

RELATIVE ADVANTAGE OF TOUCH OVER VISION  
IN THE EXPLORATION OF TEXTURE

A Thesis

by

YOON H. BAI

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of  
MASTER OF SCIENCE

August 2008

Major Subject: Computer Engineering

RELATIVE ADVANTAGE OF TOUCH OVER VISION  
IN THE EXPLORATION OF TEXTURE

A Thesis

by

YOON HO BAI

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

|                     |                   |
|---------------------|-------------------|
| Chair of Committee, | Yoonsuck Choe     |
| Committee Members,  | Jim Ji            |
|                     | John Keyser       |
| Head of Department, | Valerie E. Taylor |

August 2008

Major Subject: Computer Engineering

## ABSTRACT

Relative Advantage of Touch over Vision  
in the Exploration of Texture. (August 2008)

Yoon H. Bai, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Yoonsuck Choe

Texture segmentation is an effortless process in scene analysis, yet its mechanisms have not been sufficiently understood. Several theories and algorithms exist for texture discrimination based on vision. These models diverge from one another in algorithmic approaches to address texture imagery using spatial elements and their statistics. Even though there are differences among these approaches, they all begin from the assumption that texture segmentation is a visual task.

However, considering that texture is basically a surface property, this assumption can at times be misleading. An interesting possibility is that since surface properties are most immediately accessible to touch, texture perception may be more intimately associated with texture than with vision (it is known that tactile input can affect vision). Coincidentally, the basic organization of the touch (somatosensory) system bears some analogy to that of the visual system. In particular, recent neurophysiological findings showed that receptive fields for touch resemble that of vision, albeit with some subtle differences.

The main novelty and contribution of this thesis is in the use of tactile receptive field responses for texture segmentation. Furthermore, we showed that touch-based representation is superior to its vision-based counterpart when used in texture boundary detection. Tactile representations were also found to be more discriminable (LDA and ANOVA). We expect our results to help better understand the nature of texture perception and build more powerful texture processing algorithms.

The results suggest that touch has an advantage over vision in texture processing. Findings in this study are expected to shed new light on the role of tactile perception of texture and its interaction with vision, and help develop more powerful, biologically inspired texture segmentation algorithms.

To my family and friends far and wide

## ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Yoonsuck Choe for providing me with the opportunity to work under his guidance. He constantly taught me how to perform high-quality research and generously provided extensive academic support throughout this research. I am also grateful to benefit from his keen scientific insights, and from his knack for solving intractable difficulties into simple terms and solutions.

I would like to thank my committee members, Dr. John Keyser and Dr. Jim Xioquan Ji for their invaluable guidance and patient teaching during the initial stages of my research. I am fortunate and grateful to learn from their profound thoughts.

I would also like to acknowledge contributions from my fellow research group members. I would like to thank Donghyub Han, for providing tremendous support in many aspects. I would like to thank Henry Choi, who provided invaluable critiques and fascinating discussions. I am also thankful for other fellow members of the Brain Networks Lab (BNL) and Neural Intelligence Lab (NIL) for their intriguing discussions and contributions.

My sincerest thanks go to my parents, Donggyu Bai (Class of '89) and Hae-soon Seo, who have been loving, caring, patient and supportive throughout my whole life. I would also like to give a huge thanks to my brother and best friend, Matthew Bai, who patiently encouraged and supported me throughout rough periods.

This research was supported in part by the National Institute of Health/National Institute of Neurological Disorders and Stroke (#R01-N554252, PI: Yoonsuck Choe), and by the Industrial Affiliates Program (IAP) in the Department of Computer Science at Texas A&M University.

## TABLE OF CONTENTS

| CHAPTER |  | Page |
|---------|--|------|
| I       | INTRODUCTION . . . . .   | 1    |
|         | A. Problem Overview . . . . .  | 2    |
|         | 1. Texture Analysis . . . . .  | 3    |
|         | 2. Related Work . . . . .  | 4    |
|         | 3. Motivation . . . . .  | 5    |
|         | B. Proposed Approach . . . . .   | 6    |
|         | C. Outline of the Thesis . . . . .   | 7    |
| II      | BACKGROUND . . . . .   | 8    |
|         | A. Feature Generation . . . . .  | 8    |
|         | B. Receptive Fields . . . . .  | 9    |
|         | 1. Simple Cells and the Gabor Filter . . . . .                               | 9    |
|         | 2. Somatosensory Area 3b Receptive Fields and the<br>Tactile Model . . . . . | 12   |
|         | C. Performance Assessment . . . . .  | 15   |
|         | 1. Classification Tools . . . . .  | 15   |
|         | a. kNN Classification . . . . .  | 15   |
|         | b. Support Vector Machine . . . . .  | 17   |
|         | 2. Data Separability . . . . .   | 19   |
| III     | METHODS . . . . .  | 21   |
|         | A. Experimental Methods . . . . .  | 21   |
|         | 1. Input Preparation . . . . .   | 21   |
|         | 2. Generating RF Responses . . . . .   | 23   |
|         | 3. Performance Evaluation . . . . .  | 25   |
| IV      | RESULTS . . . . .  | 29   |
|         | A. Data Characteristics . . . . .  | 29   |
|         | 1. Visualization from Feature Selection . . . . .                            | 30   |
|         | 2. Visualization from Feature Extraction . . . . .                           | 36   |
|         | B. Performance Evaluation . . . . .  | 39   |
| V       | DISCUSSION . . . . .   | 41   |

| CHAPTER                               | Page |
|---------------------------------------|------|
| A. Contribution . . . . .             | 41   |
| B. Limitations of the Model . . . . . | 42   |
| C. Future Work . . . . .              | 43   |
| VI CONCLUSION . . . . .               | 46   |
| REFERENCES . . . . .                  | 47   |
| VITA . . . . .                        | 54   |

## LIST OF TABLES

| TABLE |   | Page |
|-------|---|------|
| I     | Measures of data separability based on isolated parameters.. . . . .  | 38   |
| II    | kNN classification rates of both models based on isolated parameters. | 39   |
| III   | SVM classification rates of both models based on isolated parameters. | 40   |
| IV    | Summary of model performance from independent parameters. . . . .     | 44   |

## LIST OF FIGURES

| FIGURE | Page  |
|--------|---|
| 1      | Main structure of texture discrimination models. . . . . 4  |
| 2      | The major pathway that visual information goes through from the eye to the primary visual cortex is shown. Signals are produced by rod and cone receptors in the retina and are then transferred to a major relay station, the LGN (lateral geniculate nucleus) via the optic nerve. Signals then travel to selected areas of the primary visual cortex (V1). Signals are sent to higher areas of cortex from there on. Adapted from [34]. . . . . 10   |
| 3      | Representations of the visual receptive field profiles. (a) Difference-of-Gaussians (DoG) corresponds to the LGN. (b) Gabor Filter (oriented at $0^\circ$ .) corresponds to a V1 simple cell. . . . . 13  |
| 4      | Profiles of the visual and tactile receptive fields are illustrated. (a) V1 simple cell RFs (Gabor filters) showing $157.5^\circ$ , $0^\circ$ , and $22.5^\circ$ orientation preference. (b) Somatosensory 3b RFs (tactile model) in five scanning directions and its corresponding model predictions. The center figure depicts the RF before scanning. The two fixed components are shown in gray ellipses and the lagging component is illustrated as a dotted (before scanning) ellipse and white ellipse (after scanning). Following each scanning direction indicated in boxed arrows, the lagging component trails each scan. For each resulting set of three adjacent models, the first shows where the lagging component started (red, dotted ellipse) and ended (black, single lined ellipse). The second adjacent figure includes the path of lagging while the right-most figure shows the final model. The initial position and orientation of Gaussian lobes followed the physiological data provided by DiCarlo and Johnson [44]. . . . . 16 |

| FIGURE | Page   |
|--------|--|
| 5      | Optimal hyperplane of a support vector machine (SVM). Training a SVM consists of finding the optimal hyperplane, that is, the one with the maximum distance from the nearest training patterns. The support vectors are those (nearest) patterns, a distance $b$ from the hyperplane. Adapted from [51] . . . . . 18   |
| 6      | Sample textures without a texture-defined boundary. . . . . 22   |
| 7      | Sample texture sets with a texture-defined boundary located in between two textures. . . . . 22  |
| 8      | The square markers indicate the trajectory of window patches that were applied to RF models. (a) shows a scanning oriented at zeros degrees. (b) shows a scanning oriented at $45^\circ$ degrees. Each marker results in a single scalar value and the whole trail of markers eventually produces a response vector. A texture-defined boundary lies the middle of the input image. Both scans had 12 frames, resulting in a 12 element vector. . . . . 24 |
| 9      | Plots in (a) and (b) respectively depict response vectors generated the tactile and visual RFs. Both tactile and visual RFs were tuned to different orientation. Each colored vector response correspond to the orientation of the same color. . . . . 25  |
| 10     | An outline of how response vectors are generated. For each modality, response vectors from a target-present (boundary) and target-absent (no boundary) input are generated to expose effects of texture boundaries and ultimately, texture discrimination. . . . . 26  |
| 11     | Three sets of one-way ANOVA were applied to test significant differences between: (1) left-hand side homogenous texture, (2) texture boundary, (3) right-hand side homogenous texture. The $y$ -axis and $x$ -axis relatively depicts the sample index and mean values. The plot shows that the means of measurements are significantly different. Corresponding $p$ -values were under 0.01. . . . . 31   |

| FIGURE | Page  |    |
|--------|---|----|
| 12     | <p>Three sets of one-way ANOVA was applied to test significant differences between: (1) left-hand side homogenous texture <i>vs.</i> texture boundary, (2) left-hand side homogenous texture <i>vs.</i> right-hand side homogenous texture, and (3) texture boundary <i>vs.</i> right-hand side homogenous texture. The <math>y</math>-axis and <math>x</math>-axis relatively depicts the sample index and mean values. The plot indicate that the means of measurements are not significantly different. Corresponding <math>p</math>-values were over 0.5. . . . .</p>   | 32 |
| 13     | <p>(a) A sample of visual feature vectors from various scans are superimposed on top of an example input image with a texture boundary in the middle. (b) A sample of tactile feature vectors from various scans are superimposed onto an example input image with a texture boundary in the middle. The response features shown here were generated from various images in the sample set and not from the single example input image. The background image was only used to depict where the texture boundary lies throughout the images in the sample set. Note that tactile feature vectors respond abruptly across the texture-defined boundary while visual feature vectors are less sensitive to texture boundaries. This figure implies that touch-based models are more correlated to texture boundaries than vision-based models. . . . .</p> | 34 |
| 14     | <p>(a) A scatter plot of 1-D projections (from LDA) of the tactile feature space and its probability density estimation is illustrated. For the scatter plot, the <math>x</math>-axis depicts the sample index of each texture response vector, and the <math>y</math>-axis indicates the linear discriminant (values of the 1-D projection values). In each plot, magenta(<math>\bullet</math>) and black(<math>\times</math>) represents boundary and non-boundary features, respectively. (b) Equivalent plots of the visual feature space. The projected feature space (LDA scatter plots) and corresponding density estimations indicate that the tactile feature space contains more class separability for texture discrimination. . . . .</p>   | 35 |
| 15     | <p>Overall PCA scatter plots are listed for each modality. The first two eigen-vectors were used to visualize the initial feature space. (a) 2-D scatter plot of the visual feature space. (b) 2-D scatter plot of the tactile feature space. . . . .</p>   | 37 |

## CHAPTER I

## INTRODUCTION

When observing our environment, scenes appear to be seamless. Generally, abundant objects and spatial events are easily and rapidly perceived. Although this seems effortless, images that are clearer and more detailed than any digitized screen reaching the visual apparatus are continuously updated and perceived so rapidly that we do not suffer from any discontinuities or breaks in vision. The efficiency and completeness of visual perception is unparalleled in comparison with any piece of apparatus or instrumentation ever invented. This remarkable achievement of the visual system begins with visual scene analysis which involves basic perceptual processes such as segmentation of a visual scene into a set of coherent patterns and recognizing objects [1].

In particular, visual perception starts from segregation of scenes based on cues related to luminance, color, contours, and texture of object surfaces. However, in nature, images or scenes that reach our eyes are nearly a conglomerate of geometric elements passed on through a medium of light. This physically means that any observable object is merely a reflection of light and the semantics of that object—if it is indeed an object—is only determined by the agent. Only when we convey this information from the sensory world and distinguish a region of interest through segmentation can we apprehend the context of the visual scene. In the same sense, it has been found that the human visual system uses texture information to automatically—or pre-attentively—segregate parts of the natural scene [2].

This biological inspiration motivated numerous approaches to achieve texture

---

The journal model is *IEEE Transactions on Automatic Control*.

segmentation in computer vision and image analysis [3][4][5][6][7]. These models diverge from one another in algorithmic approaches to address texture imagery using spatial elements and their statistics. Even though there are differences among these approaches, they all begin from the assumption that texture segmentation is a visual task. However, considering that texture is essentially a surface property, this assumption can at times be misleading [8][9]. An interesting possibility is that since surface properties are most immediately accessible to touch, tactile perception may be more intimately associated with texture than vision.

#### A. Problem Overview

The main argument of this thesis is triggered from the fact that although texture is in the first place a surface property, texture analysis and segmentation has been tacitly assumed to be a computational vision task[10]. Instead, this thesis relies on a touch-based biological model that is congruent with the essence of texture, since it is most immediately susceptible to surface characteristics. As an alternative to a vision-based approach, the main expectation of this thesis is that touch-based features will be more ecologically suited for textures. More specifically, the main hypothesis of this thesis is that touch-based features provide more discriminative information than those of the vision-based approaches in exploring textures. Subsequently, once this statement is confirmed, it is expected to provide a more powerful foundation for further texture analysis.

To aid the problem addressed in this thesis, related concepts and literature review of texture analysis, and motivations of the problem statement are provided. Afterwards, a proposed approach of this thesis and a brief outline for the remaining chapters are presented.

## 1. Texture Analysis

Texture perception is one of the pillars in the study of early visual perception. Much of the psychophysical work concentrates on texture discrimination, or detecting whether two texture patches can be discerned rapidly by human observers [6][11][12]. Based on neurophysiological studies, both visual and tactile modalities are known to detect texture boundaries in early texture perception stages [2][13]. Moreover, latter studies have shown that human texture processing is likely to start from detecting texture boundaries at relatively early stages of cortical processing [14][15][16][17].

The subsequent stage in generic texture segmentation partitions the image into regions that are homogenous with respect to one or more features. A basic issue to be considered is that of cell unit size, i.e. the resolution of the area measured in order to test for homogeneity.

Regarding homogeneity with respect to certain characteristics is related to the study of perceptual grouping in the 1920s that led to the formulating of what came to be called the Gestalt laws. This field of study suggests individual tokens appear to group according to a set of principles [18]. These principles include proximity, similarity, continuation, symmetry, and closure. The tendency of seeing elements with similar characteristics as belonging to an approximately homogenous group suggests that the same mechanism might be at work for Gestalt and texture segmentation. In the context of image segmentation, the ideal goal is to partition the image into different objects and regions [10]. But if texture is what constitutes distinctive regions, descriptive perception becomes even more ambiguous. Considering a brick wall, should the proper partition be individual bricks or the entire wall as a whole? Such obscure interpretations led to widely diverse approaches in texture segmentation. However, perceptual grouping is beyond the scope of this thesis and the primary

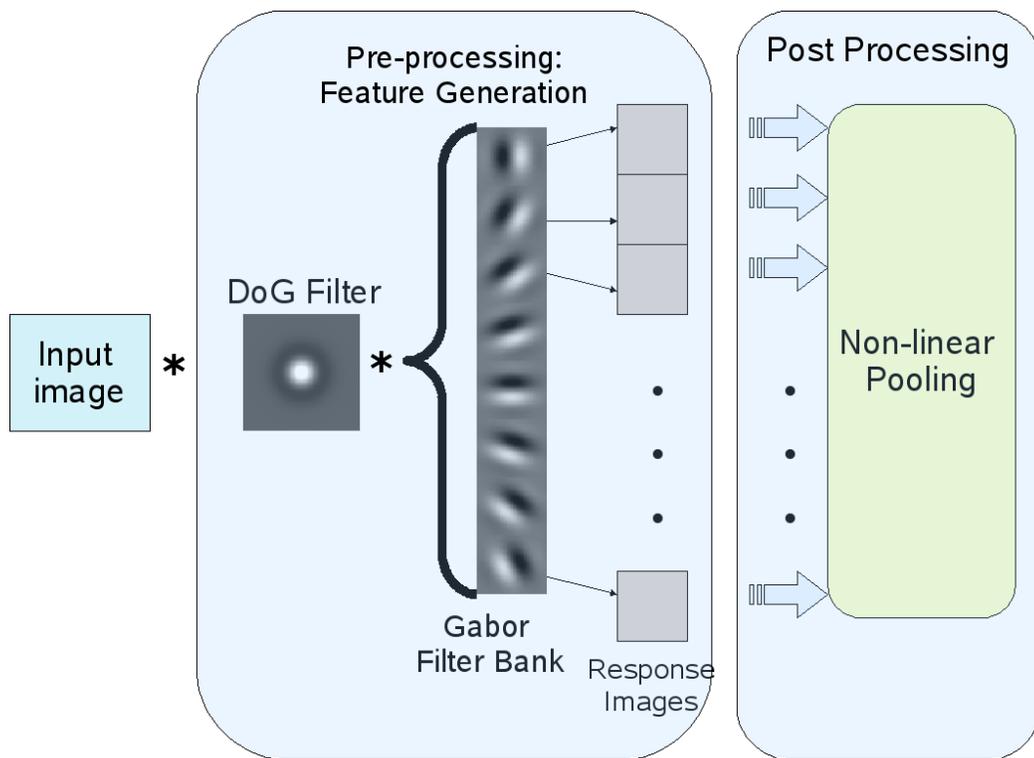


Fig. 1. Main structure of texture discrimination models.

focus will be related to the primitive stage that provides a foundation for grouping occurring during post-processing in texture analysis.

## 2. Related Work

A pioneering model in texture perception was provided by Beck [5]. According to his model, the first stage performs local feature detection with receptive fields in the visual system. The second stage extracts the total differences in color, luminance, orientation, and size between neighboring texture elements. The last stage segregates an image into regions of the same texture on the basis of the magnitude and distribution of response vectors.

Another pivotal approach by Julesz and his colleagues incorporated a conceptual

claim that *top-down* basic elements called textons define the underlying criterion for texture processing[3][4]. According to their theory, textures are discriminated if they differ in the density of certain simple, local textural features, or textons. Textons are often considered as elementary blobs defined by color, orientation, size, line gratings, etc. Although textons were theorized by Julesz as perceptual atoms, it has been criticized that this approach is based on a verbal description of image features [19].

More recently, texture discrimination has received considerable attention from the feature-based computational perspective, and many models have been proposed. Some noticeable studies are from Caelli [20], Turner [21], Voorhees and Poggio [22], Fogel and Sagi [23], Malik and Perona [24], Graham, Beck and Sutter [25], and Liu and Wang [26]. Although these models differ in details, they all share a common two-fold structure when discriminating textures (Figure 1). Typically, an image is first filtered with a set of Gabor filters of different preferred orientations and spatial frequencies that cover appropriately the spatial frequency domain, and the features obtained from a feature vector field is further used for analysis, classification, or segmentation. The latter stage is usually referred to “post-Gabor” or “post-processing”. Most studies initiate models from the feature-space from Gabor filters and usually differ from one another in later post-processing stage.

### 3. Motivation

Although texture analysis has been focused in computer vision literature, the essence of texture comes from the sensation of surface characteristics. Therefore, a set of problems encountered within the context of texture analysis, i.e. texture classification, description, and segmentation, can be simulated by touch-based feature generation. Coincidentally, the basic organization of the tactile (somatosensory) system bears analogy to that of the visual system [27], and it is known that tactile input can

affect vision [28]. Particularly, recent neurophysiological findings showed that receptive fields for touch resemble that of vision, albeit with some subtle differences [29]. These psychophysiological studies call for further investigation to explore how effective a touch-based system is in texture analysis relative to the mainstream approach inspired from the human visual system.

## B. Proposed Approach

The above discussion leads to an interesting and intuitive expectation that tactile perception might have higher capacity related to surface characteristics than visual texture perception in order to capture attention from abrupt sensations. This biological inspiration motivated a new feature space generated by tactile RFs instead of using the orthodox feature generation by visual receptive fields (RFs), or Gabor filters.

Having a new feature established, the next question that arises is to evaluate this new foundation and compare it to the orthodox visual representation. Likewise, based on previous studies in texture discrimination, it is meaningful to evaluate the effect and usefulness of features for post-processing models. Most studies use classification result comparison for feature comparisons [30]. However, we should keep in mind that this characterizes the joint performance of the pre-processed feature space and its subsequent post-processing stages as well as its subsequent classifier. In other words, widely used classification performance comparisons do not sufficiently address the fundamental representative power of the feature space. In order to overcome this issue, further statistical measurements will be incorporated to measure the data separability as well as observing the immediate classification rates of feature vectors. As a summary, this thesis presents a new feature space motivated by tactile perception

and conducts additional analyses of the model as well as the usefulness of the feature space in terms of data separability and data discrimination capability. These findings are expected to shed new light on the role of tactile perception of texture and its interaction with vision, and help develop more powerful, biologically inspired texture segmentation algorithms.

### C. Outline of the Thesis

This thesis is organized as follows. In the following Chapter, I will provide a background overview including biological inspirations and their implementation methods. In Chapter III, experiments in generating texture features are described as well as methods to evaluate characteristics of the tactile model. Next, in Chapter IV, results based on statistical characteristics and classification performance are presented. Finally, discussions and conclusions about the results will be presented in Chapter V.

## CHAPTER II

### BACKGROUND

This chapter describes the properties of biological models that inspired previous texture perception models. First, the mainstream inspirations based on the early visual system is depicted followed by touch-based perception studies. Afterwards, a brief overview of statistical and computational methods provides the foundation of measuring and evaluating texture discrimination within the corresponding feature space.

#### A. Feature Generation

As mentioned previously, when we see, we are not interpreting the pattern of light intensity that falls on our retina; we are interpreting the pattern of spikes that the millions of ganglion cells sent through our optic nerve to the brain. Likewise, when we hear, we are not interpreting the patterns of amplitude and frequency modulation that characterize the acoustic waveform; we are interpreting the patterns of spikes from roughly thirty thousand auditory nerve fibers [31]. All of the tasks our brain perform relies on the incoming sensory signals that are encoded as neural codes, or sequences of neural spikes.

If neural code is the language of our brain, quantified signals or vector responses from a *feature generator* is the colloquial equivalent in computer vision literature. In the context of signal processing and signal detection theory, imagery is ultimately defined by means of these features. Based on psychological and neurophysiological data, it is widely accepted that the human visual system transforms a retinal image into a local spatial/frequency representation. This representation is generated by receptive fields (RF) such as in simple, complex, and hypercomplex cells as identified by Hubel and Wiesel [32]. Such representations can be simulated by a bank of filters with tuned

frequencies and orientations—such as Gabor filters—and they find applications in many areas of visual scene analysis.

Typically, an image is filtered with a filter bank consisting of different spatial factors that appropriately covers the general spatial and spatial frequency domain. This feature (response vectors) domain generated from filters is referred to as the feature space that is further used for analysis, classification, or segmentation.

## B. Receptive Fields

Image decomposition is usually achieved by filtering the input image with Receptive Field (RF) profiles and the generated feature vectors at this stage correspond to neural encoding of sensory signals. The following two sections introduce the two RF profiles of interest in this thesis: early visual RF (simple cells) and tactile RF (somatosensory 3b).

### 1. Simple Cells and the Gabor Filter

Computational models inspired from the early human visual system start by simulating the visual pathway (Figure 2). An image is generated by photoreceptors in the retina, a layer of cells at the back of the eye. The information leaves the eye by way of the optic nerve, and there is a partial crossing of axons at the optic chiasm. After the optic chiasm, the axons form the optic tract. The optic tract wraps around the midbrain to get to the lateral geniculate nucleus (LGN), where all the axons must synapse. From there, the LGN axons fan out through the deep white matter of the brain as the optic radiations, which will ultimately travel to the primary visual cortex, at the back of the brain [33].

The computational model that corresponds to the mechanism up to the primary

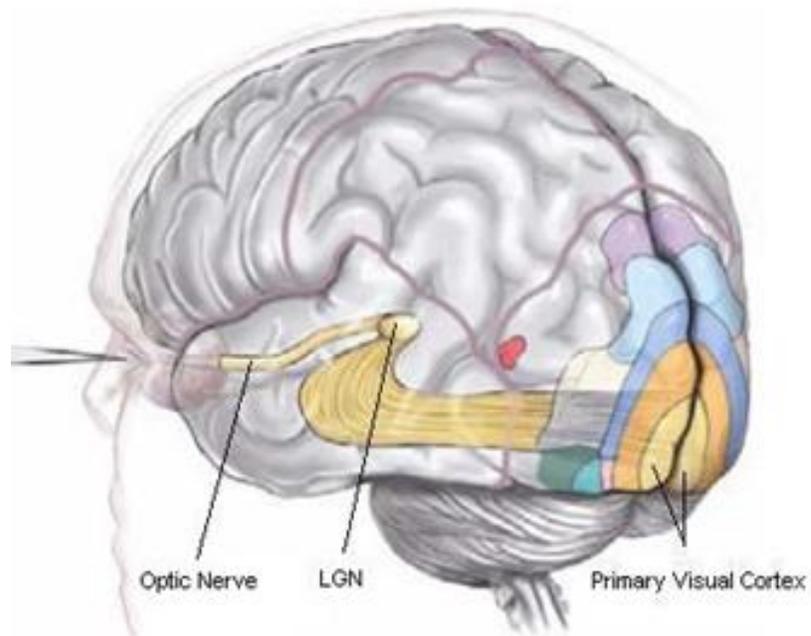


Fig. 2. The major pathway that visual information goes through from the eye to the primary visual cortex is shown. Signals are produced by rod and cone receptors in the retina and are then transferred to a major relay station, the LGN (lateral geniculate nucleus) via the optic nerve. Signals then travel to selected areas of the primary visual cortex (V1). Signals are sent to higher areas of cortex from there on. Adapted from [34].

visual cortex is represented with the difference-of-Gaussians (DoG) filters (Figure 3-(a)). Kuffler discovered the antagonistic center-surround organization of the receptive fields of cat retinal ganglion cells [35]. This model for the antagonistic center-surround organization of both retinal ganglion cells and LGN cells has been popularized by Marr, who incorporated the DoG representation into his theory of early visual processing [36].

The subsequent stage is where filters corresponding to RFs generate feature vectors in computational models. Hubel and Wiesel defined simple cells in the visual cortex as early relay stations, orientation sensitive, where the RF is made up of *on* (excitatory) and *off* (inhibitory) regions [32]. They then showed that a simple cell's response to bars and edges could be predicted from its corresponding receptive fields [37]. This profile is illustrated in Figure 3-(b).

Jones and Palmer, Porat and Zeevi, Clark and Bovik worked on the Gabor representation for early vision [38][39][40]. A distributed architecture, made up of multiple spatially and spectrally localized RFs and defined as Gabor filters, yields an early low-level representation of the visual input, namely the V1 simple cell. Typically, when generating Gabor features, an input image  $I(x, y)$ ,  $(x, y) \in \Phi$  ( $\Phi$  is the set of pixel locations) is convolved with a 2-D Gabor function  $g(x, y)$  that can be represented as,

$$g_{\gamma, \theta, \phi} = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \cos 2\pi \frac{x'}{\lambda} + \phi, \quad (2.1)$$

$$x' = x \cos \theta + y \sin \theta,$$

$$y' = -x \sin \theta + y \cos \theta,$$

where  $\theta$  is the orientation preference that is normal to the parallel excitatory and

inhibitory ellipses.  $\phi$  is the symmetry phase,  $\sigma$  is the standard deviation (width) of the Gaussian,  $\lambda$  is the wavelength, and  $\gamma$  is the aspect ratio. In experiments listed in this thesis, these parameters were set to  $\sigma = 0.56\lambda$ ,  $\gamma = 1.0$ , and  $\phi = 0.5\pi$ . These free parameters are defines the overall shape of the Gabor filter (RF profile). The specific values were adapted from widely accepted literature. Moreover, through numerous experiments, these set of parameters produced the most data separability which is preferable to show any effectiveness of the visual RF profile.

Regarding the convolution operation, we should note that this is equivalent to filtering and that a multichannel, or filter bank, approach results in multichannel filters. Specifically, a bank of Gabor filters with eight equidistant preferred orientations ( $\theta = k\frac{\pi}{8}, k = 0, 1, \dots, 7$ ) was constructed.

## 2. Somatosensory Area 3b Receptive Fields and the Tactile Model

The counterpart of the visual (V1 simple cell) receptive field (RF) model is the tactile (somatosensory area 3b) model. Using this tactile RF model for any texture segmentation of computer vision related algorithms is unprecedented, but the main motivation to incorporate this modality was based on numerous neurophysiological studies related to texture perception. Although the two modalities are different, the underlying mechanisms share remarkably similar attributes. Jiang et al. suggested that texture responses from the somatosensory cells reflect a form of feature extraction, signaling a presence of texture change as well as showing similar sequential texture information processing as observed in the visual system [13]. Another study by Fitzgerald and Hsiao suggests that orientation features, similarly to the visual system, are a possible mean that enables texture discrimination. Alternative studies indicate that for the most part, each level in the visual system can be matched to a corresponding level in the somatosensory system [41][42][43]. These statements al-

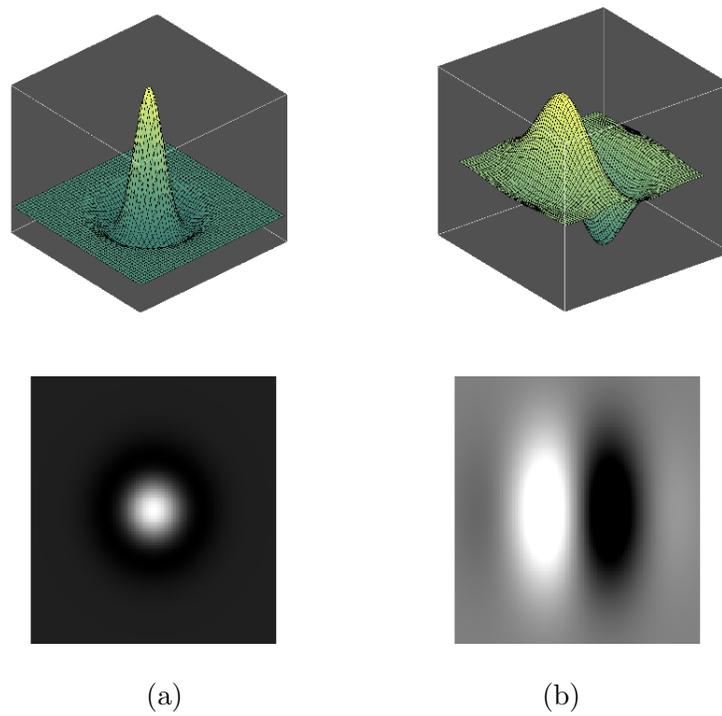


Fig. 3. Representations of the visual receptive field profiles. (a) Difference-of-Gaussians (DoG) corresponds to the LGN. (b) Gabor Filter (oriented at  $0^\circ$ .) corresponds to a V1 simple cell.

low direct usage of the somatosensory RF in lieu of visual RF when defining texture features.

DiCarlo and Johnson derived the tactile RF model by recording area 3b neural responses to dot patterns using reverse correlation. The main structure of the RFs consists of three Gaussian subfields: a central excitatory region accompanied by an inhibitory lobe and a temporally, dynamically lagging inhibitory lobe with respect to the excitation center [44]. Each subfield can be expressed as

$$G(x, y) = ae^{-\frac{1}{2}(L^T S^{-1} L)} \quad (2.2)$$

$$L = \begin{bmatrix} x - \mu_x - \mathbf{v}_x \tau \\ y - \mu_y - \mathbf{v}_y \tau \end{bmatrix}, S = \begin{bmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^w \end{bmatrix}$$

where  $(\mu_x, \mu_y)$  represents the center of the subfield,  $(\mathbf{v}_x, \mathbf{v}_y)$  represents the stimulus velocity vector, and  $\tau$  represents the delay of the peak excitation or inhibition with respect to skin stimulation. The center of excitation was fixed to stay at the middle of all tactile models while the complementary inhibition and lagging inhibition centers varied with respect to the excitatory center. The parameters  $a$ ,  $\sigma_x$ ,  $\sigma_y$ , and  $\rho$  specify the amplitude, spread, orientation, and elongation (aspect ratio) of the excitatory ( $a > 0$ ) or inhibitory ( $a < 0$ ) component represented by the Gaussian function. Finally, the three Gaussian subfields are linearly summed to represent the tactile model. In Figure 4, the outline in the middle depicts the initial RF before scanning. The arrows represent scanning directions of the fingertip. From each scan, the resulting RF is illustrated through three diagrams: (1) The excitatory and fixed inhibitory lobe are outlined in gray ellipses and the lagging component is illustrated as dotted (before scanning) and black (after scanning) ellipses; (2) the lagging inhibitory lobe

is displaced in the opposite direction of the scan; (3) and the linear summation of arrows listed as fixed orientation components.

Since the Gabor filter describes an excitatory mass flanked by an inhibitory one, the Gabor model is analogous to the tactile model, but with fewer degrees of freedom [45]. However, the uniqueness of the tactile filter relies on this extra degree of freedom of the lagging component that affects the level of occlusion of the excitatory lobe that eventually leads to the selectivity of certain spatial features such as orientation.

### C. Performance Assessment

#### 1. Classification Tools

Classifying data is a fundamental problem in machine learning. Given different categories or classes, the goal is to decide which category a new data point belongs to. Classification methods, as well as regression, rely on previous data sets which make it supervised learning. Among major classification methods, Fix and Hodges introduced a non-parametric classification method that was later named as the  $k$ -nearest neighbor (kNN) algorithm [46]. This is widely used when there is little or no prior knowledge about the distribution of the data due to the robustness in terms of Bayes error rate [47][48]. Another widely used algorithm in classification applications is the support vector machine (SVM). This method is attractive in data classification because it achieves maximum separation between two classes [49].

##### a. kNN Classification

The *k-Nearest Neighbor* (kNN) classification is a very intuitive method that classifies unlabeled examples based on their similarity to examples in the training set. Classification typically involves partitioning samples into training and testing categories.

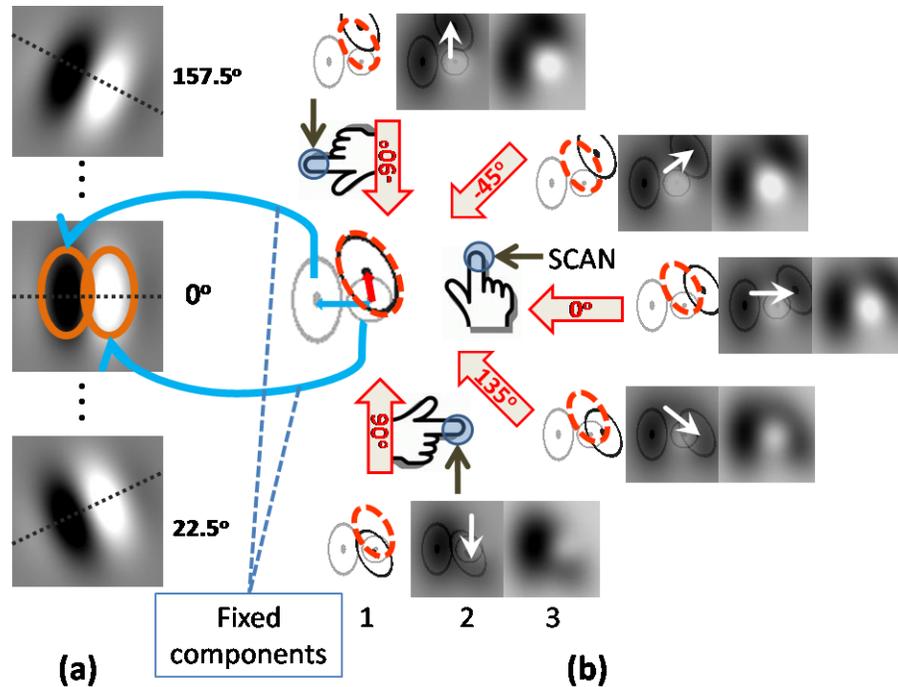


Fig. 4. Profiles of the visual and tactile receptive fields are illustrated. (a) V1 simple cell RFs (Gabor filters) showing  $157.5^\circ$ ,  $0^\circ$ , and  $22.5^\circ$  orientation preference. (b) Somatosensory 3b RFs (tactile model) in five scanning directions and its corresponding model predictions. The center figure depicts the RF before scanning. The two fixed components are shown in gray ellipses and the lagging component is illustrated as a dotted (before scanning) ellipse and white ellipse (after scanning). Following each scanning direction indicated in boxed arrows, the lagging component trails each scan. For each resulting set of three adjacent models, the first shows where the lagging component started (red, dotted ellipse) and ended (black, single lined ellipse). The second adjacent figure includes the path of lagging while the right-most figure shows the final model. The initial position and orientation of Gaussian lobes followed the physiological data provided by DiCarlo and Johnson [44].

Let  $\mathbf{x}_i$  be a training sample and  $\mathbf{x}$  be a test sample, and let  $\omega$  be the true class of a training sample and  $\hat{\omega}$  be the predicted class for a test sample ( $\omega, \hat{\omega} = 1, 2, \dots, N$ ). Here,  $N$  is the total number of classes.

During the training process, we use only the true class  $\omega$  of each training sample to train the classifier, while during testing we predict the class  $\hat{\omega}$  of each test sample. It warrants noting that kNN is a "supervised" classification method in that it uses the class labels of the training data. Unsupervised classification methods, or "clustering" methods, on the other hand, do not employ the class labels of the training data.

With 1-nearest neighbor rule, the predicted class of test sample  $\mathbf{x}$  is set equal to the true class  $\omega$  of its nearest neighbor, where  $\mathbf{m}_i$  is a nearest neighbor to  $\mathbf{x}$  if the distance

$$d(\mathbf{m}_i, \mathbf{x}) = \min_j \{d(\mathbf{m}_j, \mathbf{x})\}. \quad (2.3)$$

For k-nearest neighbors, the predicted class of test sample  $\mathbf{x}$  is set equal to the most frequent true class among k nearest training samples. This forms the decision rule,  $D : \mathbf{x} \rightarrow \hat{\omega}$ .

## b. Support Vector Machine

*Support vector machines* (SVMs) classify data from two categories by an optimal hyperplane. SVM is motivated by mapping the data to a sufficiently high dimension—typically much higher than the original feature space—in order to find the optimal hyperplane that separates the two class data [50].

Suppose we have a set of  $N$   $d$ -dimensional samples  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ . We assume each pattern  $\mathbf{x}_k$  has been transformed to  $\mathbf{y}_k = \varphi(\mathbf{x}_k)$ , where  $\varphi(\cdot)$  is a nonlinear mapping operator. For each of the  $N$  patterns,  $k = 1, 2, \dots, N$ , we let  $z_k = \pm 1$ , according to whether pattern  $k$  is in  $\omega_1$  or  $\omega_2$ . A linear discriminant in an augmented  $\mathbf{y}$  space is

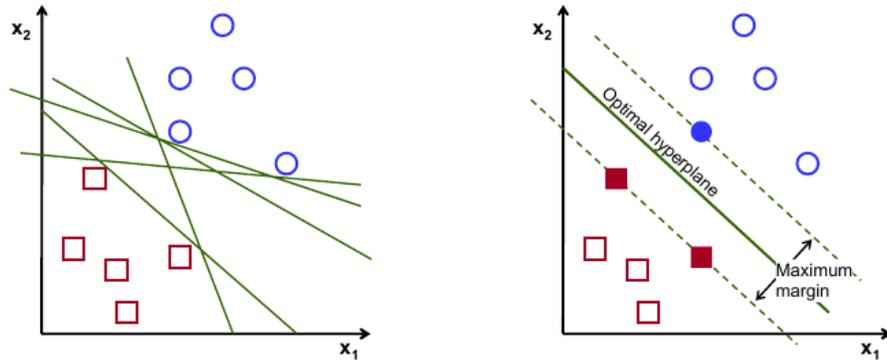


Fig. 5. Optimal hyperplane of a support vector machine (SVM). Training a SVM consists of finding the optimal hyperplane, that is, the one with the maximum distance from the nearest training patterns. The support vectors are those (nearest) patterns, a distance  $b$  from the hyperplane. Adapted from [51]

$$g(\mathbf{y}) = \mathbf{a}^t \mathbf{y} \quad (2.4)$$

where both the weight vector and the transformed pattern vector are augmented. Thus a separating hyperplane ensures

$$z_k g(\mathbf{y}_k) \geq 1, k = 1, \dots, N \quad (2.5)$$

as shown in Figure 5.

The goal in training a Support Vector Machine is to find the separating hyperplane with the largest margin. This is because we can obtain a better generalization of the classifier from a larger the margin. From algebraic derivations of the geometric distance from a hyperplane to a (transformed) pattern  $\mathbf{y}$  is  $\frac{|g(\mathbf{y})|}{\|\mathbf{a}\|}$ , and assuming that a positive margin  $b$  exists, equation 2.5 implies,

$$\frac{z_k g(\mathbf{y}_k)}{\|\mathbf{a}\|} \geq b, k = 1, \dots, N \quad (2.6)$$

and from the above equation, we can reach our goal by finding the weight vector  $\mathbf{a}$  that maximizes  $b$ .

## 2. Data Separability

The feature vectors computed in different points of a texture image are not identical; they rather form a cluster in the multidimensional feature space. The larger the distance between two clusters that correspond to two different types of texture, the better the discrimination properties of the feature generator that defines a texture.

In order to determine the distance between two clusters of feature vectors, we can observe their projections onto a 1-D space, i.e., a line, under the assumption that this projection maximizes the separability of the clusters in the 1-D space. A linear transform that, under certain conditions, realizes such a projection was first introduced by Fisher [52] and later called the *Fisher linear discriminant function* and has the following form:

$$y = (\vec{\mu}_1 - \vec{\mu}_2)^T S^{-1} \vec{x} \quad (2.7)$$

where  $\vec{\mu}_1$  and  $\vec{\mu}_2$  are the means of the two clusters,  $S^{-1}$  is the inverse of the pooled covariance matrix of the two clusters,  $\vec{x}$  is a feature vector, and  $y$  is its 1-D projection.

This projection of the feature vectors into the 1-D space maximizes the *Fisher criterion* [53], which measures the separability of the two concerned clusters in the reduced space:

$$J(W) = \frac{(\tilde{m}_1 - \tilde{m}_2)^2}{(\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)} = \frac{W^T S_B W}{(W^T S_W W)}, \quad (2.8)$$

$W = \text{projection vector},$  (1)

$S_B = \text{between class matrix},$

$S_W = \text{within class matrix}$

where  $\tilde{\sigma}_1^2$  and  $\tilde{\sigma}_2^2$  are the standard deviations of the distributions of the projected feature vectors of the two clusters:  $\tilde{\sigma}_1^2 = W^T \Sigma_1 W$ ; and  $\tilde{m}_1$  and  $\tilde{m}_2$  are the projections of the means  $\vec{\mu}_1$  and  $\vec{\mu}_2$ , respectively ( $\tilde{m}_i = W^T m_i$ ). Thus, the Fisher criterion expresses the distance between two clusters relative to their size in one single quantity. Therefore, the larger the value of the Fisher criterion, computed for two clusters, the better the separability of the two clusters. In some sense, this is analogous to the Bhattacharya distance that is also a measure of divergence between two probability densities. For subsequent analyses, the Fisher criterion will be the primary measure of the accuracy of a feature set in discriminating two classes.

## CHAPTER III

## METHODS

## A. Experimental Methods

This chapter describes experimental methods of evaluating tactile and visual texture exploration. Representations of texture through both tactile and visual modalities are generated from models based on the preceding section. These representations are referred to as feature responses (or feature vectors). Many psychophysical and neurophysiological studies have shown that texture processing and discrimination begins with texture boundary detection [14][15][16]. In particular, experiments in this thesis primarily measured sensitivity levels and detection rates of the *target* (texture-defined boundary) to compare relative advantages of the feature space generated by the tactile and visual RF models.

## 1. Input Preparation

In this thesis, all textures listed in the Brodatz texture collection [54] were used to accommodate texture variations as much as possible. Specifically, 30 *target-absent*, or non-boundary texture, inputs were used (Figure 6). To expose effects of boundary-present responses, we conducted a double-blind test by adding a virtually identical, but controlled experiment with the *target*, or texture-defined boundary. Since we have 30 non-boundary textures, we have  $\binom{30}{2} = 435$  unique texture pairs with a boundary in between the paired texture (Figure 7).

Afterwards, each texture was preprocessed by a DoG (or Laplacian of Gaussian) filter before applying the visual RF and tactile RF. This stage is equivalent to the LGN in the human visual system, and is widely used for extracting edges. Similarly,

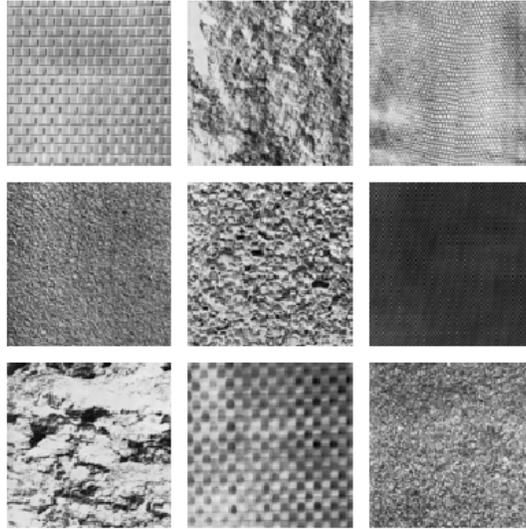


Fig. 6. Sample textures without a texture-defined boundary.

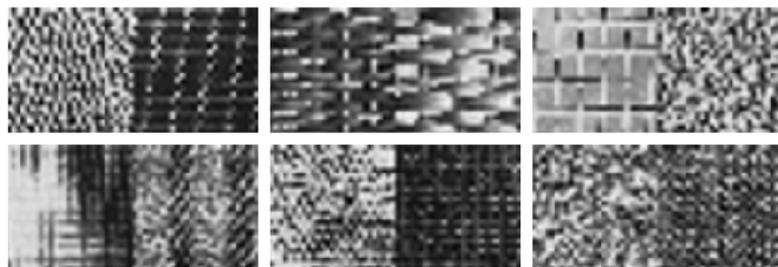


Fig. 7. Sample texture sets with a texture-defined boundary located in between two textures.

for the somatosensory system, there are indications of edge detecting mechanisms such as the LGN in the human visual system [55][56].

## 2. Generating RF Responses

In order to guarantee a fair and objective comparison between two modalities, both models (filters) had identical inputs and were applied under same conditions. However, since the tactile RF has an extra component that dynamically adjusts to the scanning direction, and while we have no prior knowledge of the effects of the scanning direction, scanning directions had to be varied as well to exhaust all possible cases related to various scanning directions. But this extra parameter does not alter the performance of Gabor filters, or the visual RF, since Gabor filters are fixed regardless of scanning direction and the resulting corresponding response vector should be identical. Moreover, to reproduce stimuli from tactile sensation from a finger, we examined a certain number of consecutive window patches (frames) sliding across a predefined scanning direction inside the input image. The pixel intensity in the image played the role of surface height in texture surfaces. For each input image, a local window  $\mathbf{W}$  sized at 17 pixels was used. This was determined by selecting the window size with the most consistent and responsive results from numerous experimental explorations. An example of a scan trajectory superimposed on an input image (containing a texture boundary) is illustrated in Figure 8.

The next step was to construct a bank of filters,  $\{F^{(i)}, i = 1, 2, \dots, 8\}$ , of both tactile and visual RF filters were constructed. The bank of filters had eight equidistances from  $0^\circ$  to  $180^\circ$  degrees in terms of orientation due to the fact that  $22.5^\circ$  degrees, or  $\frac{\pi}{8}$  was the least orientation step size for the RF to show distinctive responses [29]. For each filter in the bank, the corresponding subband response,  $r^{(i)}$  was generated by linearly convolving the image window  $\mathbf{W}$  and an RF filter from the filter bank having

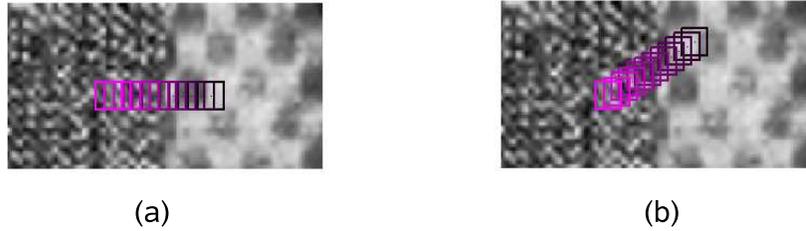


Fig. 8. The square markers indicate the trajectory of window patches that were applied to RF models. (a) shows a scanning oriented at zeros degrees. (b) shows a scanning oriented at  $45^\circ$  degrees. Each marker results in a single scalar value and the whole trail of markers eventually produces a response vector. A texture-defined boundary lies the middle of the input image. Both scans had 12 frames, resulting in a 12 element vector.

the same size, i.e.  $r^{(i)} = F^{(i)} * \mathbf{W} = \sum_u F^i(u) \mathbf{W}(v - u)$ , and finally summing up the pixel values in order to get a single scalar value. Since the image patch and filters have the same window size, this process is equivalent to a dot product operation between the filter and image patch. Repeating this process throughout the traversal (or scan) within the input image, we get a sequence of scalar values which is the response vector, i.e.,  $\{R = r^{(1)}, r^{(2)}, \dots, r^{(n)}\}$  where  $n$  is the number of frames or patches in the traversal. The number of frames determines the number of vector elements.

We should note that a homogenous texture still has variations due to patterns of surface properties. Therefore, a single scan across a texture is likely to omit some texture attributes. In order to overcome this pitfall, multiple scans were traversed across multiple regions in an input image, and 3 to 7 scans were averaged to get the mean response vector in the following experiments. However, the location of the texture boundary within the response vector was fixed and the sequence of traversal was the same as well. Hereafter, a response vector refers to the mean response vector of multiple scans across the same input image. An example response vector from a

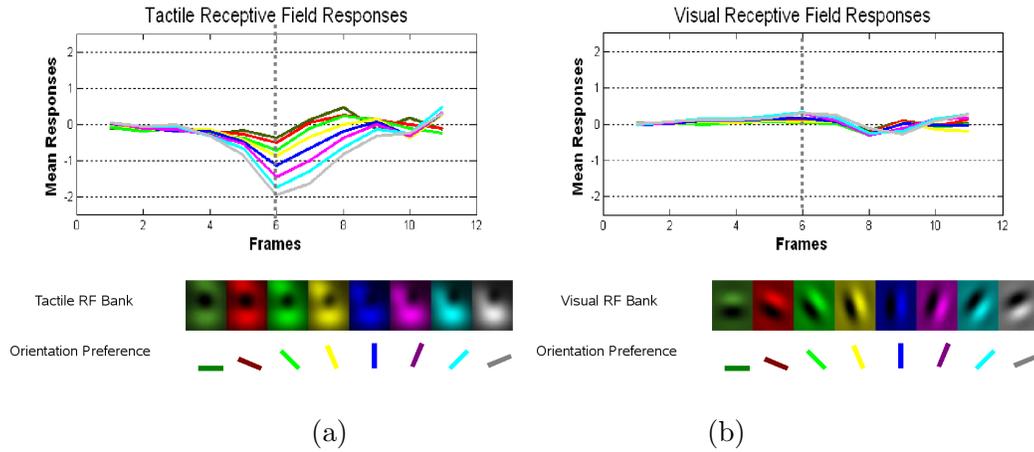


Fig. 9. Plots in (a) and (b) respectively depict response vectors generated the tactile and visual RFs. Both tactile and visual RFs were tuned to different orientation. Each colored vector response correspond to the orientation of the same color.

bank from both modalities are shown in Figure 9. For a last note to conduct a valid double-blind test, the target (texture boundary) should be properly controlled. In other words, a sample containing the target should have the texture boundary at the middle of response vectors, and a sample set without the target should not have any texture-defined boundaries. Specifically, each image in the sample set without the target (texture-defined boundary) basically consists a single homogenous texture. As a summary, Figure 10 depicts the overall system.

### 3. Performance Evaluation

Recalling back to the original hypothesis, our primary interest is to see if touch-based RF generates more discriminative texture features than that of the vision-based approach. In order to compare the representational power of the feature space, we need a method that quantifies the measurement of representational power in the context of texture boundary detection. Since this leads to the problem of characterizing and analyzing a data set, or an unknown population, we can find a logical solution from

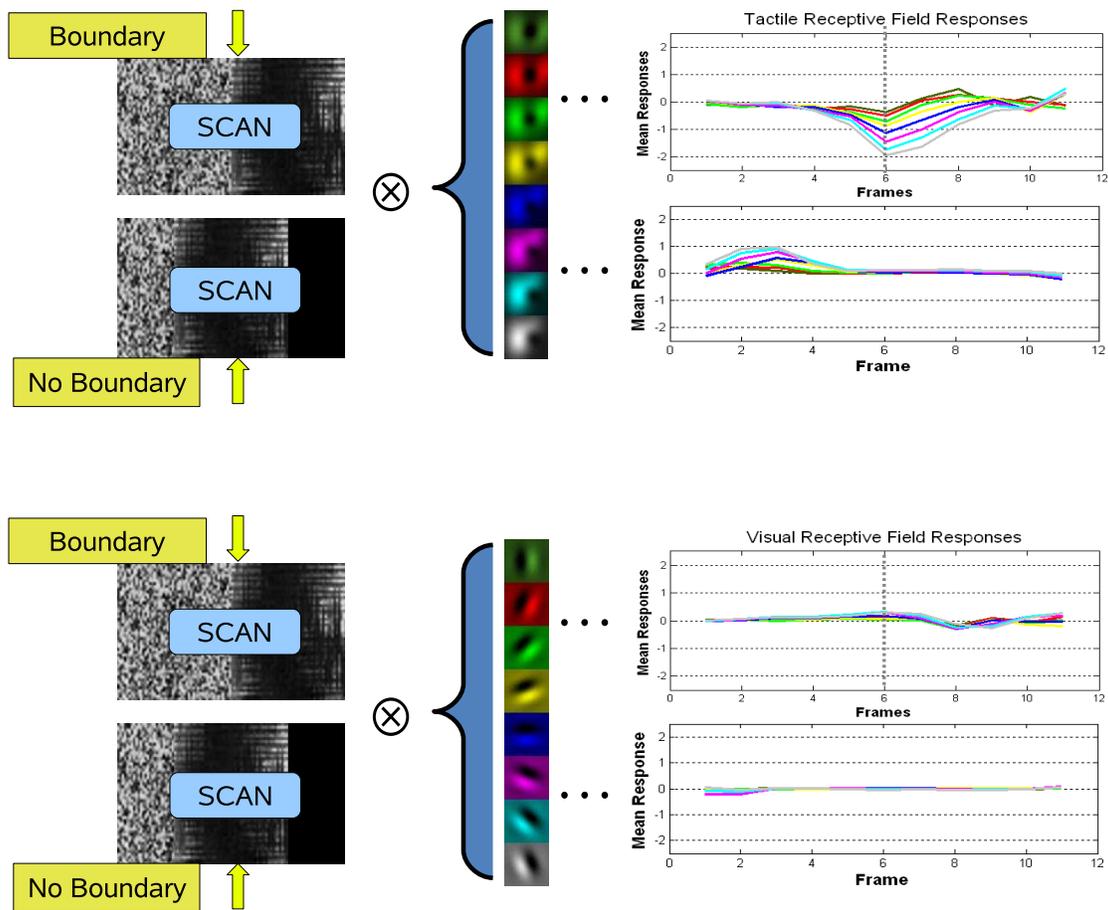


Fig. 10. An outline of how response vectors are generated. For each modality, response vectors from a target-present (boundary) and target-absent (no boundary) input are generated to expose effects of texture boundaries and ultimately, texture discrimination.

statistical approaches.

A widely used measure of the representational power of a feature space is Fisher’s Analysis of Variance (ANOVA). Technically, we would be using a *one-way* ANOVA to assess the effect of having a texture boundary versus not having a texture boundary. This can be a criterion to say which approach is more suitable. However, sensory noise also introduces variability and unexpected artifacts that may violate one-way ANOVA assumptions and produce misleading information. Specifically, if the sample is not independent, one-way ANOVA is simply not appropriate and a variation of ANOVA needs to be used instead. Also, if the assumption of normality is violated, or if outliers are present, one-way ANOVA may be vulnerable to a misleading population mean. However, at a stage of exploring the data set, we do not have sufficient proof or confidence to rely on the assumption of normal distribution. In this case, a non-parametric test or a transformation is required for a more robust test [57].

As an alternative resolution, we can take advantage of classification rates. In this approach, unprocessed feature vectors can be used to train and test classifiers such as the kNN classifier and SVM. If significant differences appear, this method simply proves the superiority of one over the other. However, if this is not the case, then this approach can suffer as well from the absence of statistical invariance of the data. If separability of the feature space is embedded in a certain projection or transformation, then a direct usage of feature vectors can contaminate classification rates. This can be significant in the field of texture analysis, since subsequent non-linear pooling and processing is usually applied to the raw feature space.

Another aspect of using classifiers in evaluating a system is that subsequent processes alter the feature space and are liable to represent the joint performance. Nonetheless, the overall enhancement, i.e. transformation or dimensionality reduction, is still subject to characteristics of the initial space. In other words, feature

spaces with higher separability is the dominant factor of data classification and post-processing does not compensate for differences from the initial feature space. Therefore, a measurement of data separability can be a robust measure of determining performance. From discriminant analysis, we can objectively compare the features by measuring cluster separability according to the Fisher criterion. However, the Fisher criterion also contains limitations since it only commits to the mean and variance of the projected features. In other words, this function becomes less useful if the mean and variance do not necessarily explain the distribution characteristics [58][59]. However, experimental results indicate that the above measurements can appropriately conduct performance comparisons. The following chapter addresses these limitations as well as determining which modality generates more discriminative texture representations.

## CHAPTER IV

## RESULTS

In this chapter, the characteristics and underlying structure of feature spaces generated by the tactile and visual receptive fields (RF) are explained. Afterwards, overall performance results are listed. The feature space was separately generated by the tactile and visual RFs. For each feature space,  $8 \times 8 \times 8$  subsets had to be generated according to the three parameters that define the RFs:

- (1) orientation preference ( $\frac{\pi}{8} \cdot 1, \frac{\pi}{8} \cdot 3, \dots, \frac{\pi}{8} \cdot 8$ );
- (2) lagging center location ( $\frac{\pi}{8} \cdot 1, \frac{\pi}{8} \cdot 2, \frac{\pi}{8} \cdot 3, \dots, \frac{\pi}{8} \cdot 8$ );
- (3) scanning direction ( $\frac{\pi}{8} \cdot 1, \frac{\pi}{8} \cdot 2, \frac{\pi}{8} \cdot 3, \dots, \frac{\pi}{8} \cdot 8$ ).

Therefore, if a subset indicates its orientation preference was set to 0 degrees, then 8 *lagging center locations*  $\times$  8 *scanning directions* were averaged to create this subset. In this manner, we can expose a particular parameter, i.e. orientation preference, for subsequent analyses. Hereafter, a subset refers to a data set that was generated by a particular combination of the three parameters, and the feature space refers to the entire group of subsets. Also, *feature vectors* are interchangeably termed as *response vectors*.

## A. Data Characteristics

In this thesis, feature vectors that were generated from tactile and visual RFs exist in higher dimensions depending on the number of frames per scan. For instance, if a scan across a texture boundary consists of 12 local window patches, the resulting response vector contains 12 scalar elements as well. When exploring these measurements, it is preferable to represent data in a reduced number of dimensions. Reasons for doing this may be easier subsequent analysis, improved classification performance through

a more stable representation, removal of redundant or irrelevant information, or an attempt to discover underlying structure by obtaining a graphical representation. In pattern recognition literature they are termed *feature selection* and *feature extraction* methods.

### 1. Visualization from Feature Selection

Given a set of measurements, dimensionality reduction was achieved in essentially two different ways. The first identified those variables that do not contribute to the classification task, and is termed *feature selection*. Moreover, since the primary concern of this thesis is a discrimination problem, we would neglect those variables that do not contribute to class separability. The second approach transformed the measurements to a lower dimensional space. This is *feature extraction* and is depicted in the next section.

The first step is to choose feature sets for which the discrimination characteristic is enhanced. This is where ANOVA aids the overall process. By using ANOVA, we can first visualize if the feature representation of the two classes, for boundary and non-boundary feature sets, are significantly different. If ANOVA shows significant differences, a data separability measure can simply be applied to characterize the data. However, if ANOVA fails, we need to isolate outliers or feature sets with unstable (i.e. high  $p$ -values) statistics.

From Figures 11 and 12, we can expect that the tactile feature space contains more class separability than that of the visual feature space, but this does not define the overall discriminative power of the feature space. Moreover, outliers in the visual feature space might have contaminated the overall set.

Intuitively, the above results imply that some structural differences may exist that trigger the discriminative power of both modalities. In particular, from ANOVA,

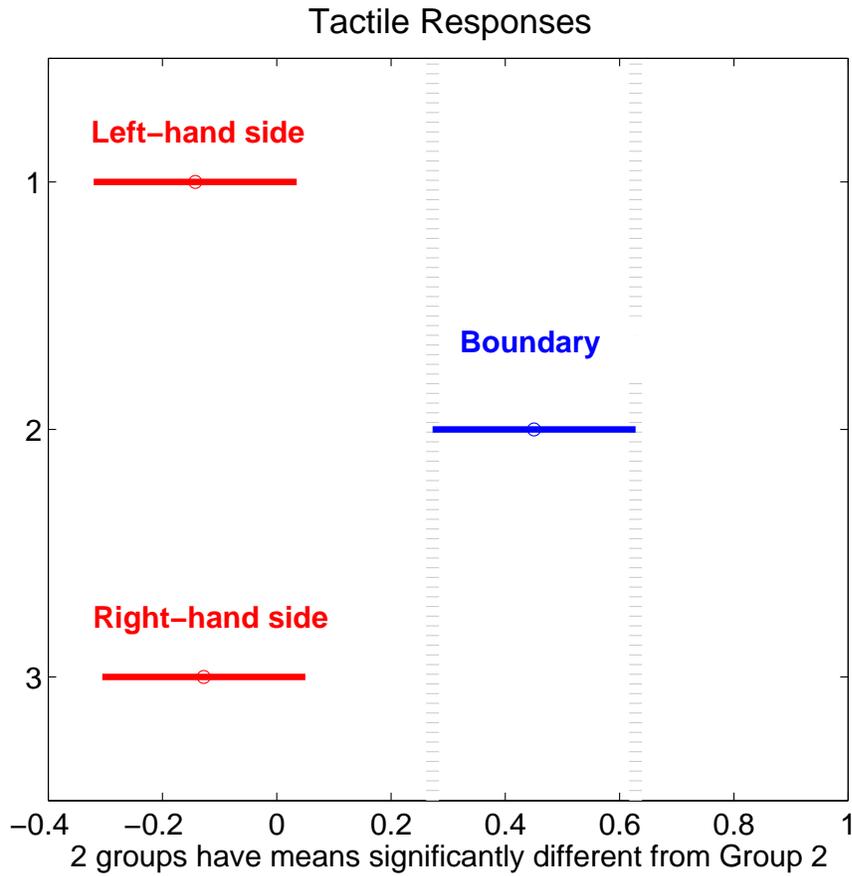


Fig. 11. Three sets of one-way ANOVA were applied to test significant differences between: (1) left-hand side homogenous texture, (2) texture boundary, (3) right-hand side homogenous texture. The  $y$ -axis and  $x$ -axis relatively depicts the sample index and mean values. The plot shows that the means of measurements are significantly different. Corresponding  $p$ -values were under 0.01.

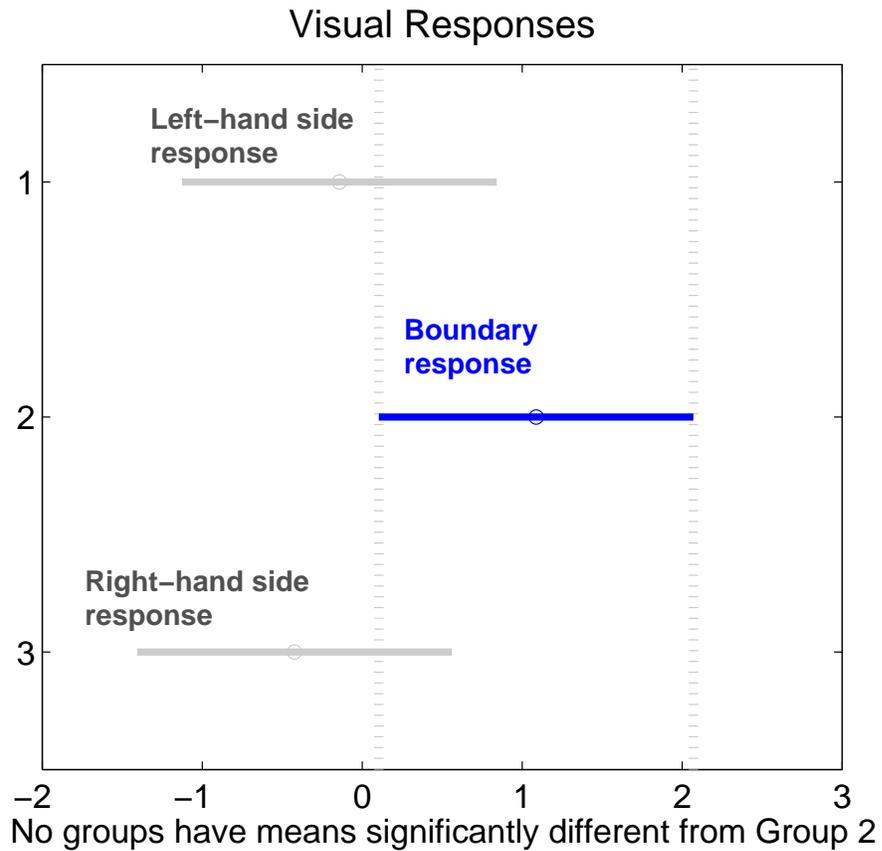


Fig. 12. Three sets of one-way ANOVA was applied to test significant differences between: (1) left-hand side homogenous texture *vs.* texture boundary, (2) left-hand side homogenous texture *vs.* right-hand side homogenous texture, and (3) texture boundary *vs.* right-hand side homogenous texture. The  $y$ -axis and  $x$ -axis relatively depicts the sample index and mean values. The plot indicate that the means of measurements are not significantly different. Corresponding  $p$ -values were over 0.5.

the tactile feature space is likely to have more separable structure of features compared to the visual feature space. Figure 13 illustrates this idea by stacking sample response vectors from both modalities. In the figure, response vectors represent a texture boundary, and the superimposed image was placed to indicate where texture boundaries would be located within the response vectors. As expected, tactile features are more sensitive to texture boundaries whereas visual features have a weak representation of the texture boundary. This expectation was captured in ANOVA but we still need further proof to provide a more robust comparison between tactile and visual feature spaces.

This leads to the second step in feature selection. The second approach estimated the overlap between the distributions from which the data are drawn and favor those feature sets for which this overlap is minimal and maximizes separability. This approach is normally referred to as Fisher’s *Linear Discriminant Analysis* (LDA) and usually assumes Gaussianity of the underlying structure of measurements. Another attractive aspect of LDA involves feature space dimensionality reduction while preserving as much of the class discriminatory information as possible [57]. This enables us to visualize the overall structure of the high-dimensional feature space as in Figure 14.

From the scatter plots, we can observe that the both tactile and visual feature space contain discriminatory information across the mean locations. Although tactile space seems to generate a much wider gap between classes, the visual feature space is also contains separable discriminatory information as well, but in a much denser region. Therefore, we need to account for the within-class spread as well as the between-class spread in order to objectively evaluate separability among different scales of clusters. The Fisher criterion is equipped for such cases and as stated in previous chapters, the larger the value of the Fisher criterion the better the separa-

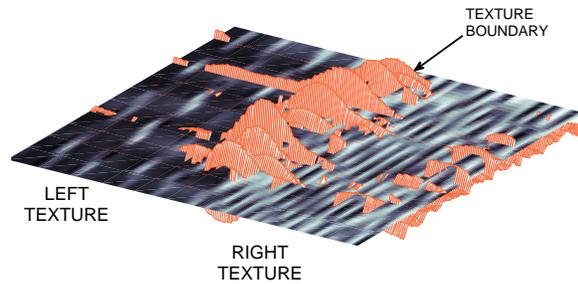


Figure (a)

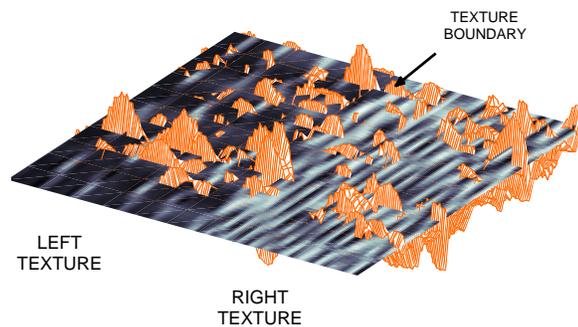


Figure (b)

Fig. 13. (a) A sample of visual feature vectors from various scans are superimposed on top of an example input image with a texture boundary in the middle. (b) A sample of tactile feature vectors from various scans are superimposed onto an example input image with a texture boundary in the middle. The response features shown here were generated from various images in the sample set and not from the single example input image. The background image was only used to depict where the texture boundary lies throughout the images in the sample set. Note that tactile feature vectors respond abruptly across the texture-defined boundary while visual feature vectors are less sensitive to texture boundaries. This figure implies that touch-based models are more correlated to texture boundaries than vision-based models.

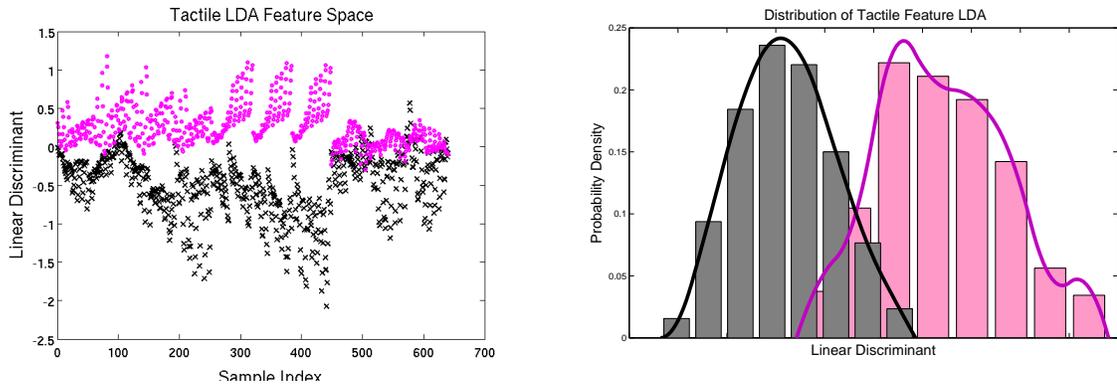


Figure (a)

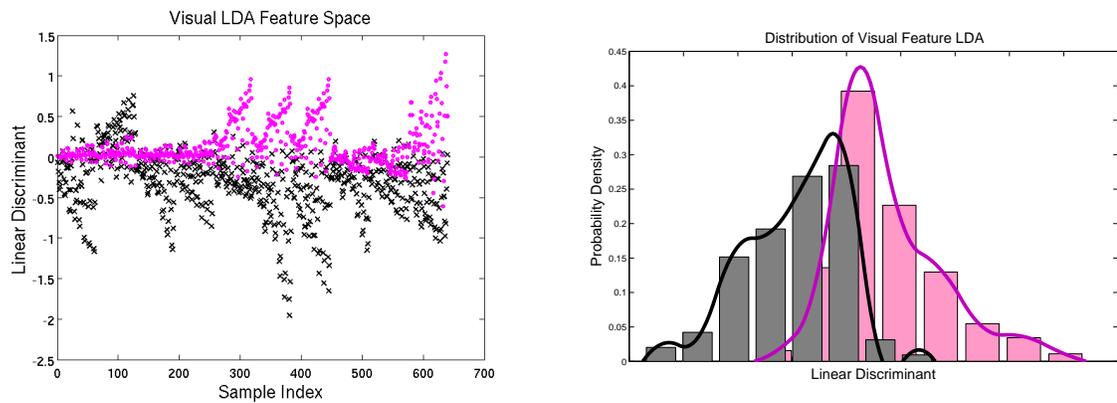


Figure (b)

Fig. 14. (a) A scatter plot of 1-D projections (from LDA) of the tactile feature space and its probability density estimation is illustrated. For the scatter plot, the  $x$ -axis depicts the sample index of each texture response vector, and the  $y$ -axis indicates the linear discriminant (values of the 1-D projection values). In each plot, magenta(●) and black(×) represents boundary and non-boundary features, respectively. (b) Equivalent plots of the visual feature space. The projected feature space (LDA scatter plots) and corresponding density estimations indicate that the tactile feature space contains more class separability for texture discrimination.

bility of the two clusters. Specific results of the tactile and visual feature space are listed in Table 1.

Table I shows the resulting Fisher criterion of the touch and vision-based feature space with respect to three major parameters, i.e., scan direction, lagging center location, and orientation preference varying in 8 different angles. For each column depicted with a major parameter, the parameter is fixed to the values in each row while the other two parameters are varied and the overall Fisher criterion values are averaged along the corresponding major parameter (column). Higher Fisher criterion values are highlighted in the table.

Note that the two clusters that the Fisher criterion measures is between boundary and non-boundary features. Since the RF models are defined by three major parameters—scan direction, lagging center location, and orientation preference—Fisher criterion had to be measured accordingly across different parameter values ( $0^\circ, 22.5^\circ, \dots, 157.5^\circ$ ) to exhaust all possible combination of the RF profile. From Table 1, we can observe that the majority of higher Fisher criterion values belong to the tactile feature space that was generated from various setting and combinations of the tactile RFs (Note that the visual responses in the second major column are all identical since the visual RF lacks the concept and component of a lagging center as in the tactile model). Since the Fisher criterion is a measure of class separability, this result suggests that the tactile feature space contains more discriminative information about texture boundaries than the visual feature space.

## 2. Visualization from Feature Extraction

Early in this chapter, two main visualizations were introduced and this section considers the latter method of *feature extraction*. The most common method of feature extraction is *Principal Component Analysis (PCA)* originated by Pearson [60]. Geo-

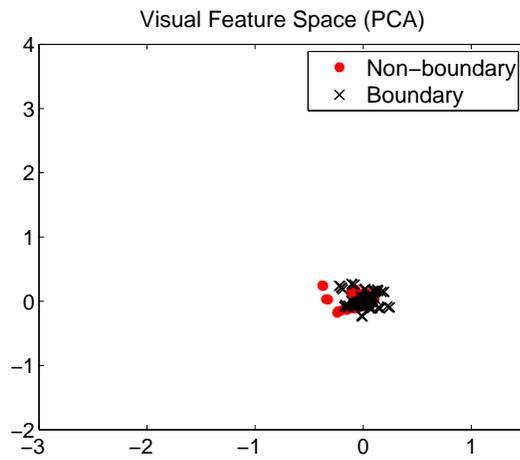


Figure (a)

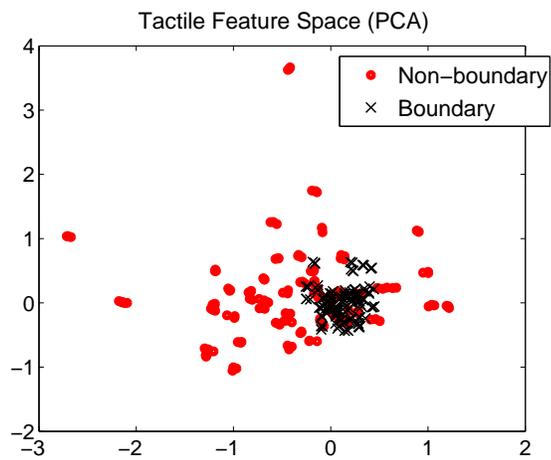


Figure (b)

Fig. 15. Overall PCA scatter plots are listed for each modality. The first two eigen-vectors were used to visualize the initial feature space. (a) 2-D scatter plot of the visual feature space. (b) 2-D scatter plot of the tactile feature space.

Table I. Measures of data separability based on isolated parameters..

| Fisher Criterion | Orientation | Scan Orientation |            | Lagging Centers |            | Orientation Preference |            |
|------------------|-------------|------------------|------------|-----------------|------------|------------------------|------------|
|                  |             | <i>TRF</i>       | <i>VRF</i> | <i>TRF</i>      | <i>VRF</i> | <i>TRF</i>             | <i>VRF</i> |
|                  | 0           | 0.28             | 0.38       | 0.70            | 0.05       | 0.74                   | 0.43       |
| 22.5             | 0.19        | 0.43             | 0.65       | 0.05            | 0.76       | 0.32                   |            |
| 45               | 0.17        | 0.49             | 0.63       | 0.05            | 0.26       | 0.12                   |            |
| 67.5             | 0.21        | 0.16             | 0.47       | 0.05            | 0.24       | 0.09                   |            |
| 90               | 0.26        | 0.06             | 0.34       | 0.05            | 0.90       | 0.11                   |            |
| 112.5            | 0.36        | 0.18             | 0.16       | 0.05            | 0.90       | 0.10                   |            |
| 135              | 0.47        | 0.43             | 0.06       | 0.05            | 0.50       | 0.05                   |            |
| 157.5            | 0.57        | 0.40             | 0.14       | 0.05            | 0.37       | 0.22                   |            |

metrically, PCA can be thought of as a rotation of the axes of the original coordinate system to a new set of orthogonal axes that are ordered in terms of the amount of variation of the original data they account for. Therefore, PCA simply performs orthogonal coordinate rotation that aligns the transformed axes with directions of the maximum variance and is widely used to visualize high-dimensional data sets. However regarding class separability, PCA does not consider class separability (as LDA) and does not guarantee that the directions of maximum variance will contain good features for discrimination [51].

Figure 15 depicts PCA of both feature spaces. The resulting overall structure indicates that the discriminatory information is not contained in the variance of the data. Moreover, since LDA succeeded to show more class separability, we can conclude that the discriminatory information is not contained in the variance of the data but in the distributions of the mean of the clusters.

Table II. kNN classification rates of both models based on isolated parameters.

|  | Orientation  | Scan Orientation |             | Lagging Centers |             | Orientation Preference |             |
|--|--------------|------------------|-------------|-----------------|-------------|------------------------|-------------|
|  |              | touch            | vision      | touch           | vision      | touch                  | vision      |
| <b>Classification<br/>Rate using<br/>kNN</b> | <i>0</i>     | <b>0.84</b>      | <b>0.86</b> | <b>0.89</b>     | <b>0.77</b> | <b>0.90</b>            | <b>0.93</b> |
|  | <i>22.5</i>  | <b>0.84</b>      | <b>0.88</b> | <b>0.88</b>     | <b>0.77</b> | <b>0.96</b>            | <b>0.92</b> |
|  | <i>45</i>    | <b>0.84</b>      | <b>0.89</b> | <b>0.87</b>     | <b>0.77</b> | <b>0.93</b>            | <b>0.92</b> |
|  | <i>67.5</i>  | <b>0.88</b>      | <b>0.89</b> | <b>0.87</b>     | <b>0.77</b> | <b>0.94</b>            | <b>0.90</b> |
|  | <i>90</i>    | <b>0.86</b>      | <b>0.89</b> | <b>0.85</b>     | <b>0.77</b> | <b>0.94</b>            | <b>0.96</b> |
|  | <i>112.5</i> | <b>0.86</b>      | <b>0.88</b> | <b>0.87</b>     | <b>0.77</b> | <b>0.96</b>            | <b>0.92</b> |
|  | <i>135</i>   | <b>0.86</b>      | <b>0.87</b> | <b>0.87</b>     | <b>0.77</b> | <b>0.94</b>            | <b>0.90</b> |
|  | <i>157.5</i> | <b>0.89</b>      | <b>0.87</b> | <b>0.85</b>     | <b>0.77</b> | <b>0.93</b>            | <b>0.90</b> |

## B. Performance Evaluation

From previous results, we should expect that any subsequent classifier would favor the touch over vision-based features due to the relatively enhanced class separability. This section shows that classification results are analogous to our expectations. Recall that *k-Nearest Neighbors* (kNN) classifiers are robust in exploring unknown data sets, and *Support Vector Machines* (SVM) guarantees maximum separability.

Tables II and III show kNN and SVM classification rates, respectively. Similar to Table I, this table also evaluated classification rates from three major parameters. For each column, a parameter is fixed to one of the orientation values in each row while the other two parameters are varied and the overall classification rates are averaged along the corresponding major parameter (column). Higher Fisher criterion values are highlighted in the table. Overall, these results validate the previous statement (last paragraph of Chapter III) that post-processing classification does not compensate for discrimination capacity of the initial feature space. Therefore, the separability

Table III. SVM classification rates of both models based on isolated parameters.

|                               | Orientation  | Scan Orientation |             | Lagging Centers |             | Orientation Preference |             |
|-------------------------------|--------------|------------------|-------------|-----------------|-------------|------------------------|-------------|
|                               |              | touch            | vision      | touch           | vision      | touch                  | vision      |
| Classification Rate using SVM | <i>0</i>     | <b>0.85</b>      | <b>0.80</b> | <b>0.87</b>     | <b>0.79</b> | <b>0.92</b>            | <b>0.91</b> |
|                               | <i>22.5</i>  | <b>0.85</b>      | <b>0.80</b> | <b>0.86</b>     | <b>0.79</b> | <b>0.95</b>            | <b>0.92</b> |
|                               | <i>45</i>    | <b>0.84</b>      | <b>0.68</b> | <b>0.85</b>     | <b>0.79</b> | <b>0.95</b>            | <b>0.89</b> |
|                               | <i>67.5</i>  | <b>0.82</b>      | <b>0.88</b> | <b>0.84</b>     | <b>0.79</b> | <b>0.92</b>            | <b>0.93</b> |
|                               | <i>90</i>    | <b>0.82</b>      | <b>0.90</b> | <b>0.85</b>     | <b>0.79</b> | <b>0.92</b>            | <b>0.84</b> |
|                               | <i>112.5</i> | <b>0.81</b>      | <b>0.88</b> | <b>0.85</b>     | <b>0.79</b> | <b>0.92</b>            | <b>0.94</b> |
|                               | <i>135</i>   | <b>0.81</b>      | <b>0.71</b> | <b>0.85</b>     | <b>0.79</b> | <b>0.92</b>            | <b>0.90</b> |
|                               | <i>157.5</i> | <b>0.81</b>      | <b>0.81</b> | <b>0.85</b>     | <b>0.79</b> | <b>0.94</b>            | <b>0.91</b> |

of the initial feature space simply forecasts the classification performance as well. Moreover, the underlying data structure and the Fisher criterion suggests analogous explanations of this outcome.

As a summary, data visualization and statistical metrics suggest that (1) texture discriminatory information is enhanced in the tactile feature space, and (2) touch-based approach is more robust for texture segmentation than the visual (Gabor) featured approach.

## CHAPTER V

### DISCUSSION

This chapter describes the contribution of this thesis related to several works, as well as relevant issues from previous chapters, and future work based on the work presented in this thesis.

#### A. Contribution

As one of the major fields in visual scene analysis, texture perception has been approached and motivated by various fields such as psychophysical studies [3][5][6], machine learning, set theory [61], information theory [7] and so on. Overall, the foundation that contributed to pioneering findings in texture perception has roots in biologically plausible concepts. Seemingly, it would be imprudent to ignore inspirations from biology. However, treating texture as a surface property, mainstream texture analysis has been confined to perceive texture as a visual task. Although neurophysiological findings have shown that both human tactile and visual systems hold significant similarities [2], this thesis focuses on the tactile modality that is more intimately associated with the essence of texture. This thesis proposed that touch is more ecologically suited for texture discrimination than the mainstream vision-based approach.

To leverage on this, a tactile RF model was derived in a similar fashion to the visual RF model. The biologically plausible model was first derived and presented by DiCarlo and Johnson to study the neural representation of tactile perception [44]. Based on their work, a similar tactile RF model was implemented for the purpose of achieving a more powerful foundation for texture segmentation. Given a computational model of the tactile RF, a prior condition was to prove its relative usefulness

based on the widely used visual RF model. In order to validate this statement, statistical measures and pattern recognition techniques were used to compare the RF models' representation performances in discriminating textures. The overall experimental evaluation was conducted by generating data with a controlled target signal (texture boundary) and target absent measurements. Fundamental feature selection (LDA) and extraction (PCA) methods were used to analyze the underlying structure of feature spaces, and the Fisher criterion (from LDA) was used to measure discriminatory information from each feature space. Afterwards, kNN classification and SVM was used to ensure that the feature space separability is reflected in classification performance as well. These results suggest that tactile RF features contain more discriminatory information of textures than the visual RF. Therefore, considering texture analysis from a touch-based approach is more intimately associated with texture than a vision-based approach and complies with our initial expectations that texture as a surface representation is indeed essentially more suitable than representing it as a visual property.

## B. Limitations of the Model

The vision and touch-based models are basically filters that are constructed based on the properties of cortical RFs. Throughout this thesis, the touch-based model has been shown to produce enhanced discrimination among natural textures than the vision-based model. The touch-based model consists of three Gaussian filters, the first corresponding to the excitatory center, the other two to inhibitory lobes (specifically, one flanked from the excitatory center, and the other temporally lagged with respect to it) that account for much of the actual properties of a somatosensory RF. Because the Gabor filter (vision-based model) describes an excitatory mass flanked by an

inhibitory one, the Gabor model is analogous to the touch-based model but with fewer degrees of freedom. Therefore, the extra lagging component of the touch-based model is the fundamental factor that distinguishes it from the vision-based model, and provides even higher capacity of handling spatial information due to this extra degree of freedom.

One should also note that this additional component *lags* and therefore relocates itself throughout scannings (active explorations). Moreover, since exploration is locally directed, the touch-based filter dynamically adjusts itself and the resulting combination of all three components quantifies texture features. Contrary to this dynamic mechanism, the vision-based model explicitly defines a one-to-one correspondence to its preference and sensitivity according to the orientation and frequency of the Gaussian lobes. At this point, unlike the vision-based model, fundamental factors of the touch-based model that convey preference and sensitivity of certain spatial features are ambiguous and are likely to be inconsistent due to dynamic interactions.

### C. Future Work

Texture representation, or texture feature generation, as a surface property was the primary interest of this thesis. Although texture representation has been widely viewed as a visual task, statistical evidence shows that touch-based approaches contain more discriminatory information of texture representations. However, despite this advantageous aspect, more elaborate investigations are required to fully understand which properties of the touch-based model facilitate texture discrimination. Although there exist several ambiguities of the model to unveil, a preliminary comparison provides an insight to extract leading factors that affect model performance. The main support of this test comes from the benchmark factors of the visual model that define special

Table IV. Summary of model performance from independent parameters.

| <b>PARAMETER ISOLATION TEST</b><br>(One free parameter while other two are fixed) |                       |               |                               |               |                           |               |
|---|-----------------------|---------------|-------------------------------|---------------|---------------------------|---------------|
|   | <b>Lagging Center</b> |               | <b>Orientation Preference</b> |               | <b>Scanning Direction</b> |               |
| <b>Criterion</b>  | <b>touch</b>          | <b>visual</b> | <b>touch</b>                  | <b>visual</b> | <b>touch</b>              | <b>visual</b> |
| <b><i>Fisher Criterion</i></b>  | 18.6                  | 14.7          | 40.6                          | 30.9          | 7.8                       | 11.2          |
| <b><i>Classification Rate</i></b>   | 73%                   | 57%           | 71%                           | 67%           | 63%                       | 61%           |

characteristics of the input space, i.e., orientation preference and sensitivity depends on the disposition of the *on* and *off* Gaussians. Based on these benchmark factors, we can conduct an experiment to explore which factors enhance (or deteriorate) the advantageous performance of the touch-based model.

Table IV evaluates model performance by means of the Fisher criterion and average classification rates between the kNN classifier and SVM. Using these two evaluation criteria, three major parameters were adjusted to explore how they affect the model performance. Specifically, the three parameters are: (1) orientation preference (orientation of Gaussian lobes), (2) lagging component location (initial lagging component location respect to excitatory region; this only applies to the touch-based model) and, (3) scanning direction. Corresponding columns under the parameter tests the touch and vision-based models while fixing itself (parameter of interest) and varying the other two. The resulting values of each criterion is an average of the varied instances.

Since the two models have analogous aspects in filter structures but differ only in the lagging component, we can expect that the lagging component and orientation preference are affected by this component and eventually contribute to the overall

model performance. Table IV confirms this statement where the lagging component and orientation preference leads to enhanced performances of the touch-based model.

On the other hand, we can expect that scanning direction has minimum effect in differentiating two model performances since when the extra lagging component is fixed, the touch-based model is essentially identical to the vision-based model with an extra constant segment. The last column under *scanning direction* confirms this statement by means of class separability and classification rates. Therefore, these initial findings suggest that similar attributes of the vision-based model also aid the touch-based model, but it should be noted that these results implicitly contain a conglomerate of unforeseen factors and require further analyses.

Another aspect that has not been delivered in this thesis is the subsequent *post-processing* stage in texture segmentation literature. This area includes a variety of approaches that includes hierarchical mechanisms inspired from biological models, Bayesian inference, information theory and so forth. These approaches commonly calls upon fundamental aspects of neural encoding and decoding issues that is left for future work. However, based on findings in this thesis, tactile perception of texture is expected to provide a powerful foundation for subsequent texture analysis.

## CHAPTER VI

### CONCLUSION

The goal of this thesis was to understand texture perception based on intuitive, biological inspirations through computational modeling. This thesis was initiated from studying visual scene analysis and advanced to visual texture perception. However, considering the nature of texture, the overall assessment of texture perception was redirected from a visual task into a surface related exploration. Coincidentally, psychophysical and neurophysiological studies provided a breakthrough in addressing both modalities for the common issue of texture perception.

Based on the above inspirations, I simulated both touch and vision-based models from cortical receptive field profiles in order to generate texture features. Furthermore, controlled experiments on both modalities were conducted to investigate which approach provided advantageous information regarding texture discrimination. Through these computational models, statistical techniques were applied to prove that an unprecedented touch-based approach was more ecologically suited for texture perception rather than the widely accepted visual-based approach.

In the future, touch-based texture perception is expected to provide an effective foundation for texture analysis.

## REFERENCES

- [1] A. Rosenfeld and M. Thurston, “Edge and curve detection for visual scene analysis,” *IEEE Transactions on Computers*, vol.C-20, no. 2, pp.562–569, May 1971.
- [2] A. Thielscher and H. Neumann, “A computational model to link psychophysics and cortical cell activation patterns in human texture processing,” *Journal of Computational Neuroscience*, vol.22, no. 3, pp. 255-282, 2006.
- [3] B. Julesz, “Texture and visual perception,” *Scientific American*, vol. 212, pp. 38–48, Feb 1965.
- [4] B. Julesz, “Textons, the elements of texture perception, and their interactions,” *Nature*, vol. 290, no. 9, pp. 9197, 1981.
- [5] J. Beck, “Effect of orientation and shape similarity on perceptual grouping,” *Perception and Psychophysics*, vol. 1, no. 3, pp. 300–302, 1966.
- [6] X. Liu and D. Wang, “A spectral histogram model for texton modeling and texture discrimination,” *Vision Research*, vol. 42, pp. 2617–2634, 2002.
- [7] J. Kim, J. W. Fisher III, A. Yezzi, M. Cetin, and A. Willsky, “A nonparametric statistical method for image segmentation using information theory and curve evolution,” *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1486–1502, October 2005.
- [8] S. Oh and Y. Choe, “Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparison of 2D vs. 3D representations,” in *Proceedings of the 2004 International Conference on Development and Learning [electronic]*, 2004, University of California at San Diego Institute for Neural Computation, pp. 175-182.

- [9] S. Oh and Y. Choe, "Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparison of 2D vs. 3D representations," *Neurocomputing*, vol. 70, no. 2, pp. 2245-2255, 2007.
- [10] J. R. Bergen and M. S. Landy, "Computational modeling of visual texture segregation," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, Eds. Cambridge, MA: MIT Press, 1991, pp. 253-272.
- [11] B. Julesz, "Visual pattern discrimination," *IRE Transactions on Information Theory*, vol. 8, no. 3, pp. 84-92, 1962.
- [12] J. Beck, "Perceptual grouping produced by changes in orientation and shape," *Science*, vol. 154, pp. 538-540, 1966.
- [13] W. Jiang, F. Tremblay, and C. E. Chapman, "Neural encoding of texture changes in the primary and the secondary somatosensory cortical areas of monkeys during passive texture discrimination," *Journal of Neurophysiology*, vol. 77, no. 3, pp. 1656-1662, 1997.
- [14] Z. Li, "Pre-attentive segmentation in the primary visual cortex," *Spatial Vision*, vol. 13, no. 1, pp. 25-50, 2000.
- [15] Z. Li, "A saliency map in primary visual cortex," *Trends in Cognitive Science*, vol. 6, no. 1, pp. 9-16, January 2002.
- [16] M. Leventhal, Y.C. Wang and Y. Zhou, "Neural correlates of boundary perception," *Visual Neuroscience*, vol. 15, no. 6, pp. 1107-1118, 1998.
- [17] S. Oh and Y. Choe, "Segmentation of textures defined on flat vs. layered surfaces using neural networks: Comparison of 2D vs. 3D representations," *Neurocomputing*, vol. 70, no. 13, pp. 2245-2255, 2007.

- [18] M. Wertheimer, "Principles of perceptual organization," In D.C Bardslee and M. Wertheimer (eds.) *Readings in Perception.*, Princeton, NJ: Van Nostrand, 1957, pp. 115–135.
- [19] J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature*, vol.333, no. 6171, pp. 363–367, 1988.
- [20] T. M. Caelli, "Three processing characteristics of visual texture segmentation," *Spatial Vision*, vol. 1, no. 1, pp. 19–30, 1985.
- [21] M. R. Turner, "Texture discrimination by Gabor functions," *Biological Cybernetics*, vol. 55, no. 2, pp. 71–82, 1986.
- [22] H. Voorhees and T. Poggio, "Computing texture boundaries from images," *Nature*, vol. 333, no. 6171, pp. 364–367, 1988.
- [23] I. Fogel and D. Sagi, "Gabor filters as texture discriminators," *Biological Cybernetics*, vol. 61, no. 2, pp. 103–113, 1989.
- [24] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *Journal of the Optical Society of America*, vol. 7, no. 5, pp. 923–932, 1990.
- [25] N. Graham, J. Beck, and A. Sutter, "Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast," *Vision Research*, vol. 32, no. 4, pp. 719–743, 1992.
- [26] X. Liu and D. Wang, "Image and texture segmentation using local spectral histograms," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 70–73, October 2006.

- [27] E. Deibert, M. Kraut, S. Kremen, and J. Hart, “Neural pathways in tactile object recognition,” *Neurology*, vol. 52, no. 7, pp. 1413–1417, 1999.
- [28] C. Spence, F. Pavani, and J. Driver, “Crossmodal links between vision and touch in covert endogenous spatial attention,” *Journal of Experimental Psychology. Human Perception and Performance*, vol. 26, no. 41, pp. 298–319, 2000.
- [29] S. Bensimaia, P. Denchev, J. Dammann III, J. Craig, and S. Hsiao, “The representation of stimulus orientation in the early stages of somatosensory processing,” *Journal of Neuroscience*, vol. 28, no. 3, pp. 776–786, 2008.
- [30] R. W. Connors and C.A. Harlow, “A theoretical comparison of texture algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 3, pp. 204–222, 1980.
- [31] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes, Exploring the Neural Code*. Cambridge, MA: MIT Press, 1999.
- [32] D.H. Hubel and T.N. Wiesel, “Receptive fields of single cells in the cats striate cortex,” *Journal of Physiology*, vol.148, no. 3, pp. 574-591, 1959.
- [33] G. Montgomery, “Seeing, hearing and smelling the world,” 1997, Howard Hughes Medical Institute, URL <http://www.hhmi.org/senses>.
- [34] N. Logothetis, “The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imageing signal,” *Philosophical Transactions Royal Society. Biological Sciences*, vol.357, no. 5, 1003-1005, 2002.
- [35] S.W. Kuffler, “Neurons in the retina: Organization, inhibition and excitation problems,” *Symposium of Quantitative Biology*, vol.27, no. 3, pp. 281-292, 1952.

- [36] D. Marr, *Vision*. San Francisco, CA: W.H. Freeman, 1982.
- [37] D.H. Hubel and T.N. Wiesel, “Receptive fields, binocular interaction and functionarchitecture in the cat’s visual cortex,” *Journal of Physiology*, vol.160, no. 2, pp. 106-154, 1962.
- [38] J. Jones and L. Palmer, “An evaluation of the two-dimensional Gabor filter model of simple receptive fields in the cat’s striate cortex,” *Journal of Neurophysiology*, vol.58, no. 9, pp. 1233-1258, 1987.
- [39] M. Porat and Y.Y. Zeevi, “The generalized Gabor scheme of image representation in biological and machine vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 452-468, 1988.
- [40] M. Clark and A.C. Bovik, “Experiments in segmenting texton patterns using localized spatial filters,” *Pattern Recognition*, vol. 22, no. 6, pp. 707–718, 1989.
- [41] A. Amedi, R. Malach, and T. Hendler, “Visuo-haptic object-related activation in the ventral visual pathway,” *Nature Neuroscience*, vol. 4, no. 3, pp.324–330, 2001.
- [42] P.J. Fitzgerald, J.W. Lane, P. H. Thakur, and S. Hsiao, “Receptive field properties of the macaque second somatosensory cortex: Representation of orientation on different finger pads,” *Journal of Neuroscience*, vol. 26, no. 24, pp.6473–6484, June 2006.
- [43] P.J. Fitzgerald, J.W. Lane, P. H. Thakur, and S. Hsiao, “Receptive field (RF) properties of the macaque second somatosensory cortex: RF size, shape, and somatotopic organization,” *Journal of Neuroscience*, vol. 26, no. 24, pp.6485–6495, June 2006.

- [44] J. DiCarlo and K. Johnson, “Structure of receptive fields in primate somatosensory area 3b: Effects of stimulus scanning direction and orientation,” *Journal of Neuroscience*, vol. 20, no. 20, pp.495–510, 2000.
- [45] S.J. Bensmaia, P.V. Denchev, J.F. Dammann III, J.C. Craig, and S. Hsiao, “The representation of stimulus orientation in the early stages of somatosensory processing,” *Journal of Neuroscience*, vol. 28, no. 3, pp. 776–786, 2008.
- [46] E. Fix and J.L. Hodges, “Discriminatory analysis, nonparametric discrimination: consistency properties,” Technical Report 4, USAF School of Aviation Medicine, Randolph Field, TX, 1951.
- [47] T.M. Cover and P.E. Hart, “Nearest neighbor pattern classification,” *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [48] K. Fukunaga and L. Hostetler, “k-nearest-neighbor Bayes-risk estimation,” *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 285–293, 1975.
- [49] C. Cortes and V. Vapnik, “Support-vector networks”, *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [50] R. Duda, Peter E. Hart, and D. Stork, *Pattern Classification*, New York: John Wiley & Sons, 2001.
- [51] R. Gutierrez–Osuna, PRISM Lecture Notes, Texas A&M University, 2008, URL: <http://research.cs.tamu.edu/prism/lectures.htm>
- [52] A. Fisher, *The Mathematical Theory of Probabilities*, Vol. 1, New York: Macmillan, 1923.
- [53] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, New York: Academic, 1990.

- [54] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York: Dover Publications, 1966.
- [55] E. Sanchez, S. Barro, and A. Canedo, “Edge detection and motion discrimination in the cuneate nucleus,” *Artificial Neural Networks*, vol. 2415, no. 12, pp. 198-203, 2002.
- [56] R. E. Fyffe, S. S. Cheema, and A. Rustioni, “Intracellular staining study of the feline cuneate nucleus. I. Terminal patterns of primary afferent fibers.” *Journal of Neurophysiology*, vol. 56. no. 5, pp. 1268–1283, 1986.
- [57] A. Webb, *Statistical Pattern Recognition*, Second Edition, Malvern, UK: John Wiley & Sons, 2002.
- [58] J.M. Jolion, P. Meer, and S. Bataouche, “Robust clustering with applications in computer vision”, *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 791-802, 1991.
- [59] A.J. Miller, *Subset Selection in Regression*, London, UK: Chapman & Hill, 1990.
- [60] K. Pearson, “On lines and planes of closest fit to systems of points in space,” *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.
- [61] N. Paragios and R. Deriche, “Geodesic active regions and level set methods for supervised texture segmentation”, *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

## VITA

Yoon H. Bai was born in Seoul, Korea. He received his B.S. in computer engineering in the Department of Electrical Engineering at the Texas A&M University in May of 2006. He continued pursuing computer engineering in the Department of Computer Science and received his M.S. in August 2008.

## Permanent Address:

Department of Computer Science  
Texas A&M University  
TAMU 3112  
College Station, TX 77843-3112

## Email Address:

yoombai@gmail.com

The typist for this thesis was Yoon H. Bai.