A Geometric Hidden Markov Tree Wavelet Model

Justin Romberg, Michael Wakin, Hyeokho Choi, Richard Baraniuk
Dept. of Electrical and Computer Engineering, Rice University
6100 Main St., Houston, TX 77005

ABSTRACT

In the last few years, it has become apparent that traditional wavelet-based image processing algorithms and models have significant shortcomings in their treatment of edge contours. The standard modeling paradigm exploits the fact that wavelet coefficients representing smooth regions in images tend to have small magnitude, and that the multiscale nature of the wavelet transform implies that these small coefficients will persist across scale (the canonical example is the venerable zero-tree coder). The edge contours in the image, however, cause more and more large magnitude wavelet coefficients as we move down through scale to finer resolutions. But if the contours are smooth, they become simple as we zoom in on them, and are well approximated by straight lines at fine scales. Standard wavelet models exploit the grayscale regularity of the smooth regions of the image, but not the geometric regularity of the contours.

In this paper, we build a model that accounts for this geometric regularity by capturing the dependencies between complex wavelet coefficients along a contour. The Geometric Hidden Markov Tree (GHMT) assigns each wavelet coefficient (or spatial cluster of wavelet coefficients) a hidden state corresponding to a linear approximation of the local contour structure. The shift and rotational-invariance properties of the complex wavelet transform allow the GHMT to model the behavior of each coefficient given the presence of a linear edge at a specified orientation — the behavior of the wavelet coefficient given the state. By connecting the states together in a quadtree, the GHMT ties together wavelet coefficients along a contour, and also models how the contour itself behaves across scale. We demonstrate the effectiveness of the model by applying it to feature extraction.

Keywords: Wavelets, edges, geometry, feature extraction

1. INTRODUCTION

Images have two salient features that any model should account for: they contain smooth, homogeneous regions, and these regions are separated by smooth edge contours. That is, images exhibit *grayscale regularity* in smooth regions and *geometric regularity* along edge contours. Both grayscale regular regions and geometric regular contours are readily characterized by their multiscale behavior. In a sense, they both become "uninteresting" as we zoom in on them; at fine resolutions, smooth regions are essentially flat and contours are essentially straight.

The *wavelet transform* handles the grayscale regularity in images naturally, and as a result has emerged as the preeminent tool in image processing. Wavelet models operate under the paradigm "wavelet coefficients representing smooth regions in the image tend to be small". This idea is the basis for many state-of-the-art wavelet domain processing algorithms for applications including compression [1,2], denoising [3,4], and segmentation [5].

An effective model for edge structure in the wavelet domain has been more elusive. At fine scales in the wavelet domain, many coefficients are needed to build up an edge (the number roughly doubles from scale to scale). But since the edge contour itself is simple (smooth), the values of the wavelet coefficients are "geometrically coherent" in that they exhibit strong dependencies, a fact which current models do not exploit.

To address this issue, recent research in applied harmonic analysis has focussed on developing alternate representations that are better suited for edges [6–10]. While these new representations have theoretically nice properties, they do not yet enjoy the same widespread success in practical applications as wavelets.

Email: {jrom, wakin, choi, richb}@rice.edu. Web: dsp.rice.edu. 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.

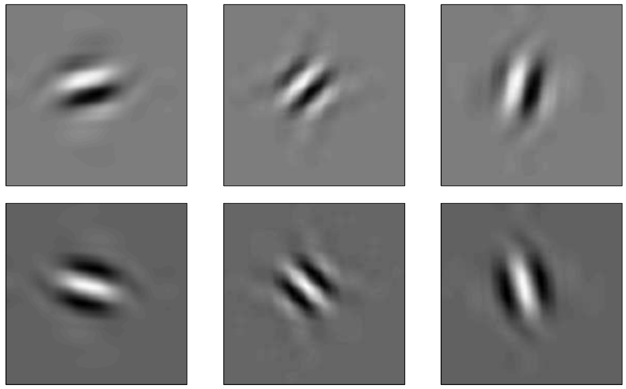

Figure 1. Real parts of the six 2D wavelet basis functions for a given scale and location. The basis functions are local in space, frequency, and direction; each responds only to edges at certain orientations.

In this paper, instead of proposing an alternate representation, we explore a method of compensating for contour regularity directly in the wavelet domain. The approach combines three recent developments in image modeling: *complex wavelets* [11], *hidden Markov trees* [3], and *multiscale geometry modeling* [12].

The complex wavelet transform (CWT) [11,13], like the real wavelet transform, decomposes an image using basis functions that are local in time and frequency. Unlike the real wavelet transform, complex wavelets are also *local in orientation* (see Figure 1), approximately *shift-invariant* and approximately *steerable*. As a result, geometrical coherency among a group of complex wavelet coefficients is easy to identify and exploit.

In Section 3, we will develop the geometric hidden Markov tree (GHMT), a statistical model for images in the complex wavelet domain. The GHMT assigns a hidden state to each group of six complex wavelet coefficients that corresponds to the local linear edge structure present in the image (see Figure 6). The presence of an edge *constrains* the CWT coefficients to behave in a certain manner, just as they are "constrained" to be small when representing smooth regions.

The GHMT captures the contour structure in the image by imposing multiscale *dependencies* between these hidden states. Since each contour in the image is smooth, their local linear behavior is closely related across scale (see Figure 2). These relationships are quantified with transition probabilities between the hidden states of parent and child CWT coefficients on the wavelet quadtree. An appropriate choice of transition probabilities allows the GHMT to favor hidden state configurations that are *geometrically faithful*; at fine scales, the contour is essentially straight, and we expect little innovation in the hidden state sequence.

In Section 4, we apply the model to the problem of *feature extraction*. Given an image, we use a fast optimization algorithm to find the state configuration with greatest *joint* likelihood. The resulting hidden state configuration gives us an estimate of the contour structure in the image across a range of scales; an example is shown in Figure 7. The pictures in Figure 7 are completely parameterized; they were redrawn from the "vector graphics" information (line segment orientations) inherent in the state configuration.

2. COMPLEX WAVELETS AND LOCAL GEOMETRY

The complex wavelet transform decomposes an image $X(s), s \in \mathbb{R}^2$ in terms of basis functions that are shifted in dilated versions of *six* different mother wavelets ψ^b:

$$X(s) = \sum_{k\in\mathbb{Z}^2} u_{j_0,k} + \sum_{b\in\mathcal{B}} \sum_{j\geq j_0} \sum_{k\in\mathbb{Z}^2} c^b_{j,k}\psi^b_{j,k}(s) \tag{1}$$

where $\psi^b_{j,k}(s) = 2^j\psi^b(2^j s - k)$. The $c^b_{j,k}$ and $\psi^b_{j,k}(s)$ are complex valued, with the real and imaginary parts of $\psi^b_{j,k}$ forming an approximate Hilbert pair. The CWT uses six subbands $\mathcal{B} = \{15, 45, 75, -15, -45, -75\}$, labeled for the directional information they give. Like the real wavelet transform, each subband can be arranged on a

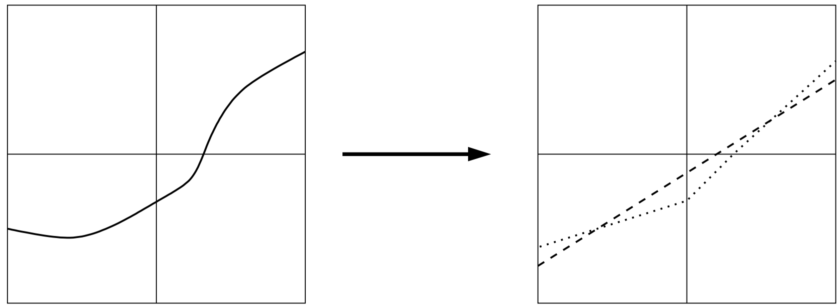

Figure 2. The linear behavior of the curve on the right is closely related across scale. The dashed line on the left shows the linear fit at one scale, and the dotted line shows the fits at the next finer scale.

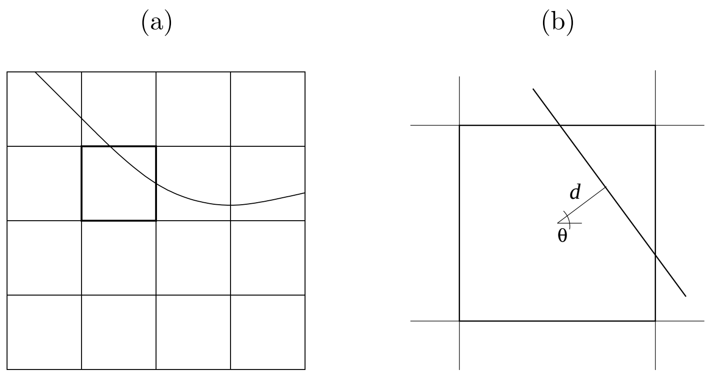

Figure 3. The local linear behavior of a contour inside the bold box in (a) is parameterized by (b) an angle θ and offset d.

quadtree with nodes indexed by scale and location. The real parts of the $\psi_{j,k}^b$ for one scale and location are shown in Figure 1.

The CWT has several properties which make it attuned to characterizing geometrical structure*:

Shift-invariance: Let $\{c_{j,k}^b\}_{b,k}$ be the CWT coefficients of an image X at a scale j. Then the corresponding scale j CWT coefficients for any *shift* of X can be calculated by interpolating the $\{c_{j,k}^b\}_{b,k}$.

Rotational-invariance: Similarly, the CWT coefficients for any *rotation* of X can be calculated by interpolating the $\{c_{j,k}^b\}_{b,k}$.

Directional selectivity: Along with being local in time and frequency, the $\psi_{j,k}^b$ are local in *orientation*.

As a result of these three properties, the CWT coefficients at one scale and location can provide a characterization of local geometrical structure. For a CWT basis function centered on location k in the plane, local edge positions can be parameterized by a distance d and an angle θ relative to k (see Figure 3). The (d, θ) parameterization provides a nice separation in the *edge response* of the six CWT coefficients (see Figure 4): d controls the *phase*, while θ controls the magnitude. The mapping from (d, θ) to the corresponding edge response is (almost) invertible. The local geometry is manifest in the CWT coefficients; simply by looking at their values, we can not only decide whether or not an edge is in the vicinity, but we can also estimate the distance d and angle θ. An example of local edge structure being "read off" from CWT coefficients is shown in Figure 5.

*While not strictly true, theses properties hold for nearly all intents and purposes.

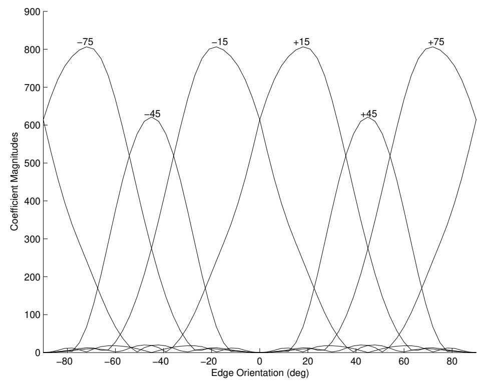

Figure 4. Magnitude edge response for the six CWT basis functions at a given scale and location as a function of θ for $d = 0$. The phase response varies linearly with d. The relative magnitudes are tied to the angle θ of local geometrical structure while the phase is tied to the offset d.

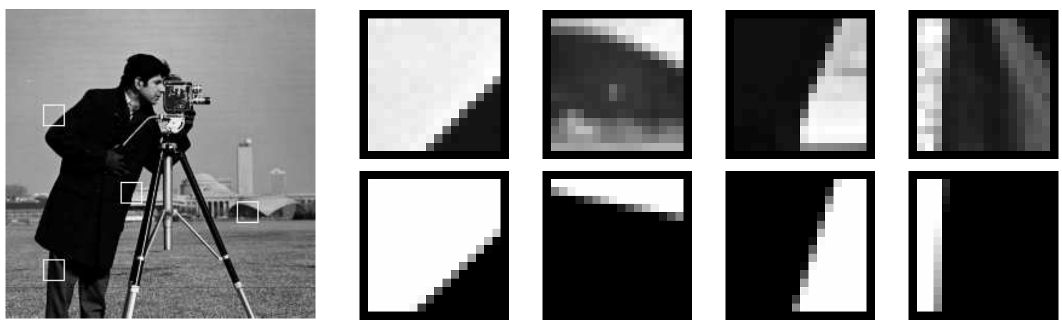

Figure 5. Local edge detection using the CWT. Highlights of different regions in the "cameraman" image are shown on the left, on the right are edge estimates from the corresponding complex wavelet coefficients. The upper row shows the original region of the image, the lower row shows the edge estimates.

3. GEOMETRIC HIDDEN MARKOV TREE MODELING

The wavelet-domain hidden Markov tree [3] (HMT) is a general purpose statistical image model underlying powerful denoising [4] and segmentation [5] algorithms. The HMT separates the wavelet coefficients $c^b_{j,k}$ of an image into two categories, S and L, and uses a different statistical model for each. Over smooth regions in the image, the wavelet coefficients are expected to have small magnitude; the HMT models these type S coefficients as coming from a low-variance Gaussian distribution. Over regions with edges or texture, the wavelet coefficients can have larger magnitude; these type L coefficients are modeled with a high-variance Gaussian. Of course, the type of each wavelet coefficient is not known a priori; it can be interpreted as a *hidden state* $q_{j,k}$ that controls the marginal distribution† of $c^b_{j,k}$

$$P(c^b_{j,k}|q_{j,k} = m) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-c^b_{j,k}/(2\sigma^2_m)\right) \quad m \in \{\mathsf{S}, \mathsf{L}\}. \tag{2}$$

To model the dependencies between the wavelet coefficients, the HMT sets up a Markov-1 dependency structure on hidden states across scale. The relationships between parent and child coefficients‡ are parameterized

†Typically, the variances σ^2_m will also vary with the scale j

‡The wavelet coefficients of an image can be naturally arranged on a quadtree, where "parent nodes" at a scale j split into four "child" nodes at scale $j + 1$.

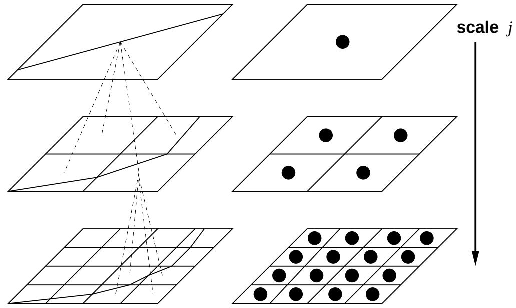

Figure 6. The geometric hidden Markov tree model. The tree of hidden states, shown on the left, models the multiscale behavior of the contours. Associated with each hidden state is a set of six complex wavelet coefficients (drawn as black dots on the right).

by transition probability matrix

$$A_{mn} = \text{Prob}\left(\text{child state} = m|\text{parent state} = n\right). \tag{3}$$

Since smooth regions of the image at one scale will break down into smooth regions at the next scale, we expect A_{SS} to be close to 1.

Rather than using a single "Gaussian with large variance" state to model contours in the image, the geometric HMT (GHMT) breaks the L state down into many *geometry states*, each corresponding to a different type of local linear contour behavior. The states are drawn from a dictionary of possible (d, θ) values, $q_{j,k} \in \ \mathsf{S} \cup \{(d_m, \theta_m)\}$, and we use the same state for all six CWT coefficients $\{\psi^b_{j,k}\}_b$ for a given scale and location (j, k).

As discussed in Section 2, each type of local linear contour structure produces distinct behavior in the CWT coefficients. Given a geometrical state $q_{j,k} = (d_m, \theta_m)$, we model the six CWT coefficients as coming from an improper Gaussian whose "mean" is the one dimensional subspace $E^{(d_m,\theta_m)}_{j,k}$ of response values of $\{\psi^b_{j,k}\}_b$ for edges of all different heights at (d_m, θ_m). We have

$$P\left(\{c^b_{j,k}\}_b|q_{j,k} = (d_m, \theta_m)\right) \propto \exp\left(-\,\text{dist}(\{c^b_{j,k}\}_b, E^{(d_m,\theta_m)}_{j,k})^2/(2\sigma^2_g)\right) \tag{4}$$

where $\text{dist}(\{c^b_{j,k}\}_b, E^{(d_m,\theta_m)}_{j,k})$ is the distance of the CWT coefficients $\{c^b_{j,k}\}_b$ to the subspace $E^{(d,\theta)}_{j,k}$ [§] and σ^2_g controls the variance around this subspace.

The transition probabilities will favor small innovations in geometry over large. To each state (d_m, θ_m), there corresponds a line ℓ_m (as in Figure 3(b)). The A_{mn} are assigned based on the Hausdorff distance $\text{HD}(\ell_m, \ell_n)$ between lines ℓ_m and ℓ_n restricted to a square in the plane. A small Hausdorff distance corresponds to little change in geometry between parent and child, and results in a high probability. We use $A_{mn} \propto e^{-\,\text{HD}(\ell_m,\ell_n)}$.

The GHMT, illustrated in Figure 6, is a general purpose statistical model for the complex wavelet coefficients of images, and thus can be applied to a range of image processing problems. In the next section, we present a simple feature extraction algorithm.

4. APPLICATION: FEATURE EXTRACTION

The hidden states in the GHMT are more than just a modeling device; they correspond to *semantic* contour information. Since edge contours are perceptually the most important features in an image, knowledge of the underlying states is interesting in its own right.

Given an image, the Markov-1 quadtree structure of the GHMT allows us to find the joint maximum-likelihood hidden state sequence $\{\widehat{q_{j,k}}\}$ using the Viterbi algorithm [14]. The result is a set of (d, θ) estimates at every

[§] We could make the distribution in (4) proper by putting a prior on the *heights* of the edges in the image.

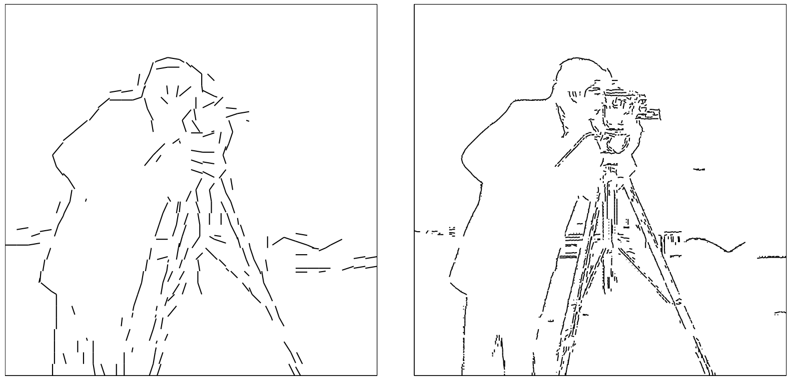

Figure 7. Contour sketches at two different scales of the "cameraman" image shown in Figure 5. The sketches were created directly from the (d, θ) values of the geometrical states estimated by the Viterbi algorithm.

node tied together by the geometry model. The collection of estimate states $\{\widehat{q_{j,k}}\}_k$ at a scale j can be used to "sketch" the contour structure in an image; the finer the scale, the more details are included.

An example of geometrical feature extraction using GHMT state estimation is shown in Figure 7. The contour sketches are re-created from the (d, θ) values inferred by the Viterbi algorithm. The estimated hidden state values are equivalent to a *vector graphics* representation of the contour sketch, and can be used to form an extremely low bitrate semantic approximation to the image.

5. CONCLUSIONS

Despite their success, traditional wavelet-based models have significant shortcomings in their treatment of edge structure in images. In this paper, we have introduced a modeling framework that accounts for geometrically smooth edge structure. The geometrical HMT models spatial clusters of complex wavelet coefficients according to an underlying hidden state specifying a linear approximation of the local contour structure. By imposing dependencies on these hidden states, we are able incorporate a characterization of geometrical structure into the model. Finally, a feature extraction algorithm, based on a joint estimation of the hidden states, was proposed that creates simple line sketch approximations to an image.

REFERENCES

1. J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing* **41**, pp. 3445–3462, Dec. 1993.
2. Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Trans. Image Processing* **6**(5), pp. 677–693, 1997.
3. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Processing* **46**, pp. 886–902, Apr. 1998.
4. H. Choi, J. K. Romberg, R. G. Baraniuk, and N. G. Kingsbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. — ICASSP '00*, (Istanbul, Turkey), June 2000.
5. H. Choi and R. G. Baraniuk, "Multiscale image segmentation using wavelet-domain hidden Markov models," *IEEE Trans. Image Processing* **10**, pp. 1309–1321, Sept. 2001.
6. E. Candés and D. Donoho, "Curvelets – a surpringly effective nonadaptive representation for objects with edges," in *Curve and Surface Fitting*, A. C. et. al, ed., Vanderbilt University Press, 1999.
7. E. L. Pennec and S. Mallat, "Image compression with geometrical wavelets," in *IEEE Int. Conf. on Image Proc. – ICIP '01*, (Thessaloniki, Greece), Oct. 2001.
8. M. N. Do, P. L. Dragotti, R. Shukla, and M. Vetterli, "On the compression of two-dimensional piecewise smooth functions," in *IEEE Int. Conf. on Image Proc. – ICIP '01*, (Thessaloniki, Greece), Oct. 2001.

9. R. Shukla, P. L. Dragotti, M. Do, and M. Vetterli, “Rate distortion optimized tree structured compression algorithms,” *IEEE Trans. Image Processing* , submitted for publication.
10. M. B. Wakin, J. K. Romberg, H. Choi, and R. G. Baraniuk, “Rate-distortion optimized image compression using wedgelets,” in *IEEE Int. Conf. on Image Proc. – ICIP ’02*, (Rochester, NY), September 2002.
11. N. G. Kingsbury, “Image processing with complex wavelets,” *Phil. Trans. Royal Society London A* **357**, pp. 2543–2560, Sept. 1999.
12. J. K. Romberg, M. B. Wakin, and R. G. Baraniuk, “Multiscale wedgelet image analysis: fast decompositions and modeling,” in *IEEE Int. Conf. on Image Proc. – ICIP ’02*, (Rochester, NY), Sept. 2002.
13. I. W. Selesnick, “Hilbert transform pairs of wavelet bases,” *IEEE Signal Processing Lett.* **8**, pp. 170–173, June 2001.
14. L. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc. IEEE* **77**, pp. 257–285, Feb. 1989.