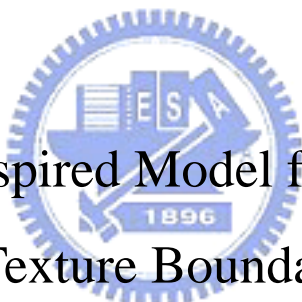# 國 立 交 通 大 學

電機與控制工程學系

碩 士 論 文

生物啟發之色彩與質感混合邊界偵測模型

Biologically-Inspired Model for Hybrid-Order Chromatic Texture Boundary Detection

研究生：蔡宗恆

指導教授：林進燈博士、

周志成博士

中華民國 九十四 年 七 月

生物啟發之色彩與質感混合邊界偵測模型

Biologically-Inspired Model for Hybrid-Order Chromatic Texture Boundary Detection

研 究 生：蔡宗恆   Student: Tsung-Heng Tsai

指導教授：林進燈博士、   Advisor: Dr. Chin-Teng Lin

周志成博士   Dr. Chi-Cheng Jou

國立交通大學

電機與控制工程學系

碩士論文

A Thesis

Submitted to Department of Electrical and Control Engineering

College of Engineering and Computer Science

National Chiao-Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master

in

Electrical and Control Engineering

July 2005

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 四 年 七 月

生物啟發之色彩與質感混合邊界偵測模型

研究生：蔡宗恆　　　　　　　　　　　指導教授：林進燈 博士
　　　　　　　　　　　　　　　　　　　　　　　周志成 博士

國立交通大學電機與控制工程研究所

中文摘要

　　本論文提出一個由生物觀點所啟發的多階次質感邊界偵測演算法。在演算法的發展階段，我們成功地整合了三個重要的視覺元素：明度、質感、色彩。現今相關研究的盲點在於僅從應用觀點出發，無法對整個議題做出完整的探討。有鑑於此，本論文對於人類視覺系統的基本運作模式乃至系統化的整合過程進行相關的基礎研究，針對質感邊界偵測的議題完成了通盤的探討。此外，在多階次的質感特徵萃取過程中，一些為過去研究所忽略但卻極其重要的議題，例如：偽反應（false response）的產生，權重值的選定等等，我們也做了徹底地討論並提出解決的方案。

　　人類視覺系統能夠有效率地處理視覺資訊的關鍵在於拮抗式的傳送機制，諸如接收域（receptive field）的組成方式以及對比色（opponent color）的形成。本論文參考視覺系統的編碼方式，並以系統化的方式建構出完整的質感邊界偵測演算法。輸入的彩色影像首先被解構為三組對比色軸，並經由高斯（Gaussian）濾波器以及賈伯（Gabor）濾波器萃取質感的一階特徵及二階特徵，輔以本文所提出之適應性權重值決定法則得到兩者對應邊界之權重值，我們可以結合出多階次的質感邊界。經由大量的測試結果，我們發現均勻質感之間的邊界都可以成功而精確地被標定，而對於較不規則或不均勻的質感圖形，演算法仍會找出一些符合我們人眼感受的特性。除了令人滿意的測試結果，本演算法的處理過程極為簡單且直觀，不需導入過多的假設以及任何的訓練過程。相較現有研究，本論文深具應用潛力。

Biologically-Inspired Model for Hybrid-Order Chromatic Texture Boundary Detection

Student: Tsung-Heng Tsai

Advisor: Dr. Chin-Teng Lin

Dr. Chi-Cheng Jou

Department of Electrical and Control Engineering
National Chiao-Tung University

ABSTRACT

In this thesis, a hybrid-order texture boundary detection technique inspired from human visual system (HVS) was presented. The proposed algorithm integrates three important visual primitives: luminance, texture, and color into a functional system. At present, the related works were developed for specific applications such that an overall investigation of the texture segregation process would be inaccessible. Therefore, the thesis focuses on relevant fundamental researches on HVS and systematic integration to investigate the task of texture boundary detection thoroughly. Moreover, some critical but ignored issues from the procedure of hybrid-order feature extraction, such as false response, weights selection, *etc.*, were also discussed and solved in this thesis.

Transmission with antagonism such as receptive field profile and opponent color is the critical point that HVS can effectively process visual information. This thesis employs the encoding form in HVS with systematic integration to build up a complete algorithm for texture boundary detection. Color images are firstly decomposed into three opponent axes and the 1st- and 2nd- order features are extracted by a Gaussian filter and Gabor filters. With the proposed adaptive weights selecting mechanism, the hybrid-order boundary can be obtained. Among extensive tests, boundaries between uniform textures can be detected successfully and accurately. For textures that are non-uniform or non-regular, the results also reflect some meaningful properties which are consistent to human visual sensation. In addition to satisfying testing results, processing employed in this algorithm is very simple and intuitive with only few assumptions and no training procedure involved. Compared with the present researches, the proposed algorithm has a good application potential.

誌謝

　　大學畢業時的印象仍鮮明，轉眼兩年已過，即將要步出交大的校園。首先要感謝的是我的指導教授－林進燈老師與周志成老師兩年來在研究上給予的指導，在我面臨瓶頸的時候，也往往提供非常寶貴而正確的方向。另外要特別感謝的是清大動機的彭明輝教授，當年不厭其煩地引導我們思索生活的態度，以及嚴謹的研究方式。這兩年來不斷反思，至今仍是受益良多。

　　感謝兩年碩士生涯陪著我的同學們，不管是實驗室的學長學弟與同窗，或是高中及大學的朋友們。認識你們帶給我太多的鼓勵與歡樂，互吐苦水、漏氣求進步、吃喝玩樂，讓我在繁重的研究之餘，仍有多采多姿且不凡的生活。

　　家人長年對我的付出，不管是實質上或是心理上的支持，往往是我平日最容易忽略掉的一環。感恩為我們教育傷透腦筋的爸爸，總是鼓勵代替一切的媽媽，工作繁忙卻時時關心我的大姐，以及二姐與哥哥越洋傳達的關懷。面對你們凌駕一切的親情，我想再次向你們說聲謝謝！

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATION AND ABBREVIATIONS

R	real numbers
C	complex numbers
exp	Euler's number
j	imaginary unit, $j = \sqrt{-1}$
(x, y)	space axes
(u, v)	frequency axes
θ	orientation
Δx, Δy	spatial resolution in *x* and *y* axes
Δu, Δv	frequency resolution in *u* and *v* axes
$\Delta\theta$	orientation resolution
g(x)	one-dimensional Gaussian function
G(u)	Fourier transform of one-dimensional Gaussian function
h(x)	one-dimensional Gabor function
H(u)	Fourier transform of one-dimensional Gabor function
g(x, y)	two-dimensional Gaussian function
σ_x, σ_y	std of two-dimensional Gaussian function in *x* and *y* axis
σ_u, σ_v	std of two-dimensional Gaussian function in *u* and *v* axis
h(x, y)	two-dimensional Gabor function
U	center frequency of two dimensional Gabor function in *u* axis
V	center frequency of two dimensional Gabor function in *v* axis
F	magnitude of center frequency of two-dimensional Gabor function
ϕ	phase of center frequency of two-dimensional Gabor function
$h_c(x, y)$	real part of Gabor function
$h_s(x, y)$	imaginary part of Gabor function
$h'(x, y)$	mother Gabor wavelet
S	number of selected frequencies in Gabor wavelets scheme
K	number of selected orientations in Gabor wavelets scheme
m	index of center frequencies in Gabor wavelets scheme
n	index of orientations in Gabor wavelets scheme
U_l	lower center frequency
U_h	upper center frequency
σ_g	std of post Gaussian filter
LGN	lateral geniculate nucleus

V1	primary visual cortex
RF	receptive field
CSF	contrast sensitivity function
FIT	feature integration theory
CNN	cellular nonlinear/neural networks
VLSI	very large scale integration

Chapter 1

Introduction

1.1 Motivation

Vision is one of the most important senses of human beings. By vision, we can, for example, locate a cup of tea or coffee to reach for, recognize a familiar person's face, and enjoy a majestic sight or a masterful art work. Our experiences from surroundings are mainly originated by our visual systems. Various visual tasks can be done immediately and effortlessly, almost without consciousness of any nerves. We never realize how hard these tasks are until sometimes attempting to make use of our visual properties in some applications of computer vision or image processing. The critical problem in these cases is object segmentation. The first stage of many engineering applications is to segment objects from background first and then apply corresponding operations for objects and background. It seems very easy and intuitive to segregate objects from background for us in our experiences. However, there is still no implementation dealing with the task well due to limited hardware devices and inappropriate software processing procedure.

Human visual system is capable of integrating basic primitives to form our rich sensation and perception, despite the fact that visual information is discretely sampled by the retina and cortex. Both biological and computational evidences have suggested that some kinds of feature extraction and data compression occur at very early stages in visual processing. How human visual system extracts features and compresses information in the stages is very ingenious and widely investigated. However, the

1

application-oriented researches seldom made good use of rich evidences from fundamental researches about human visual system. For many application tasks with criterions which are based on how human sees or how much they are like to human being, it is a reasonable approach analyzing visual information from biological point of view. Thought, in actual, some researchers approached related problems by procedures replicating human behavior and many important visual primitives, such as luminance, color, texture, and so on. However, how to integrate these primitives well to describe visual behavior is still a difficult problem. In this thesis, we will propose a novel approach which mimics the early stages of human vision integrating hybrid-order features.

1.2 Problem Statements

Early vision, also called as preattentive-stage vision, includes those mechanisms that subserve the first stages of visual processing. These mechanisms operate in parallel across the visual field, and are believed to be used for extracting the basic visual primitives. These primitives detected or extracted during early visual stages constitute our sensation and perception, and the following level called cognition combines meaningful features and compares them to patterns in our memory. In other words, the sensation/perception stage faithful reports basic visual primitives and the cognition stage combines these primitives to shape various forms upon needs. The cognition involving attentions can not yet be described and modeled briefly, such that it is unsuitable to represent the stage under unclear goal or assumption of applications. Thus, it is better to formulate a framework investigating sensation/perception stage thoroughly first and then adapt the framework for different purposes.

Empirical sciences such as psychophysics and psychology have strongly advanced our knowledge of the underlying visual processes. Also, a computational

approach is helpful to integrate multiple findings into a common framework which is accessible to be analyzed and evaluated. As a biology-inspired approach, of course, the criterion for evaluating the proposed algorithm is undoubtedly to verify if the results are consistent to how human visual system works. Take the textural pairs in Fig. 1-1(a) and Fig. 1-1(b) for example [1]. Both the images consist of two regions and each of which is made up of distinct textural tokens. The fact is obvious in Fig. 1-1(a), while a close scrutiny is necessary to observe it in Fig. 1-1(b). By immediate examination of Fig. 1-1(b), it does not result in the perception of two different textured regions, but only one uniformly textured region instead. The textural pair in Fig. 1-1(a) is called preattentively distinguishable, while the one in Fig. 1-1(b) is called preattentively indistinguishable. As the proposed algorithm is developed for considering how the human visual system processes texture at early stages, consistency between preattentive vision and results by the algorithm is a rough but fair criterion for the algorithm. We will introduce preattentive vision more detailed in Section 2.4.

(a)   (b)
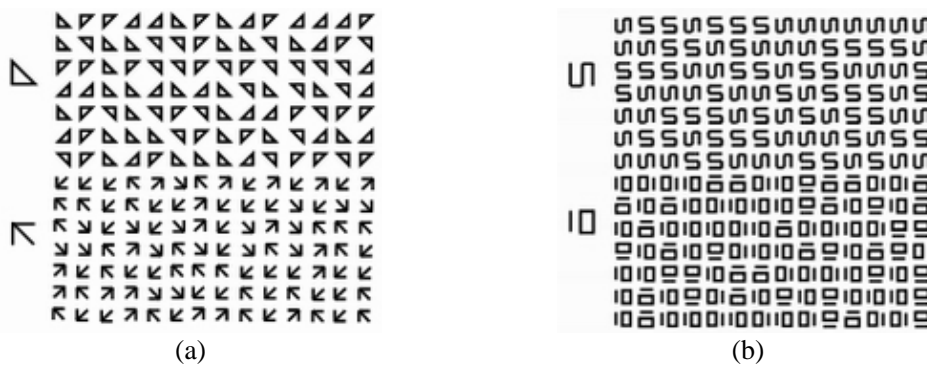
Figure 1-1: Two images made up of (a) preattentively distinguishable patterns; (b) preattentively indistinguishable patterns. The figure is adapted from [1].

Evaluation of Computational Approaches

Besides the experimental results, to tell if a computational approach describes visual behavior well, we should first consider how to evaluate it. In the well-known book "*Vision*" [2], Marr elaborated a judgment on computational approaches. As an

overall framework for specific visual task is formulated, there must be an attempt to describe some phenomenon. Marr proposed that a complete understanding of a visual process would involve explanations at three levels including: (a) the computational theory; (b) the representation and algorithm; and (c) the hardware implementation. At the first level, abstract of the computational approach should be characterized well as a mapping from one kind of information to another. Also, the appropriateness and adequacy for the task at hand should be demonstrated at this level. At the second level, it should characterize well how the computational approach can be implemented, in particular, representation of the input and output, and algorithm for the transform. At the last level, details of how the algorithm and representation are realized physically have to be described. In other words, the detailed hardware architecture needs to be explained well. Such a three-level explanation provides us a practical evaluation for computational approaches illustrating visual behavior. For a complete and realizable description about visual processing, the relevant computational approaches should meet the requirements at all three levels.

1.3 Related Works

1.3.1 Texture Analysis

In a static image, an essential primitive besides luminance is texture. Its property among different illumination conditions seldom changes as acutely as luminance does. Thus texture information can help to prevent many erroneous estimations resulting from illumination variation. Moreover, texture reveals some surface properties of objects which are useful to separate an object from background or other objects. In this section, a brief review about some important theories and properties of texture are given.

Definition of Texture

Texture, compared with other visual primitives such as color, luminance, stereo, *etc.*, is a relatively new one being discussed with. The researches about texture began about 1960s. Before we start to talk about texture, we should give a definition to texture first as to any other tasks. Unfortunately, even though the property of texture is too essential to be neglected, there has not been a clear and acknowledged definition to texture. A definition to texture in Webster's *New World Dictionary* is: *"The character of the woven fabric resulting from the arrangement, size, quality, etc. of the fabric's threads: as, a fine or coarse texture, a ribbed or twilled texture."* We can immediately imagine what texture is by the definition. However, a more qualitative or quantitative description is still inaccessible. The inaccessibility to a precise definition mainly results from two reasons. Firstly, texture is one kind of perception during early vision stages. It involves not only physical factors of stimuli but how the brain reacts to these stimuli. Thus it cannot be easily defined at physical level. The other reason is that there are large amounts of factors involved in texture perception and there exist lots of nonlinearities in the interactions among them. The two reasons above make a precise and identical definition to texture (especially, a physical meaning) hardly to be given so far [3]. Therefore, the "definition" of texture is usually formulated depending on specific application and there is no general agreement on definition.

In this thesis we referred to some perceptually motivated descriptions [4]-[6] and refined the state. Our definition can be given as follows: (a) Texture is characterized by properties of a local region and there should be adequate spatial-relationships between elements or primitives within the region. In this thesis spatial-relationships simply mean the orientation and spatial frequency. (b) The homogeneous texture discussed in this thesis means that there are similar features over single textured pattern, and the scale- and rotation- invariant issues are not considered in this thesis.

1.3.2 Theories of Texture Perception

In advances of texture theories, two major contributors: Bela Julesz and Jacob Beck, had developed two main trends of texture perception called the Nth- order statistics theory [7] and the texture element theory [8]. For completeness of this thesis, it is worth reviewing their theories first to facilitate the understanding of the theories nowadays before we discuss present theories.

Julesz et al. [7]

As literal descriptions about textures above, it is believed that the statistical distribution is suitable to model the spatial arrangement of texture. In Julesz's approach [7], [9], he tried to find the highest statistical distribution which human visual system can still discriminate and proposed: In most cases, the determination for whether human visual system can distinguish two textured patterns or not is likely the global 2nd- order correlation function. Examine the approach by signal analyzing method, Fourier transform of the global 2nd- order correlation function is the power spectrum reporting the frequency components only. Thus, considering the power spectrum without taking local information into account could not observe some spatial arrangements making up textural properties. Therefore, considering the global 2nd- order distribution merely would inevitably meet counterexamples refuting the necessity [9], [10] and sufficiency [11] of the global 2nd- order distribution in the approach.

Beck et al. [8]

From another point of view, Beck focused on the grouping phenomenon of textural elements to formulate another approach for texture perception [8]. As an advocate for Gestalt psychology, he held the main announcement: *"The whole is more than the sum of its parts."* to analyze the process of texture perception. Compared

with the approach by Julesz, the texture element theory by Beck is more similar to low-level organic operation in human visual system which employs the bottom-up procedure to construct the approach step by step. The grouping elements are usually basic visual primitives rather than cognitive forms (Fig. 1-2). As described in [12], for local features, such as orientation, contrast, size, closure, *etc.*, which could not be described well by global statistical properties, the bottom-up procedure would provide a better analyzing approach. However, when more local features were utilized jointly for experiment and discussion, the incompleteness of the theory appeared. It was found that the grouping form is not merely by summing all local feature effects and the illusory conjunction phenomenon would occur in some case [13] as shown in Fig. 1-3. The problem is due to that receptive fields in human visual system are not merely feature detectors (it will be discussed thoroughly later) and an approach would truly characterize our perception for texture only if it is developed while referring to the basic information processing in human visual system.
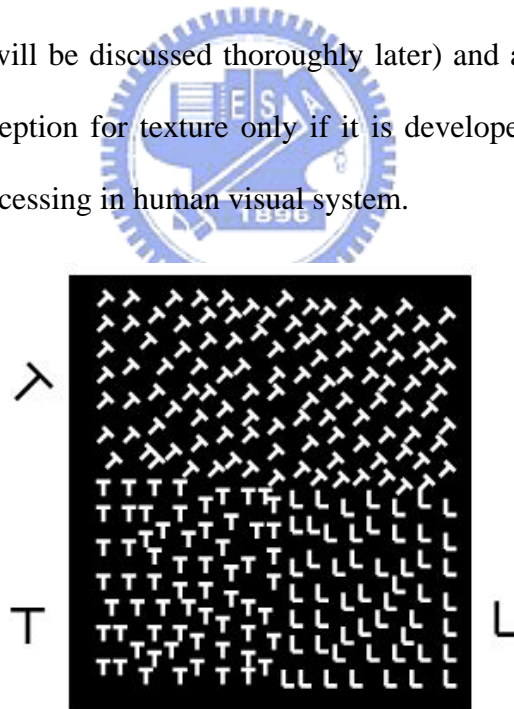
Figure 1-2: Regional perception results from some kinds of grouping properties. For texture perception, the basic primitive (orientation, in this case) plays a more important role than cognitive form (L or T, in this case).The figure is adapted from [8].
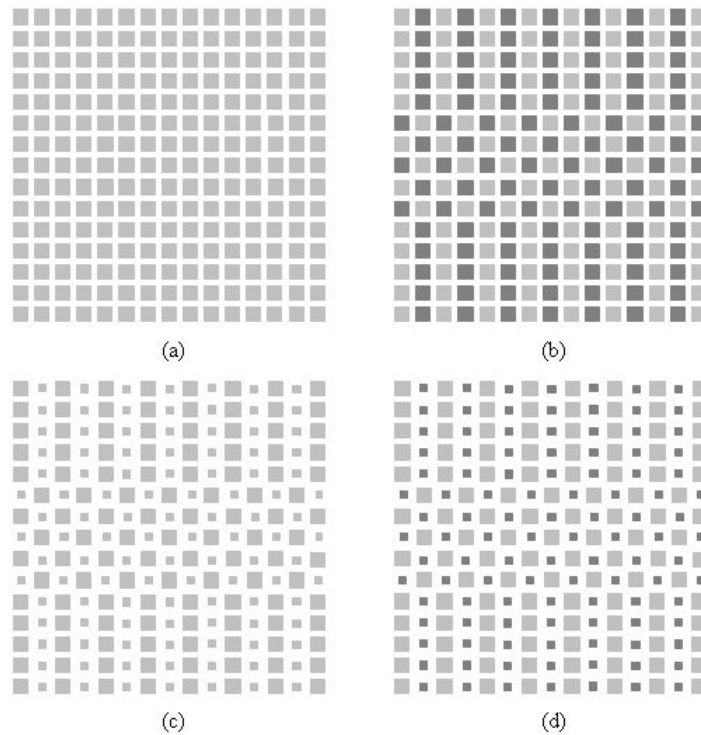
Figure 1-3: Three regions could be easily segmented by controlling factors (b) the contrast to background; and (c) the size. However, the degree of discrimination does not increase but decrease as both factors are applied simultaneously (d). The figure is adapted from [13].

Examine the two theories by Marr's three-level explanation about computational approach. Though both approaches define the process and goal more explicitly, they does not give clear descriptions about algorithm and hardware implement at upper two levels, e.g. how to extract the defined features. There was no approach completely describing all the three levels: computational theory; representation and algorithm; and hardware implementation until the linear filtering theory.

1.3.3 Linear Filtering Theory

As the texture element theory approached the texture perception subject by local feature description, it led to the critical issue considering how local features could be extracted quantitatively and fairly. To specify local features more accurately, literal meanings of texture are not sufficient. Instead, it would be more adequate to

investigate how neurons managing visual information "see" an image. The processes in visual pathway have been verified as spatial-frequency filtering and formulated by general functions, e.g. Gaussian function and Gabor function [14]. The receptive field profile at pre-cortical stage which has been regarded as the main source to our luminance sensation can be fitted well by Laplacian of Gaussian (LoG) function [15]. Also, the receptive field profile in visual cortex, which is considered the main source to texture perception from its orientation selective and spatial-frequency sensitive properties, was verified to be fitted well by two-dimensional Gabor function [16], [17]. The two-dimensional Gabor function has been regarded as a spatial-frequency filtering mechanism for texture perception. Moreover, the Gabor-based computational model for texture perception has been examined and discussed entirely. Here we will discuss some related approaches by Gabor filtering scheme. Details about human visual system and our modeling strategy will be further discussed in Chapter 2 and Chapter 3 respectively.

The overall process of texture feature extraction can be divided into three parts: Gabor filtering, rectification, and Gaussian smoothing [18]-[21]. Among them, Gabor filtering is the critical part, many researches focused on the topics in last ten years. We can group them into two categories: filter-design approaches, where the Gabor filters are designed for specific tasks; and filter-bank approaches, where the Gabor filters are selected from predominated partitions in the frequency domain.

Filter-Bank Approaches

For a quantitative description about literal meaning for texture properties such as fine, coarse, vertical-orienting, *etc.*, the Gabor function with profile narrowly tuned to specific orientation and spatial frequency is a very suitable candidate. Daugman [16] firstly present a framework composed of multiple filters with specific orientations and

spatial frequencies. Other similar frameworks were proposed by Turner [22] and Malik and Perona [23] for psychophysical verification and by Jain and Farrokhnia [24], Manjunath and Chellapa [25], and Manjunath and Ma [26] for application tasks. How the framework woks is like a feature vector transform that provide a nearly complete coverage of the spatial-frequency domain. Besides, corresponding to the spatial-frequency property in cortex, the multiresolution filtering theory [27] was utilized for a more compact representation. Sometimes we call the set of Gabor filters with octave-bandwidth and octave-frequency profiles as Gabor wavelets.

In these approaches, the frequency domain is divided by predominated filters that are not necessarily optimal for a specific task. The number of filters would sometimes result in huge computational loads. Furthermore, the filter vector with large dimensions may lead the undesirable phenomenon: "curse of dimensionality." Though approaches by Unser and Eden [28] and Tang *et al.* [29] reduced the number of filters with feature refining procedures, their methods focused on cases with only two textures. Also, procedure of feature refinement implies the feature properties would become unobservable and hard to analyze. Furthermore, the optimization involved in such scheme would be less flexible than the filter-design approaches presented below.

Filter-Design Approaches

To cope with drawbacks of filter-bank approaches, many researches focused on how to select an optimal filter set to faithfully obtain useful information and discard other meaningless information. Usually, these approaches need a training procedure or database set in advance. Some typical ones are listed below.

Bovic *et al.* [19] gave a very detailed analysis of texture perception. They proposed three supervised procedures to select center frequencies of filters by using empirical information based on the power spectrum characteristics of individual

textures. A similar but unsupervised approach was proposed by Tan [30]. It detects a global spectral peak once, and repeatedly detects conspicuous peaks with erasing operation on the frequency plane. That is, the power spectrum of a small neighborhood (e.g. 5×5) around the detected peak would be set to zero. The iterative peak detecting procedure terminates if the ratio of the current peak to the highest peak is less than a pre-specified value. Another filter-design approach developed by Dunn and Higgins [31] employed a procedure for designing a single filter to segment two textures. This is a totally supervised approach measuring textural statistics by a Rician statistical model and then using the statistics as a predictor for segmentation error. A further study by Weldon and Higgins [32] generalized the approach for multiple textures and thoroughly discussed the trade-off between classification and localization. However, since the procedure requires some assumptions about textural distribution and empirical values during filter selection, and databases of textural statistics also have to be built in advance, practicability of the approach is highly limited. Teuner *et al.* [33] proposed another approach to select the optimal filter set by an iterative pyramid Gabor representation. They indicated that human visual system does not direct its attention simply to features which occur frequently but ones which stand out significantly. Therefore, instead of selecting spectral peaks, they defined a spectral feature contrast criterion with progressive dyadic stage to select center frequencies of filters and the reciprocals of spectral feature contrast values could also be used as weights for extracted features [34]. The approach has fairly represented the procedure of texture processing. However, the efficiency and bandwidth selection are still limited due to the pyramid Gabor representation [35], [36].

For some purposes, filter-design approaches indeed showed better performance with less on-line computations. However, for developing the filter selecting procedure, more assumptions and off-line computations were required. Moreover, these

approaches were optimized for pure textured image that many factors were not taken into consideration, e.g. ambiguous luminance variation, which would result in large amounts of spectral components. That is, for more general purposes, even a complete filter-design approach like the one by Teuner *et al.* [33] could not perform well due to many uncertainties. It is better to extract textural information by more original procedure faithfully acquiring information without ambiguous one, and the filter-bank approach is a more suitable choice.

1.3.4 Chromatic Texture

A summary of texture segregation by Bergen [20] *"Texture based segregation is a perceptual phenomenon in which regions differing only in their spatial structure, not in color or brightness, and without any physical contour segregating them."* revealed that researchers seldom considered issues about color contrast [37] and color was usually employed as a regional primitive like luminance, that discussions about texture mainly focused on gray-scale images. At present, there are still few relevant researches about chromatic texture as follows.

Papathomas *et al.* [38] first built a computational model employing the concept of double opponency and used a psychophysical procedure to verify the responses between simulation and experiment. A similar approach by Jain and Healey [39] combined the filter-bank procedure and opponent color feature to develop a model for chromatic texture recognition. From another point of view, Mirmehdi and Petrou [40] generated a multiscale representation by a multiband smoothing algorithm cooperating with 3D histograms and probabilistic assignments to analyze the interaction between color and spatial frequency. Recently, an approach by Wanderley and Fisher [41] presented a feature set by color angles to deal with the illumination invariant issues.

These models attempted to implement the multiple chromatic spatial frequency channels [42]-[44] by computational approaches. However, some issues were not considered well and these approaches seemed not describe the phenomenon properly. First of all, opponent chromatic features were usually represented over-completely. There are some basic odds such as resolution and sensitive spatial frequency of luminance and chromaticity that it is inappropriate to extract luminance texture and chromatic texture by the same scheme. In addition to, the color decomposition employed did not correspond well to the opponent-process theory [45] which was considered as color delivering form at neuron stage, that it could not correctly model the chromatic contrast. At last, these approaches focused on specific issues and lacked some feasibilities in common use, e.g. only one sensitivity frequency considered in the approach by Papathomas *et al.* [38], that impracticability and issues of curse of dimensionality (at least three times to filter-bank features) would be inevitable.

Along the summary above, we would like to emphasize our purpose again: To develop a computational approach for general uses; that is, to extract meaningful features while preserving analyzable essences. In Chapter 3, we will build the scheme from the fundamental concept on chromatic texture.

1.4 Research Scope

At present, as we know there is no definitive model for dealing with the 1st- and 2nd- order information simultaneously for chromatic texture boundary detection, where the 1st- order information describes the global content; and the 2nd- order information describes the local content within regions. In this thesis, a computational approach for hybrid-order texture segregation will be proposed. We focused on mimicking the preattentive stage of visual perception, and thus there will be no clustering or classification procedure. In order to overcome insufficiency of only

considering single order feature, we integrated the 1st- and 2nd- order features into a functional system. Three important visual primitives: luminance, texture, and color were combined properly and adaptively. By the evaluation of Marr's three-level description, the proposed model can describe the texture segregation task well at all three levels. This thesis will mainly present the goal and development of algorithm for texture segregation, and the hardware implementation can be found in another research [46] by our colleague.

1.5 Outline of the Thesis

To achieve the goals described in Section 1.4, this thesis is organized as follows:

Chapter 2 introduces the knowledge from physiology and psychophysics about vision. Some acknowledged evidences of mechanisms of human visual system which reveal effective visual processing procedure will be employed during our algorithm development.

Chapter 3 proposes the modeling strategy in this thesis. Three primitives: luminance, texture, and color can be combined into a functional system properly and adaptively. Some ignored issues during individual modeling for specific goals will also be described well and solved for the hybrid-order scheme.

Chapter 4 gives a large number of experimental results and discussions among them. Discussions and comparisons for considered issues in Chapter 3 will also be present.

Chapter 5 concludes the innovations and contributions of this thesis and gives suggestions for future researches.

Chapter 2

Knowledge about Human Visual System

Human visual system is a powerful and elaborate system that is capable of extracting features and integrating them effectively. From physiological and psychophysical findings, there are large amounts of evidences revealing that human visual system carries out the task at its early stages [47], [48]. The initial stages of visual processing are very important in respect of detecting and grouping various types of visual primitives, such as curvature, line orientation, color, spatial frequency, *etc*. In this chapter, some important knowledge about vision will be introduced, and the modeling strategy to be introduced in the next chapter is based on these evidences to develop the overall framework.

2.1 Anatomical Structure of Human Visual System

2.1.1 The Visual Pathway
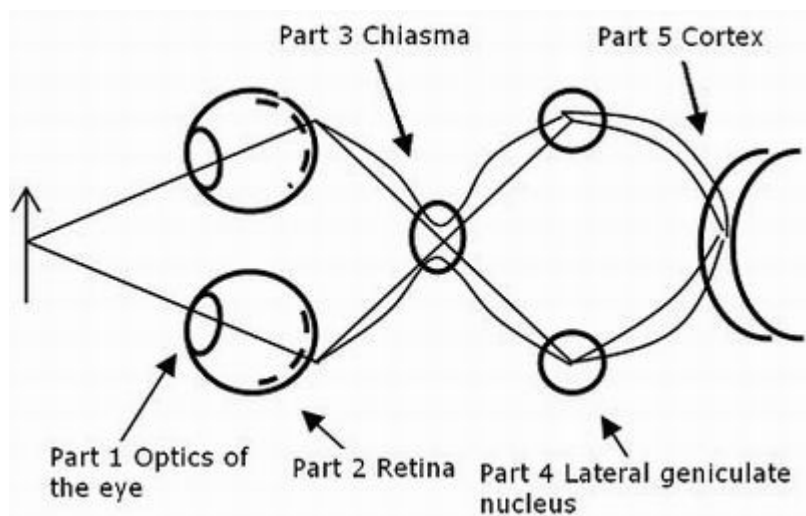
Figure 2-1: Schematic diagram of human visual system model.

Fig. 2-1 shows a probable model of human visual system at early stages. Part 1 is the optics of eye including cornea and lens which focus a scene onto the retina (part 2) where many receptors (cones and rods) spread over. The properties and uses of two types of receptors are quite different: rods, which are about 10 times as sensitive as cones, are the only functioning receptors at very low light levels. Cones, on the other hand, do not respond to dim light but be responsible for our ability to see fine details and for our color perception. There are three types of cones and each of them is categorized by the wavelength sensitivity. It is well believed that color perception is originated from the differences in wavelength selectivity of the three types of cones. After a scene is projected on retina, the receptors will translate the light signals into neuro-electrical signals (the process is usually called transduction) and transmit them from back to front of eyes. Passing through layers of horizontal cells, bipolar cells, and amacrine cells, visual signals then arrive the layer of ganglion cells whose axons pass across the surface of the retina, collect in a bundle, and leave the eye to form the optic nerves. The optic nerves of two eyes join and split in the optic chiasma (part 3) and then reach the central part of visual information, called lateral geniculate nucleus (LGN, part 4). In LGN, visual information is divided into two pathways: mangocellular pathway mainly dealing with motion perception and spatial information; and parvocellular pathway mainly dealing with color, shape, texture, *etc.* Next to LGN, the visual information is delivered to the striate cortex (V1, part 5). The numbers of cells in V1 (about 250 million) are much more than ones in LGN (about 1 million) and the functions are more complex. V1 preserves the most precise topography map in cortex. After some basic processes in V1, various kinds of visual information are delivered to corresponding pathways for further processes. Roughly speaking, the visual pathway from retina to V1 is usually called the early vision level where large amounts of previous analyses and investigations focus on. Most of our

computational model was developed based on biological evidences at this level.

2.1.2 Receptive Fields in Visual Pathway

In retina, between the layer of receptors and the layer of ganglion cells, there are three types of nerve cells: bipolar cells, horizontal cells, and amacrine cells. Among the cells, receptive fields of them reveal most direct messages about their functional processing. The region where specific receptors feed into a given sensory neuron is usually called the cell's receptive field (RF). Receptive fields have a substructure that stimulating different parts of the receptive fields will give different responses qualitatively and quantitatively thus two similar receptors in a cell's receptive field might feed onto the cell diversely due to their spatial positions in the receptive field. Besides, stimulating a large area will result in cancellation from the subdivisions rather than summation. The antagonism can be found in kinds of our sensory systems to avoid ambiguous sensations. The physical brightness amount of the "black" word in sunlight is more than the amount of "white" paper at low light levels. However, we never feel difficulty to discriminate the white paper and black word printed on. The discrimination can not be fulfilled by recording absolute information but relative one. The main concept of receptive field is not only the connection but also the opponent form, and it is obvious that to understand functions in visual pathway, we should refer to receptive fields of nerve cells in visual pathway in advance.

The bipolar cells occupy a strategic position in the retina, since all signals originated from the receptors then transmitted toward the ganglion cells must pass through them. Visual signals are delivered from receptors to bipolar cells in two separate paths: a direct path where the receptors synapse onto the bipolar; and an indirect path where the receptors contact the horizontal cells which in turn synapse onto the bipolar cell. That is, each bipolar cell is connected to receptors in a two-path

form. In the direct path, a bipolar cell obtains inputs from the receptors in a circle-shape area of retina; from the indirect connections via horizontal cells the bipolar cell receives inputs from a larger, overlapping, and concentric disk-shape area (Fig. 2-2). The two paths deliver opposite tendency to the bipolar cell. That is, for identical stimuli, one path will deliver excitatory response while the other will deliver inhibitive one. The substructure of bipolar cell's receptive field is called as the opponent center-surround mechanism. Similarly, the ganglion cell's receptive field also has such substructure, and in actual, the opponent center-surround mechanism was firstly discovered in ganglion cell [49], [50]. The layer of ganglion cells is the last stage of visual signals in eyes (the output of eyes); that is, the center-surround antagonism provides a preliminary understanding of processes in retina and helps the exploration for higher level visual processing.
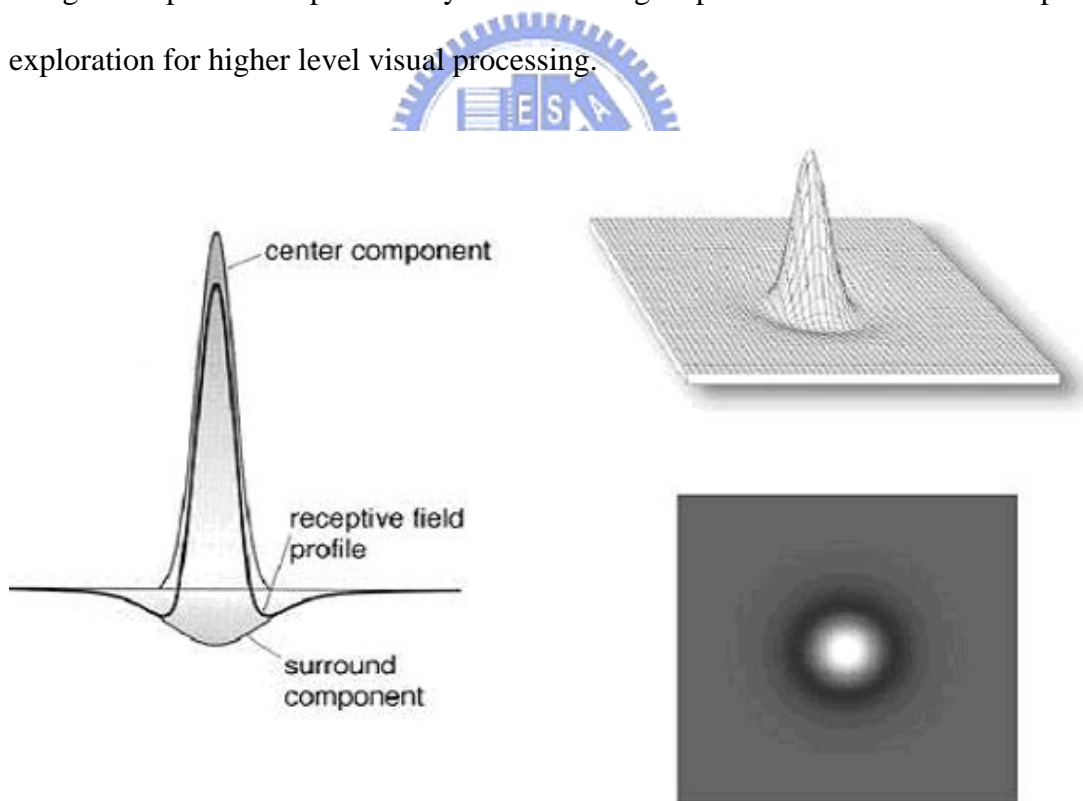
Figure 2-2: Schematic diagram of receptive field of the bipolar cell.

From retina to posterior stages in visual pathway, main processing for spatial vision task does not change until V1. On the whole, the LGN is like a relay station for

visual signals, and receptive field profile maintains the form in retina. As visual

signals are delivered to V1, the receptive field appears more elaborate properties. V1

cells have several characteristics not seen earlier: binocularity, direction selectivity,

and much narrow orientation and spatial frequency selectivity. The pioneers in this

field are Hubel and Wiesel (Nobel Prize, 1981). They discovered most V1 cells do not

respond to isotropic stimuli (e.g. point) but to specific line stimuli [51]-[53].

According to responses to various types of stimuli, Hubel and Wiesel classified V1

cells into two categories: simple cells and complex cells. Simple cells respond most to

stimuli with specific preferred orientations. Receptive fields of simple cells, like cells

in retina, also have excitatory region and inhibitive region. There are two types of

simple cells determined by the arrangement of excitatory region and inhibitive region.

One is organized with inhibition- excitation- inhibition arrangement (even symmetric)

while the other is organized with excitation- inhibition arrangement (odd symmetric)

as shown in Fig. 2-3. The other category of V1 cells, complex cells, respond most to

line segments with specific orientations moving along specific directions. In V1 and

prior stages, the receptive fields of most neuron cells appear simple profiles and can

be easily understood. Still, there seems not an integral analysis considering the

properties more thoroughly. The next section to be introduced is the main integrating

theory: linear filtering theory.
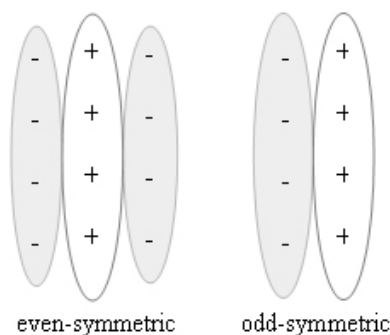
Figure 2-3: Schematic diagram of receptive field of the V1 cells.

2.2 Linear Filtering Theory

As more and more evidences from physiological and psychophysical experiments about receptive fields were recorded, many researchers attempted to further discover the fundamental processing of visual information. Among them, linear filtering theory [47], [54], [55] has made a great impact on recent visual researches. Spectral analyzing procedure is the principal concept of linear filtering theory. Researchers in the field indicated that recording visual information by spatial frequency decomposition is a more efficient way and large amounts of experiments have been performed to verify the argument. We would like to point out that the concept of linear filtering theory is a little bit different to the ideal Fourier analysis because we cannot exactly calculate frequency components in a spatially delimited image. Moreover, no neuron cell has a receptive field extending to unlimited region; that is, information from a receptive field can reach neither the most precise frequency resolution nor spatial resolution. The issue can be explained more thoroughly by introducing another important concept: multiple spatial frequency channels.

The receptive field of a unit at pre-cortical stages is possessed of center-surround antagonism which can be interpreted as a band-pass filter extracting a specific range of spatial frequencies. The contrast sensitivity function (CSF) of human visual system supports the assumption that the CSF attenuates at low and high frequencies (Fig. 2-4). Until the late 1960s, it was assumed that all ganglion cells have the same broad sensitivity profile as the CSF. In 1968, however, Campbell and Robson [54] made a revolutionary suggestion that the visual system might contain a group of independent, band-pass filters, which are narrowly tuned for ranges of frequencies (Fig. 2-4). In other words, human visual system does not employ a single mechanism to deal with all spatial frequencies but a group of mechanisms, and each of them is responsive to

only some fraction of total range.
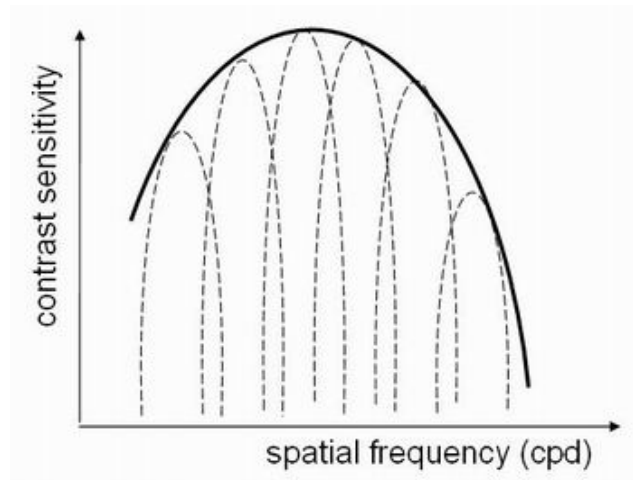
Figure 2-4: The overall luminance contrast sensitivity function (CSF) consists of multiple spatial frequency channels.

The assumption was soon supported and verified by many physiological and psychophysical evidences. Anatomical records showed that there exist cells with different sizes of receptive fields corresponding to different sensitivity frequencies; and responses of cells measured by micro-electrode also showed much narrower sensitivities than the overall CSF. In addition, in many psychophysical experiments such as pattern adaptation, frequency masking, subthreshold summation, *etc.* [56], stimulus at specific frequency did not result in an overall effect on CSF but a local influence near the stimulating frequency that supported the assumption about multiple spatial frequency channels as well. These evidences all revealed that there is no truly Fourier analyzer in human visual system, and of course, what vision system functions is not global analysis that requires extremely narrow channels but a group of channels operating spatial-frequency filtering. Besides these evidences, from the viewpoint of signal analysis, it is more economical and suitable to represent the contents on surroundings (e.g. objects, illuminations, *etc.*) by local spatial frequency filtering. In human visual system, such phenomenon appears in ganglion cells, LGN cells, and V1 cells. An image in V1 is decomposed into not only spatial frequencies but also

orientations. A schematic model of columnar organization in V1 shown in Fig. 2-5 represents that various two-dimensional spatial frequencies are considered to be in a polar arrangement, with spatial frequencies increasing from the center. By choosing appropriate basis, the organization can be represented well by spatial-frequency analysis, e.g. wavelets transform [27]. In actual, it had been verified that Gabor function [17], [57] could fit well the receptive field profile of V1 cell and a two-dimensional Gabor representation could also characterize an image completely [35], [36]. We will discuss the Gabor function more detailed in Chapter 3.
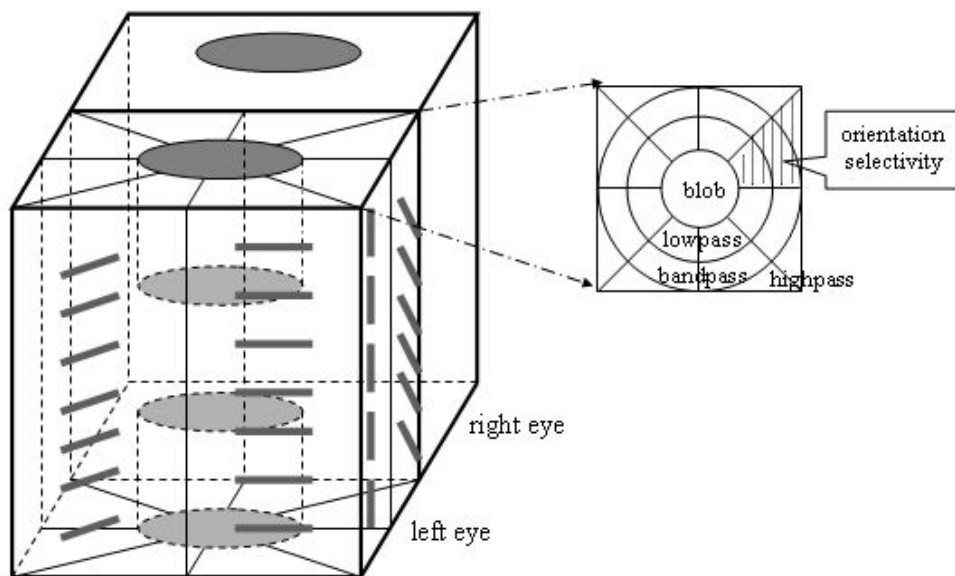
Figure 2-5: Schematic diagram of columnar organization in V1 (adapted from [47]).

2.3 Color Vision

The backgrounds described above mainly discussed the luminance content in an image, and in fact, those biological evidences were derived from experiments with gray-scale images. For color images, the processing complexities do not merely add one feature. Here I will briefly introduce the color perception and issues about spatial vision from color.

Like texture, color is not a physical quantity but perception. As described in

Section 2.1, there are three types of cones categorized by their wavelength sensitivity. We usually call the three types of cones as L-cone, M-cone, and S-cone (Long-, Median-, and Short- wavelength sensitive). Unlike luminance information, which directly corresponds to responses of receptors, color information is a manufactured output of visual processes that responses of three types of cones are integrated within a region and delivered. Vision system can roughly analyze the content of perceived spectrum to make up our color perception. The opponent-process theory [45] proposed by Hering is a very important theory characterizing color information processes at neuron stages. Like luminance representation from receptive fields with antagonism, vision system transmits color information by a similar way. After the layer of receptors, color information is delivered in three opponent channels including red-to-green channel, blue-to-yellow channel (two chromatic channels), and white-to-black channel (one luminance channel). Such representation was directly supported by records in LGN cells [58] and some investigations on complementary colors. Like receptive field profiles, under limited amounts of neurons and nerves, the opponent representation could provide a more economical and robust transmission.
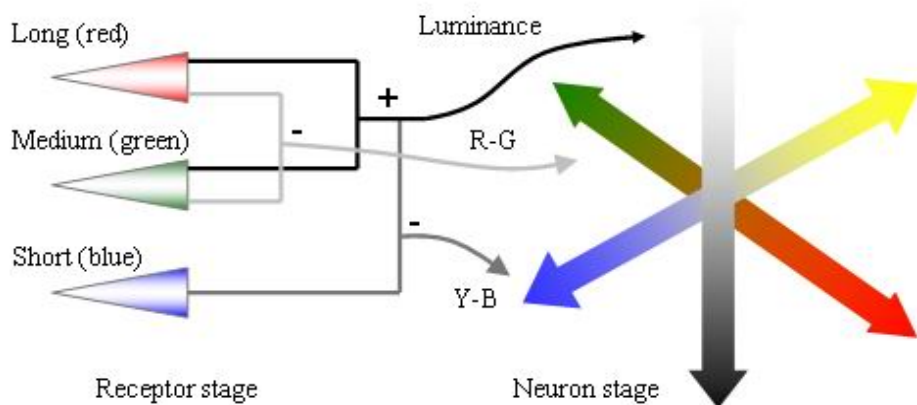
Figure 2-6: Representation of Hering's opponent-process theory.

Here comes another issue, at neuron stages, color information is encoded in one

luminance and two chromatic channels, and the luminance is represented in a group of spatial frequency channels. Are there similar mechanisms in the other two chromatic channels? Fig. 2-7 is the schematic plot of overall CSF of chromatic and luminance stimuli. Compared with luminance CSF, sensitivity range of chromatic CSF tends toward lower frequency and there is no significant attenuation at low frequencies. Also, the sensitivity amounts are less than those in luminance CSF. Texture discrimination had been seldom attributed to color information for two reasons: (a) Chromatic features are extracted within regions that inevitably lead to coarser resolution and lower sensitivity frequency in chromatic channels. (b) Due to transmission form, cells for chromatic information are possessed of opponent mechanisms containing excitatory and inhibitive wavelength ranges, that the contrast range in chromatic channel is more limited than the range in luminance channel. Some researchers [59] even asserted that color information provides nothing for texture discrimination. In fact, as long as chromatic stimuli are manipulated within proper bandwidth and range, cells for chromatic information still preserve operations for texture perception. From some experiments with isoluminant stimuli [42]-[44], chromatic information also revealed the representation of spatial frequency channels as luminance information. Moreover, records of V1 cells revealed orientation selectivity for pure chromatic stimuli. That is, except for some basic odds of sensitive frequency and resolution, the multiple spatial-frequencies filtering scheme can describe well all three opponent channels. The economical and significant visual processing scheme reveals that an efficient and simple implementation is possible.
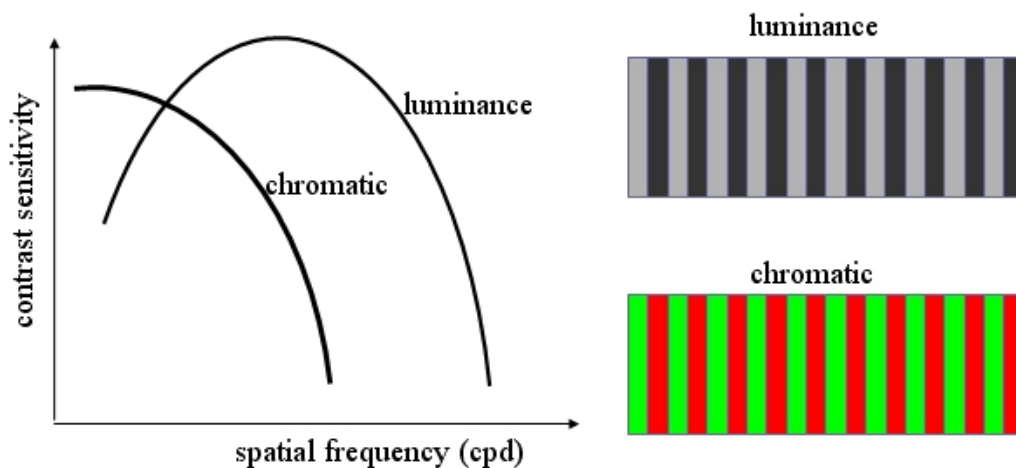
Figure 2-7: Contrast sensitivity function of luminance and chromatic stimuli.

2.4 Feature Integration Theory

In Section 1.2, we mentioned that if textures can be discriminated immediately is a rough but important judgment on whether the proposed approach functions correctly or not. The preattentive visual task is completed at very early stages without attention involved; that is, there will be no top-down process thoroughly analyzing visual primitives. The definition is clear except for the word "preattentive." Purpose of this thesis would not be definite until a clear description about preattentive processing could be given.

Feature integration theory (FIT) by Treisman and Gelade [60] gave an intuitive and critical definition to demarcation between preattentive stage and attentive stage. Fig. 2-8 is a schematic diagram of FIT composed of two stages of visual processing: At first, visual primitives of objects are analyzed in parallel and coded in feature maps. At the second stage, focal attention serially deploys to particular positions and serves to "glue" visual primitives into object representation. Some features glued from basic primitives by attention will cost more time to be perceived since the gluing procedure is not parallel but serial. Thus, to judge what stage a visual feature is processed at, the reaction time is an indicative clue. Treisman indicated that at preattentive stage, the

reaction time is fast (pop-out) no matter how many distractors are present on display. At attentive stage, however, increasing the number of distractors will increase the reaction time as shown in Fig. 2-9.
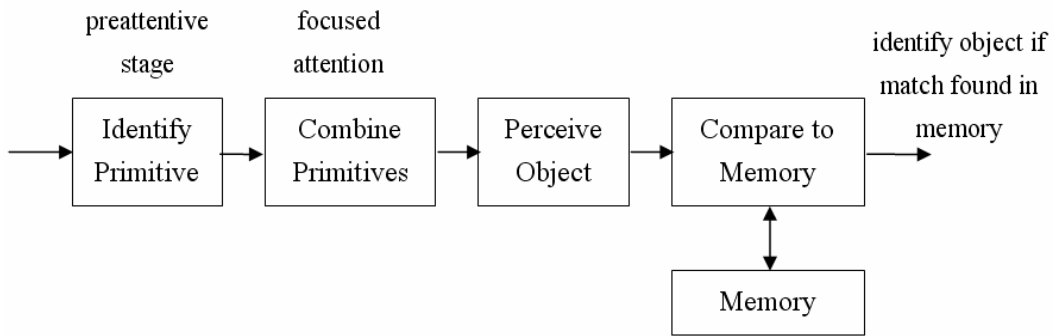
Figure 2-8: Schematic diagram of feature integration theory.
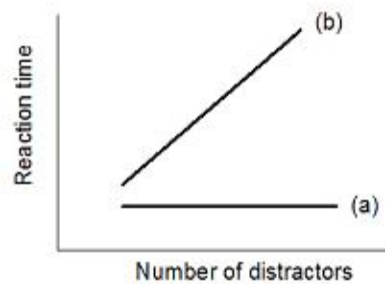
Figure 2-9: Typical results of a visual search experiment: (a) the result when pop-out occurs; (b) the result without occurrence of pop-out.

So far, we have introduced some relevant biological backgrounds about this research. In Chapter 3, a computational model will be developed based on these backgrounds.

Chapter 3

Modeling Strategy

The physiological and psychophysical evidences introduced in the preceding sections did not lead to a convenient computational model representing visual primitives. In this chapter, a novel boundary detection algorithm will be proposed. This algorithm combines the 1st- and 2nd- order features to model the texture segregation task at preattentive stage of human visual system.

Fig. 3-1 shows the overall framework. In the beginning, a color image is decomposed into one luminance and two chromatic channels in CIELAB color space. We apply Gaussian function to extract the 1st- order features, and Gabor filters to extract the 2nd- order features, respectively. In the two chromatic channels, only the lowest vertical and horizontal Gabor filters are applied due to coarser resolution and lower sensitivity frequency in chromatic channels. The 2nd- order features still need some operations like rectification and Gaussian smoothing after Gabor filtering, and the issues from hybrid-order scheme should also be considered. A typical issue is false responses to non-texture region (e.g. sharp edge) in the 2nd- order features which can be detected and removed by the proposed criterion. Another critical issue is the computational loads from the Gabor filter-bank approach. To relieve the problem, only significant features determined by variance will be reserved. After feature extraction, we then apply a local variance calculation to get the 1st- and 2nd- order boundaries respectively. Finally, with an adaptive weights selection, the merged boundary can be obtained. We may go a step further to thin the boundary by local peak detection, and get boundary similar to human visual system.
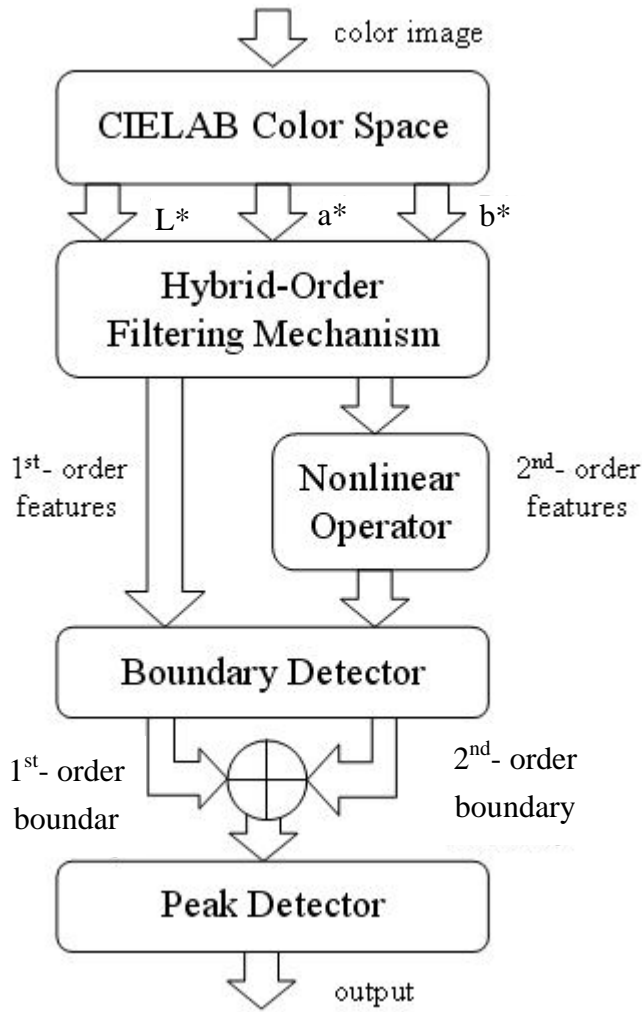
Figure 3-1: Flow chart of the proposed framework.

The proposed hybrid-order boundary detection algorithm will be presented in detail. In Section 3.1, the way to extract two important features in gray-scale images, luminance and texture, will be reviewed and discussed. Section 3.2 will introduce the strategy to extract hybrid-order features and some issues in luminance and chromatic channels. The nonlinear operations for the 2nd- order features will be described in Section 3.3, and in Section 3.4, the way to find the boundary will be illustrated.

3.1 Luminance and Texture Features Extraction

As mentioned, there are similar feature extracting mechanisms for boundary detection in three color channels. Here we will review and discuss two important

features: luminance and texture in gray-scale images first. In Section 3.2, it will be extended to color images.

3.1.1 Luminance

Among most image segmentation methods, the main criterion is to consider how and/or how much the luminance changes. There are already numerous literatures concerning the problem and a review can be found in the book "*Digital Image Processing*" by Gonzalez and Woods [61]. Some well-known methods such as Roberts, Prewitt, and Sobel operators were also reviewed in the book. The most two typical methods are Laplacian of Gaussian (LoG) operator by Marr and Hildreth [15] and Canny operator by Canny [62]. Laplacian of Gaussian operator was inspired from the anatomical structure transferring visual information in pre-cortical visual pathway. The receptive fields of bipolar cells and ganglion cells in retina and LGN cells can be fitted well by the LoG function. Marr and Hildreth used the LoG operator to find the zero-crossing of second derivative of an image smoothed by Gaussian function as the position of edge. By another approach, Canny defined three performance criteria for edges as (a) good detection; (b) good localization; and (c) only one response to a single edge, and then derived that the first derivative of Gaussian could be the optimal edge detector.

Though the analyzing notions of both operators are a little bit different, they employed the same concept to detect luminance discontinuities in two stages: Gaussian convolution and gradient. Gaussian convolution is somehow like extracting the local mean called the 1st- order feature here, and in the second stage, gradient is a measurement for variation of the 1st- order feature. In this thesis, to build up a hybrid-order scheme, only Gaussian convolution will be applied in feature extracting stage and the gradient process will be accomplished later in boundary detecting stage.

3.1.2 Texture

The orientation and spatial frequency selective property of receptive field in V1 is considered a main source to texture perception and the linear filtering approach has been widely suggested and verified for spatial vision tasks. However, as described in Section 2.2, receptive field of each neuron cell extends within a limited region. Also, most images in nature are neither pure periodic nor comprised of a small set of singular functions (e.g. Dirac function). For these reasons, spatial-frequency filtering is a proper approach for texture analysis. In images, two-dimensional spatial filters are constrained by general uncertainty relations that limit some attainable resolution for spatial position, spatial frequency, and orientation. For luminance feature extraction, it is commonly suggested that Gaussian functions are proper to extract local averaging information. Besides close match between two-dimensional Gaussian functions and pre-cortical receptive field profiles, another important reason is that Gaussian functions can achieve the lower bound of joint uncertainty in two conjoint domains; that is, the trade-off between resolutions of position and spatial frequency can be mediated optimally.

For texture feature extraction, there are similar considerations. The two-dimensional Gabor functions, as defined: complex sinusoidal gratings modulated by two-dimensional Gaussian functions, are commonly suggested for texture feature extraction [18], [19], [22], [23], [63]. The neural model by Gabor functions was originally proposed by Daugman [16] in two-dimensional form and Marcelja [64] in one-dimensional form in 1980. Subsequent physiological findings also indicated the validity of the Gabor receptive field model [17], [65]-[68]. Though the receptive fields of the simple cells in V1 all differ with each other, they have some common characteristics. The receptive field profiles consist of spatial frequency, orientation selective characteristics. Also, the investigation by Pollen and Ronner [65] showed

that adjacent pair of simple cells with matched preferred spatial frequency and orientation has a quadrature phase relation. From neurophysiological measurements by Jones and Palmer [17], [57], the two-dimensional Gabor function could fit well the anisotropic receptive profiles describing neurons in mammalian visual cortex, in the sense of satisfying chi-squared tests. From another point of view, the two-dimensional Gabor functions uniquely minimize the two-dimensional space-frequency uncertainty principle for complex valued functions on $\boldsymbol{R}^2$ [17]. Furthermore, the quadrature relationship in complex components of the Gabor function allows a useful and unique approach to texture analysis.

Two-Dimensional Gabor function

Now, to formulate texture feature extraction more precisely, we would like to introduce the two-dimensional Gabor function thoroughly. Corresponding to our definition to texture, the two-dimensional Gabor filters are appropriate for texture segregation/segmentation tasks in the sense: they have tunable spatial frequency, orientation, and bandwidth. Thus, the analysis of texture can be reduced to the analysis of outputs of selected filters which carry specific contents in images. The Gabor function was originally suggested for processing and communication of speech signals [14]. It is well-known that the one-dimensional Gaussian function $g(x)$ with its Fourier transform $G(u)$ is the only $\boldsymbol{R} \rightarrow \boldsymbol{R}$ function pair achieving the lower bound of uncertainty relationship $\Delta x \cdot \Delta u \geq 1/4\pi$. The one-dimensional Gabor function $h(x) = g(x) \cdot \exp(j2\pi Ux)$ with its Fourier transform $H(u)$ is a more general function pair ($\boldsymbol{R} \rightarrow \boldsymbol{C}$) that also achieves the lower bound [14]. Daugman [17] further extended the uncertainty relationship to two-dimensional form and formulated that the two-dimensional Gabor function $h(x, y)$ is the only function mapping $\boldsymbol{R}^2 \rightarrow \boldsymbol{C}$ while achieving the both lower bounds of uncertainty relationship $\Delta x \cdot \Delta u \geq 1/4\pi$ and

31

$\Delta y \cdot \Delta v \geq 1/4\pi$ simultaneously. Similar to one-dimensional Gabor functions, the two-dimensional Gabor functions have the general form ($j = \sqrt{-1}$):

$$h(x, y) = g(x', y') \cdot \exp[j2\pi(Ux + Vy)], \qquad (3\text{-}1)$$

where $(x', y') = (x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$ are rotated spatial domain coordinates, and the two-dimensional Gaussian functions $g(x, y)$ have the form:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]. \qquad (3\text{-}2)$$

The frequency response of Gaussian function (3-2) is

$$G(u, v) = \exp\left\{-2\pi^2\left[(\sigma_x u')^2 + (\sigma_y v')^2\right]\right\}. \qquad (3\text{-}3)$$

where $(u', v') = (u\cos\theta + v\sin\theta, -u\sin\theta + v\cos\theta)$. The frequency response of Gaussian function is also a Gaussian form. The shape of two-dimensional Gaussian function is determined by aspect ratio (σ_x/σ_y) where σ_x and σ_y are standard deviation of Gaussian function in x axis and y axis. In some cases without knowledge of the context to extract texture feature, it is reasonable to select an isotropic Gaussian modulation, namely $\sigma_x = \sigma_y = \sigma$. The impulse response $h(x, y)$ is a complex sinusoid with center frequency (U, V) that is modulated by a Gaussian envelope. A complex sinusoid in spatial domain corresponds to a position shift in frequency domain. In other words, the complex exponential components determine the place where the main frequency response components of Gabor functions lie. The complex components of two-dimensional Gabor functions determine the central spatial frequency (F) and orientation (ϕ). The frequency response of the Gabor function (3-1) is given by

$$H(u, v) = \exp\left\{-2\pi^2\left[\sigma_x^2(u'-U')^2 + \sigma_y^2(v'-V')^2\right]\right\}, \qquad (3\text{-}4)$$

where $(u', v') = (u\cos\theta + v\sin\theta, -u\sin\theta + v\cos\theta)$ and (U', V') is a similar rotation

of the center frequency (U,V). Equation (3-4) shows that the frequency response of Gabor function $H(u,v)$ is a Gaussian function with representation of center position where frequency $F = \sqrt{U^2 + V^2}$ and orientation $\phi = \tan^{-1}(V/U)$ in frequency domain. An example of Gabor function with orientation $30°$ in spatial domain is presented in Fig. 3-2, where Fig. 3-2 (a) is the real part of a standard Gabor function in the spatial domain, and Fig. 3-2 (b) is the spectral response.

(a)   (b)
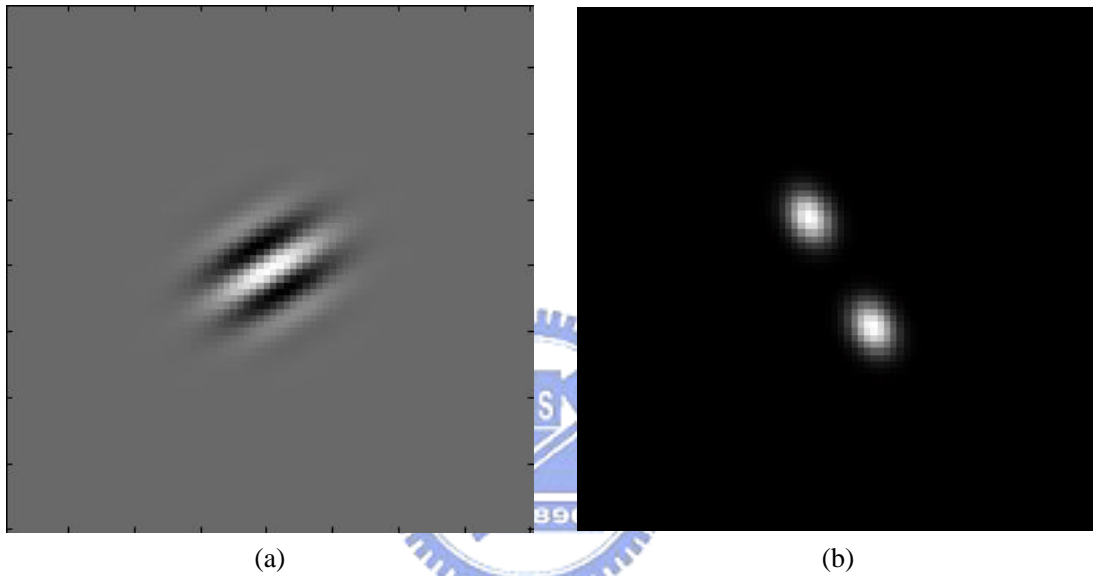
Figure 3-2: An example of 2D Gabor function in (a) spatial domain; (b) frequency domain.

The two-dimensional Gabor uncertainty principle is illustrated in Fig. 3-3. It shows the schematic representation of the real part of a two-dimensional Gabor function on the left side and the schematic representation of its Fourier transform on the right side. As the spatial resolution (Δx) gets coarser by a factor α, it allows a more accurate determination of frequency by a factor $1/\alpha$. Another illustration for trade-off between spatial resolution and orientation resolution is shown in Fig. 3-4. As the resolution in y axis (Δy) decreases, it will provide greater sensitivity to orientation, which can be seen in the frequency domain, where $\Delta\theta \propto 1/(F\Delta y)$ gets smaller.
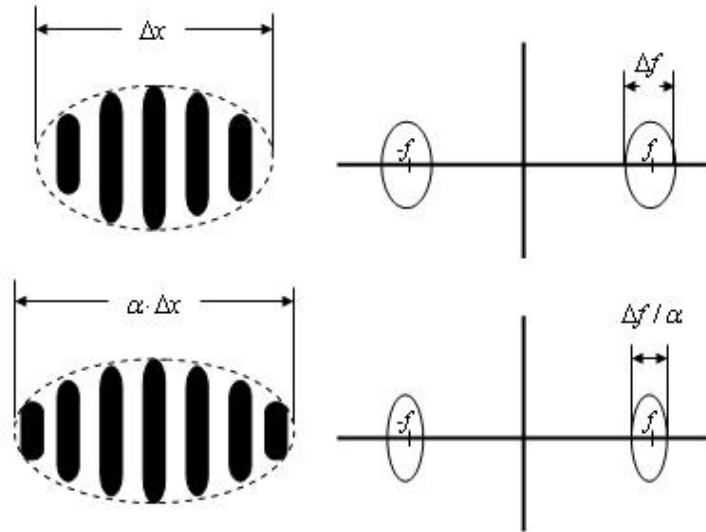
Figure 3-3: Trade-off between spatial resolution and frequency resolution (adapted from [17]).
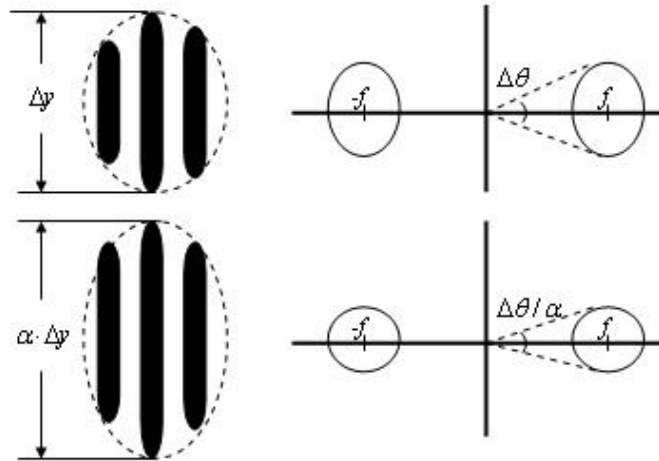
Figure 3-4: Trade-off between spatial resolution and orientation resolution (adapted from [17]).

In practice the Gabor function can be divided into real (even) part $h_c(x, y)$ and imaginary (odd) part $h_s(x, y)$ as:

$$h_c(x, y) = g(x, y) \cdot \cos[2\pi(Ux + Vy)] \tag{3-5}$$

and

$$h_s(x, y) = g(x, y) \cdot \sin[2\pi(Ux + Vy)]. \tag{3-6}$$

The real part and imaginary part are requisite for complete Gabor scheme of image representation [17], [35], [36]. For texture feature extraction, however, what principally characterizes the form of texture is not phase spectral component defining

spatial relationship between textural elements but magnitude spectral component describing textural elements (e.g. orientation, repetitiveness, *etc.*). Therefore, for texture feature extraction it is sufficient to consider magnitude information only. Some justifications by Malik and Perona [23] also showed that it is sufficient to extract significant textural information by even-symmetric or odd-symmetric filters only. From consideration of the overall framework, still, to avoid the "curse of dimensionality" at discrimination stage (boundary detection), it is more appropriate to select a compact and critical set of Gabor functions than a complete but redundant set. The scheme to be introduced, called Gabor wavelets is employed in our procedure due to its compact representation and convenient parameter selection step.

Gabor Wavelets

Wavelets, considering the joint uncertainty issue, have been presented as an alternative to Gabor functions as a basis set for representation in the visual system [69]. Multiresolution processing is a main concept of wavelets transform that function can be represented as linear superposition of strictly local elementary functions. A family of wavelets is a complete set of functions, all generated from a mother wavelet by the operations of dilation and translation. The efficient procedure by Mallat [27] provided a fast and economical way for signal decomposition. Besides pyramid schemes of Gaussian [70] and Gabor functions [36] for a compact image code, Daugman [35] unified Gabor functions and wavelets and then defined Gabor wavelets which fit well the neurophysiological and psychophysical findings indicating a log-polar distribution of response selectivity in V1 cells. These anisotropic wavelets are generated from a Gabor elementary function called mother Gabor wavelet by dilation, translation, and rotation. The mother Gabor wavelet $h'(x, y)$ with preferred orientation 0° can be presented as:

$$h'(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + j2\pi Ux \right]. \tag{3-7}$$

Rest of the filters set can be obtained by the generating function

$$h'_{mn}(x, y) = a^{-m} g(x', y'), \quad m, n \in \boldsymbol{I}. \tag{3-8}$$

$$(x', y') = a^{-m}(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$$

where $a = (U_h / U_l)^{\frac{1}{S-1}}$ and S is the number of selected spatial frequencies for decomposition; $\theta = \left(\frac{n\pi}{K} \right)$ and K is the number of orientations. The number $m = 0, 1, \cdots, S-1$ and $n = 0, 1, \cdots, K-1$ present the index of central frequencies and orientations. By the design strategy [26], to ensure that the half peak magnitude of the filter response in the frequency domain touches with each other for a compact representation. The formulas for computing the filter parameters σ_u and σ_v are given as [26]:

$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}, \sigma_v = \tan(\frac{\pi}{2K})\left[U_h - 2\ln(\frac{2\sigma_u^2}{U_h}) \right]\left[2\ln 2 - \frac{(2\ln 2)^2 \sigma_u^2}{U_h^2} \right]^{-\frac{1}{2}}. \tag{3-9}$$

The scheme has only four parameters to be selected: lower central frequency U_l; upper central frequency U_h; total number of frequencies S; and total number of orientations K. In Section 3.2, the scheme will be employed for the 2nd- order feature extraction. As mentioned, some issues in three opponent color channels have to be taken into consideration, we will also discuss those in Section 3.2.
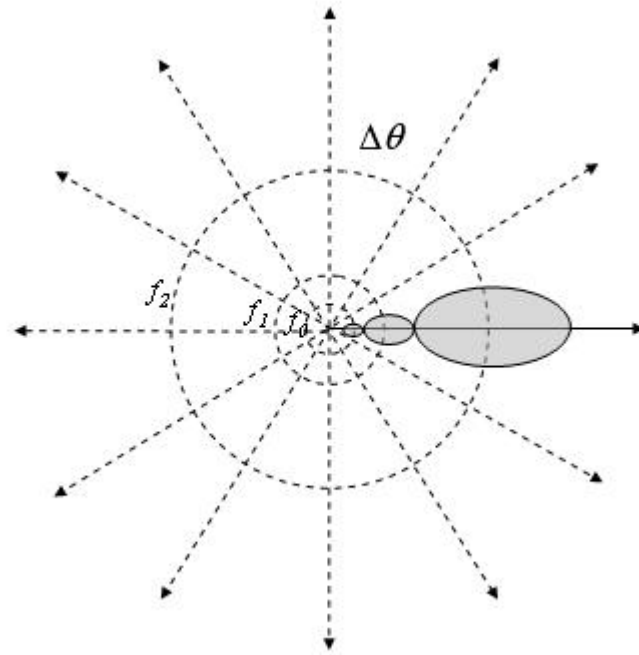
Figure 3-5: Representation of Gabor wavelets, where shaded ellipses represent envelopes of 2D Gabor filters.

3.2 Hybrid-Order Feature Extraction

3.2.1 Color Decomposition

As mentioned in Section 2.3, color information is delivered in a three-opponent-channel form and representations for spatial vision in the three channels are very similar. In 1976, Commission Internationale de l'Eclairage (CIE) defined two approximately uniform color spaces, CIELUV and CIELAB [71]. They are nearly linear to visual perception, and among both, CIELAB was developed based on the opponent-process theory employed for many psychophysical experiments. Also, the defined color difference, distance in the color space is very similar to human perception. From the reasons above, we choose CIELAB color space for color decomposition that we can be assured that biological evidences recorded in many literatures will not be misapplied. In fact, many previous approaches employed information of color contrast but applied improper color representation, that

37

reasonable results would be inaccessible.

The CIELAB color space has a luminance component $L*$ and two chromatic components $a*$ and $b*$. $L*$ is derived as:

$$L* = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & Y/Y_n > 0.008856 \\ 903.3(Y/Y_n) & Y/Y_n \leq 0.008856 \end{cases},$$ (3-10)

where Y_n is the CIEXYZ Y value for the reference white. $L*$ ranges from 0 to 100 where 0 is perfect black, 50 is average gray, and 100 is the reference white. $a*$ and $b*$ are defined as:

$$a* = 500[f(X/X_n) - f(Y/Y_n)],$$ (3-11)

$$b* = 200[f(Y/Y_n) - f(Z/Z_n)],$$ (3-12)

where X_n, Y_n, and Z_n are the tristimulus values for the reference white, and where

$$f(t) = \begin{cases} t^{1/3} & t > 0.008856 \\ 7.787 \cdot t + 16/116 & t \leq 0.008856 \end{cases}.$$ (3-13)

After a color image is decomposed, we then apply corresponding operations to the three components.

3.2.2 Feature Extraction

In Section 3.1, the schemes for extracting luminance and texture features were introduced. For color images, we represent color information by three values and employ similar schemes as those for gray-scale images. Thus we can represent the 1st- and 2nd- order features in one luminance and two chromatic channels. The mentioned luminance and texture in gray-scale images correspond to the 1st- and 2nd- order features in luminance channel. Before we adopt the schemes introduced in Section 3.1 for color images, we should discuss some issues about luminance and chromatic pathways in advance to prevent a crude representation.

For spatial vision, color information usually plays an important role for object

localization. Regions with the same color are more likely belong to the same object while regions with different colors are more likely belong to different objects. In our daily experiences, color usually presents local information as luminance does. Some textured patterns composed of color information are less mentioned or observed, and texture perception is usually ascribed to gray-scale information (in luminance pathway). In Section 2.3, some researches have been reviewed that there are some similar properties in luminance channel and chromatic channels. Manipulating isoluminant stimuli, it was shown that visual system is capable of discriminating spatial frequency and orientation [42], [44] in chromatic grating. Also, some psychophysical experimental procedures applied for validating multiple spatial frequencies filtering in luminance channel were also employed for investigating properties in chromatic channels [43], [72]. Besides previous psychophysical experiments, in recent years there were more physiological evidences supporting these properties of mechanism for processing color information [73]- [76].

Numerous researches revealed that there are similar mechanisms in luminance and chromatic channels. How visual system deals with chromatic texture was also discussed generally [37], [38], [77], [78]. Still, there seem some conflictions to our daily experiences and we seldom observe so called chromatic texture. The phenomenon is mainly due to: (a) less contrast sensitivity, lower sensitivity frequency, and coarser resolution in chromatic channel; (b) few natural cases exhibiting chromatic texture without luminance texture. Most researches investigated processes for pure chromatic stimuli in human visual system. However, such stimuli only appear under specific manipulations. Moreover, compared to mechanisms for luminance information, mechanisms for chromatic information carry less significant properties. Thus, it is a very uneconomical way to represent a visual context with identical scheme in luminance and chromatic channels. In our modeling strategy, only

information sensitive to lower frequency with vertical and horizontal orientations will be extracted as the 2nd- order features in chromatic channels. Another issue frequently discussed is about the interactions between luminance and chromatic pathways [38], [72], [77]-[80]. Based on individual experimental procedure and interpretation of results, there are still some debates about the issue. Besides, artifact stimuli in these experiments have to be manipulated precisely and many models were likely proposed to interpret some specific stimuli. Thus, to develop a general-purpose and economic approach under uncontrolled condition, the issue was not taken into consideration in this thesis.

The schematic diagram of hybrid-order feature extraction is shown in Fig. 3-6. After a color image is decomposed into one luminance component $L*$ and two chromatic components $a*$ and $b*$, the three components are convoluted with a two-dimensional Gaussian function to extract the 1st- order features. The 2nd- order features are extracted by Gabor wavelets scheme. A three-scales and six-orientations Gabor wavelets scheme is applied in the luminance channel, and only the lowest spatial frequency Gabor filters with vertical and horizontal orientations in Gabor wavelets scheme are applied in two chromatic channels. That is, we have totally three the 1st- order features and twenty-two the 2nd- order features ($6 \times 3 + 2 + 2 = 22$). The 2nd- order features still need some operation to exactly characterize texture properties. Also, some issues have to be discussed further and solved in hybrid-order scheme. In Section 3.3, there will be some discussions about further operations for the 2nd- order features.
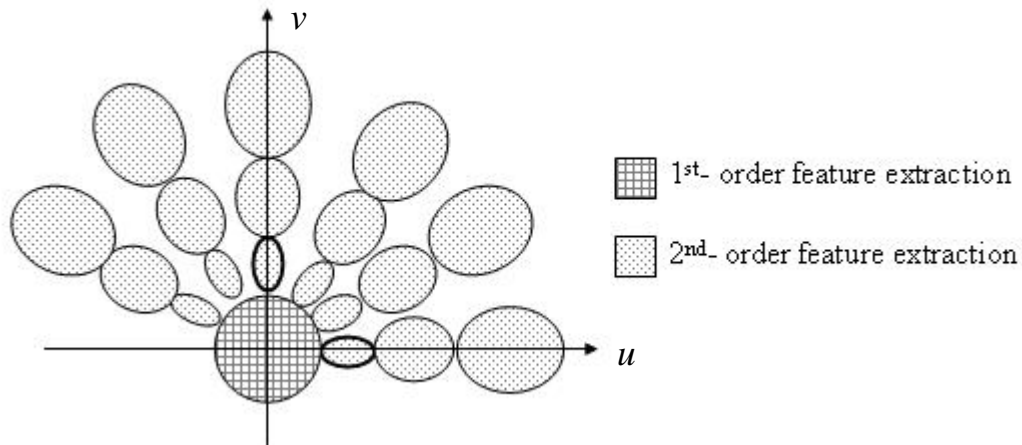
Figure 3-6: Schematic diagram of hybrid-order feature extraction presented in frequency domain, for 2nd- order feature extraction in chromatic channels, only the lowest frequency Gabor filters with vertical and horizontal orientations (bold contour ones) will be applied.

3.3 Operations for the Second-Order Features

The overall process of texture feature extraction can be divided into three parts: Gabor filtering, rectification, and Gaussian smoothing (G-R-G) [18]-[21]. In this section, to continue the discussion in Section 3.2 considering Gabor filter for the 2nd- order feature extraction, the other two stages: rectification and Gaussian smoothing will be introduced. Another important issue about the false 2nd- order features will also be discussed.

3.3.1 Full-Wave Rectification and Gaussian Smoothing

Full-Wave Rectification

Up to the present, what our procedure executes is like a set of parallel operators extracting parts of spectral contents. The review of texture theory in Chapter 1 has shown that the global power spectrum cannot provide sufficient texture properties. If the processing procedure only extracts spectral contents in another way, the main concept of the operation will not be different to global spectrum analysis. The principal difference between the linear filtering theory and Fourier transform to

41

texture analysis is due to the rectification operation. After Gabor filtering, the outputs of the filters are transformed with rectification to ensure that the fine-grain positive and negative portions of the carriers will not cancel one another when smoothed by a Gaussian post filter. Researches by Graham *et al.* [81] and Heeger [82] presented the "half-squaring," i.e., the outputs of the filters are first half-wave rectified and then squared. Simulation results with this kind of nonlinearity correlate well with neurophysiological data. In actual, it is well believed that V1 cells hold the half-wave rectifying operation [83]. Malik and Perona [23] also provided some justification for equivalent results by using even-symmetric and odd-symmetric Gabor filters.

In this thesis, only even-symmetric Gabor filters are employed with intervening full-wave rectification. Texture is a regional representation; that is, what a textural region appears is not fine-grain portions but local properties. For texture representation, the textural element is more significant than its relative location. In Fig. 3-7, for a reasonable representation of the grating texture, the responses in location (a) and (b) with 180° phase shift should be the same and thus a full-wave rectification should be applied after Gabor filtering. Moreover, the responses in (a), (b), and (c) should also be similar. The following stage, Gaussian smoothing will support the state.
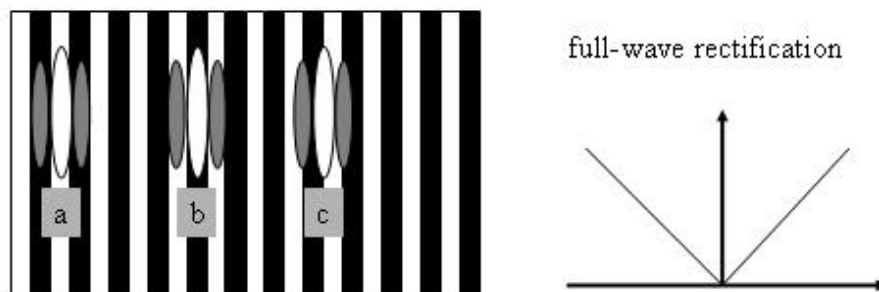
Figure 3-7: Applying full-wave rectification to make sure responses at (a) and (b) with 180° phase shift, would be the same.

Figure 3-8 (b) shows the output after Gabor filtering without rectification, and

Fig. 3-8 (c) shows the output after rectification. White pixels in the image reflect that Gabor filter has detected the matching features at the pixels, and some pixels with negative responses are not so visible in Fig. 3-8 (b). Textural contents of invisible pixels are very similar to those of white pixels. After rectification, in Fig. 3-8 (c), the two regions can be separated more apparently due to the rectification turning the negative responses to positive.
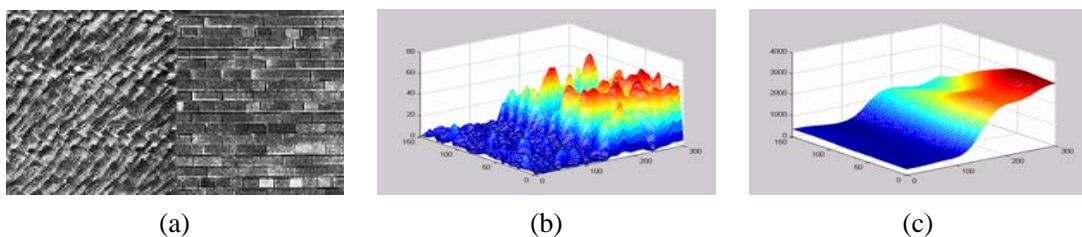

(a)

(b)

(c)
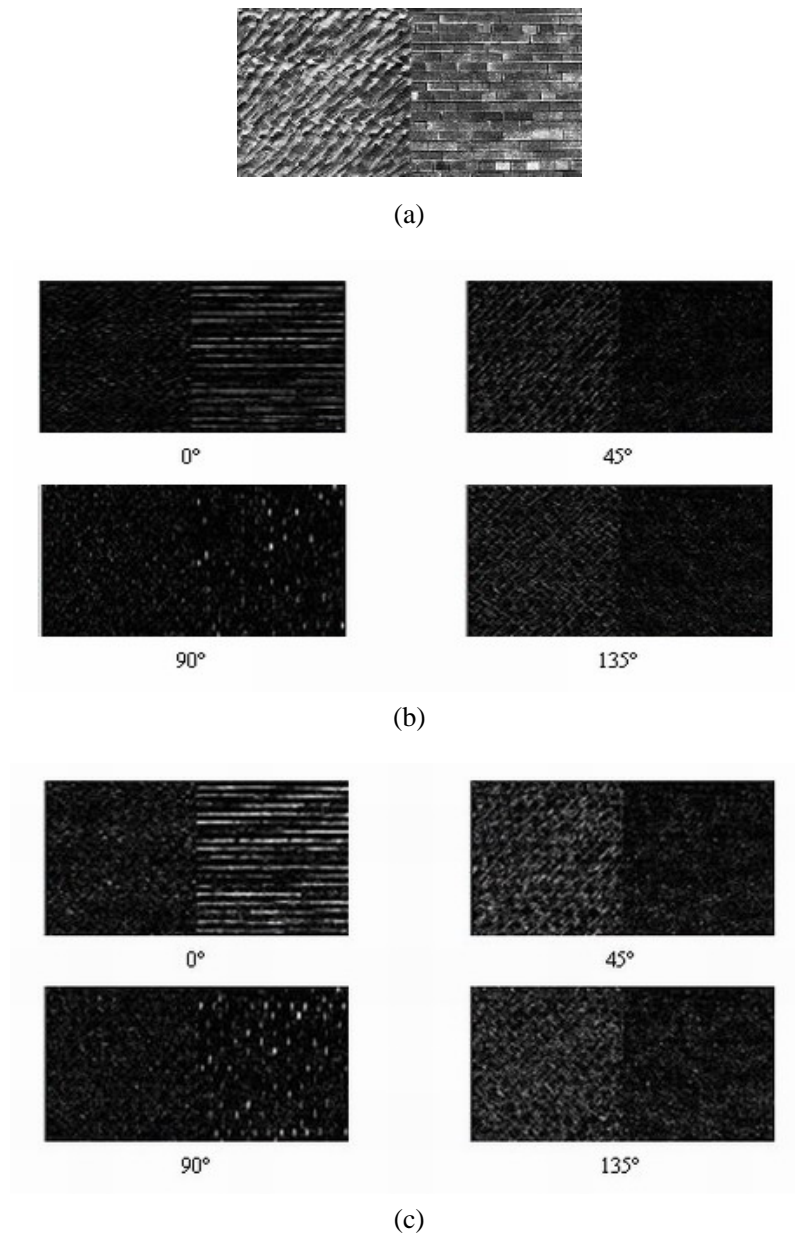
Figure 3-8: An example demonstrating the effect of rectification: (a) input; (b) responses without rectification; (c) responses with rectification.

Gaussian Smoothing

To manifest regional properties of textures, a further operation, Gaussian smoothing needs to apply after rectification. The outputs of areas with opposite phase show identical responses after full-wave rectification; however, regions with identical textural properties but different locations (phases) still exhibit some different responses. The features are inappropriate for boundary detection localizing positions where features change sharply as boundaries. For texture segregation/segmentation task, over-segmentation will occur frequently if there is no regional enhancement mechanism applying for textural features. Compared to human visual system, the outputs of V1 cells responding to similar orientation tend to aggregate together. The region with the same property will respond stronger than regions which consist of elements with different properties. The process is consistent with the localization property of textures. Such property can be simulated by a Gaussian post filter, and some spurious weak responses can also be eliminated. Fig. 3-9 (b) shows the result after rectification without Gaussian smoothing, and Fig. 3-9 (c) is the output of Gaussian post filter applying to Fig. 3-9 (b). In Fig. 3-9 (c), there is a ramp-like feature profile that provides more reliable distinction between different regions and resemblance within the same region.

(a)   (b)   (c)

Figure 3-9: An example demonstrating the effect of Gaussian smoothing: (a) input; (b) responses without Gaussian smoothing; (c) responses after Gaussian smoothing.

3.3.2 False Responses to Non-Texture Regions
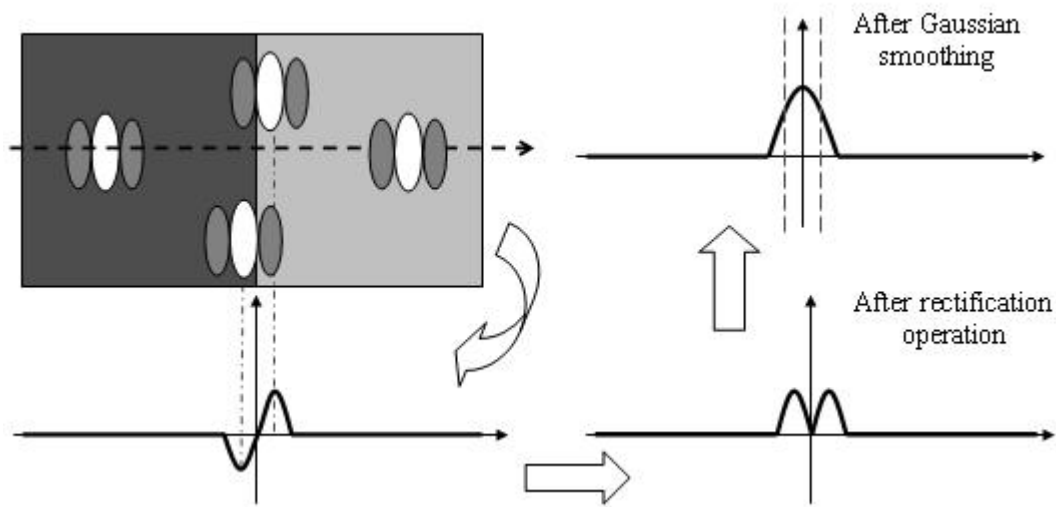
Figure 3-10: Schematic diagram demonstrating the false responses to non-texture regions.

For a pure texture segregation/segmentation task, there is no difference of average luminance between regions. The Gabor-Rectification-Gaussian process is a standard and acknowledged procedure [18]-[21]. However, if there are luminance differences between regions, the procedure will induce some false response. Usually, the false response occurs near the boundary. Fig. 3-10 is a schematic diagram demonstrating an extreme case. If only the 1st- order feature exists, after applying a Gabor filter with the same orientation to the boundary, there will be a peak-valley pair apart from the width of Gabor filter's mainlobe. Along the procedure of texture feature extraction: rectification and Gaussian smoothing, the 2nd- order feature will appear a significant peak on the boundary. Applying local variance calculation to find the boundary, we will detect the boundaries which are located at two sides of real boundary (as the dotted lines in Fig. 3-10).

Some researchers [84], [85] attempted to solve the problem by establishing a database, with training procedure they could avoid the false segregation. It is obvious that their procedure is not suitable for general-purpose tasks and we never know if the training procedure faithfully characterizes the relationship between the 1st- and 2nd-

order features. Kruizinga and Petkov [86], [87] and Grigorescu *et al.* [88] approached the issue by grating cell operator feature, which is inspired from the function of a visual neuron. Grating cell is a specific type of neuron found in areas V1 and V2 of the visual cortex and is also selective for orientation. Unlike the majority of orientation selective cells found in V1, grating cells do not react to single lines or edges but only respond when a set of bars with appropriate orientation and spacing in its receptive field. Based on the knowledge, Kruizinga and Petkov [86], [87] proposed a computational model for mimicking the response of grating cell. Their model consists of two stages. The first stage is constructed to respond at any position where a set of three parallel bars with specific orientation locates, and the second stage strengthens the output of the first stage if more than three parallel bars are present. Without considering the relevance of grating cells, the computational model indeed solved parts of the issues successfully and generalized the texture segmentation task. However, the types of textures are far more than the grating type texture, that the model is not a complete and economic approach for general cases.

Because there is still no acknowledged evidence considering about the issue from available biological models, in this thesis, we do not attempt to employ any biological process to deal with the problem under some limitations. Instead, we directly correct the conflicting phenomena from the source resulting in some false responses. The schematic diagram in Fig. 3-10 is an extreme case that only the 1st- order feature exists. In most cases, however, the 1st- and 2nd- order features exist simultaneously. Moreover, the ranges of feature values are different by cases. Take Fig. 3-11 and Fig. 3-12 for example. In Fig. 3-11, the 1st- order feature difference between regions is little. However, the 2nd- order feature is much weaker over the whole. Hence the false response will occur at parts of the boundary even thought the 1st- order feature changes slightly. Another case shown in Fig. 3-12 exhibiting obvious 1st- order feature

difference between regions while appearing a more intense 2nd- order feature does not lead to false response near the boundary. Considering the above cases, the absolute 1st- order feature difference between regions does not straightly provide a reliable judging criterion. A more appropriate criterion is the relative significance between the 1st- and 2nd- order features. If the 1st- order feature overwhelms the 2nd- order feature which is sensitive to the orientation of boundary, an obvious peak-valley pair will appear and the response will be much larger than the 2nd- order features at other locations.

| (a) | (b) | (c) |
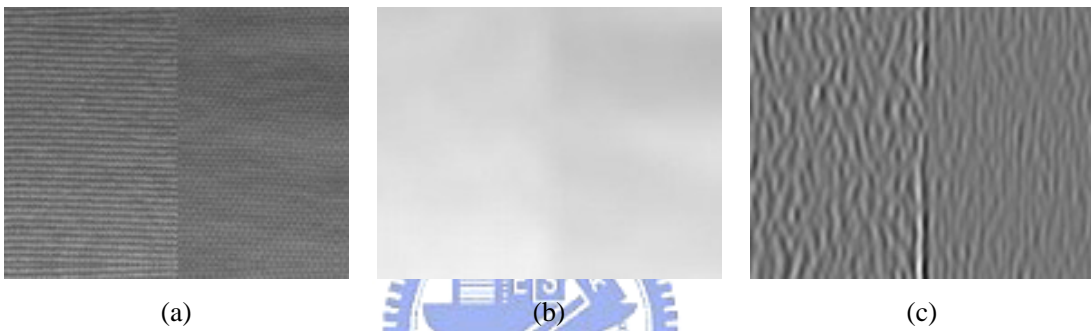
Figure 3-11: An example (Ⅰ) demonstrating the false response issue: (a) Input image; (b) output image after Gaussian filtering; (c) output image after Gabor filtering.

| (a) | (b) | (c) |
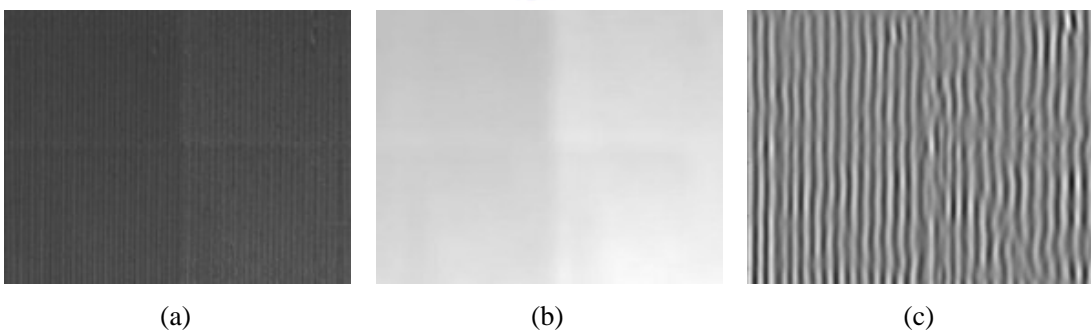
Figure 3-12: An example (Ⅱ) demonstrating the false response issue: (a) Input image; (b) output image after Gaussian filtering; (c) output image after Gabor filtering.

In Fig. 3-10 we showed that the peak and valley would be apart from the width of Gabor filter's mainlobe. We can go further to enhance the peak-valley pattern to make it more discriminable by applying a filter with positive-negative peaks pair apart from the same width of Gabor filter's mainlobe. After applying the filter, the output is

like Fig. 3-13. If positive-negative pair matches the peak-valley pair, a strong positive peak with two alike negative peaks will appear. Similarly, if the positive-negative pair matches the valley and peak, a strong negative peak with two alike positive peaks will appear. The critical feature was employed as a criterion for dealing with the issue of false response. Our procedure for solving the false response problem could be given as follows:

(i)   Apply the filter with positive-negative peaks pair apart from the width of Gabor filter's mainlobe to the direction orthogonal to the selective orientation of Gabor filter.

(ii)   Scan the peaks and valleys.

(iii) Compare absolute value of each peak (valley) to its previous one and next one. If the peak (valley) value is larger than 2 times to its previous one and next one, it means the false response appears at the location that the 2nd- order features near the region have to be replaced.

(iv)   The false 2nd- order responses are replaced by padding values linearly from both sides.

Fig. 3-14 shows the results by our method. The false responses can be replaced while preserving other correct texture information.

<center>(a)   (b)   (c)</center>
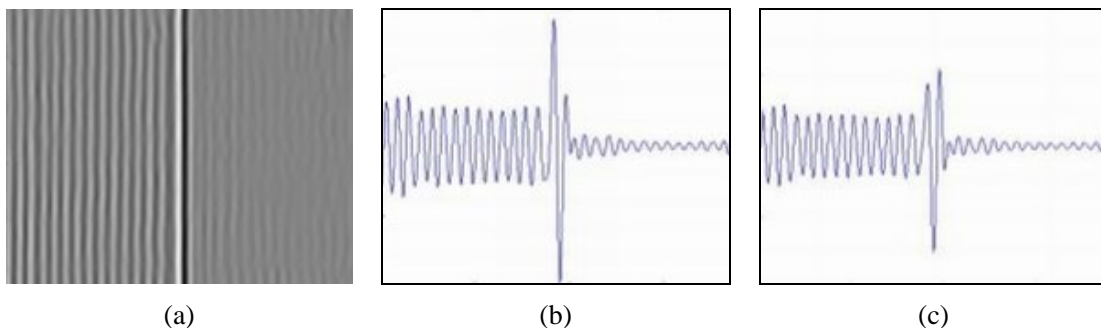
Figure 3-13: A demonstration of our procedure enhancing the features of false responses: (a) output after Gabor filtering; (b) a cross line of (a); (c) output after applying the filter with positive-negative peaks pair apart from the width of Gabor filter's mainlobe.

<center>48</center>
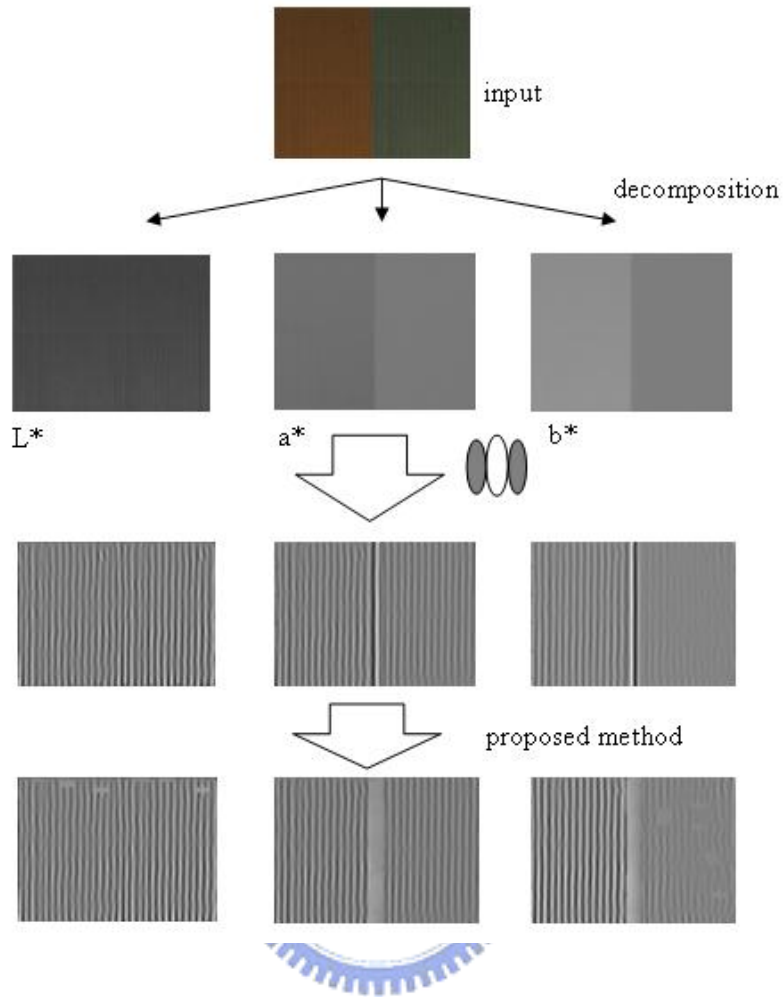
Figure 3-14: An example demonstrating the results of our procedure for the false responses.

3.3.3 Features Reduction

The main problem of filter-bank approach is "the curse of dimensionality," and the filter-design approach attempted to solve the problem with some criterion to select suitable filters fitting well with the image contents. As mentioned in Section 1.3, for a general-purpose procedure, the filter-bank approach is a more appropriate choice. However, the approach inevitably has to face the trade-off between feature completeness (more filters) and redundancy reduction (less filters). In our approach, to eliminate some insignificant features (non-responding filters), the variance of each feature will be calculated. The ideal features providing some discriminable

information must exhibit some value differences among the overall region. If a filter responds or not responds to the overall content, features extracted by the filter cannot provide sufficient information for discrimination. Employing the concept, the variance of each 2nd- order feature will be checked and the feature with variance less than 35% of the maximum will be discarded.

3.4 Hybrid-Order Boundary Detection

In this thesis, the features are extracted in parallel. At boundary detection stage, the proposed approach does not combine the two kinds of features into one feature space due to some issues: (a) the procedures extracting the 1st- and 2nd- order features are different such that the meanings of the two representations are also unequal. (b) Significance between the 1st- and 2nd- order boundary is not always the same, and thus how to select proper weighting would be a problem. Instead of roughly combining the features, we propose a procedure to find the 1st- and 2nd- order boundaries individually and then combine the two boundaries. The proposed method can combine the boundaries by determining the weights adaptively.

3.4.1 The First- and Second-Order Boundary Detection

Our procedure to find the boundary is based on the concept that the locations where features change obviously would be more likely the boundaries. After extracting features of all regions, the features can be described in a vector. The degree of how much the vector changes can be considered as an indicator of boundary. In our approach, all features from three channels are combined into two feature spaces as the 1st- order feature space and the 2nd- order feature space. By applying local variance criterion, we can localize two kinds of boundaries individually by both feature spaces. A schematic diagram is shown in Fig. 3-15, after all the features are arranged well,

local variances within 5×5 mask are calculated in both the 1st- and 2nd- order feature spaces. The two boundaries carry individual information of their own such that the combination can be implemented properly based on a criterion considering the reliability and compactness of boundary defined by each feature space.
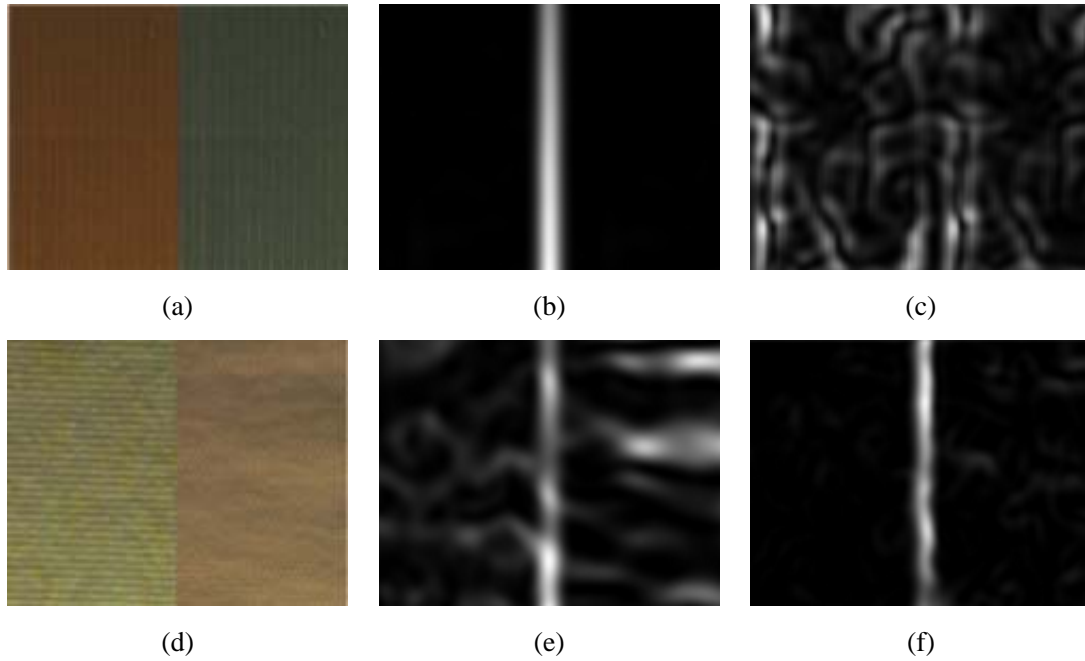

feature space

variance of 25 vectors
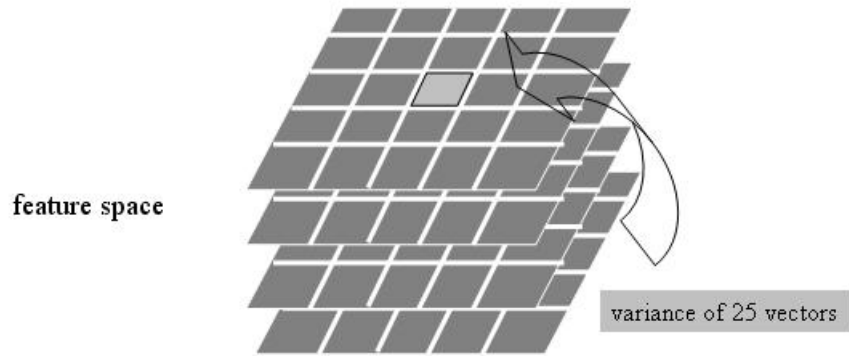
Figure 3-15: Schematic diagram of boundary detector.

3.4.2 Boundary Combination

For segregation task, previous researches revealed that there should be a common site integrating boundaries defined by different attributes [89]-[91]. Also, it was verified that combining multiple attributes posterior to a decision of localizations is advantageous to get more precise boundaries [89], [90]. If there are conflictions between boundaries defined by different features, the weights of all features are mediated by the measurements of reliabilities. A larger weight would be assigned to a boundary which is more reliable and linear summation for combining multiple features was suggested [90], [91]. Fig. 3-16 shows two opposing cases to demonstrate that the reliability of each feature is not always the same, as the case may be. Thus a criterion for determining the reliability is required.
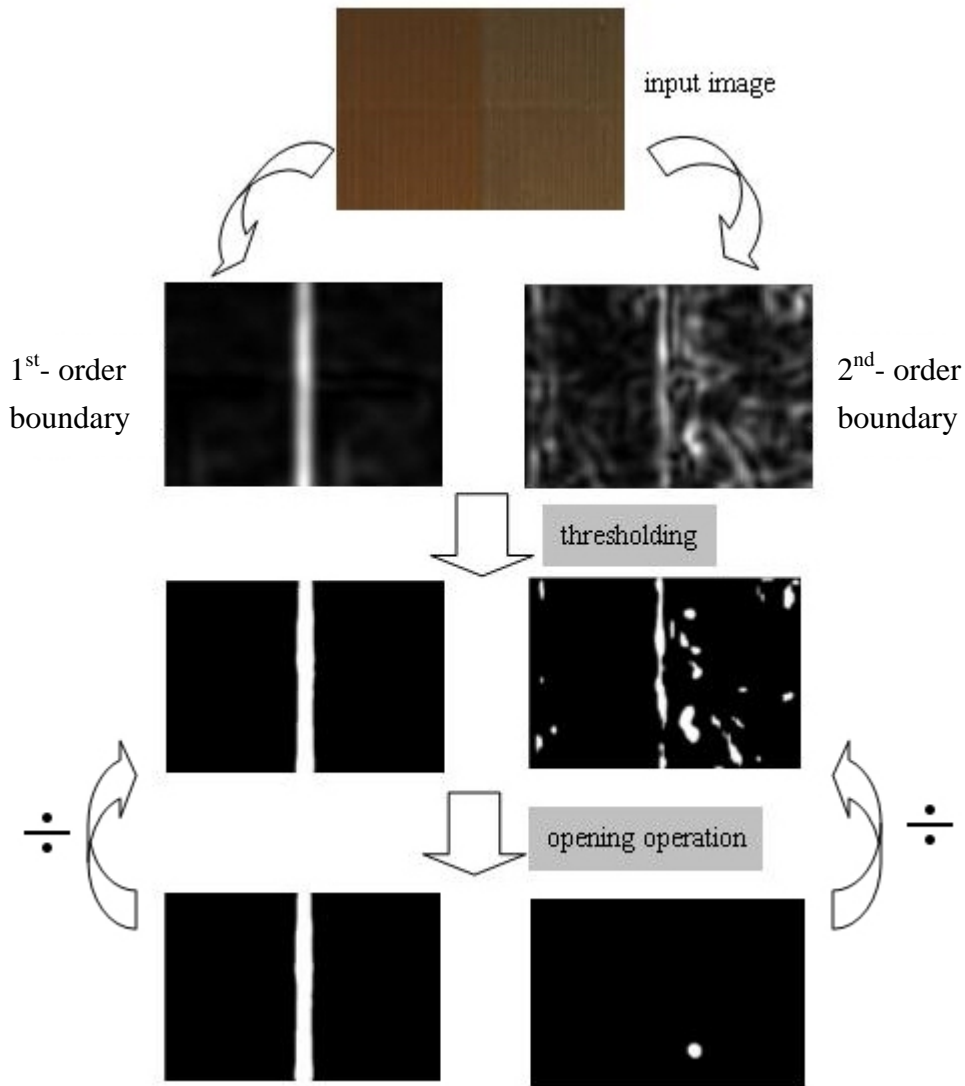
Figure 3-16: Two examples demonstrating the dominating role changes between the 1st- and 2nd- order features (a), (d) input images; (b), (e) 1st- order boundaries of (a), (d) respectively; (c), (f) 2nd- order boundaries of (a), (d) respectively.

In our approach, we utilize the opening operator to determine the weights adaptively. After applying a threshold to the raw data of boundary, we will get a compact and continuous black-white boundary if the feature space is reliable. Oppositely, for a feature space which does not provide sufficient discriminable information, we will get only some noise-like fractals. Opening operation is one kind of morphological process that it can be used as a geometrical selection. By applying opening operation to the black-white boundary image, for more compact and reliable boundary, more region areas will be reserved after opening operation. Thus the weights for combination can be selected from the region area ratio before and after opening operator. The whole procedure for boundary combination could be given as follows:

(i)   Normalize the 1st- and 2nd- order boundaries to values of 0-1.

(ii)   Apply thresholds to make the two boundary images binary.

(iii)   Apply opening operator to the two black-white boundary images.

52

(iv)   Calculate region area ratios before and after opening operator as weights.

(v)   Combine the normalized boundaries by the determined weights and then normalize the combined boundary to 0-1.

Figure 3-17: An example demonstrating the weights selection by opening operation.

In Fig. 3-17, it is demonstrated that the left side and right side of input image are segregated mainly by the 1st- order feature and the texture information is nearly uniform over the image. By applying local variance calculation, we can find the 1st- order boundary is much more compact and reliable than the 2nd- order boundary. After threshold and opening operation, the 1st- order boundary will get much larger weight

by calculating the region area ratios. The combined boundary can fairly preserve the information from the 1st- order boundary in this case. In the next chapter, more general cases will be shown that the significance between the 1st- and 2nd- order features is not always the same thus it is inappropriate to assign fixed weights for combination. The proposed method for feature combining issue can indeed assign weights adaptively and preserve the significant feature fairly.

3.4.3 Local Peak Detection

The coarse boundaries can be thinned by applying local peak detection. It is assumed that the difference between distinct patterns is maximal at their boundary. Algorithm of local peak detection is given below:

(i) *Scan row by row and column by column to find local maxima in x and y axes.*

(ii) *Sort the peaks in descending order and keep points with higher order in each row and column. These points are regarded as boundaries.*

The number of peak-points we keep in *(ii)* depends on the complexity of input image. In our testing cases to be shown in Section 4.3, we selected two peaks.

3.5 Summary

In this chapter, we built up the model from evidences about human visual system and relevant researches. Three visual primitives: luminance, texture, and color were properly combined into a functional system for the chromatic texture segregation task. Issues of false responses and weights selection were also discussed thoroughly and solved by proposed method. In the next chapter, the proposed model will be applied for extensive testing images. Besides, experimental comparisons about employed processes will also be demonstrated.

Chapter 4

Experimental Results & Discussions

In this chapter we will apply our algorithm to a number of testing patterns. All of them are synthesized by textures from Outex database [92], which contains a large collection of textures, in both form of surface textures and natural scenes. The collection of surface textures exhibits well defined variations to a given reference in terms of illumination, rotation, and spatial resolution. The synthesized images for experiments are 746×746 pixels in 24-bit RGB. When we compute the texture features for pixels near the image boundary, the regions which are not totally covered by filtering mask will be discarded. In Section 4.1, we will first introduce parameters selection. In Section 4.2, some important properties exhibiting on experiments will be discussed. A wide test on synthesized patterns by the proposed algorithm will be shown in Section 4.3. At last, the error estimation for the algorithm will be discussed in Section 4.4.

4.1 Parameters Selection

There are some parameters need to be selected: (a) the number of Gabor filters and the parameters $(U, V, \sigma_x, \sigma_y)$ determining the shape and orientation of Gabor filters in the frequency domain. Gabor filtering is computation intensive, and increasing the number of Gabor filters will increase computational loading dramatically. On the other hand, unnecessary and useless features extracted by ill-designed Gabor filters may result in incorrect boundaries. (b) The standard deviation of post Gaussian filter σ_g, which determines the smoothing level. Increasing

σ_g and the standard deviation of Gabor filter (σ_x, σ_y) can eliminate more noise, but the accuracy of boundary might decrease.

Choice of the above parameters is an important but sophisticated problem and choice of center frequencies of Gabor filters is the most frequently discussed issue in filter-design approaches (Section 1.3). Algorithm in this thesis employed the Gabor wavelets scheme to build an unsupervised procedure to deal with more general cases. As described in Section 3.1, the Gabor wavelets scheme can extract significant texture information with less filters, and all the parameters for the 2nd- order feature extraction reduce to only four selections: the number of frequencies S, the number of orientation K, the lower central frequency U_l, and the upper central frequency U_h. The values of parameters for all simulations are shown in Table 4-1.

Table 4-1: Parameters for Experiments

| Parameters	Values
Pattern Size	746 x 746 pixels
Number of Orientations	6 ($0°$, $30°$, $60°$, $90°$, $120°$, $150°$)
Number of Frequencies	3
Lowest Frequency U_l	0.06 cycles/pixel
Highest Frequency U_h	0.24 cycles/pixel
STD of Gaussian Filter	25 pixels

4.2 Experimental Comparisons

In this section, some experimental comparisons will be given to show the necessity of representation employed in this thesis. Foremost of them is the effect of multiple sensitivity frequencies, which underlines the Gabor wavelets scheme. In this case, to avoid misleading the discussion, a gray-scale image will be used for comparison. To emphasize the essentiality of color information, a comparison between results with and without color information will also be given. Finally, some

examples will be shown to demonstrate the effects of hybrid-order features, including incompleteness of each feature and collaboration of hybrid-order features.

4.2.1 Experiment Ⅰ: Effect of Multi-Band Gabor Filters

In this experiment we will demonstrate the reason why multi-band Gabor filters are employed. During this experiment we did not consider the 1st- order features but the 2nd- order features only while detecting boundaries. For clearly demonstrating the effects, the feature reduction process will be skipped. From Fig. 4-2 to Fig. 4-4, we show the raw data of the 2nd- order features with selective orientations of $0°$, $30°$, $60°$, $90°$, $120°$, $150°$ in each single band. The center frequencies are 0.06, 0.12, 0.24 cycles/pixel. Boundaries detected by each band are shown in Fig. 4-5 (a)-(c).

Fig. 4-5 (a)-(c) demonstrate that there are still some boundaries undetected; that boundaries can not be detected fully by a single band in this example. Fig. 4-5 (d) shows the result of boundary detected by three bands simultaneously; all the boundaries can be detected successfully.
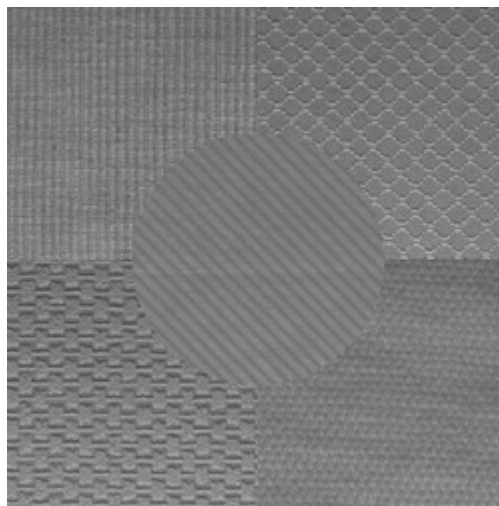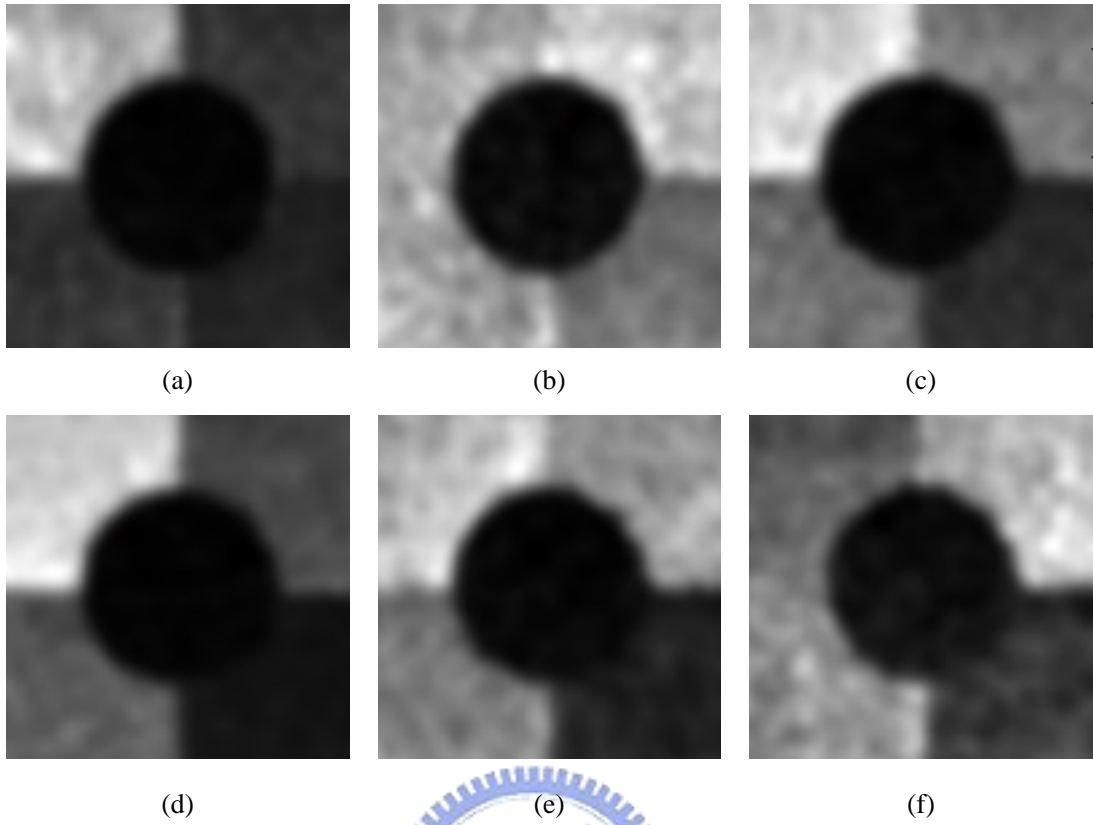

Figure 4-1: Input image for demonstrating the effect of multi-band Gabor filters.
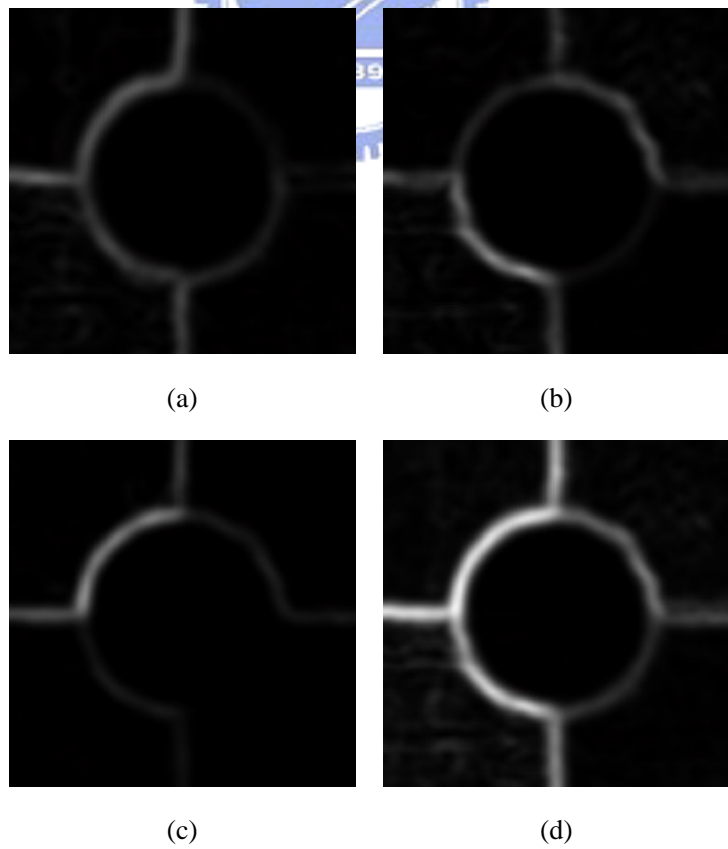
(a)

(b)

(c)

(d)

(e)

(f)
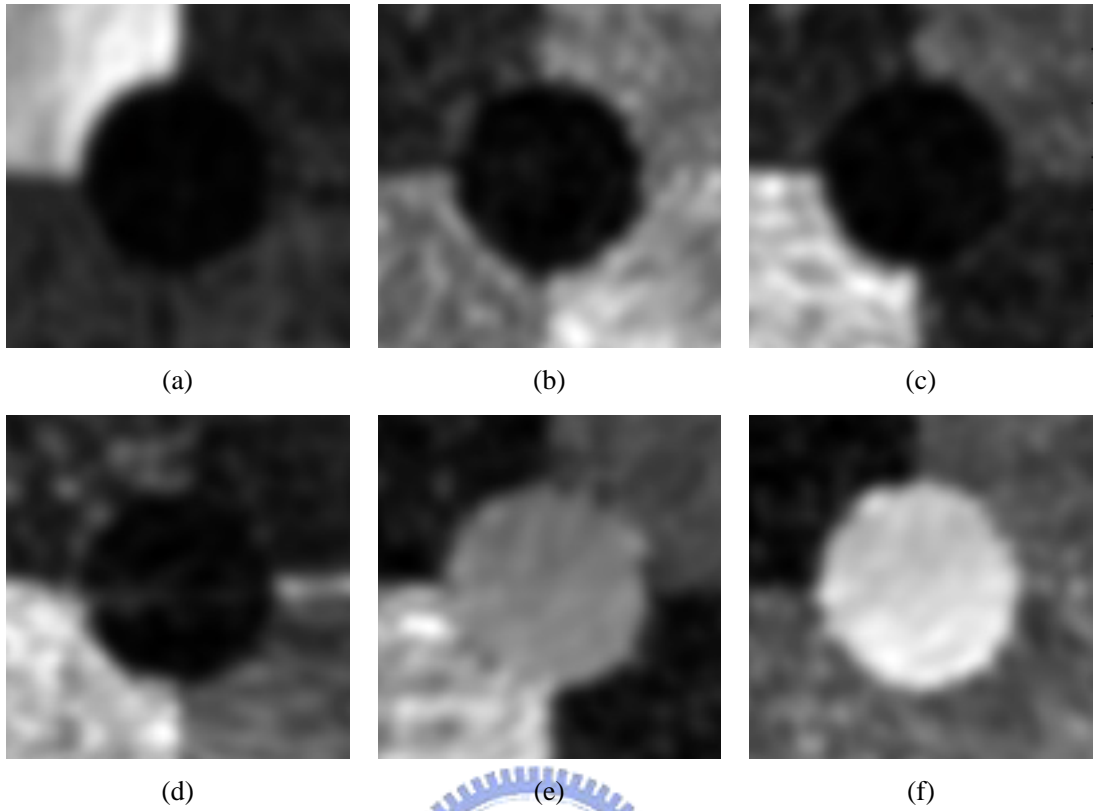
Figure 4-2: Raw data of features in band Ⅰ: (a) $0°$, (b) $30°$, (c) $60°$, (d) $90°$, (e) $120°$, (f) $150°$.

(a)

(b)

(c)

(d)

(e)

(f)
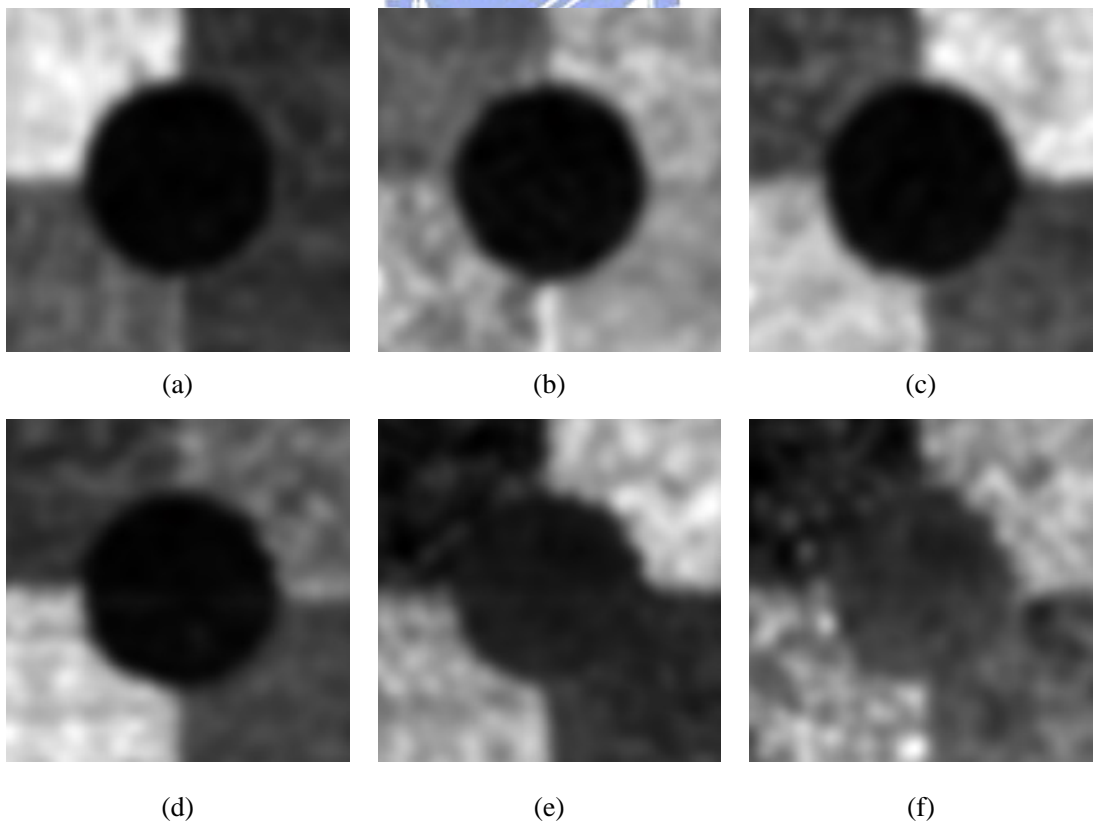
Figure 4-3: Raw data of features in band Ⅱ: (a) $0°$, (b) $30°$, (c) $60°$, (d) $90°$, (e) $120°$, (f) $150°$.

(a)   (b)   (c)

(d)   (e)   (f)

Figure 4-4: Raw data of features in band Ⅲ: (a) $0°$, (b) $30°$, (c) $60°$, (d) $90°$, (e) $120°$, (f) $150°$.

(a)   (b)

(c)   (d)

**Figure 4-5:** Boundaries detected by (a) band Ⅰ; (b) band Ⅱ; (c) band Ⅲ; (d) 3 bands simultaneously.

## 4.2.2 Experiment Ⅱ: Effect of Color Information

In this subsection, we shall first demonstrate an extreme case to emphasize the essentiality of color information. In the proposed model, a color image is decomposed into one luminance ($L*$) and two chromatic ($a*$ and $b*$) components. The demonstrated case consists of two regions where luminance and arrangements (texture) are nearly the same as shown in Fig. 4-7, and good segregation cannot be achieved without information from chromatic components (Fig. 4-8). In this case, the effects of our method for adaptive weights selection can also be demonstrated, *i.e.*, more reliable boundaries can be reserved. In Fig. 4-9, the combined boundary reserves more information from the 1st- order boundary which is more reliable and compact in this case. Two regions are segregated successfully with considering color information.
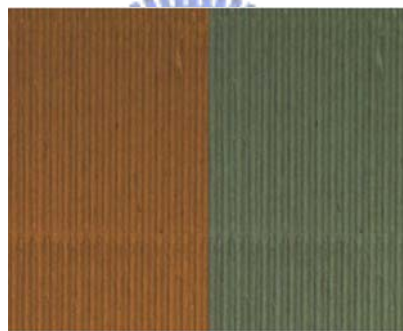
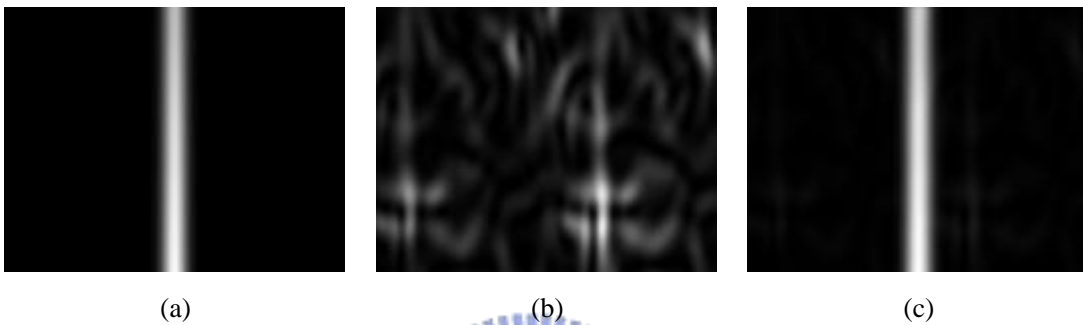**Figure 4-6:** Input image for demonstrating the effect of color information.

| (a) | (b) | (c) |
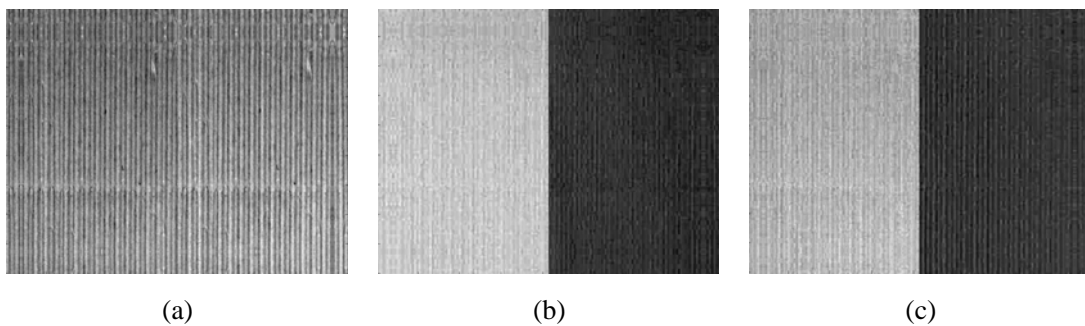
Figure 4-7: Raw data of (a) luminance $L*$ component; (b) chromatic $a*$ component; (c) chromatic $b*$ component.

|		
(a)	(b)	(c)
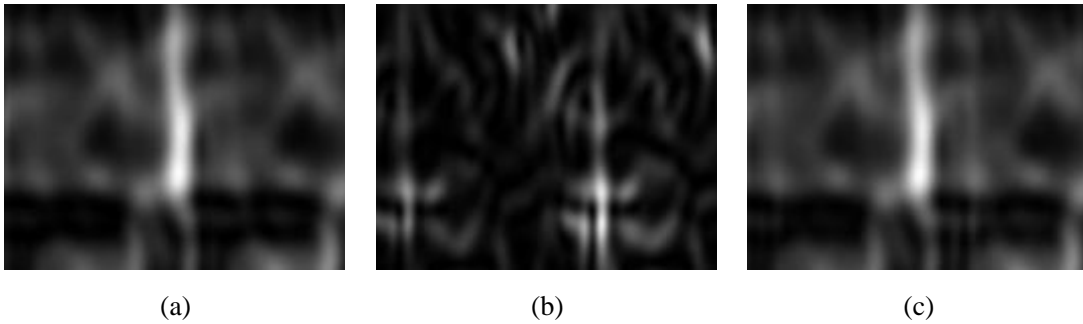
Figure 4-8: Boundaries detected by information from luminance channel only: (a) 1st- order boundary; (b) 2nd- order boundary; (c) combined boundary.

|		
(a)	(b)	(c)

Figure 4-9: Boundaries detected by information from luminance and chromatic channels: (a) 1st-order boundary; (b) 2nd- order boundary; (c) combined boundary.

In practice, we seldom find cases having specific visual primitive difference only. Another example to be demonstrated is more similar to general cases. Some illumination shading in the example would result in some ambiguities in luminance, and more reliable features could be provided by color information despite appearance of shading. In Fig. 4-12, the 1st- order features in two chromatic channels are more uniform within single texture, and a better segregation result can be obtained by considering color information (Fig. 4-13).
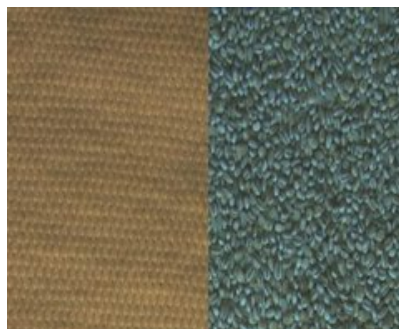
Figure 4-10: Input image for demonstrating the robustness to shading by color information.

(a)	(b)	(c)
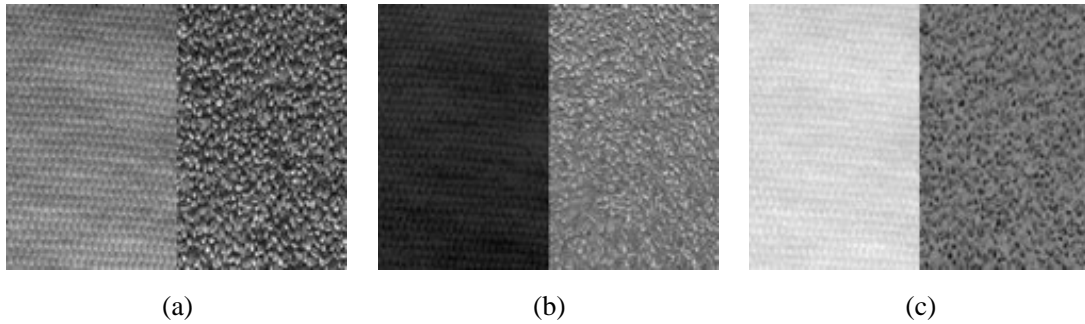
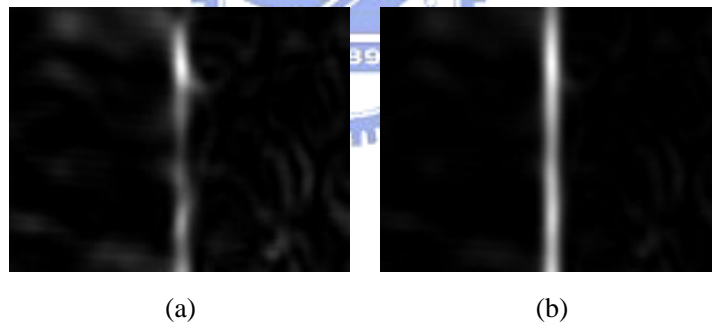**Figure 4-11:** Raw data of (a) luminance $L*$; (b) chromatic $a*$ component; (c) chromatic $b*$ component.

(a)	(b)	(c)
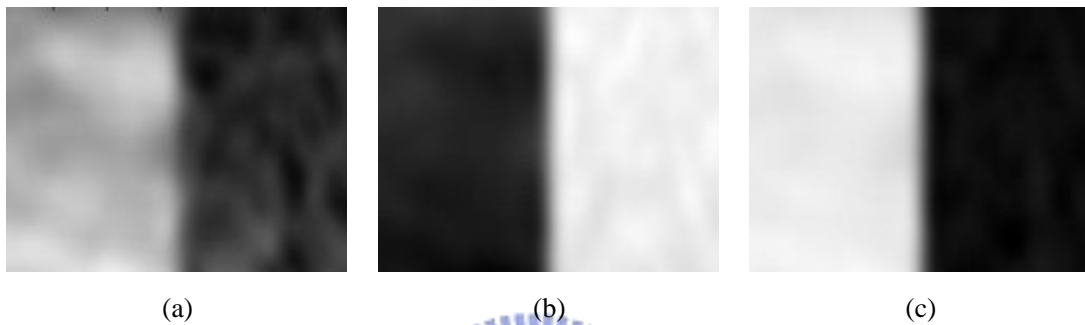
Figure 4-12: The 1st- order features in (a) luminance $L*$ component; (b) chromatic $a*$ component; (c) chromatic $b*$ component.

(a)	(b)

Figure 4-13: The 1st- order boundaries detected by (a) luminance information only; (b) luminance and chromatic information.

4.2.3 Experiment Ⅲ: Effect of Hybrid-Order Features

In Section 3.4.2 we have demonstrated that the dominating role changes between the 1st- and 2nd- order features frequently such that it is inappropriate to select fixed weights for boundary combination, and it is obvious that a model is certainly incomplete without considering hybrid-order features simultaneously. In this subsection, we shall straightly demonstrate the effect of hybrid-order features with an

example. Fig. 4-14 (b) and Fig. 4-14 (c) are the raw data of detected boundaries by the 1st- and 2nd- order features respectively. In Fig. 4-14 (c), only the boundaries in the lower left part are detected where there is obvious difference in the 2nd- order features. In Fig. 4-14 (b), it only detected the 1st- order boundaries in the upper part. In the raw data of two detected boundaries, it is obvious that it is insufficient to detect all boundaries by a single order feature. In Fig. 4-14 (d), hybrid-order features are considered simultaneously and all boundaries are detected successfully. A superposition of input image and hybrid-order boundary after local peak detection is also shown in Fig. 4-15, where detection results with good accuracy are also demonstrated.

(a)   (b)

(c)   (d)
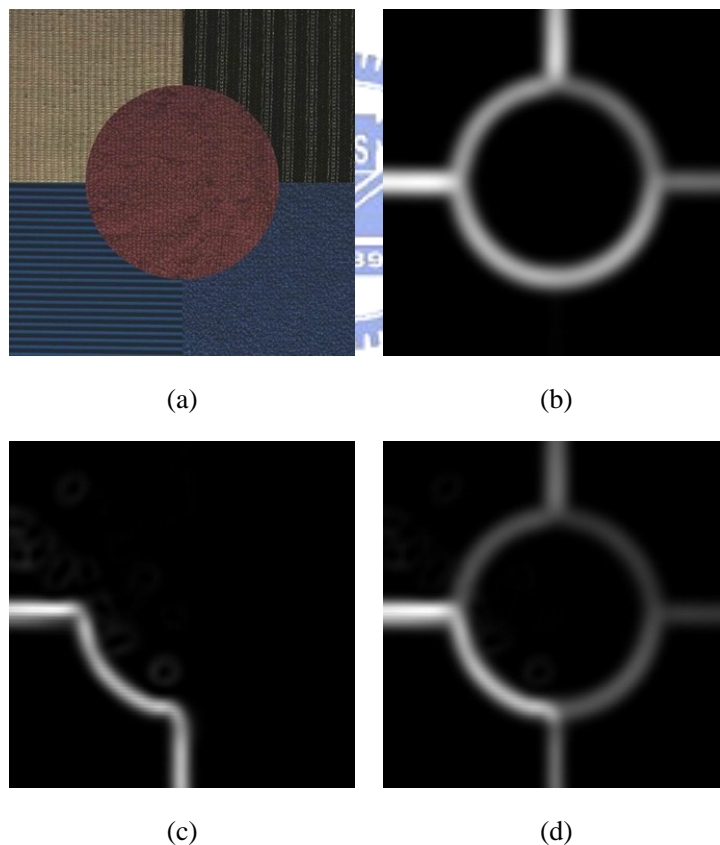
Figure 4-14: An example for demonstrating the effect of hybrid-order features: (a) input image; (b) 1st- order boundary; (c) 2nd- order boundary; (d) hybrid-order boundary.
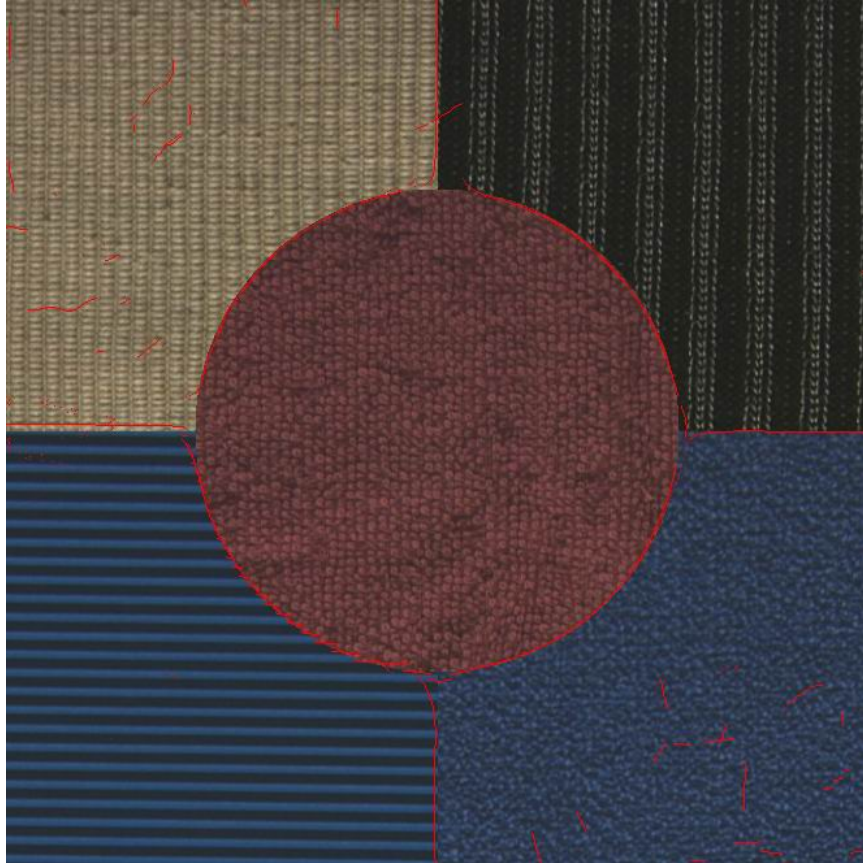
Figure 4-15: Superposition of input image and boundary after local peak detection.

4.3 Collection of Testing Results by Hybrid-Order Boundary Detection

In this section the proposed algorithm is tested by a large amount of textures randomly chosen from "Outex database." We synthesized five textures in each image, which results in eight boundaries. We will show the raw data of boundaries, boundaries after peak detection, and superposition of boundaries and testing images in order. There are totally 35 testing results, and all parameters we used are the same as we mentioned in Section 4.1.

In this section, we classified our experimental results into two categories roughly. In Section 4.3.1 we collect the results whose all boundaries are successfully detected. The results whose some boundaries are missed are collected in Section 4.3.2.

4.3.1 Fully Boundary Detection

In this subsection we focus on uniform texture which consists of similar textural elements. We will demonstrate some results (from Fig. 4-16 to Fig. 4-20) where all the boundaries between different textures are detected and weak edges within single texture are also detected. These results are consistent to our visual perception. As mentioned in Section 4.2.3, it cannot perform well without considering hybrid-order features simultaneously. Large amounts of testing images demonstrated below contain discriminable 1st- and/or 2nd- order features, and thus all boundaries are successfully detected. And, of course, a region with more regular texture and/or more uniform luminance or color can provide better discriminability than others, and boundaries surrounding it will be much more obvious and compact than others.
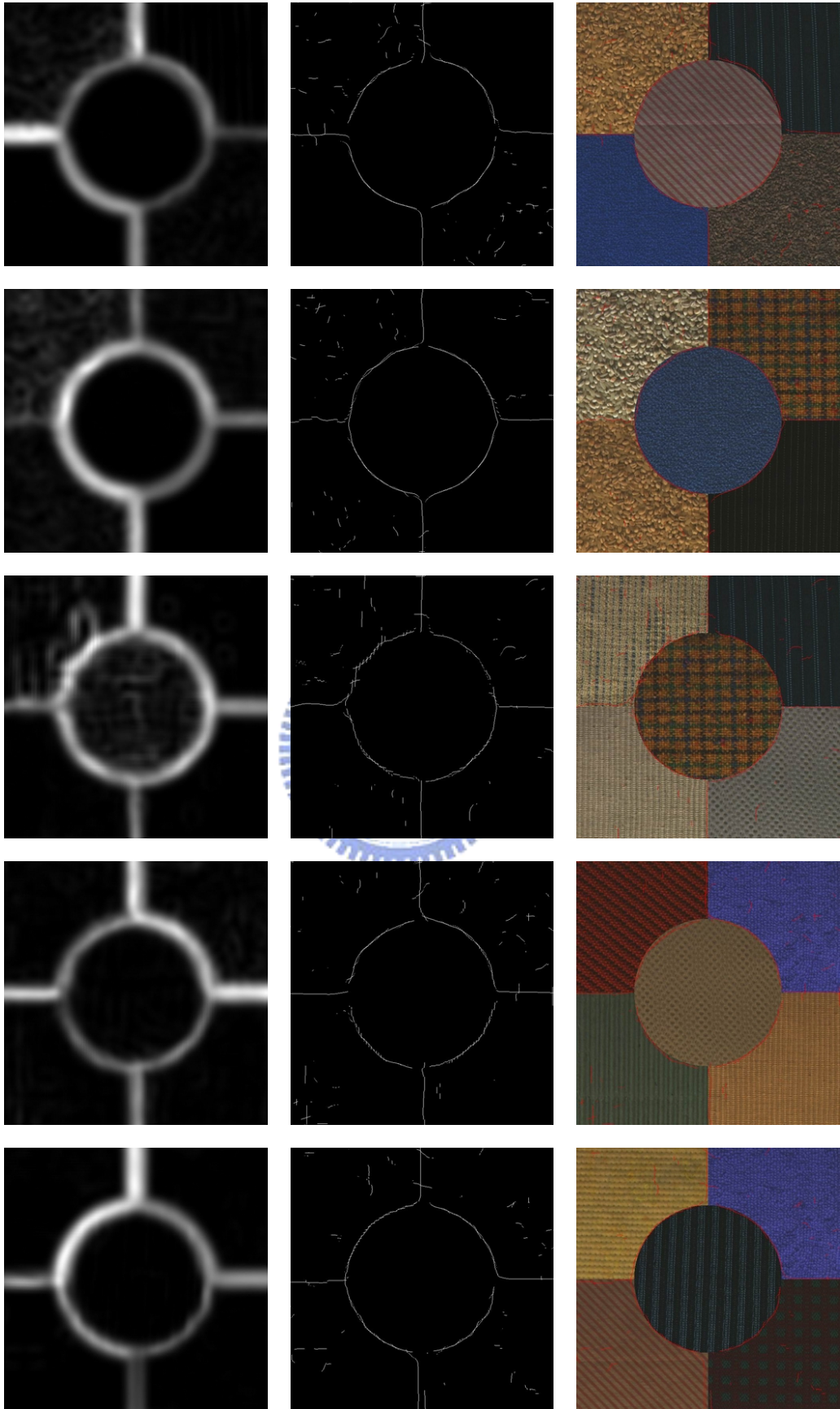
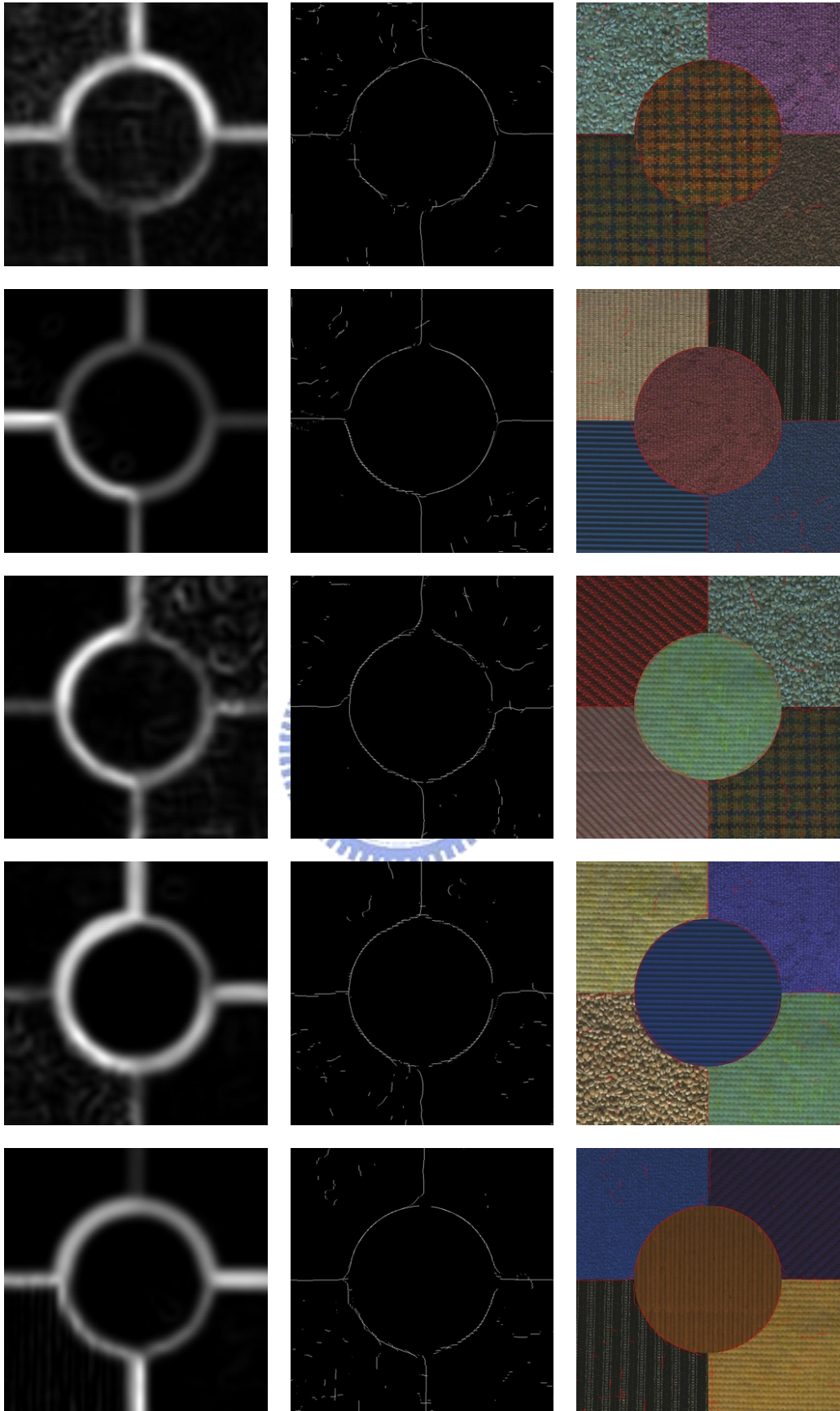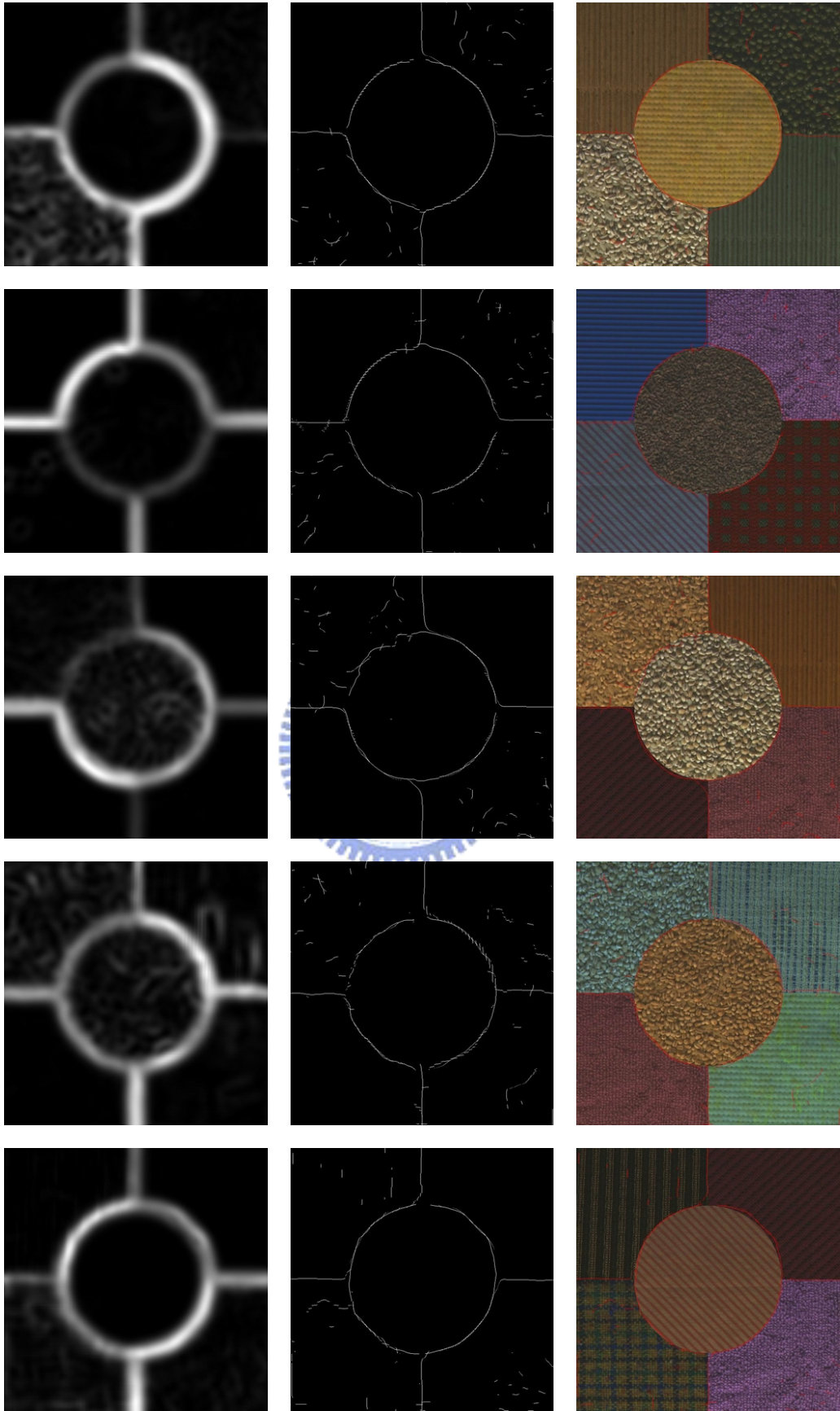**Figure 4-16:** Examples of fully boundary detection.
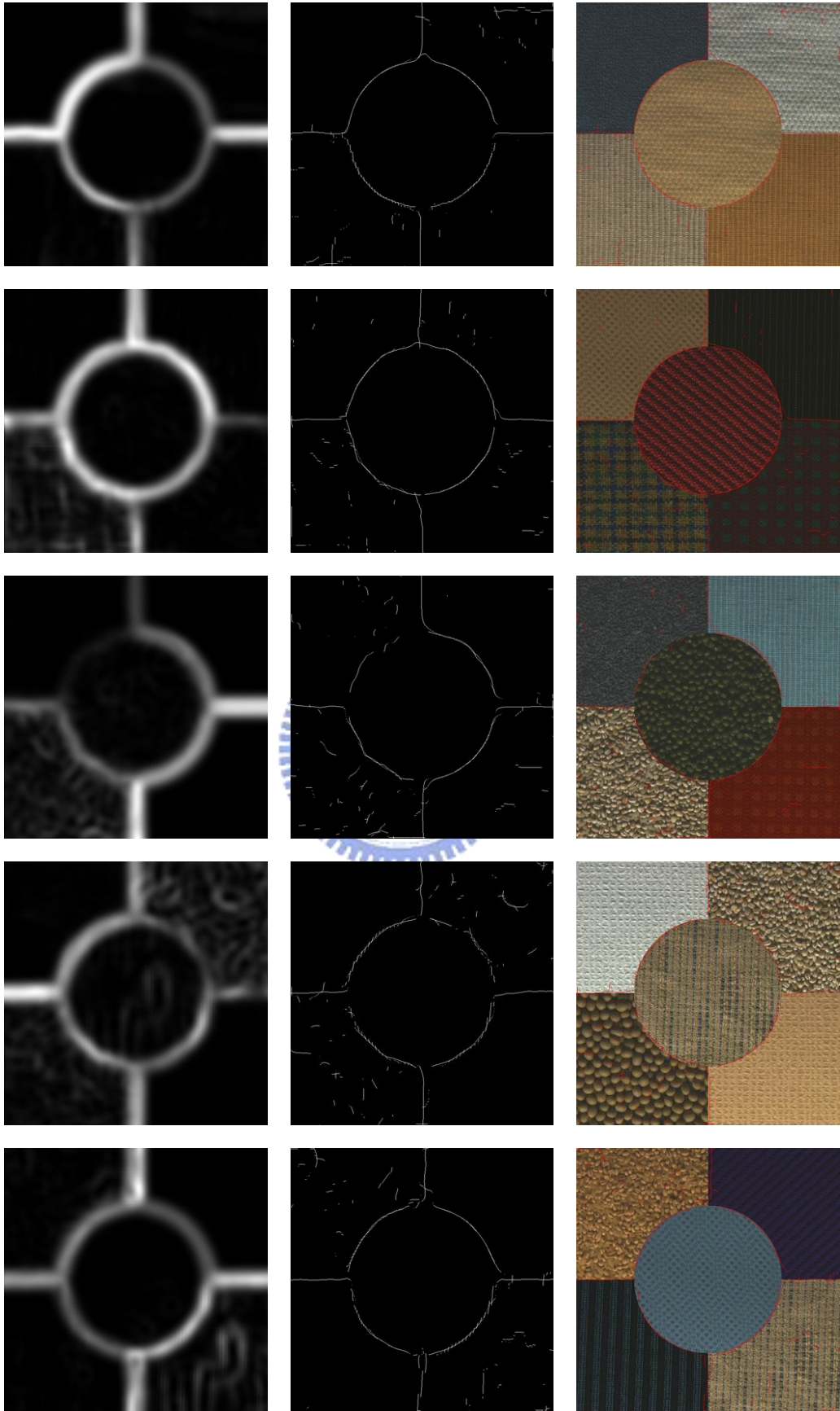
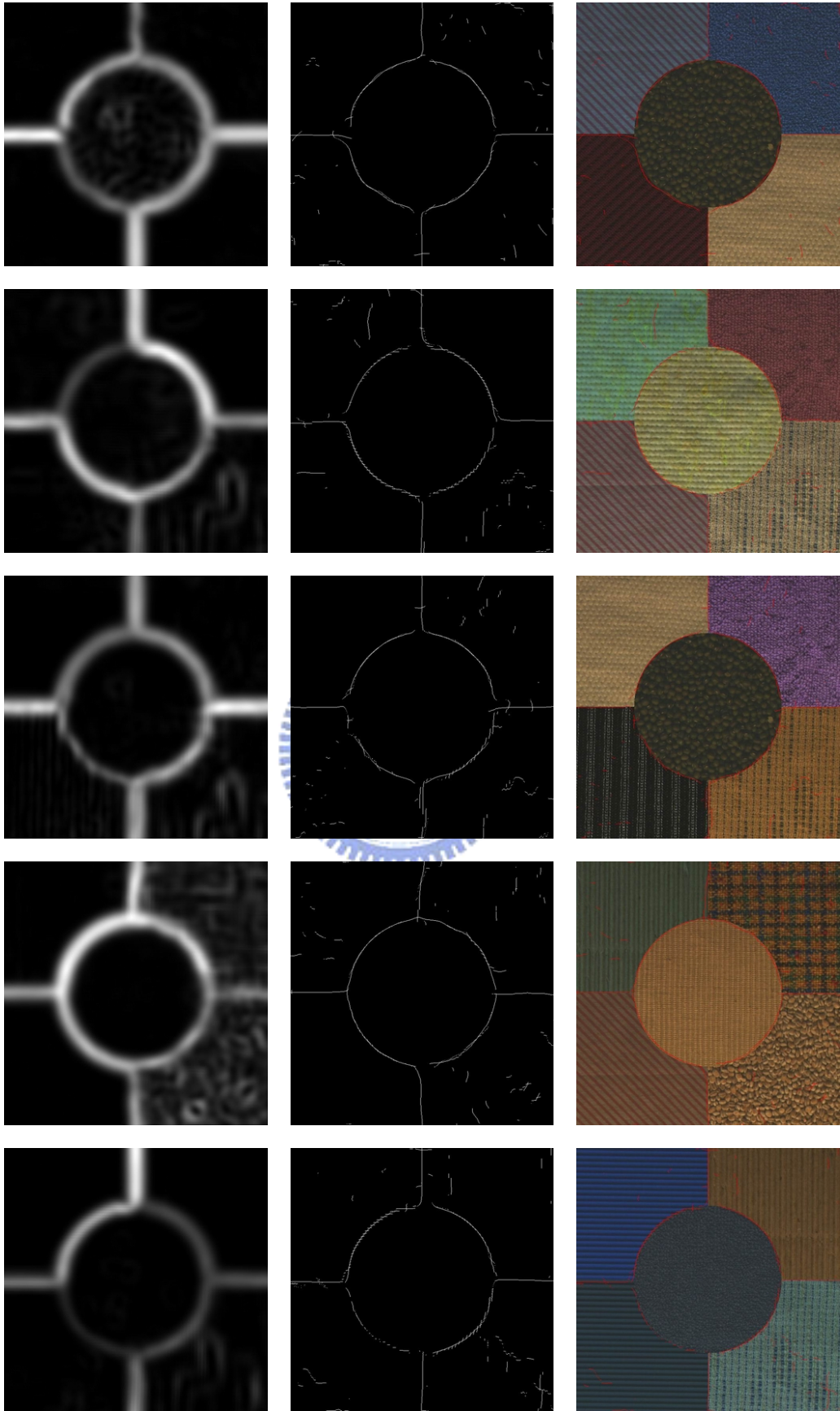**Figure 4-17:** Examples of fully boundary detection.

Figure 4-18: Examples of fully boundary detection.

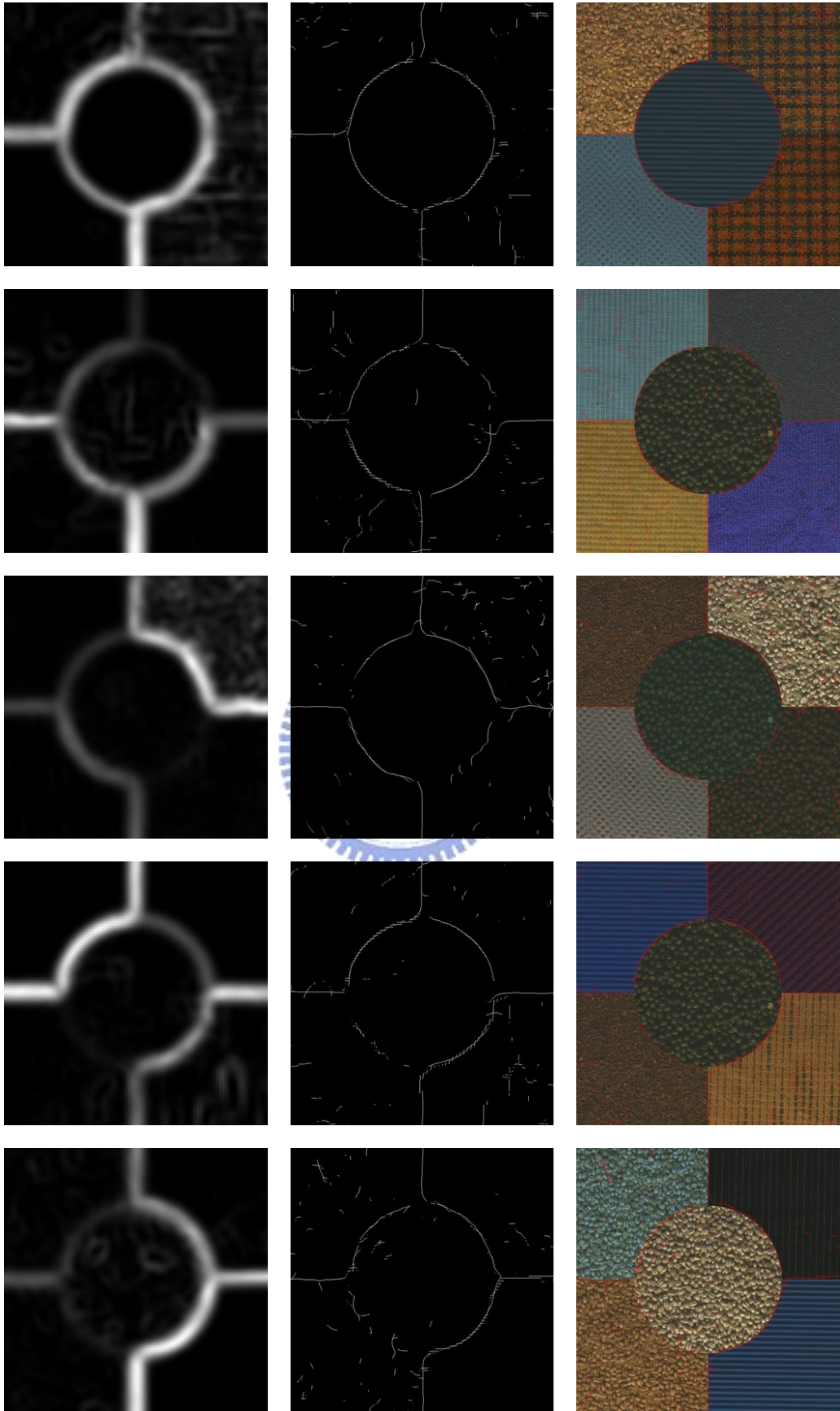**Figure 4-19:** Examples of fully boundary detection.

Figure 4-20: Examples of fully boundary detection.

4.3.2 Partially Boundary Detection

In this subsection we will demonstrate the results (Fig. 4-21 and Fig. 4-22) where some boundaries are not correctly detected. Some textures in this subsection are not so uniform that boundaries within single texture are more obvious than boundaries between different textures. In some cases, two textures have similar features, and the boundaries cannot be detected as we even cannot distinguish them at our first sight.

The visual information in the proposed approach is represented in a relative form as the main concept of receptive field. Therefore among eight boundaries in a testing image, the more obvious a boundary is, the larger values of raw data are. Although an adaptive weights selection can mediate the effects between the 1st- and 2nd- order features, the weights assigned by the mechanism are global; that is, the adaptive selected weights are determined by the most significant 1st- and 2nd- order boundaries. In other words, the context will influence the segregation results. Two regions with less difference of the 1st- order feature and non-uniform textural elements, the information of 1st- order boundary should be depressed during normalization before weights selection and sometimes the weak edges from the 2nd- order features would dominate the result. The phenomenon is inevitable due to the proposed algorithm mimicking one sight of our vision system; that is, to segregate more significant parts in a view. Even two regions could be easily segregated if they were the only contents in a view. In a multi-texture case, however, boundary between them might be ignored if the boundary is less obvious compared to other ones. To verify the statement, see the testing image with blinking eyes and some boundaries seem not as obvious as we "focus" on them. In actual, as we focus on something, it takes a few sights already and vision system has tuned a good way to observe since the weights have been specifically mediated for the focused view.

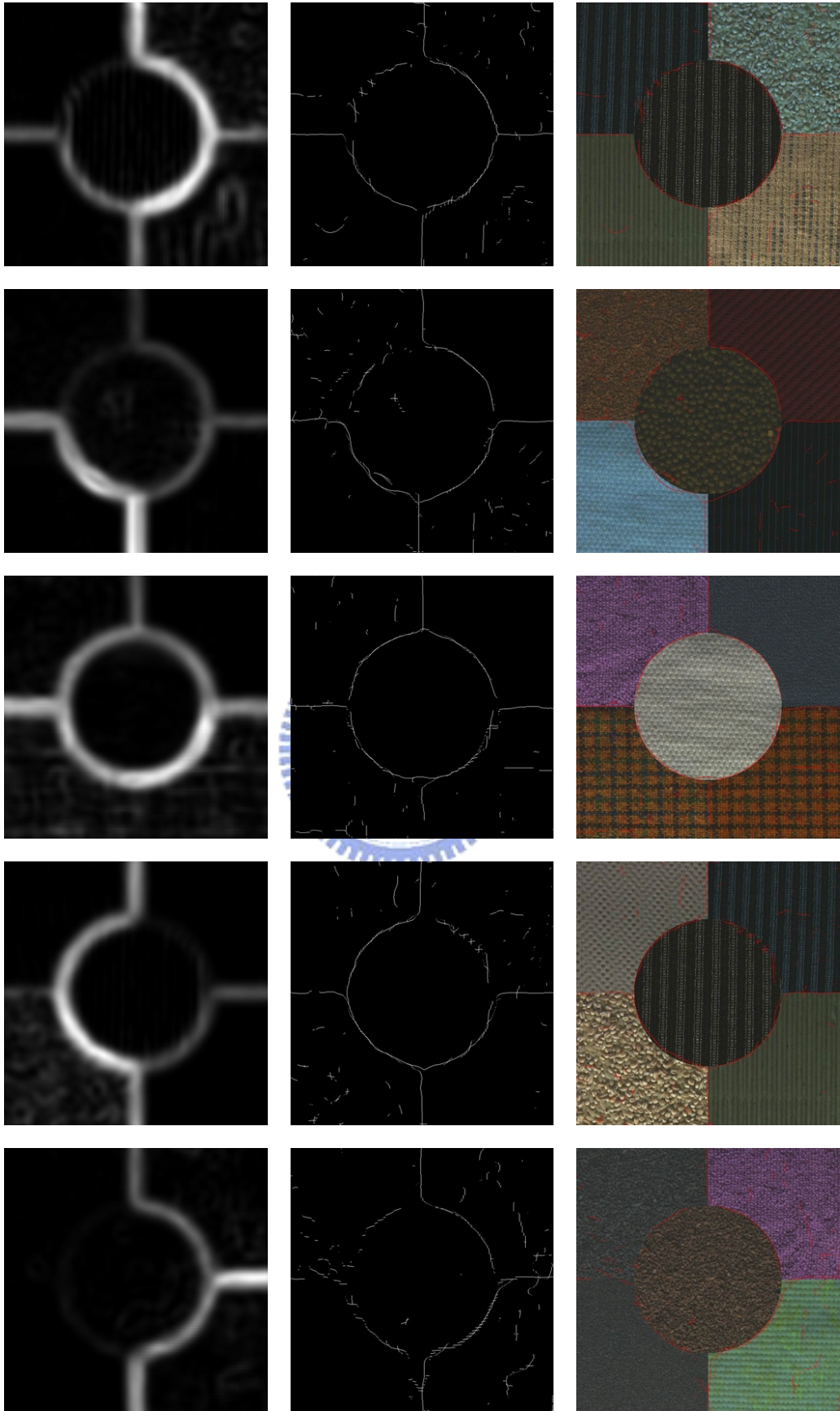**Figure 4-21:** Examples of partially boundary detection.

Figure 4-22: Examples of partially boundary detection.

4.4 Error Estimation

In this section the accuracy of the proposed method is discussed. The way we estimated the error can be shown in Fig. 4-23. The five-texture pattern was not tested for error estimation because the direction to calculate the error of circle contour is hard to assign. Moreover, since there are totally eight boundaries, if some boundaries are detected and some are not, the error estimated cannot clearly represent how the error results from (missed boundary or inaccurate location). The error we defined is:

$$Error = \frac{1}{H}\sum_{i=1}^{H}\frac{\Delta W_i}{W} + \frac{1}{W}\sum_{j=1}^{W}\frac{\Delta H_j}{H} . \qquad (4-1)$$

The distance between the answer and result detected by the proposed algorithm is measured and then divided by the number of total pixels to obtain the estimated error.
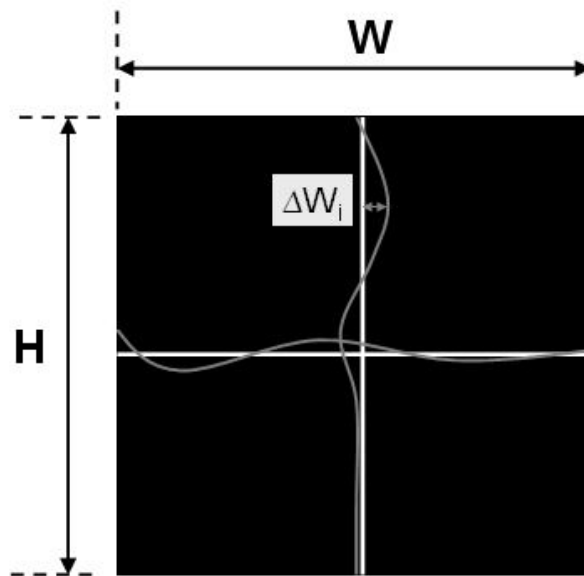
Figure 4-23: Schematic diagram of error estimation.

Fig. 4-24 is the histogram of errors estimated in our experiments. The mean of error is about 3.60% and standard deviation is 0.0197. The mean error of 3.60% corresponds to approximate 12-pixel error which is very reasonable since the size of textural element is usually larger than 15 pixels in the testing image. There are some cases with errors larger than 8%; Fig. 4-25 to Fig. 4-27 are three typical cases. As

mentioned in Section 4.3.2, the proposed algorithm mimics one-sight vision, and thus the boundary between non-uniform textures with similar 1st- order features sometimes would be ignored and weak edge within single texture would be detected. The cases with larger errors in our experiments all belong to this condition.
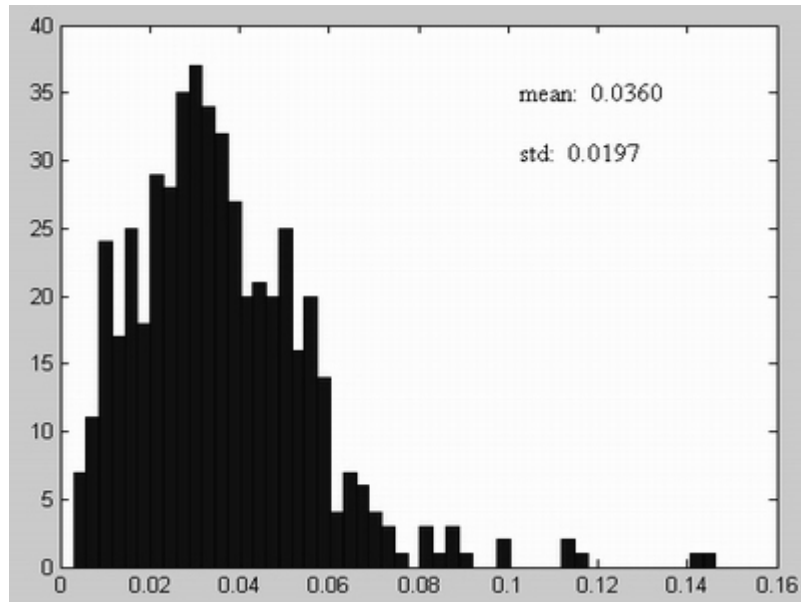
Figure 4-24: The histogram of estimated errors.
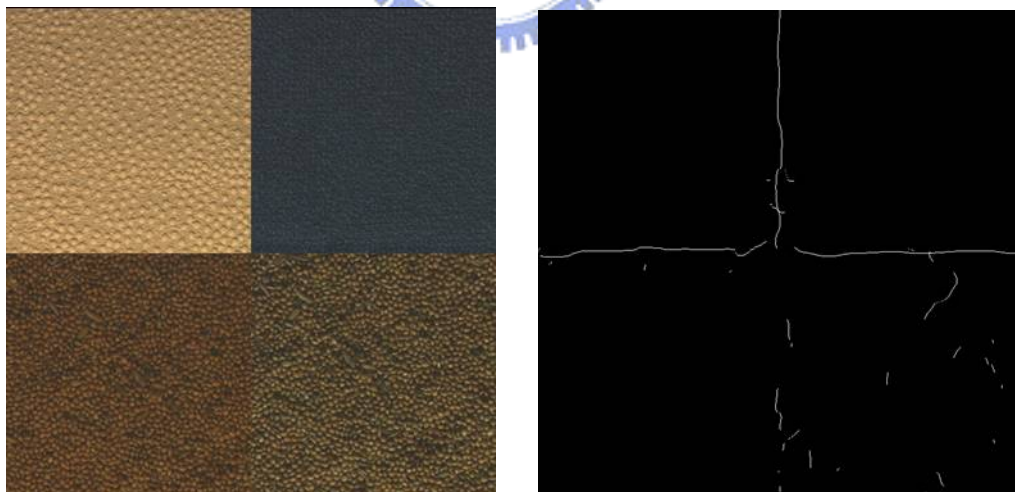
(a)   (b)

Figure 4-25: An example（Ⅰ）with large estimation error.
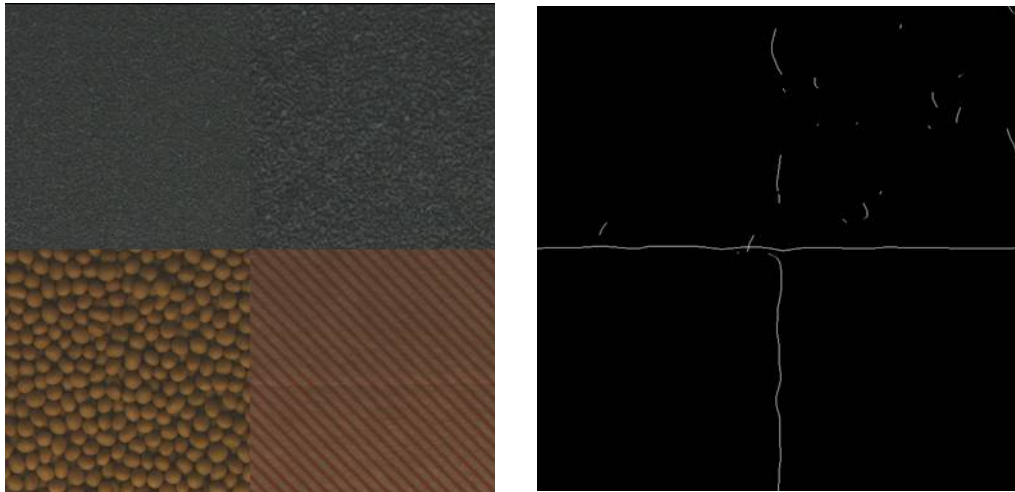
(a)   (b)

Figure 4-26: An example (Ⅱ) with large estimation error.
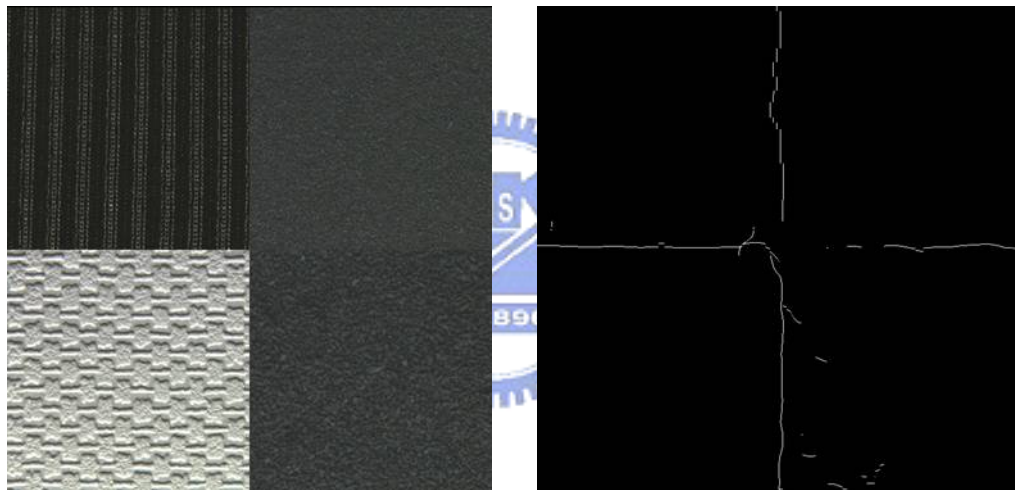
(a)   (b)

Figure 4-27: An example (Ⅲ) with large estimation error.

To demonstrate the essentiality of color information, a hybrid-order model without color information is also applied to the set of testing images and the estimated errors were also recorded. The histogram of estimated errors without color information is shown in Fig. 4-28, and the mean of errors obviously rises from 3.60% to about 5.91% (about 65% increases). Some patterns could be easily segregated with color information seem ambiguous without considering color information as shown from Fig. 4-29 to Fig. 4-31.
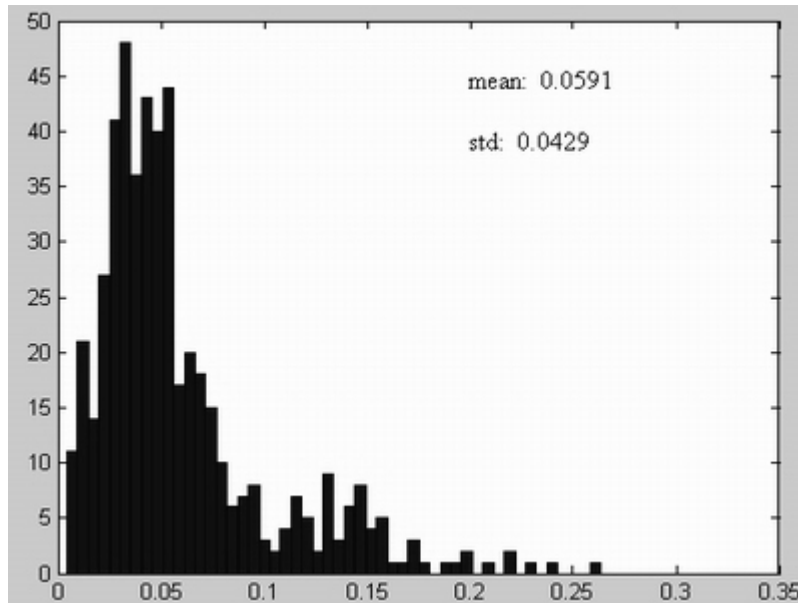
Figure 4-28: The histogram of estimated errors without considering color information.

(a)   (b)

(c)   (d)
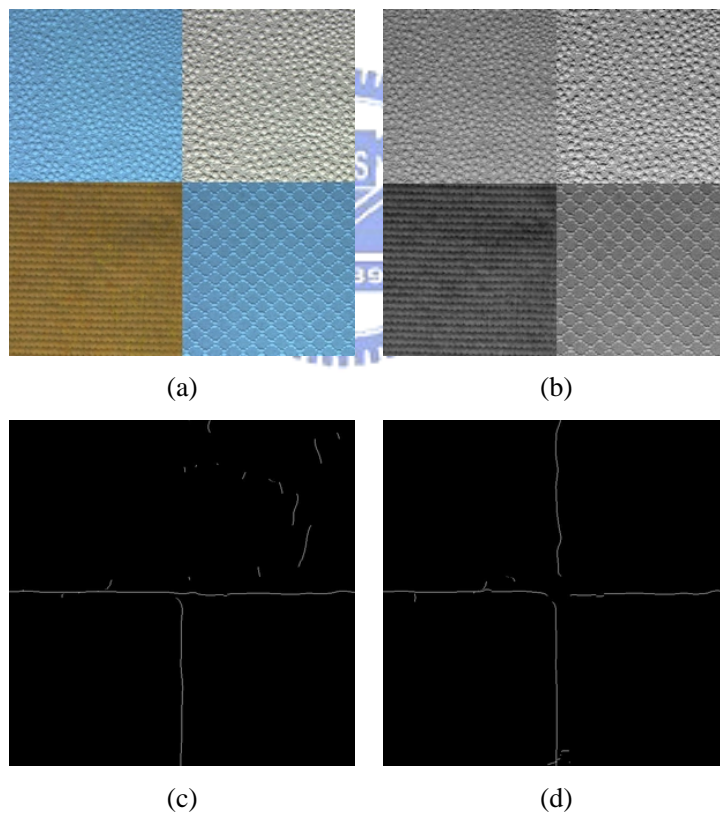
Figure 4-29: An example demonstrating the essentiality of color: (a) input image; (b) luminance;

(c) detected boundary without color information; (d) detected boundary with color information.

(a)   (b)

(c)   (d)
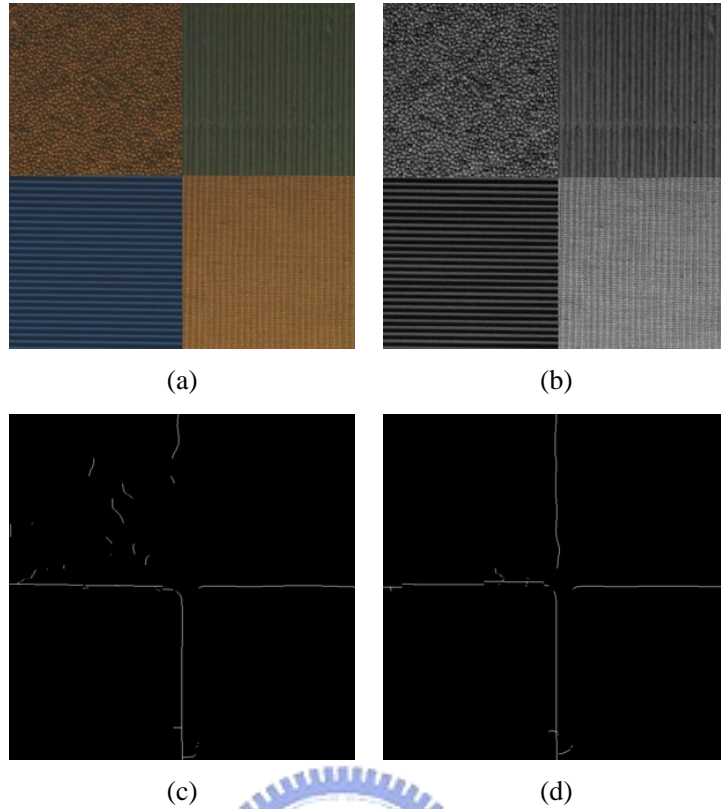
Figure 4-30: An example demonstrating the essentiality of color: (a) input image; (b) luminance; (c) detected boundary without color information; (d) detected boundary with color information.

(a)   (b)

(c)   (d)
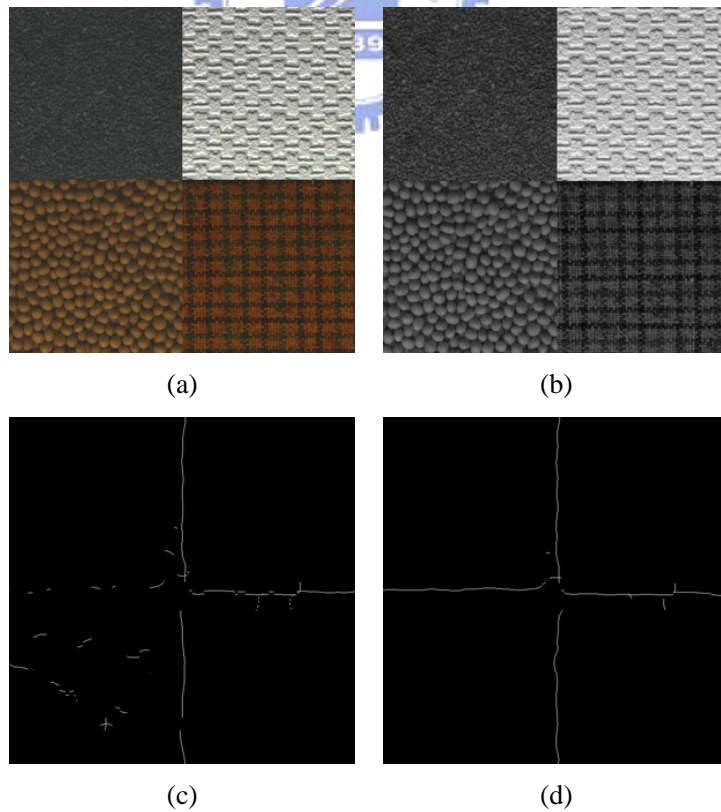
Figure 4-31: An example demonstrating the essentiality of color: (a) input image; (b) luminance; (c) detected boundary without color information; (d) detected boundary with color information.

Chapter 5

Conclusions & Future Works

In this thesis, a framework for hybrid-order boundary detection has been proposed. It mimics visual processing at very early stages of human vision where sensation and perception appear. The main contribution of this thesis is to combine three important visual primitives: luminance, texture, and color properly and adaptively. Issues for general-purpose procedure ignored before were also discussed thoroughly and solved, such that the proposed model can represent visual information for segregation task in a complete and economic way. There are still few researches reaching this goal up to the present, and previous researches usually had specific applications and employed some assumptions. The proposed approach dealt with the task with few assumptions and no training procedure involved, so a larger application space is available.

The experimental results show consistency with sensation of human visual system and the detected boundaries with adequate accuracy demonstrate application potential for other image processing tasks such as stereo, pattern recognition, retrieval, *etc*. A related work [46] considering about the hardware implementation of the framework has successfully realized the model for gray-scale images on the architecture called Cellular Nonlinear/Neural Networks (CNN). CNN is capable of parallel processing and realizable by VLSI circuits such that the computational time can greatly decrease. The real-time processing capability is critical in some applications such as tracking and surveillance system.

The proposed algorithm was widely tested to detect the boundaries of synthesized textures and the experimental results appeared similar responses to our

sensation. However, there are still some problems necessary to be overcome:

(i) The foremost one is the parameters selection of Gabor filters. The Gabor wavelets scheme has substantially relieved the condition that there are only few parameters need to be determined. However, there is still no systematic and efficient way to represent the relationship between image content and parameter. In previous ten years, many researches focused on the issue for pure texture. For more complex contents consist of hybrid-order features, the parameters selection task will be more challenging.

(ii) Another issue to be improved is about the adaptive weights selecting mechanism. In Section 4.3.2, we discussed some cases where some boundaries were not detected. The main cause is that the present approach mimics one sight of vision system and extracts most significant parts in one view. To deal with more complex cases, a mechanism capable of multiple focuses is undoubtedly required, such that weights of the 1st- and 2nd- order boundaries can be determined rather locally than globally. By someway analyzing local information, it is believed that weights can be assigned more properly to different parts of an image.

References

[1] B. Julesz, E. N. Gilbert, L. A. Shepp, and H. L. Frisch, "Inability of humans to discriminate between visual textures that agree in 2nd-order statistics--revisited," *Perception*, vol. 2, pp. 391-405, 1973.

[2] D. Marr, *Vision*. San Francisco, CA: W. H. Freeman, 1982.

[3] W. Richards, "Quantifying sensory channels: Generalizing colorimetry to orientation and texture, touch, and tones," *Sensory Processes*, vol. 3, pp. 207-229, 1979.

[4] J. K. Hawkins, "Textural properties for pattern recognition," in *Picture Processing and Psychopictorics*, B. Lipkin and A. Rosenfeld Eds., Academic Press, New York, 1969.

[5] H. Tamura, S. Mori, and Y. Yamawaki, "Texture features corresponding to visual perception," *IEEE Transactions on System, Man, and Cybernetics*, vol. 8, pp. 460-473, 1978.

[6] J. Sklansky, "Image segmentation and feature extraction," *IEEE Transactions on System, Man, and Cybernetics*, vol. 8, pp. 237-247, 1978.

[7] B. Julesz, "Visual pattern discrimination," *IRE Transactions on Information Theory, IT-8*, pp. 84-92, 1962.

[8] J. Beck, "Perceptual grouping produced by changes in orientation and shape," *Science*, vol. 154, pp. 538-540, 1966.

[9] B. Julesz, "Experiments on the visual perception of texture," *Scientific American*, vol. 232, pp. 34-43, 1975.

[10] B. Julesz, "Visual texture discrimination using random-dot patterns: Comment," *Journal of the Optical Society of America*, vol. 69, pp. 268-270, 1978.

[11] B. Julesz, "Textons, the elements of texture perception and their interactions," *Nature*, vol. 290, pp. 91-97, 1981.

[12] J. Beck, "Texture segmentation," in *Organization and Representation in Perception*, J. Beck Ed., Hillside, NJ: Erlbaum, 1982.

[13] J. Beck, A. Sutter, and R. Ivry, "Spatial frequency channels and perceptual grouping in texture perception," *Computer Vision Graphics Image Processing*, vol. 37, pp. 299-325, 1987.

[14] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers*, vol. 93, pp. 429-457, 1946.

[15] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London (B)*, pp. 187-217, 1980.

[16] J. D. Daugman, "Two dimensional spectral analysis of cortical receptive field profiles," *Vision Research*, vol. 20, pp. 847-856, 1980.

[17] J. D. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, vol. 2, pp. 1160-1169, 1985.

[18] J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature*, vol. 333, pp. 363-364, 1988.

[19] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 55-73, 1990.

[20] J. R. Bergen, "Theories of visual texture perception," in *Vision and Visual Dysfunction, vol. 10B*, D. Regan, Ed., New York: MacMillan, 1991.

[21] N. Graham, "Complex channels, early local nonlinearities, and normalization in texture segregation," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, Eds., Cambridge, MA: MIT Press, 1991.

[22] M. R. Turner, "Texture discrimination by Gabor functions," *Biological Cybernetics*, vol. 55, pp. 71-82, 1986.

[23] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *Journal of the Optical Society of America A*, vol. 7, pp. 923-932, 1990.

[24] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, pp. 1167-1186, 1991.

[25] B. S. Manjunath and R. Chellapa, "A unified approach to boundary perception: Edges, texture, and illusory contours," *IEEE Transactions on Neural Networks*, vol. 4, pp. 96-108, 1993.

[26] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 837-842, 1996.

[27] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674-693, 1989.

[28] M. Unser and M. Eden, "Multiresolution feature extraction and selection for texture segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 717-728, 1989.

[29] H. W. Tang, V. Srinivasan, and S. H. Ong, "Texture segmentation via nonlinear interactions among Gabor feature pairs," *Optical Engineering*, vol. 34, pp. 125-134, 1995.

[30] T. N. Tan, "Texture edge detection by modeling visual cortical channels," *Pattern*

Recognition, vol. 28, pp. 1283-1298, 1995.

[31] D. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation," *IEEE Transactions on Image Processing*, vol. 4, pp. 947-964, 1995.

[32] T. Weldon and W. E. Higgins, "Designing multiple Gabor filters for multitexture image segmentation," Optical Engineering, vol. 38, pp. 1478-1489, 1999.

[33] A. Teuner, O. Pichler, and B. J. Hosticka, "Unsupervised texture segmentation of images using tuned matched Gabor filters," *IEEE Transactions on Image Processing*, vol. 4, pp. 863-870, 1995.

[34] O. Pichler, A. Teuner, and B. J. Hosticka, "An unsupervised texture segmentation algorithm with feature space reduction and knowledge feedback," *IEEE Transactions on Image Processing*, vol. 7, pp. 53-61, 1998.

[35] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1169-1179, 1988.

[36] 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, pp. 452-468, 1988.

[37] W. McIlhagga, T. Hine, G. R. Cole, and A. W. Snyder, "Texture segregation with luminance and chromatic contrast," *Vision Research*, vol. 30, pp. 489-495, 1990.

[38] T. V. Papathomas, R. S. Kashi, and A. Gorea, "A human vision based computational model for chromatic texture segregation," *IEEE Transactions on System, Man, and Cybernetics-Part B: Cybernetics*, vol. 27, pp. 428-440, 1997.

[39] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Transactions on Image Processing*, vol. 7, pp. 124-128, 1998.

[40] M. Mirmehdi and M. Petrou, "Segmentation of color textures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 142-159, 2000.

[41] J. F. Camapum Wanderley and M. H. Fisher, "Multiscale color invariants based on the human visual system," *IEEE Transactions on Image Processing*, vol. 10, pp. 1630-1638, 2001.

[42] L. G. Thorell, R. L. De Valois, and D. G. Albrecht, "Spatial mapping of monkey V1 cells with pure color and luminance stimuli," *Vision Research*, vol. 24, pp. 751-769, 1984.

[43] A. Bradley, E. Switkes, and K. K. De Valois, "Orientation and spatial frequency selectivity of adaptation to colour and luminance gratings," *Vision Research*, 28, pp. 841-856, 1988.

[44] M. A. Webster, K. K. De Valois, and E. Switkes, "Orientation and

spatial-frequency discrimination for luminance and chromatic gratings," *Journal of the Optical Society of America A*, vol. 7, pp. 1034-1049, 1990.

[45] E. Hering, *Outlines of a theory of the light sense*. Translated by L. M. Hurvish and D. Jameson. Cambridge, MA: Harvard University Press.

[46] S. A. Chen, "CNN-based texture boundary detection technique and its analog circuit implementation," Master Thesis, National Chiao-Tung University, 2004.

[47] R. L. De Valois and K. K. De Valois, *Spatial Vision*. New York: Oxford University Press, 1988.

[48] D. H. Hubel, *Eye, Brain, and Vision.* Scientific American Library. New York: W. H. Freeman, 1988.

[49] S. W. Kuffler, "Discharge patterns and functional organization of mammalian retina," *Journal of Neurophysiology*, vol. 16, pp. 37-68, 1953.

[50] H. B. Barlow, "Summation and inhibition in the frog's retina," *Journal of Physiology*, vol. 119, pp. 69-88, 1953.

[51] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurons in the cat's striate cortex," *Journal of Physiology*, vol. 148, pp. 574-591, 1959.

[52] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction, and functional architecture of the visual cortex," *Journal of Physiology*, vol. 160, pp. 106-154, 1962.

[53] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *Journal of Physiology*, vol. 195, pp. 215-243, 1968.

[54] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *Journal of Physiology*, vol. 197, pp. 551-516, 1968.

[55] K. K. De Valois, R. L. De Valois, and E. W. Yund, "Responses of striate cortex cells to grating and checkerboard patterns," *Journal of Physiology*, vol. 291, pp. 483-505, 1979.

[56] A. Bradley, E. Switkes, and K. K. De Valois, "Orientation and spatial frequency selectivity of adaptation to colour and luminance gratings," *Vision Research*, 28, pp. 841-856, 1988.

[57] J. P. Jones and L. A. Palmer, "The two-dimensional spatial structure of simple receptive fields in cat striate cortex," *Journal of Neurophysiology*, vol. 58, pp. 1186-1211, 1987.

[58] R. L. De Valois, I. Abramov, and G. H. Jacobs, "Analysis of response patterns of LGN cells," *Journal of the Optical Society of America*, vol. 56, pp. 966-977, 1966.

[59] M. S. Livingstone and D. H. Hubel, "Segregation of form, color, movement and depth: Anatomy, physiology and perception," *Science*, vol. 240, pp. 740-749, 1988.

[60] A. Treisman and G. Gelade, "A feature integration theory of attention," *Cognitive Psychology*, vol. 12, pp. 97-136, 1980.

[61] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd Ed., Prentice Hall, 2002.

[62] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698, 1986.

[63] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological Cybernetics*, vol. 61, pp. 103-113, 1989.

[64] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *Journal of the Optical Society of America*, vol. 70, pp. 1297-1300, 1980.

[65] D. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science*, vol. 212, pp. 1409-1411, 1981.

[66] J. J. Kulikowski, S. Marcelja, and P. O. Bishop, "Theory of spatial position and spatial frequency relations in the receptive field of simple cells in the visual cortex," *Biological Cybernetics*, vol. 43, pp. 187-198, 1982.

[67] B. Sakitt and H. B. Barlow, "A model for the economical encoding of the visual image in cerebral cortex," *Biological Cybernetics*, vol. 43, pp. 97-108, 1982.

[68] D. A. Pollen and S. F. Ronner, "Visual cortical neurons as localized spatial frequency filter," *IEEE Transactions on System, Man, and Cybernetics*, vol. 13, pp. 907-916, 1983.

[69] S. G. Mallat, "Wavelets for a vision," *Proceedings of the IEEE*, vol. 84, pp. 604-614, 1996.

[70] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, pp. 532-540, 1983.

[71] CIE, *Uniform Color Space—Color Difference Equations—Psychometric Color Terms.* Commission Internationale de l'Eclairage, Publication No. 15, Supplement No. 2, Paris, 1978.

[72] E. Switkes, A. Bradley, and K. K. De Valois, "Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings," *Journal of the Optical Society of America A*, vol. 5, pp. 1149-1162, 1989.

[73] M. Bach, C. Schmitt, T. Quenzer, T. Meigen, and M. Fahle, "Summation of texture segregation across orientation and spatial frequency: Electrophysiological and psychophysical findings," *Vision Research*, vol. 40, pp. 3559-3566, 2000.

[74] E. N. Johnson, M. J. Hawken, and R. Shapley, "The spatial transform of color in the primary visual cortex of the macaque monkey," *Nature Neuroscience*, vol. 4, pp. 409-416, 2001.

[75] D. Schluppeck and S. A. Engel, "Color opponent neurons in V1: A review and model reconciling results from imaging and single-uint recording," *Journal of*

Vision, vol. 2, pp. 480-492, 2002.

[76] R. Shapley and M. Hawken, "Neural mechanisms for color perception in the primary visual cortex," *Current Opinion in Neurobiology*, vol. 12, pp. 426-432, 2002.

[77] A. Li and P. Lennie, "Mechanisms underlying segmentation of colored textures," *Vision Research*, vol. 37, pp. 83-97, 1997.

[78] P. M. Pearson and F. A. A. Kingdom, "Texture-orientation mechanisms pool colour and luminance contrast," *Vision Research*, vol. 42, pp. 1547-1558, 2002.

[79] K. T. Mullen and M. A. Losada, "Evidence for separate pathways for color and luminance detection mechanisms," *Journal of the Optical Society of America A*, vol. 11, pp. 3136-3151, 1994.

[80] B. Poirson and B. Wandell, "Pattern-color separable pathways predict sensitivity to simple colored patterns," *Vision Research*, vol. 36, pp. 515-526, 1996.

[81] 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, pp. 719-743, 1992.

[82] D. J. Heeger, "Normalization of cell responses in cat striate cortex," *Visual Neuroscience*, vol. 9, pp. 181-197, 1992.

[83] D. G. Albrecht and R. L. De Valois, "Striate cortex responses to periodic patterns with and without the fundamental harmonics," *Journal of Physiology*, vol. 319, pp. 495-514, 1981.

[84] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *International Journal of Computer Vision*, vol. 43, pp. 7-27, 2001.

[85] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 530-549, 2004.

[86] P. Kruizinga and N. Petkov, "Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells," *Biological Cybernetics*, vol. 76, pp. 83-96, 1997.

[87] P. Kruizinga and N. Petkov, "Nonlinear operator for oriented texture," *IEEE Transactions on Image Processing*, vol. 8, pp. 1395-1407, 1999.

[88] S. E. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of texture features based on Gabor filters," *IEEE Transactions on Image Processing*, vol. 11, pp. 1160-1167, 2002.

[89] J. Rivest and P. Cavanagh, "Localizing contours defined by more than one attribute," *Vision Research*, vol. 36, pp. 53-66, 1996.

[90] M. S. Landy and H. Kojima, "Ideal cue combination for localizing

texture-defined edges," *Journal of the Optical Society of America A*, vol. 18, pp. 2307-2320, 2001.

[91] P. V. McGraw, D. Whitaker, D. R. Badcock, and J. Skillen, "Neither here nor there: Localizing conflicting visual attributes," *Journal of Vision*, vol. 3, pp. 265-273, 2003.

[92] University of Oulu Texture Database [Online]. Available: http://www.outex.oulu.fi/outex.php.