This paper presents an adaptive sparse representation scheme for the remote sensing image, the geometric structure of which is more complex than that of natural image. The presented scheme includes two main stages which are wavelet transform and adaptive directional filter which is designed based on a binary tree. The construction of the binary tree depends on the image geometric information in frequency domain. Besides the properties of multiscale, multidirectionality and nonredundancy, our scheme has an additional property named adaptation which is a particular characteristic. Experimental results show that, in the sense of signal to noise ratio and visual quality, the performance of the nonlinear approximation of our scheme for remote sensing image is better than most of the existing sparse representation schemes.

Index Terms— adaptive sparse representation, remote sensing image, directional filter, binary tree, adaptation

1. INTRODUCTION

Generally, the efficiency of sparse representation of a multiscale transform depends on its ability of nonlinear approximation by keeping the most significant K transform coefficients. Wavelet has been demonstrated that it is not an efficient sparse representation scheme because of its isotropic refinement. Over the past few years, M. N. Do and his research team have proposed the contourlet theory which captures the directional information by employing directional filter banks (DFB), and achieves an efficient sparse representation for the natural image. The contourlet theory includes contourlet [1], nonsubsampled contourlet [2], etc. In [1], the authors proposed a double filter bank structure to obtain sparse representation for natural image having smooth contours. In [2], the authors developed a nonsubsampled contourlet transform, which consists of a nonsubsampled pyramid structure and a nonsubsampled DFB. More recently, S. Zhang et al. proposed a framework of filter banks named nonredundant contourlet transform [3]. R. Eslami et al. proposed a scheme named hybrid wavelets and DFB (HWD) [4] which employs a wavelet transform in conjunction with the modified version of DFB.

These schemes mentioned above unite the advantages of multiresolution and multidirectionality, however, they do not integrate the image geometric information, which means the number of directions at each level of wavelet-like decomposition is nonadaptive, a disadvantage for the sparse representation of image having complex geometric structure. Inspired by [5], in this paper we propose an adaptive sparse representation for remote sensing image based on the combination of the wavelet transform and adaptive directional filter. The proposed scheme includes two main stages, the first stage is to employ a two-dimensional wavelet transform for the multiscale decomposition, and the second stage is to apply an adaptive directional decomposition in each wavelet subband. The context of this paper is organized as follows. Section 2 first introduces the directional filter, and then describes an efficient method to construct a binary tree in frequency domain, which is used to instruct the adaptive directional decomposition. Section 3 studies the method of combination of wavelet transform and adaptive directional filter to achieve the adaptive sparse representation. In Section 4, we present the experimental results of our scheme, and compare them with some other schemes. Finally, we conclude this paper in Section 5.

2. ADAPTIVE DIRECTIONAL DECOMPOSITION

2.1. Directional Filter

In 2001, M. N. Do proposed a simple formulation for the DFB based on the fan filters [6]. The four-direction frequency partition and its two-level filter structure are shown in Fig. 1, where Q1 and Q2 denote the sampling

Fig. 1. (a) Four directional frequency partition subbands. (b) The corresponding two-level DFB structure based on fan filters.

In 2001, M. N. Do proposed a simple formulation for the DFB based on the fan filters [6]. The four-direction frequency partition and its two-level filter structure are shown in Fig. 1, where Q1 and Q2 denote the sampling
matrices expressed by (1). The McClellan transformation [7] is used to design the two-dimensional fan filters, and the quincunx filter banks [8] can be used to achieve a finer directional frequency partition.

\[
Q_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (1)
\]

### 2.2. Binary Tree Construction In Frequency Domain

Edges are the dominant component of the image geometric information, and are especially attended by researchers working on image representation and compression. In this subsection, we construct a binary tree based on the image geometric information with a simple method in the frequency domain. Then the tree is used to instruct the directional decomposition in the following subsection. The specification of the construction of the binary tree is described as following steps.

1) Image edge detection:
   Since our scheme is not sensitive to the quality of edge detection, any of the standard edge detection operators, such as sobel operator and prewitt operator, can be used to detect the edges and sharp transitions of the remote sensing image.

2) Fourier transform and thresholding:
   First the edge map produced from step 1 is manipulated by the Fast Fourier Transform to obtain the corresponding frequency image. Then the frequency image is processed with a threshold, which brings convenience for the construction of the binary tree. The operation of thresholding is expressed by (2).

\[
F_i(i, j) = \begin{cases} 1 & \text{if } |F(i, j)| \geq T_e \\ 0 & \text{if } |F(i, j)| < T_e \end{cases} \quad (2)
\]

Where \(F(i, j)\) and \(F_i(i, j)\) denote the original and processed frequency coefficients at location \((i, j)\), respectively.

\[T_e = \frac{1}{M \cdot N} \sum_{(i, j)} |F(i, j)|\]

is the threshold, \(M\) and \(N\) are the width and height of the image.

3) Binary tree construction:
   In our scheme, we construct four sub-basic binary trees, corresponding to the four directional frequency partition subbands illustrated in Fig. 1 (a). Here we focus on constructing a sub-basic binary tree corresponding to the directional subband ‘1’ since the constructions for the other three directional subbands are almost the same. First, we divide the frequency image obtained from step 2 into four sub-regions which correspond to the four directional subbands, and pick out the sub-region ‘1’. This sub-region corresponds to the root node of the sub-basic binary tree. Secondly, we figure out \(S_1\), the number of nonzero frequency coefficients calculated by (3), where \((u, v)\) denotes the coordinate belonging to the corresponding sub-region. Thirdly, compare \(S_1\) with a threshold \(T_d\). If \(S_1 > T_d\), we divide the sub-region ‘1’ into two homolographic sub-regions denoted by ‘11’ and ‘12’. The two sub-regions correspond to two son nodes of the tree. Then compare \(S_{11}\) and \(S_{12}\), calculated in the same way as \(S_1\), with \(T_d\) respectively, and repeat the division and comparison courses until the number of nonzero frequency coefficients in each sub-region is less than \(T_d\). From the above description, we can find that the depth of the sub-basic binary tree is a function of \(T_d\) which is an experience value in our scheme. Fig. 2 shows an example of the sub-region division and the constructed sub-basic binary tree.

\[
S_i = \sum_u \sum_v F_i(u, v), \quad (u, v) \in \{\text{sub-region '1'}\} \quad (3)
\]

[Fig. 2. (a) Sub-region division based on the value of frequency coefficients and \(T_d\). (b) The corresponding sub-basic binary tree. In this case, \(S_{11} > T_d\), \(S_{12} > T_d\), \(S_{111} > T_d\), \(S_{112} > T_d\), \(S_{1111} \leq T_d\), \(S_{1112} \leq T_d\), \(S_{11112} \leq T_d\) and \(S_{111112} \leq T_d\). (c) The updated binary tree, the depth of which is decreased by 1 through discarding the leaves on the deepest layer. This would be used in the Section 3.]

### 2.3. Adaptive Directional Filter

The directional filter together with the four sub-basic binary trees is employed to achieve the adaptive directional filter. First, we use the filter structure shown in Fig. 1 (b) to obtain the four basic directional subbands, ‘1’, ‘2’, ‘3’ and ‘4’, illustrated in Fig. 1 (a). Secondly, according to the description aforementioned, we construct four sub-basic binary trees corresponding to the four basic directional subbands. For convenience, we employ \(T_{r_1}\), \(T_{r_2}\), \(T_{r_3}\) and \(T_{r_4}\) to denote the sub-basic binary trees corresponding to sub-region ‘1’, ‘2’, ‘3’ and ‘4’, and let \(D_1\), \(D_2\), \(D_3\) and \(D_4\) represent the depths of \(T_{r_1}\), \(T_{r_2}\), \(T_{r_3}\) and \(T_{r_4}\), respectively. Lastly, we use the quincunx filter banks, specified in Chapter 3 [6], to decompose each basic directional subband based on the corresponding sub-basic binary tree. In detail, the quincunx filter banks are used to decompose the directional subband corresponding to each node, except the leaves, of each sub-basic binary tree from top to bottom. Fig. 3 illustrates the adaptive directional decomposition based on the sub-basic binary tree shown in Fig. 2 (b), where \(R_1\) and \(R_2\) denote the resampling operators expressed by (4). According to the above depiction, the conclusion is that we carry out the directional filter based on the image geometric information.
3. ADAPTIVE SPARSE REPRESENTATION

Normally, sparse representation requires a transform possesses the properties: multiscale, multidirectionality, anisotropy and nonredundancy which is desired in compression. Besides these properties, another feature named adaptation is added in our scheme. In this section we employ the wavelet together with the adaptive directional filter to achieve adaptive sparse representation for remote sensing image whose geometric structure is more complex than that of natural image.

It is known that the major characteristic of wavelet is the ability of catching edge points, and the major characteristic of DFB is the ability of extracting directional information. Our scheme combines the two approaches, and the specification is described as follows.

1) Construct the four sub-basic binary trees using the method specified in the above section, and select the maximum among $D_1, D_2, D_3$ and $D_4$ using (5).

$$D_{\text{max}} = \max\{D_1, D_2, D_3, D_4\}$$  \hspace{1cm} (5)

2) Process the remote sensing image with $D_{\text{max}}$-level wavelet decomposition to obtain the multiscale transform subbands, $\text{HH}_i, \text{HL}_i, \text{LH}_i$ and $\text{LL}_{i\text{max}}$ where $1 \leq i \leq D_{\text{max}}$.

3) Decompose each subband of wavelet transform with the adaptive directional filter. The specification is that, first decompose each subband, $\text{HH}_i, \text{HL}_i$ and $\text{LH}_i$ where $1 \leq i \leq D_{\text{max}}$, into four basic directional subbands using the DFB illustrated in Fig. 1 (b). Secondly, since the directional filter is sensitive to the high frequency component but low frequency component, we decompose the four basic directional subbands, ‘1’, ‘2’, ‘3’ and ‘4’, of $\text{HH}_1$ based on $\text{Tr}_1, \text{Tr}_2, \text{Tr}_3$ and $\text{Tr}_4$ respectively, decompose the two basic directional subbands, ‘1’ and ‘2’, of $\text{HL}_1$ based on $\text{Tr}_1$ and $\text{Tr}_2$ respectively, and decompose the two basic directional subbands, ‘3’ and ‘4’, of $\text{LH}_1$ based on $\text{Tr}_3$ and $\text{Tr}_4$ respectively, and update each basic binary tree through discarding the leaves on the deepest layer, illustrated in Fig. 2 (c). Then as the value of $i$ increases from 2 to $D_{\text{max}}$, we decompose $\text{HH}_i, \text{HL}_i$ and $\text{LH}_i$ in the same way as $\text{HH}_1, \text{HL}_1$ and $\text{LH}_1$ respectively.

It has been demonstrated that if the image $I(i, j)$ is uniformly regular, which is measured by the fact that it is $\alpha$ times continuously differentiable, then there must be exist a constant $C$ and an optimal multiscale transform scheme such that the approximate image $I_k(i, j)$ from the corresponding $K$ transform coefficients satisfies (6) [9], whereas the approximation of contourlet transform just satisfies (7). In [5] the authors presented a transform, called bandelet, based on the image geometric information, which could provide the optimal approximation. Our scheme has a similar idea with bandelet, however, the essential difference is that we exploit the geometric information through the Fast Fourier Transform which requires $O(N^2/\log N)$ operations for an image of $N^2$ pixels, while the bandelet extracts the geometric information by forming the geometric flow which needs $O(N^2/\log N^2)$ operations. The advanced analysis of the presented scheme will be attended in our further research.

$$\|I(i, j) - I_k(i, j)\|^2 \leq CK^{-\alpha}$$  \hspace{1cm} (6)

$$\|I(i, j) - I_k(i, j)\|^2 \leq CK^{-2} (\log K)^{\alpha}$$  \hspace{1cm} (7)

4. EXPERIMENTAL RESULTS

In our scheme, $T_d$ is an important variable which controls the depth of each sub-basic binary tree. Through many experiments we find that it always provides optimal sparse representation when $T_d$ equals to the 3.5 percent of the total number of the image pixels, which we set as an experience value.

In this section, we present the experimental results of our scheme, and compare them with the performances of several existing schemes. All the schemes mentioned in the experiments employ the “9/7” filters for both multiscale transform and directional filter. The remote sensing image, illustrated in Fig. 4, size of 512×512, depth of 16 bits, is used for test.

Fig. 4. Remote sensing image acquired with NASA Airborne Visible/Infrared Imaging Spectrometer.

Here we set $T_d=512\times512\times3.5\%=9175$, and compare our scheme with HWD type1, wavelet and contourlet. Fig. 5 sets out the experimental results of nonlinear approximation of each scheme, based on keeping different $K$, the number of most significant coefficients. Fig. 6 illustrates the visual results when $K=512\times512/16=16384$. It can be seen that, our
scheme provides the best recovering SNR and fine details, and reduces the artifacts significantly. The improvements of our scheme can be attributed to the fact that the adaptive directional filter, based on the constructed binary tree, has stronger ability of extracting directional information. Specifically, if the image structure is complex adequately, there would be plenty of nonzero processed frequency coefficients in the corresponding sub-region. Then we divide the sub-region into two since we deem that it is worth having a finer directional decomposition to achieve a more precise and much sparser representation.

It is validated by the experimental results that the ability of nonlinear approximation of our scheme outperforms most of the existing schemes. Efficient compression approaches based on the joint of adaptive sparse representation and spectral correlation [10] for hyperspectral image can be found in our recent work [11], in which we employ the inter-band predictor to avoid encoding the locations of significant transform coefficients.

6. ACKNOWLEDGMENT

We would like to thank the anonymous reviewers and the support of National Grand Fundamental Research 973 Program of China under Grant No.2010CB731904.

7. REFERENCES