Tactile vs. Visual Exploration of Texture

1 Introduction and Background

The visual system utilizes multiple cues to quickly segment and process parts of the visual scene in a preattentive manner, and texture plays an important role in this process (Thielscher and Neuman 2006). The ubiquity of texture is encountered in tasks involving target detection, region segmentation and classification, and visual attention; and it extends to unexpected domains such as scene identification (see, e.g., Renninger and Malik 2004), Thus, investigating the computational and statistical nature of texture can help us better understand and model vision in general.

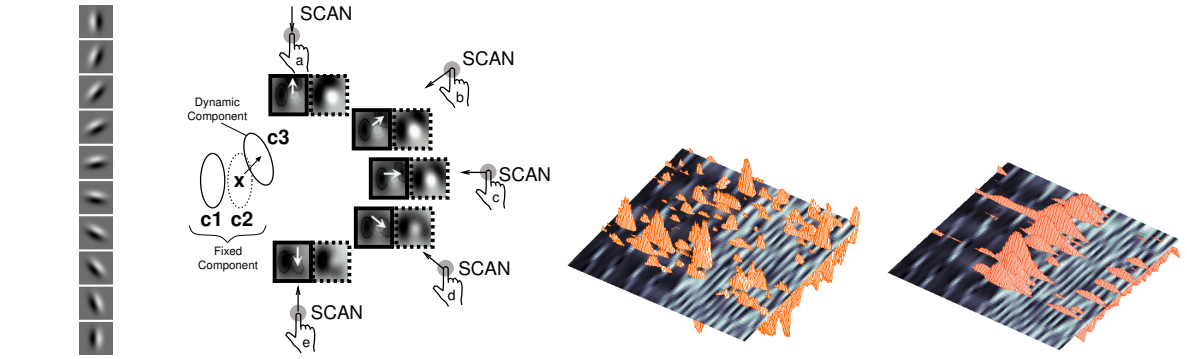
There are many methods for texture modeling and processing (Beck 1966, 1983; Efros and Leung 1999; Jain and Farrokhnia 1991; Julesz 1965, 1986; Julesz and Bergen 1982; Liu and Wang 2002; Wei and Levoy 2000). These methods employ a wide range of techniques, but they are all based on the assumption that texture is fundamentally a visual property. Hence, most of these methods are based on the use of the Gabor filter, a model of the visual cortical receptive field (see, e.g., Bovik et al. 1990; Fogel and Sagi 1989; Paragios and Deriche 2002; Zhang et al. 2002). There are more modality-neutral, statistical approaches to texture using second-order statistics (Beck 1983), Bayesian approaches (Chan and Vasconcelos 2006; Zhu et al. 2009), parametric models (Portilla and Simoncelli 2000), generative models (Heess et al. 2009), texton-based (Malik et al. 1999; Varma and Zisserman 2005), Markov Random Fields (MRF, Manjunath and Chellapa 1991; Varma and Zisserman 2003), and generalizations of the approach (Zhu et al. 1998). However, the main insights driving these models are deeply rooted in vision.

However, texture is more immediately a surface property (see e.g., He and Nakayama 1994; Nakayama et al. 1995; Oh and Choe 2007a). Objects in the 3D environment usually have uniform texture, not a mixture, as shown in Figure 1. So, texture tasks such as segmentation as we usually define it may simply be an artifact of a scene where multiple 3D objects are occluding each other. In this case, a texture boundary may appear near occlusion boundaries in the 2D projection of the 3D scene, but not in 3D. With this in mind, we can think of an interesting possibility: Since surface properties are more intimately related to touch than to vision, tactile perception can play a more immediate role in texture processing. Also, since there is no ambiguity in 3D tactile texture, it can provide a natural ground truth to the visual system.

Interestingly, the basic functional organization of the somatosensory system (especially area 3b, the hand area in macaque monkeys) exhibits a close resemblance to that of the visual system, including the receptive field characteristics and topological mapping of the sensory field (Bensimaia et al. 2008; Deibert et al. 1999; DiCarlo and Johnson 2000). As shown in Figure 2a&b, the

(a) Texture (b) Bark (c) Stone (d) Texture in a 3D context

Figure 1: **Nature of Texture.** A typical texture segmentation task is shown in (a), with its constituent textures in (b) and (c). However, the traditional texture segmentation task is non-existent in a 3D, tactile context as shown in (d) since the two objects (tree trunk and stone) are spatially separated as readily detectable by the sense of touch (the white square marks the texture shown in (a)). Note: the visual texture segmentation problem still exists in this case, but it could be alleviated via a 3D approach (Oh and Choe 2007b).



(a) Visual RFs (b) Tactile RFs (c) Visual RF response (d) Tactile RF response

Figure 2: **Visual Receptive Fields and Tactile Receptive Fields and Their Responses.** (a) Visual receptive fields (RFs) have a Gabor-like pattern, with different orientation, phase, and spatial frequency (dark represents inhibitory region and bright excitatory region). Note that visual RFs also have a temporal pattern, which is not shown here (see DeAngelis et al. 1993). (b) Tactile RFs are similar to visual RFs (e.g., C1 and C2) except for a dynamic inhibitory component (marked C3) whose position relative to the fixed components C1 and C2 changes, centered at "X", shifts to the opposite direction of the scan direction (e.g., the tip of the index finger). The arrow on the finger tip shows the scan direction; the box with a solid outline shows how the dynamic inhibitory component is shifted (white arrow) relative to the scan direction; and the box with the dotted outline shows the resulting tactile RF shape. Adapted from DiCarlo and Johnson (2000). (c) Visual RF response on a texture with a boundary in the middle is shown in 3D. (d) Tactile RF response on a texture with a boundary in the middle is shown in 3D. The tactile RF response has a more prominent hump near the texture boundary than that of the visual RF. Adapted from Bai et al. (2008).

receptive fields in the somatosensory and the visual cortical neurons have a striking similarity (with unique differences, e.g., a dynamic component dependent on tactile scanning direction).

The above observation raises an interesting question regarding the relationship between the tactile modality and texture in general (including visual texture). In concrete terms, the question is about the functional relationship between tactile receptive fields and statistical properties in textural stimuli. The visual analog of this question can be found in the various models and theories relating Gabor-like receptive fields to the statistics of natural scenes (Bell and Sejnowski 1997; Choe and Miikkulainen 2004; Hoyer and Hyvärinen 2000; Miikkulainen et al. 2005; Olshausen and Field 1996; Reinagel and Laughlin 2001; Reinagel and Zador 1999). The basic idea here is that the receptive fields (tactile vs. visual) are optimized to represent the statistical regularities in the type of stimulus they are normally exposed to (texture vs. non-textural natural scenes). The main novelty of our work is the pairing of tactile receptive fields with textural input. For example, our preliminary results in Figure 2c&d show that tactile receptive field models give better representations for texture boundary detection than visual receptive field models.

In this project, we will systematically investigate the relationship between texture and the tactile processes modeled after the somatosensory cortex. The relationship is yet largely unknown, and our project is expected to be the first comprehensive study of the relationship. As a natural outcome of the project, we will develop powerful multisensory texture processing algorithms.

2 Research Questions, Rationale, and Significance

In this project, we will tackle the five following research questions. Each question is accompanied by the rationale, brief method, and expected outcome.

Q1. Are tactile receptive fields better for texture segmentation than visual receptive fields? If texture is more intimately related to touch than vision, tactile receptive fields must be

functionally better in texture tasks than visual receptive fields. Our preliminary results show that using the tactile receptive field model as a filter gives better texture boundary detection performance than the visual receptive field model (Bai et al. 2008). We will conduct an exhaustive experiment to further validate these initial results, using well-established classifiers such as support vector machines. A positive outcome from this task will further strengthen the tie between texture and the tactile modality.

Q2. Would an identical cortical development model give rise to tactile receptive fields when exposed to textural input, while visual receptive fields when exposed to natural scene input? The mammalian sensory cortex is a fairly uniform medium, and experiments have shown that they are plastic enough to switch between different sensory modalities, e.g., auditory vs. visual (Sharma et al. 2000; von Melchner et al. 2000). Furthermore, they are tuned to the statistical regularities of the stimulus they are exposed to. Thus, it is an interesting question if textural input can drive the development of receptive fields with tactile properties as shown in Figure 2b. We will test this idea by training a visual cortical development model (LISSOM, Miikkulainen et al. 2005) with texture input. We expect the resulting receptive fields to resemble tactile receptive fields (see our preliminary results in Park et al. 2009a).

Q3. Are the representations constructed by tactile receptive field responses to different texture classes more separable than those of visual receptive field responses? Another task where the touch–texture relationship can be tested is texture classification, where different texture patterns are supposed to be labeled. Given different texture classes, we can compare tactile receptive field representations (basically a response vector) vs. visual receptive field representations. We will use kernel Isomap (Choi and Choi 2004) and kernel Fisher discriminant analysis (Mika et al. 1999a) to project the high-dimensional response vectors onto their low dimensional embedded manifolds, and measure the class separability. We expect the tactile receptive field representations to be more separable. Such an outcome would further support the link between texture and the tactile modality.

Q4. Do the tactile receptive field response distributions based on textural input have similar properties (e.g., power-law) as those of the visual receptive field response to natural scene input? We can further analyze the computational property of tactile receptive fields in the context of textural input environments. A key property of visual receptive response is the power-law response distribution (Field 1987). We showed that there is a functional significance of this property (Lee and Choe 2003; Sarma and Choe 2006), i.e., saliency detection. We will test if a similar power law exists in the tactile receptive field response to texture inputs. If it turns out to be the case, we can exploit the same downstream mechanisms (such as saliency detection) in the visual pathway and adapt it to the tactile receptive field response. This task will give us important insights on how early sensory processes of different modalities can plug into a common higher-level mechanism.

Q5. Can different dynamic components in tactile receptive fields resulting from different modes of motor exploration aid in texture processing? Finally, as shown in Figure 2b, tactile receptive fields have a dynamically changing component that depends on the direction of scan of the skin patch. This means that given the same textural input, the resulting information in the response can become different, depending on which direction the skin patch was scanned. Such an active control of information content is an interesting feature not found in visual receptive fields. We will systematically test how much deviation in information content result from different directions of scan, and test if there is an optimal direction of scan given a particular type of texture. We expect to discover ways in which active sensing can enhance the information content in texture tasks (cf. Choe and Smith 2006).

Significance: The significance of our project can be summarized as follows:

- **Novelty**: The main novelty of our work is in the discovery of the linkage between texture (defined in non-modality specific terms) and the tactile modality, and an exhaustive, systematic evaluation of this concept. Furthermore, the use of tactile mechanisms in visual texture processing tasks is also unprecedented. The novelty and significance of our pilot results have been recognized by the research community with two major awards (**best scientific paper award** at ICPR 2008, Bai et al. 2008, and **best student paper award** at IEEE CIMSVP 2009, Park et al. 2009a).
- **Thematic coherence and integration**: The single theme penetrating the entire proposed project is the intimate relationship between texture and the tactile modality. Each research question addresses different aspects of the theme, but the results from one task can reinforce others. For example, if tactile representations of texture is found to be more powerful, we can infer that it would lead to higher task performance (e.g., segmentation and classification).
- **Technical merit**: First of all, the use of tactile receptive fields in a visual task itself has technical merit. Furthermore, we will employ a diverse set of learning algorithms to conduct our research. Most of the tools we will use result from our own prior work: Bai et al. (2008); Choe and Miikkulainen (2004); Choe and Smith (2006); Choi et al. (2008); Lee and Choe (2003); Miikkulainen et al. (2005); Park et al. (2009a); Sarma and Choe (2006), where several of which appeared in top AI conferences such as AAAI (Choe and Smith 2006; Choi et al. 2008; Sarma and Choe 2006).

3 Prior Work

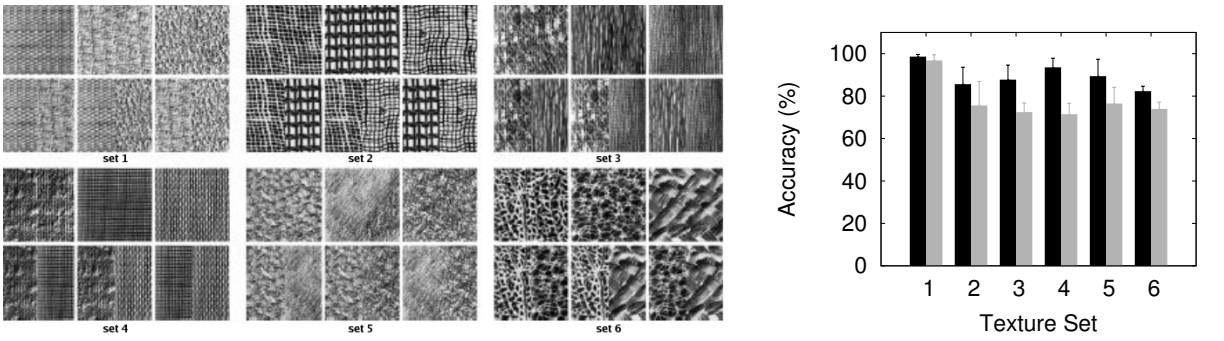
In this section, we will present our prior work that forms the basis of our proposed research: (1) tactile receptive fields used in texture boundary detection, (2) model of visual cortical development (receptive fields and lateral connections), and (3) response distribution analysis methods.

3.1 Use of Tactile Receptive Fields in Texture Boundary Detection

To test the idea that texture (in general) and tactile processes in the somatosensory system are closely related we compared texture boundary detection performance using two different types of feature vectors: one based on responses from a tactile receptive field model (DiCarlo and Johnson 2000), and the other based on those from a visual receptive field model (Daugman 1980; Jones and Palmer 1987). Figures 2a&b show examples of the two types of receptive fields. Given texture input with boundary or no boundary (Figure 3a), we generated response vectors using the tactile or the visual receptive fields by sweeping across the input image (note that tactile texture was represented as an image as well). The resulting vectors were used as feature vectors, based on which a classifier was trained to distinguish between boundary and no-boundary condition (we used a backpropagation network for this task). The results are shown in Figure 3b, where the texture boundary detection performance turns out to be significantly higher for the representation based on tactile receptive fields ($n = 100, p < 0.03$, Bai et al. 2008). These results and further analysis indicate that tactile receptive fields are better suited for texture tasks than visual receptive fields, thus illustrating the close connection between touch and texture.

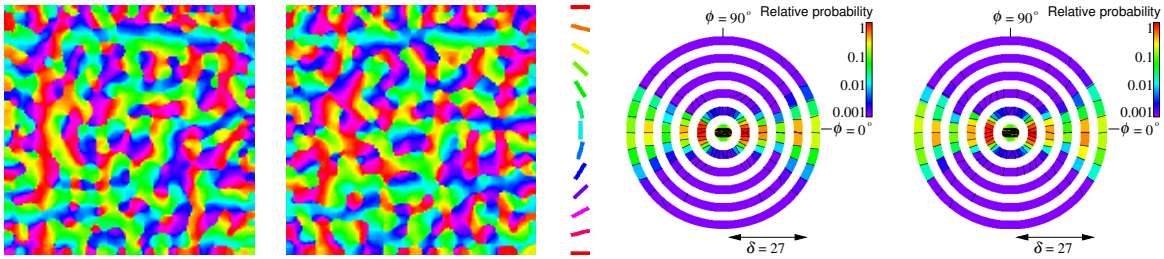
3.2 Computational Model of Visual Cortical Development

Previously, we developed a laterally connected self-organizing map model of visual cortical development and function (LISSOM, Choe 2001; Choe and Miikkulainen 2004; Miikkulainen et al. 2005). In the model, the connections learned through an activity-dependent Hebbian learning rule. In that work, we showed that difference in input distribution gives rise to differential organization of receptive field properties and lateral connection statistics, and in turn it leads to differentiation



(a) Input sets (b) Performance (black: tactile, gray: visual)

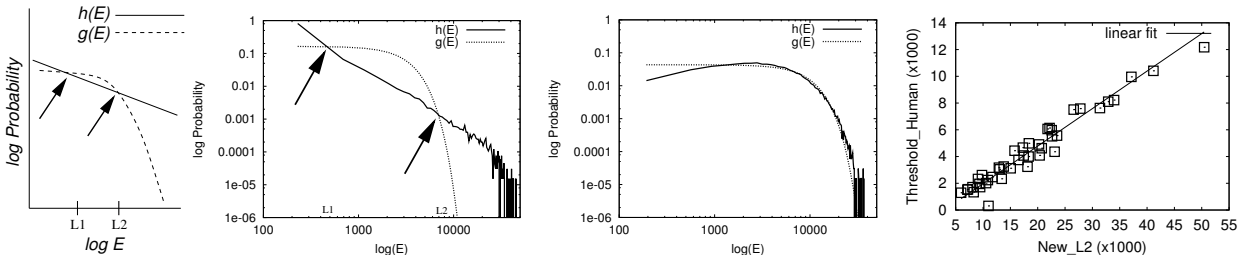
Figure 3: **Texture Input Sets and Performance Comparison.** (a) Six texture sets used in the boundary detection experiment are shown. In each set, the top row shows the no-boundary condition, and the bottom row the boundary condition. (b) Boundary detection performance (per cent correct) by tactile-receptive-field-based vs. visual-receptive-field-based representations is shown. All cases except for texture set 1, the differences were statistically significant. Adapted from (Bai et al. 2008).



(a) Ori. Map (10° curv.) (b) Ori. Map (25° curv.) (c) (d) Lat. Stat (10° curv.) (e) Lat. Stat (25° curv.)

Figure 4: **Cortical Development Model under Varying Input Distributions.** Afferent receptive field properties and lateral connection statistics of cortical maps trained under different input distributions are shown. (a) The orientation preference of afferent connections in a map trained with lines with a maximum 10° curvature is shown. Each pixel corresponds to one neuron in the map, and its orientation preference is represented in color. (b) The same as a, with the only difference being the maximum curvature ($= 25^\circ$). The two maps in a and b are clearly different. (c) The color key for a and b are shown. (d) The lateral connection statistics is shown for the map trained with 10° maximum curvature. ϕ represents the direction of a target receptive field with respect to the source receptive field centered at the origin, and δ the distance between the two. At each polar coordinate (ϕ, δ), the color represents the normalized likelihood of target receptive fields appearing at that location; and the black oriented line represents the most likely orientation of the target receptive field at that location, relative to the source receptive field. (e) The same as d, for a map trained with 25° maximum curvature is shown. In both the afferent and the lateral connections, a slight change in the input distribution causes a noticeable difference in the resulting cortical organization. Adapted from Miikkulainen et al. (2005).

of functional performance. The divergence of afferent and lateral connection properties under two different input distributions are shown in figure 4. (It is also interesting to note that even with different weight initialization, if the input presentation sequence was the same, the maps developed almost identical orientation preference (Miikkulainen et al. 2005).) Our model can be extended to test how different classes of input such as textures can give rise to different types of receptive fields, e.g., tactile-like receptive fields. In fact, Wilson et al. have been successful in extending LISSOM for modeling the whisker barrel cortex in rodents (Wilson 2007; Wilson et al. 2009), and he is working on an extension for the auditory cortex. Also, our prototype extension to enable textural inputs resulted in promising results, as briefly presented in Park et al. (2009a,b). The LISSOM model is implemented in an open-source neural map simulator called Topographica.



(a) Power-law (b) Gabor resp. (nat. scene) (c) Gabor resp. (white noise) (d) Saliency threshold

Figure 5: **Gabor Response Distribution Analysis for Saliency Detection.** (a) Response (E) distribution exhibiting a power-law property ($h(E)$) is contrasted against a normal distribution with the same variance ($g(E)$), in log-log scale. Power-law distribution shows a characteristic declining line, with a heavy tail (much higher probability than normal for $E > L2$ [marked by an arrow]). (b) Orientation energy response (E, calculated from Gabor filters) to natural images ($h(E)$) shows a power-law distribution. (c) Orientation energy response to white-noise images ($h(E)$) shows a close-to-normal distribution. (d) Orientation energy response threshold based on the $L2$ value (derived from comparison to white-noise baseline as in c) shows an almost linear fit to human-selected saliency threshold (42 input images, correlation coefficient $r = 0.98$). Adapted from Sarma and Choe (2006).

3.3 Neuronal Response Distribution Analysis for Saliency Detection

Receptive field properties of neurons determine their response properties. One characteristic statistical property of Gabor filters (a standard model of the receptive field of visual cortical neurons) is the power-law response distribution (Field 1987). Figure 5a illustrates this property. In log-log scale, power-law distribution exhibits a straight declining slope $h(E)$, where E is orientation energy calculated with Gabor filters (Geisler et al. 2001). In contrast, the normal distribution $g(E)$ with an identical variance rapidly asymptotes to 0 as E grows. Such a power-law property is almost universally observed in response distributions on natural images (e.g., Figure 5b).

What is the functional significance of such a property? Field (1987) and Olshausen and Field (1996) argued that the power-law property is due to sparse coding in the cortex. However, an immediate utility of such sparse coding is unclear. In Sarma and Choe (2006) and Lee and Choe (2003), we addressed this limitation. An interesting feature of the power-law distribution is that compared to a matching normal distribution that has the same variance, the distributions always intersect at two points. Beyond the second intersection ($L2$ in Figure 5), power-law distribution shows significantly higher probability than the baseline normal distribution. The point $L2$ turned out to be strongly correlated with the optimal saliency threshold of E chosen by human subjects (Figure 5d), and response distribution to white noise images show near-normal distribution (i.e., no salience: Figure 5c). Furthermore, $L2$ can be easily estimated from the standard deviation of the response, which can be calculated as the weighted sum of the squared response, followed by a square root nonlinearity. Thus, Gabor filters seem to enable a simple downstream computation of a meaningful functional quantity (i.e., optimal saliency threshold), and that works only when the inputs are natural images (but not when the input is a white-noise image). We expect a similar analysis can reveal an analogous relationship between tactile receptive fields vs. textures (as compared to visual receptive fields vs. natural images).

3.3.1 Results from Prior NSF Support

The PI has been supported by NSF grant #0905041 (*CRCNS data sharing: Whole Mouse Brain Neuronal Morphology and Neurovascular Browser*, 9/1/2009–8/31/2011, $182,476), #1208174 (*Open Web Atlas for High-Resolution 3D Mouse Brain Data*, 9/1/2012–8/31/2014, $200,868), and

#1256086 (*Enhanced Knife-Edge Scanning Microscopy for Sub-micrometer Imaging of Whole Small Animal Organs*, 6/1/2013–5/31/2015, $502,746). **Intellectual merit:** The grants resulted in innovative brain imaging and neuroinformatics frameworks. Combined, the projects resulted in three journal papers (Choe et al. 2011b; Chung et al. 2011; Mayerich et al. 2011b), seven conference papers (Choe et al. 2011a; Kwon et al. 2011; Mayerich et al. 2011a; Sung et al. 2013; Yang and Choe 2010, 2011a,b), two Ph.D. dissertations (Sung 2013; Yang 2011), two M.S. theses (Choi 2013; Kim 2011), and an online mouse brain atlas (see below). **Broader impacts:** Based on the data and neuroinformatics platform developed through NSF support, we launched the KESM Brain Atlas web site to deliver our submicrometer-resolution mouse brain atlases ($\sim$ 4.5 TB) and published associated code on SourceForge (project name: kesmba). The PI also gave a tutorial on brain connectivity mapping (International Joint Conference on Neural Networks, Dallas, TX, 2013) and ran open exhibits at the Society for Neuroscience annual meeting (Washington, DC, 2011; New Orleans, MS, 2012) for broader dissemination and to foster interest. The PI hosted several lab tours for the K-12 audience during Texas A&M University's *Aggieland Saturday* event and during the national *Discover Engineering* event. Three Ph.D. students, one M.S. student, and two REU students were trained, directly funded by these grants and supplements.

4 Research Plan

The research plan is organized into 5 parts, addressing the 5 research questions in Section 2:

Research Plan	Research Questions (from Section 2)
Section 4.1	Q1: Are tactile receptive fields better for texture segmentation than visual receptive fields?
Section 4.2	Q2: Would an identical cortical development model give rise to tactile receptive fields when exposed to textural input, while visual receptive fields when exposed to natural scene input?
Section 4.3	Q3: Are the representations constructed by tactile receptive field response to different texture classes more separable than those of visual receptive field responses?
Section 4.4	Q4: Do the tactile receptive field response distributions based on textural input have similar properties (e.g., power-law) as those of the visual receptive field response to natural scene input?
Section 4.5	Q5: Can different dynamic components in tactile receptive fields resulting from different modes of motor exploration aid in texture processing?

4.1 Tactile vs. Visual Receptive Fields in Texture Segmentation

As shown in Figure 3, our preliminary results showed promise in the use of the tactile receptive field model in texture boundary detection tasks (Bai et al. 2008). In this preliminary study, the task was simply to determine whether there is a texture boundary in the middle of the image or not. This is a good first step, but it is far from a full-blown texture segmentation task. We propose to extend our approach in (Bai et al. 2008) to compare tactile vs. visual receptive fields in texture segmentation.

We will compare the texture segmentation performance based on the visual receptive field model (Gabor filters) vs. the tactile receptive field model. The tactile counterpart of the primary visual cortical receptive field model (Gabor filters) is the three-component model of the somatosensory area 3b neuron's receptive field (DiCarlo and Johnson 2000). Figure 2b shows an overview of the model. The tactile receptive fields consists of three components: central excitatory region paired with an inhibitory lobe, and a temporally, dynamically inhibitory lobe, lagging in the opposite direction of finger scan, anchored at the center of the excitatory lobe. Following DiCarlo and Johnson (2000), each component can be expressed as:

$$G(x,y) = a \cdot \exp\left(-\frac{1}{2}L^T S^{-1} L\right), \text{where } L = \begin{bmatrix} x-\mu_x-\mathbf{v}_x\tau \\ y-\mu_y-\mathbf{v}_y\tau \end{bmatrix}, S = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^w \end{bmatrix} \quad (1)$$

where (x, y) represents the pixel location, $(\mu_x,\ \mu_y)$ represents the center of the lobe, $(\mathbf{v}_x, \mathbf{v}_y)$ represents the stimulus velocity vector, and τ represents the delay of the peak excitation or inhibition

with respect to skin stimulation. The center of the excitatory lobe was fixed to stay at the middle of the receptive field while the paired inhibitory lobe and lagging inhibition lobe centers varied with respect to the excitatory lobe's center. The parameters a, σ_x, σ_y, and ρ determine the amplitude, spread, orientation, and elongation (aspect ratio) of the excitatory ($a > 0$) or inhibitory ($a < 0$) lobe represented by the Gaussian function. Finally, the three Gaussian lobes are linearly summed to construct the tactile receptive field model.

We will initially use the Brodatz texture collection (Brodatz 1966) to construct the texture input, and later include textures found in natural images (e.g., close-up of natural scenes containing repeating elements). The input images will be preprocessed with a Laplacian of Gaussian filter to simulate thalamic center-surround processing. Then, we will convolve the resulting output with the tactile receptive field model, and for comparison, with the visual receptive field model. Note that for tactile receptive fields, depending on the nominal scanning direction, different receptive fields will form, as shown in Figure 2b. Whether to use only one scan direction, or combine results from multiple scan directions is a research issue we will investigate as part of this project. Post-processing for texture segmentation can be done in several different ways, such as the use of boundary vs. no-boundary classifiers (as we did in Bai et al. 2008) or texture type classifiers (see Section 4.2). Learning algorithms such as support vector machines will be used for the classifier component. The results are expected to provide further support for the texture-touch relationship.

Research issues: (1) What are the important differences between tactile and visual receptive fields? (2) Do our preliminary results scale up to larger, more diverse texture sets? (3) How can a fair comparison be made between tactile vs. visual receptive field based representations? (equalizing degree of freedom in filter parameterization, normalization of response, etc.). (4) Can accounting for the temporal dimension of V1 RF help level the performance? (5) Could the TRF's advantage be due to an end-stopping property (Bolz and Gilbert 1986; Gilbert 1994; Gilbert and Wiesel 1979; Zetzsche and Nuding 2005)? If so, would a more fair comparison be against these types of RFs in V1 (and beyond)? (6) Can actual mechanotransduction in touch (Orr et al. 2006), not only its informational content, contribute to touch-based texture processing?

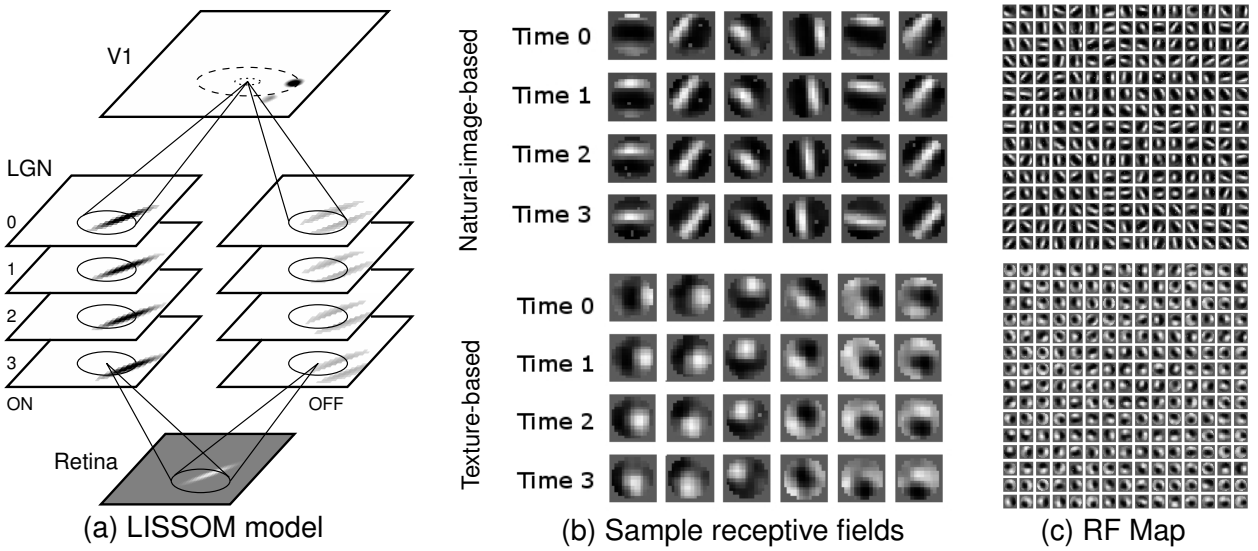
4.2 Simulated Cortical Development of Tactile vs. Visual Receptive Fields

To test the whether or not textural input could have led to tactile-like receptive fields (in the cortex), we will train, using texture images as input, our general cortical development model LISSOM (laterally interconnected synergistically self-organizing map, Choe and Miikkulainen 2004; Miikkulainen et al. 2005). Figure 6a shows the overall architecture of the model. This is a model that includes multiple time-lagged maps in the thalamus, to account for dynamic properties in the receptive field such as the sensitivity to direction of motion (DeAngelis et al. 1993). The following description closely follows Choe and Miikkulainen (2004); Miikkulainen et al. (2005).

The activation of the thalamic layers (here, we will call it the lateral geniculate nucleus [LGN], although this only applies to the visual pathway) is done through a fixed on-center/off-surround (ON channel) and off-center/on-surround (OFF channel) receptive fields made of difference of Gaussian functions. The cortical neurons are activated as follows:

$$s_{ij} = \gamma_{\mathrm{A}} \left(\sum_{ab \in \mathrm{ON}} \xi_{ab} A_{ab,ij} + \sum_{ab \in \mathrm{OFF}} \xi_{ab} A_{ab,ij} \right), \tag{2}$$

where ξ_{ab} is the activation of neuron (a, b) in the receptive field of neuron (i, j) in the thalamic ON or OFF channels, $A_{ab,ij}$ is the corresponding afferent weight, and γ_{A} is a constant scaling factor. The afferent stimulation is squashed using the sigmoid activation function, forming the neuron's initial response as



(a) LISSOM model (b) Sample receptive fields (c) RF Map

Figure 6: **Preliminiary Cortical Development Results.** (a) The general architecture of the LISSOM model is shown. Activity propagates from the bottom, from retina to lateral geniculate nucleus (LGN) to visual cortex (V1). There are multiple time-lagged maps in LGN to model dynamic receptive fields (direction sensitivity). (b) Self-organized receptive fields are shown. The top set, trained with natural scene input, shows visual-like receptive fields. The bottom set, trained with texture, shows tactile-like receptive fields. Each cortical neuron, represented in each column, is associated with 4 frames with different time lag, thus showing a dynamic sensitivity (e.g., oriented bar moving in a specific direction). (c) The full organization of the map is shown (a 15 $\times$ 15 grid sampled from a 48 $\times$ 48 full map). Note that only the first frame in the dynamic receptive field is shown. Again, the top was natural-scene-trained, and the bottom texture-trained. (Preliminary results presented in Park et al. (2009b).)

$$\eta_{ij}(0) = \sigma(s_{ij}), \tag{3}$$

where $\sigma(\cdot)$ is a piecewise linear sigmoid.

After the initial response, lateral interaction sharpens and strengthens the cortical activity over a very short time scale. At each of these subsequent discrete time steps, the neuron combines the afferent stimulation s with lateral excitation and inhibition:

$$\eta_{ij}(t) = \sigma\left(s_{ij} + \gamma_{\mathrm{E}} \sum_{kl} \eta_{kl}(t-1) E_{kl,ij} - \gamma_{\mathrm{I}} \sum_{kl} \eta_{kl}(t-1) I_{kl,ij}\right), \tag{4}$$

where $\eta_{kl}(t-1)$ is the activity of another cortical neuron (k,l) during the previous time step, $E_{kl,ij}$ is the excitatory lateral connection weight on the connection from that neuron to neuron (i,j), and $I_{kl,ij}$ is the inhibitory connection weight. All connection weights have positive values. The scaling factors γ_{E} and γ_{I} represent the relative strengths of excitatory and inhibitory lateral interactions, which determine how easily the neuron reaches full activation. The above cortical activation step is repeated several times so that the lateral interaction can further refine the cortical activity, while the afferent input stays fixed.

Once the cortical activation has settled to a stable state, afferent and lateral connection weights are adapted using a normalized Hebbian learning rule:

$$w'_{pq,ij} = \frac{w_{pq,ij} + \alpha X_{pq} \eta_{ij}}{\sum_{uv}(w_{uv,ij} + \alpha X_{uv} \eta_{ij})}, \tag{5}$$

where $w_{pq,ij}$ is the current afferent or lateral connection weight (either A, E or I) from (p,q) to (i,j), $w'_{pq,ij}$ is the new weight to be used until the end of the next settling process, α is the learning rate

for each type of connection (α_{A} for afferent weights, α_{E} for excitatory, and α_{I} for inhibitory), X_{pq} is the presynaptic activity after settling (ξ for afferent, η for lateral), and η_{ij} stands for the activity of neuron (i, j) after settling, Afferent inputs (i.e. both ON and OFF channels together), lateral excitatory inputs, and lateral inhibitory inputs are normalized separately.

For this task, we will extend the Topographica neural map simulator. We will use texture (Brodatz and textures from natural images) and non-textural natural images to train the LISSOM model. In order to develop dynamic receptive fields, we will generate a successive sequence of four inputs on the input layer ("retina" in Figure 6a) by moving it across the given texture or natural image. Training will follow a standard schedule of learning parameter update (see Miikkulainen et al. 2005). We will measure orientation preference, orientation selectivity, direction preference, and other quantities based on the learned connection weights. Figure 6b&d shows preliminary results, where texture-trained receptive fields show tactile-like property, while natural-scene-trained ones show visual-like property. We expect similar results with a more exhaustive input set.

Research issues: (1) Would different texture elements of varying scale (size) interfere with self-organization? (2) How can we quantitatively measure the "tactile-ness" and "visual-ness" of the receptive fields? One way to do this is to use a fitting algorithm to parametrically fit Gabor patterns or the tactile three-component model to the self-organized receptive fields and measure the degree of fit. (3) Are architectural adjustments necessary for LISSOM to properly model the somatosensory system? For example, is ventroposterior nucleus (VP) projecting to area 3b functionally similar to the LGN? Furthermore, how can we model the scan-direction dependence?

4.3 Separability of Tactile vs. Visual Representations of Texture (Classification)

With the formal tactile and visual receptive field models (three-component model and Gabor), and with their self-organized versions (using LISSOM), we can measure their representational power in terms of texture. One such measure is to compare the class separability in tactile or visual space. Another method is to actually check the texture classification performance.

To find the dominant factors for texture classification, we will apply kernel Fisher discriminant (KFD) (Mika et al. 1999b) to the feature spaces of the receptive field response. KFD is a generalized version of Fisher discriminant analysis (or linear discriminant analysis, LDA) using kernel trick as in support vector machines or kernel principal component analysis (Schölkopf and Smola 2002). KFD is a good first choice for visualizing high-dimensional data such as the tactile and visual response vectors, and can also be used for classification. Here, we briefly review KFD. The basis function in the feature space can be obtained by maximizing the ratio of the within-class scatter matrix in the feature space to the between-class scatter matrix in the feature space, as in LDA. Let $\mathcal{X}_i = \{x_1^i, x_2^i, ..., x_{l_i}^i\}, (i = 1, ..., C)$, be samples from C classes and $\mathcal{X} = \bigcup_i^C \mathcal{X}_i$. Suppose $\Phi(\cdot)$ is a nonlinear mapping function to the feature space, then the within-class scatter matrix in feature space, $\boldsymbol{S}_{\boldsymbol{W}}^{\Phi}$, is given by

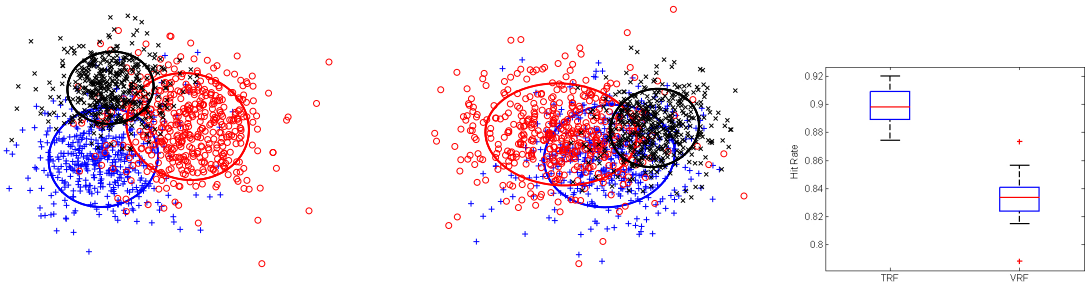
$$\boldsymbol{S}_{\boldsymbol{W}}^{\Phi} = \sum_{i=1}^{C} \sum_{x \in \mathcal{X}_i} (\Phi(x) - m_i^{\Phi})(\Phi(x) - m_i^{\Phi})^T, \tag{6}$$

where $m_i^{\Phi} = \frac{1}{l_i} \sum_{j=1}^{l_i} \Phi(x_j^i)$. The between-class scatter matrix in feature space is given by

$$\boldsymbol{S}_{\boldsymbol{B}}^{\Phi} = \boldsymbol{S}_{\boldsymbol{T}}^{\Phi} - \boldsymbol{S}_{\boldsymbol{W}}^{\Phi}, \tag{7}$$

where the total scatter matrix in feature space, $\boldsymbol{S}_{\boldsymbol{T}}^{\Phi}$, is given by $\boldsymbol{S}_{\boldsymbol{T}}^{\Phi} = \sum_{x \in \mathcal{X}} (\Phi(x) - m^{\Phi})(\Phi(x) - m^{\Phi})^T$,where $m^{\Phi} = \frac{1}{|\mathcal{X}|} \sum_{i=1}^{C} l_i m_i^{\Phi}$and $|\mathcal{X}|$ is the sample size.

We will also use kernel Isomap (Choi and Choi 2004, 2007; Choi et al. 2008). In kernel Isomap Given N objects with each object being represented by an m-dimensional vector $\boldsymbol{x}_i$, $i = 1, \ldots, N$,



(a) KFD proj. of tactile response (a) KFD proj. of visual response (c) Texture classification

Figure 7: **Kernel Fisher Discriminiant Analysis of Tactile and Visual Response Vector Space, and Texture Classification.** Projections using Kernel Fisher Discriminant analysis of (a) Tactile response and (b) visual response to three texture classes. The ellipses are equidistant traces from each class' average with 1.5 times standard deviation. TRF response space is more separable. (c) Texture classification (three different texture classes) based on projection of tactile and visual response vectors onto the first two eigenvectors from KFD show superior performance for the tactile receptive field ($n = 30$, k-NN classifier). The box-and-whisker plot shows the median, top and bottom 25-percentile, standard deviation, and outliers (+ marks). (Unpublished pilot results.)

the kernel Isomap algorithm finds an implicit mapping which places N points in a low-dimensional space. In contrast to Isomap (Tenenbaum et al. 2000), the kernel Isomap can project novel data points onto the discovered low-dimensional space, as well, through a kernel trick. The kernel Isomap mainly exploits the solution of the additive constant problem, the goal of which is to find an appropriate constant to be added to all dissimilarities (or distances), apart from the self-dissimilarities, that makes the kernel matrix positive semidefinite. Given a distance matrix, we use Dijkstra's geodesic distances (shortest paths) $\boldsymbol{D}$, and calculate the doubly centered kernel matrix as below.

$$\boldsymbol{K} = -\frac{1}{2}\boldsymbol{H}\boldsymbol{D}^2\boldsymbol{H}, \tag{8}$$

where $\boldsymbol{D}^2 = [D_{ij}^2]$ means the element-wise square of the geodesic distance matrix $\boldsymbol{D} = [D_{ij}]$, $\boldsymbol{H}$ is the centering matrix, given by $\boldsymbol{H} = \boldsymbol{I} - \frac{1}{N}\boldsymbol{e}_N\boldsymbol{e}_N^\top$ for $\boldsymbol{e}_N = [1 \ldots 1]^\top \in \mathbb{R}^N$. Then, we make the kernel matrix positive definite by adding a constant, c.

$$\widetilde{\boldsymbol{K}} = \boldsymbol{K}(\boldsymbol{D}^2) + 2c\boldsymbol{K}(\boldsymbol{D}) + \frac{1}{2}c^2\boldsymbol{H}, c = \text{largest eigenvalue of} \begin{bmatrix} \boldsymbol{0} & 2\boldsymbol{K}(\boldsymbol{D}^2) \\ -\boldsymbol{I} & -4\boldsymbol{K}(\boldsymbol{D}) \end{bmatrix}. \tag{9}$$

Eq. (9) implies substituting $\widetilde{\boldsymbol{D}}$ for $\boldsymbol{D}$ in Eq. (8), which is given by $\widetilde{D}_{ij} = D_{ij} + c(1 - \delta_{ij})$, which makes the matrix $\boldsymbol{K}$ positive semi-definite. The term δ_{ij} is the Kronecker delta. Finally, projection mapping $\boldsymbol{Y}$ is obtained by eigen-decomposition of $\widetilde{\boldsymbol{K}} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T$: $\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{\Lambda}^{\frac{1}{2}}$.The projection for novel data points is similar as in kernel PCA and is described in (Choi and Choi 2007). These methods are expected to quantitatively evaluate the representational power of the tactile vs. the visual representation of texture.

We will apply KFD and kernel Isomap to the responses of tactile and visual receptive fields (both formal models and self-organized ones) on texture inputs from several texture classes (from the Brodatz set and natural textures). Figure 7a&b shows preliminary results using KFD. We used the square root function as the kernel function for both cases. The figure shows that the tactile responses give clusters that are more separable across texture classes than those based on visual responses. We expect such results to carry over to more than three texture classes. We will then

perform classification using the projections on the KFD and kernel Isomap basis functions. Again, standard classifiers such as support vector machines (in this case, the multi-class variant) will be used. Figure 7c shows our preliminary results (with k-NN classifier in this case) where tactile representation exhibits higher classification performance compared to visual representation.

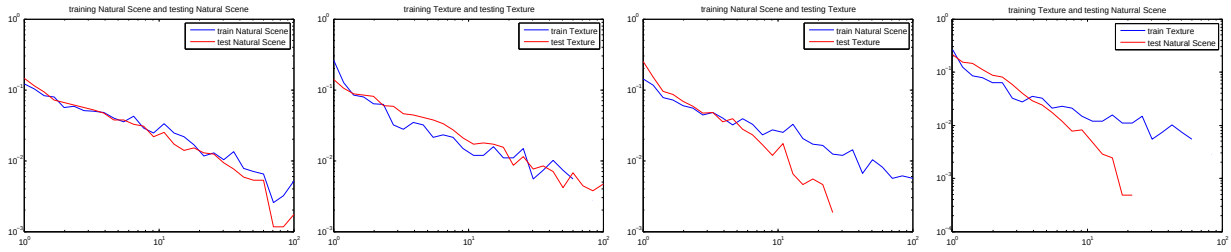
Research issues: (1) Would visual representations show higher separability on natural scenes compared to tactile representations? (2) Can a combined tactile-visual representation show higher separability on textural inputs than any of the two representations considered separately? Bimodal visuo-tactile neurons found in fMRI studies can be a good model (Tal and Amedi 2009). (See our work on manifold integration Choi et al. 2008, and also Liu and Wang 2002.) (3) Would using kernel Isomap result in similar conclusions as KFD? Our initial results with kernel Isomap (data not shown) suggest that tactile representations of texture can be more compact (i.e., fewer dominant factors).

4.4 Analysis of Tactile vs. Visual Receptive Field Response Distribution

As discussed in Section 3.3, Gabor filters, formal model of visual cortical simple cells, show a power-law property in their response distribution. Our prior work suggest that this property can have a simple yet elegant functional significance, i.e., for saliency thresholding. Furthermore, we showed that such a power-law response is only observed when the inputs were natural images. Thus, the visual receptive fields are specialized to represent the statistics in natural images. A natural question that arises from this line of thought is whether tactile receptive field response properties are tuned to textural input in a similar manner.

First, we need to estimate the response distribution, and then we can find the matching normal distribution with the same variance for comparison. The response distribution can be estimated from the responses (E) using a histogram (bin size can be variable), followed by normalization, as $h(E) = \frac{f(E)}{\sum_{x \in B_h} f(x)}$where $f(E)$ is the frequency of response value E in the histogram, B_h is the set of histogram bin locations, and $h(E)$ is the resulting probability mass function which specifies the response distribution for the filtered image. The baseline normal distribution can be calculated as follows. We calculate the raw second moment of the E distribution (i.e., the expected value of E^2) for the input image as $\sigma_h^2 = \sum_{x \in B_h} x^2 h(x)$. We use this calculated σ_h^2 to find the matching continuous normal probability density function $\mathcal{N}(x; 0, \sigma_h^2)$ with mean 0, variance σ_h^2 for all $E \in B_h$ and normalize it to find the discretized normal probability mass function $g(E)$ of the response level E: $g(E) = \frac{\mathcal{N}(E;0,\sigma_h^2)}{\sum_{x \in B_h} \mathcal{N}(x;0,\sigma_h^2)}$.Finally, we can also calculate the baseline distribution given white-noise input. The only step needed is to make the variance of the baseline distribution to be the same as the original response distribution. Since the standard deviation of a random variable scaled by the factor of c is $c \times \sigma$ where σ is the standard deviation before scaling, we simply multiply the white noise response with a constant σ_h/σ_r, where σ_h and σ_r are the standard deviations from the original natural image and the white-noise image, respectively. Then the resulting response matrix has the same variance as the original distribution calculated from a given natural image. See Figure 8 for example.

Again, we will use both the formal receptive field models and the self-organized receptive fields for this experiment. We will calculate the response distribution of tactile and visual receptive fields to texture, non-textural natural images, and white-noise images. The resulting response distributions will be compared. We expect to find that tactile receptive fields are suboptimal when given natural images, and visual receptive fields when given textural images. Initial results are encouraging, where tactile receptive fields exposed to natural image show a more normal (Gaussian) characteristic, and visual receptive fields to texture a similar trend (Figure 8).



(a) Train Nat., Test Nat. (b) Train Tex., Test Tex. (c) Train Nat., Test Tex. (d) Train Tex., Test Nat.

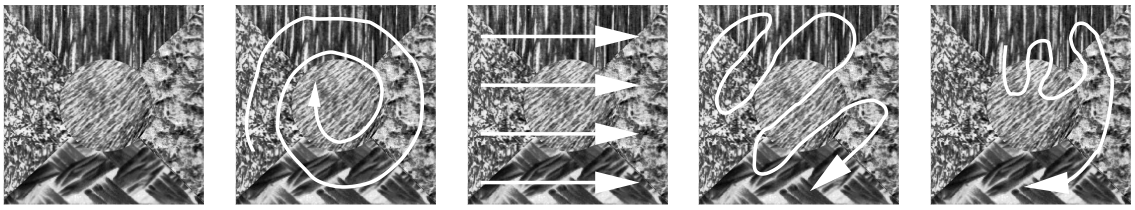
Figure 8: **Tactile and Visual Response Distributions.** Response distributions of tactile and visual receptive fields to texture and natural images are shown. The x-axis represents response level, and the y-axis the probability, in log-log scale. All four possible combinations of pairing are shown. In all plots, the blue curve shows the response of the self-organized receptive fields given the same inputs used during training (Nat. for natural images, Tex. for texture). The red curve shows the response of the same receptive fields to a different set of inputs not used during training. (a) Receptive fields trained on natural images show similar response distribution to other natural images (red). (b) Receptive fields trained on texture show similar response distribution to other textures (red). (c) Receptive fields trained on natural images show closer to normal distribution (i.e., the probability quickly drop to 0) when exposed to texture (red). (d) Receptive fields trained on texture show closer to normal distribution when exposed to natural images (red). (Unpublished pilot results.)

Research issues: (1) What would be the equivalent of white noise images for the tactile receptive fields? (2) How can the "normal-ness" of the response distribution to non-optimal stimulus be quantitatively measured? Kurtosis could be a good initial candidate (Field 1994). (3) Would textures with different scale in their texture element cause difficulty in the analysis? (4) If the difference in the optimal vs. non-optimal stimulus cases is only in the slope of in log-log scale (i.e., the scaling exponent), how can we interpret it and exploit it? (5) For the tactile response, can we find a threshold that is functionally meaningful as the saliency threshold in the case of visual response? (6) Can the point of intersection of optimal vs. non-optimal response distributions have a functional significance? (see, e.g., intersection points in Figure 8c&d).

4.5 Motor Control of Information Through Dynamic Tactile Receptive Fields

As we have reviewed in Section 1 and in Figure 2b, the tactile receptive field has a dynamic inhibitory component that shifts its position depending on the tactile scanning direction. The role of this dynamic component in terms of information processing is unknown, and in this task we will investigate it. In the tasks above, the scanning direction is assumed to be either fixed (typically, across the texture boundary), or responses from all scanning directions will be used. Allowing the direction of scan to change dynamically based on the momentary information content in the scan path can lead to more powerful feature representations (e.g., Figure 9). This theme resonates with our prior work on sensorimotor basis of understanding (Choe and Smith 2006; Choe et al. 2007). We will also take hints from the experimental and modeling research on active whisking in rats (Grant et al. 2009; Mitchinson et al. 2007): Rats are known to actively deflect their whiskers to enhance their tactile perception. Active generation of tactile cues to aid vision (Weichert et al. 2005) is another promising direction (inverse of our approach), but we will not pursue this.

As the input, we will construct plain or mixed textures with complex boundaries as shown in Figure 9a. First, we will generate response vectors based on a fixed scan path regardless of the texture region layout in the input (e.g., Figure 9b-d), and generate a response map based on the vectorized dot product of each momentary dynamic receptive field and the local image patch at that location. The resulting response maps will be tested using the methods outlined in the previous sections. We expect to find diverging performance dependent on the scan path. After



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

Figure 9: **Various Scanning Paths.** Various possible scanning paths are shown, for use with the tactile receptive field where scanning direction can alter the receptive field property. (a) An example texture segmentation problem with 5 component textures is shown. (b-d) Global scanning strategies are illustrated. (e) Local, adaptive scanning strategy is illustrated.

we have confirmed the potential utility of a customized scan path, we will develop an algorithm for adaptive scan path, using a local information measure to guide the scanning direction. This will be a form of active learning, where the algorithm itself determines the input sample. In this case, the scan path can show something like Figure 9e.

Research issues: (1) Would the motor exploration strategy affect both segmentation and classification performance? (2) Would motor exploration affect the separability of the tactile representations? (3) What could be a reasonable resolution of the scan path (how densely packed should the scan lines be)? (4) What kind of local information measure would be appropriate?

5 Education and Dissemination Plan (Broader Impact)

Curriculum development and graduate education: An important aspect of the research component of this project is that the subject requires an interdisciplinary approach, ranging from computer science, cognitive science, to neuroscience. The new interdisciplinary course the PI designed and taught in spring 2003 (Intelligent Neural Systems) and offered again in spring 2004 and spring 2005 (as Computations in Neural and Biological Systems) was received very well among students. For that course, the PI was awarded the departmental graduate teaching award in spring 2004. The course contained interdisciplinary material from artificial intelligence to cognitive science and neuroscience, and in the course the close relationship among these fields and their mutual relevance were made explicit. Now, the course is being taught as a regular course (Cortical Networks, taught in spring 2007). As part of this project, the PI will incorporate the research methodology and outcome resulting from this project into the interdisciplinary curriculum, with an emphasis on multimodal processing.

Undergraduate research and under-represented groups: The PI is also committed to equally motivating and mentoring underrepresented groups, and acknowledge that this mission is best carried out in the undergraduate classroom. The PI has been participating in the Undergraduate Summer Research Grant (USRG) program sponsored by the College of Engineering at Texas A&M. The program attracts excellent undergraduate students from across the country, and many of these students are from minority-serving institutions. The PI supervised two undergraduate students (one minority) under this program (2003 and 2005). The PI also mentored six students supported under the NSF Research Experience for Undergraduates program (REU; #0353957, PI: Valerie Taylor) in 2004, 2005, 2006, 2007, 2008, and 2009, and two students (one minority, both female) under the CRA-Women Distributed Mentor Project (DMP) for undergraduate women in computing research. Four of the above students applied to and were accepted by top graduate programs in the US. The PI will continue mentoring undergraduate students under NSF REU, TAMU USRG, and CRA-Women DMP. Undergraduate students will be given specific tasks, typi-

cally one or two small items from the "open research issues" list at the end of each section in the research plan: for example, response distribution analysis that can readily leverage existing data resulting from the main project. The methodology and tools developed as part of the proposed project will also be organized into a toolkit that will allow summer research students to have a hands-on experience in cutting-edge research in robust intelligence.

K-12 outreach: The PI has been participating in departmental events where the research labs are showcased to high school students, teachers, and parents. Since 2003, our lab hosted eight lab tours to high school student groups and other interested parties. The PI has also been mentoring high school students at the local ISD, with concrete activities such as data analysis, programming tutorial, and crash course in computational geometry. The PI will continue this in a more proactive way: teaching in the "Hour of Code" events (organized by CS Ed Week) currently on-going at the local ISD.

Multidisciplinary scientific meeting: Our project provides a convincing link between texture research in computer vision and tactile mechanism research in neuroscience. The two communities have not been in contact, and we propose to organize a workshop to facilitate this much needed interaction. The PI is experienced in organizing meetings to facilitate discussion among neuroscientists and engineers. In 2008, the PI organized a mini-symposium at the Society for Neuroscience meeting, on High-Throughput Microscopy and Computational/Theoretical Challenges in the Analysis of Neural Circuit Structure. In 2009, the PI organized a panel session on Functional Principles Underlying Biological Intelligence at the IEEE CIMSVP conference. The workshop will be of a similar format.

Software and data dissemination: All software and data (especially texture data sets) resulting from this project will be publicly released under the GNU Public License and/or BSD license. We will set up a dedicated web page for this purpose. The PI will leverage on his experience with the Topographica neural simulator project web site, the book web site for Miikkulainen et al. (2005), and the KESM Brain Atlas web site.

In summary, the following concrete activities are proposed as part of this project: (1) Graduate and undergraduate training (REU supplement, underrepresented group emphasis), (2) High school student mentoring and teaching in the "Hour of Code" events. (3) Interdisciplinary workshop on texture and tactile research. (4) Open-source software and texture data dissemination.

6 Management Plan

The timeline for the project is summarized in the table below. Note that year 1 (Y1) starts in Fall 2014 and ends in Summer 2015, etc. Request for REU supplement is embedded in this proposal for year 1 (summer 2015). Dissemination activities will take place during the summer. REU training and high school student mentoring will take place during summer 2015 and subsequent summer semesters. Workshop will be organized in fall 2016 (Society for Neuroscience satellite event) or summer 2017 (International Joint Conference on Neural Networks).

Table 1: Tasks and Timeline

Task	Y1			Y2			Y3		
4.1. TRF/VRF in Texture Segmentation	▬	▬	▬						
4.2. Cortical Development of TRF/VRF			▬	▬	▬				
4.3. Separability of TRF/VRF Representations				▬	▬	▬			
4.4. Analysis of TRF/VRF Response Distributions						▬	▬		
4.5. Motor-Based Control of Information Through TRF							▬	▬	▬
5. Dissemination			▬			▬			▬

REFERENCES CITED

Bai, Y. H., Park, C., and Choe, Y. (2008). Relative advantage of touch over vision in the exploration of texture. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008)*, 1–4, 10.1109/ICPR.2008.4760961. **Best Scientific Paper Award**.

Beck, J. (1966). Effect of orientation and of shape similarity on grouping. *Perception and Psychophysics*, 1:300–302.

Beck, J. (1983). Textural segmentation, second-order statistics and textural elements. *Biological Cybernetics*, 48:125–130.

Bell, A. J., and Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, 37:3327.

Bensimaia, S., Eenchev, P., Maass III, J. D., Craig, J., and Hsiao, S. (2008). The represenetation of stimulus orientaion in the early stages of somatosensory processing. *Neuroscience*, 28:776–786.

Bolz, J., and Gilbert, C. D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature*, 320:362–364.

Bovik, A., Clark, M., and Geisler, W. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):55–73.

Brodatz, P. (1966). *Textures: A Photographic Album for Artists and Designer.* New York, NY: Dover Publication.

Chan, A., and Vasconcelos, N. (2006). Layered dynamic textures. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems 18*, 203–210. Cambridge, MA: MIT Press.

Choe, Y. (2001). *Perceptual Grouping in a Self-Organizing Map of Spiking Neurons*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX. Technical Report AI01-292.

Choe, Y., Mayerich, D., Kwon, J., Miller, D. E., Chung, J. R., Sung, C., Keyser, J., and Abbott, L. C. (2011a). Knife-edge scanning microscopy for connectomics research. In *Proceedings of the International Joint Conference on Neural Networks*, 2258–2265. Piscataway, NJ: IEEE Press.

Choe, Y., Mayerich, D., Kwon, J., Miller, D. E., Sung, C., Chung, J. R., Huffman, T., Keyser, J., and Abbott, L. C. (2011b). Specimen preparation, imaging, and analysis protocols for knife-edge scanning microscopy. *Journal of Visualized Experiments*, 58:e3248. doi: 10.3791/3248.

Choe, Y., and Miikkulainen, R. (2004). Contour integration and segmentation in a self-organizing map of spiking neurons. *Biological Cybernetics*, 90:75–88.

Choe, Y., and Smith, N. H. (2006). Motion-based autonomous grounding: Inferring external world properties from internal sensory states alone. In Gil, Y., and Mooney, R., editors, *Proceedings of the 21st National Conference on Artificial Intelligence(AAAI 2006)*, 936–941.

Choe, Y., Yang, H.-F., and Eng, D. C.-Y. (2007). Autonomous learning of the semantics of internal sensory states based on motor exploration. *International Journal of Humanoid Robotics*, 4:211–243.

Choi, H., and Choi, S. (2004). Kernel Isomap. *Electronics Letters*, 40:1612–1613.

Choi, H., and Choi, S. (2007). Robust kernel Isomap. *Pattern Recognition*, 40(3):853–862.

Choi, H., Choi, S., and Choe, Y. (2008). Manifold integration with markov random walks. In *Proceedings of the 23rd National Conference on Artificial Intelligence(AAAI 2008)*, 424–429.

Choi, J. (2013). *Knife-Edge Scanning Microscope Mouse Brain Atlas in Vector Graphics for Enhanced Performance*. Master's thesis, Department of Computer Science and Engineering, Texas A&M University.

Chung, J. R., Sung, C., Mayerich, D., Kwon, J., Miller, D. E., Huffman, T., Abbott, L. C., Keyser, J., and Choe, Y. (2011). Multiscale exploration of mouse brain microstructures using the knife-edge scanning microscope brain atlas. *Frontiers in Neuroinformatics*, 5:29.

Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Visual Research*, 20:847–856.

DeAngelis, G. C., Ohzawa, L., and Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. 1. general characteristics and postnatal developement. *Neurophysiology*, 69:109–117.

Deibert, E., Kraut, M., Kremen, S., and Hart, J. (1999). Neural pathways in tactile object recognition. *Neurology*, 52:1413–1417.

DiCarlo, J. J., and Johnson, K. O. (2000). Spatial and temporal structure of receptive fields in primate somatosensory area 3b: Effects of stimulus scanning direction and orientation. *Neuroscience*, 20:495–510.

Efros, A. A., and Leung, T. K. (1999). Texture Synthesis by Non-parametric Sampling. In *IEEE International Conference on Computer Vision*, 1033–1038. Corfu, Greece.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4:2379–2394.

Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation*, 6:559–601.

Fogel, I., and Sagi, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics*, 61:102–113.

Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge Co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41:711–724.

Gilbert, C. D. (1994). Circuitry, architecture and functional dynamics of visual cortex. In Bock, G. R., and Goode, J. A., editors, *Higher-Order Processing in the Visual System (Ciba Foundation Symposium 184)*, 35–62. Chichester: Wiley.

Gilbert, C. D., and Wiesel, T. N. (1979). Morphology and intracortical projections of functionally identified neurons in cat visual cortex. *Nature*, 280:120–125.

Grant, R. A., Mitchinson, B., Fox, C. W., and Prescott, T. J. (2009). Active touch sensing in the rat: Anticipatory and regulatory control of whisker movements during surface exploration. *Jounral of Neurophysiology*, 101:862–874.

He, Z. J., and Nakayama, K. (1994). Perceiving textures: Beyond filtering. *Vision Research*, 34:151–62.

Heess, N., Williams, C. K. I., and Hinton, G. E. (2009). Learning generative texture models with extended fields-of-experts. In *Proceedings of the British Machine Vision Conference*. In press.

Hoyer, P. O., and Hyvärinen, A. (2000). Independent component analysis applied to feature extraction from colour and stereo images. *Network: Computation in Neural Systems*, 11:191–210.

Jain, A. K., and Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24:1167–1186.

Jones, J. P., and Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Neurophysiology*, 58:1233–1258.

Julesz, B. (1965). Texture and visual perception. *Scientific American*, 212:38–48.

Julesz, B. (1986). Texton gradients: the texton theory revisited. *Biol Cybern*, 54:245–251.

Julesz, B., and Bergen, J. (1982). Texton theory of preattentive vision and texture perception. *Journal of the Optical Society of America*, 72:1756.

Kim, D. (2011). *Automatic Seedpoint Selection and Tracing of Microstructures in the Knife-Edge Scanning Microscope Mouse Brain Data Set*. Master's thesis, Department of Computer Science, Texas A&M University, College Station, Texas.

Kwon, J., Mayerich, D., and Choe, Y. (2011). Automated cropping and artifact removal for knife-edge scanning microscopy. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 1366–1369.

Lee, H.-C., and Choe, Y. (2003). Detecting salient contours using orientation energy distribution. In *Proceedings of the International Joint Conference on Neural Networks*, 206–211. IEEE.

Liu, X., and Wang, D. (2002). A spectral histogram model for texton modeling and texture discrimination. *Vision Research*, 42:2617–2634.

Malik, J., Belongie, S., Shi, J., and Leung, T. K. (1999). Textons, contours and regions: Cue integration in image segmentation. In *ICCV(2)*, 918–925.

Manjunath, B. S., and Chellapa, R. (1991). Unsupervised texture segmentation using markov random field models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13:478–482.

Mayerich, D., Kwon, J., Panchal, A., Keyser, J., and Choe, Y. (2011a). Fast cell detection in high-throughput imagery using gpu-accelerated machine learning. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 719–723.

Mayerich, D., Kwon, J., Sung, C., Abbott, L. C., Keyser, J., and Choe, Y. (2011b). Fast macro-scale transmission imaging of microvascular networks using KESM. *Biomedical Optics Express*, 2:2888–2896.

Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. Berlin: Springer. URL: http://www.computationalmaps.org. 538 pages.

Mika, S., Ratsch, G., Weston, J., Schölkopf, B., and Müller, K. (1999a). Fisher discriminant analysis with kernels. In *Proceedings of IEEE Neural Networks for Signal Processing Workshop*, 41–48.

Mika, S., Ratsch, G., Weston, J., Schölkopf, B., and Müller, K. (1999b). Fisher discriminant analysis with kernels. In *Proceedings of IEEE Neural Networks for Signal Processing Workshop*, 41–48.

Mitchinson, B., Martin, C. J., Grant, B. A., and Prescott, T. J. (2007). Feedback control in active sensing: rat exploratory whisking is modulated by environmental contact. *Proceedings of the Royal Society B: Biological Sciences*, 274:1035–1041.

Nakayama, K., He, Z. J., and Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In Kosslyn, S. M., and Osherson, D. N., editors, *An Invitation to Cognitive Science: Vol. 2 Visual Cognition*, 1–70. Cambridge, MA: MIT Press. Second edition.

Oh, S., and Choe, Y. (2007a). Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparision of 2d vs. 3d representations. *Neurocomputing*, 70:2245–2255.

Oh, S., and Choe, Y. (2007b). Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparison of 2D vs. 3D representations. *Neurocomputing*, 70:2245–2255.

Olshausen, B. A., and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609.

Orr, A. W., Helmke, B. P., Blackman, B. R., and Schwartz, M. A. (2006). Mechanisms of mechanotransduction. *Developmental Cell*, 10:11–20.

Paragios, N., and Deriche, R. (2002). Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46:223–247.

Park, C., Bai, Y. H., and Choe, Y. (2009a). Tactile or visual?: Stimulus characteristics determine receptive field type in a self-organizing map model of cortical development. In *Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Multimedia Signal and Vision Processing*, 6–13. **Best Student Paper Award**.

Park, C., Choi, H., and Choe, Y. (2009b). Self-organization of tactile receptive fields: Exploring their textural origin and their representational properties. In *Advances in Self-Organizing Maps: Proceedings of the 7th International Workshop, WSOM 2009*, 228–236. Heidelberg: Springer.

Portilla, J., and Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40:49–71.

Reinagel, P., and Laughlin, S. (2001). Editorial: Natural stimulus statistics. *Network: Computation in Neural Systems*, 12:237–240.

Reinagel, P., and Zador, A. M. (1999). Natural scene statistics at the center of gaze. *Network: Computation in Neural Systems*, 10:341–350.

Renninger, L. W., and Malik, J. (2004). When is scene identification just texture recognition? *Vision Research*, 44:2301–2311.

Sarma, S., and Choe, Y. (2006). Salience in orientation-filter response measured as suspicious coincidence in natural images. In Gil, Y., and Mooney, R., editors, *Proceedings of the 21st National Conference on Artificial Intelligence(AAAI 2006)*, 193–198.

Schölkopf, B., and Smola, A. J. (2002). *Learning with Kernels*. MIT Press.

Sharma, J., Angelucci, A., and Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, 404:841–847.

Sung, C. (2013). *Exploration, Registration, and Analysis of High-Throughput 3D Microscopy Data from the Knife-Edge Scanning Microscope*. PhD thesis, Department of Computer Science and Engineering, Texas A&M University.

Sung, C., Woo, J., Goodman, M., Huffman, T., and Choe, Y. (2013). Scalable, incremental learning with MapReduce parallelization for cell detection in high-resolution 3D microscopy data. In *Proceedings of the International Joint Conference on Neural Networks*. In press.

Tal, N., and Amedi, A. (2009). Multisensory visualtactile object related network in humans: Insights gained using a novel crossmodal adaptation approach. *Experimental Brain Research*, 198:165–182.

Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323.

Thielscher, A., and Neuman, H. (2006). A computational model to link psychophysics and cortical cell activation patterns in human texture processing. *J of Comp Neurosci*, 22:255–282.

Varma, M., and Zisserman, A. (2003). Texture classification: Are filter banks necessary? In *Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003)*, 691–698.

Varma, M., and Zisserman, A. (2005). A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62:61–81.

von Melchner, L., Pallas, S. L., and Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, 404(6780):871–876.

Wei, L.-Y., and Levoy, M. (2000). Fast texture synthesis using tree-structured vector quantization. In *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, 479–488.

Weichert, F., Landes, C., Richards, T., Dohrmann, M., Wünderlich, R., Kaczmarczyk, R., Uebing, A., Streng, A., Liese, W., Kim, Y.-J., Groh, A., Linder, R., and Wagner, M. (2005). How to meka a haptic device help touch virtual histological slides. In *Proceedings of the Marie Curie Euroconferences MuTra: Challenges of Multidimensional Translation*, electronic. `http://www.euroconferences.info/proceedings/2005_Proceedings/2005_Wagner_Mathias.pdf`.

Wilson, S. (2007). *Self-organisation can explain the mapping of angular whisker deflections in the barrel cortex*. Master's thesis, The Universiy of Edinburgh, Scotland, United Kindom.

Wilson, S. P., Mitchinson, B., Pearson, M., Bednar, J. A., and Prescott, T. J. (2009). Learning cortical representations from multiple whisker inputs. *BMC Neuroscience*, 10(Suppl):P334.

Yang, H.-F. (2011). *Reconstruction of 3D Neuronal Structures from Densely Packed Electron Microscopy Data Stacks*. PhD thesis, Department of Computer Science and Engineering, Texas A&M University.

Yang, H.-F., and Choe, Y. (2010). Electron microscopy image segmentation with estimated symmetric three-dimensional shape prior. In *Proceedings of the 6th International Symposium on Visual Computing*.

Yang, H.-F., and Choe, Y. (2011a). Ground truth estimation by maximizing topological agreements in electron microscopy data. In *Proceedings of the 7th International Symposium on Visual Computing (LNCS 6938)*, 371–380.

Yang, H.-F., and Choe, Y. (2011b). An interactive editing framework for electron microscopy image segmentation. In *Proceedings of the 7th International Symposium on Visual Computing (LNCS 6938)*, 400–409.

Zetzsche, C., and Nuding, U. (2005). Nonlinear and higher-order approaches to the encoding of natural scenes. *Network*, 16:191–221.

Zhang, J., Tan, T., and Ma, L. (2002). Invariant texture segmentation via circular Gabor filters. *16 th International Conference on Pattern Recognition (ICPR'02)*, 2.

Zhu, L., Chen, Y., Freeman, B., and Torralba, A. (2009). Nonparametric bayesian texture learning and synthesis. In Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C. K. I., and Culotta, A., editors, *Advances in Neural Information Processing Systems 22 (NIPS 2009)*, 2313–2321.

Zhu, S. C., Wu, Y., and Mumford, D. (1998). Filters, Random Fields and Maximum Entropy (FRAME): Towards a unified theory for texture modeling. *Int. J. Comput. Vision*, 27(2):107–126.