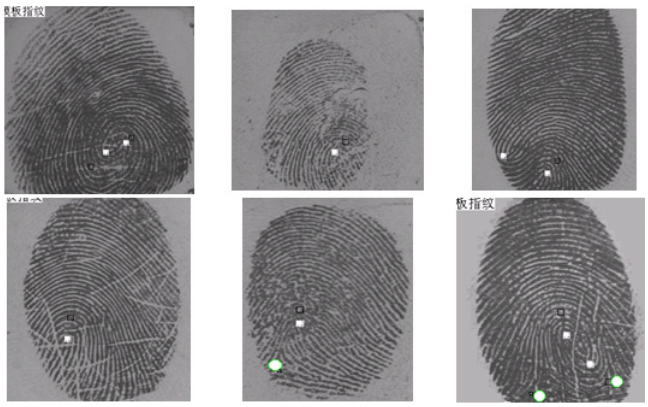


FIGURE 3. Examples of singularities detected Note: Results of the proposed method are denoted with the center of white ellipses, results of Poincaré index method based on are denoted with the center of black blocks

these fingerprint images are of varying quality in order to provide a realistic situation. In order to evaluate the performance of our algorithm, there is a comparison between the method based on the Poincaré index direct computation and our proposed method. In the method based on the Poincaré index direct computation, to improve the reliability of singularity detection, we adopt the larger block size (16×16) and smooth the orientation field. Fig. 3 shows some examples of singularity detection for typical fingerprints based on the two methods above. According to the experimental results, compared with the method of the Poincaré index direct computation, our method of singularities detection is more reliable and precise. For fingerprint images of poorer quality, our algorithm can still be used to get accurate and reliable singularities.

In this paper, a robust singularity detection algorithm is presented, which is based on the new idea of multilevel partition as well as the conformity of the singularities detected by multilevel partition and by the different methods of the orientation field estimation. Utilizing the characteristic of the corresponding relationship between the singularities detected by different approaches and measures of checking singularities, our method can improve the precision and reliability of singularity detection effectively and is robust to the low quality fingerprint images.

Although there are demerits in the singularity detection of images of quite low quality, it can meet fingerprint classification and fingerprint matching because the detected singularities are sorted by the reliability of singularity and some measures are adopted to restrict the position error to a very small degree.

References

[1] Yin YL, Ning XB, Zhang XM, *Development and application of automatic fingerprint identification technology.* Journal of Nanjing University (Natural Sciences) **38(1)** (2002), 29–35.

[2] Maltoni D., Maio D., Jain A. K. and Prabhakar S, *Handbook of Fingerprint Verification.* Springer Verlag, 2003.

[3] A. K. Jain, S. Prabhakar, and L. Hong, A Multichannel Approach to Fingerprint Classification, *Proc. of Indian Conference on Computer Vision, Graphics, and Images Processing (ICVGIP'98).* New Delhi, India, December 21–23,1998.

[4] M. Kawagoe and A. Tojo, *Fingerprint pattern classification.* Pattern Recognition **17(3)** (1984), 295–303.

[5] K. Karu and A. K. Jain, *Fingerprint Classification.* Pattern Recognition **29(3)** (1996), 389–404.

[6] A. K. Jain, S. Prabhakar and S. Pankanti, *Mathing and Classification:A Case Study in Fingerprint Domain.* Pro. INSA-A (Indian National Science Academy) **67(2)** (2001), 67–85.

[7] Lumini A., Maio D., Maltoni D., *Continuous vs Exclusive Classification for Fingerprint Retrieval.* Pattern Recognition **18(10)** (1997), 1027–1034.

[8] Cheng JG, Tian J, Ren Q, Zhang TH,Singular Point-based Fingerprint Classification, *Second Workshop on Biometrics in conjunction with the 6^{th} International Conference for Young Computer Scientists.* Hangzhou,China,S4-27–S4-34,2001.

[9] E. R. Herny, *Classification and use of Fingerprints.* George Routledge and Sons, London, 1900.

[10] V. S. Srinivasan and N. N. Murthy, *Detection of Singular Points in Fingerprint Images.* Pattern Recognition **25(2)** (1992), 139–153.

[11] Bazen A. M. and Gerez S. H., *Systematic Methods for the Computation of Directional Fields and Singular Points of Fingerprints.* IEEE Transactions on PAMI **24(7)** (2002), 905–919.

[12] Zhang WW, Wang S, Wang YS,Corner Detection Based Singularity Detection of Fingerprint Image, *Second Workshop on Biometrics in conjunction with the 6^{th} International Conference for Young Computer Scientists.* Hangzhou,China,S4-51–S4-56,2001.

[13] A. K. Jain, L. Hong and R. Bolle, *On-line Fingerprint Verification.* IEEE Trans. Pattern Anal. and Machine Intell **19(4)** (1997), 302–314.

[14] J.B. MacQueen, Some methods for classification and analysis of multivariate observations,*Pro. Symp. Math. Statistics and Probablity.* University of California Press, Berkeley 1967.

Taizhe Tan
School of Information Science & Technology at Sun Yat-sen (Zhongshan) University
510275 Guangzhou,China;
Faculty of Computer, Guangdong University of Technology
510090 Guangzhou,China
e-mail: tantaizhe@263.net

Jiwu Huang
School of Information Science & Technology at Sun Yat-sen (Zhongshan) University
510275 Guangzhou,China

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 463–475
© 2006 Birkhäuser Verlag Basel/Switzerland

Wavelet Receiver: A New Receiver Scheme for Doubly-Selective Channels

Guangjian Shi and Silong Peng

Abstract. We present a new receiver scheme, termed as *wavelet receiver*, for doubly-selective channels to combat the annoying Doppler effect. The key point is to convert the Doppler effect to Doppler diversity, taking advantage of the diversity technique to improve system performance. To this end, a new framework based on multiresolution analysis (MRA) is established. In this framework, we find that RAKE receiver which can only combat the multipath fading, is a special case of *wavelet receiver* which can exploit joint multipath and Doppler fading. Theoretical analysis and experimental simulation show that *wavelet receiver* can greatly enhance system performance.

Mathematics Subject Classification (2000). Primary 42C40; Secondary 60G35.

Keywords. Doubly-selective channel, Doppler diversity, multiresolution analysis (MRA), wavelet receiver.

1. Introduction

Time selectivity and frequency selectivity characterize the fading channel and they exhibit somewhat of similarity in terms of scattering function of channel [1]. Inspired by the success of RAKE receiver in the scenario of frequency selective channel, which resolves the received signal to multipath diversity, we seek for the possible receiver scheme for doubly-selective channels.

The idea originated by Sayeed [2] may be the first attempt of converting the Doppler effect to Doppler diversity. This idea has been extended to various systems. Its application to OFDM system can be found in [3, 4] and to space-time diversity can be found in [5]. But the method of Sayeed, which is referred as time-frequency representation (TFR) method, suffers from a drawback: the basis functions of TFR method are not orthogonal.

This work was supported by National Natural Science Foundation of China with No.60272042.

In this paper, we follow Sayeed's idea but exploit the Doppler diversity in another way. It based on the fact that usually used shaping pulses are the scaling functions of wavelet. For example, the square root raised cosine (SRRC) pulses with excess bandwidth in [0,1/3] is the scaling function of Meyer wavelet [6], [7] and the rectangle pulse is the scaling function of Harr wavelet. It stimulates us to study the Doppler diversity in multiresolution analysis (MRA) framework. In the MRA framework, we find that the RAKE receiver is a special case of *wavelet receiver*. The superiority of *wavelet receiver* over RAKE receiver is in that *wavelet receiver* can exploit orthogonality of wavelet space in addition to that of scaling space. On the other hand, *wavelet receiver* can be implemented by a series of parallel RAKE receivers, which will be detailed in section 4.

The paper is organized as follows. System model and preliminary knowledge of MRA are introduced in the next section. In Section 3, the diversity concept is reviewed and the sufficient and necessary conditions for the existence of Doppler diversity are also established. Our proposed *wavelet receiver* is the focus of Section 4. In Section 5, Numerical experiments are provided. At last conclusion is drawn in Section 6.

1.1. Notation

Let us introduce some notational conventions used in this paper. Z denotes the integer field, Z^- the nonpositive integer set and $L^2 = \{x(t) : \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty\}$ square integrable space. $Q(x) = 1/\sqrt{2\pi}\int_x^{-\infty} e^{-x^2/2}dx$ denotes the Q-function, $\Gamma(\cdot)$ denotes Γ-function and $\delta_{i,j}$ is Kronecker delta function. $card\{\cdot\}$ is reserved for cardinality of set, $(\cdot)^\perp$ for orthogonal complement. We denote the inner product as $< \cdot, \cdot >$, L^2 norm as $\| \cdot \|$, projection operator of space X as $P_X(\cdot)$. $E(\cdot)$ is expectation of random variable.

2. Preliminaries

2.1. System Model and RAKE Receiver

Consider a typical wireless system and channel is assumed to be doubly-selective. The equivalent low pass model can be expressed as

$$r(t) = s(t) + z(t), \tag{2.1}$$

where

$$s(t) = \int_{-\infty}^{\infty} h(t, t-\tau)x(\tau)d\tau \tag{2.2}$$

and $z(t)$ is the additive white Gaussian noise (AWGN) with power spectrum density (PSD) $\sigma^2/2$ per dimension, $x(t)$ is pulse shaping function and $h(t,\tau)$ denotes the time-varying impulse response of the channel. $h(t,\tau)$ is assumed to be wide-sense stationary with uncorrelated scatter (WSSUS) complex Gaussian stochastic process. In a system equipped with bandwidth W, $s(t)$ can be expressed as [1]

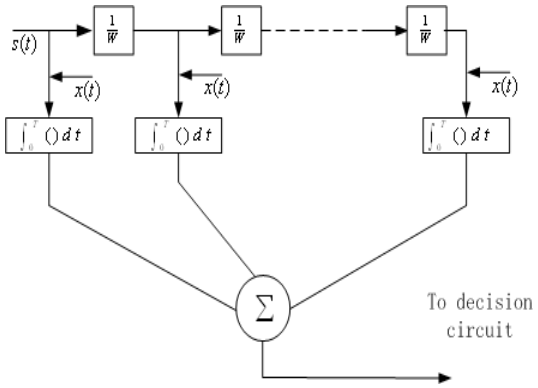


FIGURE 1. The structure of RAKE receiver

$$s(t) = \sum_{n=-\infty}^{\infty} h_n(t)x(t - n/W). \tag{2.3}$$

If fading is slow, we can get the following signal model

$$s(t) = \sum_{n=-\infty}^{\infty} h_n x(t - n/W). \tag{2.4}$$

In such case, there exists a kind of optimum receiver originated by Price and Green. Its structure is shown in Fig.1, which is somewhat analogous to the ordinary garden rake, so it's often referred to as RAKE receiver [1].

2.2. Multiresolution Analysis(MRA)

This paper is established on the basis of multiresolution analysis, so in this subsection we will introduce some fundamental knowledge of MRA although it is presented in the most compact form.

Theorem 2.1 (Mallat). [8, p.129] *If a sequence of successive approximation spaces V_j satisfies the following*

1. $\cdots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$
2. $\overline{\bigcup\limits_{j\in Z} V_j} = L^2(R)$
3. $\bigcap\limits_{j\in Z} V_j = \{0\}$
4. $f \in V_0 \Rightarrow f(\cdot - n) \in V_0$ *for all* $n \in Z$
5. $f \in V_j \Leftrightarrow f(2^j \cdot) \in V_0$
6. *there exists* $\phi(x)$ *so that* $\{\phi_{0,n} : n \in Z\}$ *is an orthonormal basis in* V_0

Then conclusions *can be drawn that there exist wavelet* $\psi(x)$*, whose translations and scaled versions* $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$ *satisfy:*

$$\begin{aligned} \textit{Let} \quad W_j &= span\{\psi_{j,k}(x)\ k \in Z\},\ then \\ V_{j-1} &= V_j \bigoplus W_j\ \ for\ j \in Z \\ L_2 &= span\{\psi_{j,k}(x)\ j,k \in Z\}. \end{aligned}$$

In the next subsection, we will see that the scaling space V_0 is just the space spanned by the shaping pulses. For further reference convenience, we introduce a corollary as below

Corollary 2.2. $L_2 = V_0 + \overset{0}{\underset{j=-\infty}{\bigoplus}} W_j$ *and* $V_0^{\perp} = \overset{0}{\underset{j=-\infty}{\bigoplus}} W_j$.

This corollary gives the exact structure of the orthogonal complement of V_0. Now we arrive at the following signal representation

$$s(t) = \sum_{n=-\infty}^{\infty} a_n \phi_{0,n}(t) + \sum_{j=-\infty}^{0} \sum_{k=-\infty}^{\infty} b_{j,k}\psi_{j,k}(t). \tag{2.5}$$

2.3. Nyquist Pulse and Wavelet

William [7] has noted that Square Root Raised Cosine (SRRC) pulse with excess bandwidth in [0,1/3] is the scaling function of Meyer wavelet. In fact, scaling function of any wavelet can be used as Nyquist pulse. Nyquist pulse $x(t)$ satisfies $\sum_{m=-\infty}^{\infty} X(f + m/T) = T$, which is just the equivalent counterpart of theorem 2.1's condition 6 in frequency domain [1, p. 558] and [8, p. 132]. It implies scaling function of any wavelet satisfies Nyquist pulse condition. On the other hand, the generally used shaping pulse is exact scaling function of wavelet. For example, the Square Root Raised Cosine (SRRC) pulse with excess bandwidth ranged at [0,1/3] is the scaling function of Meyer wavelet function [6], [7] and the rectangle pulse is that of Haar wavelet.

3. Multipath and Doppler Diversity

In this section, diversity concepts are reviewed from a new perspective and at the same time the necessary and sufficient condition for the existence of the Doppler diversity is presented.

Diversity with order L traditionally is viewed as independent transmission of the same information L times with a specified but the same waveform

$$r_i(t) = \alpha_i \exp(-j\theta_i)s(t)\ i = 1 \cdots L. \tag{3.1}$$

In fact, the same information can be transmitted by different waveforms as long as the receivers have the knowledge about them. So we can extend the concept of diversity to more general form

$$r_{i_k}(t) = \alpha_{i_k} \exp(-j\theta_{i_k}) s_k(t) \; i_k = 1, \cdots, L_k \; k = 1, \cdots K, \tag{3.2}$$

with order

$$L = \sum_{k=1}^{K} L_k.$$

With such extension, we now can define multipath and Doppler diversity in the doubly-selective channels.

Definition 3.1. Multipath diversity is defined as $m(t) = P_{V_0} s(t)$, Doppler diversity as $d(t) = P_{V_0^{\perp}} s(t)$.

Definition 3.2 (Multipath Diversity Order (MDO)). According to the expression

$$m(t) = \sum_{n=-\infty}^{\infty} a_n \phi_{0,n}(x), \tag{3.3}$$

define MDO as the number: $card\{n : a_n \neq 0, \; n \in Z\}$.

Definition 3.3 (Doppler Diversity Order (DDO)). According to the expression

$$d(t) = \sum_{j=-\infty}^{0} \sum_{k=-\infty}^{\infty} b_{j,k} \psi_{j,k}(x), \tag{3.4}$$

define DDO as the number: $card\{(j,k) : b_{j,k} \neq 0, \; j \in Z^-, \; k \in Z\}$.

3.1. Existence Theorem for Doppler Diversity

Theorem 3.4 (Existence Theorem for Doppler Diversity). *The Doppler Diversity exists if and only if* $card\{(j,k) : b_{j,k} \geq \sigma^2, \; j \in Z^-, \; k \in Z\}$ *is nonzero, where* $b_{j,k} = E(| < r(t), \psi_{j,k}(t) > |^2)$

Proof. The proof is presented in Appendix 7.1 □

4. Wavelet Receiver

Besides the multipath diversity, the Doppler diversity now is well structured. In this section, we take advantage of this structure to maximize the joint multipath and Doppler diversity. This comes to our proposed *wavelet receiver*.

The structure of *wavelet receiver* is shown in Fig. 2, which is composed of a series of parallel RAKE receivers. Each RAKE receiver is characterized by three parameters: the matched filter, the interval of translation and the order of the diversity. *Wavelet receiver* can be viewed as a kind of 2D-RAKE receiver.

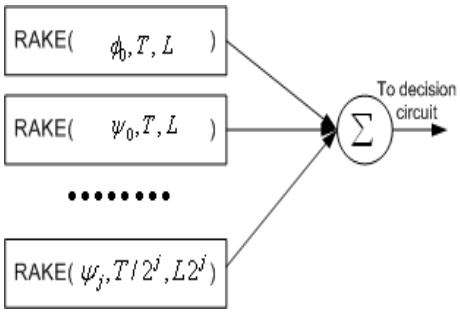


FIGURE 2. The structure of wavelet receiver

The decision variable for coherent detection of *wavelet receiver* with maximal ratio combiner can be expressed as

$$\begin{aligned} D &= Re[\int_0^T r(t)s^*(t)dt] \\ &= Re[\sum_{n=-\infty}^{\infty}\int_0^T r(t)h_n^*(t)x^*(t-n/W)dt] \\ &= Re[\sum_{n=-\infty}^{\infty} a_n \int_0^T r(t)x^*(t-n/W)dt \\ &+ \sum_{j=-\infty}^{0}\sum_{k=-\infty}^{\infty} b_{j,k}\int_0^T r(t)\psi_{j,k}^*(t)dt] \\ &= Re[2\varepsilon\sum_{n=-\infty}^{\infty}|a_n|^2 + 2\varepsilon\sum_{j=-\infty}^{0}\sum_{k=-\infty}^{\infty}|b_{j,k}|^2 \\ &+ \int_0^T z(t)s^*(t)dt], \end{aligned} \tag{4.1}$$

where ε denotes the energy of the shaping pulse. It is clear that the diversity of a system combines multipath diversity item $E_{mul} = 2\varepsilon\sum_{n=-\infty}^{\infty}|a_n|^2$ and Doppler diversity item $E_{dop} = 2\varepsilon\sum_{j=-\infty}^{0}\sum_{k=-\infty}^{\infty}|b_{j,k}|^2$. If the channel is slow fading, all $b_{j,k}$ tend to be zeros, the *wavelet receiver* reduces to the RAKE receiver. The additional performance gain E_{dop} enhances the signal-to-noise-ratio (SNR) from

$$SNR_r = \frac{E_{mul}}{\sigma^2}$$

to

$$SNR_w = \frac{E_{mul} + E_{dop}}{\sigma^2},$$

where SNR_r and SNR_w denote the SNR of the RAKE receiver and *wavelet receiver* respectively. We term this phenomenon as SNR enhancement effect of *wavelet receiver*.

4.1. Independence Analysis

Independence is crucial to diversity technique. The statistical characteristics of the noise components determine the optimal combining algorithm. If the maximum ratio combiner is used, the independence of noise components is expected. In this subsection, we present noise independence analysis. The independence can be obtained by the orthogonality of the basis function. For notation simplicity, we stack the scaling function in V_0 and wavelets in $\{W_j,\ j \in\ Z^-\}$ into the B_i with uniform index i. The noise component projected to B_i is obtained as

$$N_i =< z(t), B_i(t) >, \tag{4.2}$$

so

$$\begin{aligned} E(N_i N_j^*) &= E\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z(t)z^*(\tau)B_i^*(t)B_j(\tau)dtd\tau \\ &= \sigma^2\int_{-\infty}^{\infty} B_i^*(t)B_j(t)dt \\ &= \sigma^2\delta_{i,j}. \end{aligned} \tag{4.3}$$

4.2. Performance Analysis

In this subsection, the performance analysis of *wavelet receiver* is presented. For performance comparison, we give the BER [1] analysis of the BPSK modulated signaling using the *wavelet receiver* and RAKE receiver, although the result for other modulated signaling is also available. For *wavelet receiver*, the BER is

$$p_w = Q(\sqrt{SNR_w}), \tag{4.4}$$

and

$$p_r = Q(\sqrt{SNR_r}) \tag{4.5}$$

for the RAKE receiver.

In case of statistical fading channel, system performance P not only depends on the SNR, but also the statistical distribution of the E_{mul} and E_{dop}. That is

$$P = \int_0^{\infty} P(\gamma)p(\gamma)d\gamma, \tag{4.6}$$

where

$$\gamma = 2\varepsilon\sum_{n=-\infty}^{\infty}|a_n|^2/\sigma^2 + 2\varepsilon\sum_{j=-\infty}^{0}\sum_{k=-\infty}^{\infty}|b_{j,k}|^2/\sigma^2. \tag{4.7}$$

[1]The probability that a bit-error will occur in a given time is referred to as the bit-error rate (BER).

Following the derivation of the [1, p.846-847], we get the error probability P in the Rayleigh fading channel can be approximated by

$$P \approx \left(\begin{array}{c} 2T-1 \\ T \end{array} \right) (\frac{1}{4\gamma_1})^L (\frac{1}{4\gamma_2})^{T-L}, \tag{4.8}$$

where T, L denote the total and multipath diversity order respectively and $\gamma_1 = \varepsilon/\sigma^2 E(a_n)$, $\gamma_2 = \varepsilon/\sigma^2 E(b_{j,k})$ are average SNR per channel for multipath and Doppler diversity gain. And the corresponding performance P_r of RAKE receiver is

$$P_r \approx \left(\begin{array}{c} 2L-1 \\ L \end{array} \right) (\frac{1}{4\gamma_1})^L. \tag{4.9}$$

4.3. Channel Estimation

The *wavelet receiver* challenges new channel estimation algorithms. In coherent detection, not only the multipath diversity gains, but also the Doppler diversity gains are needed to be estimated. It obviously complicates the channel estimation algorithm. The variable to be estimated, in fact, is exponentially grown as the scale parameter limits to infinity ($j \to \infty$). Let the multipath diversity order (MDO) be L, the number of variables to be estimated for Doppler diversity gain is $(2^{J+1}-1)L$ if the maximal scale is 2^J.

To avoid the channel gain estimation for the Doppler diversity branches, we can use the noncoherent detection for Doppler diversity. This come to our *wavelet receiver Form 2*, we termed it as *half coherent wavelet receiver*. To distinguish with each other, we termed above *wavelet receiver* as *wavelet receiver Form 1* or *coherent wavelet receiver*. In this paper, *half coherent wavelet receiver* is realized simply by square-law combiner for all Doppler diversity branches. Of course if more prior knowledge is available, weighted square-law combiner can be applied to optimize the performance.

5. Experiments

In experiments, the time variation of channel is model as Jake's model. The Doppler power spectrum of mobile radio channel is

$$\zeta(\Delta t) = 1/2E[(h_n(t)h_n^*(t+\Delta t))] = J_0(2\pi f_m \Delta t),$$

where $J_0(t)$ is the zero-order Bessel function of the first kind and f_m is the maximum Doppler frequency [1, p.809]. The maximal value of Δt is T. We usually use the normalized maximum Doppler spread $f_m T$ as the parameter of the Jake's model

$$\zeta(t) = J_0(2\pi f_m T t), \tag{5.1}$$

where t ranges [0,1]. In this paper, Jake's model is simulated by IFFT method [9]. The shaping pulse is SRRC with excess bandwidth 1/3 and the maximal scale parameter J is set to 5.

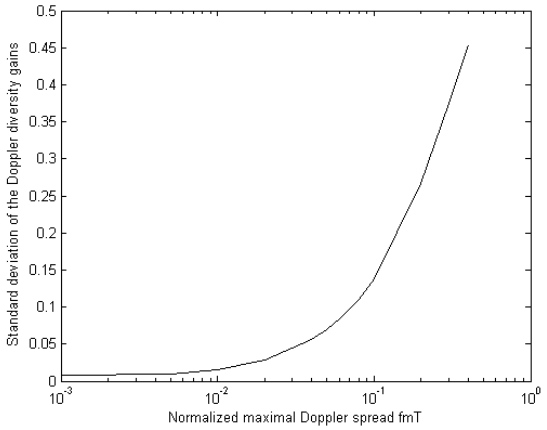


FIGURE 3. Plot of the standard deviation of Doppler diversity gains over normalized maximum Doppler spread f_mT.

5.1. Simulation Result 1

The variance of Doppler diversity gain $b_{j,k}$ depends on the normalized maximum Doppler spread f_mT. In this experiment, we will present the numerical result. The result is tabulated in Table 1, where $S(a)$,$S(b)$ denotes standard deviation of gain in multipath diversity branch and Doppler diversity branch respectively, and $m(b)$ is the mean of Doppler diversity gain. The normalized maximum Doppler spread is set as $0.001, 0.005, 0.01, 0.02, 0.04, 0.05, 0.06, 0.08, 0.1, 0.2, 0.3$ and 0.4 respectively. At any specified value, 1000 Monte Carlo simulation is run. Fig. 3 illustrates the standard deviation of the Doppler diversity gain increases monotonously with f_mT, which verifies the intuition that more Doppler diversity is obtained when the channel statistics varies faster. Although the value in Table 1 is obtained at $\varepsilon = 8$, the ratio between the $S(a)$ and $S(b)$ has no dependence on ε.

5.2. Simulation Result 2

Fig. 4 shows the performances of RAKE receiver and corresponding *wavelet receiver* in different normalized maximum Doppler spread f_mT. Coherent detection with maximum ratio combiner is assumed. In the Fig. 4, SNR per bit is evaluated as $\varepsilon/\sigma^2 E(a_n)$, which is identical to the case of RAKE receiver. *Wavelet receiver* has significant performance improvement compared with RAKE receiver. These improvement is obviously contributed to the SNR enhancement effect of *wavelet receiver*. The performance of *half coherent wavelet receiver* with square-law combiner for Doppler diversity gains is also presented in Fig. 5. We can see if lower BER is desired, *half coherent wavelet receiver* is also a better choice.

TABLE 1. Relation between normalized maximum Doppler spread f_mT and standard deviation of multipath diversity gain $S(a)$, mean and standard deviation of Doppler diversity gain $m(b)$, $S(b)$

f_mT	0.001	0.005	0.01	0.02	0.04	0.05	0.06	0.08	0.1	0.2	0.3	0.4
$S(a)$	5.5714	4.418	4.5605	4.417	4.6089	4.5446	4.5336	4.4874	4.499	4.8212	5.1648	5.4374
$m(b)$	6.459	-16.91	-24.24	-6.763	18.54	17.48	-3.323	-1.061	23.56	41.91	58.19	11.92
$S(b)$	0.007573	0.01001	0.01575	0.02872	0.05590	0.06951	0.08349	0.1106	0.1381	0.2665	0.3750	0.4544

$*10^{-5}$

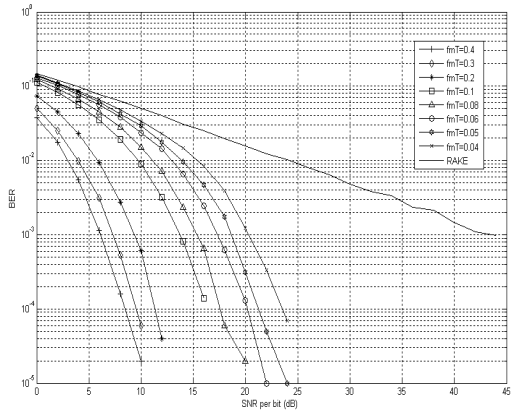


FIGURE 4. Performance comparison between RAKE receiver and *coherent wavelet receiver* with different normalized maximum Doppler spread $f_m T$.

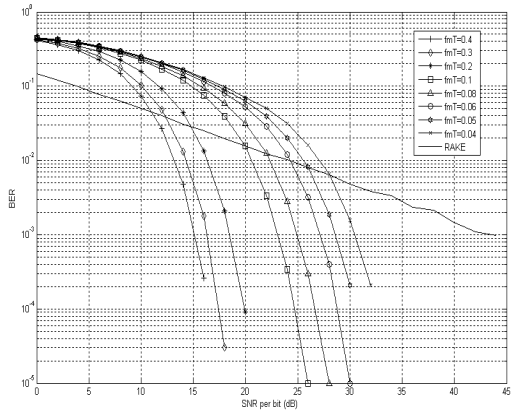


FIGURE 5. Performance comparison between RAKE receiver and *half coherent wavelet receiver* with different normalized maximum Doppler spread $f_m T$.

6. Conclusion

We have proposed a novel framework to exploit the Doppler diversity, which provides a new way to investigate mobile communication systems with fast fading channel. Based on the well-established wavelet theory, we have derived the structure of Doppler diversity. Two forms of *wavelet receiver* are established. It shows

that the time variations can be totally captured by *wavelet receiver*. And the quantitive relations between Doppler diversity and normalized maximum Doppler spread furthers the understanding of the fast fading communications systems.

Multiresolution analysis framework unites the signal processing of the fast and slow fading channel. In this united framework, we can exploit the achievements of the slow fading channel to benefit the fast fading channel processing. This promises great application scenarios of the *wavelet receiver*, such as its application to CDMA or OFDM systems etc., which will be our future work.

7. Appendices

7.1. Proof of the Theorem 3.4

Proof. According to (2.5)

$$\begin{aligned} B_{j,k} &= < r(t), \psi_{j,k} > = < s(t), \psi_{j,k} > + < z(t), \psi_{j,k} > \\ &= b_{j,k} + b_{j,k}^{noise}, \end{aligned}$$

we can get

$$E(|B_{j,k}|^2) = E(|b_{j,k} + b_{j,k}^{noise}|^2) = b_{j,k}^2 + \sigma^2, \tag{7.1}$$

due to the fact that $z(t)$ is AWGN with PSD σ^2. Now using the conclusion of Definition 3.3, we arrive at conclusion of this theorem. □

7.2. The Derivation of the SNR of the Wavelet Receiver

The output of maximum ratio combiner of *wavelet receiver* can be expressed as a decision variable formed in (4.1)

$$D = E_{mul} + E_{dop} + \int_0^T z(t)s^*(t)dt. \tag{7.2}$$

The noise power is

$$E(|\int_0^T z(t)s^*(t)dt|^2) = \|s(t)\|^2\sigma^2, \tag{7.3}$$

and $\|s(t)\|^2 = E_{mul} + E_{dop}$, so the SNR of the *wavelet receiver* is

$$SNR_w = \frac{(E_{mul} + E_{dop})^2}{(E_{mul} + E_{dop}) * \sigma^2} = \frac{E_{mul} + E_{dop}}{\sigma^2} \tag{7.4}$$

References

[1] J. G. Proakis *Digital Communications,* 4th ed. New York: MaGraw-Hill, 2001.

[2] A. M. Sayeed and B. Aazhang "Joint multipath-Doppler diverity in Mobile wireless comunications," *IEEE Trans. Commun.*, vol. 49, pp. 123-132, Jan. 1999.

[3] B. C. Kim and I. T. Lu "Doppler diversity for OFDM wireless mobile communications Part I: frequency domain approach," in *proc. 2003 IEEE Vehicular Technology Conf.(VTC 57th)*, 22-25 Apr. 2003 vol. 4, pp. 2677-2681.

[4] B. C. Kim and I. T. Lu "Doppler diversity for OFDM wireless mobile communications. Part II: time-frequency processing," in *proc. 2003 IEEE Vehicular Technology Conf.(VTC 57th)*, 22-25 Apr. 2003, vol. 4, pp. 2682-2685.

[5] A. G. Feng, Q. Yin, J. G. Zhang and Z. Zhao "Joint space-multipath-Doppler RAKE receiving in DS-CDMA systems over time-selective fading channels," in *proc. IEEE Int. Symp. on Circuits Syst.* vol.1, pp. 601-604, May 2002.

[6] W. Chen, Q. Yang, W. J. Jiang, and S. L. Peng "On bandlimited scaling function," *Journal of Computionational Mathematics*, vol. 20, pp. 373-380, Apr. 2002.

[7] W. J. Willime and C. D. Jeffrey "The square root raised cosine wavelet and tts relation to the Meyer functions," *IEEE Trans. Singal Processing*, vol. 49, pp. 248-251, Jan. 2001.

[8] I. Daubechies, *Ten Lectures on Wavelets,* Philadelphia, SIAM, 1992.

[9] C. Komninakis and J. F. Kirshman "Fast Rayleigh fading simulation with an IIR filter and polyphase interpolation," available at http://rfdesign.com/ar/407rfdf2x.pdf

Guangjian Shi
National ASIC Design center
Institute of Automation, Chinese Academy of Sciences, Beijing, China
e-mail: guangjianshi@yahoo.com.cn

Silong Peng
National ASIC Design center
Institute of Automation, Chinese Academy of Sciences, Beijing, China
e-mail: silong.peng@ia.ac.cn

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 477–487
© 2006 Birkhäuser Verlag Basel/Switzerland

Face Retrieval with Relevance Feedback Using Lifting Wavelets Features

Chon Fong Wong, Jianke Zhu, Mang I Vai,
Peng Un Mak and Weikou Ye

Abstract. By using support vector machine (SVM), this paper presents a novel face retrieval scheme in face database based on lifting wavelets features. The relevance feedback mechanism is also performed. The scheme can be described in three stages as follows. First, lifting wavelets decomposition technique is employed because it not only can extract the optimal intrinsic features for representing a face image, but also can accelerate the speed of the wavelets transform. Second, Linear Discriminant Analysis (LDA) is adopted to reduce the feature dimensionality and enhance the class discriminability. Third, relevance feedback using SVM is applied to learn on user's feedback to refine the retrieval performance. The experimental evaluation has been conducted on ORL dataset in which the results show that our proposed approach is effective and promising.

Keywords. face retrieval, face recognition, Content-based image retrieval (CBIR), Lifting Wavelets Transform (LFWT), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), relevance feedback.

1. Introduction

Face retrieval can be considered as a problem of similar face searching in the feature space by integrating content-based image retrieval (CBIR) [1] and face recognition [2] techniques. Among many approaches for facial representations, wavelets decomposition technique [3, 4] is a powerful method that is used to extract the intrinsic features for face recognition and image retrieval. However, due to the semantic gap existing between low-level features and high-level concepts [5], it is often difficult to achieve satisfactory retrieval performance by using rigid similarity measure on low-level features in a larger face database. One of the feasible ways to bridge the gap is to utilize relevance feedback to refine the retrieval results [5].

In this paper, we study the problem of face retrieval with SVM-based relevance feedback. We propose to use lifting wavelets features extraction algorithm because it can find the optimal bases for the representation of a face image [13]. Moreover, lifting scheme provides wavelets transform to be performed in spatial space and maps integers to integers, which accelerates the speed of the wavelets transform [6]. As the dimension of the feature vectors extracted by LFWT is still high, this is not proper for the retrieval task. To tackle this problem, Linear Discriminant Analysis (LDA) [8] is utilized to further reduce the feature dimesionality and enhance the class discriminability. Because LDA directly applied to feature vectors may cause the singularity problem at the within-class scatter matrix, Principal Component Analysis (PCA) [7] is commonly pre-performed for dimension reduction [8]. While the discriminant and lower-dimensional feature vectors are extracted, the remaining key element of face retrieval is to design a robust classifier. Previously, the face classifier based on the nearest feature line (NFL) [10] and nearest feature space (NFS) was exploited [3]. However, in the case of small training samples in databases, the ability of these classifiers for making confident decisions under different class boundaries is still needed to be explored. It is well-known that SVM [9] classifier has a good generalization performance in tacking small sample size in pattern recognition. When incorporated with relevance feedback, SVM classifier can be learned from training data of relevance face images and irrelevance face images marked by the users. Then the model can be used to find more relevance face images in the whole database [11, 12].

The rest of the paper is organized as follows. In Section 2, detailed description of lifting wavelets transform is presented. Subspace analysis methods, such as PCA and LDA, will be studied in Section 3. Section 4 briefly introduces relevance feedback technique. SVM classification method is reviewed in Section 5. Section 6 discusses the experimental results. Conclusions are drawn in Section 7.

2. Lifting Wavelets Transform

Any discrete wavelets transform (DWT) or two subband filtering with finite filters can be decomposed into a finite sequence of simple filtering steps, which are called the lifting steps. This decomposition corresponding to a factorization of the polyphase matrix of wavelet or subband filters into elementary matrices is described as follows.

The polyphase representation of a discrete-time filter $h(z)$ is defined as

$$h(z) = h_e(z^2) + z^{-1}h_o(z^2) \tag{2.1}$$

where h_e denotes the even coefficients, and h_o denotes the odd coefficients:

$$h_e(z) = \sum_k h_{2k} z^{-k}, \text{ and } h_o(z) = \sum_k h_{2k+1} z^{-k} \tag{2.2}$$

The synthesis filters $h(z)$ and $g(z)$ (low-pass and high-pass filters respectively) can thus be represented by their polyphase matrix $P(z) = \begin{bmatrix} h_e(z) & g_e(z) \\ h_o(z) & g_o(z) \end{bmatrix}$ and $\tilde{P}(z)$ can be also defined for the analysis filters analogously. The filters $h_e(z)$,$h_o(z)$,$g_e(z)$ and $g_o(z)$, along with their analysis counterparts, are Laurent polynomials. As the set of all Laurent polynomials exhibits a commutative ring structure, within which polynomial division with remainder is possible, long division between Laurent polynomials is not a unique operation [6].

The Euclidean algorithm [6] can be used to decompose $P(z)$and $\tilde{P}(z)$as

$$P(z) = \prod_{i=1}^{m} \begin{bmatrix} 1 & s_i(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_i(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} \tag{2.3}$$

$$\tilde{P}(z) = \prod_{i=1}^{m} \begin{bmatrix} 1 & 0 \\ -s_i(z^{-1}) & 1 \end{bmatrix} \begin{bmatrix} 1 & -t_i(z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/K & 0 \\ 0 & K \end{bmatrix} \tag{2.4}$$

As this factorization is not unique, several pairs of $\{s_i(z)\}$ and $\{t_i(z)\}$ filters are admissible; however, in case of DWT implementation, all possible choices are equivalent.

Due to the computational complexity, it has been shown in [6] that using the lifting scheme instead of the standard filter bank algorithm, can asymptotically lead to a relative speed-up to 64% for Daubechies 9/7 filter which is implemented by the scheme shown in Figure 1.

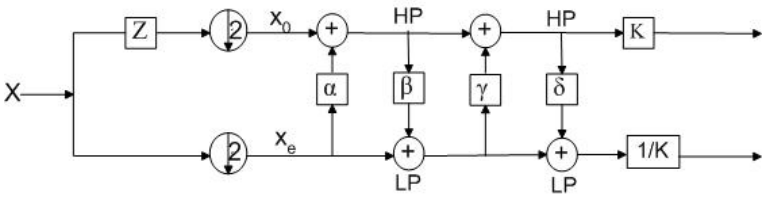


FIGURE 1. Block diagram of the forward wavelet transform using lifting scheme: the implementation of Daubechies 9/7 filter α=-1.586134342; β=-0.05298011854; γ=0.8829110762; δ=0.4435068522; K=1.149604398)

The best selection of wavelet filter and transform basis has been proposed for face recognition [13]. Similar to [13], we always select the low frequency subimage for further decomposition as it is less sensitive to vary images. The two-level wavelet lowest frequency subimage is used for PCA training in this paper. For example, an original image resolution of 128x128, size of the sub-image is reduced by 16 times. Figure 2 shows that a two-level decomposition using Daubechies 9/7 filter, the length of feature vector is $128 \times 128/2^{2\times 2} = 1024$.

FIGURE 2. An original image with resolution 128×128, and the two-level wavelet decomposition

3. Subspace Analysis

3.1. Principle Component Analysis (PCA)

The Principle Component Analysis (PCA) is a popular and powerful tool to reduce the dimensionality of the given data under unsupervised settings [7]. Consider a set of L feature vectors $\{\vec{g}_1, \vec{g}_2, \cdots, \vec{g}_L\}$ taking values in an n dimensional feature space. $\Sigma_{\vec{g}} \in R^{nxn}$ is defined as the covariance matrix of the feature vector $\vec{g}$: $\Sigma_g = \{[\vec{g} - \varepsilon(\vec{g})][\vec{g} - \varepsilon(\vec{g})]^T\}$, where ε is the expectation operator. Then PCA of a random vector $\vec{g}$ factorizes its covariance matrix $\Sigma_{\vec{g}}$ into the following form:

$$\Sigma_{\vec{g}} = \Phi\Lambda\Phi \text{ with } \Phi = [\vec{\varphi}_1, \vec{\varphi}_2, \cdots, \vec{\varphi}_n] \tag{3.1}$$

$$\Lambda = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\} \tag{3.2}$$

where $\Phi \in R^{nxn}$ is an orthogonal eigenvector matrix and $\Lambda \in R^{nxn}$ is a diagonal eigenvalues matrix with diagonal elements in decreasing order ($\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n$). The projection can be obtained via $\vec{y} = T_{pca}^T[\vec{g} - \varepsilon(\vec{g})]$, where eigenvectors $T_{pca} = [\vec{\varphi}_1, \vec{\varphi}_2, \cdots, \vec{\varphi}_p]$, $p < n$ and $T \in R^{nxp}$. The lower dimensional vector captures the most expressive features of the original data.

3.2. Linear Discriminant Analysis (LDA)

Different from the unsupervised learning of PCA, LDA seeks to find a linear transformation that maximizes the between-class scatter and minimizes the within-class scatter on the given data and class labels. Empirically, LDA has shown better performance than PCA for dimension reduction [14].

LDA training is carried out via scatter matrix analysis. Assume each face image belongs to c-class, the within-and between-class scatter matrices $\Sigma_\omega \in R^{pxp}, \Sigma_b \in R^{pxp}$ are computed as follows:

$$\Sigma_\omega = \sum_{i=1}^{c} P_r(\omega_i)\varepsilon\{(\vec{y} - \vec{m}_i)(\vec{y} - \vec{m}_i)^T | \omega = \omega_i\} \tag{3.3}$$

$$\Sigma_b = \sum_{i=1}^{c} P_r(\omega_i)(\vec{m}_i - \vec{m}_0)(\vec{m}_i - \vec{m}_0)^T \tag{3.4}$$

where $P_r(\omega_i)$ is the priori class probability and usually is replaced by $1/c$, and $\vec{m}_i$ is mean vector, $\vec{m}_0$ overall mean vector. Various measures are available for quantifying the discriminatory power. One commonly used is:

$$J(A) = \arg\max_A \frac{A\Sigma_b A^T}{A\Sigma_\omega A^T} \tag{3.5}$$

where A is an $m \times n$ matrix with $(m \leq n)$. The advantage of using this ratio is that it has been proven that if Σ_ω is a non-singular matrix then this ratio is maximized when $\Sigma_b A^* = \Sigma_\omega A^* \Lambda_A$ and A^* is the eigenvector matrix of $\Sigma_\omega^{-1}\Sigma_b$, Λ_A is a diagonal matrix. There are at most $c-1$ nonzero generalized eigenvectors. Let T_{fld} denote the reduction matrix of A^*, the reduced feature vectors for classification can then be derived as $\vec{u} = T_{fld}^T \vec{y}$.

4. Relevance Feedback

Relevance feedback [5] is a technique that takes advantage of human-computer interaction to refine high level queries represented by low level features. It is used to incorporate user's concept with the learning process [15, 16] for CBIR.

In our face retrieval system, after a user submits a query by a given sample, the system will return a set of similar images to the user. The returned images may be relevant (positive) or irrelevant (negative) to the user's target. Thus, relevance feedback is engaged as a query refinement technique for helping the retrieval task. The relevance feedback mechanism solicits the user to mark the relevance on the retrieved images and then refines the results by learning the feedbacks by the user.

In order to learn the user's feedback effectively, we employ a popular yet powerful machine learning technique, i.e., Support Vector Machine, to attack the problem. SVM is a state-of-the-art classification technique with excellent generalization performance. It has been shown with successful applications in relevance feedback [11, 17, 18].

5. Support Vector Machine

Support Vector Machine (SVM) [9] is a popular technique for classification. The basic idea of SVM is to look for the optimal separating hyperplane (OSH) which best separates the data points into two classes with a maximum margin in a projected feature space based on the Structure Risk Minimization principle.

Consider the problem of separating the set of training vectors into two separated classes, $(\vec{x}_1, y_1), ...(\vec{x}_N, y_N)$, where $\vec{x}_i \in R^d$,$y_i \in \{-1, +1\}$ with a hyperplane $\vec{w} \cdot \vec{x} + b = 0$. The set of vectors is said to be optimally separated by the hyper-

plane if it is separated without error and the margin is maximal. For all $\vec{x}_i$, if the following constraints are satisfied,

$$y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1, \quad i = 1, \cdots, N \tag{5.1}$$

the distance from the closest point to the hyperplane is $1/\|\vec{w}\|$. Hence, the hyperplane that optimally separates the data is the one that minimizes the objective function:

$$\Phi(\vec{w}) = \frac{1}{2}\|\vec{w}\|^2 = \frac{1}{2}(\vec{w} \cdot \vec{w}) \tag{5.2}$$

The solution to the optimization problem of (5.2) under the constraints of (5.1) is given by the saddle point of the Lagrange functional:

$$L(\vec{w}, b, \alpha) = \frac{1}{2}\|\vec{w}\|^2 - \sum_{i=1}^{N} \alpha_i \{y_i [(\vec{w} \cdot \vec{x}_i) + b] - 1\} \tag{5.3}$$

where α_i are the Lagrange multipliers. The parameters can be found by solving the following quadratic programming problem:

$$\max L(\alpha) = \sum_{i-1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \vec{x}_i \cdot \vec{x}_j \tag{5.4}$$

subject to:

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \text{ and } \alpha_i \geq 0 \tag{5.5}$$

The solution to $\vec{w}$ can be expressed $\vec{w} = \sum_i \alpha_i y_i \vec{x}_i$ in terms of a subset of training patterns, called support vectors, which lie on the margin. The decision function can thus be written as

$$f(\vec{x}) = sign(\sum\nolimits_i \alpha_i y_i \vec{x}_i \cdot \vec{x} + b) \tag{5.6}$$

So far the discussion has been restricted to the case where the training data is linearly separable. To generalize the OSH to the non-separable case, SVM introduces slack variables and a penalty factor such that the objective function can be modified as

$$\Phi(\vec{w}, \xi) = \frac{1}{2}(\vec{w} \cdot \vec{w}) + C(\sum_{i=1}^{N} \xi_i) \tag{5.7}$$

The input data can also be mapped through some nonlinear mapping into a high-dimensional feature space in which the OSH is constructed. Thus the dot production can be represented by $k(\vec{x}_i, \vec{x})$ when the kernel k satisfy Mercer's condition [19]. Table 1 shows three typical kernel functions [20]. Finally, we obtain the decision function

TABLE 1. Types of kernel functions

Kernel function	$k\left(\vec{x}_i, \vec{x}\right), i = 1, 2, \cdots N$
Polynomial	$\left(\gamma \vec{x}_i^T \vec{x} + r\right)^d, \gamma > 0$
Radial basis function (RBF)	$\exp\left(-\gamma \left\|\vec{x}_i - \vec{x}\right\|^2\right), \gamma > 0$
Sigmoid	$\tanh\left(\gamma \vec{x}_i^T \vec{x} + r\right)$

γ, d and r are kernel parameters.

$$f(\vec{x}) = sign(\Big(\sum\nolimits_i \alpha_i y_i \cdot k(\vec{x}_i \cdot \vec{x}) + b\Big) \tag{5.8}$$

Because SVM enjoys solid theoretical foundations from the perspective of statistical learning theory, it has shown many advantages when applied to problems with limited training samples. Consequently, it can achieve excellent empirical performance when applied to real-world problems.

6. Experimental Result

6.1. Experimental Setup

To evaluate the performance of our face retrieval system, we perform experiment using face dataset from ORL [21]. The ORL dataset consists of 400 frontal face images corresponding to 40 subjects, each subject with 10 different images. The size of each image is 92x112 pixels, with 256 grey levels per pixel, as shown in Figure 3.

FIGURE 3. Example of face images used for retrieval

6.2. Implementation Details

In our experiments, we use 20 classes each with 5 images as the training set, all images are decomposed by LFWT, the lowest subband forms facial feature vectors (Dimension=644). Afterwards, we process these facial feature vectors with PCA and LDA to further reduce the dimension to 19 for retrieval task. To present face retrieval approach, the metric of evaluation is taking average retrieval accuracy which is defined as the average ratio of the number of relevant face images in top N returns over the total number of relevant face images in the face dataset.

In the retrieval process, each face image is assumed as user's query. The system will measure it with all the remaining face images in the dataset, the returned face images are ranked by their distances from the boundary of SVM classifier. If the result is non-satisfactory, user is suggested to mark the returned images as relevant face images (+1) or irrelevant face images (-1) in top N, then perform further rounds of feedback. SVM uses both relevant (+1) and irrelevant (-1) face images as the training data to construct a more informative classifier, thus to predict more relevant face images from the dataset.

To implement the relevance feedback algorithm with SVM, we modify the codes of public *libsvm* library [22].

6.3. Average Retrieval Accuracy

The whole process is repeated for 400 times to generate the average retrieval accuracy. The Euclidean distance method is viewed as the baseline in this experiment. From Figure 4, it can be seen that our proposed LFWT+LDA+SVM (kernel is Radial basis function) method outperforms LFWT+LDA+Euclidean distance method. Moreover, with only one round feedback, the accuracy can reach to 99.6%.

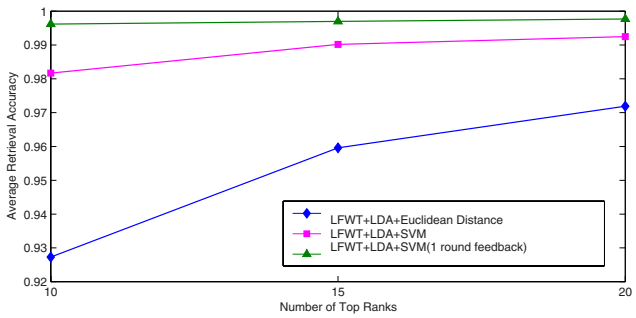


FIGURE 4. Retrieval performance of different methods on ORL dataset

In order to demonstrate our proposed scheme is effective and robust, we make a comparison with recent face recognition methods which use ORL dataset for performance evaluation, as shown in Table 2.

TABLE 2. Performance of different methods on ORL dataset

Methods	**Eigenvectors**	**Recognition Rate**
LFWT+LDA+NN	19	92.7%
PCA+NFL [10]	40	96.9%
LFWT+LDA+NFS [3]	60	96.1%
LFWT+LDA+SVM	**19**	**98.2%**

We observe that our proposed method has a lower-dimensional feature vectors and better recognition rate. In fact, average retrieval accuracy in top 10 on ORL dataset can be viewed as recognition rate. Furthermore, with relevance feedback involved, the system can achieve 100% accuracy. Therefore, we can summarize that our proposed scheme is promising for the challenging face retrieval task in a larger face database. Our experiment is implemented in a personal computer with Pentium 4 1.6GHz CPU with 256MB RAM.

7. Conclusions

This paper proposed a novel scheme for face retrieval using the lifting wavelets features and presented an effective relevance feedback algorithm using support vector machine for improving the retrieval performance. The LFWT is used to extract facial features because it increases the discriminability meanwhile decreases the dimensionality of the features. Moreover, it is faster than other wavelets transform approaches as it computes the wavelets transform in real domain and others in complex domain. In our experiment, we observed that the dimension of the feature vectors and the numbers of training samples used for SVM learning are two important factors to impact the overall performance of our face retrieval system. Further, we noted that the retrieval accuracy can be significantly improved through learning user's interaction with relevance feedback. This makes our scheme more effective for the applications with larger face databases. Experimental results have shown that our proposed scheme enjoys advantages in speed and retrieval performance, which makes it promising for practical face retrieval applications.

References

[1] Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R, "Content-based image retrieval at the end of the year," *IEEE Trans. Pattern Analysis and Machine Intelligence* vol. 22, pp. 1349-1380, 2000

[2] R. W. Chellappa, C.L. and Sirohey, S. "Human and machine recognition of faces: a survey," *Proc. of IEEE* 83 (1995) 705-741

[3] J.-T. Chien and C.-C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, pp. 1644-1649, 2002

[4] R. Foltyniewicz, "Automatic face recognition via wavelets and mathematical morphology," *Proc. Int'l Conf. Pattern Recognition*, pp. 13-17,1996.

[5] Y.Rui, T.S.Huang, et al., "Relevance feedback: A power tool for interactive content-based image retrieval", *IEEE trans. Circuits and video technolog.*, 8(5):644-655, Sep.1998

[6] Daubechies, I. and W. Sweldens. "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.,* Vol. 4, Nr. 3, 1998.

[7] M.A. Turk and A.P. Pentland, "Face recognition using eigenfaces," presented at *Computer Vision and pattern Recognition*, Proceedings CVPR'91, pp. 586-591, 1991

[8] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. PAMI*, vol. 19, pp. 711-720, 1997.

[9] V. Vapnik, The Nature of Statistical Learning Theory, *Springer-Verlag*, New York, 1995

[10] S. Z. Li and J. Lu, "Face recognition using the nearest feature line method," Neural Networks, IEEE Transactions on, vol. 10, pp. 439-443, 1999.

[11] P.Hong, Q.Tian, and T.S.Huang, "Incorporate support vector machines to content-based image retrieval with relevant feedback," *IEEE Proc. Int'l Conf. Image Processing (ICIP'00),* pp.750-753, 2000

[12] Simon Tong and Edward Chang, "Support vector machine active learning for image retrieval," *Proc. ACM Multimedia (MM'01),* pp.107-118, 2001

[13] A.Z.Kouzani, F.He, and K.Sammut, "Wavelet packet face representation and recognition," *Proc. IEEE Conf.* Systems, Man, and Cybernetics, pp.1614-1619,1997

[14] A.M. Martinez and A.C. Kak, "PCA versus LDA," PAMI, *IEEE Trans. on*, vol. 23, pp. 228-233, 2001

[15] Z. Su, H.J. Zhang, S. Li, and S.P. Ma, "Relevance Feedback in content-based Image Retrieval: Bayesian Framework, Feature Subspaces, and Progressing Learning," *IEEE Trans. Image Processing,* Vol. 12, No. 8, pp. 924-937, 2003.

[16] Y. Rui, T.S. Huang, and S. Mehrotra, "Content-based Image Retrieval with Relevance Feedback in MARS," Proc. of the International Conf. on Image Processing, pp. 815-818, 1997.

[17] C.-H. Hoi and M.R. Lyu., "Group-based relevance feedback with support vector machine ensembles", *Proc. Int'l Conf. Pattern Recognition*, (ICPR'04), Cambridge, UK, pp. 874-877, 2004

[18] C.-H. Hoi and M.R. Lyu., "A novel log-based relevance feedback technique in content-based image retrieval", *Proc. ACM Multimedia (MM'04)*, pp. 24-31, 2004

[19] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery*, 2(2) pp. 1-47, 1998.

[20] M.O.Stitson, J.A.E.Weston, et al., "Theory of Support Vector Machine", Technical report, CSD-TR-96-17, Univ. of London

[21] Olivetti Oracle Research Lab (ORL), "ORL website: http://mambo.ucsc.edu/psl/olivetti.html."

[22] C.-C. Chang and C.-J. Lin, LIBSVM : a library for support vector machine, 2001. Software available at http://www.csie.ntu.edu.tw/∼cjlin/libsvm/

Chon Fong Wong
Department of Electrical and Electronics Engineering
Faculty of Science and Technology
University of Macau
Macau SAR, PR. China
e-mail: planks@macau.ctm.net

Jianke Zhu
Room 101, Floor 1, Ho Sin-Hang Engineering Building
Compupter Science and Engineering Department
C.U.H.K.,Shatin,
Hong Kong PR. China
e-mail: jkzhu@cse.cuhk.edu.hk

Mang I Vai
Department of Electrical and Electronics Engineering
Faculty of Science and Technology
University of Macau
Macau SAR, PR. China
e-mail: fstmiv@umac.mo

Peng Un Mak
Department of Electrical and Electronics Engineering
Faculty of Science and Technology
University of Macau
Macau SAR, PR. China
e-mail: fstpum@umac.mo

Weikou Ye
Department of Physics
Shaoxing University
Shaoxing, PR. China
e-mail: wlywg@mail.sxptt.zj.cn

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 489–502
© 2006 Birkhäuser Verlag Basel/Switzerland

High-Resolution Image Reconstruction Using Wavelet Lifting Scheme

Shengwei Pei, Haiyan Feng and Minghui Du

Abstract. High-resolution image reconstruction refers to reconstruction of high-resolution images from multiple low-resolution, shifted, blurred samples of a true image. By expressing the true image as a square integrable function, Point Spread Function (PSF) can be used to construct biorthogonal wavelet filters directly and some algorithms for high-resolution image reconstruction were proposed based on the filters. However, the filters are the piecewise linear spline and corresponding primal and dual wavelet functions are all one vanishing moments. In order to improve the quality of reconstructed high-resolution images, we propose a method in this paper which can increase the numbers of vanishing moments of the wavelet functions so as to improve the performance of the biorthogonal filters using wavelet lifting scheme. Experiment results show that the method can improve the quality of reconstructed high-resolution images effectively. Also, we derive a fast algorithm that can reconstruct high-resolution images efficiently when blurring matrix is block-circulant-circulant-block (BCCB) matrix or Toeplitze-plus-Hankel system with Toeplitze-plus-Hankel block (THTH) matrix.

Mathematics Subject Classification (2000). Primary 65T60; Secondary 68U10.

Keywords. high-resolution reconstruction, super-resolution reconstruction, wavelet, lifting scheme, image processing.

1. Introduction

In some special scenes, for example, satellite imaging and some military infrared imaging, high-resolution digital images are necessary. Due to the limitations of digital camera, such as integrated circuit technology, cost, etc., usually desired high-resolution images are not available. In order to increase the resolution of

This research is supported by Ministry of Education, Government of Guangdong Province (Research Grant No.8303069).

images, the information from low-resolution images is collected and used to reconstruct the desired high-resolution images. This process is called high-resolution image reconstruction. Recently, high-resolution and super-resolution image reconstruction have been one of the most active research fields. Several methods for high-resolution image or super-resolution image reconstruction were proposed [1, 2, 3, 4, 5, 6]. Tsai and Huang [1] gave an approach with Fourier Transform to reconstruct super-resolution images. Nguyen and Milanfar [2] proposed an efficient method that considered sampling lattice and employed wavelet interpolation to reconstruct high-resolution images. A regularized method based on multi-channel sampling was proposed by Hong [3]. ML (Maximum Likelihood) method was proposed by Tom in [4]. At present, the MAP (Maximum a-Posteriori) method [5, 6] is researched widely and has good performance in the quality of reconstructed images, but the method is very complicated. Projection Onto Convex Set (POCS) method was proposed by H. Stark in [7], which described an iterative approach to incorporate prior knowledge into the reconstruction process.

Here we follow the research work in Bose and Boo [8] and Chan [9] and consider reconstructing high-resolution images from the low-resolution images of the same scene by wavelet lifting scheme. By expressing a true image as a square integrable function [9], Point Spread Function (PSF) can be used to construct biorthogonal filters directly. Some algorithms for high-resolution image reconstruction [9, 10] were proposed based on the filters. However, the filters are the piecewise linear spline and the corresponding primal and dual wavelet functions are all one vanishing moments [11, 15]. The quality of reconstructed images is related tightly with the numbers of vanishing moment of wavelet functions. Thus, in order to get better quality of reconstructed images, wavelet lifting scheme can be employed to increase the numbers of vanishing moment of wavelet functions so as to improve the performance of the biorthogonal wavelet filters. In this paper, we analyze the high-resolution reconstruction approach from the wavelet and multiresolution point of view [9], then, based on wavelet lifting scheme, we propose a method for high-resolution image reconstruction. Also, a fast algorithm for solving $Lf = g$ is derived when L is BCCB or THTH matrix.

This paper is arranged as follows: In section 2, a mathematical model of high-resolution image reconstruction is given. In section 3, we give the algorithm proposed by Chan [9] and then derive our algorithm based on the wavelet lifting scheme. Numerical experiment results are given in section 4 to illustrate the effectiveness of the method. The conclusion about the algorithm will be shown in section 5.

2. Mathematical Model

Here we give a brief introduction to the mathematical model of the high-resolution image reconstruction problem. Details about the model can be found in [8]. Suppose that the imaging sensor array has $N_1 \times N_2$ pixels. Let the actual length

and width of each pixel be T_1 and T_2 respectively. These sensors are called low-resolution sensors. Our aim is to construct a high-resolution image of the same scene by using an array of $K_1 \times K_2$ low-resolution sensors. More precisely, we want to reconstruct an image with $M_1 \times M_2$ pixels, where $M_1 = K_1 N_1$ and $M_2 = K_2 N_2$. So, the length and width of each of these high-resolution pixels will be T_1/K_1 and T_2/K_2, respectively. Here, Let $K_1 = K_2 = K$ and $f(x_1, x_2)$ be the intensity of the scene at any point. By reconstructing the high-resolution image, we mean to find or approximate the values:

$$\frac{K^2}{T_1 T_2} \int_{iT_1/K}^{(i+1)T_1/K} \int_{jT_2/K}^{(j+1)T_2/K} f(x_1, x_2) dx_1 dx_2; \quad 0 \le i \le M_1, 0 \le j \le M_2.$$

In order to get enough information to resolve the high-resolution image, there must be subpixel displacements between the sensors in the sensor arrays. Ideally, for sensor $(k_1, k_2), 0 \le k_1, k_2 \le K$, its horizontal and vertical displacements $d^x_{k_1,k_2}$ and $d^y_{k_1,k_2}$ with respect to the point (0, 0) are given by:

$$d^x_{k_1,k_2} = (k_1 + \frac{1-K}{2})\frac{T_1}{K}; \quad d^y_{k_1,k_2} = (k_2 + \frac{1-K}{2})\frac{T_2}{K}.$$

For this low-resolution sensor, the average intensity registered at its (n_1, n_2)th pixel is modeled by:

$$g_{k_1,k_2}[n_1, n_2] = \frac{1}{T_1 T_2} \int_{T_1 n_1 + d^x_{k_1,k_2}}^{T_1(n_1+1)+d^x_{k_1,k_2}} \int_{T_2 n_2 + d^y_{k_1,k_2}}^{T_2(n_2+1)+d^y_{k_1,k_2}} f(x_1, x_2) dx_1 dx_2 + \eta_{k_1,k_2}[n_1, n_2]. \quad (2.1)$$

According to (2.1), if we intersperse all the low-resolution images g_{k_1,k_2} to form an $M_1 \times M_2$ image g by assigning:

$$g[Kn_1 + k_1, Kn_2 + k_2] = g_{k_1,k_2}[n_1, n_2].$$

then g is the approximate of the f. Here, g can be called observed high-resolution image. In order to get an image that is better than the observed high-resolution image, we have to get f.

In (2.1), we can see that $g_{0,0}[0, 0]$ involves the points outside the region of image we have known, so the system is underdetermined. To compensate for this, one imposes the boundary conditions on f for x_i outside the domain. Usually, periodic boundary condition is a common one, but the symmetric boundary condition and zero boundary condition can also be imposed. The symmetric boundary condition is a new one and has good performance in boundary artifacts. In the following text, we will give algorithms for periodic and symmetric boundary conditions respectively.

Using any above boundary condition and ordering the discretized values of f and g in a row-by-row fashion, we obtain an $M_1 M_2 \times M_1 M_2$ linear system as:

$$Lf = g. \quad (2.2)$$

where L is the convolution operator (blurring matrix), g is a vector formed from the low-resolution images and f is the desired high-resolution image.

According to the separable property of Point Spread function (PSF), we can get the following equation:

$$L = L^x \otimes L^y. \tag{2.3}$$

where $\otimes$ is the Kronecker tensor product. For periodic and symmetric boundary conditions, the system (2.2) is ill-conditioned and susceptible to noise.

3. Reconstruction

3.1. The Basic Principle

For (2.1), we can consider the process as a lowpass filter imposed on image f. So $L_2^p, L_4^p, L_2^s, L_4^s$ can also be considered as lowpass filters. L_2^p, L_4^p, L_2^s and L_4^s denote the L in (2.2) under the condition of periodic and symmetric boundary conditions for 2×2 sensor arrays and 4×4 sensor arrays respectively. Here, we recall the scaling function $f(x_1, x_2) \in L_2(R^2)$ in wavelet transform:

$$\phi = 4 \sum_{\alpha \in Z} a(\alpha)\phi(2 \cdot -\alpha).$$

where sequence $a(\alpha)$ is a refinable mask (also known as a scaling coefficient or a lowpass filter).

According to the blurring matrix L given in (2.3), we can obtain the refinable mask for 2×2 sensor arrays from PSF directly [9]. It is listed as follows: $m_2(\alpha) = 1/4, 1/2, 1/4; \alpha = -1, 0, 1$. For 4×4 sensor arrays, the refinable mask is the stable refinable function ϕ with dilation 4. It is listed as follows: $m_4(\alpha) = 1/8, 1/4, 1/4, 1/4, 1/8; \alpha = -2, \cdots, 2$. We can see that the masks (filters) are all piecewise linear splines and Deslauriers-Dubuc filters [11] and their corresponding scaling functions are Deslauriers-Dubuc interpolating scaling functions [15, 16].

The refinable masks, wavelet masks and their duals can be found in [8]. The tensor product dual pairs of the refinement symbols are given by $\hat{a}(\omega) = \hat{m}(\omega_1)\hat{m}(\omega_2)$ and $\hat{a}^d(\omega) = \hat{m}^d(\omega_1)\hat{m}^d(\omega_2)$. The corresponding wavelet symbols b_v and b_v^d can also be gotten (See [9]).

Let $f \in S^1(\phi^d)$, then:

$$\begin{aligned} f &= \sum_{\alpha \in Z^2} \langle f, 2\phi(2 \cdot -\alpha)\rangle 2\phi^d(2 \cdot -\alpha) \\ &= 2 \sum_{\alpha \in Z^2} v(\alpha)\phi^d(2 \cdot -\alpha). \end{aligned} \tag{3.1}$$

where $v(\alpha), \alpha \in Z^2$ are the pixel values of the high-resolution image we are seeking, and they form the discrete representation of f under the basis $2\phi^d(2 \cdot -\alpha), \alpha \in Z^2$. Then $(a * v)(\alpha)$ is the observed high-resolution image. According to biorthogonal

wavelet theory, we know that refinement masks and wavelet masks satisfy the following equation [9, 12]:

$$\bar{\hat{a}}^d\hat{a} + \sum_{v\in Z^2\setminus\{(0,0)\}} \bar{\hat{b}}_v^d\hat{b}_v = 1.$$

So, according to the above equation, an iterative algorithm for high-resolution image reconstruction is obtained as follows:

$$\hat{v}_{n+1} = \bar{\hat{a}}^d\hat{a}\hat{v} + (1-\beta)(\sum_{v\in Z_2^2\setminus\{(0,0)\}} \bar{\hat{b}}_v^d\hat{b}_v)\hat{v}_n. \tag{3.2}$$

where factor $(1-\beta)$ is used to denoise in each iterative cycle. The equation (3.2) can be written as follows using matrices:

$$(L^dL + \beta/(1-\beta)I)f = 1/(1-\beta)L^dg. \tag{3.3}$$

where L, L^d are the matrices generated by $\hat{a}, \bar{\hat{a}}^d$ respectively.

3.2. Wavelet Lifting

According to the descriptions above, we can get analysis and synthesis filters m, m^d, r and r^d directly from PSF [9]. After simple computation, we know that the vanishing moment of the wavelet functions corresponding to the filters are all ones. Since the quality of recovered images is related tightly with the numbers of vanishing moment, larger numbers of vanishing moment of the wavelet functions are desirable [13]. Lifting is an elementary modification of perfect reconstruction filters, which is used to improve the wavelet properties such as vanishing moment. In order to improve the quality of reconstructed images, we can use lifting scheme to increase the numbers of vanishing moments of wavelet functions corresponding to r, r^d. Here, we use primal lifting and dual lifting to modify m, r, m^d and r^d in turn. For 2×2 sensor arrays, according to the lifting theory [14], we have:

$$\begin{aligned} \hat{m}^l(\omega) &= \hat{m}(\omega) + \hat{r}(\omega)\hat{l}_l^*(2\omega); \\ \hat{r}^{dl}(\omega) &= \hat{r}^d(\omega) - \hat{m}^d(\omega)\hat{l}_l(2\omega); \end{aligned} \tag{3.4}$$

$$\begin{aligned} \hat{m}^{dl}(\omega) &= \hat{m}^d(\omega) - \hat{r}^d(\omega)\hat{l}_d(2\omega); \\ \hat{r}^l(\omega) &= \hat{r}(\omega) + \hat{m}(\omega)\hat{l}_d^*(2\omega); \end{aligned} \tag{3.5}$$

where $\hat{m}^l(\omega), \hat{r}^{dl}(\omega), \hat{m}^{dl}(\omega)$ and $\hat{r}^l(\omega)$ are the Fourier transforms of lifted m, r^d, m^d, r respectively; $\hat{m}^d(\omega), \hat{r}^d(\omega), \hat{r}(\omega), \hat{m}(\omega)$ are the Fourier transforms of m^d, r^d, r, m respectively; $\hat{l}_l(\omega), \hat{l}_d(\omega)$ are Fourier transforms of two FIR filters corresponding to primal lifting and dual lifting respectively, and $\hat{l}_l^*(\omega), \hat{l}_d^*(\omega)$ are their complex conjugates. Here we want to create two vanishing moments wavelet functions corresponding to r^d and r. After computation, the shortest filters l_l and l_d have Fourier transforms:

$$\hat{l}_l(2\omega) = \frac{1}{8}e^{iw}(\cos(\omega) - 1). \tag{3.6}$$

$$\hat{l}_d(2\omega) = -\frac{3}{8}e^{iw}(\cos(\omega) - 1). \tag{3.7}$$

Inserting (3.6) and (3.7) into (3.4) and (3.5) respectively, gives:

$$m_2^l(\alpha) = \frac{1}{128}, 0, \frac{23}{128}, \frac{5}{8}, \frac{23}{128}, 0, \frac{1}{128}; \quad \alpha = -3, -2, \cdots, 3.$$

$$m_2^{dl}(\alpha) = -\frac{5}{64}, \frac{1}{16}, \frac{33}{32}, \frac{1}{16}, -\frac{5}{64}; \quad \alpha = -2, -1, \cdots, 2.$$

For 4×4 sensor arrays, the PSF can be considered as a refinable function with dilation 4. For filters m_4, m_4^l, their lifted filters m_4^l, m_4^{dl} are all symmetric and can be written as follows [16, 17, 18]:

$$m_4^l(\alpha) = -\frac{1}{3072}, -\frac{1}{3072}, \frac{7}{3072}, \frac{47}{384}, \frac{389}{1536}, \frac{377}{1536}, \frac{389}{1536}, \frac{47}{384}, \frac{7}{3072}, -\frac{1}{3072}, -\frac{1}{3072};$$

$$m_4^{dl}(\alpha) = -\frac{1}{512}, -\frac{29}{512}, \frac{15}{128}, \frac{165}{512}, \frac{61}{256}, \frac{165}{512}, \frac{15}{128}, -\frac{29}{512}, -\frac{1}{512}.$$

Now, we obtain the lifted filters.

3.3. Reconstruction Computation

We assume that the PSF is symmetric and the size of high-resolution images is $M \times M$. For L and L^d in (3.3), $L = L^x \otimes L^y, L^d = L^{d,x} \otimes L^{d,y}$. For 2×2 sensor arrays with periodic boundary condition, $L^x, L^y, L^{d,x}$ and $L^{d,y}$ can be expressed as $L_2^{p,x}, L_2^{p,y}, L_2^{p,d,x}, L_2^{p,d,y}$. Since PSF is symmetric and all the images are square, we have $L_2^{p,x} = L_2^{p,y}$ and $L_2^{p,d,x} = L_2^{p,d,y}$. For 2×2 sensor arrays with symmetric boundary condition, we can also get $L_2^{s,x}, L_2^{s,y}, L_2^{s,d,x}$ and $L_2^{s,d,y}$. Similarly, we have $L_2^{s,x} = L_2^{s,y}, L_2^{s,d,x} = L_2^{s,d,y}$. For 4×4 sensor arrays, similarly, $L_4^{p,x}, L_4^{p,y}, L_4^{p,d,x}, L_4^{p,d,y}$ and $L_4^{s,x}, L_4^{s,y}, L_4^{s,d,x}, L_4^{s,d,y}$ can also be gotten for periodic boundary condition and symmetric boundary condition respectively. So, according to equation (2.3), the following equations hold:

$$\begin{aligned}
L_2^p &= L_2^{p,x} \otimes L_2^{p,y}; \qquad L_2^{p,d} = L_2^{p,d,x} \otimes L_2^{p,d,y};\\
L_4^p &= L_4^{p,x} \otimes L_4^{p,y}; \qquad L_4^{p,d} = L_4^{p,d,x} \otimes L_4^{p,d,y};\\
L_2^s &= L_2^{s,x} \otimes L_2^{s,y}; \qquad L_2^{s,d} = L_2^{s,d,x} \otimes L_2^{s,d,y};\\
L_4^s &= L_4^{s,x} \otimes L_4^{s,y}; \qquad L_4^{s,d} = L_4^{s,d,x} \otimes L_4^{s,d,y}.
\end{aligned} \tag{3.8}$$

According to the assumptions above, $L_2^{p,x}, L_2^{p,d,x}, L_2^{s,x}, L_2^{s,d,x}, L_4^{p,x}, L_4^{p,d,x}, L_4^{s,x}$ and $L_4^{s,d,x}$ are all $M \times M$ matrices. We now give $L_2^{p,x}$ and $L_2^{s,x}$ as follows for an example:

$$L_2^{p,x} = \begin{bmatrix} \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} & & & & \frac{1}{128} & 0 & \frac{23}{128} \\ \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} & & & & \frac{1}{128} & 0 \\ 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} & & & & \frac{1}{128} \\ \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} & & & \\ & \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \frac{1}{128} \\ \frac{1}{128} & & & & \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 \\ 0 & \frac{1}{128} & & & & \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} \\ \frac{23}{128} & 0 & \frac{1}{128} & & & & \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} \end{bmatrix};$$

$$L_2^{s,x} = \begin{bmatrix} \frac{103}{128} & \frac{23}{128} & \frac{1}{128} & \frac{1}{128} & & & & & & \\ \frac{23}{128} & \frac{81}{128} & \frac{23}{128} & 0 & \frac{1}{128} & & & & & \\ \frac{1}{128} & \frac{23}{128} & \frac{5}{8} & \frac{23}{128} & 0 & \ddots & & & & \\ \frac{1}{128} & 0 & \frac{23}{128} & \frac{5}{8} & \ddots & \ddots & \frac{1}{128} & & & \\ & \frac{1}{128} & 0 & \ddots & \ddots & \frac{23}{128} & 0 & \frac{1}{128} & & \\ & & \ddots & \ddots & \ddots & \frac{5}{8} & \frac{23}{128} & \frac{1}{128} & & \\ & & & \frac{1}{128} & 0 & \frac{23}{128} & \frac{81}{128} & \frac{23}{128} & & \\ & & & & \frac{1}{128} & \frac{1}{128} & \frac{23}{128} & \frac{103}{128} & & \end{bmatrix}. \tag{3.9}$$

If the boundary condition is chosen to be periodic, then L and L^d in (3.3) are all block-circulant-circulant-block (BCCB) matrices and (3.3) can be solved by three FFTs in $O(M^2 \log(M^2))$ operations [19, section 5.1.3]. If the boundary condition is symmetric, then L and L^d in (3.3) are all block Toeplitze-plus-Hankel system with Toeplitze-plus-Hankel blocks (THTH) matrix and (3.3) can be solved by three 2D fast cosine transforms (FCTs) in $O(M^2 \log(M^2))$ operations [17].

3.4. Fast Algorithm for Solving $Lf = g$

In this subsection, we give a fast algorithm which can decrease computation complexity and speed up computation efficiently in solving (3.3). During computation, $L_2^{p,x}, L_2^{p,d,x}, L_4^{p,x}, L_4^{p,d,x}, L_2^{s,x}, L_2^{s,d,x}, L_4^{s,x}$ and $L_4^{s,d,x}$ can be stored in the memory of a computer in the manner of sparse matrices, but it will be complex and difficult to store L, L^d and $L \otimes L^d$ as sparse matrices. So, we derive a method from (3.3) for reducing memory occupancy and thus speeding up the computation. According to [19, 20], we know that $L_2^{p,x}, L_2^{p,d,x}, L_4^{p,x}, L_4^{p,d,x}, L_2^{s,x}, L_2^{s,d,x}, L_4^{s,x}$ and $L_4^{s,d,x}$ can be diagonalized by matrix A (Fourier transform or Cosine transform matrix) and its inverse matrix A^{-1}. Here, we take $L_2^{p,x}, L_2^{p,y}$ as an example. Assume that the diagonal matrix of $L_2^{p,x}$ is $D_2^{p,x}$ and the diagonal matrix of $L_2^{p,y}$ is $D_2^{p,y}$, then:

$$L_2^{p,x} = AD_2^{p,x}A^{-1}. \tag{3.10}$$

$$L_2^{p,y} = AD_2^{p,y}A^{-1}. \tag{3.11}$$

Assume that $A \otimes A = A_{2n}$, $A^{-1} \otimes A^{-1} = A_{2n}^{-1}$ and $D_2^p = D_2^{p,x} \otimes D_2^{p,y}$. According to (3.10) and (3.11) and assumptions above, L_2^p can be written:

$$\begin{aligned} L_2^p &= L_2^{p,x} \otimes L_2^{p,y} = AD_2^{p,x}A^{-1} \otimes AD_2^{p,y}A^{-1} \\ &= (A \otimes A)(D_2^{p,x}A^{-1} \otimes D_2^{p,y}A^{-1}) \\ &= (A \otimes A)(D_2^{p,x} \otimes D_2^{p,y})(A^{-1} \otimes A^{-1}) \\ &= A_{2n}D_2^pA_{2n}^{-1}. \end{aligned} \tag{3.12}$$

Similarly, following equation also holds:

$$L_2^{p,d} = A_{2n}D_2^{p,d}A_{2n}^{-1}. \tag{3.13}$$

According to (3.12), we know that the diagonal matrix of L_2^p is D_2^p. Similarly, the diagonal matrix of $L_2^{p,d}$ is $D_2^{p,d}$ and $D_2^{p,d} = D_2^{p,d,x} \otimes D_2^{p,d,y}$. So, (3.3) can be rewritten as follows:

$$f = (L_2^{p,d}L_2^p + \frac{\beta}{1-\beta}I)^{-1}(\frac{1}{1-\beta})L_2^{p,d}g. \tag{3.14}$$

As we know, $\beta/(1-\beta)I$ is a diagonal matrix, so following equation holds:

$$A_{2n}((\frac{\beta}{1-\beta})I)A_{2n}^{-1} = (\frac{\beta}{1-\beta})I$$

According to the descriptions above, we can get:

$$(L_2^{p,d}L_2^p + \frac{\beta}{1-\beta}I)^{-1} = A_{2n}((D_2^{p,d,x} \otimes D_2^{p,d,y})(D_2^{p,x} \otimes D_2^{p,y}) + (\frac{\beta}{1-\beta})I)^{-1}A_{2n}^{-1}. \tag{3.15}$$

Inserting (3.15), (3.13) into (3.14), gives:

$$f = A_{2n}(D_2^{p,d}D_2^p + \frac{\beta}{1-\beta}I)^{-1}(\frac{1}{1-\beta})D_2^{p,d}A_{2n}^{-1}g. \tag{3.16}$$

For any $MM \times 1$ matrix L, $M \times M$ matrix A, B and $P = A \otimes B$, following equation holds (See appendix for mathematical proof):

$$PL = S^{-1}(BS(L)A^T). \tag{3.17}$$

where S is an operator to order L into an $M \times M$ matrix in a row-by-row fashion, and S^{-1} is an operator to order $BS(L)A^T$ into an $MM \times 1$ matrix in a row-by-row fashion. So (3.16) can be rewritten as follows:

$$\begin{aligned} f &= S^{-1}(AS((D_2^{-p,d}D_2^{-p} + \\ &\quad (\frac{\beta}{1-\beta})^{-1}I)(\frac{1}{1-\beta})\ D_2^{p,d}S^{-1}(A^{-1}S(g)(A^{-1})^T))A^T). \end{aligned} \tag{3.18}$$

where $D_2^{-p}, D_2^{-p,d}$ are inverse matrices of D_2^p and $D_2^{p,d}$. From (3.18), we can see that it need not to construct $L_2^{p,d}$ and this will reduce memory occupancy and thus speed up computation greatly. According to the definition of FFT, IFFT, FCT and IFCT, (3.18) can be solved by two FFTs and two IFFTs in $O(M^2\log(M^2))$

operations when boundary condition is periodic or by two FCTs and two IFCTs when boundary condition is symmetric, but not three 2D FFTs or FCTs given in [9, 20]. Due to the decrease in memory occupancy, the computation can also be speeded up, especially when the size of images is large.

4. Numerical Experiments

Numerical experiments were conducted to test the algorithm proposed in this paper. We will illustrate effectiveness of the algorithm by the experiment results. In these experiments, high-resolution images are all size 256×256. For 2×2 sensor arrays, we use four low-resolution images of size 128×128 to conduct experiments. For 4×4 sensor arrays, sixteen low-resolution images of size 64×64 are used in test experiments. The peak signal-to-noise ratio (PSNR), which compares the reconstructed image $\mathbf{f}_c$ with the original image $\mathbf{f}$, is defined by:

$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2 NM}{\|\mathbf{f} - \mathbf{f}_c\|_2^2}.$$

We give the experiment results in Table 1-Table 4. In these tables, β^* means the optimal values of β; "biorthogonal" means the biorthogonal method proposed in [9]; "dual" means to only apply dual lifting to the reconstruction process, and "primal and dual" means to apply primal lifting and dual lifting to the reconstruction process at the same time. Table 1 is the experiment results for 2×2 sensor arrays when boundary conditions is periodic. Table 2 is the results for 2×2 sensor arrays with symmetric boundary condition. We can see that the "dual" and "primal and dual" methods improve the quality of reconstructed images significantly. Especially, the "primal and dual" is more efficient than "dual". We can also see that the method we proposed is more efficient for the images with greater SNR. Table 3 and Table 4 are the experiment results for 4×4 sensor arrays with periodic boundary condition and symmetric boundary condition respectively. From the two tables, we can see the same effect as it shown in Table 1 and Table 2. In comparison with 2×2 sensor arrays, the efficiency of the method for 4×4 sensor arrays is slightly reduced. From the Table 1 and the Table 2, we can also see that the method is susceptible to noise, namely, the method is more effective on the low-resolution images with greater SNR value. Visual experiment results are shown in Figure 1 and Figure 2.

5. Conclusion

In this paper, we proposed an algorithm for high-resolution image reconstruction based on wavelet lifting scheme. By the method, we improved the work of chan [9] greatly. The method can also be easily extended to super-resolution image reconstruction. Using the experiment results, we have shown that our method is effective in high-resolution image reconstruction. However, for 4×4 sensor arrays, performance of the method is worse than it for 2×2 sensor arrays. The experiment

TABLE 1. 2×2 sensor arrays with periodic boundary condition

		biorthogonal		dual		primal and dual	
image	SNR	PSNR	β^*	PSNR	β^*	PSNR	β^*
Boat	30db	34.89	0.013	34.991	0.016	35.11	0.021
	40db	41.40	0.017	41.43	0.018	41.516	0.019
Lena	30db	35.92	0.027	35.969	0.029	35.992	0.031
	40db	42.70	0.025	42.737	0.034	42.779	0.035

TABLE 2. 2×2 sensor arrays with symmetric boundary condition

		biorthogonal		dual		primal and dual	
image	SNR	PSNR	β^*	PSNR	β^*	PSNR	β^*
Boat	30db	36.048	0.012	36.44	0.039	36.598	0.22
	40db	44.874	0.0013	44.981	0.00027	45.265	0.031
Lena	30db	37.403	0.027	37.592	0.075	37.699	0.31
	40db	45.612	0.0001	45.75	0.00043	46.13	0.035

TABLE 3. 4×4 sensor arrays with periodic boundary condition

		biorthogonal		dual		primal and dual	
image	SNR	PSNR	β^*	PSNR	β^*	PSNR	β^*
Boat	30db	34.89	0.013	34.991	0.016	35.11	0.021
	40db	41.40	0.017	41.43	0.018	41.516	0.019
Lena	30db	35.92	0.027	35.969	0.029	35.992	0.031
	40db	42.70	0.025	42.737	0.034	42.779	0.035

TABLE 4. 4×4 sensor arrays with symmetric boundary condition

		biorthogonal		dual		primal and dual	
image	SNR	PSNR	β^*	PSNR	β^*	PSNR	β^*
Boat	30db	34.910	0.011	34.939	0.018	34.957	0.018
	40db	43.00	0.015	43.14	0.0145	45.21	0.02
Lena	30db	36.127	0.027	36.17	0.026	36.35	0.029
	40db	44.00	0.031	44.18	0.033	44.37	0.035

results also show that the method is susceptible to noise. At the same time, a fast algorithm was derived in this paper for solving $Lf = g$ under the periodic and symmetric boundary conditions.

(a) Original high-resolution "boat" image (256×256)

(b) Low-resolution image with 30db white noise is added (128×128)

(c) Reconstructed HR image by the method by "Chan" with periodic boundary condition

(d) Reconstructed HR image by "primal and dual lifting" method with periodic boundary condition

(e) Reconstructed HR image by the method by "Chan" with symmetric boundary condition

(f) Reconstructed HR image by "primal and dual lifting" method with symmetric boundary condition

FIGURE 1. Experiment results for 2×2 sensor arrays

Appendix

Assume that A, B are all $M \times M$ matrices and $P = A \otimes B$. Let g be an $MM \times 1$ arbitrary matrix, then following equation holds:

$$P \cdot g = S^{-1}(B \cdot S(g) \cdot A^T). \tag{A.1}$$

where S is an operator to order g to an $M \times M$ matrix in a row-by-row fashion, and S^{-1} is an operator to order an $M \times M$ matrix to an $MM \times 1$ matrix in a row-by-row fashion.

Proof. Assume that the (k, i)th element of A is $A(k, i)$. According to the property of Kronecker tensor product, P can be represented as an $M \times M$ block matrix, and its (m, n)th element can be represented as $A(m, n) \cdot B$. So we can write $Pg = f$

(a) Original high-resolution "boat" image (256×256)

(b) Low-resolution image with 30db white noise is added (64×64)

(c) Reconstructed HR image by the method by "Chan" with periodic boundary condition

(d) Reconstructed HR image by "primal and dual lifting" method with periodic boundary condition

(e) Reconstructed HR image by the method by "Chan" with symmetric boundary condition

(f) Reconstructed HR image by "primal and dual lifting" method with symmetric boundary condition

FIGURE 2. Experiment results for 4×4 sensor arrays

as:

$$\begin{bmatrix} A(1,1)\cdot B & A(1,2)\cdot B & \cdots & A(1,M)\cdot B \\ A(2,1)\cdot B & A(2,2)\cdot B & \cdots & A(2,M)\cdot B \\ \vdots & \vdots & \ldots & \vdots \\ A(M,1)\cdot B & A(M,2)\cdot B & \cdots & A(M,M)\cdot B \end{bmatrix} \begin{bmatrix} g_{1,1} \\ \vdots \\ g_{M,1} \\ g_{1,2} \\ \vdots \\ g_{M,M} \end{bmatrix} = \begin{bmatrix} f_{1,1} \\ \vdots \\ f_{M,1} \\ f_{1,2} \\ \vdots \\ f_{M,M} \end{bmatrix}$$

According to (A.1), the (m, j)th element of f can be represented by:

$$f_{m,j} = \sum_{k=1}^{M} A(j,k) \sum_{p=1}^{M} B(m,p) \cdot g_{p,k}. \tag{A.2}$$

Let $y = BGA^T$, where

$$G = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,M} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,M} \\ \vdots & \vdots & \cdots & \vdots \\ g_{M,1} & g_{M,2} & \cdots & g_{M,M} \end{bmatrix}.$$

According to the definition of y, the (m, j)th element of y is

$$y_{m,j} = \sum_{k=1}^{M} A(j,k) \sum_{p=1}^{M} B(m,p) \cdot g_{p,k}. \tag{A.3}$$

we can see that (A.3) has the same representation as (A.2), so $y_{m,j} = f_{m,j}$ and $P \cdot g = S^{-1}(B \cdot S(g) \cdot A^T)$.

□

Acknowledgment

Many thanks to Education Ministry of Guangdong Province, P.R. China.

References

[1] R.Y. Tsai and T.S. Huang, *Multipleframe image restoration and registration*, Advances in Computer Vision and Image Processing. **1**(1984),317-339.

[2] N. Nguyen and P. Milanfer, *An efficient wavelet-based algorithm for image super-resolution*, Proc. International Conference of Image Processing. 2000, 351-354.

[3] M.C. Hong, M.G. Kang, and A.K. Katsaggelos, *A regularized multi-channel restoration approach for globally optimal high-resolution video sequence*, SPIE VCIP, **3024** (1997), 1306-1317.

[4] B.C. Tom etc., *Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images*, Proc. 1995 IEEE Int. Conf. Image Processing, 1995, 539-542.

[5] R.R. Schulz and R.L. Stevenson, *Extraction of high-resolution frames from video sequence*, IEEE Trans. Image Processing, **5** (1996), 996-1011.

[6] Di zhang and Minghui Du, *Fast hybrid approach to large-magnification super-resolution image reconstruction*, Optical Engineering, SPIE, **44** (2003), 037005-1-9.

[7] H. Stark and P. Oskoui, *High resolution image recovery from image plane arrays, using convex projections*, J. Optical. Soc. AM. A, **6** (1989), 1715-1726.

[8] N. Bose and K. Boo, *High-resolution image reconstruction with multisensors*, International J. Imaging System and Technology, **9** (1998), 294-304.

[9] Raymond H. Chan and Tony F. Chan, etc., *Wavelet Algorithms for High-resolution Image Reconstruction*, SIAM, J. Sci. Comput., **24** (2003), 1408-1432.

[10] Raymond H. Chan, and Tony F. Chan, Lixin Shen and Zhouwei Shen, *Wavelet Deblurring Algorithm for Spatially Varying Blur from High-resolution Image Reconstruction*, Research report, CUHK-2000-20, Department of Mathematics, The Chinese University of HongKong, 2000.

[11] G.Deslauriers and S. Dubuc, *Interpolation dyadique, in "Fractals, dimensions non enti" res et applications*, Masson, Paris, 1987.

[12] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF Regional Conf. Ser. In Appl. Math., Vol.61. SIAM, 1992.

[13] David S. Taubman and Michael W. Marcelin, *JPEG 2000 Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2001.

[14] Stephane Mallat, *A Wavelet Tour of Signal Processing*, Elsevier Academic Press, 1999.

[15] G.Deslauriers and S. Dubuc, *Symmetric iterative interpolation processes*, Constr. Approx. **5** (1998),49-68.

[16] Wim Sweldens, *The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets*, Applied and Computational Harnonic Analysis, **3** (1996), 186-200.

[17] W. Lawton and S. L. Lee, and Z. Shen, *Stability and orthonormality of multivariate refinable functions*, SIAM J. Math. Anal., **28** (1997), 999-1014.

[18] Z. Shen, *Extension of matrices with Laurent polynomial entries*, proceedings of the 15th IMACS World Congress 1997 on Scientific Computation, Modeling and Applied Mathematics, A. Syclow, 1997, 57-61.

[19] R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1993.

[20] Michael K. Ng, Raymond H. Chan and Wun-Cheung Tang, *A Fast Algorithm for Deblurring Models with Neumann Boundary Conditions*, SIAM J. Sci. Comput., **21** (1999), 851-866.

Shengwei Pei
Research & Development Center,
China Academy of Space Technology,
Beijing,
P.R. China
e-mail: eeswpei@yahoo.com.cn, eeswpei@scut.edu.cn

Haiyan Feng
Division of Science and Technology.
South China University of Technology
Guangzhou, Guangdong Province, 510640
P.R. China
e-mail: adhyfeng@scut.edu.cn

Minghui Du
School of Electronic and Information Engineering.
South China University of Technology
Guangzhou, Guangdong Province, 510640
P.R. China
e-mail: ecmhdu@scut.edu.cn

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 503–513
© 2006 Birkhäuser Verlag Basel/Switzerland

Mulitiresolution Spatial Data Compression Using Lifting Scheme

B. Pradhan, K. Sandeep, Shattri Mansor,
Abdul Rahman Ramli and Abdul Rashid B. Mohamed Sharif

Abstract. In many applications referring to terrain visualization, there is need to visualize terrains at different levels of detail. The terrain should be visualized at different levels for different parts; for example, a region of high interest should be in a higher resolution than a region of low or no interest. The lifting scheme has been found to be a flexible method for constructing scalar wavelets with desirable properties. In this paper, it is extended to the GIS data compression. A newly developed data compression approach to approximate the land surface with a series of non-overlapping triangles has been presented. Over the years the TIN data representation has become a case in point for many researchers due its large data size. Compression of TIN is needed for efficient management of large data and good surface visualization. This approach covers following steps: First, by using a Delaunay triangulation, an efficient algorithm is developed to generate TIN, which forms the terrain from an arbitrary set of data. A new interpolation wavelet filter for TIN has been applied in two steps, namely splitting and elevation. In the splitting step, a triangle has been divided into several sub-triangles and the elevation step has been used to 'modify' the point values (point coordinates for geometry) after the splitting. Then, this data set is compressed at the desired locations by using second generation wavelets. The quality of geographical surface representation after using proposed technique is compared with the original terrain. The results show that this method can be used for significant reduction of data set.

Keywords. Delaunay triangulation, Triangulated irregular network (TIN), Geographical Information System, Lifting scheme, Second generation wavelet, Image compression, Multiresolution Analysis

1. Introduction

Recently, most of the methods for image compression are based on wavelets and related techniques. Wavelet approaches for image compression tend to outperform Fourier approaches because of its ability to represent both spatially localized features and smooth regions in an image. The superior compression capability of wavelets combined with their natural multiresolution structure makes them a good representation for storing images. While working with dyadic wavelet decomposition digital images are represented by wavelet coefficients. These types of representation in dyadic wavelet decomposition are known as linear decomposition over a fixed orthogonal basis. The non-linearity in the approximation of images by wavelets is introduced by the thresholding of the wavelet coefficients. This type of approximation can be viewed as *mildly* nonlinear. Recently, several *highly* nonlinear methods for capturing the geometry of images were developed, such as wedgelets [1] as well as edge-adapted nonlinear multiresolution and geometric spline approximation [2]. This paper presents a new approach for non-linear image compression method using second generation wavelets. A random set of points has been approximated to represent a surface by Delaunay triangulation. The theory, computations, and applications of Delaunay triangulations and Voronoi diagrams have been described in detail n the literature [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. The present work describes a fast algorithm based on Tsai's Convex Hull Insertion algorithm [10] and [11] for the construction of Delaunay triangulations and Voronoi diagrams of arbitrary collections of points on the Euclidean plane. The original algorithm has been improved further for a faster computation of geometric structures. The source code has been written in FORTRAN compiler. Once the triangulated irregular network has been created from the random set of points was further subjected to compression by using second generation wavelets. Results were shown in a comparative study basis for the TIN data compression at different level of resolution

2. Delaunay Triangulation

Many researchers [13] and [14] have suggested different ways to construct triangulations with the local equilateral property. A well known construction called the Delaunay Triangulation simultaneously optimizes several of the quality measures such as max-min angle, min-max circumcircle, and min-max min-containment circle. The Delaunay triangulation "DT" of a point set is the planar dual of the famous Voronoi diagram. The Voronoi diagram is a partition of the plane into polygonal cells one for each input point so that the cell for input point $'a'$ consists of the region of the plane closer to $'a'$ than to any other input point. So long as no four points lie on a common circle then each vertex of the Voronoi diagram has degree three and the DT which has a bounded face for each Voronoi vertex and vice versa will indeed is a triangulation. If four or more points do lie on a common circle then these points will be the vertices of a larger face that may then be triangulated to give a triangulation containing the DT Voronoi diagrams and Delaunay

triangulations have been generalized in numerous directions. For more information on Delaunay triangulations and Voronoi diagrams see the surveys by [15] and [16]. There is a nice relationship between Delaunay triangulation and three dimensional convex hulls. Lift each point of the input to a paraboloid in three-space by mapping the point with coordinates (x, y) to the point $(x, y, x^2 + y^2)$. The convex hull of the lifted points can be divided into lower and upper parts: a face belongs to the lower convex hull if it is supported by a plane that separates the point set from $(0, 0, -\infty)$. It can be shown that the DT of the input points is the projection of the lower convex hull onto the xy-plane as depicted in Figure 1. Finally a direct characterization: if a and b are input points the DT contains the edge $\{a, b\}$ if and only if there is a circle through a and b that intersects no other input points and contains no input points in its interior Moreover each circumscribing circle (circumcircle) of a DT triangle contains no input points in its interior.

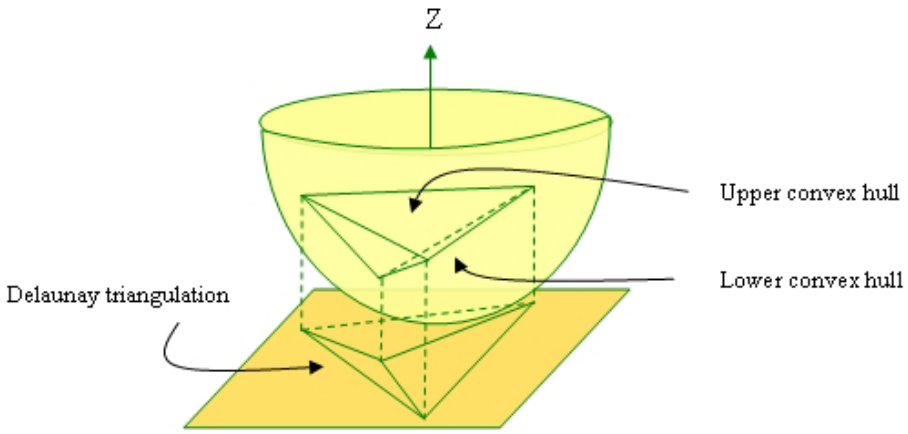


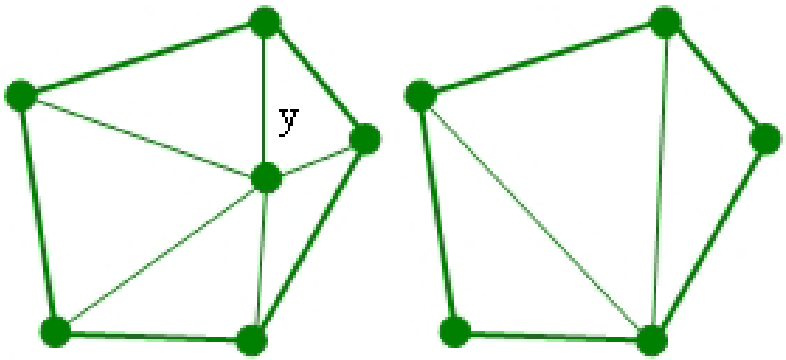
FIGURE 1. The lifting transformation maps the DT to the lower convex hull.

Some properties of Delaunay triangulations have been discussed as follows. Let Y denote a finite planar point set.

- A Delaunay triangulations $D(Y)$ of Y is one, such that for any triangle in $D(Y)$, the interior of its circumcircle does not contain any point from Y. This specific property is termed as Delaunay property.
- The Delaunay triangulation $D(Y)$ of Y is unique, provided that no four points in Y are co-circular. Since neither the set X of pixels nor its subsets satisfy this condition, we initially perturb the pixel positions in order to guarantee unicity of the Delaunay triangulations of X and of its subsets. Each perturbed pixel corresponds to one unique unperturbed pixel. From now on, we denote the set of perturbed pixels by X, and the set of unperturbed pixels by $\tilde{}X$.
- For any $y \epsilon Y$, $D(Y \setminus y)$ can be computed from $D(Y)$ by a *local* update. This follows from the Delaunay property, which implies that only the cell $C(y)$ of

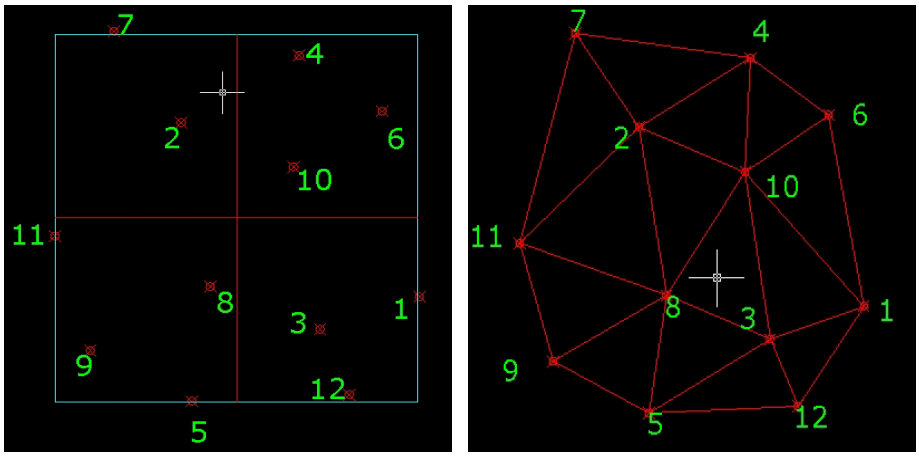
y in $D(Y)$ needs to be retriangulated. Recall that the cell $C(y)$ of y is the domain consisting of all triangles in $D(Y)$ which contain y as a vertex. Figure 1 shows a vertex $y \epsilon D(Y)$ and the Delaunay triangulation of its cell $C(y)$.

- $D(Y)$ provides a partitioning of the convex hull $[Y]$ of Y.



(a) (b)

FIGURE 2. Removal of the vertex $y \epsilon D(Y)$, and the Delaunay triangulation of its cell $C(y)$. The five triangles of the cell $C(y)$ in (a) are replaced by the three triangles in (b)



(a) (b)

FIGURE 3. Figure 3(a) shows irregular set of points and Figure 3(b) shows the TIN data structure using Delaunay triangulation

3. Interpolation Wavelet Filters for TIN

An interpolation wavelet filter for TIN lies in subdivision process which has two steps [17]. One is a splitting step; the other one is an elevation step. In the splitting step, a triangle is divided into several sub-triangles. The elevation step is to calculate the point values (point coordinates for geometry) after the splitting. Let us discuss this partition step mathematically. Consider a data set to be partitioned into two groups, called S and e. Using P to express the point coordinates, one can construct an estimation of P_s based on P_e.

$$\tilde{P}_s = E(P_e) \tag{1}$$

The estimation function (filter) (E) can be a local estimation or global estimation. A global estimation is generally computationally expensive; therefore, a local estimation using only neighboring points is preferred. After the estimation step, a wavelet term, W_s and an approximation term, A_s for the original data can be constructed as:

$$\begin{aligned} W_s &= P_s - \tilde{P}_s \\ A_e &= P_e + C(W_s) \end{aligned} \tag{2}$$

The correction function C is a customizable function based on different optimization requirements. An inverse transform can be constructed as:

$$\begin{aligned} P_e &= A_e - C(W_s) \\ P_s &= W_s + E(P_e) \end{aligned} \tag{3}$$

If the original point set can be partitioned into a nested group, then the above process can be iteratively applied to different sets in this group. A nested group has the flowing structure:

$$\begin{aligned} &e^0 \subset e^1 \subset e^2 \subset \cdots\cdots \subset e^N \\ &e^i \bigcup s^i = e^{i+1}, i = 0, \ldots, N-1 \end{aligned} \tag{4}$$

e^N denotes the finest representation of the geometry. e^N can be partitioned into e^{N-1} and s^{N-1}; then e^{N-1} can be partitioned into e^{N-1} and s^{N-1}, and so on, until e^1 is partitioned into e^0 and s^0. Note that the superscripts are used to represent different resolutions (larger numbers represent finer resolution). Based on this nested (or hierarchical) structure of the partition, one can construct wavelets and approximations of the data as:

$$\begin{gathered} A_{e^N} = P_{e^N} \\ e^i = e^{i-1} \bigcup s^{i-1}, B_{e^{i-1}} = A_{e^i}(e^{i-1}), B_{s^{i-1}} = A_{e^i}(s^{i-1}) \\ W_{e^{i-1}} = B_{e^{i-1}} - E(B_{e^{i-1}}), i = N, N-1, \ldots, 1 \\ A_{e^{i-1}} = B_{e^{i-1}} + C(W_{s^{i-1}}) \end{gathered} \tag{5}$$

Here, B is an intermediate symbol to represent the partitioning result. A_{e^i} is partitioned into two components: $A_{e^i}{}^{(e^{i-1})}$ and $A_{e^i}{}^{(s^{i-1})}$, which belong to e^{i-1} and s^{i-1} respectively. Based on equation (5), the original data P_{e^N} is decomposed into A_{e^0}, W_{s^0}, W_{s^1}, ..., $W_{s^{N-1}}$. Equation (5) is the analysis transform, which decomposes the finer representation into a coarser representation plus details. The synthesis transforms in the inverse transform and is shown in equation (6). The reconstructed A_{e^0}, A_{e^1}, ..., $A_{e^N}(= P_{e^N})$ yield a multiresolution representation of the original data.

$$\begin{gathered} B_{e^i} = A_{e^i} - C(W_{s^i}) \\ B_{s^i} = W_{s^i} + E(B_{e^i}), i = 0, 1, \ldots, N-1 \\ e^{i+1} = e^i \bigcup s^i, A_{e^{i+1}}(e^i) = B_{e^i}, A_{e^{i+1}}(s^i) = B_{s^i} \\ P_{e^N} = A_{e^N} \end{gathered} \tag{6}$$

In the above derivation of a wavelet representation [18, 19, 20, 21] the process does not depend on a regular setting for the data; therefore, it can be used in both the regular and irregular setting cases. This is an important advantage of the lifting scheme [22] and [23]. If the filters E and C and are the same for every point at a given level, the scheme is a uniform scheme. If they also do not change with the resolution, i, the scheme is a stationary scheme as well. However, equations (5) and (6) are general formulas. Non-stationary and non-uniform schemes can be written in this form with indices on E and C. Nevertheless, those schemes could cost more computing resources and may be less effective for data compression in GIS applications.

4. An Example

Figure 4 shows an example that illustrates the basic idea of the above construction process. Here E and C depend on e0 and s0; therefore, it is a non-stationary and non-uniform transform. Given irregular data on points $e^1 = \{t_0, t_1, t_2, t_3, t_4, \}$, the one step partition and wavelet transform are:

$$\begin{gathered} A_{e^1} = P_{e^1}, e^0 = \{t_0, t_2, t_4\}, s^0 = \{t_1, t_3\} \\ e^1 = e^0 \bigcup s^0, B_{e^0} = A_{e^1}(e^0), B_{s^0} = A_{e^1}(s^0), \lambda_1 = \tfrac{t_1 - t_0}{t_2 - t_0}, \lambda_1 = \tfrac{t_3 - t_2}{t_4 - t_2} \\ W_{s^0} = B_{s^0} - (B_{e^0}) = \begin{bmatrix} P_{t_1} \\ P_{t_3} \end{bmatrix} - \begin{bmatrix} 1-\lambda_1 & \lambda_1 & 0 \\ 0 & 1-\lambda_2 & \lambda_2 \end{bmatrix} \begin{bmatrix} P_{t_0} \\ P_{t_2} \\ P_{t_4} \end{bmatrix} \\ A_{e^0} = B_{e^0} + C(W_{s^0}) = \begin{bmatrix} P_{t_0} \\ P_{t_2} \\ P_{t_4} \end{bmatrix} + \begin{bmatrix} 1-\lambda_1 & 0 \\ \lambda_1 & 1-\lambda_2 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} W_{t_1} \\ W_{t_3} \end{bmatrix} \end{gathered} \tag{7}$$

In this example, a linear estimator and corrector are used for E and C. If λ_1 and λ_2 are 0.5, the correction step makes the norm defined in equation (8) take a similar value than the uncorrected value. (Note that in equation (8), A_{e^0} (t_j) for $t_j \notin e^0$ refers to the estimation value $E(A_{e^0})$ at t_j. This is an example of the

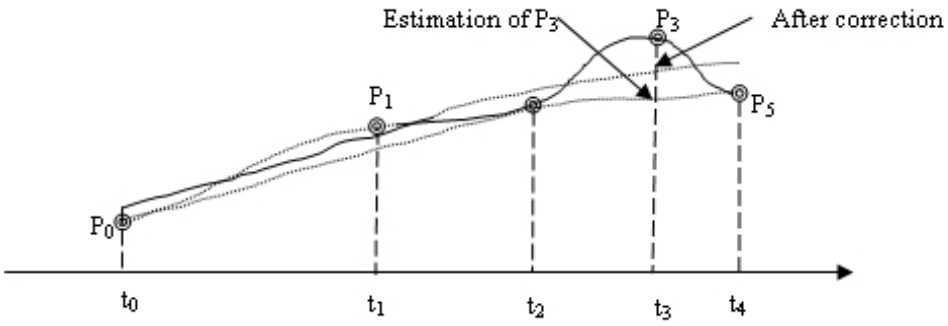


FIGURE 4. Example that illustrates the basic idea of construction of lifting scheme process.

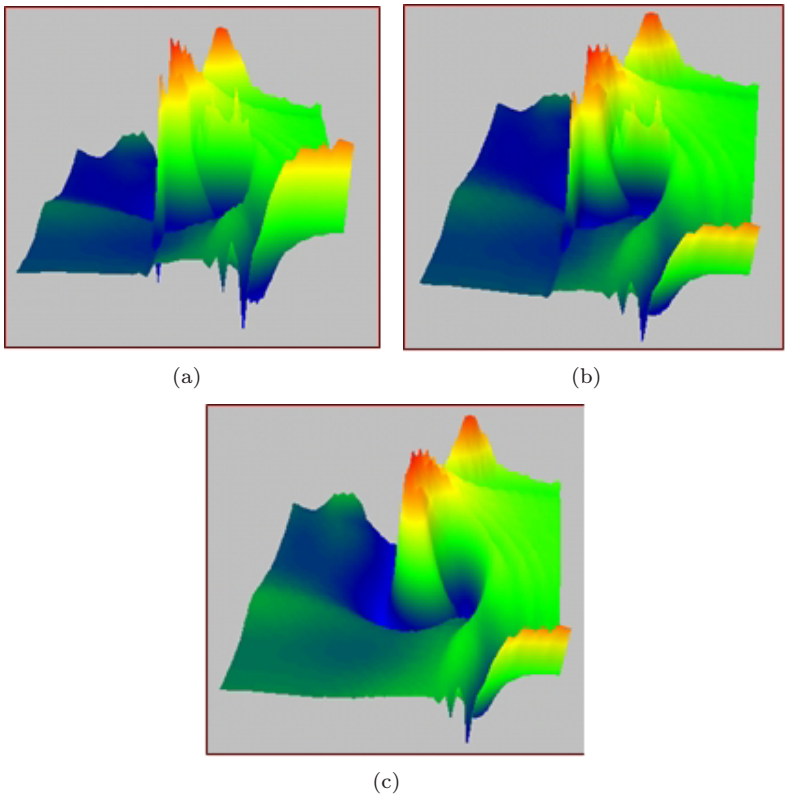
optimization in the correction step. Equation (7) is the analysis transform for the first step. The synthesis transform can be derived by replacing the estimator and the corrector in equation (6).

$$norm = \sum_{t_j \in e_1} (A_{e^1}(t_j) - A_{e^0}(t_j))^2 \tag{8}$$

The above scheme is one type of filter based on the lifting scheme. This is called an approximation filter, in which every point value will change after each iteration. The other type of filter based on the lifting scheme is an interpolation filter, in which a point value reaches its final position once it is calculated. If the correction term is omitted in equation (5), the filter becomes an interpolation filter. An approximation filter may be optimal in a defined global norm. However, interpolation filters have the advantage of interpolating point values, which may be useful in some GIS applications. Therefore, choosing the type of filter depends on the application. For example, approximation filters are generally used in processing image because the whole scene of an image is more important to the viewer than any individual pixel. For GIS terrain data, however, interpolation is generally preferred because point values are often more useful than a general shape. Therefore, interpolation wavelet filters will be used in this research for processing three-dimensional terrain data.

5. Results and Discussion

Results for irregular triangulated network data for three-dimensional GIS terrain have been presented. The original data was in an irregular fashion. Delaunay triangulation method for the creation of the TIN has been applied to the original data. A new algorithm has been developed for the creation of the TIN. Further, TIN



(a) (b)

(c)

FIGURE 5. Figure (a): shows the original TIN data; Figure (b): shows the wavelet compression at the level of 12% and Figure (c): shows the wavelet compression at level of 23%

was compressed using the second generation wavelets. A new algorithm for second generation wavelet compression is proposed. Based on the initial configuration of the original TIN, different resolutions are constructed during wavelet analysis. Figures 7 and 8 above show the results computed using the wavelet based lifting scheme algorithm. Figure 6 shows the original data. Figure 7 shows the results after 12% compression and figure 8 shows after 23% compression. Different compression schemes, such as Huffman coding can be applied to these wavelet coefficients to further reduce the storage size. This work also provides current implementation of wavelet coefficients during the compression operation. The proposed algorithm has the multiresolution capability and easy to compress due to large number of wavelet coefficients with small magnitudes which is suitable for distributed GIS applications such as web displaying. The research is ongoing. The future work

includes the wavelet compression analysis for different sets of terrain at different level of resolution to verify the results with existing compression techniques.

Conclusion

The construction of Triangulated Irregular Network using Delaunay triangulation has been shown. This approach uses fast and efficient second generation wavelets algorithm for multiresolution analysis of GIS data compression. This algorithm is easy to perform the mathematical and computational operation with minimal time, irrespective of the large data. Our algorithm scheme preserves high-gradient regions that might exist in a given data set. We have tested our method with various data sets. The computational cost of our algorithm depends on the different approaches used. The initial triangulation can be done in $O(n \log n)$, the gradient approximation can be done in $O(n \log n)$. The individual refinement step has to check all the original data points lying in the involved triangles, so the complexity of each step is $O(n)$. How often the iteration step is executed depends on the error value given in the input. As a general rule, the authors have assumed that no more iteration should be done than they are original data sites. So the overall complexity is $O(n^2)$. We are currently investigating the detailed error analysis for the different sets of data sets at different scales.

Acknowledgement

Authors would like to thank Addin Osman and Lawal Billa for their useful support concerning the manuscript putting in current Latex format.

References

[1] Donoho, D., 1999, "Wedgelets: nearly-minimax estimation of edges", *Annals of Stat.*, vol. 27, pp. 859-897

[2] Demaret, L., Dyn, N., Floater, M. S. and Iske, A., 2004, "Adaptive thinning for terrain modelling and image compression", in *Advances in Multiresolution for Geometric Modelling*, N. A. Dodgson, M. S. Floater, and M. A. Sabin (eds.), Springer-Verlag, Heidelberg, pp. 321-340

[3] Kao, T., Mount, D. M., and Saalfeld, A., 1991, "Dynamic maintenance of Delaunay triangulations": *Proc. Auto-Carto I0*, Baltimore, Maryland, p. 219-233

[4] Lawson, C. L., 1972, "Generation of a triangular grid with application to contour plotting", California Institute of Technology, Jet Pollution Laboratory, Technical Memorandum No. 299

[5] Lee, D. T., and Schachter, B. J., 1980, "Two algorithms for constructing a Delaunay triangulation", *Intern. Jour. Computer and Information Sciences*, v. 9, no. 3, p. 219-242

[6] Macedonio, G. and Pareschi, M. T., 1991, "An algorithm for the triangulation of arbitrary distributed points: applications to volume estimate and terrain fitting", *Computers & Geosciences*, v. 17, no. 7, p. 859-874

[7] Mirante, A. and Weingarten, N., 1982, "The radial sweep algorithm for constructing triangulated irregular networks", *IEEE Computer Graphics and Applications*, v. 2, no. 3, p. 11-21

[8] Puppo, E., Davis, L., DeMenthon, D. and Teng, Y. A., 1992, "Parallel terrain triangulation", *Proc. 5th Intern. Symposium on Spatial Data Handling*, v. 2, Charleston, South Carolina, p. 632–641

[9] Sibson, R., 1978, "Locally equiangular triangulations", *Computer Jour.*, v. 21, no. 3, p. 243-245

[10] Tsai, V. J. D., and Vonderohe, A. P., 1991, "A generalized algorithm for the construction of Delaunay triangulations in Euclidean n-space", *Proc. GIS/LIS '91*, v. 2, Atlanta, Georgia, p. 562-571

[11] Tsai,V. J. D., 1993, "Fast topological construction of Delaunay triangulations and Voronoi diagrams", *Computer & Geosciences*, Vol. 19, No. 10, pp. 1463- 1474

[12] Watson, D. F., 1981, "Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes", *Computer Jour.*, v. 24, no. 2, p. 167-172

[13] Evans, W., Kirkpatrick, D. and Townsend, G., 2001, "Right-Tringulated Irregular Networks. Algorithmica", *Special Issue on Algorithms for Geographical Information Systems*, 30, 2, pp. 264-286

[14] Absolo, M. J., Blat, J., and De Giusti, A., 2000, "A Hierarchial Tringulation for Multiresolution Terrain Models." *The Journal of Computer Science & Technology (JCS&T)*, 1, 3

[15] Fortune, S., 1991, "Voronoi diagrams and Delaunay triangulations", *In Computing in Euclidean Geometry*, ed. F. K. Hwang and D. Z. Du, World Scientific

[16] Aurenhammer, F., 1991, "Voronoi diagrams–a survey of a fundamental geometric data structure", *ACM Computing Surveys*, v. 23, no. 3, p. 345-405

[17] Kiema, J. B.K., and Bhr, H.-P., 2001, "Wavelet compression and the automatic classification of urban environments using high resolution multispectral imagery and laser scanning data", *Geoinformatics*, 5, 165-179

[18] Cohen, A., 2001, "Applied and computational aspects of nonlinear wavelet approximation", *Multivariate Approximation and Applications*, N. Dyn, D. Leviatan, D. Levin, and A. Pinkus (eds.), Cambridge University Press, Cambridge, pp. 188-212

[19] Dyn, N., Levin, D. and Gregory, J. A., 1990, "A butterfly subdivision scheme for surface interpolation with tension control", *ACM Transaction on Graphics*, 9, 160-169

[20] Mallat, S., 1989, A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 11, 674-693

[21] Wu, J. and Amaratunga, K., 2003, "Wavelet triangulated irregular networks", *International Journal of Geographical Science*, vol. 17, no. 3, pp. 273- 289

[22] Sweldens, W., 1994, "Construction and Applications of wavelets in Numerical Analysis", Unpublished PhD thesis, Dept. of Computer Science, Katholieke Universiteit Leuven, Belgium

[23] Sweldens, W., 1996, "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Appl. Comput. Harmon. Anal.*, 3(2):186- 200

B. Pradhan
Faculty of Engineering, University Putra Malaysia, 43400, UPM, Serdang
Selangor Darul Ehsan, Malaysia
Tel. +603-89467543, Fax. +603-86566061
Contact Author:
e-mail: biswajeet@mailcity.com

K. Sandeep
Department of Mechanical Engineering
Institute of Technology
Banaras Hindu University (BHU)
Tel. +91-542-2368427, Fax.+91-542-2368428

Shattri Mansor
Faculty of Engineering, University Putra Malaysia, 43400, UPM, Serdang
Selangor Darul Ehsan, Malaysia
Tel. +603-89467543, Fax. +603-86566061
Contact Author:
e-mail: biswajeet@mailcity.com

Abdul Rahman Ramli
Faculty of Engineering, University Putra Malaysia, 43400, UPM, Serdang
Selangor Darul Ehsan, Malaysia
Tel. +603-89467543, Fax. +603-86566061
Contact Author:
e-mail: biswajeet@mailcity.com

Abdul Rashid B. Mohamed Sharif
Faculty of Engineering, University Putra Malaysia, 43400, UPM, Serdang
Selangor Darul Ehsan, Malaysia
Tel. +603-89467543, Fax. +603-86566061
Contact Author
e-mail: biswajeet@mailcity.com

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 515–523
© 2006 Birkhäuser Verlag Basel/Switzerland

Ridgelet Transform as a Feature Extraction Method in Remote Sensing Image Recognition

Yuanyuan Ren, Shuang Wang, Shuyuan Yang and Licheng Jiao

Abstract. Using ridgelet transform to do the feature extraction, and RBFNN to do the recognition and classification, a remote sensing image recognition method is put forward in this paper. We do mathematical implementation and experimental investigation of ridgelet transform to analyze its characteristic and show its performance. Since ridgelet transform outperforms wavelet transform in extracting the linear features of objects, the proposed method has higher efficiency than that of wavelets. The simulation in remote sensing image shows its feasibility..

Keywords. Ridgelet transform,Radon transform,Dyadic wavelet,Image recognition and classification.

1. Introduction

Feature extraction is a necessary process in image classification and recognition. The aim of it is to reduce the complexity of the original problem by means of projecting the input patterns to a lower-dimensional space, i.e. to represent the target with fewer items without loss in another space. This calls for a sparser representation of the objective. Indeed, it is one of the main objectives of computational harmonic analysis (CHA) [1]. Wavelets, with its good time-frequency characteristic and its sparseness in representing the 0-dimensional (0-D) singularities of signals (i.e. the point-like targets), is an efficient tool for feature extraction, and has been widely used in the detection and recognition of various digital images. However, it is not the optimal basis to represent higher dimensional singularities such as linear singularities in 2-dimensional (2-D) space, and hyperplane singularities in higher dimensional spaces [2]. An obvious example of linear singularities in 2-D are edges, which usually contain abundant information, especially in textual

This work was supported by the National Science Foundation under grant no. 60133010.

images and SAR images. Ridgelets are superpositions of ridge functions or simple elements that are in some way related to ridge functions. With its orientation, ridgelets can represent objects with linear singularity in 2-D sparsely, and result in fewer features to represent more information and consequently improve accuracy of classification. Throughout the development of ridgelets, monoscale ridgelets [3], orthonormal ridgelets [4], finite ridgelets [5], and ridgelets frame [6] have been proposed sequentially. In remote sensing signal processing, recognition of 2-D radar targets is an important task. Radar images, such as planes and ships, often have stronger linear features. It can make contribution to the subsequent disposals of recognition and classification when a proper feature extraction tool is used. In this paper finite ridgelets are adopted for feature extraction. We firstly apply ridgelet transform to the pretreated images and then recognize and classify the targets with a simple RBFNN. Feature extractions based on the wavelet transform are also discussed for comparison. Experimental results illustrate the superiority of ridgelets to wavelets in extracting the linear features.

This paper is arranged as follows: the next section briefly explains the concept of the ridgelet transform and sparseness conception. In section 3 the detailed implementation of the referred algorithm is presented. Experimental results are shown in section 4 and the conclusion in section 5.

2. Implementation of Ridgelet Transform

The concept of ridgelets was firstly proposed by Cands [7] in 1998. Compared with wavelets, ridgelets have orientation and thus can sparsely represent linear singularities in 2-D, plane-like singularities in 3-D.

2.1. 2-D Ridgelet Transform in Image Processing

2-D ridgelet transform[2, 7] is defined as follows:

Let $\psi : \Re \to \Re$ satisfy $\int \psi(t)dt = 0$, for all $a > 0, b \in \Re, \theta \in [0, 2\pi)$, define bivariant function $\psi_{a,b,\theta} : \Re^2 \to \Re^2$:

$$\psi_{a,b,\theta}(x) = a^{-1/2} \cdot \psi((\cos(\theta)x_1 + \sin(\theta)x_2 - b)/a) \tag{1}$$

Where ψ is a 1-D wavelet function. Function $\psi_{a,b,\theta}$ is constant along "ridges", while a wavelet form transverse to these ridges, hence the name ridgelets.

For a bivariant function, its ridgelet coefficients are defined as:

$$R_f(a, b, \theta) = \int \bar{\psi}_{a,b,\theta}(x)f(x)dx =< \psi_{a,b,\theta}, f > \tag{2}$$

Suppose the Radon coefficients $Rf(u, t) = \int f(x)\delta(u'x - t)dx$ of $f(x)$ be the integral of f along $u'x = t$, and then we get:

$$R_f(a, b, u) =< \psi_{a,b}, Rf(u, \cdot) > \tag{3}$$

where $\psi_{a,b}(t) = \psi((t - b)/a)/\sqrt{a}$ is 1-D wavelets.

It can be inferred from above that the ridgelet transform is precisely the application of 1-D wavelet transform to the slices of the Radon transform where u is constant and is varying [2]. And its coefficients of a function are equal to the

orthogonal multiplication of its Radon transform with 1-D wavelets. Linear singularities in images are translated into point-like singularities by Radon transform, and these singularities can be best represented by wavelet transform. Through such transform, the ability of ridgelets to represent linear targets in higher dimensional space can be performed.

2.2. Dyadic Wavelet Transform

Dyadic wavelets [8, 9] lie between continuous wavelets and discrete wavelets. In dyadic wavelets, the scales are discretized while the shifts are continuous, which is different from discrete wavelets whose scales and shifts are both discretized, and the continuous wavelets whose scales and shifts are both continuous. Dyadic wavelets have the time translation invariance and are important in singularity detection. In this paper, 1-D dyadic wavelet transform is adopted. It is defined as follows:

$$\mathrm{W}_{2^j} f(t) = f * \psi_{2^j}(t) = \frac{1}{2^j} \int_R f(x) \psi(\frac{t-x}{2^j}) dx \tag{4}$$

Suppose $\psi(t)$ is a wavelet satisfying the permission condition. Let its scale varies along the dyadic sequence in time domain, hence the name dyadic wavelet. Since dyadic wavelet transform does not allow sub-sampling, it can better preserve the whole properties as well as its completeness, stability and redundancy. The singularity of a signal with its magnification can be confirmed by the variation of dyadic wavelet transform coefficients with scale parameter in one place. Since the local maxima are corresponding to the singularity points, the magnification and location of singularities can be detected by the dyadic wavelet transform [9].

2.3. The Analysis of Sparseness Conception

The conception about the "sparseness" can be traced back to a thesis written by Olshausen and Field [10]. They proposed that images can be presented by superstitions of base functions in different sampling blocks and searched for a sparse representation for the sampling block to reconstruct the original images. For a sampling block X^p, its sparse representation through transform is as below:

$$X^{\mathrm{p}} \approx \sum_{\mu} \theta_{\mu}^{p} \phi_{\mu} \tag{5}$$

where ϕ_μ are the base functions and θ_μ^p the corresponding coefficients. The higher the degrees of approximation of the coefficients are, the sparser the expression is. Ridgelet transform has the higher degrees of approximation than wavelet in 2-D, and thus can get sparser representation for linear singularities.

It can also be seen from the distribution graph below (Fig.1 (a) (b)) that after Radon transform, though there are many low values around the several distinct higher peeks, the number of peaks corresponding to the linear features in the original image which containing information are few. In contrast, though the distribution of wavelet transform coefficients (Fig.1 (c) (d)) is compact, the number of its peeks is more than the above graph. And in such a case, we propose the "coefficients convergence" method. Points around the peeks are converged.

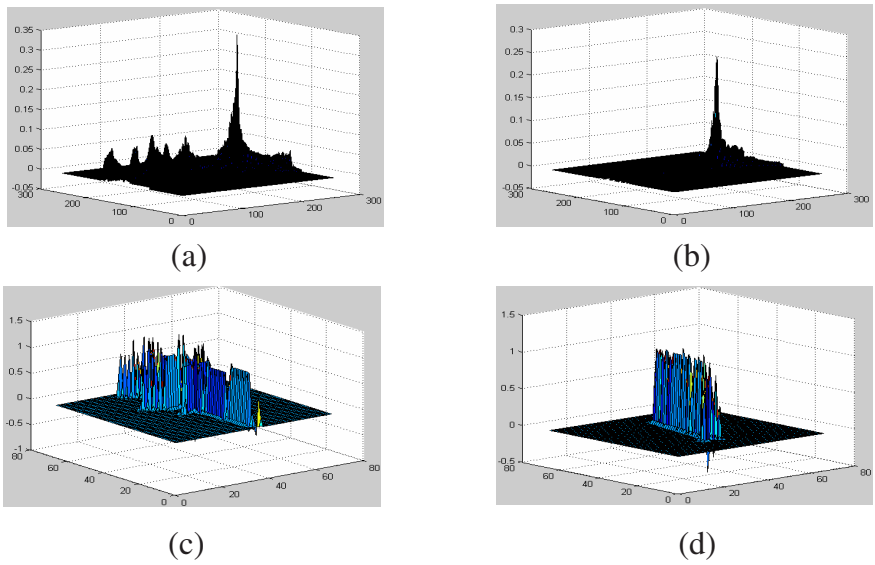


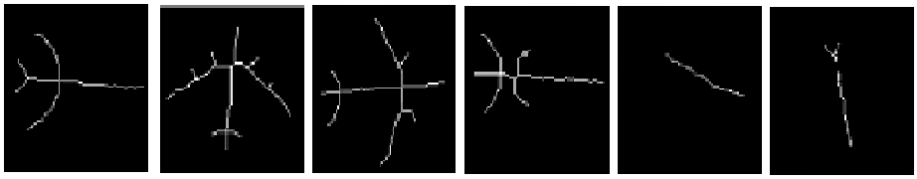
FIGURE 1. Distribution of coefficients (3-D) (a) Radon coefficients of planes, (b) Radon coefficients of ships; (c) wavelet coefficients of planes; (d) wavelet coefficients of ships

3. Recognition of Remote Sensing Image Based on Ridgelet Transform

For the remote sensing targets, the screening process is inevitably affected by the screening condition and weather, thus appear distortion, blurring, dithering. So these have to be pretreated first. After pretreatment, they have been changed into binary images.

3.1. Analysis of Algorithm

Before transform, samples are thinned to extrude their linear features. The skeletons of ships and planes are shown in Fig.2. After thinning, the redundancies of images are slashed while primary information is retained. We then transform images with ridgelet transform, by the way of dyadic wavelet transform implemented by *ÃTrous* algorithm after Radon transform implemented by FSS (Fast-Slant-Stark) [11]. It can be seen from the explanatory drawing that the skeletons of the same type of targets (Fig.2 (a) (b) (c) (d)) resemble each other, while skeletons of different objects differ. Since the number of basic skeletons after thinning is no more than 6, we converge the transformed coefficients to 6 points at most as features. Choose the approximation coefficients as wanted information and use these information to converge, thus produce the feature matrix of the images.



(a) (b) (c) (d) (e) (f)

FIGURE 2. (a)-(d): Skeletons of different type of planes (e) (f): skeletons of ships

3.2. Detailed Algorithm

Based on the analysis mentioned above, we adopt the following scheme to extract features with ridgelet transform:

1. Pretreatment of images with mathematical morphology: for the binary input image $f(m,n), 1 \leq m \leq 128, 1 \leq n \leq 128$ of 128×128, dilation is first used, and then thinning, thus get the treated image f_1;
2. Radon transform: Radon transform is taken for with the Fast-Slant-Stark method, and the transformed image is denoted by f_2 is 256×256;
3. Dyadic wavelet transform: dyadic wavelet transform is done to each slice of f_2, with $\tilde{A}Trous$ algorithm, thus each column in the radon transformed matrix is changed into $2^i - 1$ columns according to the preset scale i, each of which stores the corresponding frequency component coefficients at corresponded scale and the first stores the approximating component, select the low frequency coefficients as features. The size of the transformed matrix is 256×256, denoted by R_i;
4. Feature extraction and selection:
   (a). For the discrimination: converge the ridgelet transform coefficients R_i of each image into several representing points, pick-up the first 6 largest values as features and form the feature matrix;
   (b). For the classification: singular value decomposition is applied on the ridgelet transform coefficients matrix R_i of each image, Select the first 25 values to compose the feature vector;
5. An RBFNN is adopted as a classifier to implement the classification. Samples obtained through the above steps are sent into the network for training and testing.

4. Experimental Results

Experiments of target recognition and classification based on the ridgelet transform for extracting the features combined with RBFNN are put into practice: First are the experiments on recognition. The samples are shown in Fig. 3.

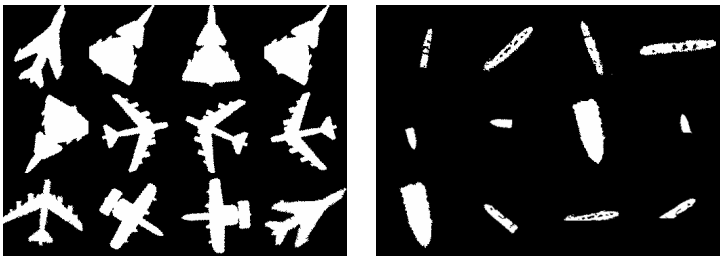


FIGURE 3. Samples 1:Planes and ships

TABLE 1. Comparison of wavelet and the referred method for recognition

	Level[1]	Wavelet transform		The referred method	
		SMR[2]	RPR[3]	SMR	RPR
Experiment 1:	L=1	86	90.49 percent	23	97.46 percent
Training samples: 160	L=2	100	88.94 percent	27	97.01 percent
Testing samples: 904	L=3	73	91.92 percent	43	95.24 percent
Experiment 2:	L=1	36	91.82 percent	13	97.05 percent
Training samples: 160	L=2	49	88.86 percent	14	96.82 percent
Testing samples: 440	L=3	40	90.91 percent	30	93.18 percent
Experiment 3:	L=1	35	90.28 percent	11	96.94 percent
Training samples: 240	L=2	37	89.72 percent	14	96.11 percent
Testing samples: 360	L=3	35	90.28 percent	28	92.22 percent

Experiment 1 is carried out under the following condition: there are totally 1064 2-D pictures, 80 from which about planes and ships respectively are randomly selected for training, while 904 are for detection. Experiment 2 is under the same condition with reference [12]: 280 ships and 320 planes composing totally 600 samples are included. 80 respectively are randomly selected for training, while 440 for detection. Experiment 3 is under the same condition with the reference [13]: 280 ships and 320 planes composing totally 600 samples are included, 40 percent of which are for training, while 60 percent for detection([1] Level: decomposition level;[2] SMR: Samples of Miss Recognized; [3] RPR: Rate of Proper Recognized).

The experimental results show the higher rate of ridgelet transform than wavelet through the identical procedure in recognition. And the results in experiment 2 are better than the rate of 94.5 percent listed in reference [12]; the results in experiment 3 are better than that listed in reference [13].

The RPR of the referred method decreases with the increasing decomposition level. The reason for this phenomenon is that, as the decomposition level increasing, scale becomes coarser, and information contained in the obtained image decreased, lead to the lower recognition rate. Therefore the following experiments will carried out only in the level 1. Then are the experiments on classification of planes, samples are shown in Fig. 4.

TABLE 2. Results of classification of 4 different kinds of planes

Experiment 4:		SMR	RPR
Training samples: 81	Wavelets	24	79.13 percent
Testing samples: 115	The referred method	3	97.39 percent

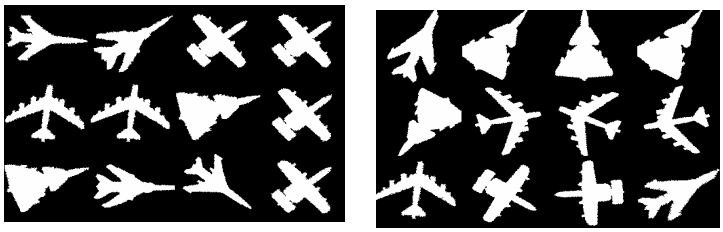


FIGURE 4. Samples 2:4 different kinds of planes

Experiment 4 is done under the following condition: there are totally 196 2-D pictures about 4 kinds of planes, in which 81 are randomly selected for training and 115 for detection. Adjust the parameters and get the experimental results in Table 2:

Finally are the experiments on the bridges in SAR images of Washington.D.C from Sandia National Laboratory, which includes 12 bridges.SAR images, for their complicated background and speckle noise, are difficult to distinguish between different bridges. So the standard deviations of each frequency component as well as mean values are extracted in the features selection process. Fig.5 shows the bridges used in the experiments. To make up the deficiency of the samples, images are rotated at intervals of 6 degrees with the bicubic interpolation method, and this produces 780 images in all, 260 of which are taken for training and 520 for testing. The experimental results are shown in table 3.

TABLE 3. Results of classification of 12 bridges

Experiment 5:		SMR	RPR
Training samples: 260	Reference [14]	-	85.76 and 93.65 percent
Testing samples: 520	The referred method	32	93.85 percent

The results in reference [14] are listed as comparison. In the above chart, the data 85.76 percent in reference [14] is got by the Radon transform with RBFNN as classifier while 93.65 percent is got by the Radon transform with SVM as classifier.

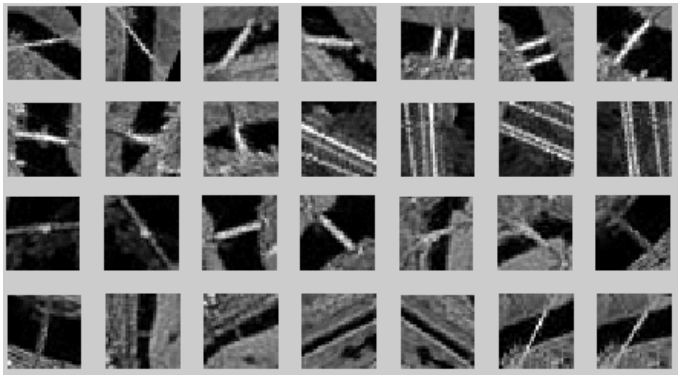


FIGURE 5. Samples 3:12 bridges in SAR image (where the two adjacent images are of the same kind. And the 25th and 26th bridges are fake targets seen from top to down and from left to right).

5. Conclusion and Discussions

Ridgelet transform used in feature extraction is introduced in this paper. Feature extractions are implemented based on the ridgelet transform, and then neural network is used for the subsequent recognition and classification. The experimental results show the ridgelet transform is more efficient than wavelet transform in recognizing and classifying the remote sensing images for its fewer decomposition coefficients, as well as its better application prospect. The drawback of it is that it requires longer calculation time in feature extraction than wavelet transform. This is caused by the concrete way of implementation of ridgelet transform, i.e. the radon transform combined with dyadic wavelet transform. The further research is to increase the implementation efficiency of the ridgelet transform. When a fast and efficient method of ridgelet transform is obtained, this problem can be solved.

References

[1] E.J. Cands, *Ridgelets and Their Derivatives: Representation of Image with Edges,* Department of Statistics, Stanford University 1999.

[2] E.J. Cands, and D.L. Donoho, Ridgelets: the Key to Higher-dimensional Intermittency?, Phil.Trans. R.Soc. Lond. A. (1999). 357:2495.

[3] E.J. Cands, *Monoscale Ridgelets for the Representation of Images with Edges*, Department of Statistics, Stanford University 1999.

[4] D.L Donoho, *Orthonormal Ridgelets and Linear Singularities,* SIAMJ. Math. Anal, 2000, 31 (5): 1062.

[5] M.N.Do and M.Vetterli, *The Finite Ridgelet Transform for Image Representation*, Image Processing. IEEE Transactions on 12.(2003) :16-28.

[6] T.Shan,L.C. Jiao and X.C.Feng,*Ridgelets Frame,* In: Aurlio, C.C, Mohamed, S.K. (Eds.): Proceedings of Image Analysis and Recognition. Lecture Notes in Computer Science, Vol. 3211. Springer-Verlag, Berlin Heidelberg New York (2004) 479-486.

[7] E.Cands, *Ridgelets: theory and applications,* Ph.D. thesis, Department of Statistics, Stanford University, 1998.

[8] S.Mallat, *A Wavelet Tour of Signal Processing*, Academic Press.1999.

[9] B.Hou, F.Liu and L.C. Jiao, *Linear detection based on ridgelet transform*, China science, Vol.46 No.2, April. 2003: 73 105.

[10] B.A.Olshausen and D.J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature, 381 607-609, 1996.

[11] A.Averbuch, R.R.Coif man,D.L. Donoho, M.Israeli and J.Walden, *Fast Slant Stack: A Notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computable, Algebraically Exact, Geometrically Faithful an Invertible*, Department of Statistics, Stanford University (2001).

[12] Y.N.Zhang, J.b. Hong, X.h.Wang and R.c Zhao, *An efficient image target recognition method for remote sensing,* Signal Processing,Vol.18. No.1 Feb.2002

[13] Y.N.Zhang, J.b. Hong,Y. Liao and R.c Zhao,, *Remote sensing image recognition based on SVM*, Journal of Northwestern Polytechnical University, Nov.2002. Vol.20 No.4.

[14] L.C.Jiao, *AIntelligence signal and image processing*, Journal of xidian university, 2003.9,144-150.

Yuanyuan Ren
National Key Lab for Radar Signal Processing and Institute of Intelligent Information Processing
Xidian University, 710071, Xi'an, China
e-mail: snowyraining@sohu.com

Shuang Wang
National Key Lab for Radar Signal Processing and Institute of Intelligent Information Processing
Xidian University, 710071, Xi'an, China

Shuyuan Yang
National Key Lab for Radar Signal Processing and Institute of Intelligent Information Processing
Xidian University, 710071, Xi'an, China

Licheng Jiao
National Key Lab for Radar Signal Processing and Institute of Intelligent Information Processing
Xidian University, 710071, Xi'an, China

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 525–542
© 2006 Birkhäuser Verlag Basel/Switzerland

Analysis of Frequency Spectrum for Geometric Modeling in Digital Geometry

Zhanchuan Cai, Hui Ma, Wei Sun and Dongxu Qi

Abstract. In order to explore the properties of frequency spectrum for geometric modeling, a complete orthogonal piecewise k-degree polynomials in $L^2[0,1]$, so-called U-system, is introduced. The expansion in U-series has advantageous properties for approximations in both quadratic norm and uniform. Using U-system with finite items, it can be realized to exactly represent geometric modeling. This paper analyzes the properties of frequency spectrum for geometric modeling in theory and gives some interesting results in geometric transform.By comparing U-system with Fourier system,experiments indicate that U-system is more suitable for analyzing frequency spectrum for geometric modeling than Fourier system is.

Mathematics Subject Classification (2000). Primary 99Z99; Secondary 00A00.

Keywords. U-system, geometric modeling, energy, frequency spectrum, normalization.

1. Introduction

With the increasing use of geometric modeling in industry, algorithms processing geometric signals become more and more important[1] . In the field of digital geometric processing, it is vital that the different geometric graphs can be represented by some basic graphical elements. In many cases of CAGD, objects are always described as one/ several group of graphs, such as machine parts and components, cartoons, surfaces of buildings and so on [2][3]. It is remarkable that there is close relation between geometric problems and image processing. For example, in pattern recognition, people need to match the graph based on its characters after

This project is supported by "Mathematics mechanization method and its application on Information Technology" (the National Grand Fundamental Research 973 Program of China under Grant No.2004CB3180000), the National Natural Science Foundation of China (No. 60133020), Guangdong Provincial Science and Technology Foundation (No. 2004A10302005, No. 051Z205005) and The China Scholarship Council. Corresponding author:Wei Sun(Email:lwhite@263.com).

getting the geometric graph of a certain object by edge detection. So, it is necessary to discuss the effective methods on how to make frequency spectrum analysis with the data coming from these geometric graphs. In the domain of CAGD [4] [5], discussion of frequency spectrum will help to classify graphs [15] [16].

Firstly, in the theories and applications of communication and signal processing, the class of complete orthogonal functions in $L^2[0,1]$ is the basic and important mathematical tool. Secondly, if there is a felicitous class of complete orthogonal functions system, which can exactly express a widely kind of groups of parametric geometric modeling by finite functions, we can take them as the most basic graphical elements. Thus we can get a new expression of the group of geometric graphs based on those functions. Moreover, we may use some theories on the class of orthogonal complete functions to do further analysis and synthesis of geometric information and other relative important research.

Currently, despite the recent promising progress of digital geometric processing, it still lacks analysis of frequency spectrum like digital image. The research of digital geometry has been rarely conjoint with the analysis of frequency spectrum, mainly because of the choice of orthogonal complete functions. In fact, not all class of orthogonal complete functions is suitable for the analysis and synthesis of geometric information. It is well known that Fourier system, Legendre system and Chebyshev system are the class of orthogonal functions system [6][7]. However, they cannot exactly express commonly geometric information (curves, surfaces) by using finite items, such as the phenomena of Gibbs (See FIGURE 9). So they are not suitable for analysis of frequency spectrum for geometric modeling of digital Geometry [15][17]. Moreover, because of the strong discontinuity of Walsh and Harr systems, they cannot be used as a tool for those applications, which need the finite items to approximate continuous function [6].

Twenty years ago, Qi and Feng proposed the U-system and gave the relatively completed theories [8]. While in [19], they study a similar idea on the triangle. Using these theories, in [9], motivated by in the work in [8], Micchelli and Xu constructed orthogonal multiwavelets for any invariant sets in R^n and later they used the wavelet to develop Galerkin methods and collocation methods to solve integral equations with weakly singular kernels. This indicates the application foreground of U-system [9-13].

The $k-$degree U-system consists of a series of piecewise $k-$degree polynomials. It includes not only differentiable functions but also piecewise ones, which include all kinds of discontinuous points on [0, 1]. It means that discontinuous points in different levels are conducive to express geometric information.

In this paper we investigate a special class of complete orthogonal functions in $L^2[0,1]$, so-called U-system, U-system can exactly express common geometric information (curves, surfaces) by using finite items .So U-system are suitable for analysis of frequency spectrum for geometric modeling in digital geometry [15][16][17]. Moreover, we analyze frequency and normalized frequency properties of geometric modeling in theory and experiment. We find that normalized frequency and normalized energy for the same geometric modeling are invariant in

geometric transform, such as translation, scale and rotation. In the end, some experimental results are given.

2. Piecewise Linear Complete Orthogonal U-system

As preparing, this part firstly introduce the definition of piecewise linear U-system.

The piecewise linear complete orthogonal function system on the interval [0, 1], so called U-system,is constructed as following:

Consider the linear function space on the interval [0, 1], denoted by S10, and then the following two function:

$$U_0(x) = 1, \quad 0 \le x < \frac{1}{2} \qquad (1)$$

$$U_1(x) = \sqrt{3}(1-2x), \quad \frac{1}{2} < x \le 1 \quad (2)$$

Structure the standard orthogonal base of S10. Further consider the piecewise linear function space divided by the point $x = \frac{1}{2}$ on the interval [0,1], denoted by S11. Obviously, the dimension of S11 is 4. Besides the two functions showed in (1) $\sim$ (2), we need two piecewise linear functions:

$$U_2(x) = \begin{cases} \sqrt{3}(1-4x), & 0 \le x < \frac{1}{2} \\ \sqrt{3}(4x-3), & \frac{1}{2} < x \le 1 \end{cases} \qquad (3)$$

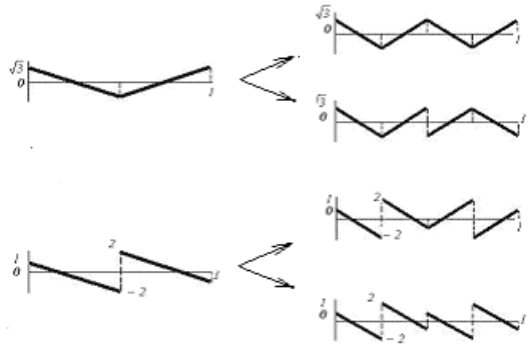
$$U_3(x) = \begin{cases} 1-6x, & 0 \le x < \frac{1}{2} \\ 5-6x, & \frac{1}{2} < x \le 1 \end{cases} \qquad (4)$$

Where the first three function in (1) $\sim$ (3) is orthogonal each other. The fourth function in (4) can be easily obtained based on the piecewise linear condition and the orthogonality with the second function. Note that the fourth function is odd symmetric on the point $x = \frac{1}{2}$, but it is not symmetric on the interval [0, $\frac{1}{2}$] and [$\frac{1}{2}$, 1]. Summarily, the four functions in (1) $\sim$ (4) constitute a group of standard orthogonal base.

Next, consider the piecewise liner function space S12 on the interval [0,1], divided by the point $x = \frac{1}{4}$,$x = \frac{1}{2}$,$x = \frac{3}{4}$. Because the dimension of S12 is 8, there must be 8 orthogonal piecewise linear functions. That is to say that, four piecewise linear function divided by the point $x = \frac{1}{4}$,$x = \frac{1}{2}$,$x = \frac{3}{4}$ must be constructed besides the four function above. Therefore, using "compressing and copying" generates the other four functions.

So-called "compressing and copying" can be divided into two parts. First, compress the third and the fourth function on the half of the interval, [0, $\frac{1}{2}$]; Next, direct-copy and opposite-copy it on [$\frac{1}{2}$, 1]. Then, every function generates two new function on S12, which are contrary parity on the point $x = \frac{1}{2}$. "Compressing and copying" is showed in the following figure:

"Compress and copy" the third and the fourth function, and then the four new functions can be expressed as following:



$$U_4(x) = \begin{cases} U_2(2x), & 0 \le x < \frac{1}{2} \\ U_2(2-2x), & \frac{1}{2} < x \le 1 \end{cases} \quad (5)$$

$$U_5(x) = \begin{cases} U_2(2x), & 0 \le x < \frac{1}{2} \\ -U_2(2-2x), & \frac{1}{2} < x \le 1 \end{cases} \quad (6)$$

$$U_6(x) = \begin{cases} U_3(2x), & 0 \le x < \frac{1}{2} \\ U_3(2-2x), & \frac{1}{2} < x \le 1 \end{cases} \quad (7)$$

$$U_7(x) = \begin{cases} U_3(2x), & 0 \le x < \frac{1}{2} \\ -U_3(2-2x), & \frac{1}{2} < x \le 1 \end{cases} \quad (8)$$

It is easy to be verify that the eight piecewise linear functions in (1)∼(8) constitute a group of standard orthogonal base on S12. It is stressed that, the "compressing and copying" mentioned above which generates the new function have the generality. That is to say that, when consider eight equal sub-interval on [0, 1] and the dimension of piecewise linear function space S13, 16, more 8 new function need to be added. These functions can be generated by "compressing and copying" mentioned above.

Generally, construct the $2n+1$ orthogonal base function on S1n, "compress and copy" the latter 2n function of these functions to obtain 2n+1 functions, and add them to constitute the standard orthogonal base on space S1 (n+1). $n = 0, 1, 2, 3, ...$ It can be proved that, when $n \to \infty$, the obtained infinite piecewise linear function group have the standard orthogonality and completeness.

Piecewise linear standard orthogonal complete system in $L^2[0,1]$ can be expressed by:

$$U_0(x) = 1, U_1(x) = \sqrt{3}(1-2x), 0 \le x \le 1$$

$$U_2^{(1)}(x) = \begin{cases} \sqrt{3}(1-4x), & 0 \le x < \frac{1}{2} \\ \sqrt{3}(4x-3), & \frac{1}{2} < x \le 1 \end{cases}$$

$$U_2^{(1)}(x) = \begin{cases} 1-6x, & 0 \le x < \frac{1}{2} \\ 5-6x, & \frac{1}{2} < x \le 1 \end{cases}$$

.....

$$U_{n+1}^{2k-1}(x) = \begin{cases} U_n^k(2x), & 0 \le x < \frac{1}{2} \\ U_n^k(2-2x), & \frac{1}{2} < x \le 1 \end{cases}$$

$$U_{n+1}^{2k}(x) = \begin{cases} U_3(2x), & 0 \le x < \frac{1}{2} \\ -U_3(2-2x), & \frac{1}{2} < x \le 1 \end{cases}$$

$$k = 1, 2, 3, ..., 2^{n-1}, n = 2, 3, ...$$

The value of the function takes average limited value on two sides on the divided point. Such 16 defined function on piecewise linear U system is showed in FIGURE 1. (b)

3. The k-degree U-system

In this section, the k-degree U-system is introduced [8].

3.1. Construction of k-degree U-system

Step 1: Taking first k+1 polynomials in Legendre orthogonal system, and denote them as,$U_0(x), U_1(x), ..., U_i(x), x \in [0, 1]$

Step 2: Construct k+1 new functions $f_i(x), i = 1, 2, ..., k+1, x \in [0, 1]$,and let them satisfy conditions:

(i) $f_i(x)$ is a k-degree piecewise polynomial divided by the point $x = \frac{1}{2}$;
(ii) $< f_i(x), f_j(x) >= \delta_{ij}, i, j \in \{1, 2, ..., k+1\}$;
(iii) $< f_i(x), x^j >= 0, i \in \{1, 2, ..., k+1\}, i \in \{0, 1, 2, ..., k+1\}$.

Where$< \bullet, \bullet >$ denotes the inner product in L^2[0,1]. Thus, we get the series of functions: $U_0(x), U_1(x), ..., U_k(x), f_1(x), f_2(x), ..., f_{k+1}(x)$;

Step 3: Using "squish-repeat", or "direct-copy and opposite-copy" [14] constructs the other $2 \cdot (k+1)$ functions.

Beginning from $f_1(x)$, each function generates two new ones as follows,

$$f_{i,1}(x) = \begin{cases} f_i(2x), & 0 \le x < \frac{1}{2} \\ f_i(2-2x), & \frac{1}{2} < x \le 1 \end{cases}$$

$$f_{i,2}(x) = \begin{cases} f_i(2x), & 0 \le x < \frac{1}{2} \\ -f_i(2-2x), & \frac{1}{2} < x \le 1 \end{cases}$$

The rest may be deduced by analogy, and we can obtain the class of k-degree U-system.

Obviously, when $k = 0$, the 0-degree U-system is just the series of Walsh functions. That means Walsh functions are the special case of U-system.

When $k = 0, 1, 2, 3$, the relative functions are shown in FIGURE 1 (a), (b), (c), (d). It is noticeable that, in FIGURE 1, functions of U-system are denoted as $U_l^{(j)}$, where l denotes the number of fragments and j does the j^{th} function[8].

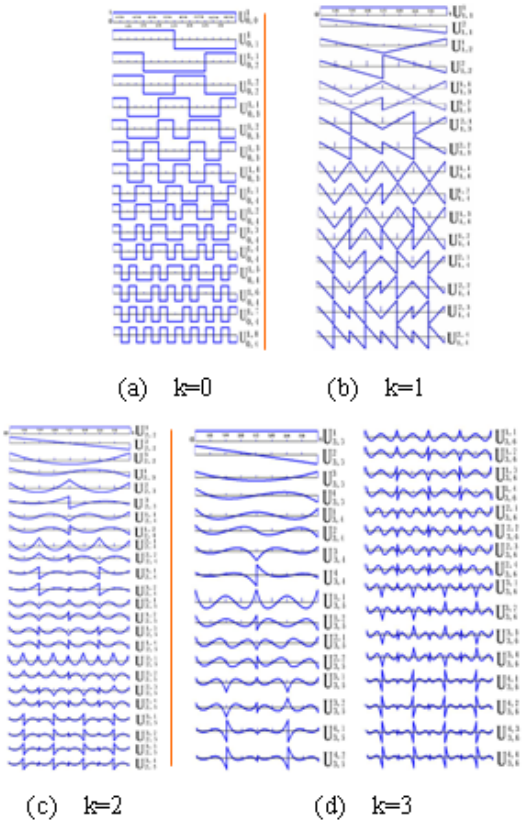


FIGURE 1. $k-$degree U-system (when k=0,1,2,3)

3.2. Properties of k-degree U-system

K-degree U-System has some properties [8]:

(i) Orthonormality $< f_i(x), f_j(x) >= \delta_{ij}, i, j = 0, 1, 2, ...$

(ii) Convergence of Fourier-U Series by Group
For a given function F, let

$$P_{n+1}F = \sum_{i=0}^{n} a_i U_{k,i}$$

Then let $P_{n+1}F$ is the best L_2-approximation to F from the space $span(U_{k,j})_0^n$..
Thus we have

$$\lim_{n \longrightarrow \infty} ||F - P_n F||_2 = 0, F \in L_2[0, 1]$$

$$\lim_{n\longrightarrow\infty} ||F - P_n F||_\infty = 0, F \in C[0,1]$$

These denote that Fourier-U series have the properties of L_2-convergence, completeness and convergence uniform by group.

(*iii*) Fourier-U Series Reproduction

If F is a piecewise k-degree polynomial, which has some discontinuous points on $x = \frac{q}{2^r}$ (q and r are integers), it can be exactly express with finite items of Fourier-U series. The k-degree U-system has much abundant discontinuity information. Especially using convergence uniform by group and Fourier-U series reproduction, we can analyse frequency spectrum of geometric modeling.

4. Orthogonal Representation of Parametric Curves and Surfaces [21] [22] of Geometric Modeling

Let $s, t \in [0,1]$, for simplicity, we take one degree U-system as an example. Assuming that the interval $[0,1]$ is divided into eight sub-intervals. Then, approximation function is expressed:

$$F_f(s,t) = f_i(s,t)$$

where $s \in [\frac{m}{8}, \frac{m+1}{8}), t \in [\frac{n}{8}, \frac{n+1}{8}), m, n = 0, 1, ..., 7$

Let

$$P(s,t) = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix} = \sum_{i=0}^{15}\sum_{j=0}^{15} \lambda_{i,j} U_{1,i} U_{1,j}(t)$$

The curves or surfaces in the given geometric modeling are expressed parametric polynomials:

$$\begin{cases} x(s,t) = F_x(s,t) \\ y(s,t) = F_y(s,t) \\ z(s,t) = F_z(s,t) \end{cases}$$

Where $\lambda_{i,j} = \int_0^1 P(s,t) U_{1,i} U_{1,j} dsdt, i, j = 0, 1, 2, ..., 15. P(s,t)$ can be exactly expressed by the given $F_{i,j}(s,t)$ with finite items. So we call $\lambda_{i,j}$ as frequency spectrum of $P(s,t)$.

The "energy" of $P(s,t)$ is defined as following:

$$E = (\sum_{i=0}^{15}\sum_{j=0}^{15} \lambda_{i,j}^2)^{\frac{1}{2}}$$

5. Frequency Spectrum Theorem

According to the above defined $\lambda_{i,j}$, we can obtain these frequency spectrum theorems in below. These theorems are very helpful and strongly mathematic tools for analyzing frequency spectrum for geometric modeling.

Theorem 5.1 (Superposed Theorem). *If frequency spectrum of function $F_1(s,t)$ is $\lambda_{i,j}^{(1)}, i,j = 0,1,2,...,n$, and frequency spectrum of function $F_2(s,t)$ is $\lambda_{i,j}^{(2)}, i,j = 0,1,2,...,n$, then frequency spectrum of function $F(s,t) = F_1(s,t) + F_2(s,t)$ should be $\lambda_{i,j}^{(1)} + \lambda_{i,j}^{(2)}$.*

Proof.

Frequency spectrum of function $F(s,t)$ is

$$\begin{aligned}\lambda_{i,j} &= \int_0^1 \int_0^1 P(s,t)U_{1,i}U_{1,j}dsdt \\ &= \int_0^1 \int_0^1 (P_1(s,t)U_{1,i}U_{1,j}dsdt + \int_0^1 \int_0^1 P_2(s,t)U_{1,i}U_{1,j}dsdt \\ &= \lambda_{i,j}^{(1)} + \lambda_{i,j}^{(2)}\end{aligned}$$

□

Corollary 5.2. *If frequency spectrum of function $F_m(s,t)$ is $\lambda_{i,j}^{(m)}, i,j=0,1,2,...,n$; $m = 0,1,2,...$; where $\lambda_{i,j}^{(m)} = ((\lambda_{i,j}^{(m)})_x, (\lambda_{i,j}^{(m)})_y, (\lambda_{i,j}^{(m)})_z)$,then frequency spectrum of function*

$$F(s,t) = \sum_{j=1}^{n} F_j(x,y)$$

should be

$$\sum_{m=1}^{n} \lambda_{i,j}^{(m)}$$

Theorem 5.3 (Retarded Theorem). *If frequency spectrum of function $F(x,y)$ is $\lambda_{i,j}, i,j = 0,1,2,...,n$, then frequency spectrum of function $F(x \pm a, y \pm b) \pm c$ should be $\lambda_{i,j} + (\mp a\delta(k), \mp b\delta(k), \pm c\delta(k))$,where $\delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad or \quad j \neq 0 \\ 1, & i = 0 \quad and \quad j = 0 \end{cases}$ a,b,c is random constant.*

Proof.

Supposed that $F(x,y)$are expressed parametric polynomials:

$$P(s,t) = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix}$$

then frequency spectrum of function $F(x,y)$ is $\lambda_{i,j} = \int_0^1 \int_0^1 P(s,t)U_{1,i}U_{1,j}dsdt$

$F(x \pm a, y \pm b) \pm c$ are expressed parametric polynomials:

$$P^{'}(s,t) = \begin{pmatrix} x^{'}(s,t) \\ y^{'}(s,t) \\ z^{'}(s,t) \end{pmatrix} = \begin{pmatrix} x(s,t) \pm a \\ y(s,t) \pm b \\ z(s,t) \mp c) \end{pmatrix}$$

So $\lambda^{'}_{i,j} = \int_0^1 \int_0^1 P^{'}(s,t) U_{1,i} U_{1,j} dsdt = \lambda_{i,j} + (\mp a\delta_{i,j}, \mp b\delta_{i,j}, \mp c\delta_{i,j})$,

where $\delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad \text{or} \quad j \neq 0 \\ 1, & i = 0 \quad \text{and} \quad j = 0 \end{cases}$ a,b,c is random constant.

□

Corollary 5.4. *If frequency spectrum of function $F(x,y)$ is $\lambda_{i,j}, i, j = 0, 1, 2, ..., n$, then frequency spectrum of function $F(x-a, y-a)+a$ should be $\lambda_{i,j} + (a\delta(k), b\delta(k), c\delta(k))$, where $\delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad or \quad j \neq 0 \\ 1, & i = 0 \quad and \quad j = 0 \end{cases}$ a is random constant. That is to say, frequency spectrum of function $F(x-a, y-a)+a$ is $\lambda_{i,j} + a (i = 0 \quad and \quad j = 0$ or $i \neq 0 \quad or \quad j \neq 0)$.*

Theorem 5.5 (Inverse Ratio Theorem). *If frequency spectrum of function $F(x,y)$ is $\lambda_{i,j}, i, j = 0, 1, 2, ..., n$, then frequency spectrum of function $K_1 F(K_2 x, K_3 y)$ should be $(\frac{1}{K_2}(\lambda_i)_x, \frac{1}{K_3}(\lambda_i)_y, K_1(\lambda_i)_z)$. where K_1, K_2, K_3 is random constant.*

Proof.

Supposed that $F(x,y)$are expressed parametric polynomials:

$$P(s,t) = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix}$$

then frequency spectrum of function $F(x,y)$ is $\lambda_{i,j} = \int_0^1 \int_0^1 P(s,t) U_{1,i} U_{1,j} dsdt$

$K_1 F(K_2 x, K_3 x)$ are expressed parametric polynomials:

$$P^{'}(s,t) = \begin{pmatrix} x^{'}(s,t) \\ y^{'}(s,t) \\ z^{'}(s,t) \end{pmatrix} = \begin{pmatrix} \frac{1}{K_2} x(s,t) \\ \frac{1}{K_3} y(s,t) \\ K_1 z(s,t) \end{pmatrix}$$

So $\lambda^{'}_{i,j} = \int_0^1 \int_0^1 P^{'}(s,t) U_{1,i} U_{1,j} dsdt = (\frac{1}{K_2}(\lambda_i)_x, \frac{1}{K_3}(\lambda_i)_y, K_1(\lambda_i)_z)$, where K_1, K_2, K_3 is random constant.

□

Corollary 5.6. *If frequency spectrum of function $F(x,y)$ is $\lambda_{i,j}, i, j = 0, 1, 2, ..., n$, then frequency spectrum of function $\frac{1}{K}(Kx, Ky)$ should be $\frac{1}{K}\lambda_{i,j}$, where K is random constant.*

6. The Properties of Frequency Spectrum

According to the above defined

$$\lambda_{i,j} = \int_0^1 \int_0^1 P(s,t) U_{1,i} U_{1,j} dsdt, \quad i,j = 0,1,2,...,15.$$

have the following properties. (See FIGURE 8).

Here, we only give the proof of translation transform. The others are similar.

Theorem 6.1. *if* $P^{'}(s,t) = P(s,t) + P_0$ *then* $\lambda^{'}_{i,j} = \lambda_{i,j} + P_0\delta_{i,j}$, *where*

$$\delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad or \quad j \neq 0 \\ 1, & i = 0 \quad and \quad j = 0 \end{cases}$$

Proof.

if $P^{'}(s,t) = P(s,t) + P_0$ then

$$\begin{aligned} \lambda^{'}_{i,j} &= \int_0^1 \int_0^1 (P(s,t) + P_0) U_{1,i} U_{1,j} dsdt \\ &= \int_0^1 \int_0^1 (P(s,t) U_{1,i} U_{1,j} dsdt + \int_0^1 \int_0^1 P_0 U_{1,i} U_{1,j} dsdt \\ &= \lambda_{i,j} + P_0\delta_{i,j} \end{aligned}$$

where $\delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad \text{or} \quad j \neq 0 \\ 1, & i = 0 \quad \text{and} \quad j = 0 \end{cases}$ □

Theorem 6.2. *Energy is invariant in rotation for the same geometric modeling.*

Proof.

Supposed geometric modeling rotated by θ circled coordinates axes z, then $x^{'} = x\cos\theta - y\sin\theta, y^{'} = y\cos\theta + x\sin\theta, z^{'} = z$,

So $(\lambda^{'}_{i,j})_x = \int_0^1 \int_0^1 (x(s,t)\cos\theta - y(s,t)\sin\theta) U_{1,i} U_{1,j} dsdt$

$= \cos\theta \int_0^1 \int_0^1 x(s,t) U_{1,i} U_{1,j} dsdt - \sin\theta \int_0^1 \int_0^1 y(s,t) U_{1,i} U_{1,j} dsdt$

$= (\cos\theta)(\lambda_{i,j})_x - (\sin\theta)(\lambda_{i,j})_y$.

The same as, $(\lambda^{'}_{i,j})_y = (\cos\theta)(\lambda_{i,j})_y - (\sin\theta)(\lambda_{i,j})_x$.

So $((\lambda^{'}_{i,j})_x)^2 + ((\lambda^{'}_{i,j})_y)^2 = ((\lambda_{i,j})_x)^2 + ((\lambda_{i,j})_y)^2$.

That is, $E^{'} = E$.

Experimental results see FIGURE 5. □

In order to make the frequency satisfy the invariance in rotation, translation, scale transforms, it can be defined normalized frequency as following:

$NF_{i,j} = \frac{||\lambda_{i,j}||}{||\lambda_{1,1}||}, i \neq 0$ or $j \neq 0$. Especially, when $i = 0, j = 0$, we call $\lambda_{0,0}$ as "DC" term.

Next, we prove the invariance in rotation, translation, scale transform.

Theorem 6.3. *Normalized frequency is invariant in rotation, translation, scale transforms for the same geometric modeling.*

Proof.

Let $\lambda'_{i,j}$ be obtained from$\lambda_{i,j}$, $\lambda_{i,j}$translated byP_0 and scaled by γ. The corresponding $\lambda'_{i,j}$ is

$$\begin{aligned}\lambda'_{i,j} &= \int_0^1 \int_0^1 \gamma(P(s,t)+P_0)U_{1,i}U_{1,j}dsdt \\ &= \gamma[\int_0^1 \int_0^1 (P(s,t)U_{1,i}U_{1,j}dsdt + \int_0^1 \int_0^1 P_0U_{1,i}U_{1,j}dsdt] \\ &= \gamma(\lambda_{i,j}+P_0\delta_{i,j}), \text{ where } \delta_{i,j} = \begin{cases} 0, & i \neq 0 \quad \text{or} \quad j \neq 0 \\ 1, & i = 0 \quad \text{and} \quad j = 0 \end{cases}\end{aligned}$$

So $NF'_{i,j} = \frac{||\lambda'_{i,j}||}{||\lambda'_{1,1}||} = \frac{||\lambda_{i,j}||}{||\lambda_{1,1}||} = NF_{i,j}$.

that is, normalized frequency is invariant in translation, scale transform.

Similarly, it is easily proved that normalized frequency is invariant in rotation transform. □

Accordingly, the "normalized energy" of $P(s,t)$ is defined as following:

$$NE = (\sum_{i=0}^{15}\sum_{j=0}^{15}(NF_{i,j}^2))^{\frac{1}{2}}, i \neq 0 \quad \text{or} \quad j \neq 0.$$

It is obvious that normalized energy satisfies the invariance in rotation, translation, scale transforms. So we obtain the theorem as following:

Theorem 6.4. *Normalized energy is invariant in rotation, translation, scale transforms for the same geometric modeling.*

7. Experimental Examples

For analysis and theorem the above, we have done lots of experiments to verify them. The experiment indicates that U-system is a new method for signal analyzing

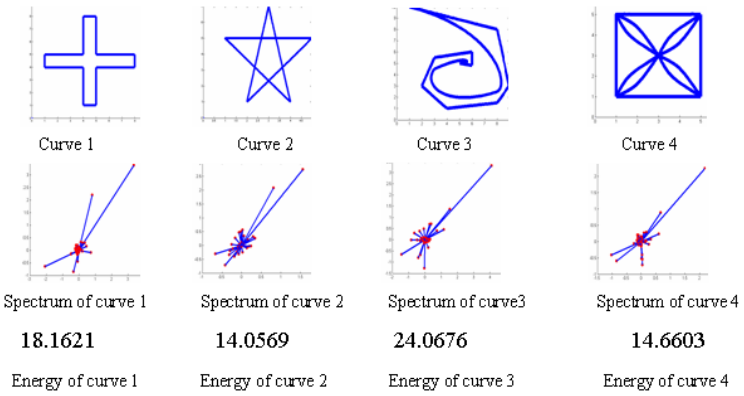


FIGURE 2. Spectral Analysis of Curves

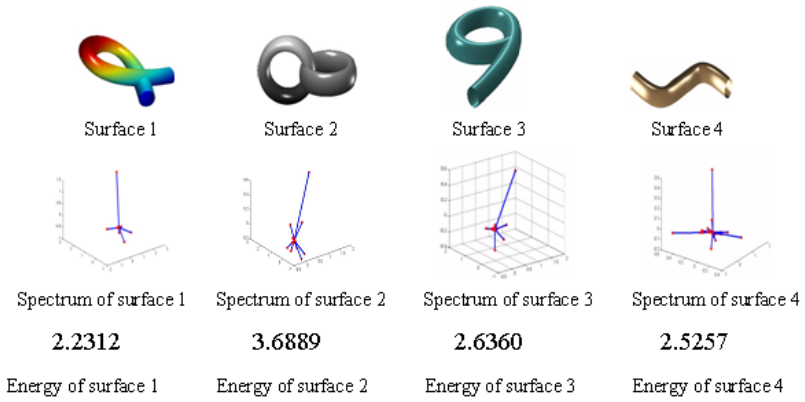


FIGURE 3. Spectral Analysis of Surfaces

and interesting exploring for analyzing frequency spectrum for geometric modeling. In this section, we will give some experimental results.

Experiment 1: In this experiment, frequencies spectrum of curves are denoted as vectors (λ_x, λ_y). We give the results of frequency and energy analysis of some curves by using three-degree U-system (See FIGURE 2). Using these spectrums,

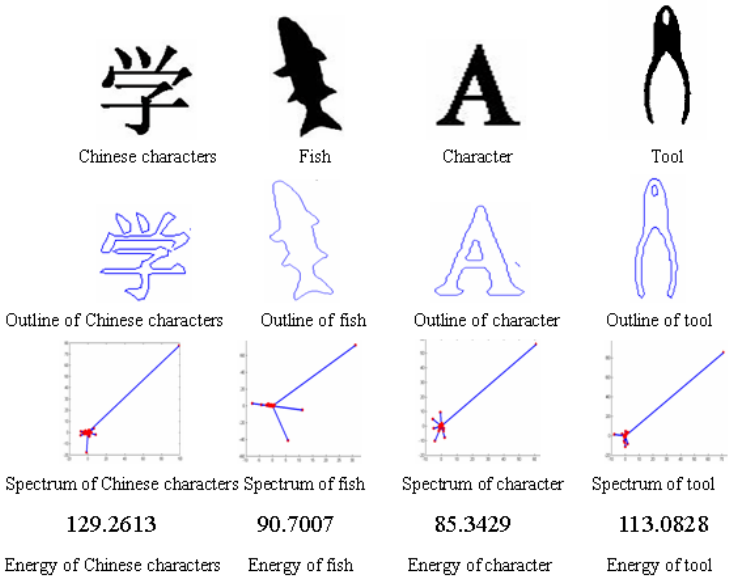
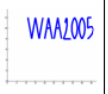
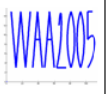
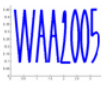
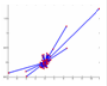
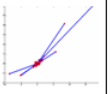
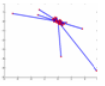
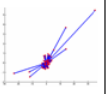
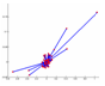
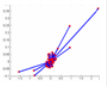
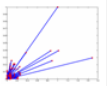
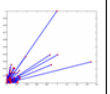
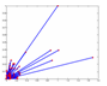
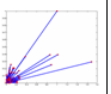
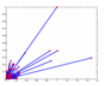
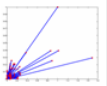


FIGURE 4. Spectral Analysis of Modeling

	Initial	Translation $P_0(9,7)$	Translation $P_0(-10,-12)$	Scale $(\alpha = 3)$	Scale $(\alpha = 0.09)$	Scale $(\alpha = 0.2)$
Curves						
Frequency						
Energy	76.3464	101.0656	59.5451	229.0393	6.8712	15.2693
Normalized Frequency						
Normalized Energy	3.1222	3.1222	3.1222	3.1222	3.1222	3.1222

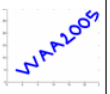
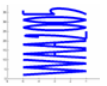
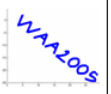
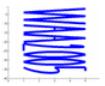
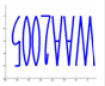
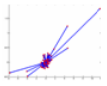
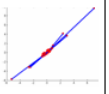
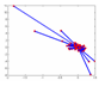
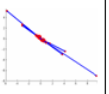
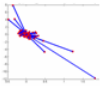
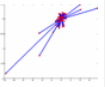
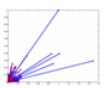
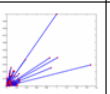
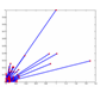
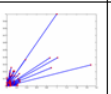
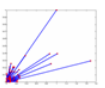
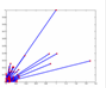
	Initial	Rotation $(\theta = \pi/4)$	Rotation $(\theta = \pi/2)$	Rotation $(\theta = -\pi/4)$	Rotation $(\theta = -\pi/2)$	Reflection $(\theta = \pi)$
Curves						
Frequency						
Energy	76.3464	76.3464	76.3464	76.3464	76.3464	76.3464
Normalized Frequency						
Normalized Energy	3.1222	3.1222	3.1222	3.1222	3.1222	3.1222

FIGURE 5. Normalized Frequency Analysis of Curves in U-system

we can exactly reconstruct these curves. However, the other orthogonal functions system, such as Fourier system, Legendre system and Chebyshev system, cannot completely reconstruct these curves by using finite items.

Experiment 2: In this experiment, frequencies spectrum of surfaces are denoted as vectors $(\lambda_x, \lambda_y, \lambda_z)$. We give the results of frequency and energy analysis of some surfaces by using three-degree U-system (See FIGURE 3). Using these

spectrums, we can exactly reconstruct these surfaces. However, the other orthogonal functions system, such as Fourier system, Legendre system and Chebyshev system, cannot completely reconstruct these surfaces by using finite items.

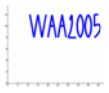
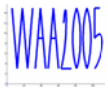
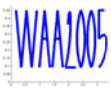
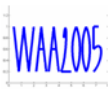
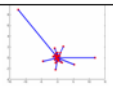
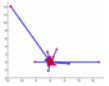
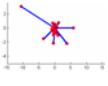
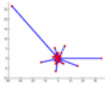
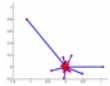
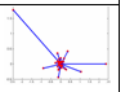
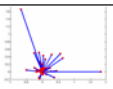
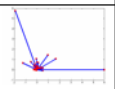
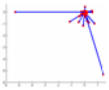
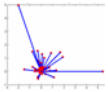
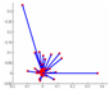
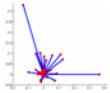
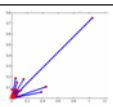
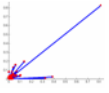
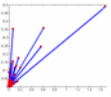
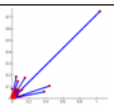
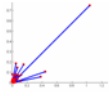
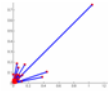
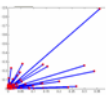
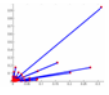
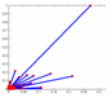
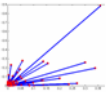
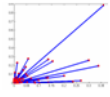
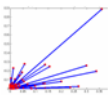
	Initial	Translation $P_0(9,7)$	Translation $P_0(-10,-12)$	Scale ($\alpha = 3$)	Scale ($\alpha = 0.09$)	Scale ($\alpha = 0.2$)
Curves						
Frequency X						
Frequency Y						
Energy	18.4158	24.5705	14.7281	54.5032	1.9652	3.9678
Normalized Frequency X						
Normalized Frequency Y						
Normalized Energy	2.0586	1.7859	2.4283	2.0586	2.0586	2.0586

FIGURE 6. Normalized Frequency Analysis of Curves in Fourier System

Experiment 3: In this experiment, frequencies spectrum of geometric modeling are denoted as vectors (λ_x, λ_y). We give the results of frequency and energy analysis of geometric modeling by using one-degree U-system (See FIGURE 4). Using these spectrums, we can exactly reconstruct these outlines of geometric modeling.However, the other orthogonal functions system, such as Fourier system, Legendre system and Chebyshev system, cannot completely reconstruct these outlines by using finite items.

Experiment 4: FIGURE 5. lists the results of normalized frequency analysis of curves by using three-degree U-system. In FIGURE 5, curves are denoted as vectors (λ_x, λ_y). Some results about "frequency and normalized frequency" and "energy and normalized energy" of group of geometric graphics are shown in FIGURE 5. FIGURE 5 indicates normalized frequency and normalized energy for the same geometric modeling are invariant in geometric transform, such as translation, scale and rotation.

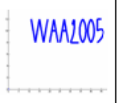
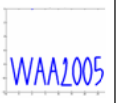
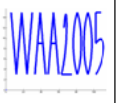
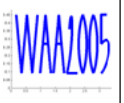
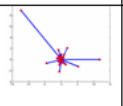
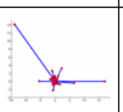
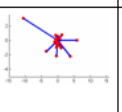
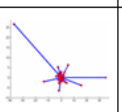
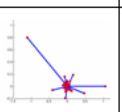
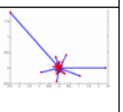
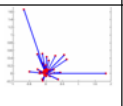
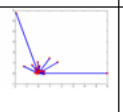
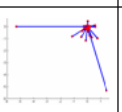
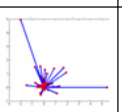
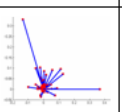
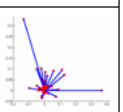
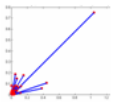
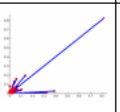
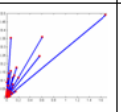
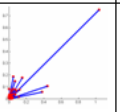
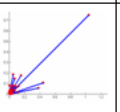
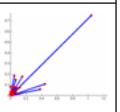
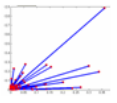
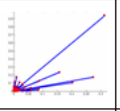
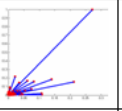
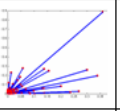
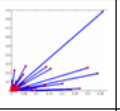
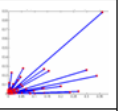
	Initial	Translation $P_0(9,7)$	Translation $P_0(-10,-12)$	Scale $(\alpha = 3)$	Scale $(\alpha = 0.09)$	Scale $(\alpha = 0.2)$
Curves						
Frequency X						
Frequency Y						
Energy	18.4158	24.5705	14.7281	54.5032	1.9652	3.9678
Normalized Frequency X						
Normalized Frequency Y						
Normalized Energy	2.0586	1.7859	2.4283	2.0586	2.0586	2.0586

FIGURE 7. Normalized Frequency Analysis of Curves in Fourier System

Geometric Transform	The curves or surfaces	Frequency
Identical Transform	$P(s,t)$	$\lambda_{i,j} = \int_0^1 \int_0^1 P(s,t) U_{1,i}(s) U_{1,j}(t) ds dt$
Translation	$P'(s,t) = P(s,t) + P_0$	$\lambda'_{i,j} = \lambda_{i,j} + P_0 \delta_{i,j}, where \quad \delta_{i,j} = \begin{cases} 0, i \neq 0, or, j \neq 0 \\ 1, i = 0, and, j = 0 \end{cases}$
Scale	$P'(s,t) = \alpha P(s,t)$	$\lambda'_{i,j} = \alpha \lambda_{i,j}$
Rotation	$[P'(s,t)^T,1] = [P(s,t)^T,1]T$	$[\lambda'_{i,j},1] = T[\lambda_{i,j},1]$

FIGURE 8. Frequency Spectrum Properties of Curves and Surfaces in the U-system

Experiment 5: FIGURE 6,7. lists the results of normalized frequency analysis of curves by using Fourier system. FIGURE 6,7, curves are denoted as vectors (λ_x, λ_y) . Some results about "frequency and normalized frequency" and "energy

Initial geometric modeling	Reconstructive curves using Fourier coefficient	Reconstructive curves using 3-degree U system
		
		

FIGURE 9. Comparing Fourier Transform with U system

and normalized energy" of geometric modeling in the Fourier system are shown in FIGURE 6,7. FIGURE 6,7 indicates normalized frequency and normalized energy for the same geometric modeling in the Fourier system are not all invariant in geometric transform, such as translation, scale and rotation.

FIGURE 5,6,7 indicate that using U-system with finite items, we can obtain an exact expression of geometric modeling. So it is more suitable for analysis of frequency spectrum for geometric modeling than Fourier system.

FIGURE 8. Frequency spectrum properties of curves and surfaces in the U-system.

FIGURE 9. Comparing Fourier transform with U system.

8. Conclusions and Future Work

From analysis the above, it can be concluded that, because of the strong continuity of Fourier system and Legendre system and Chebyshev system, they cannot exactly express common geometric modeling by using finite items, such as the phenomena of Gibbs (See FIGURE 9). So they are not suitable for analysis of frequency spectrum for geometric modeling of digital geometry [15][16]. Moreover, because of the strong discontinuity of Walsh and Harr systems, they cannot be used as a tool for those applications. However, using U-system with finite items, it can be realized exact expression of geometric modeling. According to the expression, we use frequency spectrum to calculate their "energy". So U-system is suitable for analysis of frequency spectrum for geometric modeling .In theory, we have proved that normalized frequency and normalized energy for the same geometric modeling are invariant in geometric transforms, such as rotation, translation, scale transforms. We also give some experimental results to support our conclusion.

In the theories and applications of communication and signal processing, U-system is the basic and important math tool. So, In future, we will further explore the properties of frequency spectrum of geometric modeling and applications in the field such as pattern recognition and digital signal processing.

Acknowledgment

This project is supported by "Mathematics mechanization method and its application on Information Technology" (the National Grand Fundamental Research 973 Program of China under Grant No.2004CB3180000), the National Natural Science Foundation of China (No. 60133020), Guangdong Provincial Science and Technology Foundation (No. 2004A10302005, No. 051Z205005), The China Scholarship Council and the Research Foundation of Sun Yat-sen University.

References

[1] Zhou Kun, Bao Hujun, Shi Jiaoying, *A Unified Framework for Digital Geometric Processing.* J. of Chinese Computers. **25** (2002), 904–909.

[2] Donald Hearm, M.Pauline Baker, *Computer Graphics with OpenGL.* 3rd Edition, Pearson Press, 2000.

[3] James D. Foley etc, *Computer Graphics, Principles and Practice.* Second Edition, Beijing, Engineering Industry Press, 2002.

[4] Wang Guojin, etc, *Computer Aided Geometric Design.* Higher Education Press and Springer Press, 2001. 1-17 (in Chinese).

[5] Shi Fazhong, *Computer Aided Geometric Design and Non-Uniform Rational B-Spline.* Higher Education Press, 2001(in Chinese)

[6] Chong Zhongkan,*Orthogonal functions and its applications.*1982.

[7] Henning F. Harmuth, *Sequency Theory Foundations and Applications.*1977.

[8] Yuyu Feng, Dongxu Qi, *A Sequence of Piecewise Orthogonal Polynomials.* J. of SIAM J. Math. Anal. **15(4)** (1984), 834–844.

[9] Charles A. Micchelli and Yuesheng Xu, *Using the Matrix Refinement Equation for the Construction of Wavelets on Invariant Sets.* Applied and Computational Harmonic Analysis. **1** (1994), 391–401.

[10] G. Battle, *Wavelets of Federbush-Lemarie type.* preprint, 1993

[11] P. Federbush, *A mass zero cluster expansion.* Part 1, The expansion, Comm. Math. Phys. **81** (1981), 327–340.

[12] Qi Dongxu, Feng Yuyu, *On the Convergence of Fourier-U Series.* J. of University of Science and Technology of China,Issue of Math.... **5** (1983), 7–17.

[13] Qi Dongxu, Feng Yuyu, *On the Orthogonal Complete system U.* Acta.Scientiarum Naturalium Universitatis Jilinensis **2** (1984), 21–31.

[14] Zhang Qishan, Zhang Youguang, *The Theory of Bridge Function and Its Application.*Beijing: National Defence Industry Press, 1992 (in Chinese).

[15] Dongxu Qi, *Frequency Spectrum Analysis in Digital Geometry (Special report).*Hefei City,2005.

[16] Dongxu Qi, Chenjun Tao,Ruixia Song, Hui Ma, Wei Sun, Zhanchuan Cai, *Representation for a Group of Parametric Cures Based on the Orthogomal Complete U-system.*J. of Computer,2005 (accepted).

[17] Hui Ma, Ruixia Song, Chenjun Tao, Dongxu Qi, *Orthogonal Complete U-system and Its Application in CAGD.*the first Korea-Chin Joint Conference on Geometric and Visual Computing,2005(accepted).

[18] Ruixia Song, Hui Ma, Dongxu Qi, *A New Class of Complete Orthogonal Functions and Its Applications Based on U-system.*The 4th International Conference on Wavelet Analysis and Its Applications 2005 (accepted).

[19] Y.Y.Feng and D.X.Qi, *On the Harr and Walsh Systems on a Triangle.*MRC Technical Summary Report No.2235, University of Wisconsin-Madison, 1981.

[20] Knut Morken, *Computer Aided Geometric Design.*Computer Aided Geometric Design. **22(9)** (2005), 838–848.

[21] Dinesh Manocha and John F. Canny, *Algorithm for implicitizing rational parametric surfaces.*Computer Aided Geometric Design. **9(1)** (1992), 25–50.

Zhanchuan Cai
School of Information Science and Technology,
Sun Yat-sen University Guangzhou,
510275 P.R. China

Hui Ma
Faculty of Information Technology,
Macao University of Science and Technology,
Macao

Wei Sun
School of Information Science and Technology,
Sun Yat-sen University Guangzhou,
510275 P.R. China

Dongxu Qi
School of Information Science and Technology,
Sun Yat-sen University Guangzhou,
510275 P.R. China;
Faculty of Information Technology,
Macao University of Science and Technology,
Macao

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 543–559
© 2006 Birkhäuser Verlag Basel/Switzerland

Detection of Spindles in Sleep EEGs Using a Novel Algorithm Based on the Hilbert-Huang Transform

Zhihua Yang, Lihua Yang and Dongxu Qi

Abstract. A novel approach for detecting spindles from sleep EEGs (electroencephalograph) automatically is presented in this paper. Empirical mode decomposition (EMD) is employed to decompose a sleep EEG, which are usually typical nonlinear and non-stationary data, into a finite number of intrinsic mode functions (IMF). Based on these IMFs, the Hilbert spectrum of the EEG can be calculated easily and provides a high resolution time-frequency presentation. An algorithm is developed to detect spindles from a sleep EEG accurately, experiments of which show encouraging detection results.

Mathematics Subject Classification (2000). Primary 94A13; Secondary 42A16.

Keywords. Empirical mode decomposition, Hilbert-Huang transform, signal detection.

1. Introduction

Sleep is a complicated physiological process. Research on sleep is very important to both clinical diagnosis and curative effect evaluation in nervous psychiatry. Generally, sleep consists of two phases: no-rapid eye movement (NREM) and rapid eye movement(REM). The NREM phase can be decomposed into 4 stages according to sleep depths [6]. A crucial clue for the sleep depth to be at the second or the third stages is that the sleep-spindles, whose frequencies are between 12 and 20Hz, take place in the EEG (electroencephalo-graph) [7].

Traditionally, sleep-spindles are detected visually by neurologists or sleep experts. Research on automated sleep analysis can be traced back to as early as the 1970s [12, 13, 9, 11, 2, 8, 10, 5]. In recent years, two novel algorithms for automated

This work was supported by NSFC (Nos. 60475042, 60133020), the National Key Basic Research Project of China (No.2004CB318000) and the Scientific and Technological Planning Project of Guangzhou city.

detection of spindles in sleep EEG were developed by using classical time-frequency analysis [7, 3]. Since EEGs are typically nonlinear and non-stationary signals and the duration of a sleep-spindle is usually very short, it is usually difficult to obtain satisfactory results in automated detection of sleep-spindles by using traditional time-frequency analysis.

Recently, a novel analysis method for nonlinear and non-stationary data, which is called *Hilbert-Huang Transform (HHT)*, was developed [4]. Its key part is the so-called *empirical mode decomposition (EMD)*, with which any complicated data set can be decomposed into a finite and often small number of *intrinsic mode function(IMF)* that admit well-behaved Hilbert transforms. EMD is adaptive, and therefore, highly efficient. It is based on the local characteristic time scale of the data and is applicable to nonlinear and non-stationary processes. With the Hilbert transform, the IMFs yield instantaneous frequencies as functions of time that give sharp identifications of embeded structures. The final presentation of the results is a time-frequency-energy distribution, designated as the Hilbert spectrum, which has high time-frequency localization.

Because of these properties, in this paper, HHT is employed to analyze sleep EEGs and an algorithm to detect spindles from sleep EEGs automatically is developed. Experiments show an encouraging detection rate which is higher than those developed in [7, 3].

This paper is organized as follows: Section 2 is a brief summary on the Hilbert-Huang Transform; Analysis of EEG data based on HHT is given in Section 3; In Section 4, a novel algorithm for automated detection of sleep-spindles is proposed. Experiments to support the algorithm are conducted and the anti-noise ability is discussed in Section 5; Finally, Section 6 is the conclusion of this paper.

2. Hilbert-Huang Transform

The Hilbert-Huang Transform (HHT) was proposed by Huang et al [4]. It consists of two parts: (1) *Empirical Mode Decomposition (EMD)*, and (2) *Hilbert Spectral Analysis*. With EMD, any complicated data set can be decomposed into a finite and often small number of *intrinsic mode functions (IMF)*. An IMF is defined as a function satisfying the following conditions:

(a) The number of extrema and the number of zero-crossings must either be equal or differ at most by one;
(b) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

An IMF defined as above admits well-behaved Hilbert transforms. EMD decomposes signals adaptively and is applicable to nonlinear and non-stationary data (Fundamental theory on nonlinear time series can be found in [1]). In this section, a brief introduction is given to make this paper somewhat self-contained. The readers are referred to [4] for details.

For an arbitrary function $X(t)$ in L_p-class [14], its Hilbert transform $Y(t)$ is defined as

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{X(t')}{t - t'} \mathrm{d}t', \tag{2.1}$$

where P indicates the Cauchy principal value. Consequently an analytic signal $Z(t)$ can be produced by

$$Z(t) = X(t) + iY(t) = a(t)e^{i\theta(t)}, \tag{2.2}$$

where

$$a(t) = [X^2(t) + Y^2(t)]^{\frac{1}{2}}, \ \theta(t) = \arctan(\frac{Y(t)}{X(t)}) \tag{2.3}$$

are the instantaneous amplitude and the phase of $X(t)$. Since the Hilbert transform $Y(t)$ is defined as the convolution of $X(t)$ and $1/t$ by Eq. (2.1), it emphasizes the local properties of $X(t)$ even though the transform is global. In Eq. (2.2), the polar coordinate expression further clarifies the local nature of this representation. With Eq. (2.2), the instantaneous frequency of $X(t)$ is defined as

$$\omega(t) = \frac{d\theta(t)}{dt}. \tag{2.4}$$

However, there is still considerable controversy on this definition. A detailed discussion and justification can be found in [4].

EMD is a necessary pre-processing of the data before the Hilbert transform is applied. It reduces the data into a collection of IMFs and each of them represents a simple oscillatory mode that is a counterpart of a simple harmonic function, but is much more general. With this definition, one can decompose any function according to the following algorithm.

Algorithm 2.1. Let $X(t)$ be a signal.

Step 1 Initialize: $r_0(t) = X(t), \ i = 1$;

Step 2 Extract the i-th IMF as follows:

- **(a)** Initialize: $h_0(t) = r_{i-1}(t), \ j = 1$;
- **(b)** Extract the local minima and maxima of $h_{j-1}(t)$;
- **(c)** Interpolate the local maxima and the local minima by cubic splines to form $u_{j-1}(t)$ and $l_{j-1}(t)$ as the upper and lower envelops of $h_{j-1}(t)$ respectively;
- **(d)** Calculate $m_{j-1}(t) = \frac{u_{j-1}(t)+l_{j-1}(t)}{2}$ as an approximation of the local mean of $h_{j-1}(t)$ at t;
- **(e)** Let $h_j(t) = h_{j-1}(t) - m_{j-1}(t)$;
- **(f)** If the stopping criterion is satisfied, i.e., $h_j(t)$ is an IMF, set $\mathrm{imf}_i(t) = h_j(t)$; Else go to (b) with $j = j + 1$.

Step 3 Let $r_i(t) = r_{i-1}(t) - \mathrm{imf}_i(t)$;

Step 4 If $r_i(t)$ still has at least 2 extrema, go to Step 2 with $i = i + 1$; otherwise the decomposition is finished and $r_i(t)$ is the residue.

By Algorithm 2.1, any signal $X(t)$ can be decomposed into finite IMFs, $\mathrm{imf}_j(t)$ $(j = 1, \cdots, n)$, and a residue $r(t)$, where n is nonnegative integer depending on $X(t)$, i.e.,

$$X(t) = \sum_{j=1}^{n} \mathrm{imf}_j(t) + r(t). \tag{2.5}$$

For each $\mathrm{imf}_j(t)$, let $X_j(t) = \mathrm{imf}_j(t)$. Its corresponding instantaneous amplitude $a_j(t)$ and instantaneous frequency $\omega_j(t)$ can be computed with Eqs. (2.3) and (2.4). By Eqs. (2.2) and (2.4), $\mathrm{imf}_j(t)$ can be expressed as the real part, Re, in the following form:

$$\mathrm{imf}_j(t) = \mathrm{Re}\left[a_j(t) \exp\left(i \int \omega_j(t) dt\right)\right]. \tag{2.6}$$

Therefore, by Eqs. (2.5) and (2.6), $X(t)$ can be expressed as the IMF:

$$X(t) = \mathrm{Re} \sum_{j=1}^{n} a_j(t) \exp\left(i \int \omega_j(t) dt\right) + r(t). \tag{2.7}$$

It is interesting to compare the representation above with the following classical Fourier expansion:

$$X(t) = \sum_{j=1}^{\infty} a_j e^{i\omega_j t}, \tag{2.8}$$

where both a_j and ω_j are constants. Contrasting Eq. (2.7) with Eq. (2.8), it is apparent that the IMF expansion provides a generalized Fourier expansion. It relieves the restriction of the constant amplitude and fixed frequency of Fourier expansion, and arrives at a variable amplitude and frequency representation. With the IMF expansion, the amplitude and frequency modulations are clearly separated. Its main advantage over Fourier expansion is that it accommodates nonlinear and non-stationary data perfectly.

Equation (2.7) enables us to represent the amplitude and the instantaneous frequency as functions of time in a three-dimensional plot, in which the amplitude is contoured on the time-frequency plane. The time-frequency distribution of amplitude is designated as the Hilbert amplitude spectrum or simply Hilbert spectrum, denoted by $H(\omega, t)$. It can also be defined equivalently in mathematics as follows: Let $X(t)$ be decomposed into finite IMFs $\mathrm{imf}_j(t)$ $(j = 1, \cdots, n)$ and a residue $r(t)$ by Algorithm 2.1, then,

$$H(\omega, t) = \begin{cases} 0 & \text{if } J_{\omega,t} \text{ is an empty set,} \\ \sum_{j \in J_{\omega,t}} a_j(t), & \text{otherwise,} \end{cases} \tag{2.9}$$

where $J_{\omega,t} = \{j | 0 \le j \le n \text{ satisfying } \omega_j(t) = \omega\}$.

$H(\omega, t)$ gives a time-frequency-amplitude distribution of a signal $X(t)$. If amplitude squared is more desirable commonly to represent energy density, then the squared values of amplitude can be substituted to produce the Hilbert energy spectrum.

3. HHT for EEG Data

To observe the performance of the HHT for EEG data, a 6 second segment, denoted by $X(t)$, is selected from a sleep EEG which is sampled when the sleep is at the 2nd stage of a NREM phase. It contains two sleep-spindles, marked by 'A' and 'B' respectively, as shown in Fig.1. With Algorithm 2.1, it is decomposed into seven IMFs and a residue, which are shown in Fig. 2(c1) $\sim$ (c8) respectively from top to bottom. From Fig.2(c2), it is easy to see that there are two sub-segments of high amplitude in the periods from about 20 to 160 and from about 750 to 1050 respectively, whose frequencies are around 13Hz (the first consists of about 9 waves within 0.7s(140 points) and the second consists of 18 waves within 1.5s(300 points)). Such a sub-segment whose amplitude is high and frequency is between 12 $\sim$ 20Hz is a possible spindle wave we want to detect, which is called a *PSW* for simplicity. Based on the analysis of this example, it seems possible to detect sleep-spindles based on features in some of the IMFs, such as the second one shown in Fig.2(c2). To examine the observation, let us consider another example as shown in Fig.3. It is also a sleep EEG segment of 6 seconds, in which only one sleep-spindle, marked by 'A', is included. Similarly, with Algorithm 2.1, it is decomposed into seven IMFs and a residue, which are shown in Fig. 4(c1) $\sim$ (c8) respectively from top to bottom. It is observed that the second IMF does not contain PSWs. However, in the third IMF as shown in Fig.4(c3), there is a PSW starting at about the 800 and ending at the 1000 (corresponding the part marked by 'A' in Fig.3). By conducting more experiments like Fig.2 and Fig.4, we conclude that, for a sleep EEG segment which contains spindles, in general, one cannot determine in which IMF the interesting sub-segments may appear. In fact, a spindle consists of a number of oscillatory modes which are the same or similar in local characteristic time scale (an oscillatory mode is a wave between a pair of successive maxima or a pair of successive minima [4]), when a EEG segment is decomposed by EMD, these oscillatory modes may be dispersed to different IMFs. Therefore, the PSWs cannot be detected based on some IMFs and consequently

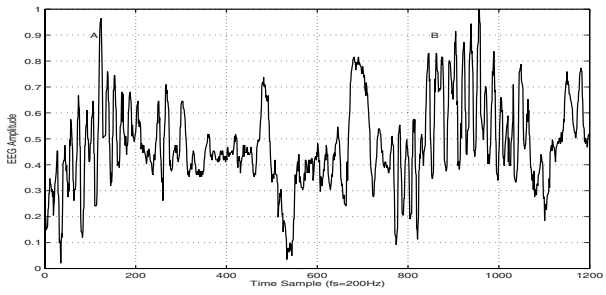


FIGURE 1. A sleep EEG segment of 6 seconds at the second sleep stage, in which two sleep-spindles are included as marked by 'A' and 'B' respectively.

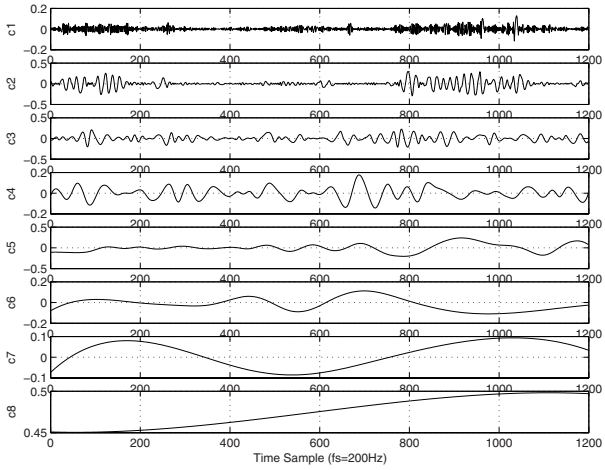


FIGURE 2. The resulting EMD components from the EEG data of Fig. 1. The last component, c8, is not an IMF, it is the residue.

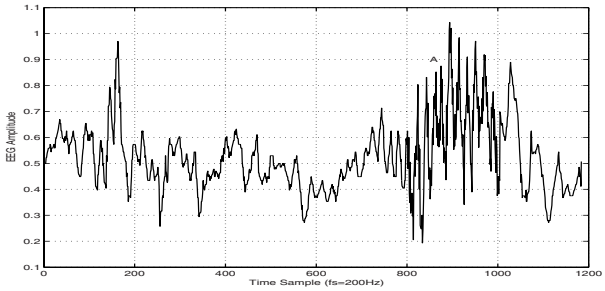


FIGURE 3. A sleep EEG segment of 6 seconds long at the second sleep stage, in which a sleep-spindle is included as marked by 'A'.

the Hilbert spectrum, which is the global time-frequency-energy distribution, is considered as a natural substitution for the detection of sleep-spindles from EEGs.

Since IMFs have good time-frequency resolution as described in Section 2 (see [4]), one can predict that, in the Hilbert spectrum of a sleep EEG segment, high energies will take place within its PSWs. This prediction is verified by the graphs of the Hilbert spectrum of the sleep EEG segments: Fig. 5 and Fig. 6 are respectively the contour maps of the Hilbert spectrums of the sleep EEG segments shown in Fig. 1 and Fig. 3. In Fig. 5, one can find two high energy bands whose frequencies are $8 \sim 20$Hz or so: one starts at about 0 and ends at 200, the other starts at about 750 and ends at 1050, as highlighted by the two rectangles in the figure. Similarly, in Fig. 6, there is a high energy band whose frequencies are $8 \sim 20$Hz or so, starting at about 730 and ending at 1000, as highlighted by the

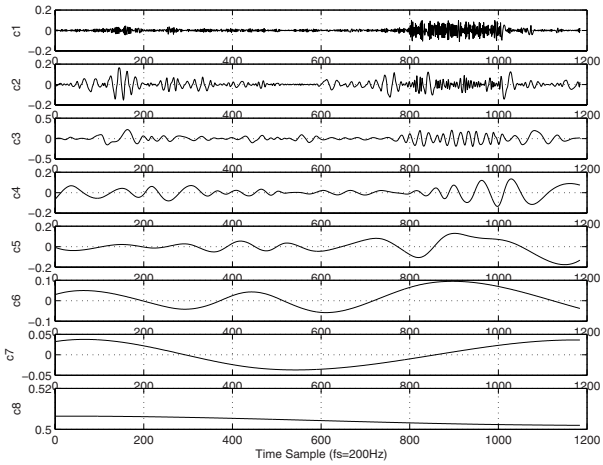


FIGURE 4. The resulting EMD components from the EEG data of Fig. 3. The last component, c8, is not an IMF, it is the residue.

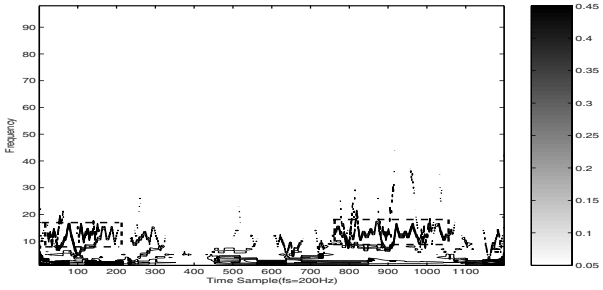


FIGURE 5. The contour map of Hilbert spectrum for the EEG data of Fig. 1.

rectangle in Fig. 6. Therefore, the locations and durations of the sleep-spindles can be detected from the Hilbert spectrum of sleep EEGs successfully.

4. A Novel Algorithm for Automated Detection of Sleep-Spindles

In this section, an auto-detection algorithm for sleep-spindle detection is developed and, consequently, experiments are conducted to support the algorithm.

Based on the discussion above and the fact that the duration of a sleep-spindle is usually longer than 0.5s in practice, a novel method for automated detection of spindles from a EEG is developed in this section.

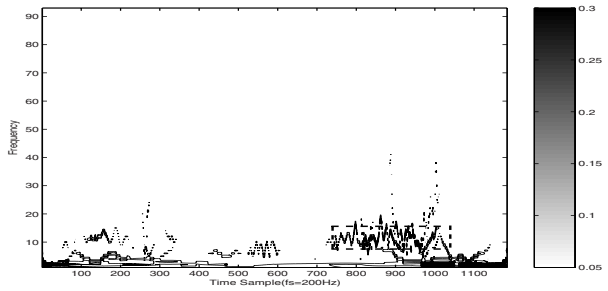


FIGURE 6. The contour map of Hilbert spectrum for the EEG data of Fig. 3.

In practice, a sleep EEG contains a large amount of data. It is terribly time-consuming to calculate the Hilbert spectrum of a global sleep EEG signal. Experiments show that the time needed for calculating the Hilbert spectrum, which is called the CPU time hereafter for simplicity, does not depend linearly on the data length of the signal. When the length of data is longer than 1500 or so, the CPU time needed increases rapidly. The five dotted curves in Fig. 7 illustrates how the CPU times depend on the datum lengths for five signals of 2000 random data ranging from 0 to 1, when a PC of Pentium IV-1.7GHz is used. It should be pointed out that the graph changes somewhat if the signals is replaced by another signal.

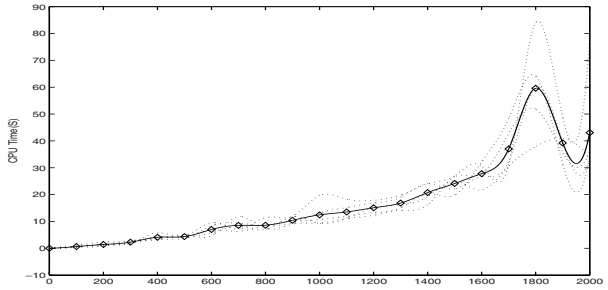


FIGURE 7. CPU time as a function of the length of the data set

The solid curve in Fig. 7 is the average of the five dotted curves, which illustrates how the CPU time depends on the datum length in the rough. Thus, to save CPU time, a global EEG signal should be divided into many short segments. The length of the short segments should be as long as possible to keep enough integrality of data if the CPU time is acceptable. According to Fig. 7, it is a reasonable tradeoff to divide the global EEG signal into segments of $1200 \sim 1600$ data (about $6 \sim 8$ seconds for frequency of sampling = 200Hz). In our experiments, the global sleep EEG is divided into segments of 1200 data points. The corresponding time for each segment is about 6 seconds.

After a sleep EEG is divided as before, we decompose each segement with EMD into IMFs. Since the first IMF usually consists of the highest frequency components of the segment, such as noise (see Fig. 2(c1) and Fig. 4(c1)), it is discarded in our algorithm. For each segment, the sleep-spindles are detected by the following algorithm.

Algorithm 4.1. Let $x(t)$ be a segment of data of length 1200. The sleep-spindles are detected as follows:

Step 1 Decompose $x(t)$ with EMD into IMFs, then remove the first IMF. For each other IMF, calculate its instantaneous frequency and instantaneous amplitude by Eqs. (2.4) and (2.3). Quantify the instantaneous frequency into integers between 1 and 100Hz.

Step 2 Compute the Hilbert spectrum $H(\omega, t)$; here it is a matrix of 100 rows and 1200 columns. Then normalize the amplitude of $H(\omega, t)$ linearly such that the values of $H(\omega, t)$ range from 0 to 255.

Step 3 Extract the 8th to 20th rows of $H(\omega, t)$ to form a sub-matrix, denoted by M, of 13 rows and 1200 columns.

Step 4 Calculate the maximum of each column of M to generate an array, $C = (C[1], \cdots, C[1200])$. It is an energy measure of the data on frequencies ranging from $8 \sim 20$Hz at each local time. Then, define a smoothed version of C as:

$$C_1[k] = \frac{1}{L} \sum_{i=k-L/2}^{k+L/2} C[i],$$

where L, an even integer, is the width of the smoothing window ($L = 50$ in the experiments of this paper) and the boundary extension is conducted as: $C[i] = C[1]$ for $i \leq 0$ and $C[i] = C[1200]$ for $i > 1200$.

Step 5 Let T be a threshold. Then, we search $1 \leq k \leq 1100$ and $I \geq 100$ such that

$$C_1[k + i - 1] \geq T \text{ for } i = 1, 2, \cdots, I,$$

and

$$C_1[k + I] < T \text{ or } k + I = 1200.$$

Then a sleep-spindle that starts at k and has duration I is detected. We set $T = 50$ in the experiments of this paper.

5. Experiments

In this section, experiments are conducted to support our algorithm on sleep-spindle detection.

For the segment of a sleep EEG shown in Fig. 1, according to Steps 1 and 2 of Algorithm 4.1, its IMFs and Hilbert spectrum $H(\omega, t)$ are calculated as shown in Figs. 2 and 5 respectively. By Step 3 of Algorithm 4.1, the sub-matrix M of the 8th to 20th rows of $H(\omega, t)$ is generated as shown in Fig. 8. Then the energy array

C of M and the smoothed version, C_1, are calculated in accordance with Step 4 of Algorithm 4.1, which are shown in Fig. 9: the top is C and the bottom is C_1. Finally, by Step 5, two spindles of a sleep EEG are detected in this segment as shown in Fig. 10, in which the starting points, the durations, and the end points are marked by the dotted lines. The first starts at about the 20th datum (namely: the 0.1th second) with a duration of about 0.8s and the second starts at about the 750th datum (namely: the 3.75th second) with a duration of about 1.5s.

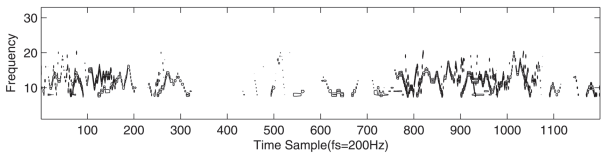


FIGURE 8. The sub-matrix M generated by the 8th – 20th rows of the Hilbert spectrum $H(\omega, t)$ shown in Fig. 5

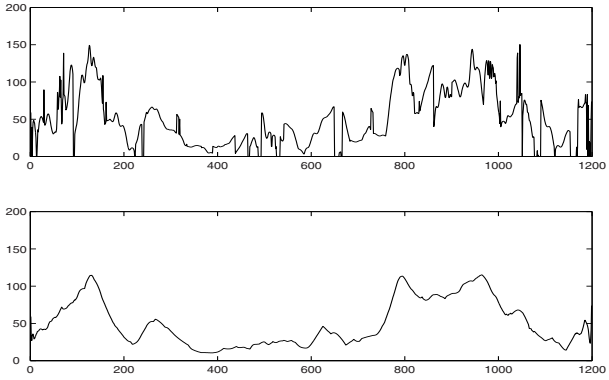


FIGURE 9. The top: The energy array of M shown in Fig. 8, C, each of whose component is the maximum of the elements in the corresponding column of M. The bottom: The smoothed version C_1 of C calculated according to Step 4 of Algorithm 4.1.

Similarly, Figs. 4 and 6 are the IMFs and the contour map of the Hilbert spectrum $H(\omega, t)$ of the segment shown in Fig. 3 in accordance with Steps 1 and 2 of Algorithm 4.1. The corresponding sub-matrix M calculated by Step 3 are shown in Fig. 11. Then, the energy array C of M and the smoothed version, C_1, calculated by Step 4, are displayed graphically in Fig. 12: the top is C and the bottom is C_1. Fig. 13 is the detection result by Step 5, in which a sleep-spindle is detected and marked with dotted lines. It starts at about the 750th datum (namely: the 3.75th second) with a duration of about 1.25s.

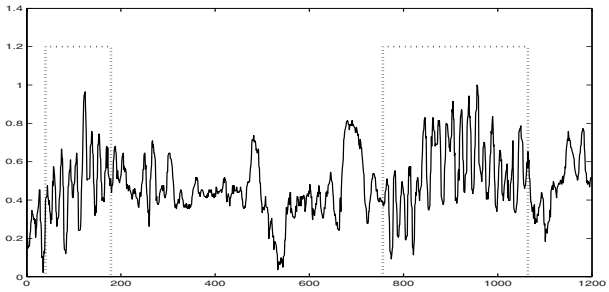


FIGURE 10. The detection result for the segment in Fig. 1. Two sleep-spindles are detected and marked with dotted lines. The first starts at about the 20th datum (namely: the 0.1th second) with a duration of about 0.8s and the second starts at about the 750th datum (namely: the 3.75th second) with a duration of about 1.5s.

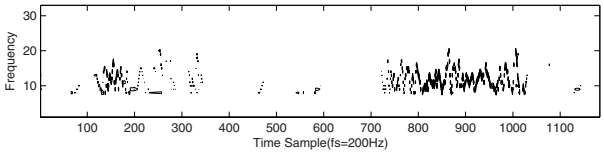


FIGURE 11. The sub-matrix M generated by the 8th – 20th rows of the Hilbert spectrum $H(\omega, t)$ shown in Fig. 6

To test our detection algorithm, 100 segments, each of which consists of 1200 data (about 6 seconds with frequency of sampling 200Hz) and all of which contain 183 spindles, are selected from a sleep EEG database. The locations and durations of these sleep-spindles have been determined visually by experts. For an automated detection algorithm, its detection accuracy depends on two aspects: (1) the accuracy of the location detected and (2) the accuracy of the duration detected. To estimate them quantitatively, a mis-detection degree is introduced as follows.

Definition 5.1. Let X(t) be a sleep EEG segment which contains a sleep-spindle starting at t_b and ending at t_e. For an automated detection method for sleep-spindles, the mis-detection degree, simply denoted by MD, is defined as follows:

1. if one sleep spindle is detected from X(t), with starting point t_b' and end point t_e', then
$$MD = \frac{L_\vee - L_\wedge}{L}, \tag{5.1}$$
where, $L_\vee = \max(t_e, t_e') - \min(t_b, t_b')$, $L_\wedge = \min(t_e, t_e') - \max(t_b, t_b')$ and $L = t_e - t_b$ as shown in Fig. 14.
2. if no spindle or more than one spindle is/are detected, then $MD = \infty$.

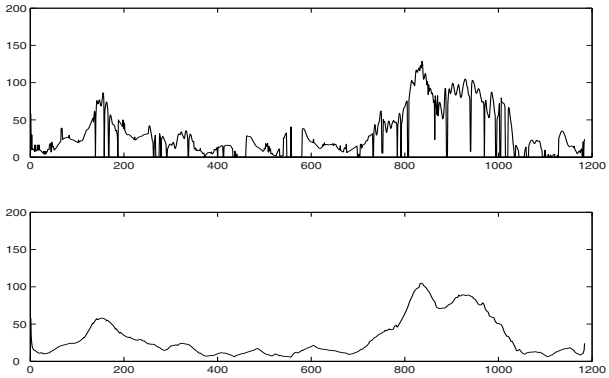


FIGURE 12. The top: The energy array of M, shown at the bottom of Fig. 11, denoted by C, each of whose component is the maximum of the elements in the corresponding column of M. The bottom: The smoothed version C_1 of C calculated according to Step 4 of Algorithm 4.1.

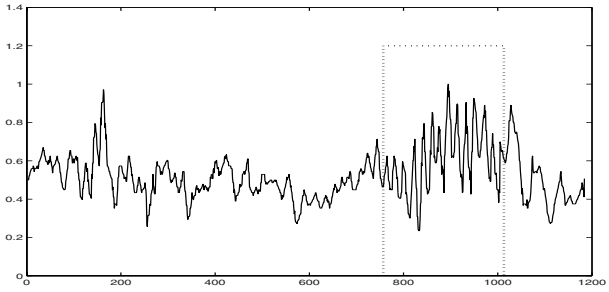


FIGURE 13. The detection result for the segment in Fig. 3. A sleep-spindle is detected and marked with dotted lines, which starts at about the 750th datum (namely: the 3.75th second) with a duration of about 1.25s.

It is easy to see that MD is a nonnegative number and MD=0 if and only if the sleep-spindle is detected accurately, and the smaller MD is, the more accurately the detection does. Table 1 lists the distribution of the MDs produced by Algorithm 4.1 for all the 183 samples and the corresponding histogram is displayed in Fig. 15, in which all the MDs greater than 1 is included into that of MD=1.1. It is encouraging to see that most of the MDs are between 0 and 0.2, which shows that our algorithm arrives at satisfying detection results in both locations and durations.

Let us compare our technique with those in [3, 7]. The same dataset is employed to conduct the experiments. The distributions of the corresponding MDs

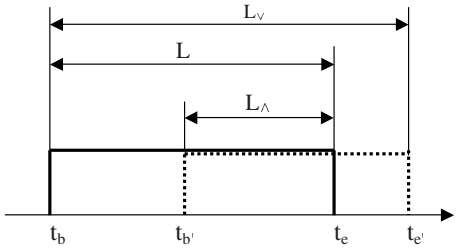


FIGURE 14. $L_\vee, L_\wedge$ and L in Definition 2

TABLE 1. The distribution of the MDs produced by Algorithm 4.1, NS is the number of spindles whose MDs are within the given interval.

MD	0 ∼0.1	0.1 ∼0.2	0.2 ∼0.3	0.3 ∼0.4	0.4 ∼0.5	0.5 ∼0.6	0.6 ∼0.7	0.7 ∼0.8	0.8 ∼0.9	0.9 ∼1	> 1
NS	1101	443	86	51	43	25	17	7	3	3	51

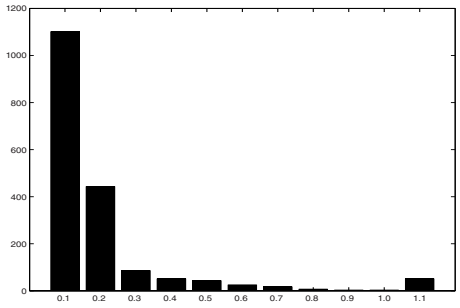


FIGURE 15. The histogram of the MDs corresponding to Table 1

calculated according to the algorithms in [7] and [3] are listed in the 3th and 4th rows of Table 2 respectively. To compare more conveniently, the distribution of the MDs by our algorithm is also listed in the 2nd row of Table 2. The corresponding histogram is shown in Fig. 16, in which all the MDs greater than 1 is included into that of MD=1.1. It is noticed that the MDs produced by our algorithm concentrate much closer to 0 than those by [3] and [7], which implies that our detection does better than theirs. To describe the detection result quantitatively, we define the detection rate as the ratio of the number of MDs which are less than some given threshold, T, to the total number of spindles. With this definition and $T = 0.2, 0.5$ and 1 respectively, the detection rates according to Algorithm 4.1, the algorithms in [7] and [3] are listed in Table 3.

Before the end of this section, let us discuss the detection of spindles from noisy sleep EEGs by Algorithm 4.1. It is easy to understand that Algorithm 4.1

TABLE 2. The distributions of the MDs calculated according to Algorithm 4.1, the algorithms in [7] and [3]. NS_j are the numbers of spindles whose MDs are within the given intervals, with j=1, 2, and 3 corresponding to Algorithm 4.1, the algorithms in [7] and in [3] respectively.

MD	0 ∼0.1	0.1 ∼0.2	0.2 ∼0.3	0.3 ∼0.4	0.4 ∼0.5	0.5 ∼0.6	0.6 ∼0.7	0.7 ∼0.8	0.8 ∼0.9	0.9 ∼1	> 1
NS_1	1101	443	86	51	43	25	17	7	3	3	51
NS_2	893	466	164	88	42	23	12	9	11	9	113
NS_3	752	367	231	97	44	48	26	12	10	13	230

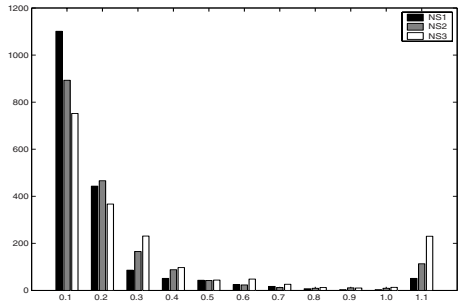


FIGURE 16. The histogram of the MDs corresponding to Table 2

TABLE 3. The detection rates corresponding to Algorithm 4.1, the algorithms in [7] and [3] for $T = 0.2, 0.5$ and 1

Detection rate	$T = 0.2$	$T = 0.5$	$T = 1$
Algorithm 4.1	84.4%	94.2%	97.2%
Algorithm in [7]	74.3%	90.3%	93.8%
Algorithm in [3]	61.1%	81.5%	87.4%

can do well for noisy sleep EEG signals since the frequencies of sleep spindles, about 13Hz, are usually much lower than those of noise. Another reason for this anti-noise ability is that the first IMF has no contribution to the Hilbert spectrum. To demonstrate our view, a great number of experiments have been conducted and excellent results are obtained. Fig. 17 are three noisy versions of Fig. 3 by adding Gaussian white noises with SNR (Signal Noise Ratio, which is defined as the ratio of signal variance to noise variance in dB.) 20dB, 25dB and 30dB respectively from top to bottom. The detection results are shown in Fig. 18, corresponding to those in Fig. 17 from top to bottom. It is easy to see that, spindles can be detected successfully with a minor change in the locations and durations even though the SNRs are very low.

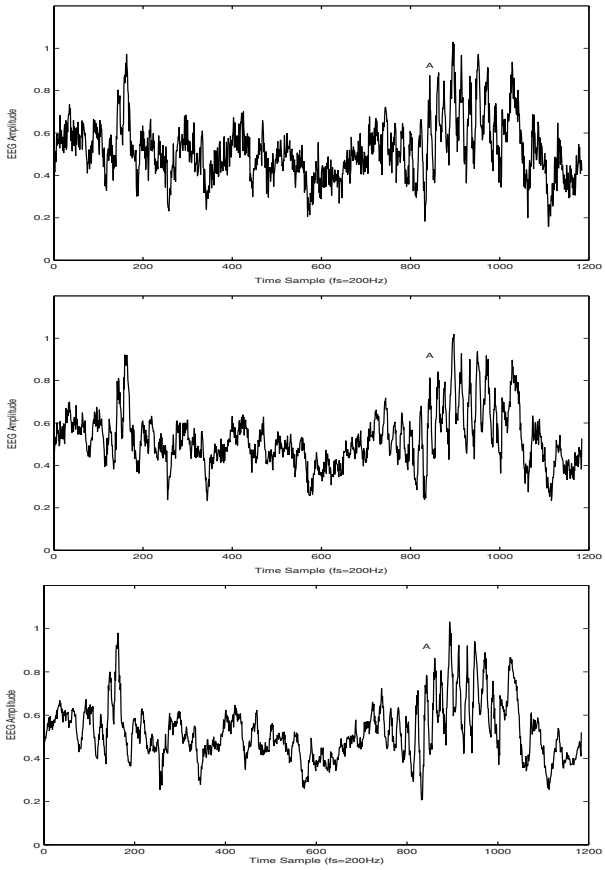


FIGURE 17. Three noisy versions of Fig. 3 by adding Gaussian white noises with SNR 20, 25 and 30 respectively from top to bottom.

6. Conclusion

In this paper, a novel approach for detecting spindles automatically from a sleep EEG based on the Hilbert-Huang transform is developed. Empirical mode decomposition is employed to decompose sleep EEGs into a finite and often small number of intrinsic mode functions. Then the Hilbert spectrum $H(\omega, t)$ is used to give a high resolution time-frequency presentation and extract features of EEGs. Consequently, an algorithm is proposed to detect spindles automatically from a sleep EEG. Experiments show more accurate detection results than those in [7, 3]. Finally, the anti-noise ability of the algorithm is demonstrated theoretically and experimentally.

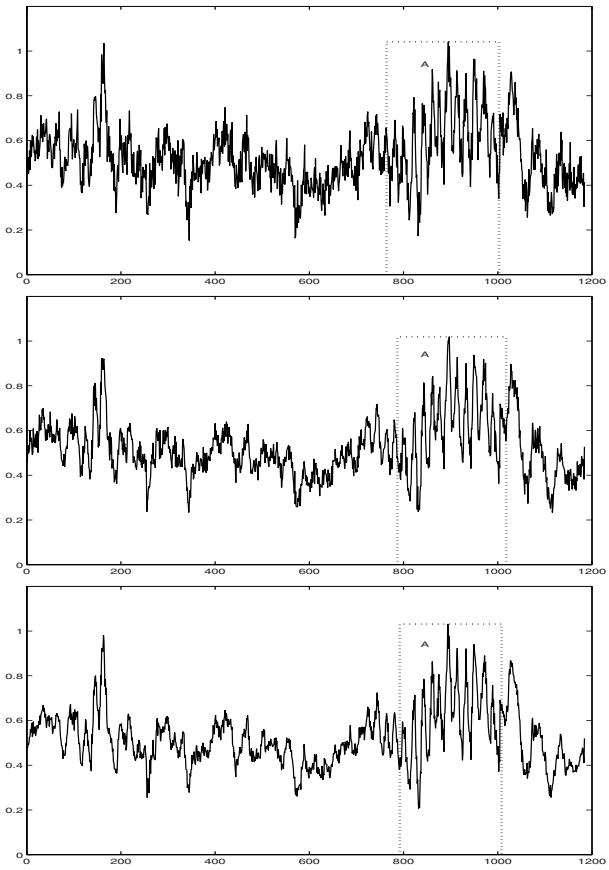


FIGURE 18. The detection results of Fig. 17 correspondingly from top to bottom by Algorithm 4.1

References

[1] Hongzhi An and Min Chen. *Non-linear Time series Analysis*. Shanghai Science & Technology Press, China, 1998.

[2] J. Doman, C. Detka, and T. Hoffman et al. Automating the sleep laboratory: Implementation and validation of digital recording and analysis. *International Journal of Biomedical Computing*, 38(3):277–290, 1995.

[3] Fei Huang and Chongxun Zheng. Automated recognition of spindles in sleep electroencephalogram base on time-frequency analysis. *Journal of Xi'An Jiao Tong University*, 36(2):218–220, 2002.

[4] N. E. Huang, Z. Shen, and S. R. Long et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, A(454):903–995, 1998.

[5] M. J. Korenberg. A robust orthogonal algorithm for system identification and time-series analysis. *Biological Cybernetics*, 60:267–276, 1989.

[6] Leukel. *Essential of Physiological Psychology*. The CV Company, USA, 1978.

[7] Jianping Liu, Shiyong Yang, and Chongxun Zheng. High resolution time-frequency analysis method for extracting the sleep spindles. *Journal of Biomedical Engineering*, 17(1):50–55, 2000.

[8] N. Pradhan and P. K. Sadasivan. The nature of dominant lyapunov exponent and attractor dimension curve of eeg in sleep. *Computers in Biology and Medicine*, 26(5):419–428, 1996.

[9] J. Pricipe, S. K. Gala, and T. G. Chang. Sleep staging automation base on the theory of evidence. *IEEE Transactions on Biomedical Engineering*, 36(5):503–509, 1987.

[10] Shannahoff-Khalsa David S., Gillin J. Christian, and etc. Ultradian rhythms of alternating cerebral hemispheric eeg dominance are coupled to rapid eye movement and non-rapid eye movement stage 4 sleep in humans. *Sleep Medicine*, 2(4):333–346, 2001.

[11] N. Schaltenbrand, R. Lengelle, and J. P. Macher. Neural network model: Application to automatic analysis of human sleep. *Computer and Biomedical Research*, 26(2):157–171, 1993.

[12] J. R. Smith, I. Karacan, and M. C. K. Yang. Automated analysis of the human sleep EEG. *Waking and Sleeping*, 2:229–237, 1978.

[13] E. Stanus, B. Lacroix, and M. Kerkhofs et al. Automated sleep coring: A comparative reliability study of two algorithms. *Electroencephalography and Clinical Neurophysiology*, 66(4):448–454, 1987.

[14] E. C. Titchmarsh. *Introduction to the Theory of Fourier Integrals*. Oxford University Press, 1948.

Zhihua Yang
Information Science School
GuangDong University of Business Studies,
Guangzhou 510320, P. R. China
e-mail: yangyangzh@tom.com

Lihua Yang
School of Mathematics and Computing Science
Sun Yat-sen University, Guangzhou 510275, P. R. China
e-mail: mcsylh@mail.sysu.edu.cn

Dongxu Qi
Faculty of Information Technology
Macao University of Science and Technology, Macao
School of Information Science and Technology
Sun Yat-sen University, Guangzhou 510275, P. R. China
e-mail: dxqi@must.edu.mo

Wavelet Analysis and Applications
Tao Qian, Mang I. Vai and Xu Yuesheng, Eds.
Applied and Numerical Harmonic Analysis, 561–574
© 2006 Birkhäuser Verlag Basel/Switzerland

A Wavelet-Domain Hidden Markov Tree Model with Localized Parameters for Image Denoising

Minghui Yang, Zhiyun Xiao and Silong Peng

Abstract. Wavelet-domain hidden Markov tree models have been popularly used in many fields. The hidden Markov Tree (HMT) model provides a natural framework for exploiting the statistics of the wavelet coefficients. However, the training process of the model parameters is computationally expensive. In this paper, we propose a HMT model with localized parameters which has a fast parameter estimation algorithm with no complex training process. Firstly, Wold decomposition is used to reduce the influence on the estimation of image noise variance due to texture. Secondly, coefficients in each subband are classified into two classes based on spatially adaptive thresholds. Thirdly, parameters of different class are estimated using the local statistics. Finally, the posterior state probability is estimated with an up-down step like the traditional HMT model. We apply this model to image denoising and compare it with other models for several test images to demonstrate its competitive performance.

Mathematics Subject Classification (2000). Primary 42C40; Secondary 94A08.
Keywords. Hidden Markov tree, wavelet transform, wold decomposition.

1. Introduction

In recent years, wavelet transform has become a popular tool for statistical signal processing, since it has many attractive properties, such as locality, multiresolution and compression. In many algorithms, wavelet coefficients have been modeled either as jointly Gaussian, or as non-Gaussian but independent. Jointly Gaussian models can capture the correlations between wavelet coefficients, but they are in conflict with the compression characteristic of wavelet transform. Non-Gaussian models assume the coefficients are statistical independent. However, the wavelet transform cannot completely decorrelate real-world signals, so that the assumption that the coefficients are independent is unrealistic.

This work was supported by National Natural Science Fund of China with No.60272042.

Sendur et al. [1] considered the dependencies between the wavelet coefficients and their parents. For this purpose, some non-Gaussian bivariate distributions were proposed and corresponding nonlinear shrinkage functions were derived from the models using Bayesian estimation theory. The performance can also be improved by estimating simple bivariate model in a local neighborhood [2].

Crouse et al. [3] proposed a hidden Markov tree as a statistical model for such properties of wavelet coefficients. This model first introduced the hidden state variables to match wavelet coefficients, then used Markovian dependencies between the hidden state variables to characterize the dependencies between the wavelet coefficients. For estimating the model parameters, an iterative Expectation Maximization (EM) approach was used which assumed that all wavelet coefficients and state variables within a common scale be identically distributed. Nevertheless, the training process was still computationally expensive.

Romberg et al. [4] introduced a simplified HMT model which specified the HMT parameters with just nine parameters (independent on the size of the image and the number of wavelet scales). A class of universal parameters was obtained by training a class of images, then these parameters were fixed and modeled a wide range of images. This universal HMT (uHMT) model can be very simple, because it requires no training. But the fixed parameters may be not exact for another kind of images.

Xiao et al. [5] tried to find a fast estimation technique for the HMT model parameters. They used adaptive thresholds to classify the wavelet coefficients into two classes, and estimated the model parameters by computing local statistics. This method gave better results with less run time.

We propose an image denoising algorithm for natural images. This new method not only introduces Wold decomposition into the estimation of the noise variance, but also improves Xiao's method.

The paper is organized as follows: Section 2 presents the estimation of the noise variance. In section 3, we details the improved HMT model with localized parameters. Section 4 shows the experimental results and compares with other algorithms. Finally, we conclude in section 5.

2. Estimation of the Noise Variance

2.1. Pervious Work

Image denoising is just to estimate the original signal f from a corrupted image g by removing noise n . Generally, the noise is assumed to be additive white Gaussian noise,

$$g = f + n \tag{2.1}$$

The compression property indicates that wavelet coefficients are energy compacted. Because of the orthogonal property of the wavelet transform, the noise in wavelet-domain is identical independent distributed with small magnitude. Thus,

small coefficients are more likely due to noise and large coefficients due to important signal features. Therefore, by setting small wavelet coefficients to zero, one effectively removes noise without degrading the signal.

Many wavelet thresholding methods have been discussed before. Donoho et al. [6] described a principle for spatially adaptive method which worked by shrinkage of the wavelet coefficients. They gave an asymptotic optimal threshold as the universal $\sigma_n\sqrt{2\log(N)}$, where σ_n was the standard deviation of the additive noise. These thresholds depended on the sample size N. As described in [7], deviding the threshold by 2 yields a soft threshold estimator that is close to the true value. Chang et al. [8] proposed a spatially adaptive wavelet thresholding method based on context modeling. Each wavelet coefficients was modeled as a random variable of a generalized Gaussian distrtibution with an unknown parameter estimated by its context. A good approximation to the optimal threshold which minimized the mean squared error of the soft-thresholding estimator was given as below:

$$T_B = \sigma_n^2/\sigma_x \tag{2.2}$$

where σ_n^2 was the additive noise variance and σ_x was the localized standard deviation of the signal.

The noise variance used in the threshold estimation methods is usually unknown. People often estimate it by using the robust median estimator in the highest subband $HH1$ of the wavelet transform

$$\hat{\sigma_n} = median(|Y[i,j]|)/0.6745 \tag{2.3}$$

where $Y[i,j] \in$ subband $HH1$.

2.2. Wold Decomposition Used in Estimation

To estimate the variance σ_n^2 of the noise n from the corrupted image g, we need to suppress the influence of the original image f . In the previous methods for noise variance estimation, the original figure is supposed to be a piecewise smooth signal. In the highest subband, the influence of f can be ignored and the coefficients are approximately Gaussian random variables of the noise variance σ_n^2 .

However, the assumption is not suitable for natural scene. Many real images are composed of both smooth and textured regions. So the influence of the signal f cannot be ignored in the highest subband of the wavelet transform. One can first segment the noisy image into two parts: texture region and piecewise smooth region. Then the noise variance can be estimated in the piecewise smooth region. But the segmentation is computationally expensive, and is always disturbed by the additive noise.

In the texture analysis, Francos et al. [9] had presented a unified texture model which was applicable to a wide variety of texture types found in natural images. On the basis of the two-dimensional Wold decomposition for homogeneous random fields, the texture field was decomposed into a sum of two mutually orthogonal components: a purely indeterministic component and a deterministic component.

Now we introduce the Wold decomposition into noisy natural images analysis. The noisy natural images can be decomposed into two components: a indeterministic component and a deterministic component. The indeterministic component includes the additive noise and the stochastic part in texture model. The deterministic component includes the piecewise smooth region and the deterministic part in texture model. Then the noise variance is estimated based on the indeterministic component.

The first step of the estimation algorithm is to compute the spectrum of the noisy image. The image is first zero-meaned, then the magnitude of the signal discrete Fourier transform is computed. Assumed the size of the image is N by N, the magnitude image is translated by $(N/2, N/2)$ so that the zero frequency is at the center of the frequency plane.

As the spectral measure associated with the deterministic field is singular, we need to locate the singularities in the magnitude image. The second step is to find the local maxima of the magnitude by searching a 5 by 5 neighborhood of each frequency sample.

Thirdly, the local maxima are examined for their harmonic peaks if their values are above 10% of the magnitude range. As the magnitudes of the additive noise may also act as local maxima, the threshold should be set larger than 5% which is often used in texture analysis in order to save the harmonic peaks.

Fourthly, starting from each harmonic peak, a region is grown outwards continuously until the value of the magnitude is lower than a small portion of this peak value (10% in our experiments). This region is regarded as the support region of the local peak in question. However, there still have many noise in the local peaks and their support regions. Two morphological transformations, namely closing and opening, can be used to construct a mask, which can wipe off the isolated points. The resulting support regions are obtained by selecting in the mask.

Fifthly, after the spectral support of the deterministic component is determined, the component frequencies are separated from the rest, which comprises the indeterministic component.

Finally, the noise variance is estimated from the indeterministic component.

Figure 1 shows an example of Wold decomposition using in a noisy image. The deterministic component and the indeterministic component after inverse discrete fourier transform are displayed as Figure 1 (g) and Figure 1 (h). The results of noise variance estimation compared with traditional estimation results are listed in table 1.

3. The Localized HMT Model

3.1. Wavelet-Domain HMT Model

The wavelet-domain HMT model was first proposed for statistical signal processing and analysis in [3]. The wavelet decomposition transforms the image into a multiscale representation with a quad-tree structure. The HMT model was developed by

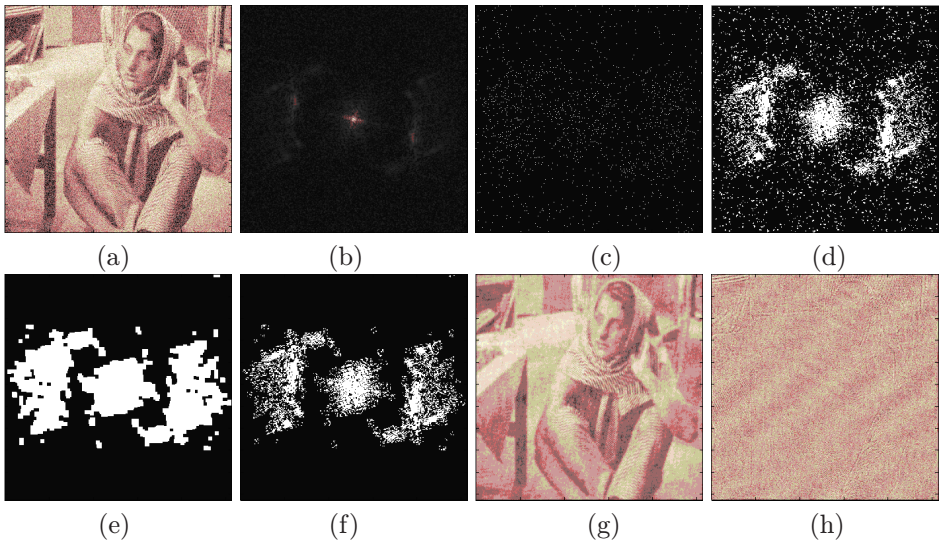


FIGURE 1. Example of Wold decomposition using in noisy image. (a) original image corrupted by noise ($\sigma_n = 0.1$), (b) the Fourier magnitude image, (c) the local maxima, (d) the support region of the local maxima, (e) the mask to wipe off the isolated points, (f) the support region in the mask, (g) the deterministic component, (h) the indeterministic component.

connecting hidden state variables vertically across scales to capture the statistical inter-scale dependencies between the wavelet coefficients of signals and images.

In the tree structure, we denote the node index as i . The scale covering the node i is denoted as j . The node i has a parent node denoted as $\rho(i)$ at the scale $j+1$, and four children nodes denoted as $c(i)$ at the scale $j-1$. Smaller j corresponds to higher resolution analysis. The set of descendants of a node i, including itself, is denoted as $T(i)$. The root node index is 1, then $T(1)$ denote the whole tree.

The wavelet coefficient and the hidden state variable of the node i are denoted as w_i and S_i respectively. In the HMT model, the wavelet coefficient w_i is conditionally independent of all other random variables given its state S_i .

To capture the non-Gaussian nature of the wavelet coefficients, we model the marginal probability $f(w_i)$ of each coefficient as a two-state, zero-mean Gaussian mixture model in this paper as

$$f_W(w_i) = \sum_{m=S,L} P_{S_i}(m) * f_{W|S}(w_i|S_i = m) \tag{3.1}$$

where $P_{S_i}(m) = p(S_i = m|w, \theta)$ denotes the conditional probability mass function (pmf) of the node i given the model parameters θ . $m = S, L$ means the state m of coefficient is small or large. The parameters for the HMT model include:

1. $P_{S_1}(m)$, the pmf for the root node S_1 .
2. $\varepsilon_{i,\rho(i)}^{m,n} = P_{S_i|S_{\rho(i)}}(m|S_{\rho(i)} = n)$, the state transition probability from the parent state variable $S_{\rho(i)}$ to its child's states S_i, where $S_i = m$ and $S_{\rho(i)} = n$.
3. $\mu_{i,m}$,$\sigma_{i,m}^2$, the mean and variance of the wavelet coefficient w_i given $S_i = m$. Generally $\mu_{i,m} = 0$.

These parameters can be grouped into a model parameter vector θ as:

$$\theta = \{P_{S_1}(m), \varepsilon_{i,\rho(i)}^{m,n}, \sigma_{i,m}^2\} \tag{3.2}$$

In image denoising, the estimation problem can be expressed in the wavelet domain as :

$$w_i = y_i + n_i \tag{3.3}$$

where w_i, y_i, and n_i denote the wavelet coefficients of the observed data, the signal and the noise, respectively.

If the HMT model $\theta = \{P_{S_1}(m), \varepsilon_{i,\rho(i)}^{m,n}, \gamma_{i,m}^2\}$ for the noise signal is estimated, $\gamma_{i,m}^2$ is the variance of the noisy wavelet coefficient w_i, then the conditional mean estimate of y_i given w_i is

$$\hat{y_i} = E[y_i|w_i, \theta] = \sum_{m=S,L} P_{S_i(m)} \frac{\sigma_{i,m}^2}{\sigma_{i,m}^2 + \sigma_n^2} w_i \tag{3.4}$$

where $\sigma_{i,m}^2 = max(0, \gamma_{i,m}^2 - \sigma_n^2)$. The final signal estimate is computed as the inverse wavelet transform of these estimates of the signal wavelet coefficients.

3.2. The Localized HMT Model

In the HMT model, the basic problem is parameter estimation. As the state of the wavelet coefficients is unobserved, direct maximum likelihood estimation is intractable. Therefore, the iterative EM approach was used in the parameters estimation.

For an n-pixel image, generally, the HMT model has approximately $4n$ parameters. These parameters need a lot of signal samples in the training step. Crouse et al. [3] used an robustly estimation by assuming that the model parameters are the same in each scale. This approach reduced the total number of the HMT model parameters from $4n$ to $4J$, where J presents the number of wavelet scales. Romberg et al. [4] reduced the number of parameters to nine, by exploiting the inherent self-similiarity of real-world images. Fan et al. [10] developed an initialization scheme for the EM algorithm to achieve more efficient HMT model training. These methods all assumed that the HMT model are identically distributed in each scale. However, the assumption are not reliable for natural images. For example, the pmf of wavelet coefficients' state in the piecewise smooth region should not equal the pmf in the texture region. Xiao et al. [5] proposed a fast parameter estimation

technique for the HMT model. The parameters in this model are localized. We improved this algorithm and incorporate it with the noise variance estimation.

3.2.1. Wavelet Coefficients Classification. As described in the previous section, Chang et al.[8] proposed a spatially adaptive wavelet thresholding method based on context modeling. These thresholds can be used in shrinkage of wavelet coefficients. Considering the compression property of wavelet transform, we can classify the wavelet coefficients into two states, large and small, by introducing the adaptive thresholds computed by equation 2.3. The estimation of noise variance has been discussed in section 2.

The needed localized standard deviation σ_x is computed as:

$$\sigma_x^2[i,j] = max(0, \frac{1}{< N[i,j] >} \sum_{[k,j]\in N[i,j]} w^2[k,j] - \sigma_n^2) \tag{3.5}$$

where $N[i,j]$ denotes the neighborhood of the location i in scale j, and $< N[i,j] >$ denotes the number of coefficients including in $N[i,j]$.

We use a binary mask $B[i,j]$ to denote the state of the wavelet coefficients, with '1' denotes 'large' state and '0' denotes 'small' state.

$$B[i,j] = \begin{cases} 0, & w[i,j] < T_B[i,j] \\ 1, & w[i,j] \geq T_B[i,j] \end{cases} \tag{3.6}$$

Because large values of wavelet coefficients tend to propagate across scales, $B[i,j]$ is modified by its parent state.

$$B[i,j] = B[i,j] * B[\rho(i), j+1] \tag{3.7}$$

3.2.2. Localized Parameters Estimation. When we have classified the wavelet coefficients into two states, we can estimate the localized parameters $\theta = \{P_{S_1}(m), \varepsilon_{i,\rho(i)}^{m,n}, \gamma_{i,m}^2\}$.

1. The pmf for the root rode S_1 in the coarsest scale J

$$P_{S_1}(m) = \frac{1}{< N[i,J] >} \sum_{[k,J]\in N[i,J]} B_m[k,J], \qquad m = 0,1 \tag{3.8}$$

   where $B_m[k,J]$ denotes the pixel whose state value is m in $B[k,J]$.

2. The state transition probability

$$\varepsilon_{i,\rho(i)}^{m,n} = \frac{\sum_{[k,j]\in N[i,j]} B_m[k,j] * B_n[\rho(k), j+1]}{\sum_{[k,j]\in N[i,j]} B_n[\rho(k), j+1]} \tag{3.9}$$

   We use the relation of two binary masks of two scales to simulate the true state transition probability.

3. The state variance of the noisy signal

$$\gamma_{i,m}^2 = \frac{\sum_{[k,j]\in N[i,j]} w^2[k,j] * B_m[k,j]}{\sum_{[k,j]\in N[i,j]} B_m[k,j]} \tag{3.10}$$

   The variance of each state is calculated in the state mask.

The estimation of state variance $\gamma^2_{i,m}$ are not consistent with Xiao's method, since the denominator is the sum of $B_m[k,j]$ instead of the whole number of $N[i,j]$.

3.2.3. Posterior State Probability Estimation. Given initial localized parameters θ, we can calculate the posterior state probability $P_{S_i}(m) = p(S_i = m|w,\theta)$ with an upward-downward step described in [3]. The state variance $\gamma^2_{i,m}$ can also be updated at the same time. Then the signal state variance $\sigma^2_{i,m}$ is computed given $\gamma^2_{i,m}$.

3.2.4. Posterior State Probability Estimation. Using equation 3.4, the signal wavelet coefficients are estimated, then the final signal estimate is computed by the inverse wavelet transform.

Because wavelet transform lack translation invariance, the denoised image exhibits Gibbs phenomena in the neighborhood of discontinuities. We use the Cycle-spinning method proposed by Coifman [11] to suppress such artifacts. Three images by shifting original noisy image in horizontal, vertical and diagonal directions respectively are denoised in the proposed method, then total four denoised images are averaged after registration. The average image is a desired image which suppresses the artifacts.

4. Experiments

We use three standard testing images, Lena, Baboon and Barbara, to demonstrate our method's effectiveness. The variances of the additive noise are 0.01, 0.04 and 0.0025. The results are compared with other denoising algorithms in [2] [3] [4] [5] [6].

The peak signal-to-noise ratio (PSNR), the estimated noise variance $\hat{\sigma_n}$ and time cost are presented in table 1 for each method. From the table, we can find that Crouse's HMT model gives good results, but with much time cost. Romberg's uHMT model gives good performance to Lena, but not very good for the other two images, because the universal parameters are not fit these two images. Donoho's soft thresholding method gives the fastest process but the worst result in these five method. Sendur's bivariate shrinkage model and Xiao's algorithm denoise the noisy images in good performances with low time costs. Our method, after including the Wold decomposition in the estimation of the noise variance, can gain the best performance in a middle time cost.

Figure 2, 3 and 4 show the comparison of the different methods on a local region. Our method outperforms all the other methods in both visual quality and PSNR performance.

5. Conclusions

In this paper, we have proposed a wavelet-domain HMT model with localized parameters and applied it to natural image denoising. The localized parameters are

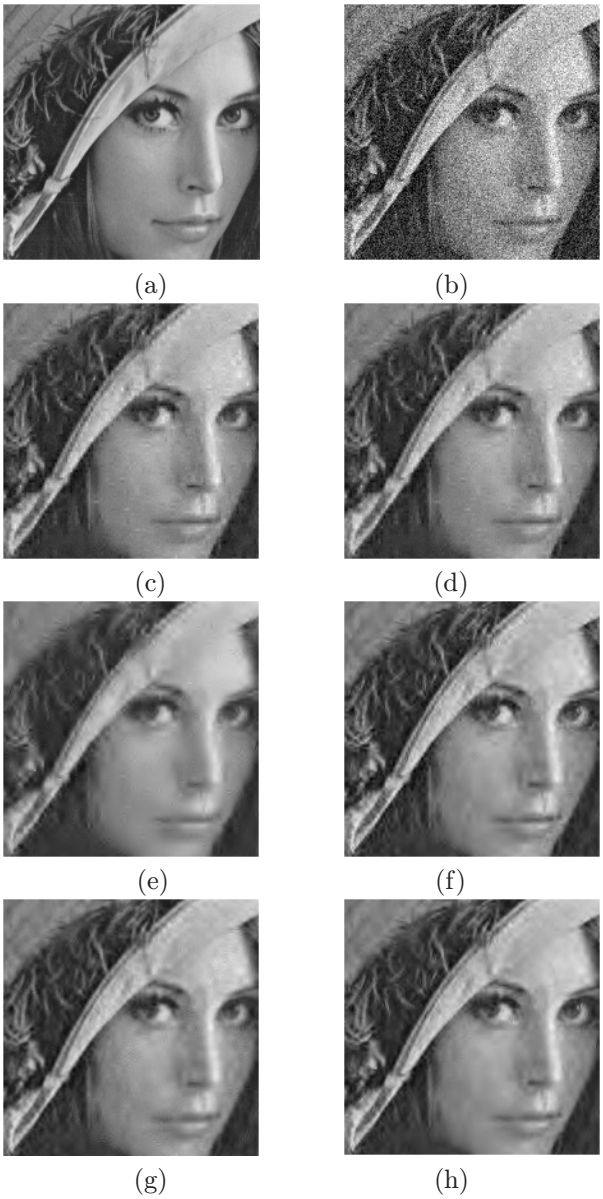


FIGURE 2. Comparing the results of various denoising methods, for Lena corrupted by noise $\sigma_n = 0.1$. (a) original image, (b) noisy image, (c) Crouse's HMT model, (d)Romberg's uHMT model, (e)Donoho's soft thresholding (f) Sendur's bivariate model (g) Xiao's fast estimation (h) our method.

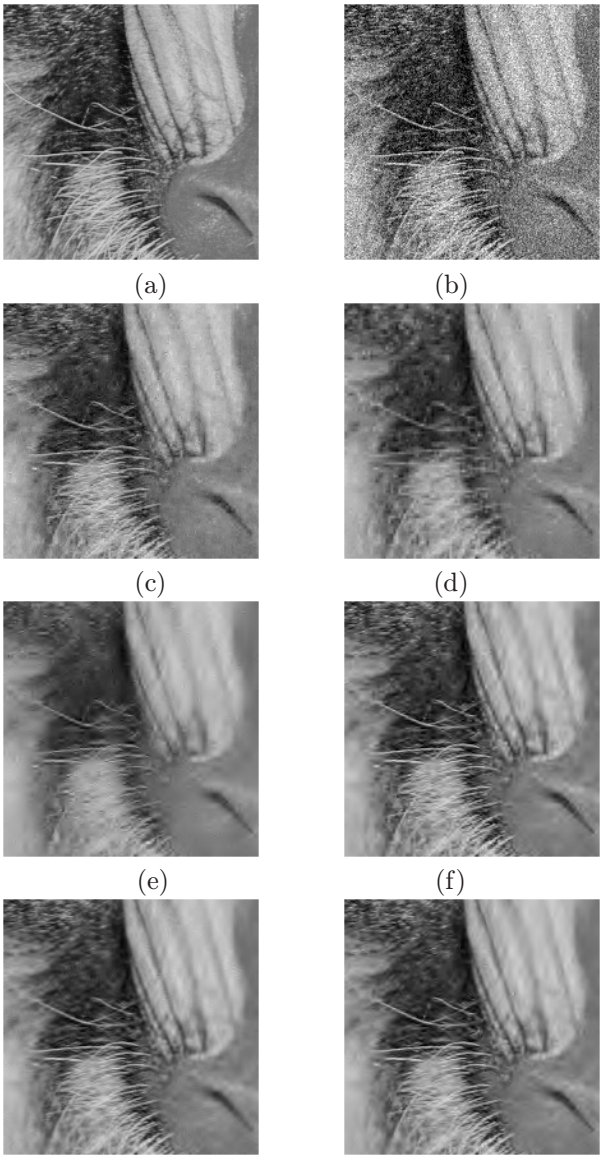


FIGURE 3. Comparing the results of various denoising methods, for Baboon corrupted by noise $\sigma_n = 0.1$. (a) original image, (b) noisy image, (c) Crouse's HMT model, (d)Romberg's uHMT model, (e)Donoho's soft thresholding (f) Sendur's bivariate model (g) Xiao's fast estimation (h) our method.

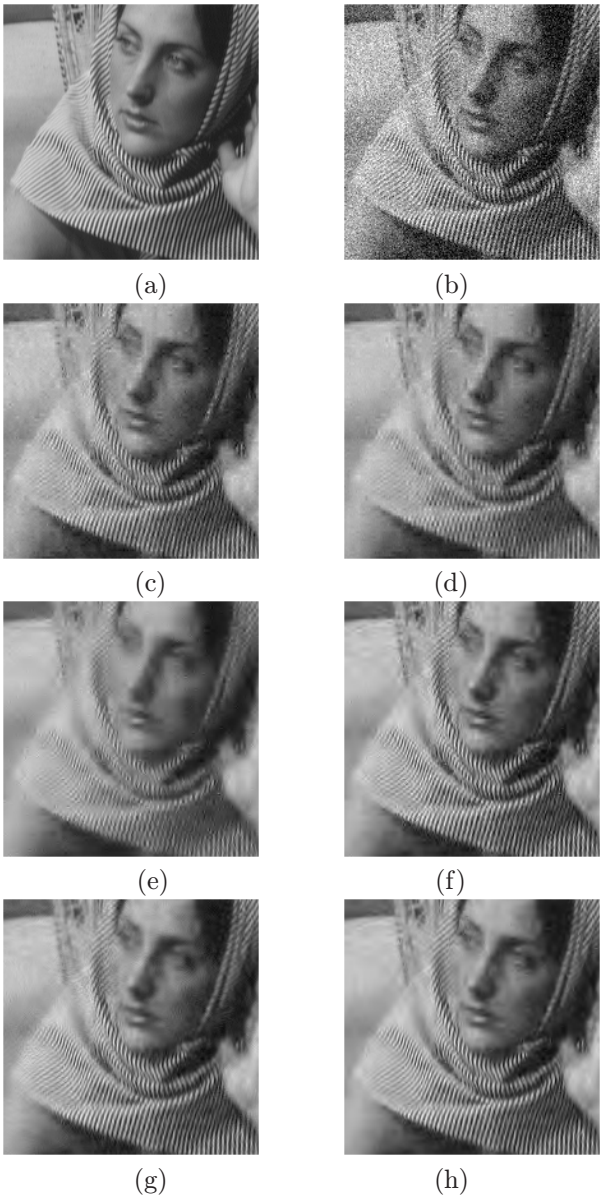


FIGURE 4. Comparing the results of various denoising methods, for barbara corrupted by noise $\sigma_n = 0.1$. (a) original image, (b) noisy image, (c) Crouse's HMT model, (d)Romberg's uHMT model, (e)Donoho's soft thresholding (f) Sendur's bivariate model (g) Xiao's fast estimation (h) our method.

TABLE 1. Comparison of different algorithms

Test images	Qualities	Crouse's HMT model	Romberg's uHMT model	Donoho's soft thresholding	Sendur's bivariate model	Xiao's fast estimation	Our method
Lena $\sigma_n = 0.1$ PSNR = 20.0205db	PSNR (db)	29.1688	29.2112	27.7566	29.6220	29.6101	30.2547
	$\hat{\sigma}_n$	0.1013	0.1013	0.1006	0.1006	0.1006	0.0995
	Time cost (s)	137.8	8.701	1.204	1.468	8.765	25.347
Lena $\sigma_n = 0.2$ PSNR = 13.9636db	PSNR (db)	25.9545	25.5650	25.2048	26.2453	25.9562	27.0709
	$\hat{\sigma}_n$	0.2022	0.2022	0.1993	0.1993	0.1993	0.1879
	Time cost (s)	106.9	9.077	1.377	1.313	8.893	24.578
Lena $\sigma_n = 0.05$ PSNR = 26.0313db	PSNR (db)	32.5526	28.4025	30.4443	32.9460	32.9087	33.2430
	$\hat{\sigma}_n$	0.0521	0.0521	0.0512	0.0512	0.0512	0.0511
	Time cost (s)	126.3	9.717	1.298	1.125	8.925	23.251
Baboon $\sigma_n = 0.1$ PSNR = 19.9942db	PSNR (db)	24.4567	22.2999	21.8771	24.0663	24.2922	24.5978
	$\hat{\sigma}_n$	0.1095	0.1095	0.1084	0.1084	0.1084	0.1035
	Time cost (s)	123.6	8.576	1.283	1.156	8.706	24.638
Baboon $\sigma_n = 0.2$ PSNR = 13.9996db	PSNR (db)	21.5765	20.4227	20.4180	21.2228	21.4799	21.8824
	$\hat{\sigma}_n$	0.2057	0.2057	0.2031	0.2031	0.2031	0.1840
	Time cost (s)	86.78	8.639	1.314	1.469	9.015	24.823
Baboon $\sigma_n = 0.05$ PSNR = 26.0461db	PSNR (db)	27.8668	27.3293	23.7595	27.4995	27.5742	27.8846
	$\hat{\sigma}_n$	0.0645	0.0645	0.0626	0.0626	0.0626	0.0625
	Time cost (s)	97.32	8.857	1.325	1.14	8.702	24.417
Barbara $\sigma_n = 0.1$ PSNR = 19.9875db	PSNR (db)	26.6408	25.2643	24.5679	27.1052	27.6340	27.9056
	$\hat{\sigma}_n$	0.1041	0.1041	0.1023	0.1023	0.1023	0.1012
	Time cost (s)	124.1	8.826	1.247	1.234	8.927	23.729
Barbara $\sigma_n = 0.2$ PSNR = 13.9940db	PSNR (db)	23.2911	22.6676	22.6780	23.7879	24.2670	24.5139
	$\hat{\sigma}_n$	0.2041	0.2041	0.2013	0.2013	0.2013	0.1918
	Time cost (s)	78.23	8.326	1.325	1.406	8.676	24.182
Barbara $\sigma_n = 0.05$ PSNR = 26.0127db	PSNR (db)	30.2944	28.1496	27.2639	30.9790	31.4332	31.8015
	$\hat{\sigma}_n$	0.0557	0.0557	0.0536	0.0536	0.0536	0.0526
	Time cost (s)	142.7	9.248	1.296	1.203	8.910	22.589

estimated by computing the statistics of the classified wavelet coefficients based on spatial adaptive thresholding method. Wold decomposition method is introduced in this method to estimate the noise variance. After separating the deterministic component of texture regions from noise, the estimation is more accurate. Experiments show that the proposed algorithm not only reduces the computational complexity, but also gives superior performance than other methods.

However, as shown in the figures, the result of the denoised image is less than satisfactory if there are too many texture regions in the image. This kind of images need further discussion in order to find some suitable algorithms for these images.

Further work will concentrate on extending this localized HMT model to other applications, such as segmentation, classification and compression.

References

[1] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," IEEE Transactions on Signal Processing, Volume 50, No. 11, pp. 2744-2756 November 2002.

[2] L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," IEEE Signal Processing Letters, Volume 9, No. 12, pp. 438-441 December 2002.

[3] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," IEEE Transactions on Signal Processing, Volume 46, pp. 886-902, April 1998.

[4] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree -structured image modeling using wavelet domain hidden markov models," IEEE Transactions on Image Processing, Volume 10, No.7, pp. 1056-1068, July 2001.

[5] Zhiyun Xiao, Wei Wen, and Silong Peng. "A fast classification-based parameter estimation technique for wavelet-domain HMT model," Advanced Concepts for Intelligent Vision Systems (ACIVS'04), August 2004, Brussels, Belgium.

[6] D. L. Donoho, I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," Biometrika, Volume 81, pp. 425-455, 1994.

[7] S. Mallat, "A wavelet tour of signal processing," second edition, Academic Press, 1998.

[8] S. G. Chang, B. Yu, and M. Vetterli, "Image denoising via lossy compression and wavelet thresholding," in Proc. IEEE Int. Conf. Image Processing, Volume 1, pp. 604-607, October 1997.

[9] J. M. Francos, A. Z. Meiri, and B. Porat, "A unified texture model based on a 2-D Wold-like decomposition," IEEE Transactions on Signal Processing, Volume 41, No.8, pp. 2665-2678, August 1993.

[10] G. L. Fan, X. G. Xia, "Improved hidden markov models in the wavelet-domain," IEEE Transactions on Signal Processing, Volume 49, Issue 1, pp. 115-120, January 2001.

[11] R. R. Coifman, D. L. Donoho, "Translation-invariant de-noising," in Wavelets and Statistics, A. Antoniadis and G. Oppenheim, Eds. New York. Lecture Notes in Statistics, Springer-Verlag, pp. 125-150, 1995.

Minghui Yang
National ASIC Design Engineering Center
Institute of Automation
Chinese Academy of Sciences
Beijing 100080
China
e-mail: minghui.yang@ia.ac.cn

Zhiyun Xiao
National ASIC Design Engineering Center
Institute of Automation
Chinese Academy of Sciences
Beijing 100080
China
e-mail: zhiyun.xiao@ia.ac.cn

Silong Peng
National ASIC Design Engineering Center
Institute of Automation
Chinese Academy of Sciences
Beijing 100080
China
e-mail: silong.peng@ia.ac.cn