A Hidden Markov Tree Model for the Complex Wavelet Transform

Justin K. Romberg, Hyeokho Choi, Richard G. Baraniuk *

Department of Electrical and Computer Engineering, Rice University

6100 S. Main St., Houston, TX 77005, USA

Nick Kingsbury †

Department of Engineering, University of Cambridge

Trumpington St., Cambridge, CB2 1PZ, UK

Submitted to IEEE Transactions on Signal Processing, September 2002

EDICS Numbers: 2-HIER, 2-REST

Abstract

Multiresolution models such as the *hidden Markov tree* (HMT) aim to capture the statistical structure of signals and images by leveraging two key wavelet transform properties: wavelet coefficients representing smooth/singular regions in a signal have small/large magnitude, and small/large magnitudes persist through scale. Unfortunately, the HMT based on the conventional (fully decimated) wavelet transform suffers from shift-variance, making it less accurate and realistic. In this paper, we extend the HMT modeling framework to the *complex wavelet transform*, which features near shift-invariance and improved directionality compared to the standard wavelet transform. The complex HMT model is computationally efficient (with linear-time computation and processing algorithms) and applicable to general Bayesian inference problems as a prior density for images. We demonstrate the effectiveness of the model with two applications. In a simple estimation experiment, the complex wavelet HMT model outperforms a number of high-performance denoising algorithms, including redundant wavelet thresholding (cycle spinning) and the redundant HMT. A multiscale maximum likelihood texture classification algorithm produces fewer errors with the new model than with a standard HMT.

*Contact author: Justin Romberg. Email: jrom@rice.edu, choi@ece.rice.edu, richb@rice.edu. Web: www.dsp.rice.edu. Fax: 713.348.6196. This work was supported by the National Science Foundation grants CCR–9973188, ONR grant N00014-02-1-0353, AFOSR grant F49620-01-1-0378, and the Texas Instruments Leadership University Program.

†Email: ngk@eng.cam.ac.uk. Supported by the Rice-Trinity Exchange Program and the UK Engineering and Physical Sciences Research Council (EPSRC).

1 Introduction

The compact representation of the discrete wavelet transform (DWT) has led to many successful signal and image processing algorithms, including wavelet-based estimation (denoising), detection and classification, segmentation, fusion, and synthesis. Wavelets can be interpreted as multiscale singularity detectors, allowing us to describe the wavelet coefficients for a large class of signals and images in terms of two key properties [1, 2]:

2Populations: Smooth signal/image regions are represented by small wavelet coefficients, while edges, ridges, and other singularities are represented by large coefficients.

Persistence: Large and small wavelet coefficient values cascade along the branches of the wavelet tree.[1]

Since many real-world signals (photograph-like images, for example) consist mostly of smooth regions separated by a few singularities, the 2Populations property tells us that the wavelet coefficients are a *sparse* representation for these signals (this notion of sparsity can be made mathematically precise; see [4, 5] for example). Most of the wavelet coefficient magnitudes (representing the smooth regions) are small, while a few (representing the singularities) are large. The Persistence property codifies how these large values are structured: smooth regions/singularities are represented by small/large values at every scale.

Interpreting these properties statistically, 2Populations implies that wavelet coefficients have non-Gaussian marginal statistics; that is, their marginal probability density functions have a large peak at zero (many small coefficients) and heavy tails (a few large coefficients). Persistence leads to wavelet coefficient values that are statistically dependent along the branches of the wavelet tree. In this paper, we will use the 2Populations and Persistence properties to design a statistical model for signals and images. This model defines a *prior* for the joint probability distribution of the wavelet coefficients, with wavelet coefficients that exhibit the 2Populations and Persistence probabilities being more likely than those that do not. The model is based on two recent developments in multiscale signal processing: the wavelet-domain *hidden Markov tree* (HMT) model [1] and the *dual-tree complex wavelet transform* [6].

[1]Due to the multiresolution nature of the wavelet transform, the wavelet coefficients can be naturally arranged in a binary tree (1D) or a quadtree (2D), as shown in Figure 3. We discuss this briefly in Section 2; see [3] for more details.

Image processing algorithms inspired by the 2Populations and Persistence properties have proven simple and effective. For example, denoising by thresholding [7], which is in many ways the canonical wavelet-domain processing algorithm, exploits the sparseness of the wavelet representation: since a close approximation of the image can be built out of a few large coefficients, the remaining small magnitude coefficients contain mostly noise and can safely be discarded. The revolutionary zero-tree wavelet compression algorithm [8] leverages the Persistence property by coding entire subtrees of small wavelet coefficients with just a few bits. The wavelet-domain HMT model [1] explicitly incorporates both of these properties and has proved very effective in denoising [2] and segmentation [9] applications.

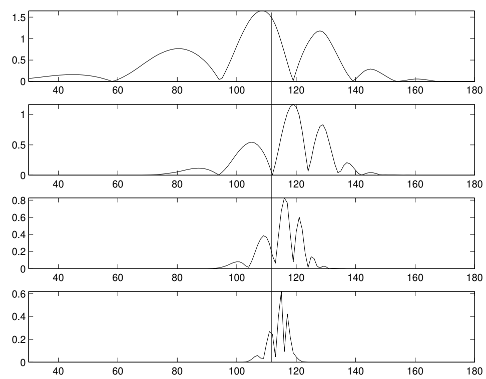
However, current wavelet-domain image processing algorithms have a tendency to produce images with mild ringing artifacts around the edges. At the heart of the problem is the fact that the orthogonal wavelet transform is not *shift-invariant* — singularities at different locations in the image are represented differently with wavelets.

The negative impact of shift-variance is best illustrated by investigating the step response of a particular wavelet basis function. Figure 1(a) plots the different values the magnitude of a real orthogonal DWT coefficient takes for different shifts of a one-dimensional (1D) step edge. As the location of the step changes, the magnitude of the wavelet coefficient shown representing this edge oscillates — the wavelet basis function oscillates, so its step response also oscillates. As a result, the wavelet coefficients representing the edge are not always "large", and the branches down the wavelet tree that describe discontinuities are often not the unbroken progressions of large coefficients that algorithms based on the Persistence property expect.

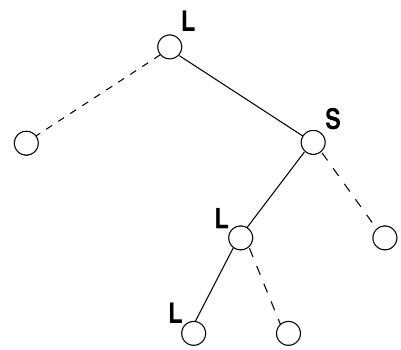
One way to counteract the effects of shift-variance is to use the undecimated wavelet transform [2, 10]. Denoising results using this method are typically better than the standard, orthogonal DWT, but the performance gain comes at a price: the $O(N \log N)$ undecimated algorithms are more expensive computationally than the $O(N)$ algorithms based on the orthogonal DWT.[2] Also, the redundancy of the undecimated wavelet transform ($N \log N$ coefficients compared to N for the orthogonal DWT) makes it inappropriate for certain applications like image compression.

In this paper, we address the problem of shift-variance by using the dual tree complex wavelet transform (CWT) [6, 11]. The CWT expands a signal or image in terms of a complex wavelet with complementary real and imaginary parts. In Section 3, we will show that when the real and

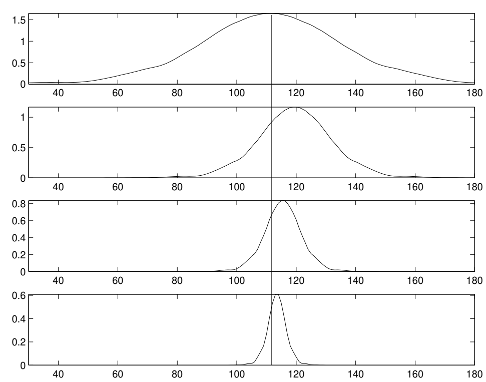
[2] Throughout this paper, we will use N to denote the number of data points in the signal in question; for an image, N is the number of pixels.



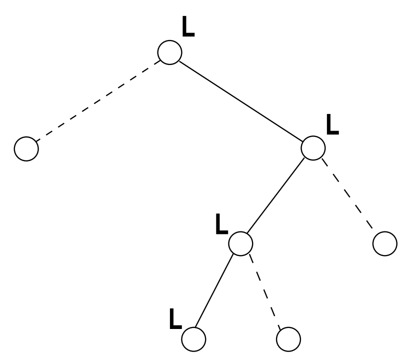
(a) D8 magnitude step response



(b) D8 state sequence



(c) complex magnitude step response



(d) complex state sequence

Figure 1: *(a) Magnitude of the 1D step response of Daubechies-8 (D8) wavelet coefficients at four consecutive scales. The horizontal axis represents the location of the step edge and the vertical axis represents the magnitude of the corresponding wavelet coefficient. (b) State sequence of the D8 coefficients for the edge location corresponding to the vertical line. (c) Step response of the complex wavelet coefficients with the same locales. (d) State sequence for complex coefficients. Note that the* L-S-L-L *state sequence for the D8 coefficients is a "broken" chain of large states, while the* L-L-L-L *state sequence for the complex coefficients is an unbroken chain.*

imaginary wavelet basis functions are chosen appropriately, the complex wavelet coefficients will be approximately shift-invariant.[3]

The magnitude step response of a complex wavelet basis function, shown at four consecutive scales in Figure 1(c), is a smooth envelope. In contrast to the real DWT, the CWT step response magnitude varies monotonically with the distance of the edge from the location of the wavelet basis function. This leads to strong persistence of large magnitudes around edges independent of their alignment, as shown in Figure 1(c). As a result, edge structure is more regular in the CWT coefficients than in the DWT, making the "local edge detection" heuristic (and hence the 2Populations and Persistence properties) more accurate.

The CWT is slightly redundant (by a factor of two in 1D and four in 2D), and the CWT basis functions form a tight frame. The redundancy, however, is far less than the undecimated DWT, which is $\log N$ times redundant in 1D and $3 \log N$ in 2D. The computational cost of the CWT remains $O(N)$, the same as the orthogonal DWT.

In Section 4, we will extend the wavelet-domain HMT [1] to model complex wavelet coefficients. To capture the 2Populations property, we model each wavelet coefficient as a two-state Gaussian mixture; a controlling hidden state records the magnitude of the coefficient as either S (small) or L (large). To capture the Persistence property, we introduce dependencies across scale (between each parent and its children) in the form of transition probabilities between these hidden states. We ensure Persistence by assigning high probabilities to L-L and S-S transitions. Each path down the wavelet tree forms a Markov-1 chain in such a way that the unbroken progressions L-L-L-L and S-S-S-S are more likely than broken ones such as L-S-L-L and S-S-L-S.

As a prior density on the wavelet coefficients of the noise-free data, the HMT model plays the rôle of a statistical *grammar*. Under this grammar, unbroken strings of L and S states are preferable (more "grammatically correct") than broken strings. Numerous experiments with real data have verified that the HMT is an effective model for a large class of real-world signals and images [1,2,9].

The complex wavelet HMT yields an even more powerful model. To demonstrate the effectiveness of the complex wavelet hidden Markov tree (CHMT) in modeling images, we present in Section 5 results for image denoising and texture segmentation. The results show a clear improvement over standard HMT methods ($\sim$ 1dB gain in denoising performance over the real HMT, as shown in Figure 9 and Table 2).

[3]By "shift-invariant" we mean that wavelet coefficients at a given scale carry the same information for all shifts of the signal being analyzed.

We start in Section 2 by briefly reviewing the real orthogonal wavelet transform and examining its lack of shift-invariance. Section 3 shows how to address the shift-invariance problem by using complex wavelets. In Section 4, we construct the CHMT model. Section 5 presents the improved denoising and classification results, and Section 6 contains discussion and conclusions.

2 Review of Real Wavelets

In this section, we give a concise review of the real orthogonal wavelet transform in order to set up notation and to illustrate its lack of shift-invariance.

2.1 The wavelet transform

The 1D discrete wavelet transform (DWT) decomposes a signal $f(t) \in L^2(\mathbb{R})$ in terms of a shifted and dilated "mother" wavelet $\psi(t)$ and scaling function $\phi(t)$:

$$f(t) = \sum_{\ell \in \mathbb{Z}} u_{j_0,\ell}\, \phi_{j_0,\ell}(t) + \sum_{j \geq j_0} \sum_{\ell \in \mathbb{Z}} w_{j,\ell}\, \psi_{j,\ell}(t) \tag{1}$$

where $\phi_{j_0,\ell}(t) := 2^{j_0/2}\phi(2^{j_0}t - \ell)$ and $\psi_{j,\ell}(t) := 2^{j/2}\psi(2^j t - \ell)$ [12]. If the $\{\phi_{j_0,\ell}, \psi_{j,\ell}, j \geq j_0, \ell \in \mathbb{Z}\}$ form an orthonormal basis for $L^2(\mathbb{R})$, then the *scaling coefficients* $u_{j_0,\ell}$ and *wavelet coefficients* $w_{j,\ell}$ can be calculated using the standard L^2 inner product: $u_{j_0,\ell} = \langle f, \phi_{j_0,\ell}\rangle$ and $w_{j,\ell} = \langle f, \psi_{j,\ell}\rangle$.

The wavelet transform is computed recursively using the filter bank structure shown in Figure 2(a). Given the scaling coefficients $\{u_{j+1,\ell},\ \ell \in \mathbb{Z}\}$ at scale $j+1$, we compute the scaling coefficients $\{u_{j,\ell},\ \ell \in \mathbb{Z}\}$ and wavelet coefficients $\{w_{j,\ell},\ \ell \in \mathbb{Z}\}$ at scale j by passing the $\{u_{j+1,\ell}\}$ through digital filters H_0 and H_1 and downsampling by a factor of two. The impulse responses $h_0[n]$, $h_1[n]$ of H_0 and H_1 are related to the scaling and wavelet basis functions in (1) by $\phi(t) = \sqrt{2}\sum_n h_0[n]\phi(2t-n)$ and $\psi(t) = \sqrt{2}\sum_n h_1[n]\phi(2t-n)$ [3]. To reconstruct the signal from its wavelet (and scaling) coefficients, we apply the filter bank structure shown in Figure 2(b), which computes the scaling coefficients $u_{j+1,\ell}$ at scale $j+1$ by upsampling $u_{j,\ell}$ and $w_{j,\ell}$, filtering with G_0 and G_1 (whose impulse responses are just time reversed versions of h_0 and h_1 in the orthogonal case), and adding the results.

In discrete time, we can interpret each sample of an N-point signal as a scaling coefficient at the finest scale $j = \log_2 N$, and apply the filter bank structure in Figure 2 recursively to efficiently compute the wavelet transform with $O(N)$ computational complexity [3].

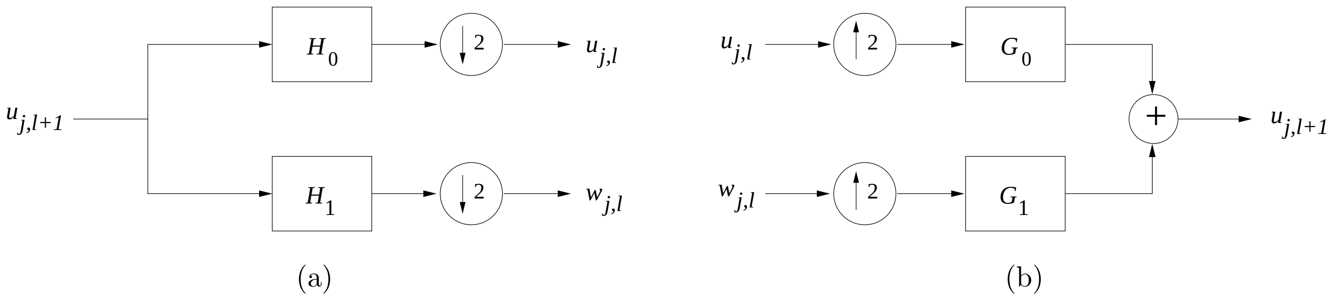


Figure 2: *Wavelet filter bank structure in 1D.*

In 2D, the discrete wavelet transform decomposes an image[4] $x(s) \in L^2(\mathbb{R}^2)$ in terms of a set of shifted and dilated wavelet functions $\{\psi^{0°}, \psi^{90°}, \psi^{\pm 45°}\}$ and scaling function $\phi(s)$:

$$x(s) = \sum_{k\in\mathbb{Z}^2} u_{j_0,k}\, \phi_{j_0,k}(s) + \sum_{b\in\mathcal{B}} \sum_{j\geq j_0} \sum_{k\in\mathbb{Z}^2} w^b_{j,k}\, \psi^b_{j,k}(s) \tag{2}$$

with $\phi_{j_0,k}(s) := 2^{j_0}\phi(2^{j_0}s - k)$, $\psi^b_{j,k}(s) := 2^j\psi^b(2^j s - k)$, and $b \in \mathcal{B} := \{0°, 90°, \pm 45°\}$. The $0°$, $90°$, and $\pm 45°$ denote the *subbands* of the wavelet decomposition (labeled as such for the directional information they give; see below). In this paper, we will consider only separable 2D wavelet transforms, meaning we can write $\psi^{0°}(s) := \psi^{0°}(s_1, s_2) = \phi(s_1)\psi(s_2)$, $\psi^{90°}(s) = \psi(s_1)\phi(s_2)$ and $\psi^{\pm 45°}(s) = \psi(s_1)\psi(s_2)$ where ϕ, ψ are 1D scaling and wavelet functions as in (1). A separable 2D DWT can be computed efficiently in discrete time by applying the associated 1D filter bank to each column of the image, and then applying the filter bank to each row of the result.

A 2D wavelet basis function $\psi^b_{j,k}$ analyzes a local region of the image x. This region is a square centered on $2^{-j}k + 2^{-(j+1)}$, $k = (k_1, k_2) \in \mathbb{Z}^2$, with width and height proportional to 2^{-j} (see Figure 3). We will call this region the *locale* of basis function $\psi^b_{j,k}$ (or the corresponding coefficient $w^b_{j,k}$). Salient features of the image in this locale (such as singularities) affect the behavior of $w^b_{j,k}$. At the next finest scale $j+1$, the locale of $w^b_{j,k} := w^b_{j,(k_1,k_2)}$ breaks into the four locales of $w^b_{j+1,(2k_1,2k_2)}$, $w^b_{j+1,(2k_1,2k_2+1)}$, $w^b_{j+1,(2k_1+1,2k_2)}$ and $w^b_{j+1,(2k_1+1,2k_2+1)}$. This allows each subband of the wavelet transform to be naturally arranged in a quad tree (in 1D we use a binary tree), illustrated in Figure 3(c), with each parent wavelet coefficient at scale j giving rise to four child coefficients at scale $j+1$ occupying the same spatial region.

Roughly speaking, we can think of wavelets as local edge detectors in the horizontal ($0°$ subband), vertical ($90°$) and diagonal ($\pm 45°$) directions. Using this heuristic, we see that the singularity structure in images leads to the 2Populations and Persistence properties presented in the Intro-

[4]We use the term "image" to mean a function of a two dimensional variable $s = (s_1, s_2)$

duction. If an image singularity is located inside the support of a wavelet basis function, then the corresponding wavelet coefficient is large (2Populations). If a basis function lies over a singularity, then its parent basis function lies over the same singularity; hence both the parent and child coefficients have large magnitude (Persistence).

For convenience, we will often enumerate the wavelet coefficients with one index, writing (for example) w_i in place of $w^b_{j,k}$. We will also use bold ($\mathbf{w}$) to denote the vector of all wavelet coefficients.

2.2 The wavelet transform and shift-invariance

A major drawback of the wavelet representation is that it is not shift-invariant. Ideally, we would like the representation of a 1D input signal[5] $f(t)$ in terms of the set $\{w_{j,\ell},\ \ell \in \mathbb{Z}\}$ of wavelet coefficients at a certain scale j to be the same for all shifts of $f(t)$. Restating, given the wavelet coefficients at scale j for one shift of the input signal, we would like to be able to predict (interpolate) the wavelet coefficients at scale j for *any* shift of the input.

To see why this is not true in general for the orthogonal wavelet transform, we examine (as in [13]) the computation of the wavelet coefficients of $f(t)$:

$$w_{j,\ell} = \langle f, \psi_{j,\ell} \rangle \sim \int_{-\infty}^{\infty} f(t)\, \psi(2^j t - \ell)\, dt = \left. (f * \psi(2^j \cdot))(t) \right|_{t=\ell}. \tag{3}$$

The $\{w_{j,\ell},\ \ell \in \mathbb{Z}\}$ are *samples*, spaced every 2^{-j} in time, of the convolution of $f(t)$ with $\psi(2^j t)$. To be able to interpolate the wavelet coefficients at different shifts, we need the Fourier product $F(\omega)\Psi(2^{-j}\omega)$ to be strictly bandlimited. Since wavelets used in practical applications have finite support in time, $F(\omega)\Psi(2^{-j}\omega)$ will in general have infinite support and be only approximately bandlimited. As a result, the $\{w_{j,\ell}\}$ will alias, and although we can design H_0, H_1 so that this aliasing cancels when we reconstruct, we cannot use $\{w_{j,\ell}\}$ to interpolate the wavelet coefficients for arbitrary shifts of the input.

Figure 1(a) illustrates the consequences of the DWT's lack of shift-invariance in the context of the 2Populations and Persistence properties. As small changes are made in the position of the 1D step edge, the energy in the edge "leaks" between scales. The profile of the wavelet coefficient magnitudes across scale is not a string of large coefficients as predicted by Persistence; rather it follows a pattern of large and small magnitudes that is different at every location.

In Section 3, we will show that we can combat this lack of shift-invariance by using complex valued wavelet filters (leading to complex valued wavelet basis functions). By analyzing only the

[5] For clarity, we present the discussion on shift-invariance in 1D. The extension to 2D is straightforward.

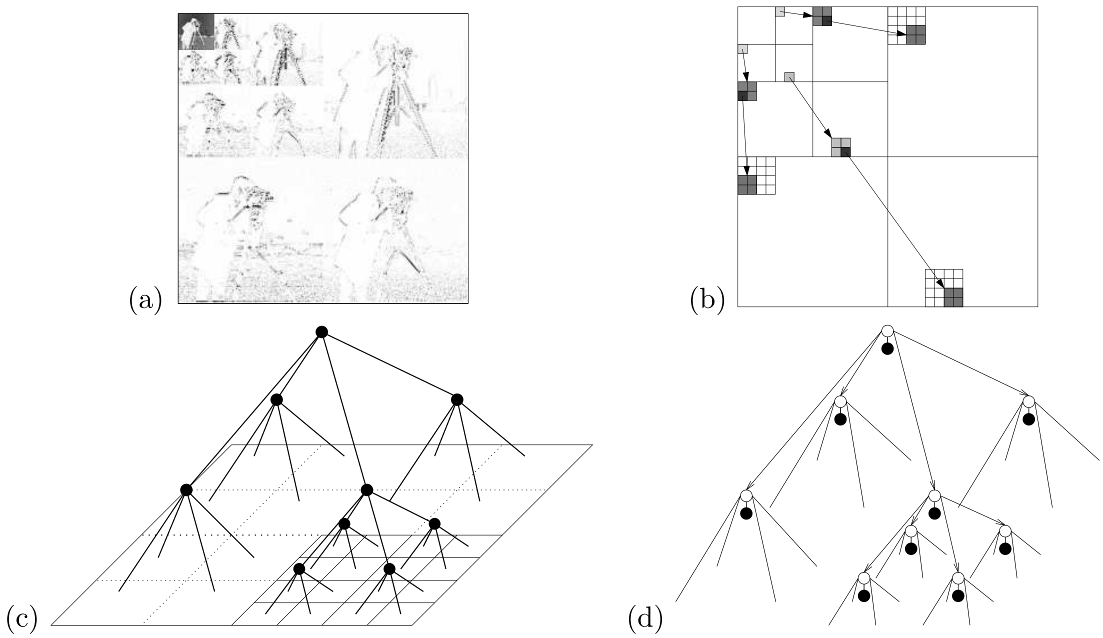


Figure 3: *(a) Wavelet transform of the* cameraman *test image; wavelet coefficients with large magnitude are black, while those with small magnitude are white. (b) Each wavelet coefficient analyzes a local region in the image. The region is broken into four subregions at the next finest scale. This leads to (c) a natural quad tree organization of the wavelet coefficients in each subband. (d) The HMT model associates a hidden state (white circle) with each wavelet coefficient (black circle). The dependencies between the wavelet coefficients are modeled as dependencies between the hidden states.*

positive frequencies in the input signal, we essentially cut its bandwidth in half and greatly reduce the aliasing.

3 Complex Wavelets

Although wavelet based algorithms have enjoyed a certain amount of success, they often produce images with ringing artifacts around the edges. The lack of shift-invariance in the orthogonal wavelet transform is the culprit: since an edge is essentially a singularity moving around in space, the real DWT represents it differently at every location. In this section, we will show that complex wavelets [6,11,14,15], by providing an approximately shift-invariant multiscale representation, strengthen the 2Populations and Persistence assumptions and hence provide a more accurate characterization of edge structure.

3.1 The dual-tree complex wavelet transform

The 1D dual-tree complex wavelet transform (CWT) [6] decomposes a signal $f(t)$ in terms of a *complex* shifted and dilated mother wavelet:

$$f(t) = \sum_{\ell \in \mathbb{Z}} u_{j_0,\ell}\, \phi_{j_0,\ell}(t) + \sum_{j \geq j_0} \sum_{\ell \in \mathbb{Z}} c_{j,\ell}\, \psi_{j,\ell}(t) \tag{4}$$

with $\phi_{j_0,\ell}$ and $\psi_{j,\ell}$ complex: $\phi_{j_0,\ell} = \phi^{\mathsf{r}}_{j_0,\ell} + \sqrt{-1}\phi^{\mathsf{i}}_{j_0,\ell}$, $\psi_{j,\ell} = \psi^{\mathsf{r}}_{j,\ell} + \sqrt{-1}\psi^{\mathsf{i}}_{j,\ell}$. The $\psi^{\mathsf{r}}_{j,\ell}$ and $\psi^{\mathsf{i}}_{j,\ell}$ are themselves real wavelets (see Figure 4(a)): the complex wavelet transform is a combination of two real wavelet transforms, in 1D the $\{\phi^{\mathsf{r}}_{j_0,\ell}, \phi^{\mathsf{i}}_{j_0,\ell}, \psi^{\mathsf{r}}_{j,\ell}, \psi^{\mathsf{i}}_{j,\ell}\}$ form a tight frame with $2\times$ redundancy. The real and imaginary parts of the dual-tree CWT are computed using separate filter bank structures (hence the name "dual-tree") with wavelet filters H^{r}_0 and H^{r}_1 for the real part and $H^{\mathsf{i}}_0, H^{\mathsf{i}}_1$ for the imaginary part.

In 2D, the CWT decomposes an image $x(s), s = (s_1, s_2) \in \mathbb{R}^2$ using dilations and translations of a complex scaling function and six complex wavelet functions

$$x(s) = \sum_{k \in \mathbb{Z}^2} u_{j_0,k}\, \phi_{j_0,k}(s) + \sum_{b \in \mathcal{B}} \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}^2} c^b_{j,k}\, \psi^b_{j,k}(s). \tag{5}$$

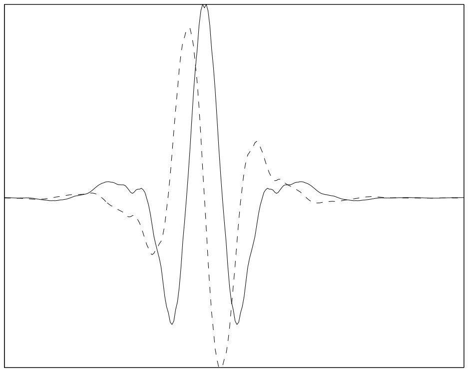
The six subbands of the 2D CWT are labeled $\mathcal{B} = \{+15^\circ, +45^\circ, +75^\circ, -15^\circ, -45^\circ, -75^\circ\}$ for the oriented direction of the wavelet function as shown in Figure 4(b). For a separable 2D CWT based on 1D complex ϕ, ψ, we have $\psi^{+15^\circ}(s) = \phi(s_1)\psi(s_2)$, $\psi^{+45^\circ}(s) = \psi(s_1)\psi(s_2)$, $\psi^{+75^\circ}(s) = \psi(s_1)\phi(s_2)$, $\psi^{-15^\circ}(s) = \phi(s_1)\overline{\psi(s_2)}$, $\psi^{-45^\circ}(s) = \psi(s_1)\overline{\psi(s_2)}$, and $\psi^{-75^\circ}(s) = \psi(s_1)\overline{\phi(s_2)}$ where $\overline{\psi}$ is the complex conjugate of ψ. The 2D CWT expansion represents $x(s)$ using a tight frame with $4\times$ redundancy.

As in the real case, we will index the complex wavelet coefficients using one number, writing c_i in place of $c^b_{j,k}$.

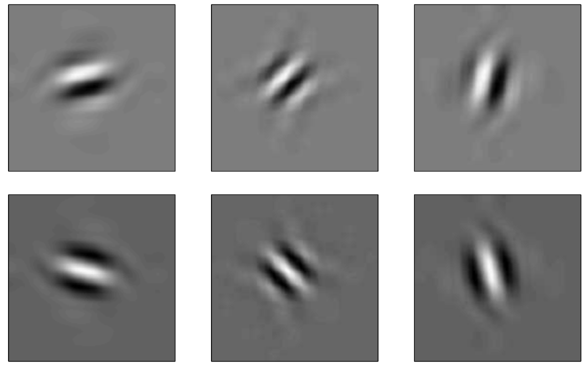
3.2 The CWT and shift-invariance

In Section 2.2, we saw that the wavelet coefficients of a signal $f(t)$ enjoy shift-invariance if either the scaling coefficients $\{u_{j,\ell}\}$ at every j or the wavelet filters H_0, H_1 are appropriately bandlimited. In this section, we will see that by using complex filters that suppress negative frequencies, we can achieve (approximate) shift-invariance.

One approach is to analyze only the positive frequencies in the input signal $f(t)$ using a DWT with real mother wavelet ψ^{r} [15–17]. Let $f^+(t)$ be the projection of f onto the positive frequencies;



(a) 1D dual tree complex wavelet basis functions



(b) 2D dual tree complex wavelet basis functions

Figure 4: *(a) Real (solid) and imaginary (dashed) parts of a complex wavelet basis function. (b) Real parts of the six 2D wavelet basis functions for a given scale and location.*

if $F(\omega)$ and $F^{+}(\omega)$ are the Fourier transforms of f and f^{+}; then

$$F^{+}(\omega) = \begin{cases} F(\omega) & \omega \geq 0 \\ 0 & \omega < 0. \end{cases} \tag{6}$$

For real $f(t)$, $F(\omega)$ is symmetric, and we can recover $f(t)$ from $f^{+}(t)$. Since $F^{+}(\omega)\Psi^{\mathsf{r}}(2^{-j}\omega)$ has half the bandwidth of $F(\omega)\Psi^{\mathsf{r}}(2^{-j}\omega)$, the aliasing in each subband is greatly reduced, and the wavelet coefficients are approximately shift-invariant. This approach analyzes a complex signal $f^{+}(t)$ with real wavelets $\psi^{\mathsf{r}}_{j,k}(t)$.

Equivalently, we can analyze $f(t)$ with two wavelets $\psi^{\mathsf{r}}, \psi^{\mathsf{i}}$ that are a Hilbert transform pair ([18] gives conditions on the filters $H^{\mathsf{r}}_0, H^{\mathsf{i}}_0$ for this, and [14, 17, 19] provide some simple design techniques). Then the $\psi^{+}_{j,k}(t) = \psi^{\mathsf{r}}_{j,k}(t) + \sqrt{-1}\psi^{\mathsf{i}}_{j,k}(t)$ only contain positive frequencies, so again $F(\omega)\Psi^{+}(2^{-j}\omega)$ has half the bandwidth of $F(\omega)\Psi^{\mathsf{r}}(2^{-j}\omega)$, leading to approximate shift-invariance. This approach analyzes a real signal $f(t)$ with complex wavelets $\psi^{+}_{j,k}(t)$.

Figure 4(a) shows an example of a 1D dual-tree complex wavelet basis function. The real and imaginary basis functions share the same locale, but are "90° out of phase"; ψ^{r} is large when ψ^{i} is small and vice versa. Also of note are the symmetry properties: if we design ψ^{r} (as in Figure 4) to be even symmetric, then ψ^{i} will be odd symmetric.

The improved shift-invariance of the CWT renders the 2Populations and Persistence heuristics much more accurate. Returning to Figure 1, recall that the magnitude of a real wavelet coefficient varies widely with the location of the singularity inside the basis function's locale — the basis functions $\psi_{j,\ell}$ are oscillatory, so their step responses are also oscillatory. As the 1D step edge in

(a) (b) (c)

Figure 5: *An edge contour (a) in the plane reconstructed using wavelet coefficients at one scale j from (b) the real DWT and (c) the CWT. The CWT reconstruction is consistent at all locations and orientations, while the DWT reconstruction varies from location to location.*

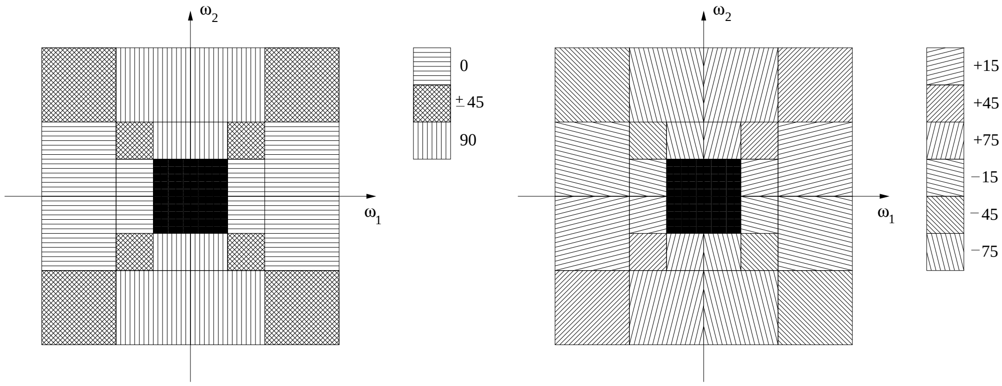
Figure 1(a) is shifted within the locale of the Daubechies-8 (D8) wavelet, the resulting wavelet coefficients swing wildly and come very close to zero for some shifts. We can also see in the figure that we do not have true persistence of magnitudes; some coefficients with small magnitudes have children with large magnitudes.

In contrast, as we can see from Figure 1(c), the magnitudes of the CWT coefficients vary smoothly with the location of the 1D step edge inside the basis function's locale. The closer the edge is to the center of the locale, the larger the corresponding magnitude. As the edge moves from one locale to the next, the energy transfers smoothly between the corresponding CWT coefficients. Thus models based on the 2Populations and Persistence properties, such as the HMT discussed in Section 4, are more accurate for images when placed on the CWT coefficient magnitudes rather than the real wavelet coefficient magnitudes.

Figure 5 shows a simple 2D contour reconstructed from one scale of its real and complex wavelet transforms. As the contour moves through space, the reconstruction using the real DWT coefficients changes erratically, while the CWT reconstructs all local shifts and orientations in the same manner.

3.3 The CWT and directional selectivity

In 1D we can interpret the complex wavelet transform as performing a subband analysis on only the positive frequencies of the signal. In 2D, this allows us to distinguish between a feature oriented at an angle α and one oriented at angle $-\alpha$. For example, the CWT has separate subbands for



(a) DWT tiling

(b) CWT tiling

Figure 6: *Frequency plane tilings for the real and complex wavelet transforms. Each shade represents the frequency localization for a subband. The three subbands of the real wavelet transform, shown in part (a), captures features along lines at angles of* $\{0°, \pm 45°, 90°\}$. *The complex wavelet transform, shown in part (b), can discriminate between features at positive and negative frequencies. Hence there are six subbands capturing features along lines at angles of* $\{+15°, +45°, +75°, -15°, -45°, -75°\}$.

features at $\pm 15°$ whereas the real wavelet transform lumps these into one "horizontal" subband (see Figure 6). An edge in an image at a $15°$ angle causes the CWT coefficients in the "$+15°$" subband to be large, while an edge at $-15°$ causes the CWT coefficients in the "$-15°$" subband to be large. In contrast, both of these edges would cause the real coefficients in the "$0°$" (horizontal) subband of the DWT to be large. Thus, the 2D complex wavelet transform provides us with greater *directional selectivity* than the DWT.

By providing explicit information about singularities at a broader range of orientations, the CWT allows us to distinguish between and characterize images that are different in more subtle ways. From a model selection standpoint, the CWT allows us to choose from a richer set of models. For example, an image with features oriented at $+\alpha°$ and another image with the same features oriented at $-\alpha°$ will always have the same (or at least very similar) DWT HMT models. But using the complex HMT, they will in general have different models, allowing us to statistically discriminate between the two images.

4 The Complex Wavelet HMT Model

The wavelet-domain hidden Markov tree (HMT) [1] models the joint probability density function of the wavelet coefficients of an image[6] by explicitly incorporating the 2Populations and Persistence properties. The modeling proceeds in two stages: modeling the marginal density of each wavelet coefficient and modeling the dependencies between the wavelet coefficients. The HMT model for the real DWT has been successfully employed for image denoising [2] and segmentation [9]. In this section we will extend the HMT to model dual-tree complex wavelet coefficients. Because of the CWT's improved Persistence and better directional selectivity, the complex HMT (CHMT) is a more accurate image model, evidence of which is shown in the applications of Section 5.

4.1 The Wavelet-domain hidden Markov tree

The HMT models the marginal distribution of each real DWT coefficient as a *Gaussian mixture.* To each w_i, we associate a discrete hidden state q_i that takes on values $m = \mathsf{S}, \mathsf{L}$ with probability mass function (pmf) $p(q_i)$. Conditioned on $q_i = m$, w_i is Gaussian with mean $\mu_{i,m}$ and variance $\sigma^2_{i,m}$. Thus, its overall marginal pdf is given by

$$g(w_i) = \sum_{m=\mathsf{S},\mathsf{L}} p(q_i = m)\, g(w_i | q_i = m), \tag{7}$$

with $g(w_i|q_i = m) \sim \mathcal{N}(\mu_{i,m}, \sigma_{i,m})$. While each wavelet coefficient w_i is conditionally Gaussian given its state variable q_i, overall it has a non-Gaussian density. Since the wavelet coefficients of images (and other piecewise smooth signals) are typically zero mean, we constrain $\mu_{i,m} = 0$.

To match the Persistence property of large and small coefficients, the HMT applies a Markov-1 dependency structure to the hidden states across scale (a hidden Markov model). Using links to denote dependencies, the HMT model takes the form of the tree in Fig. 3(d). The state transition probabilities between the connected states model the persistence of large/small coefficient magnitudes across scale. Using $\rho(i)$ to denote the index of the parent of node i, the parameter $p(q_i = m|q_{\rho(i)} = n)$ gives the probability that a child coefficient w_i has hidden state m when its parent $w_{\rho(i)}$ has state n.

The HMT model parameters consist of the Gaussian mixture means and variances $\mu_{i,m}, \sigma^2_{i,m}$, the transition probabilities $p(q_i|q_{\rho(i)})$, and the pmf $p(q_0)$ for the root state q_0. Each of the three DWT subbands is modeled with a separate (independent) model. Group these into the parameter

[6] We will discuss the HMT in 2D for concreteness, although the ideas apply to data of any dimension.

vector Θ. In practice, we usually *tie*, sharing the same HMT parameters within each scale band [1]. A reduced-parameter HMT that exploits the self-similarity of image data is derived in [2].

The HMT can be trained to capture the wavelet-domain features of the image(s) of interest using the iterative expectation-maximization (EM) algorithm [1]. For a given set of training data, the trained model $g(\mathbf{w}|\Theta)$ approximates the joint pdf of the wavelet coefficients. See [1, 2, 9] for more details.

The 2Populations and Persistence properties induce a certain structure on the hidden state quadtree of an image. By 2Populations, we expect most of the nodes to be in the S state. From Persistence, smooth regions of the image should be modeled by entire subtrees in the S state, while in singular regions there should be chains of L states branching down through scale. However, referring back to the 1D example in Figure 1(b), we see that a simple step edge causes the "broken" state sequence L-S-L-L. Since the HMT is designed to favor unbroken sequences over broken ones, the relative likelihood of the step edge will be low, and the edge will not be properly treated by HMT based algorithms.

In the next section, we will show that using the HMT model with the complex wavelet transform remedies the problem of broken state sequences around singularities.

4.2 Complex wavelet HMT

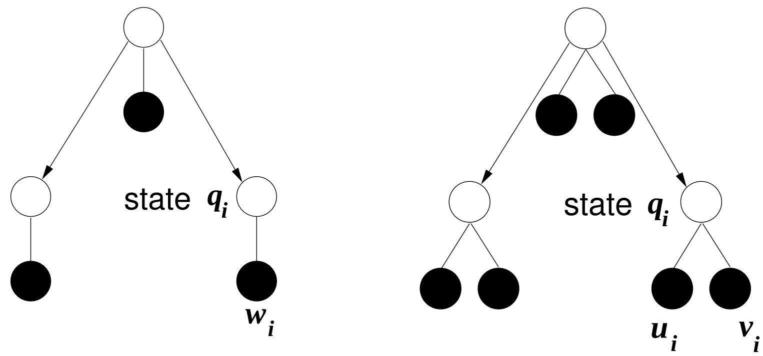
For the complex HMT, we associate with each CWT coefficient c_i a hidden state q_i taking value S,L depending on whether $|c_i|$ is small or large. Consider each complex coefficient c_i as a random vector (u_i, v_i). We approximate the marginal density $g(c_i)$ as a two-state, 2D Gaussian mixture

$$g(c_i|q_i = m) = \frac{1}{\sqrt{2\pi}\sigma_{i,m}} \exp\left(-\frac{u_i^2 + v_i^2}{2\sigma_{i,m}^2}\right). \tag{8}$$

Factoring (8) into

$$g(c_i|q_i = m) = \frac{1}{\sqrt{2\pi}\sigma_{i,m}} \exp\left(\frac{-u_i^2}{2\sigma_{i,m}^2}\right) \exp\left(\frac{-v_i^2}{2\sigma_{i,m}^2}\right) \tag{9}$$

reveals that we model the real and imaginary parts as two-state, 1D Gaussian mixtures with *shared state variable* q_i; see Figure 7. This is quite reasonable. Small $|c_i|$ result only when both $|u_i|$ and $|v_i|$ are small simultaneously (a smooth region of the image). This is captured by $q_i = \mathsf{S}$. On the contrary, large $|c_i|$ result when either $|u_i|$ is large, $|v_i|$ is large, or both are large (edge region). This is captured by $q_i = \mathsf{L}$. (Note that, as in the DWT HMT case, small outcomes from the L-state



(a) HMT (b) CHMT

Figure 7: *(a) The 1D real HMT associates a hidden state with each wavelet coefficient. (b) In the CHMT, the real and imaginary parts share a hidden state. In 2D, the wavelet coefficients and their associated hidden states are arranged in a quadtree as in Figure 3.*

Gaussian mixture density are entirely possible — and in fact likely — since this component is zero mean.)

The scatter plot in Figure 8 demonstrates that a mixture of 2D Gaussians is a good approximation to the empirical marginal distribution of the CWT coefficients in one scale. This approach is also equivalent to modeling the magnitude $|c_i|$ as a two-state Rayleigh mixture [20] and the phase $\measuredangle c_i$ as a uniform random variable. This means that the model ignores phase information in the CWT.

The CWT HMT (CHMT) corresponding to (8) with scale-to-scale Markov transitions has an almost identical structure to the real DWT HMT. The only differences will be the substitution of (8) for (7) and the use of six subband trees instead of three. The result is an HMT that supports efficient $O(N)$ training and inference algorithms while featuring better shift invariance and angle selectivity. Hence the training, inference, and processing algorithms developed in [1, 2, 9] are easily extended to the CHMT.

Referring again to 1D step edge example in Figure 1, we see that the step edge is captured by the unbroken state sequence L-L-L-L. As a result, the relative likelihood of an edge, a common structure in images, is greater in the complex case than in the real case. The progressions of large and small coefficient magnitudes in the CWT coefficients $c_i = u_i + \sqrt{-1}v_i$ are thus more reliable than in either the real part u_i or imaginary part v_i alone.

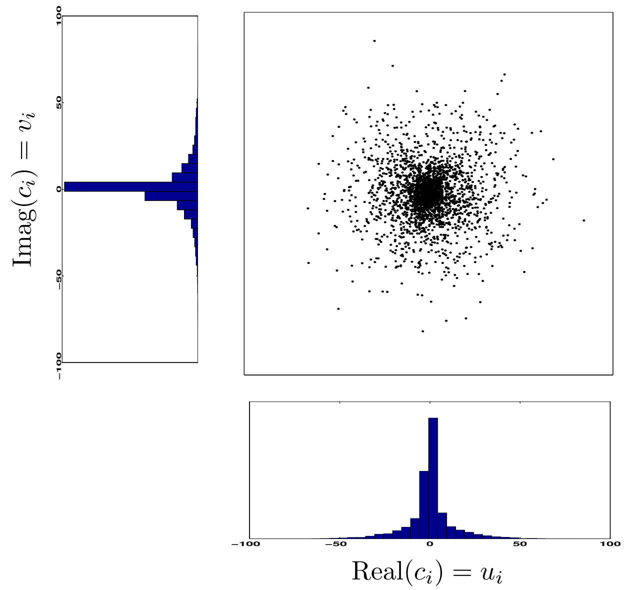


Figure 8: *Scatter plot of the joint distribution of the real and imaginary parts of the complex wavelet coefficients c_i from one subband of the "boats" test image. The marginals for the real and imaginary parts are shown as histograms along the x (real) and y (imaginary) axes. Since the joint distribution is circularly symmetric, we treat the real and complex parts as independent.*

4.3 Alternative CHMT models

As an alternative to the CHMT, we could model the real part $\mathbf{u}$ and imaginary part $\mathbf{v}$ of the wavelet coefficients separately using two standard 2-state, 1D HMTs. The combination is a 4-state, 2D HMT for each $|c_i|$ that depends on the four combinations of large/small u_i and v_i. Tying the $\mathbf{u}$ and $\mathbf{v}$ HMTs node-by-node by forcing them to share the same mixture means and variances (but not hidden state values) effectively models $\mathbf{u}$ and $\mathbf{v}$ as independent realizations of the same random process. Thus, we can apply the training and inference algorithms for multiple data observations described in [1]. There are, however, several drawbacks to this approach. First, the computational complexity increases, since we are now using a 4-state model. Second, we have lost the intuitive interpretation of state L meaning a singularity lies in the locale of the basis function and state S meaning the image is smooth over the locale of the basis function. Finally, and perhaps most importantly, having the complex and real coefficients share a hidden state is a more realistic and justifiable model.

As presented above, we model each of the subbands independently; there are 6 separate quadtrees for a CHMT image model. It is also possible to capture dependencies between subbands by using one quadtree with a vector state at each node consisting of the $\{\mathsf{S}, \mathsf{L}\}$ states for each subband [21].

4.4 A coherent CHMT model

The CHMT introduced above models the structure between only the *magnitudes* of the CWT coefficients. The phases of the coefficients are modeled as unstructured independent random variables uniformly distributed on $[0, 2\pi)$.

In [22], we show that the magnitudes and phases of a CWT coefficient and its parent follow one of three distinctive patterns depending on the image region being analyzed (see Table 1):

Smooth region: Both the coefficient and its parent have small magnitude (state S); the phase is undefined.

Isolated edge: When a single edge or other singularity resides within the locale of a CWT basis function and its parent, both coefficients have large magnitude (state L). Moreover, like the phase of a Fourier basis sinusoid, the CWT phase indicates position — in this case, the offset of the edge from the center of the locale. Since the edge positions at the two adjacent scales are simply related, so are the CWT phases. We say that due to the isolated edge, these CWT coefficients have *coherent phase*.

Texture region: When a rough textured region consisting of multiple edges or singularities resides within the locale of a CWT basis function and its parent, both coefficients have large magnitude (state L). However, due to the interaction between the edges, the CWT phases at the two adjacent scales will not be simply related. Indeed, they will appear almost statistically independent. We say that due to the texture these CWT coefficients have *incoherent phase*.

We can draw an interesting analogy between the texture phenomenon and the *speckle effect* that arises in coherent imaging systems, such as synthetic aperture radar (SAR). Speckle occurs whenever the target has features whose complexity lies on the same or smaller scale than the illumination wavelength; a complicated pattern of constructive and destructive interference occurs at the receiver due to the large phase shifts incurred on reflection. In the CWT we also have a speckle effect when we "probe" a texture region with a complex wavelet.

Thanks to the "zooming" property of the wavelet transform, as we move to finer and finer scales of a piecewise smooth signal, the texture regions will eventually disappear, since the edges

Table 1: *The three fundamental modes of CWT coefficients.*

feature	$\|c_{J,k}\|$	$\angle c_{J,k}$
smooth	S	undefined
1 edge	L	coherent
texture	L	*in*coherent

and singularities within textures will eventually isolate and begin to have coherent phase. Thus, in the limit, complex wavelet coefficients will have either small magnitude and undefined phase (smooth region) or large magnitude and coherent phase (isolated edge region).

In [22], we propose a three-state coherent CHMT model for singularity detection. The S state remains the same as in 4.2. The L state is divided in two: one state labeled T for large magnitude and incoherent phase (for rough texture regions) and one labeled E for large magnitude and coherent phase (for isolated edge regions). A CWT coefficient in hidden state T has a uniform phase distribution, while one in hidden state E has a distribution tightly centered on a value predicted from its parent. Running a state estimation (Viterbi) algorithm to identify the most likely hidden state (S, T, E) of each complex wavelet coefficient very effectively segments signals into smooth, texture, and isolated edge regions [22].

5 Applications

CHMTs are applicable to the gamut of statistical signal and image processing tasks, from estimation and detection/classification to segmentation, fusion, and synthesis. To illustrate the accuracy of the CHMT, we will extend the HMT image denoising algorithms of [1, 2] and multiscale texture classification algorithm of [9]. Results using algorithms based on the CHMT show a marked improvement over those using the standard HMT.

5.1 Image denoising

Consider the problem of estimating an image corrupted by additive white Gaussian noise of variance σ^2_{noise}. Translating the problem into the wavelet domain[7], we wish to estimate the original signal's wavelet coefficients $\mathbf{y}$ from the noisy coefficients $\mathbf{c}$:

$$\text{Estimate } y_i \text{ given } \; c_i = y_i + \eta_i, \quad \eta_i \sim \mathcal{N}(0, \sigma^2_{\text{noise}}).$$

Because the structure of the CHMT is identical to the real HMT, we can estimate the image using the following empirical Bayesian programme [1]:

1. Using the EM algorithm, train a set of CHMT parameters Θ for the complex wavelet coefficients $\mathbf{y}$ based on the observed coefficients $\mathbf{c}$.

2. Compute the posterior state probabilities $p(q_i|\mathbf{c}, \Theta)$ using the fast $O(N)$ up-down algorithm.

3. Calculate the conditional mean estimate for the noise-free complex wavelet coefficient y_i

$$\widehat{y}_i = \mathsf{E}\left[y_i|\mathbf{c}, \Theta\right] = \sum_{m=\mathsf{S},\mathsf{L}} p(q_i = m|\mathbf{c}, \Theta) \, \frac{\sigma^2_{i,m}}{\sigma^2_{\text{noise}} + \sigma^2_{i,m}} \, c_i, \tag{10}$$

   which is a linear combination of two Wiener shrinkage filters. Note that (10) shrinks the magnitude $|c_i|$ of the complex wavelet coefficient while leaving the phase $\measuredangle c_i$ unchanged.

It is also possible to forgo the training step by using a set of pre-specified *universal* parameters based on a Besov space analysis of real-world images [2].

HMT-based denoising typically outperforms standard thresholding techniques, because the degree of coefficient shrinkage is determined based not only on the size of the coefficient but also on its relationship with its neighbors across scale [1, 2].

Proper determination of the posterior state probabilities $p(q_i|\mathbf{c}, \Theta)$ is the key to the good performance. To counter variations due to shift variance when using real wavelets, these quantities should ideally be estimated based on a number of different shifts of the data. By considering all possible shifts, we obtain the shift-invariant, undecimated HMT (RHMT) model of [2]. However, the complexity rises with the number of incorporated shifts. The dual-tree CWT provides a computationally attractive alternative.

[7] Although the complex wavelet basis functions are not orthogonal, they are close enough so that the noise is approximately white in the wavelet domain.

Table 2: *Comparison of PSNRs in dB of* 256×256 *images denoised using the wavelet algorithms of Figure 9 plus* MATLAB*'s spatially adaptive Wiener filter (Wiener2) for noise variances* $\sigma_{\text{noise}} = 10$ *and* 25.

Image	*Boats*		*Lena*		*Cameraman*	
σ_n	10	25	10	25	10	25
Noisy	28.1	20.1	28.1	20.2	28.1	20.2
Wiener2	31.8	26.0	31.8	26.0	32.0	25.7
WHT	32.0	27.0	31.0	26.6	31.0	26.5
HMT	32.5	27.6	32.1	27.2	32.0	26.6
RHMT	33.2	28.0	32.7	27.7	32.6	27.1
CHMT	**33.7**	**28.5**	**33.1**	**28.4**	**32.9**	**28.0**

Table 2 and Figure 9 compare the performance of a number of different denoising algorithms. While the computational complexity of the CHMT is virtually the same as the real DWT HMT, we see that complex modeling buys us a significant performance improvement both in PSNR[8] and visually (sharper edges and smoother backgrounds). We attribute this improvement to the near shift invariance and superior directional resolution of the CWT. Indeed, the CHMT outperforms the state-of-the-art (and computationally expensive) redundant HMT denoiser of [2].

5.2 Multiscale Texture Classification

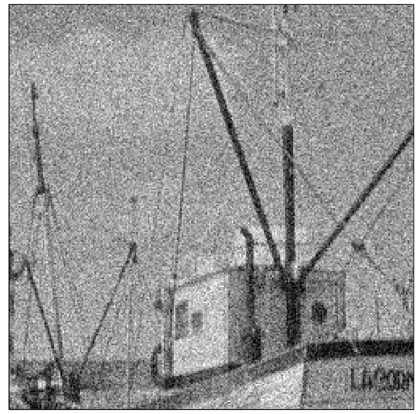
In [9], we developed a multiscale texture segmentation algorithm based on the HMT. The algorithm has two key components: a multiscale classification step (termed "raw segmentation") that uses the HMT to assign class labels at a range of scales, and a fusion step that combines these classifications into a final segmentation. In this section, we will show that the CHMT results in a better multiscale classification, helping us to arrive at a more accurate final segmentation.

Consider the standard *texture classification problem.* Given probability models $\{\Theta_l\}$ for a set of textures and an $K \times K$ data set $\mathbf{x}$, we decide which of the texture models best describes $\mathbf{x}$. This decision is based on the texture model that yields the highest likelihood; that is, we choose the index l that maximizes $g(\mathbf{x}|\Theta_l)$. As K becomes larger, the classification becomes more reliable.

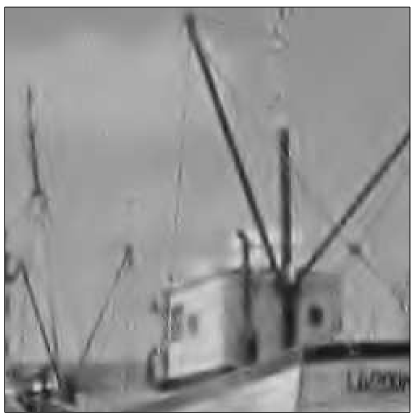
A harder problem is *supervised texture segmentation.* Given the texture models $\{\Theta_l\}$ and an

[8]Peak signal-to-noise ratio for an estimate $\widetilde{\mathbf{x}}$ of an N-pixel image $\mathbf{x}$ with values between 0 and 255 is defined as PNSR $= -20 \log_{10}\left(\frac{\|\mathbf{x}-\widetilde{\mathbf{x}}\|_2}{256\, N}\right)$.

(a) *Boats* image segment



(b) Noisy image, 20.2dB



(c) WHT denoising, 28.5dB

(d) HMT denoising, 28.6dB

(e) RHMT denoising, 28.6dB

(f) CHMT denoising, 29.3dB

Figure 9: *A 256×256 segment of the 512×512* Boats *image denoised using a variety of approaches, with PSNRs. (a) Original image with 256 gray levels. (b) Noisy image ($\sigma_n = 25$). (c) Redundant wavelet hard thresholding (WHT) using D8 wavelet [10, 23]. (d) HMT denoising using D8 DWT [1]. (e) Redundant RHMT denoising using D8 [2]. (f) CHMT denoising. The CHMT denoising algorithm gives a significant improvement both visually and in PSNR. In this particular example, the areas around the wires suffers from fewer ringing artifacts in the CHMT restored image than in the other restored images.*

image comprised of a montage of homogeneous texture regions of unknown shapes, we affix a label to each pixel in the image corresponding to the texture that best describes the surrounding area. The solution to this problem is less straightforward, since we do not have any information about the *shapes* of the regions. Given the shape information (and a model for the texture in each shape), we can use the maximum likelihood technique described above to classify each region independently.

Multiscale classification acts as a bridge between these two problems. At scale $j = 0, 1, \ldots,$ we divide the $K \times K$ image into dyadic squares of side length $2^{-j}K \times 2^{-j}K$ and then classify each dyadic square using the maximum likelihood technique. The result is a set of "raw segmentations" of the image, where we have assumed that the shapes of the texture regions are dyadic squares. The raw segmentations at coarse scales (small j) are based on many data points, and hence the decisions are more reliable than those at fine scales (large j). However, the raw segmentations at fine scales have better resolution; collections of the smaller dyadic squares can more accurately approximate the true shapes of the texture regions. Given this set of raw segmentations, a *fusion* algorithm (see [9, 24], for example) combines them into a final segmentation.

Wavelet-based texture models, such as the HMT and CHMT, are well suited for the multiscale classification problem. A subtree of wavelet coefficients with root at scale j contains information about the image inside a dyadic square of side length $\sim K2^{-j}$. By examining all subtrees of the wavelet coefficients of an image, we can assign a label to each dyadic square depending on which texture the subtree of wavelet coefficients best describes.

Wavelet HMT based methods approach the problem in two steps [9]:

1. Train HMT models for the wavelet coefficients of each texture. This requires samples (and hence a priori knowledge) of the textures found in the image.

2. Compute the likelihood of the image wavelet coefficients under each model at multiple scales, and assign the label corresponding to the model that yields the highest likelihood.

The third step is the fusion step, for which we refer the reader to [9, 24].

HMT-based multiscale classification results using real and complex wavelets are shown in Figure 10. The CHMT algorithm outperforms (gives fewer errors than) the HMT algorithm, which leads us to conclude that complex wavelets capture more of the texture's key structure. In addition, the CWT's greater directional selectivity automatically makes the classifier more robust to rotations of the texture than the HMT. In related work, [21] outlines a rotational invariant texture retrieval algorithm using a steerable HMT.

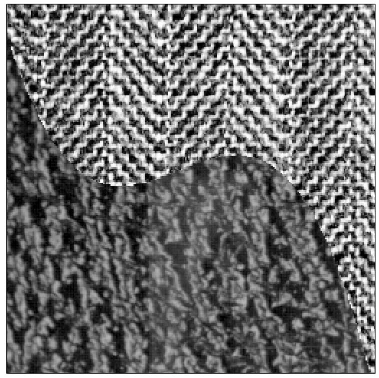
We can use this multiscale CHMT classification as a front-end to a supervised segmentation algorithm [9, 24]. These algorithms use a priori information about the shape of the texture regions to combine the raw segmentations into a final segmentation map. For example, [9] presents an algorithm based on a context model for shape, while [24] uses one based on description length. The multiscale classifier based on the complex wavelet tree can be inserted into either of these algorithms without modification. Since the multiscale classification is more accurate using the CHMT, we can expect substantial improvement in the end segmentation results. In addition, [20] uses the CHMT for unsupervised segmentation.

6 Conclusions

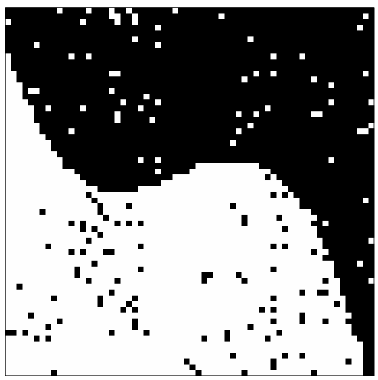
Wavelet-domain image structure can be roughly described by two general properties. The 2Populations property says that we expect small magnitude wavelet coefficients to represent smooth regions in an image, while large magnitude coefficients represent singularities (2Populations). The Persistence property describes the inter-scale relationship of the wavelet coefficients: we expect large/small magnitudes to persist across multiple scales. Although image models based on these simple rules, such as the HMT, have yielded successful processing algorithms, 2Populations and Persistence are not quite accurate for real orthogonal wavelet transforms. At the heart of the problem is the DWT's lack of shift-invariance; singularities are treated differently at different locations.

We have demonstrated that a solution to the problem of shift-variance is to use models based on the CWT. Each (complex) CWT basis function is comprised of a pair of wavelets in quadrature; roughly speaking the real and imaginary parts share the same envelope, but are 90° out of phase. The effect is a drastic reduction in the sensitivity of wavelet coefficient magnitudes to singularity locations in an image. As a result, the 2Populations and Persistence properties constitute a much more accurate description of complex wavelet domain image structure. In addition, the CWT makes phase information available, enabling models that separate the multiscale behavior of coefficients representing isolated edges from that of coefficients for rough textures.

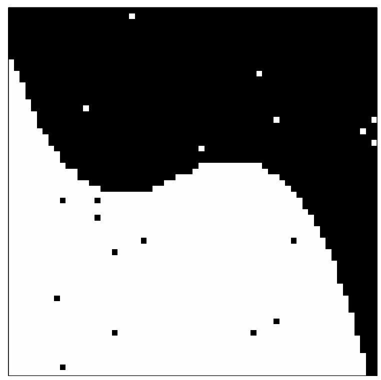
Extending the HMT — the canonical 2Populations/Persistence modeling framework — to model complex wavelet transforms is straightforward. The new model, the CHMT, inherits the HMT's fast training and inference algorithms. We have demonstrated the effectiveness of the CHMT in two applications. The denoising results show significant gains over current leading methods in both PSNR (~ 1 dB gain) and visually — the CHMT estimator produces images with crisp edges



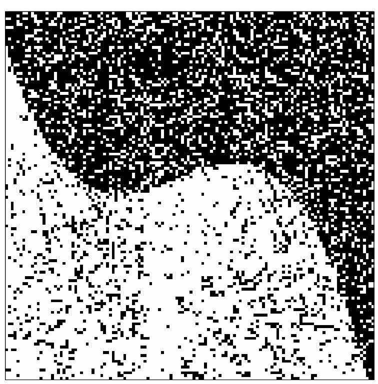
(a) texture montage



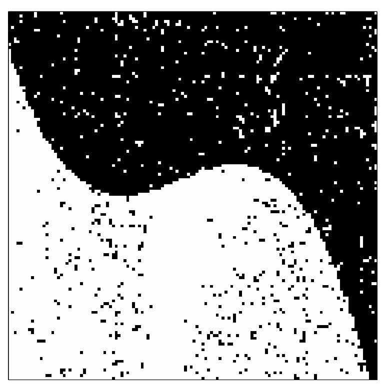
(b) 4×4 HMT



(c) 4×4 CHMT



(d) 2×2 HMT



(e) 2×2 CHMT

Figure 10: *Classification results for an image composed of two textures ("grass" and "herringbone"). (a) Texture montage image. Classification of* 4×4 *pixel regions using an HMT with (b) real wavelets, and (c) complex wavelets. Classification of* 2×2 *pixel regions using (d) real wavelets and (e) complex wavelets.*

virtually free of ringing artifacts. Multiscale classification (a front-end to supervised segmentation) using the CHMT also produces raw segmentations with fewer errors than with the usual HMT.

In this paper, we have seen that complex wavelet transform models capture singularity structure more accurately than those based on real wavelets. However, we have focussed mainly on processing algorithms that deal only with the magnitude of the wavelet coefficients. The multiscale *phase* relationships between the complex wavelet coefficients, touched on briefly in Section 4.4, also carry considerable information about singularities in the image (specifically, their spatial *location*). Building even more accurate image models by exploiting these phase relationships is a topic of current research [22].

References

[1] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Proc.*, vol. 46, pp. 886–902, Apr. 1998.

[2] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Trans. on Image Proc.*, vol. 10, pp. 1056–1068, July 2001.

[3] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 2nd ed., 1998.

[4] D. L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *Applied and Computational Harmonic Analysis*, vol. 1, pp. 100–115, Dec. 1993.

[5] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Inform. Theory*, vol. 38, pp. 719–746, Mar. 1992.

[6] N. G. Kingsbury, "Image processing with complex wavelets," *Phil. Trans. Royal Society London A*, vol. 357, pp. 2543–2560, Sept. 1999.

[7] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[8] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Proc.*, vol. 41, pp. 3445–3462, Dec. 1993.

[9] H. Choi and R. G. Baraniuk, "Multiscale image segmentation using wavelet-domain hidden Markov models," *IEEE Trans. Image Proc.*, vol. 10, pp. 1309–1321, Sept. 2001.

[10] R. Coifman and D. Donoho, "Translation-invariant de-noising," in *Wavelets and Statistics* (Antoniadis/Oppenheim, ed.), vol. 103 of *Lecture Notes in Statistics*, Springer-Verlag, 1995.

[11] N. G. Kingsbury, "Shift invariant properties of the dual-tree complex wavelet transform," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. — ICASSP '99*, (Phoenix, AZ), Mar. 1999.

[12] I. Daubechies, *Ten Lectures on Wavelets.* New York: SIAM, 1992.

[13] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable mulitscale transforms," *IEEE Trans. Inform. Theory*, vol. 38, pp. 587–607, Mar. 1992.

[14] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and Computational Harmonic Analysis*, vol. 10, pp. 234–253, May 2001.

[15] F. C. Fernandes, *Directional, Shift-Insensitive, Complex-Wavelet Transforms with Controllable Redundancy.* PhD thesis, Rice University, 2001.

[16] P. Abry and P. Flandrin, "Multiresolution transient detection," in *Proc. of the IEEE-SP Int. Symp. Time-Frequency and Time-Scale Analysis*, (Philadelphia, PA), pp. 225–228, Oct. 1994.

[17] R. L. C. van Spaendonck, *Seismic Applications of Complex Wavelet Transforms.* PhD thesis, Delft University of Technology, 2002.

[18] I. W. Selesnick, "Hilbert transform pairs of wavelet bases," *IEEE Signal Proc. Letters*, vol. 8, pp. 170–173, June 2001.

[19] 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 Process. — ICASSP '01*, (Salt Lake City, UT), May 2001.

[20] C. W. Shaffey, N. G. Kingsbury, and I. H. Jermyn, "Unsupervised image segmentation via Markov trees and complex wavelets," in *IEEE Int. Conf. on Image Proc. — ICIP '02*, (Rochester, NY), Sept. 2002.

[21] M. N. Do and M. Vetterli, "Rotation invariant texture classification and retrieval using steerable wavelet-domain hidden Markov models," *To appear in IEEE Trans. on Multimedia*, 2002.

[22] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Multiscale edge grammars for complex wavelet transforms," in *IEEE Int. Conf. on Image Proc. — ICIP '01*, (Thessaloniki, Greece), Oct. 2001.

[23] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Proc. Letters*, vol. 3, pp. 10–12, Jan. 1996.

[24] R. Neelamani, J. K. Romberg, H. Choi, R. Riedi, and R. G. Baraniuk, "Multiscale image segmentation using joint texture and shape analysis," in *Proc. SPIE Conference on Wavelet Apps. in Signal and Image Proc. VIII*, (San Diego, CA), Aug. 2000.