The NSCT-HMT Model of Remote Sensing Image Based on Gaussian-Cauchy Mixture Distribution

XIANGHAI WANG¹, RUOXI SONG², CHUANMING SONG¹, AND JINGZHE TAO²

¹The College of Computer and Information Technology, Liaoning Normal University, Dalian, 116029, China
²The School of Urban and Environmental Sciences, Liaoning Normal University, Dalian, 116029, China

Corresponding author: X.Wang (e-mail: xhwang@lnnu.edu.cn), C.Song (e-mail: chmsong@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grants: 41671439 and 61402214, and the Innovation Team Support Program of Liaoning Higher Education Department Grant: LT2017013.

ABSTRACT The Nonsubsampled Contourlet Transform (NSCT) not only retains the characteristics of Contourlet transform, but also has the good characteristic of shift-invariance, which plays a significant role in denoising, fusion and segmentation of texture-rich images. The nonsubsampled contourlet transform (NSCT) not only retains the properties of Contourlet transform but also has the important property of shift-invariance, which plays a significant role in image processing, such as denoising, fusion and segmentation of texture-rich images. This paper proposes a Gaussian-Cauchy mixture distribution based Nonsubsampled Contourlet Transform hidden Markov tree model (GC-NSCT-HMT). The specific form of Gaussian-Cauchy mixture distribution is determined by the kurtosis of the NSCT coefficients in each subband. Firstly, we study the probability density distribution of the remote sensing image NSCT coefficients and then propose the Gaussian-Cauchy mixture distribution, which can adaptively adjust according to the statistical property of NSCT coefficients through a balance function. Experimental results show that the proposed mixture distribution can achieve a good imitative effect to the NSCT coefficients. Secondly, we study the marginal statistical property and the joint statistical property of the NSCT coefficients, the persistence and aggregation properties of them are also studied in depth. We find that the 'father' NSCT coefficient can transfers to its son coefficients through a tree structure. Thirdly, we combine the above conclusions with the hidden Markov tree model (HMT) and the GC-NSCT-HMT model is proposed. Finally, we apply our model to remote sensing image denoising, the subjective and objective experimental results demonstrate the feasibility of the proposed method.

INDEX TERMS Gaussian-Cauchy mixture distribution, NSCT, HMT, GC-NSCT-HMT, remote sensing image denoising

I. INTRODUCTION

Multiscale geometric analysis provides a sparse representation for images. After the Wavelet transform effectively represents the singular points of images, multiscale geometric transforms, such as the Ridgelet transform, Bandelet transform, Curvelet transform, Shearlet transform, Contourlet transform and Directionlet transform have been proposed to provide a better representation for the high dimensional singular features of images, which can better capture the multidirectional edges and textures of the image [1-3]. Among them, the Contourlet transform has attracted the attention of the research community for its ability to capture edges and contours[4]. The Contourlet transform can decompose an image into subbands by using the multiscale and multidirectional filter banks, thus the transform has multiscale, multidirection, and anisotropy properties. However, due to the downsampling and upsampling processes, the Contourlet transform is not shift-invariant. To solve this problem, the nonsubsampled contourlet transform (NSCT) was first proposed by Cunha et al. [5]. This transform not only retains the characteristics of the contourlet transform but also has the desirable property of shift-invariance, which plays a significant role in the denoising, fusion and segmentation of complex images.
How to efficiently express and correlate the coefficients of these multiscale transforms has become a hot issue. In recent years, researchers tend to use various distributions to model the multiscale transform coefficients, and at the same time excavate the properties of them. Argenti et al. [6] used GGD to describe the wavelet coefficients, and then developed a MAP estimation denoising method based on wavelet decomposition. Gao et al. [7] presented a new SAR image denoising method based on Directionlet transform, and the Directionlet coefficients are first modeled by Cauchy transform. To model the dual-tree wavelet coefficients, Hill et al. [8] proposed a bivariate Cauchy-Cauchy transform. To model the wavelet coefficient inter-dependencies, the W-HMT based signal processing becomes increasingly complicated. To solve this problem, the W-HMT models an extremely large number of wavelet coefficient interactions, the denoising performance demonstrate the feasibility of the proposed distribution model. Then, they extended their method to images of specific types, such as medical image and color image, at the same time propose the corresponding denoising scheme.[12-14].

However, these modeling approaches are mainly concentrate on the relationship between internal coefficients of each subband, and can not accurately describe the correlation between the coefficients of different scales. To tackle with this problem, Crouse et al. [15] studied the properties of wavelet coefficients, then proposed the Wavelet Hidden Markov Tree (W-HMT) model. The HMT model can capture the inter-scale dependencies through the hidden states while assuming independence within and across the three subbands. Basically, the HMT model establishes the link between the hidden state of each coefficient and that of their children. As the W-HMT models an extremely large number of wavelet coefficient interactions, the W-HMT based signal processing becomes increasingly complicated. To solve this problem, Crouse et al. [16] improved their previous method, and proposed a wavelet domain contextual hidden Markov model by modeling wavelet coefficient inter-dependencies via contexts; as a result, the model can substantially reduce the complexity. However, since the wavelet transform has a strong limitation in capturing directional information, researchers have recently considered devising HMT models in the contourlet or NSCT domains [17-23], which have a stronger capability for capturing directional information. Ducan et al. [17] established the Contourlet HMT model (C-HMT) on the basis of the W-HMT. The model depicts the correlation of Contourlet coefficients through Markov chains with hidden states and subsequently applied it to image segmentation and texture retrieval problems. Wu et al. [18-19]introduced the D-S theory to the C-HMT model to improve the segmentation accuracy of the SAR images. Yang et al. [20] combined the C-HMT with the PCNN (Pulse-Coupled Neural Network) theory, and proposed a remote sensing image fusion method. Wang et al. [21]established a novel Contourlet HMT model with directional features (C-HMT) , and the model was applied in image segmentation. Shahdoosti et al. [22] developed an image denoising scheme by combining the one-sided exponential distribution with the HMT model, and experimental results showed the feasibility of the proposed method. Wang et al. [23] studied the distribution of the NSCT coefficients and the correlation between each subband coefficients and proposed a novel NSCT HMT model, which can well capture the correlation of the NSCT coefficients in different directions and different scales.

While, The probability density function (PDF) used to describe the coefficients is a vital part of the above mentioned C-HMT and NSCT-HMT models, the accuracy of the statistical model directly affects the validity of the constructed HMT model. Most of these models use Gaussian distribution to approximate the subband coefficients, while it is not ideal when approximating NSCT coefficients of remote sensing images with rich texture. To solve the above mentioned problems, we propose a Gaussian-Cauchy mixture distribution based Nonsubsampled Contourlet Transform hidden Markov tree model. The novelty of the proposed approach includes: (1) A kurtosis based Gaussian-Cauchy mixture distribution is proposed to fit the NSCT coefficients of remote sensing images; (2) based on the proposed Gaussian-Cauchy mixture distribution, a novel NSCT-HMT model (GC-NSCT-HMT) is proposed; (3) the proposed GC-NSCT-HMT model has been applied in remote sensing image denoising and the subjective and objective experimental results show the feasibility of our proposed method.

II. THE ADAPTIVE GAUSSIAN-CAUCHY MIXTURE DISTRIBUTION MODEL

A. STATISTICAL ANALYSIS OF NSCT COEFFICIENTS

To study the marginal statistical properties of NSCT coefficients, the gray level histograms of the finest subband of three remote sensing images (marked as: Test-RS1, Test-RS2, Test-RS3) are depicted in Figure 1. The directions of NSCT are 2, 4, 8 from the coarse scale to the finer scale. Fig.1 shows that NSCT coefficients exhibit a sharp peak with an average value zero and two heavy tails on both sides of the peak. This implies that NSCT is a image sparse representation method, that is to say, only a small number of coefficients is large, so we can use only a small number of non-zero coefficients to represent the image. It is obvious that the kurtosis of the three shown distributions are much higher than 3 (the kurtosis of Gaussian distribution), so it is difficult to make a good approximation to the image if only using the Gaussian distribution [15]. We can get the similar conclusions in other subbands or other remote sensing images.
B. Comparison of NSCT coefficients fitted by classical probability distributions

In this paper, we choose the kurtosis of NSCT subband coefficients as the experimental variable, and use the K-S value as the objective index for comparing of some classical distributions with the prior PDF of NSCT coefficients. The K-S value is defined as:

$$ks = \max_{w \in R} |F_h(w) - F_e(w)|$$  \hspace{1cm} (1)

where $F_h(w)$ and $F_e(w)$ denote the cumulative distribution function of the prior probability density function and that of the empirical cumulative distribution function, respectively.

Here, we test the fitting effect of different distributions on a series of test remote sensing images collected from online open source set, and then divide the kurtosis value of the remote sensing images into several intervals to calculate the average K-S value of each interval. Each test image is decomposed by NSCT, for the NSCT we use 2, 4, 8 directions in the scales from coarser to finer, respectively. Then, the finest subband NSCT coefficients are fitted using five classical statistical models, the Maximum Likelihood (ML) estimation method is used to estimate the corresponding parameters.

Tab.1 exhibits the fitting performance of the above mentioned classical modeling distributions. From these statistic results we can draw the conclusion that, Gaussian distribution can achieve the best performance when the kurtosis of the NSCT subband coefficients is small (3<Kurtosis<7), while when the kurtosis of the NSCT subband coefficients is large (kurtosis>7) Cauchy distribution fits the best.

Further, Fig.2(a) shows the plots of a Gaussian PDF with different parameter values, and Fig.2(b) shows that of Cauchy PDF. From Tab. 1 and Fig. 2 we can see, Gaussian distribution exhibits the Gaussian behavior, Cauchy distribution exhibits the strong heavy tail behavior.
C. Construction of the kurtosis based adaptive mixture Gaussian - Cauchy distribution model

Based on the above conclusions, we present a novel kurtosis based adaptive Gaussian-Cauchy mixture distribution model. Its probability density function is defined as:

\[
f(x) = g(k,m) f_1(x) + (1 - g(k,m)) f_2(x)
\]

where \( f_1(x) \) and \( f_2(x) \) are the PDFs of Gaussian distribution and Cauchy distribution, respectively. And the specific expression of them are:

\[
f_1(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right)
\]

\[
f_2(x) = \frac{1}{\pi} \cdot \frac{b}{(x-a)^2 + b^2}
\]

where \( \sigma, \mu \) are the variance and mean of Gaussian distribution, respectively; and \( a, b \) are the location parameter and scale parameter of Cauchy distribution, respectively.

\( g(k,m) \) is the kurtosis based adaptive balance function. The function’s value is in the range of [0,1], and the function is specified in Eq. (4).

\[
g(k,m) = \frac{1}{m \cdot (k-N)}
\]

where \( k \) is the kurtosis of the NSCT subband coefficients; \( N \) is the standard reference kurtosis; \( m \) is the decline factor which is used to control the descent speed of the model. The corresponding curve is shown in Fig. 3. In our experiment we choose \( m = \log_2 k \).

D. Effectiveness of the Gaussian-Cauchy mixture distribution modeling NSCT coefficients

To illustrate the efficiency of the proposed Gaussian-Cauchy mixture distribution for fitting NSCT subband coefficients, we take the same experiments as Sect. II.B using the proposed Gaussian-Cauchy mixture distribution. Tab. 2 shows the corresponding KS metric. From Tab. 1 and Tab. 2, we can get the conclusion that the proposed Gaussian-Cauchy mixture distribution model can achieve the best fitting effect compared with other traditional probability density distributions.

Further, we respectively use Weibull distribution, Cauchy distribution, Rayleigh distribution, Exponential distribution, Gaussian distribution and the proposed distribution to fit the finest subband NSCT coefficients. The modeling performance of three test images (Test-RS1, Test-
RS2, Test-RS3) is shown in Fig.4, where the blue line is the empirical data, and the red line is the fitting curve. It is evident from this figure that the proposed distribution fits the empirical data best.

<table>
<thead>
<tr>
<th>Test image</th>
<th>Test-RS-1</th>
<th>Test-RS-2</th>
<th>Test-RS-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull distribution</td>
<td><img src="image1" alt="Weibull distribution" /></td>
<td><img src="image2" alt="Weibull distribution" /></td>
<td><img src="image3" alt="Weibull distribution" /></td>
</tr>
<tr>
<td>Cauchy distribution</td>
<td><img src="image4" alt="Cauchy distribution" /></td>
<td><img src="image5" alt="Cauchy distribution" /></td>
<td><img src="image6" alt="Cauchy distribution" /></td>
</tr>
<tr>
<td>Rayleigh distribution</td>
<td><img src="image7" alt="Rayleigh distribution" /></td>
<td><img src="image8" alt="Rayleigh distribution" /></td>
<td><img src="image9" alt="Rayleigh distribution" /></td>
</tr>
<tr>
<td>Exponential distribution</td>
<td><img src="image10" alt="Exponential distribution" /></td>
<td><img src="image11" alt="Exponential distribution" /></td>
<td><img src="image12" alt="Exponential distribution" /></td>
</tr>
<tr>
<td>Gaussian distribution</td>
<td><img src="image13" alt="Gaussian distribution" /></td>
<td><img src="image14" alt="Gaussian distribution" /></td>
<td><img src="image15" alt="Gaussian distribution" /></td>
</tr>
</tbody>
</table>

**Table 2. The Kolmogorov-Smirnov Performance of the Gaussian-Cauchy Distribution.**

<table>
<thead>
<tr>
<th>k</th>
<th>&lt;3.0</th>
<th>3.0-4.0</th>
<th>4.0-5.0</th>
<th>5.0-6.0</th>
<th>6.0-7.0</th>
<th>7.0-8.0</th>
<th>8.0-9.0</th>
<th>9.0-10.0</th>
<th>10.0-11.0</th>
<th>11.0-12.0</th>
<th>12.0-14.0</th>
<th>&gt;16.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>ks</td>
<td>0.0185</td>
<td>0.0197</td>
<td>0.0292</td>
<td>0.0335</td>
<td>0.0498</td>
<td>0.0325</td>
<td>0.0299</td>
<td>0.0346</td>
<td>0.0214</td>
<td>0.0287</td>
<td>0.0356</td>
<td>0.0197</td>
</tr>
</tbody>
</table>
III. THE GAUSSIAN-CAUCHY MIXTURE DISTRIBUTION BASED NSCT-HMT MODEL

NSCT was proposed on the basis of Contourlet Transform, which not only retains the good properties of Contourlet transform, but also has the property of shift-invariance. In NSCT, the NSLP is used to realize the multi-scale decomposition of an image, and the NSDFB is used to realize the multi-direction decomposition. Figure 5(a) plots the brief view of NSCT, Figure 5(b) shows the idealized frequency decompositions [5]. In this paper, we define three NSCT coefficient relationships: PX (parent, same location in the coarser scale), NX (neighbor, same subband) and CX (cousin, same spatial location different direction). The detailed positional relationship is depicted in Fig. 6. Where, PX is used to depict the inter-scale dependency of NSCT coefficients, NX and CX are used to depict the intra-scale dependency of NSCT coefficients.

A. The Joint Statistics of NSCT Subband Coefficients

To describe the relationship of NSCT coefficients in different subbands, we studied the joint statistical characteristics of NSCT coefficients. And Fig. 7 plots the joint probability distribution statistics of three test remote sensing images (Test-RS1, Test-RS2, Test-RS3).

According to Fig.7, we observe the following: 1) All points are tend the center position; 2) For the large coefficient points, the probability of a large neighborhood coefficient is large, and we can get the same conclusion to its son coefficients, which indicates that the NSCT coefficients has good aggregation and persistence properties.
Besides, we use the mutual information defined below to estimate the correlation among coefficients:

$$I(X;Y) = \int \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \, dx \, dy$$

(5)

Then we can find that, for the reference coefficient $X$, the mutual information within its generalized neighborhood relationships satisfies $I(X, PX) > I(X, NX) > I(X, CX)$.

<table>
<thead>
<tr>
<th>Test Image</th>
<th>$P(X \mid PX)$</th>
<th>$P(X \mid NX)$</th>
<th>$P(X \mid CX)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-RS-1</td>
<td><img src="image1" alt="Plot" /></td>
<td><img src="image2" alt="Plot" /></td>
<td><img src="image3" alt="Plot" /></td>
</tr>
<tr>
<td>Test-RS-2</td>
<td><img src="image4" alt="Plot" /></td>
<td><img src="image5" alt="Plot" /></td>
<td><img src="image6" alt="Plot" /></td>
</tr>
<tr>
<td>Test-RS-3</td>
<td><img src="image7" alt="Plot" /></td>
<td><img src="image8" alt="Plot" /></td>
<td><img src="image9" alt="Plot" /></td>
</tr>
</tbody>
</table>

(FIGURE 7. The joint probability distribution of the NSCT subband coefficients.)

Based on the above conclusions, we determine the parent-child relationship as the transfer relationship between the NSCT coefficients, that is, after the NSCT transformation from the coarse to the finer scale, it forms a binary tree structure as shown in Fig. 8.

(FIGURE 8. The relationship of the NSCT coefficients.)

B. The marginal statistics of NSCT subband coefficients

There exists large amount of coefficients with a small value, which usually stand for the flat regions, we define them as ‘small coefficients’ here; meanwhile, there also exists small amount of coefficients with a large value, which usually stand for the edge regions, we define them as ‘big coefficients’ [24]. And the corresponding function is defined as follows:

$$S(x) = I \left( \frac{|x|}{\sigma_w} \geq T \right)$$

(15)

where $\sigma_w$ is the noise standard deviation, $T$ is a given threshold, $T$ is set in the range from 2.5 to 3.5. (15) was promoted into the vector form by Bart et al. [25], the corresponding function is:

$$S(x) = I \left( \frac{\|x\|}{C_w^{1/2}} \geq T \right)$$

(16)

According to the previous discussion we can know that: the marginal distribution of NSCT subband coefficients exhibits non-Gaussian, large kurtosis, and heavy tail features. And the proposed Gaussian-Cauchy mixture distribution can fit it well.
In the following, the PDF of Gaussian-Cauchy distribution based HMT model is calculated by Goossens’ method: the conditional probability density function of \( f(S(x) = 1) \) and \( f(S(x) = 0) \):

\[
\begin{align*}
  f(S(x) = 1) &= N(0,(zC_a + C_u)) \\
  f(S(x) = 0) &= N(0,((zC_a)^{-1} + (T^2C_u)^{-1})^{-1} + C_u)
\end{align*}
\]

(17)

where \( C_a \) is the NSCT subband coefficients, \( N(\cdot,\cdot) \) is the Gaussian function, \( z \) is the hidden multiplier.

Performing integral operation on the hidden multiplier, we can get:

\[
\begin{align*}
  f(S(x) = 1) &= \int_0^{\infty} f(S(x) = 1) f_z(z)dz \\
  f(S(x) = 0) &= \int_0^{\infty} f(S(x) = 0) f_z(z)dz
\end{align*}
\]

(6)

where \( f_z(z) = \frac{1}{\Gamma(\tau)} z^{\tau-1} e^{-\frac{z^2}{\sigma_e^2}} \), \( \Gamma(\tau) = \int_0^{\infty} z^{\tau-1} e^{-\frac{z^2}{2\sigma_e^2}}dz \), \( P \) is a parameter, and in our experiment we set \( p = 5 \).

Thus the PDF of Gaussian-Cauchy distribution based HMT model is:

\[
  f(x) = P(S(x) = 0) f(S(x) = 0) + P(S(x) = 1) f(S(x) = 1)
\]

(7)

where

\[
\begin{align*}
  P(S(x) = 0) &= \int_0^{\infty} f_z(z) \left( \frac{\sigma_e^2}{z^2 + \sigma_e^2} \right)^{d/2} dz \\
  P(S(x) = 1) &= 1 - P(S(x) = 0)
\end{align*}
\]

(8)

and \( d \) is the number of the NSCT coefficients.

The specific derivation of the above formulas can be referred to the literature[17].

C. Construction of the Gaussian-Cauchy mixture distribution based NSCT HMT model

Based on the previous conclusions, this paper proposes a Gaussian-Cauchy distribution based NSCT HMT model (GC-NSCT-HMT). Here, we use the GSM presentation of the Gaussian-Cauchy mixture distribution as the conditional probability density distribution of the HMT model, and adopt the parent-child relationship as the transfer relationship between the NSCT coefficients. The specific construction process of the model is as following:

First, decompose the remote sensing images by NLSP, and the scale vector is \( j = \{1,2,\cdots,L\} \), where \( j = 1 \) denotes the scale with the lowest resolution, \( j = L \) stands for the scale with the highest resolution. At the same time decompose each scale by NSDFB with \( D_i \) directions, for each direction \( q \in \{1,2,\cdots,D_i\} \), it can form a tree structure with \( p \) nodes, and we define this tree structure as \( T_{ij}^q \), where subscript denotes the root node of the tree. Other symbols are defined as following:

\[
N^q = \{N_1^q,N_2^q,\cdots,N_p^q\}
\]

denotes the corresponding hidden state of the NSCT coefficients, and \( T_j^q \) denotes a tree structure whose direction is \( k \) and the root node is \( j \), the tree structure includes all the descendants of the root node; \( N_i \) stands for the NSCT coefficient of node \( i \); \( S_i \) stands for the hidden state of node \( i \); \( \rho(i) \) denotes the corresponding father coefficient of node \( i \).

Thus, for a NSCT of \( J \) scales, \( K \) directions, the novel NSCT HMT model contains the following parameters:

1) \( P_{N_i}(m) \) is the root state probability of the root node, \( N_i \) where \( m \) is the quantity of the hidden state and \( m = 2 \);

2) \( \epsilon_i = f(s_i = m | s_{\rho(i)} = n) \) is the state transition probability matrix, the specific meaning is when hidden state \( s_{\rho(i)} \) of the father node \( \rho(i) \) is \( n \), the probability of hidden state \( s_i \) of the son node \( i \) is \( m \).

The parameters of our proposed NSCT HMT model can be gathered in a vector:

\[
\Theta = \{P_{N_i}(m),\epsilon_i^{m,n} | i = 1,2,\cdots,p; m,n = 0,1\}
\]

(9)

D. Parameter estimation of the proposed GC-NSCT-HMT model

The EM algorithm is used to estimate the parameters of the proposed GC-NSCT-HMT model. The detailed iterative process of EM algorithm is introduced as follows.

Step 1: E step

Calculating the joint posterior probability density function \( P(S_i = m | N,\theta) \) and \( P(S_i = m, s_{\rho(i)} = n | N,\theta) \) of hidden states by using the upward-downward algorithm. Where \( N \) is the observed NSCT subband coefficient, \( \theta \) is the initial probability distribution, and \( l \) is the number of iterations.

To obtain these probabilities, we define \( T_i \) as the observed NSCT coefficient sub-tree with root node \( i \); \( T_i \) is the subtree of \( T_j \); \( T_{ij}^q \) is the tree formed by all the NSCT coefficients remaining after \( T_i \) is removed from \( T_j \). We then define the follows.

The conditional likelihood function:

\[
\beta_i(m) = f(T_i | S_i = m,\theta)
\]

\[
\beta_{\rho(i)}^{m,n}(m) = f(T_i | S_{\rho(i)} = m,\theta)\beta_{\rho(i)}^{m,n}(m) = f(T_i | S_{\rho(i)} = m,\theta)
\]

(10)

The joint probability density distribution:

\[
\alpha_i(m) = P(S_i = m, T_i | \theta)
\]

(11)

When \( S_i \) is given, \( T_i \) and \( T_{ij}^q \) are conditionally independent, according to the characteristic of HMM we can get:
$$
p(S_i = m, T_i | \theta) = \alpha_i(m) \cdot \beta_i(m)
$$

Then the likelihood value of the NSCT coefficient N is:

$$
f(N | \theta) = f(T_i | \theta) = \sum_{m=1}^{M} p(S_i = m, T_i | \theta) = \sum_{m=1}^{M} \alpha_i(m) \cdot \beta_i(m)
$$

Finally the conditional probability can be obtained by the Bayes' theorem:

$$
P(S_i = m | N, \theta) = \frac{\alpha_i(m) \cdot \beta_i(m)}{\sum_{n=1}^{M} \alpha_i(n) \cdot \beta_i(n)}
$$

$$
P(S_i = m, S_{\rho(i)} = n | N, \theta) = \frac{\beta_i(m) \cdot \alpha_{\rho(i)}(n) \beta_{\rho(i)}(n)}{\sum_{n=1}^{M} \alpha_i(n) \cdot \beta_i(n)}
$$

Downward step

1. Starting from the lowest resolution layer \( j = 1 \), for all hidden states let \( \alpha_i(m) = p_{\eta} \cdot m \).
2. Move to the next scale along the tree structure \( j = j + 1 \).
3. At the current scale \( j \), calculate:

$$
\alpha_i(m) = \sum_{n=1}^{M} e_{i,\rho(i)} \cdot \alpha_{\rho(i)}(n) \beta_{\rho(i)}(n)
$$

4. If \( j = L \), stop, else turn to (2).

Upward step

1. Starting from the highest resolution layer \( j = L \), for all hidden states let \( p_{\eta} = 1/M \) (M=2 in this paper). At the same time, we set each value in the state transition probability matrix \( p_{\eta} = 1/M \). Calculate the conditional probability of NSCT coefficients in each state \( \beta_i(m) \).
2. Move to the previous scale along the tree structure \( j = j - 1 \).
3. Calculate \( \beta_{\rho(i)}(m) = f \times \Pi \beta_{\rho(i)}(m) \)
4. If \( j = 1 \), stop, else turn to (2).

Step 2: M step

Update \( \theta^{t+1} \) to maximize the expected likelihood function:

$$
P_{\theta_t(m)} = \frac{1}{T} \sum_{t=1}^{T} P(S_i = m | N', \theta)
$$

$$
\epsilon_{i,\rho(i)} = \frac{1}{T} \sum_{t=1}^{T} P(S_i = m, S_{\rho(i)} = n | N', \theta)
$$

Step 3: Repeat E step and M step until convergence.

IV. THE APPLICATION OF GC-NSCT-HMT MODEL IN REMOTE SENSING IMAGE DENOISING

The noise component of the remote sensing images usually affects the characteristics of the reflection spectrum, greatly reduces the accuracy of data, thus, remote sensing image denoising has attracted many researchers’ attention and has become an important issue in remote sensing image processing area. The denoising method based on HMT model can well preserve edges in denoised images. Here, we apply the proposed GC-NSCT-HMT model in remote sensing image denoising, experimental results show the feasibility of our method.

A. The denoising algorithm

First the NSCT coefficients of the noisy image is modeled by the GC-NSCT-HMT model, thus, we can get the corresponding parameter vector: \( \Theta = \{ P_{\eta}, \epsilon_{i,\rho(i)} \} \). Then estimate the noise variance \( (\sigma_{(\eta,k,i)})^2 \) by using the Monte Carlo method, where \( j, k, i \) stand for the scale, direction and the coefficient, respectively. Since the variance of the original image \( (\sigma_{(x,j,k,i)})^2 \) can be expressed as the variance of the noisy image minus the noise coefficient variance:

$$
(\sigma_{(x,j,k,i)})^2 = (\sigma_{(y,j,k,i)})^2 - (\sigma_{(\epsilon,\rho(i,j,k,i,m)})^2)
$$

Using the Bayes estimation we can get:

<table>
<thead>
<tr>
<th>State=1</th>
<th>State=2</th>
<th>State=1</th>
<th>State=2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr,m</td>
<td>0.43</td>
<td>0.57</td>
<td>0.52</td>
</tr>
<tr>
<td>A1.2k-1</td>
<td>State1</td>
<td>0.97</td>
<td>0.06</td>
</tr>
<tr>
<td>A1.2k</td>
<td>State2</td>
<td>0.03</td>
<td>0.93</td>
</tr>
<tr>
<td>A2.2k</td>
<td>State1</td>
<td>0.96</td>
<td>0.06</td>
</tr>
<tr>
<td>A2.2k</td>
<td>State2</td>
<td>0.04</td>
<td>0.94</td>
</tr>
<tr>
<td>A3.2k-1</td>
<td>State1</td>
<td>0.99</td>
<td>0.24</td>
</tr>
<tr>
<td>A3.2k</td>
<td>State2</td>
<td>0.01</td>
<td>0.76</td>
</tr>
<tr>
<td>A3.2k-1</td>
<td>State1</td>
<td>0.01</td>
<td>0.77</td>
</tr>
<tr>
<td>A3.2k</td>
<td>State2</td>
<td>0.01</td>
<td>0.75</td>
</tr>
</tbody>
</table>

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2876447, IEEE Access
Last, using the conditional probability obtained by the EM algorithm \( p(S_{j,k,i} = m \mid y_{j,k,i}, \Theta) \), we can get the estimation of the clean image \( x \):

\[
E[x_{j,k,i} \mid y_{j,k,i}, \Theta, S_{j,k,i}] = \frac{(\sigma_{(j,k,i),m}^{(x)})^2}{(\sigma_{(j,k,i),m}^{(x)})^2 + (\sigma_{(j,k,i),m}^{(\text{noise})})^2} \times y_{j,k,i}
\]

(16)

Here, we summarize the important steps of the proposed GC-NSCT-HMT model:

**Step 1.** Decompose the noisy images using NSCT;

**Step 2.** Estimate the Gaussian-Cauchy mixture PDF parameters;

**Step 3.** Construct the GC-NSCT-HMT model, and estimate the parameters using the EM algorithm;

**Step 4.** Using the denoising algorithm to obtain the clean NSCT coefficients;

**Step 5.** Reconstruct of the clarity image.

### B. Simulation results

To verify the effectiveness of the proposed method, we perform a number of simulation experiments, where

\[
\text{'maxflat'} \text{ is chosen as the nonsubsampled multiscale filter of NSCT, and 'dmaxflat' is chosen as the nonsubsampled multidirection filter, also the directions of NSCT are 2, 4, 8 from the coarse scale to the finer scale.}
\]

The test images are Test-RS-1, Test-RS-2 and Test-RS-3 (the same with sec.II.A) with size 512x512. In our experiments the test images are corrupted by simulated additive noise with a standard deviation equal to 15, 20, 30, 40, 50, respectively. Also PSNR is used as the objective evaluation index of denoising effect. Besides, we compare our model with BLS-GSM [26], LLT [27], NL-means [28], BM3D schemes [29], the NSCT-HMT model [15] and the and the C-NSST-HMT model[30]. Tab.4 shows the PSNR results for different methods, compared with the state-of-art methods, the proposed denoising method can achieve better performance in terms of PSNR.

Fig.9 shows the denoising results using different methods when the variance of the additive noise is 30; Fig.10 is the denoising results of a local region of Test-RS-3 when enlarging 3 times. It can be observed from the results that for remote sensing images with rich texture information and detailed information, the proposed denoising method can not only effectively remove the noise, but also can well preserve the edge and texture features.

#### TABLE 4. THE PSNR OF THE PROPOSED MODEL AND SOME STATE-OF-ART MODELS.

<table>
<thead>
<tr>
<th>Test images</th>
<th>Noise variance</th>
<th>PSNR(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BLS-GSM</td>
</tr>
<tr>
<td>Test-RS-1</td>
<td>15</td>
<td>28.72</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>27.35</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>25.63</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>24.53</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>23.74</td>
</tr>
<tr>
<td>Test-RS-2</td>
<td>15</td>
<td>26.21</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>25.59</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>23.45</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>22.17</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>21.30</td>
</tr>
<tr>
<td>Test-RS-3</td>
<td>15</td>
<td>27.44</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>24.89</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>23.86</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>23.14</td>
</tr>
</tbody>
</table>

**TABLE 4.** The PSNR of the proposed model and some state-of-art models.

Test images | Noise variance | PSNR(dB)          |
-------------|----------------|-------------------|
| Test-RS-1   | 15              | 28.72  | 26.11 | 25.92    | 27.21 | 28.01    | 25.79      | 28.42    |
|             | 20              | 27.35  | 25.78 | 25.85    | 26.48 | 27.24    | 25.12      | 27.69    |
|             | 30              | 25.63  | 24.88 | 23.70    | 25.76 | 26.12    | 24.33      | 26.31    |
|             | 40              | 24.53  | 23.88 | 18.26    | 24.64 | 24.34    | 23.17      | 24.64    |
|             | 50              | 23.74  | 23.85 | 14.89    | 24.44 | 23.64    | 22.98      | 23.81    |

Test images | Noise variance | PSNR(dB)          |
-------------|----------------|-------------------|
| Test-RS-2   | 15              | 26.21  | 22.35 | 25.43    | 25.01 | 26.11    | 25.14      | 26.64    |
|             | 20              | 25.59  | 22.25 | 24.87    | 24.77 | 25.47    | 24.73      | 25.74    |
|             | 30              | 23.45  | 21.90 | 21.41    | 23.84 | 24.18    | 23.41      | 24.63    |
|             | 40              | 22.17  | 21.34 | 17.28    | 22.37 | 22.87    | 22.97      | 23.58    |
|             | 50              | 21.30  | 20.71 | 14.77    | 21.57 | 22.31    | 21.85      | 22.31    |

Test images | Noise variance | PSNR(dB)          |
-------------|----------------|-------------------|
| Test-RS-3   | 15              | 27.44  | 25.18 | 25.48    | 25.92 | 26.97    | 25.13      | 27.89    |
|             | 30              | 24.89  | 24.16 | 23.09    | 25.11 | 24.98    | 23.54      | 25.52    |
|             | 40              | 23.86  | 23.31 | 17.93    | 24.07 | 24.12    | 22.11      | 24.46    |
|             | 50              | 23.14  | 22.68 | 14.77    | 23.37 | 23.33    | 21.37      | 23.51    |
The denoised images

<table>
<thead>
<tr>
<th>Method</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Image 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLS-GSM</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>LLT</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>NL-Means</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>BM-3D</td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>NSCT-HMT</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>C-NSST-HMT</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
</tr>
</tbody>
</table>
The SSIM performance when the variance of the additive noise is 30 is shown in Tab 5.

According to Figs. 9, 10 and Tabs. 3, 4, we observe the following:

- The coefficients of remote sensing image features are highly correlated. At the same time, the contourlet transform is particularly suitable to analyze images with many edges and contours features. The proposed mixture Gaussian-Cauchy distribution model can effectively model the NSCT coefficients, which lays a good foundation for the subsequent operations in image processing.

- For remote sensing images with rich texture and detail information, the proposed denoising algorithm can not only effectively remove the noise but can also well preserve the features of the edges and texture regions of the original image.
V. CONCLUSION
In this paper, we first analyze the probability density distribution of the NSCT coefficients and propose the adaptive Gaussian-Cauchy mixture distribution based on the kurtosis of each subband. In addition, we analyze the marginal statistical property and the joint statistical property of the NSCT coefficients, and draw the conclusion that the NSCT coefficients have the property of persistence and aggregation. Then, we combine these conclusions with the hidden Markov tree model and propose the GC-NSCT-HMT model, at the same time give the parameter structure of the model. Last, we propose a denoising frame based on the GC-NSCT-HMT model. The simulation results verify the effectiveness of the model, and shows a certain advantage of remote sensing images with rich textures.

REFERENCES


XIANGHAI WANG was born in 1965. He received the Ph.D. degree from JiLin University, JiLin, in 1999. Currently, he is a professor of Computer and Information Technology College at Liaoning Normal University, Dalian. He is a senior member of the China Computer Federation. His research interests include computer image processing, multimedia information processing, and remote sensing image processing.

SONG RUOXI was born in 1994. She received the B.S. degree from Mathematics Science College of Inner Mongolia Normal University. She received the M.S. degree at the Computer College of Information and Technology from the Liaoning Normal University, Dalian. She is studying for a Ph.D. degree at the School of Urban and Environmental Sciences, Liaoning Normal University, Dalian. Her research interests include Remote Sensing Image Processing and mathematical modeling.

CHUAMING SONG was born in 1980. He received the Ph.D. degree from Computer Science and Technology College of Nanjing University. He is an associate professor of Computer and Information Technology College, Liaoning Normal University, Dalian. He is a member of China Computer Federation (CCF). His research interests include image, video coding, and digital watermarking of multimedia.

JINGZHE TAO was born in 1987. He is studying for a Ph.D. degree from the College of Urban and Environmental Science, Dalian. His research interest includes hyperspectral image fusion and super-resolution.