INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
AN ADAPTIVE TEXTURE SEGMENTATION APPROACH FOR APPLICATIONS IN DIGITAL PHOTOGRAMMETRY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Dong-Cheon Lee, B.E., M.E., M.S.

* * * * *

The Ohio State University
1997

Dissertation Committee:
Dr. Anton E. Schenk, Advisor
Dr. Rongxing Li
Dr. Alan Saalfeld

Approved by
Advisor

Graduate Program in Geodetic Science and Surveying
ABSTRACT

Texture plays an important role in visual information processing since texture provides useful information about shape, orientation, and depth of the objects. The human visual system utilizes texture without difficulty as one of the visual cues for image interpretation, scene analysis and object recognition. However, to extract and to analyze texture are difficult tasks in machine perception including computer vision and digital photogrammetry. The ultimate goal of digital photogrammetry is to reconstruct surfaces automatically. Surface reconstruction from raw imagery is an ill-posed problem. In order to solve the ill-posed nature of surface reconstruction, it is crucial to use information from different visual cues.

In this study, an adaptive strategy was developed by using Gabor filters in order to extract texture information and to segment images. The Gabor filters are conceived as hypothetical structures of the retinal receptive fields. Therefore, to develop a texture analysis system which resembles the performance of human visual perception is possible using the Gabor filters A scheme to select appropriate parameters of the Gabor filters without human visual inspection was proposed. The entire framework is based on the theory of human texture perception which was introduced by Julesz (1981).
Various kinds of images were used to evaluate the performance of the proposed strategy. The results show that the proposed adaptive strategy improves performance of the Gabor filters for texture extraction and segmentation.
Dedicated to my parents, wife and daughter
ACKNOWLEDGMENTS

Studying at the Department of Geodetic Science for a long time would not have been possible for me without invaluable support and help from many persons whom I would like to gratefully acknowledge.

First of all, I would like to express my sincere gratitude to my adviser, Dr. Anton F. Schenk for his invaluable guidance, encouragement, patience and insight throughout this research for many years. Working as a research and teaching associate under his guidance has been the most valuable experience in my years of education. I would like to extend my sincere thanks to other members of my advisory committee, Dr. Rongxing Li and Dr. Alan Saalfeld for their thoughtful comments and suggestions.

A special appreciation goes to Dr. Christopher Jekeli and Dr. Ayman Habib who have provided me their dedicated time for discussion and kind concern. I would like to thank Mr. Andre Templer (Photogrammetry Laboratory supervisor) and Mr. Dan Vehr (Computing Facility system manager) for their technical support and providing me with access to various systems. My special gratitude is also extended to Dr. Christian Heipke at Technical University Munich who supported and encouraged me with his enthusiasm while staying here as a visiting professor.
I would like to thank all faculty members, staff and students for their support and help. To share our knowledge and thoughts with my colleagues was always constructive. In this regard I would like to mention Kechwan Hong (GIS), Hong-Gyoo Sohn (Remote Sensing), Jay Hyoun Kwon (GPS/INS), and Jen-Jer Jaw (Photogrammetry). Especially, discussion with my friend Grady Tuell at NOAA has broadened my knowledge from his excellent experiences. I also appreciate the work of Irene Tesfai who edited this manuscript for English without changing my style of expression.

Finally, and most important, I would like to express my deepest gratitude to my family. Love, encouragement, support, and patience from my wonderful parents, beloved wife Miran and daughter Ester Hyunjoo have truly made my fulfillment of study possible. I always thank God for allowing me such a great family and guiding me all along.
VITA

September 9, 1958 ....................................... Born - Seoul, Korea

1982 .............................................................. B.E. Civil Engineering,
Yonsei University
Seoul, Korea

1984 .............................................................. M.E. Civil Engineering
Yonsei University
Seoul, Korea

1984 - 1985 ................................................... Research Associate
Yonsei University
Seoul, Korea

1987 .............................................................. M.S. Geodetic Science and Surveying
The Ohio State University
Columbus, Ohio

1988 - 1992 ................................................... Graduate Research Associate
The Ohio State University
Columbus, Ohio

1994 - present ................................................ Graduate Teaching Associate
The Ohio State University
Columbus, Ohio

FIELDS OF STUDY

Major Field: Geodetic Science and Surveying
TABLE OF CONTENTS

PAGE

ABSTRACT ..................................................................................................................... ii
DEDICATION ................................................................................................................. iv
ACKNOWLEDGMENTS ................................................................................................ v
VITA ................................................................................................................................... vii
LIST OF TABLES ........................................................................................................... xi
LIST OF FIGURES ...................................................................................................... xii

CHAPTERS:

1. INTRODUCTION ................................................................................................... I
   1.1 Role of Texture ............................................................................................... 3
   1.2 Characteristics of Texture ............................................................................... 5
   1.3 Background and Motivation ........................................................................... 7
   1.4 Scope and Objective of the Research ............................................................. 8
   1.5 Organization of the Report ........................................................................... 10

2. STATUS OF TEXTURE ANALYSIS METHODS .............................................. 12
   2.1 Statistical Approach ...................................................................................... 13
       2.1.1 Co-occurrence Matrices ................................................................... 13
       2.1.2 Generalized Co-occurrence Matrices ............................................... 15
       2.1.3 Autocorrelation Function .................................................................. 16
       2.1.4 Texture Energy Measure .................................................................. 16
       2.1.5 Markov Random Field Models ........................................................ 17
       2.1.6 Simultaneous Autoregressive Models .............................................. 19
       2.1.7 Fractal Models ................................................................................ 19
       2.1.8 Spatial-Frequency Approach ........................................................... 20

viii
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td>Structural Approach</td>
<td>21</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Placement Rule</td>
<td>21</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Symbolic Description</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Statistical-Structural Approach</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Neural Network Approach</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>Summary</td>
<td>24</td>
</tr>
<tr>
<td>3.</td>
<td>TWO-DIMENSIONAL GABOR FUNCTIONS AS MULTI-CHANNEL FILTERS</td>
<td>26</td>
</tr>
<tr>
<td>3.1</td>
<td>Derivation of Two-Dimensional Gabor Functions</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Properties of Gabor Filters</td>
<td>30</td>
</tr>
<tr>
<td>3.3</td>
<td>Parameters of Gabor Filters</td>
<td>34</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Standard Deviation of the Gaussian</td>
<td>34</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Spatial Frequency</td>
<td>35</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Orientation</td>
<td>35</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Phase</td>
<td>36</td>
</tr>
<tr>
<td>4.</td>
<td>DESCRIPTION OF THE PROPOSED ADAPTIVE APPROACH FOR TEXTURE ANALYSIS</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Characteristics of an Image</td>
<td>39</td>
</tr>
<tr>
<td>4.2</td>
<td>Texton Theory of Human Texture Perception</td>
<td>40</td>
</tr>
<tr>
<td>4.3</td>
<td>Adaptive Strategy</td>
<td>45</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Progressive Segmentation</td>
<td>45</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Determination of the Gabor Filter Size</td>
<td>48</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Computation of Dominant Local Orientation</td>
<td>50</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Computation of Spatial Frequency</td>
<td>55</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Extraction of Texture Information</td>
<td>57</td>
</tr>
<tr>
<td>4.4</td>
<td>Texture Segmentation</td>
<td>58</td>
</tr>
<tr>
<td>5.</td>
<td>EXPERIMENTS AND RESULTS</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Description of the Test Images</td>
<td>62</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Marchetsreut Image</td>
<td>62</td>
</tr>
<tr>
<td>5.1.2</td>
<td>OSU Campus Stereopair</td>
<td>63</td>
</tr>
<tr>
<td>5.1.3</td>
<td>GPSVan Image</td>
<td>63</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Tire Image</td>
<td>63</td>
</tr>
<tr>
<td>5.1.5</td>
<td>Synthetic Aperture Radar Image</td>
<td>63</td>
</tr>
<tr>
<td>5.1.6</td>
<td>Halftone Image</td>
<td>64</td>
</tr>
<tr>
<td>5.1.7</td>
<td>Multi-Spectral Image</td>
<td>64</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Parameters of the Gabor filters of the <em>Marchesreut</em> image</td>
</tr>
<tr>
<td>5.2</td>
<td>Summary of the texture boundary images</td>
</tr>
<tr>
<td>5.3</td>
<td>Summary of the analysis</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Parameters of texture</td>
<td>6</td>
</tr>
<tr>
<td>1.2 Properties of texture</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Laws' one-dimensional texture energy masks</td>
<td>18</td>
</tr>
<tr>
<td>2.2 The four most important 5 × 5 convolution masks</td>
<td>18</td>
</tr>
<tr>
<td>3.1 (a) Real (even-symmetric), and (b) imaginary (odd-symmetric) components of a 2D Gabor functions with 16 pixel period and 45° orientation</td>
<td>32</td>
</tr>
<tr>
<td>3.2 An example of 2D Gabor filter bank</td>
<td>32</td>
</tr>
<tr>
<td>3.3 Spatial frequency responses of a 2D Gabor functions shown in Figure 3.1. (a) Real and (b) imaginary components</td>
<td>33</td>
</tr>
<tr>
<td>3.4 (a) Example of 2D Gabor filter bank in the spatial-frequency domain, and (b), (c), (d), (e) 2D Gabor filters in spatial domain</td>
<td>33</td>
</tr>
<tr>
<td>4.1 Entropy and variance of each block of the natural image</td>
<td>41</td>
</tr>
<tr>
<td>4.2 Relationship between entropy and variance of the Marchetsreut image</td>
<td>41</td>
</tr>
<tr>
<td>4.3 (a) Marchetsreut image, and (b) its blobs with $\alpha=0.55$</td>
<td>43</td>
</tr>
<tr>
<td>4.4 Progressive segmentation scheme</td>
<td>46</td>
</tr>
<tr>
<td>4.5 Problematic situation</td>
<td>46</td>
</tr>
<tr>
<td>4.6 Relationship between entropy and blob size</td>
<td>50</td>
</tr>
<tr>
<td>4.7 Computation of the local dominant orientation ($\theta_o$) in a group of blobs</td>
<td>52</td>
</tr>
</tbody>
</table>
4.8 (a) Test image (size: 256 x 256). (b) Blobs of the test image ($\sigma=0.55$) ........54
4.9 Dominant orientation of the each block (in degree) of the test image ..........54
4.10 Computation of the average local blob size ..............................................55
4.11 The framework of the proposed adaptive scheme for texture analysis by using 2D Gabor filters .................................................................61

5.1 Marchetsreut image .....................................................................................65
5.2 OSU Campus image #193 ...........................................................................66
5.3 OSU Campus image #195 ...........................................................................67
5.4 GPSVan image .............................................................................................68
5.5 Tire image .....................................................................................................69
5.6 SAR image ...................................................................................................70
5.7 Halftone image .............................................................................................71
5.8 Multi-spectral image .....................................................................................72
5.9 An example of the progressive segmentation for the Marchetsreut image .................................................................................................74
5.10 Blobs of the Marchetsreut image ...............................................................77
5.11 Blobs of the OSU Campus image #193 .....................................................78
5.12 Blobs of the OSU Campus image #195 .....................................................79
5.13 Blobs of the GPSVan image .......................................................................80
5.14 Blobs of the tire image ................................................................................81
5.15 Blobs of the SAR image .............................................................................82
5.16 Blobs of the halftone image ........................................................................83
5.17 Blobs of the IR band of the multi-spectral image .......................................84
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.18</td>
<td>Blobs of the green band of the multi-spectral image</td>
</tr>
<tr>
<td>5.19</td>
<td>Blobs of the blue band of the multi-spectral image</td>
</tr>
<tr>
<td>5.20</td>
<td>The Marchetsreut image with mirror padding</td>
</tr>
<tr>
<td>5.21</td>
<td>Segmentation with a pair of the Gabor filter for the Marchetsreut image</td>
</tr>
<tr>
<td>5.22</td>
<td>Segmentation of the Marchetsreut image with starting block size of 128x128 pixels</td>
</tr>
<tr>
<td>5.23</td>
<td>Segmentation of the Marchetsreut image with starting block size of 64x64 pixels</td>
</tr>
<tr>
<td>5.24</td>
<td>Segmentation of the Marchetsreut image with starting block size of 32x32 pixels</td>
</tr>
<tr>
<td>5.25</td>
<td>Segmentation of the OSU campus image #193 with starting block size of 128x128 pixels</td>
</tr>
<tr>
<td>5.26</td>
<td>Segmentation of the OSU campus image #193 with starting block size of 64x64 pixels</td>
</tr>
<tr>
<td>5.27</td>
<td>Segmentation of the OSU campus image #193 with starting block size of 32x32 pixels</td>
</tr>
<tr>
<td>5.28</td>
<td>Segmentation of the OSU campus image #195 with starting block size of 128x128 pixels</td>
</tr>
<tr>
<td>5.29</td>
<td>Segmentation of the OSU campus image #195 with starting block size of 64x64 pixels</td>
</tr>
<tr>
<td>5.30</td>
<td>Segmentation of the OSU campus image #195 with starting block size of 32x32 pixels</td>
</tr>
<tr>
<td>5.31</td>
<td>Segmentation of the GPSVan image with fixed $\sigma=6$</td>
</tr>
<tr>
<td>5.32</td>
<td>Segmentation of the GPSVan image with fixed $\sigma=3$</td>
</tr>
<tr>
<td>5.33</td>
<td>Segmentation of the tire image with fixed $\sigma=12$</td>
</tr>
<tr>
<td>5.34</td>
<td>Segmentation of the tire image with dynamic $\sigma$</td>
</tr>
</tbody>
</table>
5.35 Segmentation of the SAR image ................................................................. 98
5.36 Segmentation of the halftone image ............................................................ 99
5.37 Power spectrum of the halftone pattern of the scanned image .................. 100
5.38 Scanned image after removing halftone noise ............................................. 101
5.39 Halftone pattern of the scanned image ....................................................... 101
5.40 Segmentation of the scanned image after removing halftone noise .......... 102
5.41 Segmentation of the IR band of the multi-spectral image ........................... 103
5.42 Segmentation of the green band of the multi-spectral image ...................... 104
5.43 Segmentation of the blue band of the multi-spectral image ....................... 104
5.44 Edges of the Marchetsreut image .............................................................. 106
5.45 Texture boundaries of the Marchetsreut image .......................................... 106
5.46 Edges of the OSU campus image #193 .................................................... 107
5.47 Texture boundaries of the OSU campus image #193 ................................. 107
5.48 Edges of the OSU campus image #195 ..................................................... 108
5.49 Texture boundaries of the OSU campus image #195 ................................. 108
5.50 Edges of the GPSVan image ..................................................................... 109
5.51 Texture boundaries of the GPSVan image ................................................. 109
5.52 Edges of the tire image ............................................................................. 110
5.53 Texture boundaries of the tire image ......................................................... 110
5.54 Edges of the SAR image ........................................................................... 111
5.55 Texture boundaries of the SAR image ....................................................... 111
5.56 Edges of the halftone image ..................................................................... 112
5.57 Texture boundaries of the halftone image ......................................................... 112
5.58 Edges of the IR band of the multi-spectral image .............................................. 113
5.59 Texture boundaries of the IR band of the multi-spectral image ....................... 113
5.60 Edges of the green band of the multi-spectral image ........................................ 114
5.61 Texture boundaries of the green band of the multi-spectral image .................. 114
5.62 Edges of the IR blue of the multi-spectral image ............................................. 115
5.63 Texture boundaries of the blue band of the multi-spectral image .................... 115
5.64 Edges of the Marchetsreut image with \( \sigma = 1 \) .................................................. 117
5.65 Texture boundaries coincide with intensity edges with \( \sigma = 5 \) ....................... 118
5.66 Texture boundaries coincide with intensity edges with \( \sigma = 1 \) ....................... 118
CHAPTER 1

INTRODUCTION

The ultimate goal of digital photogrammetry is to automatically generate photogrammetric products, such as digital elevation models (DEMs), orthophotos, and maps, from aerial and satellite imagery. In this endeavor, the reconstruction of surfaces plays an important role. For one, the surface of the object space, represented in the form of a DEM, is a photogrammetric product used in many engineering applications. The DEM is also necessary for generating orthophotos --- another product that is increasingly being used. Moreover, surfaces are an important step toward object recognition and image understanding. Surface reconstruction from imagery is an ill-posed problem.

It is known that the human visual system, unsurpassed in its ability to reconstruct surfaces, employs different visual cues to solve this difficult task. Such cues include stereopsis, texture, color, motion, shading. The prevailing method in digital photogrammetry is stereopsis. In order to solve the ill-posed reconstruction problem, different cues should be integrated (Lee and Schenk, 1992). It is known that the combination of stereo with texture yields a more robust solution.
Even though texture is one of the most informative visual cues in shape and depth perception, object recognition, and scene analysis, there is no universally accepted formal definition of texture. Several authors have attempted to define texture qualitatively. The following descriptions provide a sample on how different authors define texture.

- **The visual or tactile surface characteristics and appearance of something.**
  The overall structure of something incorporating all or most of its part (Grove, 1976).

- **The degree of roughness or smoothness, coarseness or fineness of a surface as felt by touch** (Longman Group, 1995).

- **The repetitive or random deviation from the nominal surface that forms the 3D topography** (ANSI).

- **Complex visual patterns composed of entities, or subpatterns, that have characteristic brightnesses, colors, slopes, sizes, etc.** (Rosenfeld, 1982).

- **Detailed structure in an image that is too fine to be resolved, yet coarse enough to produce a noticeable fluctuation in the gray levels of neighboring cells** (Horn, 1986).

- **Tonal primitive properties as well as spatial relationships between them** (Haralick, 1992).

- **Elementary structural texture elements, or random distribution of image values, single or in groups** (Klette and Zamperoni, 1996).
Despite all the progress in texture analysis research during the last three decades, no standard paradigm has yet emerged. To develop a texture analysis system capable of dealing with all aspects of texture is a very difficult visual information processing task, because of the complex nature of texture. It is the lack of mathematical models that make automatic description and recognition of texture patterns a complex and, as yet, unsolved problem (Greenspan, 1996).

1.1 Role of Texture

According to Marr's (1982) computational vision theory, texture along with stereo, shading, motion, and color provides a visual cue. The visual cues are processed by independent and parallel modules. Texture plays an important role in determining shape, depth and surface discontinuities.

According to (Rao, 1990), the main tasks in texture analysis are:

- Detecting and defining texture for identification and description of texture patterns.
- Image classification for image segmentation by using texture.
- Shape-from-texture for recovering information about surface orientation, shape and depth.

Texture has been implemented in some image analysis and computer vision systems. In manual image analysis (photo-interpretation), there is general agreement that the basic elements color (or tone for black and white imagery), size, shape,
texture, pattern, and shadow are the most important factors the analyst considers when interpreting images. Within a given range of scales the texture of a group of objects may be distinctive enough to serve as a reliable clue for identifying objects (Estes et al., 1983; Lillesand and Kiefer, 1987). The MIT vision machine system (Poggio et al., 1988) integrates visual cues, notably edges, stereo, motion, texture, and color, for achieving high performance in unstructured environments for the tasks of recognition and navigation. Texture is used in this system for detecting occlusion boundaries and orientation discontinuities as an indication of physical discontinuities. Bülthoff and Mallot (1990) suggest the integration of stereo, shading, texture, highlight, motion, contour, and occlusion, to derive three-dimensional descriptors for the representation of visible surfaces. The role of texture, in their system, is to recover object orientations.

Schenk and Toth (1992) adopted Marr's computer vision paradigm in their concept of digital photogrammetric workstations. The authors argue that the integration of visual cues yields a more explicit description of surfaces. The motivation for this concept is to increase the degree of automation for solving photogrammetric tasks, such as DEM and object recognition. The role of texture in their conceptual system is to provide information about surface orientation, for recovering the shape of objects, and for the segmentation of images.
1.2 Characteristics of Texture

Texture is qualitatively and quantitatively described by its coarseness, contrast, density, orientation (i.e. directionality), frequency, repetitiveness of spatial patterns, regularity, etc. These elements are considered as texture parameters to be estimated in texture analysis. Figure 1.1 illustrates some of the important texture parameters. One of the characteristics of texture is that it has both stochastic and deterministic properties (Tamura et al., 1978; Gool et al., 1985). Processing texture information belongs to early vision. This does not mean that the development of computational texture analysis is an easy task, however.

Figure 1.2 demonstrates that texture is analyzed relative to its surrounding. Texture primitives themselves are rotation invariant. Humans segregate the same texture patterns by their orientations (see Figure 1.2 (a)). Differently perceived texture patterns as in Figure 1.2 (a) are grouped as one texture by adjacent texture (see Figure 1.2 (b)). Figure 1.2 (c) shows the scale dependency of texture primitives. Scale plays a significant role that must be considered in texture analysis, because the same texture at different scales may be perceived as different textures. The image scale causes the dual nature of texture, namely, micro- and macro-texture. However, there are no clear criteria to differentiate between micro- and macro-texture primitives, rather it is related to psychological effect and also depends on the scale and resolution of the image. The scale dependency problem characterizes texture as a hierarchical process, i.e., texture corresponds to different resolutions.
Figure 1.1 Parameters of texture.
(Adapted from Laws (1980) with permission of the author.)

Figure 1.2 Properties of texture.
Global features characterize the whole texture rather than texture elements. Therefore, a multi-resolution approach is proposed (see, e.g., Jain and Farrokhnia, 1991; Lee and Schenk, 1992).

1.3 Background and Motivation

Studies in physiology and the psychophysics of texture analysis indicate that the human visual system analyzes images by decomposing them into frequency and orientation components. Multi-resolution processing, localization in space, and orientation selectivity can be achieved by using two-dimensional Gabor filters.

The biological motivation for Gabor functions lies in their goodness of fit to receptive field profiles of simple cells in the striate cortex. The specialization of cells in the visual cortex increases. For example, there are orientation sensitive cells and end-stopped cells on cortical receptive fields. The orientation sensitive cells respond to a specific angle, i.e., orientation selectivity of the cells. End-stopped cells are sensitive to the length of the stimulus (Hubel, 1988). Multi-resolution on the retinal-ganglion-cell receptive fields and functions of specialized cells correspond to the parameters of two-dimensional Gabor filters: resolution (or scale), orientation selectivity, and spatial frequency tuning. In this study, the multi-channel responses for texture analysis are processed adaptively by using Gabor filters.
One of the advantages of using Gabor filters is that they achieve optimal resolution in both space and spatial frequency. Greenspan (1996) states an open question of using Gabor filters:

"The decision regarding the appropriate number of frequencies and orientations required for the representation of the input domain."

The main concern in texture analysis is to solve the scale dependency problem and to determine appropriate operator sizes. An operator of constant size provides fairly limited information. Laws (1980) suggests using multiple or adaptive window sizes. Schenk (1995) addresses the problem of using fixed operator sizes in another application, mobile mapping system (MMS). He suggests using varying operator sizes based on image scale and resolution. Krupnik (1994) suggests using a dynamic window size for multiple image matching in order to improve matching accuracy.

1.4 Scope and Objective of the Research

This study is concerned with developing a robust texture segmentation system. For such a system, it is important to utilize properties of texture and to understand how the human visual system works for texture discrimination and grouping. The proposed approach is based on two-dimensional Gabor filters for extracting texture information, followed by segmenting the texture image by unsupervised clustering.

One of the main issues in multi-channel processing is the selection of appropriate filter parameters. This is particularly true for two-dimensional Gabor
filters. Using an appropriate subset of filters with optimal parameters not only reduces the computational cost but also extracts more meaningful texture information. The appropriate parameters are selected based on the statistical characteristics of the images. The proposed scheme reduces the amount of human intervention (i.e., visual inspection) in selecting appropriate parameters. The proposed multi-resolution technique employs adaptive filter sizes to solve the scale dependency problem. Proper filter sizes are determined for different regions in an image.

Although texture has been recognized as a valuable characteristic of image analysis, the complexity involved in its quantification makes texture segmentation a difficult process. In this study, an unsupervised classification is implemented for texture segmentation. Unsupervised classification offers the advantage that less a priori information is required. Texture boundaries after segmentation are detected, analyzed and compared with the physical edges in the images.

In digital photogrammetry, one of the main tasks is matching to achieve automatic orientations (Cho, 1995; Greenfeld, 1987; Heipke, 1992; Krupnik, 1994; Zilberstein, 1991). Most common entities of the input for matching are points or edges. Many matching methods suffer from problematic texture areas with homogeneous or repetitive texture patterns. Most of the methods fail to find the best match in these areas. Texture processing would provide information about areas, so as to avoid the problematic areas.
Texture information can be utilized as a constraint for matching, i.e., image matching with constraints from texture pattern. Texture boundaries may be used as matching entity which serves as initial or global matching stage, then other features such points or edges can refine the matching. Consequently, utilizing texture information provides fast processing and reliable matching results.

Another aspect is that the criteria suggested in this study for determining appropriate parameters may have potential to develop a symbolic description of texture patterns. Finally, results from texture segmentation can be an input to the high level processing such as object recognition, feature extraction and image understanding.

1.5 Organization of the Report

Chapter 2 reviews the status of texture analysis methods. The methods are categorized by approach based on the property of texture --- statistical approach, structural approach, and so on. The review presents advantages and disadvantages of some of the methods.

Chapter 3 describes Gabor functions including derivation and properties of two-dimensional Gabor functions. The relationship between spatial and spatial-frequency domains is presented. Specifically, the parameters in Gabor filters are described.

Chapter 4 describes a proposed strategy for adaptive approach and theory of human texture vision. A scheme for selection of appropriate parameters is presented.

Chapter 5 presents the implementation of the proposed scheme and
experimental results. Various kinds of imagery were used including aerial images, an image taken from the GPSVan, a synthetic aperture radar (SAR) image, a halftoned image, and a multi-spectral image.

Chapter 6 provides conclusions and discussion including performance of the proposed method. The chapter concludes with future recommendations.
CHAPTER 2

STATUS OF TEXTURE ANALYSIS METHODS

According to Laws (1980),

"... the eye must use the same feature extraction methods on every texture field, regardless of source. We do not know what these methods are, although there is indirect evidence that edge detection is involved. We do not know that any retinal transform must retain enough information to distinguish different textures and suppress or ignore information distinguishing equivalent textures. If computers could achieve the same processing results as humans, it would not matter how low-level data reduction was accomplished. It is unlikely, however, that we can ever simulate the activity of the human cortex without first learning the type of data it uses as input ..."

Many texture analysis methods attempt to mimic the human visual system. But because it is not fully understood how the human visual system analyzes texture, most methods have an ad hoc character. Texture analysis methods can be categorized into statistical and structural approaches (Ballard and Brown, 1982; Gool, 1985; Jain, 1989). The majority of texture research focuses on segmenting regions rather than on characterizing textures. Some combination of texture operations lead to reasonable segmentation with a wide range of textures (Parker, 1997).
The statistical approach generates parameters to characterize the stochastic properties of the spatial distribution of gray levels in an image. Here, texture properties are characterized by statistics derived from gray value distribution or from the local feature distributions of texture patterns. Statistical texture analyses are suitable when texture primitives have a size of a few pixels. They may also be effective in cases of large texture primitives if the boundaries of the primitives are highly convoluted or the interior areas are not completely homogeneous in intensity.

The structural approach analyzes visual scenes in terms of the organization and relationship among its substructures. Structural texture analysis is suitable if complete descriptions of individual texture primitives are derivable from the image, i.e., the texture primitives comprise a relatively large number of pixels and the boundaries of the primitives are consistently discernible (Haralick, 1986; Roan et al., 1987).

2.1 Statistical Approach

Textures that are random in nature are well suited for statistical characterization. The main issue in statistical analysis of texture is to estimate texture parameters which describe statistical characteristics of texture patterns.

2.1.1 Co-occurrence Matrices

The co-occurrence matrix method, introduced by Haralick (1973), one of the classic texture analysis methods. The co-occurrence matrix (or spatial gray level dependent matrix; SGLDM) is based on the joint conditional probability density
function of gray values, \( P(i, j, d, \theta) \) with which two neighboring pixels are separated by a distance \( d \). One pixel has gray value \( i \) and its neighbor has gray value \( j \). The neighborly relationship is expressed by angle \( \theta \). The conditional probability can be used for extracting textural features.

In common practice, the angles are quantized to \( \pi/4 \) intervals; \( 0^\circ, 45^\circ, 90^\circ, \) and \( 135^\circ \) (Dyer et al., 1980; Haralick and Shapiro, 1992). Since the information is not usually retained at both \( \theta \) and \( \theta+\pi \), practically \( P(i, j, d, \theta) = 1/2 \left[ P(i, j, d, \theta) + P(i, j, d, \theta+\pi) \right] \) is used (Ballard and Brown, 1982). A number of texture features, which are used to measure statistics of the texture, can be defined. They include: uniformity of energy, entropy, maximum probability, homogeneity, contrast, and correlation.

The chief advantage of co-occurrence matrices is that they are independent of changes in the histogram of an image. Only changes in the topological relationship of gray values (Haralick, 1986), has an effect. Even though this model is superior over other texture analysis methods, it also has some disadvantages. Individual elements of a co-occurrence matrix do not make good features. Moreover, they are highly dependent on the chosen resolution (Reed and Wechsler, 1990). Some of the statistical features are correlated to each other, and they do not provide psychological correlates (Ballard and Brown, 1982).
2.1.2 Generalized Co-occurrence Matrices

Co-occurrence matrices are generalized for describing spatial distributions of local features, including lines and edges. Instead of analyzing the spatial distribution of gray values (Davis et al., 1981), texture primitives and their spatial interactions are considered in generalized co-occurrence. The primitives are characterized by size, elongation, and angular orientation.

One of the disadvantages of generalized co-occurrence matrices lies in its computational complexity. Other problems arise from the difficulty to locate texture primitives and to measure their attributes. Further, the identification of nearest neighbors of texture elements is problematic (Laws, 1980).

2.1.3 Autocorrelation Function

Texture primitives have statistical and spatial characteristics. Especially, the spatial size of the primitives may describe coarseness and periodicity of the texture. The drop-off distance of the autocorrelation function can be used to measure texture coarseness because the autocorrelation function will drop off slowly with larger texture primitives and vice versa (Haralick and Shapiro, 1992).

The autocorrelation function is not an appropriate texture descriptor since most natural textures have similar autocorrelation functions which are difficult to separate.
2.1.4 Texture Energy Measure

Laws (1980) developed a widely-used method of texture energy measures. The original image is first filtered with a set of convolution masks. The filtered images contain micro-texture features. They are then processed with a nonlinear local texture energy filter (moving average of the absolute value of the filtered images) to obtain macro-texture features. The second step is termed a texture energy transform.

The energy is originally defined as the sum of squares, however the sum of absolute values is preferable because it is computationally cheaper without loss of performance. Figure 2.1 shows Laws’ 1 × 5 center-weighted vector masks which form a variety of convolution masks, size 5 × 5. The convolution masks are generated from the one-dimensional masks by outer products. According to Laws’ experiments, the four most important masks are E5L5, R5R5, E5S5, and L5S5 which are shown in Figure 2.2.

Gong and Huang (1988) proposed to combine texture energy with hidden Markov models to segment textured regions. Hsiao and Sawchuk (1989) improved Laws’ method by applying an edge preserving noise smoothing operator. The moving window operators introduce considerable errors inside regions and along the borders between textures. The operator is a locally linear minimum mean square error estimator (LLMMSE) which smoothes out noise in homogeneous regions and leaves the observation unchanged in the vicinity of the edges. Lee and Schenk (1992) present a scale space approach, which was obtained by Gaussian image pyramid, with texture

Since the filter sizes for texture energy measures are relatively small, only the high frequency texture elements can be analyzed (Reed and du Buf, 1993). Laws' method suffers from determining appropriate operator sizes. The optimal window sizes for micro-texture masks and macro-texture operators depend on the characteristics of texture. Laws suggests multiple or adaptive window sizes, even though he uses fixed sizes of $3 \times 3$ or $5 \times 5$ for micro-texture masks and $15 \times 15$ or $31 \times 31$ for macro-texture operators.

2.1.5 Markov Random Field Models

Markov random fields (MRFs) have the potential for image analysis, texture classification, image restoration, and image compression. The MRF models for texture assume that the texture fields are stochastic, periodic, stationary and conditional independent (Germain and Germain, 1984; Haralick and Shapiro, 1992; Ohanian and Dubes, 1992).

Parameters of the MRF models are used to recognize texture. Cross and Jain (1983) show that MRF parameters control strength, direction, coarseness, gray level distribution, and sharpness of the clustering in the image. The authors focus on the application of the binomial texture model with selection of an appropriate conditional distribution. The results confirm that the MRF model fits micro-textures well.
Figure 2.1 Laws' one-dimensional texture energy masks.

\[
L5 = [1, 4, 6, 4, 1] \\
E5 = [-1, -2, 0, 2, 1] \\
S5 = [-1, 0, 2, 0, -1] \\
W5 = [-1, 2, 0, -2, 1] \\
R5 = [1, -4, 6, -4, 1]
\]

Figure 2.2 The four most important 5 x 5 convolution masks.

\[
E5L5 = \begin{bmatrix}
-1 & -4 & -6 & -4 & -1 \\
-2 & -8 & -12 & -8 & -2 \\
0 & 0 & 0 & 0 & 0 \\
2 & 8 & 12 & 8 & 2 \\
1 & 4 & 6 & 4 & 1
\end{bmatrix} \\
R5R5 = \begin{bmatrix}
1 & -4 & 6 & -4 & 1 \\
-4 & 16 & -24 & 16 & -4 \\
-4 & 16 & -24 & 16 & -4 \\
1 & -4 & 6 & -4 & 1
\end{bmatrix}
\]

\[
E5S5 = \begin{bmatrix}
-1 & 0 & 2 & 0 & -1 \\
-2 & 0 & 4 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 \\
2 & 0 & -4 & 0 & 2 \\
1 & 0 & -2 & 0 & 1
\end{bmatrix} \\
L5S5 = \begin{bmatrix}
-1 & 0 & 2 & 0 & -1 \\
-4 & 0 & 8 & 0 & -4 \\
-6 & 0 & 12 & 0 & -6 \\
-4 & 0 & 8 & 0 & -4 \\
-1 & 0 & 2 & 0 & 1
\end{bmatrix}
\]
The binomial model allows generating natural texture patterns with a smooth peak and valley in the gray value height field. Gimel'farb and Zalesny (1992) describe a probabilistic model representing piecewise-homogeneous textured images as samples of MRF. The parameters of the model are estimated by using a stochastic approximation technique.

MRFs also have been applied to integrate information from different sources of the visual cues in order to achieve unique and stable visual reconstruction. Aloimonos and Shulman (1989) suggest MRFs as a means of regularization of the ill-posed problem in visual reconstruction following a top-down approach.

2.1.6 Simultaneous Autoregressive Models

Texture can be quantified by modeling spatial interaction among pixels using a class of two-dimensional, noncausal, stochastic random field models called simultaneous autoregressive (SAR), or spatial autoregressive models. The parameters of SAR models can be estimated by least squares or maximum likelihood methods. The important issue is to select appropriate window sizes. Khotanzad and Chen (1989) propose an algorithm for automatic selection of appropriate window size by textural change measure.

2.1.7 Fractal Models

It is difficult to define texture patterns by traditional Euclidean geometry. Roughness, regularity, and multi-scale of the texture can be characterized by fractal dimension. Fractal dimension of the image intensity surface should describe the surface
of the objects because natural phenomena often produce fractal objects (Ohanian and Dubes, 1992; Russ, 1992). Fractal models have been used to synthesize textures and computer generated images using fractal models that look very real.

Fractal models of surfaces have been used in texture segmentation. The problem is that it describes the low frequency behavior well but it cannot describe the high frequency aspects (Kashyap, 1986). In order to segment textures, there should be enough difference in roughness (Medioni and Yasumoto, 1984).

2.1.8 Spatial-Frequency Approach

Spatial-frequency approaches are based on image representations that indicate the frequency content in localized regions in the spatial domain. According to Reed and du Buf, 1993), spatial-frequency methods are able to achieve high resolution in both the spatial and spatial-frequency domains, consistent with recent theories of human vision.

Gabor filters, generated by two-dimensional Gabor functions, are described as multi-scale, orientation selective, and multi-channel spatial-frequency filters. The use of Gabor filters is one of the recent and efficient methods for texture analysis (Bovik and et al., 1990; Clark and Bovik, 1989; Fogel and Sagi, 1989; Gopal and et al., 1990; Jain and Farrokhnia, 1991; Jain and Raru, 1996; Kehrer and Meinecke, 1996). Chapter 3 provides a detailed description of Gabor functions.

Bergen and Landy (1991) developed another kind of orientation selective filters for visual texture segregation. However, the filter size is limited to 5 x 5. The filters
approximate the second directional derivatives of the corresponding resolution level in the horizontal (0°), vertical (90°), right (45°) and left (135°) diagonal directions. Another application of Gabor functions is object recognition (see Buhmann et al. (1992). In their study, Gabor functions are used for wavelet transform for image coding.

2.2 Structural Approach

Textures with more regular structure or macro-texture can be analyzed better with structural methods. Here, the relationship and organization of the substructures of textures are analyzed. In order to describe the texture with structural models (so-called deterministic texture), the texture primitives must be arranged in nearly regular repetitive spatial pattern (Gool et al., 1983; Haralick, 1986). Description of the primitives, choice of the primitive and placement rules are essential issues in structural analysis of texture.

2.2.1 Placement Rule

In structural methods, texture elements are extracted and their spatial relationships are analyzed. On the structural level, a texture pattern is defined by elements which occur repeatedly according to placement rules, \( t = R(e) \) where \( t \) denotes texture pattern of macro-scopic view, macro-texture; \( R \) denotes a relation, a placement rule; and \( e \) denotes an texture element of micro-scopic view, micro-texture (Tamura et al., 1978). The placement rules can be described by grammar models, e.g.,
shape grammars, tree grammars, and array grammars (Ballard and Brown, 1982; Banks, 1990; Fu, 1977; Schalkoff, 1992). Lam and H S Ip (1993) demonstrate that Voronoi polygon tessellation for segmenting structural texture is useful.

Structural texture description should be sufficiently flexible so that a class of equivalent textures can be generated using similar primitives in similar geometric relationships. According to Rao (1990), this type of texture is referred to as "strongly ordered textures," (for example, a brick wall), and can be described by placement rules. However, some textures are ambiguous and require more than one texture primitive.

In structural methods, bottom-up and top-down approaches are employed. In bottom-up cases, texture primitives are extracted first followed by investigating the spatial arrangement of the extracted elements. In this approach, the segmentation process often becomes sensitive to various image degradations. In the top-down case, the spatial structure of texture is recognized before the primitive is extracted. Also, combined methods of bottom-up and top-down approaches can be used (Gool et al., 1985).

The placement rule is inappropriate for most natural imagery because it is difficult to identify texture primitives and to parse structure of the texture. Another major problem in implementing the placement rule is that such a method requires a priori knowledge for a preliminary texture segmentation (Laws, 1980).
2.2.2 Symbolic Description

Rao (1990) suggests a descriptive vocabulary for structural textures, i.e., symbolic representation of the texture as a visual language. The key issue is to provide symbolic descriptions for both identified primitive elements and their placements. For example, the primitives are described with their periodicity, e.g., spacing, size, and ratio of size to period. The placements can be borrowed from other disciplines such as petrography, frieze groups, and wallpaper groups.

2.3 Statistical-Structural Approach

A method that combines statistical and structural approaches is the so-called random mosaic model. The first step involves regular or random tessellation of a plane into cells, followed by assigning property values to each cell (Haralick and Shapiro, 1992). These models represent random geometric processes.

A random mosaic model could define rules for partitioning a plane into different cells, where each cell contains a geometrical figure whose features have prescribed probability distributions. Although random mosaic models and MRFs (see Section 2.1.5) are both model-based approach, in general, random mosaic models would provide higher resolution than MRFs (Jain, 1989).
2.4 Neural Network Approach

The neural network is often applied for two major purposes in texture analysis: training texture extraction operators (e.g., Jain and Karu, 1996) and texture classification (e.g., Chellappa et al., 1992). The advantage with the neural network is automatic feature selection even though there is no guarantee that the automatically selected features are working perfectly. According to Visa (1990) there is a tendency that the selected features are too well adapted to the learning images.

2.5 Summary

Texture is an important visual cue for analyzing natural scene and images. Despite its importance, there is no completely adequate formal paradigm or method available yet. Many existing approaches recognize and analyze texture with ad hoc techniques, or provide only partial solutions. The main reason is the complicated nature of texture.

As the name indicates, MRF models, SAR models, and fractal models are categorized as model-based approaches. Model-based methods hypothesize underlying processes for textures and segment based on certain parameters obtained by these processes. In recent years, there has been a great deal of activity in the refinement of existing approaches and in the development of completely new techniques (Reed and du Buf, 1993). Methods such as co-occurrence matrices, texture energy measures, and use of two-dimensional Gabor functions are popular. Structural approaches have
limited applications since many textures in nature have no well-defined texture primitives and placement rules.

A promising approach that resembles human texture processing is based on two-dimensional Gabor functions. However, there are some important issues which require special attention for implementing Gabor functions with automatic processing. Chapters 4 and 5 address these issues.
CHAPTER 3

TWO-DIMENSIONAL GABOR FUNCTIONS
AS MULTI-CHANNEL FILTERS

Texture can be characterized by orientation, frequency, and multi-scale properties of texture primitives. Therefore, a good texture analysis operator must have these properties: orientation selectivity, good localization both in the spatial and frequency domains, and multi-scale capability. Such is the case for Gabor functions. They can be used to cluster similar texture primitives (Shao and Förstner, 1994). Applied as filters, two-dimensional Gabor functions perform remarkably similar to the human visual texture analysis. Psychophysiological experiments with Gabor filters for texture analysis and segmentation show that Gabor filters are comparable to the performance of the human visual system (Fogel and Sagi, 1989; Jain and Farrokhnia, 1991; Turner, 1986). That is, Gabor functions can be conceived as hypothetical structures of neural receptive fields in the visual cortex (Cougman, 1985; Greenspan, 1996; Kehrer and Meinecke, 1995).

This is probably due to the fact that the positive parts and the negative parts of the Gabor filter correspond to the excitatory and inhibitory regions of the receptive
field, and exhibit global neighborhoods (Kehrer and Meinecke, 1995; Julesz, 1990). The visual pathway in the human visual system includes a set of channels that are orientation selective and tuned to spatial frequency. Also it is known that different sizes of the receptive fields exist on the retinal ganglion cells (Marr, 1982). The different sizes of the receptive fields control the resolution channel in the visual pathway. Visual input is processed independently by multiple channels (i.e., resolution, orientation, and spatial frequency) (Bruce and Green, 1992). Two-dimensional Gabor functions having resolution, orientation, and spatial frequency selectivity are called multi-channel filters.

3.1 Derivation of Two-Dimensional Gabor Functions

Gabor functions were introduced by Gabor (1946) and extended to two-dimension by Daugman (1985), called “two-dimensional visual cortical filters.” A Gabor function is defined by a harmonic oscillator which is a complex sinusoidal plane wave (or carrier) of some frequency and orientation within a Gaussian envelope, i.e., a sine and a cosine function modulated by a Gaussian function. The general form of two-dimensional Gabor functions is given by

$$ h(x, y) = g(x, y) \cdot \exp[2\pi i(Ux + Vy)] $$  \hspace{1cm} (3.1)

where \( i = \sqrt{-1} \), \( U \) and \( V \) are derived in equations (3.7) and (3.8), respectively. \( g(x, y) \) is two-dimensional Gaussian function given by
\[ g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp \left[ -\left( \frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} \right) \right] \]  

(3.2)

where \( \sigma_x \) and \( \sigma_y \) are standard deviations with respect to the x-axis and y-axis. In most applications, it is assumed that \( \sigma_x = \sigma_y = \sigma \), i.e., the unity aspect ratio of the Gaussian function (Dunn and Higgins, 1995). From equations (3.1) and (3.2), the following equation is obtained

\[ h(x, y) = \frac{1}{2\pi \sigma^2} \exp \left[ -\left( \frac{x^2 + y^2}{2\sigma^2} \right) \right] \cdot \exp[2\pi i(Ux + Vy)] \]  

(3.3)

By applying the Euler identity, equation (3.3) can be rewritten as

\[ h(x, y) = \frac{1}{2\pi \sigma^2} \exp \left[ -\left( \frac{x^2 + y^2}{2\sigma^2} \right) \right] \cdot \{ \cos[2\pi(Ux + Vy)] + i \sin[2\pi(Ux + Vy)] \} \]  

(3.4)

The complex exponential is a two-dimensional complex sinusoid at radial (or angular) frequency,

\[ F = \sqrt{U^2 + V^2} \]  

(3.5)

and orientation,

\[ \theta = \tan^{-1} \left( \frac{V}{U} \right) \]  

(3.6)

(Bovik et al., 1990; Dunn and Higgins, 1995).

The following relationship can be easily derived from equations (3.5) and (3.6),

\[ U = F \cos \theta \]  

(3.7)

\[ V = F \sin \theta \]  

(3.8)
Finally, the following two-dimensional Gabor functions, are obtained by substituting equations (3.7) and (3.8) into equation (3.4),

\[
\begin{align*}
    h(x, y) & = \frac{1}{2\pi\sigma^2} \exp \left[ -\left( \frac{x^2 + y^2}{2\sigma^2} \right) \right] \\
    & \{ \cos[2\pi(F\cos\theta x + F\sin\theta y)] + \sin[2\pi(F\cos\theta x + F\sin\theta y)] \}
\end{align*}
\]  

(3.9)

Two-dimensional Gabor functions also can be expressed with angular frequency

\[
\omega = 2\pi F
\]  

(3.10)

\[
G(x, y; \sigma, \omega, \theta, \phi) = \exp \left[ -\frac{(x - x_o)^2 + (y - y_o)^2}{2\sigma^2} \right] \cdot \sin\{\omega(x \cos \theta - y \sin \theta) + \phi\}
\]  

(3.11)

where \(x_o\) and \(y_o\) specify the center of the Gaussian function, and \(\phi\) is the phase shift (Turner, 1986; Fogel and Sagi, 1989). The radial frequency \(F\) with units of cycles/image, and the period \(T\) are related by:

\[
F = \frac{I}{T}
\]  

(3.12)

where \(T\) represents the wavelength in units of pixels. Therefore, the angular frequency with radian unit is computed by

\[
\omega = \frac{2\pi}{T}
\]  

(3.13)

The two-dimensional Gabor functions have real and imaginary components. The \textit{cosine\ } part and the \textit{sine\ } part in equation (3.9) represent the real component and imaginary component, respectively. The phase shift in equation (3.11) generates real components for \(\phi = 90^\circ\) and imaginary components for \(\phi = 0^\circ\). Figure 3.1 shows a perspective view of real and imaginary components of two-dimensional Gabor
functions. Figure 3.2 is a bird’s eye view of a family of Gabor filter bank in the spatial domain, consisting of various \( \varphi \) elements.

3.2 Properties of Gabor Filters

Gabor filters are tuned to spatial frequencies at different scales and orientations and perform a local Fourier analysis. Gabor filters are sometimes called Gabor wavelets because of the analogy to the generation of a wavelet basis from a single mother wavelet. The whole family of Gabor filters (i.e., Gabor filter bank) is generated by rotation and dilation of a single Gabor filter. Equation (3.14) is a general form of the dilation equation to create wavelets.

\[
W_i(x) = s \cdot W_{i-1}(ax + b)
\]  

(3.14)

where \( W(x) \) is a given function, \( s \) is the scale, \( a \) the dilation factor, and \( b \) is a shift (Glassner, 1995). The coefficients \( s, a, \) and \( b \) in equation (3.14) correspond to \( \exp\left[-\frac{(x-x_o)^2 + (y-y_o)^2}{2\sigma^2}\right] \), \( \omega \), and \( \varphi \) in equation (3.11), respectively.

The family of filters are self-similar at different levels of scale (Buhmann and et al., 1992; Hofmann et al., 1996). However, Gabor functions are near orthogonal. Therefore, there exists a trade-off between redundancy and completeness in designing Gabor filter banks. Two-dimensional Gabor filters have optimal joint resolution in that they minimize the product of effective areas occupied in spatial and frequency domains (Daugman, 1985; Turner, 1986). Gabor filters are linear, shift invariant, rotational
invariant with respect to the tuning frequency, and scale invariant. These properties, together with the orientation selectivity, multi-resolution process, and band-pass filters, make Gabor filters unique and robust for texture extraction and analysis.

The spatial frequency response, or modulation transfer function (MTF) of Gabor functions (equation (3.1)) is obtained by the two-dimensional Fourier transform:

\[ H(u, v) = \exp\left\{-2\pi \sigma^2 \left[ (u - U)^2 + (v - V)^2 \right]\right\} \]  

(3.15)

where \( u \) and \( v \) denote frequencies. The MTF specifies the amount by which the filter modulates each frequency component of the input image. Figure 3.3 describes the spatial frequency responses of the Gabor function shown in Figure 3.1. Figure 3.4 schematically illustrates the Gabor filter bank in the spatial-frequency domain. Some of the Gabor filters are also shown in the spatial domain.
Figure 3.1 (a) Real (even-symmetric), and (b) imaginary (odd-symmetric) components of a 2D Gabor function with 16 pixel period and $45^\circ$ orientation.

The filter coefficients are converted into gray values. Bright regions represent positive values and vice versa.

Figure 3.2 An example of 2D Gabor filter bank.
Figure 3.3 Spatial frequency responses of a 2D Gabor function shown in Figure 3.1. (a) Real and (b) imaginary components.

Figure 3.4 (a) Example of 2D Gabor filter bank in the spatial-frequency domain, and (b),(c),(d),(e) 2D Gabor filters in spatial domain.
3.3 Parameters of Gabor Filters

There are four parameters in two-dimensional Gabor filters: the standard deviation of the Gaussian envelope (σ), spatial frequency (ω), orientation (θ), and phase (φ). These parameters characterize certain texture patterns. Therefore, Gabor filters allow extracting pictorial information containing texture primitives.

The selection of appropriate parameters is the most important issue in applying Gabor filters. Since the parameters vary independently, optimal parameters cannot be determined uniquely (Kehrer and Meinecke, 1995). All possible combinations of the parameters result in a large set of filters. In Chapter 4, a strategy for selecting appropriate parameters is discussed.

3.3.1 Standard Deviation of the Gaussian

The standard deviation of the Gaussian envelope controls the spatial extent as well as the band width of the filters. Physiological evidence suggests a tendency toward decreasing spatial extent with increasing frequency preference (Kulikowski and Bishop, 1981; Turner, 1986). In most cases, circularly symmetric Gaussian envelopes (i.e., unity aspect ratio) are used. However, increasing the length of the filters along the axis parallel to the plane wave tends to improve the sensitivity of the filter to such local features (Turner, 1986).

With increasing filter size, larger overlapping neighborhoods at a center pixel result. If the window size is too large, the accuracy of the texture boundaries is
reduced. Therefore, the choice of the standard deviation is a tradeoff between output variation and boundary localization (Dunn and Higgins, 1995).

3.3.2 Spatial Frequency

The number of cycles per period is the spatial frequency. The period is related to the size of the input texture pattern (i.e., texture primitive). The spatial frequency describes the coarseness or roughness of texture. It also controls the resolution that is related to the amount of information. Various resolutions are obtained as different generations of wavelets are created by varying dilation factors.

3.3.3 Orientation

In Gabor filters, orientation varies from 0° to less than 180° due to symmetry (e.g., 0° ≡ 180°, 45° ≡ 225°, 90° ≡ 270°, etc.). Orientation is one of the texture parameters. Proper selection of the orientation can detect the directionality or dominant orientation angle of the texture primitives or pattern. In common practice, the orientation angle is discretized by 45°. This discretization is also found in other image processing techniques, such as co-occurrence matrices and chain code.

According to Voorhees (1987), there is evidence that the human visual system has a special sensitivity for vertical and horizontal orientations, and requires around 20° or more to preattentively perceive orientation differences. Orientation preference is a particularly important feature of simple cells in the visual cortex. Their maximum response occurs where edges are oriented at a particular angle to the visual axis. The preference is quite distinctive; rotating the stimulus by more than 20° from the
preferred direction greatly reduces the cell’s firing rate (Bruce and Green, 1992). Therefore, a certain level of discretization is reasonable, not only for computational simplicity but also from psychological evidence.

3.3.4 Phase

Two-dimensional Gabor functions are composed of two components, real and imaginary components with a 90° phase shift between them. The phase can solve the ambiguity problem. It cannot be decided unequivocally which parameter variation is responsible for the activity of the filter change.

The second component (imaginary, or odd-symmetric function) is introduced by a 90° phase shift to the first function (real, or even-symmetric function). The properties of the second component are that it must have the same preferred frequency and orientation and must cover the same part of the image (Kehrer and Meinecke, 1995).
CHAPTER 4

DESCRIPTION OF THE PROPOSED
ADAPTIVE APPROACH FOR TEXTURE ANALYSIS

This chapter describes the proposed adaptive strategy for extracting texture information with two-dimensional Gabor filters. The motivation for an adaptive approach lies in the fact that we cannot extract meaningful information from an image with one kind of operator only. Texture analysis requires a variety of parameters. This is why an adaptive approach is considered an essential strategy in this research. The strategy of using multiple filter sizes has been proven successful in other applications. A well-known example is smoothing an image. If a fixed size operator is chosen then unacceptable loss of information may occur in some parts, while in the other parts of the image noise was removed well.

The cardinal question is: What is the optimal filter size? As shown in Chapter 3, the parameters of Gabor filters (e.g., size, orientation, and spatial-frequency) depend on the image contents which are characterized by statistical properties. Now, the question is how one can define different regions properly. In this research, a
progressive subdividing scheme is suggested for the initial segmentation to define the regions. This scheme is adopted from the progressive sampling for digital elevation models (DEM) (Makarovic, 1973). A similar strategy is also employed in the split-and-merge algorithm (Gonzalez and Wintz, 1987; Pavlidis, 1982).

According to Dunn and Higgins (1995) "Distinctive discontinuities are detectable only if the Gabor filter parameters are suitably chosen." A scheme for determining appropriate parameters of two-dimensional Gabor filters is proposed in this research. The size of the Gabor filter for each region was determined based on the statistical properties of the image. The theory of human visual processing for texture analysis was applied to the entire framework. Textons (see Section 4.2) as texture primitives are extracted at the preattentive visual stage. Textons in terms of blobs were detected by using the Laplacian of the Gaussian (LoG) operator. The dominant orientation and spatial frequency were computed with blobs for each region which was defined by the progressive subdividing scheme. Each region of the image then was adaptively processed with pairs of two-dimensional Gabor filters (0° and 90° phase) with chosen parameters. The adaptive processing is considered as attentive (or focal attention) visual perception. The process was to transform the differences of the texture patterns into a detectable filter response, i.e., discontinuities at texture boundaries. Segmentation was performed by an unsupervised classification technique.
4.1 Characteristics of an Image

Mean, variance, and entropy are fundamental statistical values to characterize images. The mean measures the overall brightness of an image, and the variance expresses the contrast. Entropy and variance can be used as indicators for representing complexity and randomness of an image. There is correlation between complexity (or randomness) and property of texture. Entropy is a measure of the degree of randomness of the variables and is defined as

\[ H = -\sum_{i=0}^{n} p(i) \log_2 p(i) \]  (4.1)

where \( n \) is number of variables with probabilities, \( p(i) \), and \( i \) indicates a certain gray value. The entropy represents the amount of information required to represent an image (Gonzalez and Wintz, 1987). Since the gray value of most natural images are distributed between 0 and 255, the entropy of homogeneous to totally random image ranges from 0 to 8. The entropy of the natural images used in this study ranges from 4 to 7.

Areas with lower entropy are homogeneous and most likely less textured areas, and vice versa. Figure 4.1 shows the entropy and variance for each block of natural images. Figure 4.2 shows the relationship between entropy and variance of the image shown in Figure 4.1. The contrast measured by the variance provides the amount of information in an image. Therefore, the overall behavior of entropy and variance are similar over each image block. To use entropy may provide an advantage over
variance, since we know the range of entropy. In this study, initial segmentation of the image for an adaptive approach is based on entropy.

4.2 Texton Theory as Human Texture Perception

Julesz and Bergen (1983) introduced the notion of textons. Textons are texture primitives which are basic features such as blobs or line segments (with associated orientation, dimension, color, etc.), their terminators and crossings of blobs. There are two modes in the human visual system: preattentive vision and attentive vision (or scrutiny vision). Texton extraction is a preattentive process and corresponds to the primal sketch in Marr's vision theory (Marr, 1976). Textons form a part of other elements of the primal sketch including intensity edges (i.e., intensity changes) with their geometrical distribution and organization such as orientation, contrast, and dimensions. According to Julesz and Bergen's experiments (1983), there is evidence that preattentive texture discrimination in "the human visual system can instantaneously (160 milliseconds or less) detect differences in a few local conspicuous features (i.e., textons), regardless of where they occur." Preattentive texture discrimination can serve as a model system with which to distinguish the role of local texture element detection from global computation in visual perception (Julesz, 1981).
Overall entropy = 6.8
Overall variance = (1198)

*Numbers in squares are block numbers. Numbers in the parentheses represent variance.*

Figure 4.1 Entropy and variance of each block of the natural image.

*Note: Variances are normalized to entropy range for comparison with the entropy.*

Figure 4.2 Relationship between entropy and variance.
The issue is how to extract textons from images. Voorhees (1987) proposes to use the LoG operator for detecting blobs in images. The LoG is given by

$$\nabla^2 G(x, y) = \frac{1}{2\pi\sigma^4} \left[ \left( \frac{x^2 + y^2}{\sigma^2} \right) - 2 \right] \cdot \exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right)$$

(4.2)

The relationship between the size of the LoG operator and the standard deviation of a Gaussian is

$$w = 3 \cdot (2\sqrt{2}\sigma)$$

(4.3)

where $w$ denotes the size of the LoG operator.

The LoG operator was devised by Marr and Hildreth (1980) to construct the primal sketch. The primal sketch is a symbolic representation of the image with primitives (or tokens) such as zero-crossings, blobs, terminators and discontinuities, edge segments, virtual lines, groups, curvilinear organization, and boundaries. As mentioned earlier, these tokens of the primal sketch correspond to textons, therefore LoG is considered an appropriate operator for detecting blobs as textons. It has been shown that the receptive fields in the retina are the biological equivalents to the LoG operator. The convolution of an image, $f(x,y)$, with an LoG, provides zero-crossings which represent intensity edges of the image. Blobs may be regarded as duals of edges, i.e., positive for dark blobs and negative for light blobs (Voorhees, 1987). After convolving the image with LoG the positive regions were assigned with black to represent blobs.
Selection of the standard deviation ($\sigma$) of the LoG, which controls scale, is an important issue. Voorhees (1987) suggests using different scales over different parts of the image, even though a single scale may be sufficient. However, to use a single scale is reasonable since texton extraction is a preattentive process. The human visual system may not be able to utilize different resolution channels during the extremely short preattentive time (e.g., 160 milliseconds or less).

By and large, the selection of appropriate scales is still an open question and known as the scale dependency problem in texture analysis. Figure 4.3 shows a natural image and detected blobs. The blobs are sparsely distributed over homogeneous (or less textured) areas. Homogeneous areas have low entropy values (see Figure 4.1).

Figure 4.3 (a) Marchesreut image, and (b) its blobs with $\sigma = 0.55$. 

43
Another important issue is how to deal with noise in images. Textures in object space (i.e., real world) may differ from textures appearing in images (termed image texture) due to following reasons:

1. Texture perception is a three-dimensional visual phenomenon; however, in general, two-dimensional analyses of texture are performed on an image.
2. Noises from various sources are introduced during image formation and scanning.
3. Digitization and quantization during scanning results in information loss.
4. Images are scaled models of the real world resulting in texture primitives that are perceived differently under extreme scale change.

In this study, noises in the images are considered as part of textures. Both LoG operators and Gabor filters have the capability of removing a certain level of noise, therefore any further noise removing process is not deemed necessary in this study. However, an experiment of noise removal was discussed and presented in Chapter 5.
4.3 Adaptive Strategy

4.3.1 Progressive Segmentation

The raw image is segmented in a progressive fashion, based on its statistical characteristics. The strategy of progressive segmentation is adopted from progressive sampling of DEMs. First, the image is divided into regular square blocks. Then, entropy is computed for each image block. The advantage of using the entropy was discussed in Section 4.1.

Blocks with an entropy exceeding a predetermined threshold (one reasonable value may be an overall entropy of the image), are further subdivided into four quadrants, and the recursive process is repeated. The subdivision stops at the predetermined smallest region size. It might be reasonable if the subdivided regions are not smaller than the smallest filter size.

Adaptive filtering with Gabor filters is performed by combining the results from progressive segmentation and appropriate parameter selection. Figure 4.4 illustrates the procedure of the progressive segmentation. Figure 4.5 points to a potential problem. Small textured areas located in the center of the block and surrounded by homogeneous quadrants remain undetected, because the entropy of four quadrants are lower than the thresholding. This problem can be solved by choosing a smaller initial block size.
Thicker squares represent regions with higher entropy where subdivisions are performed.

Figure 4.4 Progressive segmentation scheme.

Figure 4.5 Problematic situation.
Obviously, it is reasonable to subdivide regions with higher entropy because those regions are most probably more textured areas. Each segmented region is now processed with Gabor filters whose parameters were chosen by a proposed strategy. Consequently, homogeneous areas are processed with larger size filters (i.e., larger standard deviation of the Gaussian) and larger spatial frequency of the Gabor filters, and vice versa. The advantages of the proposed strategy are to fully imitate human visual perception of texture analysis (i.e. preattentive and attentive visions), and to reduce computational cost.

Khotanzad and Chen (1989) discuss the problem of how to select an appropriate window size. The window size refers to the size of the initially segmented regions. According to them,

"... important problem is how to select an appropriate window size over which local texture characteristics are observed and measured. The window size should be large enough to contain several texture elements for the enclosed region to exhibit similar textural characteristics as that of the underlying region that it masks. At the same time, it should be as small as possible to enable accurate detection of the edges."

The similar problem occurs in area-based image matching (e.g., cross-correlation or least squares matching). The window size should be large enough to avoid ambiguity (i.e., multiple matching). On the other hand, it should be small enough to keep geometric distortions under control (e.g., foreshortening). There is a trade-off between
uniqueness and accuracy in determining window size, and it depends on the characteristics of the image such as scale, resolution and contents. A compromise must be found by computing a uniqueness measure for different template sizes (Schenk, 1996). The scale of the image and the area of the image (e.g., agricultural, built-up, or residential area) may provide some information to decide the initial starting block size. In this study, we chose three different initial sizes of the block, and the results are presented in Chapter 5.

4.3.2 Determination of the Gabor Filter Size

The appropriate filter size depends on the characteristic of the image in terms of its statistical properties, scale, and resolution. The standard deviation (σ) of the Gaussian controls the size of a Gabor filter, i.e., resolution. The Gaussian function is symmetric with respect to zero, and the function value cannot be zero. Equation (4.5) is used to determine the relationship between σ and filter size.

\[ \exp(-x^2/\sigma^2) = 10^{-N} \]  

Therefore, the filter size is determined by

\[ w = (2\sqrt{2N \ln 10})\sigma \]  

It is known that there are several different sizes of the retinal-ganglion-cell receptive fields (i.e., resolution channels) in the human visual system. In this study, we selected \( N = 1.5 \) and utilized four different filter sizes with \( \sigma = 12, 9, 6, \) and 3, corresponding to \( 65 \times 65, 47 \times 47, 31 \times 31, \) and \( 15 \times 15, \) respectively. The purpose of using different standard deviations is to perform multi-resolution texture analysis.
Entropy was chosen to represent homogeneousness. Therefore, a larger $\sigma$ was chosen for the region with small entropy and *vice versa*:

$\text{IF } (H \geq 6.5) \quad \sigma = 3$

$\text{IF } (6.0 \leq H < 6.5) \quad \sigma = 6$

$\text{IF } (5.5 \leq H < 6.0) \quad \sigma = 9$

$\text{IF } (H < 5.5) \quad \sigma = 12$

To select an optimal $\sigma$ value is subjective because it is scale dependent. It is suggested to use 65 x 65 for the largest filter size based on our experiments and agreement with other authors’ experiments (Turner, 1986; Fogel and Sagi, 1989). Figure 4.6 shows the relationship between entropy and blob size. Generally, regions with high entropy have small blob size and vice versa. The summation of the real component of the Gabor filter is not zero while that of the imaginary component is zero. However, the summation of the real component with size of 65 x 65 is small enough which does not change the image tone.
The blob sizes are normalized to entropy range for comparison purpose. 50 image blocks are plotted for the Marchetsreut image (see Chapter 5).

Figure 4.6 Relationship between entropy and blob size.

4.3.3 Computation of Dominant Local Orientation

Rao (1990) developed a scheme for estimating the orientation of a texture field by modifying Kass and Witkin's (1987) algorithm. In this study, Rao's scheme of "inverse arctangent method" is applied to compute the dominant local orientation of the blobs for each window (i.e., image block of initial segmentation). The underlying assumption for this scheme is that the texture has only one dominant local orientation for each window. Figure 4.7 illustrates the method to compute dominant local orientation of a set of blobs in a window. In order to compute the orientation at the \((i,j)\) pixel location, gradient vectors using finite differences are computed as
where \( f(i,j) \) represents gray values, \( G_x(i,j) \) and \( G_y(i,j) \) are x- and y-component of the gradient vector at \((i,j)\) pixel, respectively, and computed by

\[
G_x(i,j) = \frac{\partial f(i,j)}{\partial x} = \begin{bmatrix} G_x(i,j) \\ G_y(i,j) \end{bmatrix} \quad \frac{\partial f(i,j)}{\partial y}
\]

(4.6)

These equations are the same as used in the Sobel operator which is a gradient operator for edge detection (Gonzalez and Wintz, 1987). The orientation angle at the \((i,j)\) pixel location is computed by

\[
\theta_y = \tan^{-1}\left[ \frac{G_y(i,j)}{G_x(i,j)} \right]
\]

(4.8)

As Rao (1990) suggests, we used arctangent of one argument, which returns the angle in the range from \(-45^\circ\) to \(45^\circ\). According to him, "This effectively takes vectors that point in opposite directions, and maps them onto the same direction. This is a way of ensuring that gradient vectors pointing in opposite directions actually reinforce each other instead of canceling." However, the results are sometimes not reliable (Rao, 1990).
$G_{i,j}$ denotes the gradient of the blob at $(i,j)$ with $\theta_j$ orientation angle.

Figure 4.7 Computation of the local dominant orientation ($\theta_0$) in a group of blobs. (Adapted from Rao (1990) with permission of the publishers.)
Finally, the dominant local orientation is computed by a weighted average of the orientations of blobs in each window. Then, the estimated orientation is:

$$\hat{\theta} = \frac{1}{m \cdot l} \sum_{i=0}^{m-1} \sum_{j=0}^{l-1} \theta \cdot l_{ij} + \frac{\pi}{2}$$  \hspace{1cm} (4.9)$$

where $l_{ij} = [G_x^2(i,j) + G_y^2(i,j)]^{1/2}$. $m$ and $n$ determine the window size.

If the gradient vectors are smaller than a certain threshold value or the standard deviation exceeds a certain threshold value, we decided that there was not a dominant local orientation. Finally, the orientations were discretized by 45°:

- IF (0° ≤ $\theta$ < 22.5°)  \hspace{1cm} $\theta = 0°$
- IF (22.5° ≤ $\theta$ < 67.5°)  \hspace{1cm} $\theta = 45°$
- IF (67.5° ≤ $\theta$ < 112.5°)  \hspace{1cm} $\theta = 90°$
- IF (112.5° ≤ $\theta$ < 157.5°)  \hspace{1cm} $\theta = 135°$
- IF (157.5° ≤ $\theta$ < 180.0°)  \hspace{1cm} $\theta = 0°$

The discretization of the orientation, which is based on psychological experiments, was discussed in Chapter 3. Figure 4.8 shows a test image (downloaded from the Web site: www.ua.lg.pt) and its blobs, and Figure 4.9 represents computed dominant local orientations.
Figure 4.8 (a) Test image (size: 256 x 256). (b) Blobs of the test image ($\sigma = 0.55$).

Figure 4.9 Dominant orientation of each block (in degrees) of the test image.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>41 (45)</td>
<td>69 (90)</td>
<td>33° (45)</td>
<td>41 (45)</td>
<td></td>
</tr>
<tr>
<td>3 (0)</td>
<td>41 (45)</td>
<td>131 (135)</td>
<td>85 (90)</td>
<td></td>
</tr>
<tr>
<td>68 (90)</td>
<td>ND</td>
<td>20° (0)</td>
<td>1 (0)</td>
<td></td>
</tr>
<tr>
<td>64 (90)</td>
<td>1 (0)</td>
<td>46° (45)</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>

*Numbers in parenthesis are discretized orientation.*

*Note: ND denotes no dominant orientation, and * indicates wrong results.*
### 4.3.4 Computation of Spatial Frequency

The local spatial frequency is defined as $\omega = 2\pi / T$, where $T$ is the wavelength that is considered an average local dimension of blobs within a window (i.e., initially segmented block). Figure 4.9 illustrates the scheme of computing the blob size.

![Diagram of blob size computation](image)

*Arrows present the direction to compute the blob width.*

(a) orientation $= 0^\circ$, (b) orientation $= 90^\circ$, (c) orientation $= \theta$.

Figure 4.10 Computation of the average local blob size.

In order to compute the blob size for each window, it is suggested to compute the dominant local orientation before hand because the spatial frequency depends on the orientation. For $0^\circ$ orientation (see Figure 4.10(a)), the summation of black pixels ($i.e., f(x,y)=0$), which are occupied by blobs, is computed for vertical direction ($i.e.,$ column-wise). This sum is divided by the number of blobs and we obtain for

$$ T(\theta=0^\circ) = \frac{\sum_{i=0}^{n} \text{[run length for } f(i,j) = 0]}{N} \quad (4.10) $$
where \( m \) is number of rows in the window, and \( N \) is the number of the blobs. For 90° orientation (see Figure 4.10(b)), the same process is repeated, except for the horizontal direction, i.e., row-wise. In this case, the blob size is computed by

\[
T_{\theta=90^\circ} = n \frac{\sum_{j=0}^{n} \text{run length for } f(i, j) = 0}{N}
\]  

(4.11)

where \( n \) is the number of columns in the window. For 45° and 135° orientations (see Figure 4.10(c)), the vertical and horizontal blob sizes are computed and then combined by

\[
T_{\theta=45^\circ \text{or } 135^\circ} = \sqrt{[T_{\theta=0^\circ}]^2 + [T_{\theta=90^\circ}]^2}
\]  

(4.12)

If there is no dominant orientation in a window, we computed the blob widths for all directions then computed the average value. According to Voorhees (1987),

"the individual effects of blob density and blob size are harder to measure because they are interrelated."

The author performed various psychophysical experiments to determine minimum perceivable ratio of the blobs. The experiments show that the ratio is about 1.3. In this study, a ratio of 2 is chosen. Discretization of the blob size was performed as

- IF \( T \leq 4 \) \( T = 4 \)
- IF \( 4 < T \leq 8 \) \( T = 8 \)
- IF \( 8 < T \leq 16 \) \( T = 16 \)
- IF \( T > 16 \) \( T = 32 \)
4.3.5 Extraction of Texture Information

Extracting texture is performed by a convolution process. If there is a dominant orientation for a certain region of the image, the region is convolved with a pair of Gabor filters with corresponding orientation. If there is no dominant orientation, pairs of the Gabor filters with all orientations (i.e., $\theta = 0^\circ, 45^\circ, 90^\circ,$ and $135^\circ$) are used. The results of the multi-channel response are integrated as follows

$$f_n(i, j) = \sum_{k=1}^{n} \left[ f_n(i, j) * G(x, y|\sigma, \omega, \theta, \varphi = 0^\circ) \right] + \sum_{k=1}^{n} \left[ f_n(i, j) * G(x, y|\sigma, \omega, \theta, \varphi = 90^\circ) \right]$$

where $f_n(i, j)$ is a region of the image, and $*$ denotes convolution operator. In many regions of natural images, no dominant local orientations exists. Actually, computations of the absolute are approximation of the square root of sum of square of each convolution. To compute absolute provide computational simplicity without losing information. For example, Sobel operators also use the same scheme when combining horizontal and vertical edge operators. Laws (1980) states that

"It is probable that the human visual system avoids root mean square computation, and it is quite possible that simpler statistics are more appropriate for texture analysis."
The image is always processed with Gabor filter pairs ($\varphi = 0^\circ$ and $\varphi = 90^\circ$). The convolution process is performed region by region adaptively. That is, each region of the image is processed with a different set of parameters, selected by the proposed strategy.

The proposed process corresponds to Julesz and Bergen's (1983) attentive of texture discrimination because focal attention with scrutiny occurs at this stage. There is biological evidence that human observers pay more attention to regions with abrupt luminance changes than homogeneous regions of an image or natural scene (Hubel, 1988). During this process, it is without doubt that the human visual system analyzes texture adaptively by varying visual channels such as resolution, orientation, and spatial frequency.

### 4.4 Texture Segmentation

The purpose of image segmentation is to form meaningful regions by grouping features that have some common characteristics and properties distinct from their neighboring regions. Each region should be uniform and homogeneous with respect to some attributes such as tone, color, or texture (Low, 1991; Schalkoff, 1989). Image segmentation involves the procedure of classification and pattern recognition. Image segmentation methods are categorized into edge-based and region-based approaches. Texture segmentation is a region-based approach.
According to Pratt (1978),

"... Putting (texture) concept to practice, however, has been hindered by the lack of reliable and efficient means of detecting and measuring texture."

Pavlidis (1982) mentioned that

"Segmentation identifies areas of an image that appear uniform to an observer, and subdivides the image into regions of uniform appearance. It is relatively easy to define uniformity in terms of gray level or color. It is far more difficult to specify what we mean by uniform texture."

Even though the authors made these claims at a relatively early age of developing texture methods, the main issue of the statements is still true. So far, texture operators which result in uniform regions distinct from neighboring regions are not available. Therefore, in many cases, some additional classification procedure is required for texture segmentation. Unsupervised classification (or clustering) is considered a suitable technique for automatic processing, since it requires less prior information than supervised classification. Segmentation with unsupervised classification can be performed without an operator's intervention. Hence, many authors (e.g., Jain and Farrokhnia, 1991; Kervrann and Heitz, 1995; Khotanzad and Chen, 1989; Manjunath and Chellapa, 1991; Nguyen and Cohen, 1993; Teuner et al., 1995) prefer to use unsupervised texture segmentation.

In this study, ISODATA (Iterative Self-Organizing Data Analysis Technique, "A" being added for pronunciation) is applied for segmenting the images after filtering
with the Gabor filters. ISODATA is one of the iterative optimization methods for unsupervised pattern classification based on minimum distance decision. Iterative and self-organizing refer to repeated performance classification and recalculation statistics, and locating clusters with minimum user input, respectively.

ISODATA is similar to K-means clustering since the cluster centers are iteratively determined sample means in both methods (Gnanadesikan, 1977; Tou and Gonzalez, 1974). The number of cluster classes, convergence rate, maximum number of iterations, and minimum number of pixels in a cluster are input to the ISODATA clustering. Figure 4.11 illustrates the framework of the proposed procedure.
Figure 4.11 The framework of the proposed adaptive scheme for texture analysis by using 2D Gabor filters.
CHAPTER 5

EXPERIMENTS AND RESULTS

This chapter presents test results obtained with a software prototype of the proposed adaptive strategy for texture analysis and segmentation. First, the test images are described. They include aerial photographs, close-range images, a SAR image, a halftone image, and a multi-spectral image. Results are analyzed and the performance of the proposed strategy is evaluated. The resulting texture boundaries are compared with physical discontinuities of the images (i.e., intensity edges). Finally, an overall analysis of the experimental results is provided.

5.1 Description of the Test Images

5.1.1 Marchetsreut Image

The image shows a rural area around Marchetsreut near Passau in Germany (see Figure 5.1). The scale of the image is 1:15,000. It was digitized with a pixel size of 15 μm, 8 bits per pixel, yielding a ground resolution of 0.23m (Heipke et al., 1995). The overall entropy of the image is 6.8.
5.1.2 *OSU Campus Stereopair*

The images displayed in Figures 5.2 and 5.3 depict an aerial stereo pair. The photographs (Nos. 193 and 195) are taken over the central area of The Ohio State University campus. The original photo scale is 1: 4,000. The diapositives were scanned at a resolution of 30 μm pixel size by the Intergraph Corporation using the PhotoScan system. Images with a resolution of 512 × 512 pixels from the image pyramid were selected. Additionally, the images were resampled to epipolar geometry. The overall entropies of the images are 7.6 for the left and 7.3 for the right one.

5.1.3 *GPSVan Image*

The image shown in Figure 5.4 was taken with a Pulnix CCD camera installed on the GPSVan from the Center for Mapping of The Ohio State University. The number of pixels are 760 (H) × 480 (V) with a pixel size of 11.6 μm (H) × 8.0 μm (V). The nominal focal length is 8 mm. The overall entropy is 6.2.

5.1.4 *Tire Image*

An image of a tire is shown in Figure 5.5. The image is one of the sample images of MATLAB. The image size is 512 × 512 pixels. The overall entropy is 6.1.

5.1.5 *Synthetic Aperture Radar Image*

The synthetic aperture radar (SAR) image displayed in Figure 5.6 shows (according to the Alaska SAR Facility) “the tip of East Cape located in eastern Siberia near Alaska. The image covers approximately 100 km by 100 km.” The image was obtained by European Space Agency’s ERS-1 satellite and was processed at the
Alaska SAR Facility. The image was downloaded from the Alaska SAR Facility Web site: www.asf.alaska.edu. The image size is modified to 512 × 512 pixels. The overall entropy of the image is 4.2.

5.1.6 Halftone Image

The image shown in Figure 5.7 is a part of the cover page of the journal Photogrammetric Engineering and Remote Sensing (Vol. 63, No.9, 1997). It shows a color infrared image. The image was scanned with an HP ScanJet 4c color scanner at a resolution of 600 dpi in black and white (i.e., gray tone) mode. The image size is 512 × 512 pixels and the overall entropy is 4.7.

5.1.7 Multi-Spectral Image

The image of an agricultural area displayed in Figure 5.8 is a multi-spectral image with infrared (IR), green and blue bands. The image size is 512 × 512 pixels. The overall entropies for each band are 7.5, 7.3, and 7.2, respectively. The IR band contains more information than the other bands since the area is covered with vegetation.
Figure 5.1 Marchetsreut image.
Figure 5.2 *OSU Campus* image #193.
Figure 5.3 OSU Campus image #195.
Figure 5.4 GPSVan image.
Figure 5.5 Tire image.
Figure 5.6 SAR image.
Figure 5.7 Halftone image.
Figure 5.8 Multi-spectral image.
5.2 Procedure of the Adaptive Strategy

The procedure of the implementation is described step by step:

2. Performing progressive segmentation to define the image blocks.
3. Detecting blobs by the modified LoG operator.
4. Determining size of the Gabor filters for each block created in Step 2.
5. Computing dominant local orientation for each block created in Step 2.
6. Computing spatial frequency for each block created in Step 2.
7. Creating mirror padded image (see Section 5.2.4).
8. Performing adaptive filtering of the padded image with Gabor filters using the parameters selected in steps 4, 5, and 6.
10. Detecting texture boundary by the LoG operator.
11. Comparing texture boundaries with physical discontinuity (i.e., intensity edges).

The entire procedure was also illustrated in Figure 4.11 in Chapter 4.

5.2.1 Performing Progressive Segmentation

As described in Chapter 4, the purpose of progressive segmentation is to define the image blocks which form the basis of the adaptive processing. The criterion for further subdivision into four quadrants is based on the entropy. The starting block size and the criterion for stopping subdivision are discussed in Section 4.3. Some
experiments were performed with three different starting block sizes. In the case of 512 × 512 images, 128 × 128 pixels (i.e., 4 × 4 blocks), 64 × 64 pixels (i.e., 8 × 8 blocks), and 32 × 32 pixels (i.e., 16 × 16 blocks) are used as a starting point.

Figure 5.9 displays an example of the image blocks of the Marchetsreut image determined by the progressive segmentation scheme. This image was obtained by using the Sobel operator after adaptive filtering with the Gabor filters.

Note: The lines which look like contours were detected by computing gradient using the Sobel operator after processing with the Gabor filters.

Figure 5.9 An example of the progressive segmentation for the Marchetsreut image.
5.2.2 Texton Images with Detected Blobs

Blobs are detected with the modified LoG operator. The positive values of the convoluted image are considered to be blobs. A small value of the standard deviation ($\sigma$) provided reasonable size of blobs. After several experiments with different $\sigma$ values it was found that $\sigma = 0.55$ provides quite good results. This value was applied to all test images to obtain blobs. Since $\sigma$ is small (i.e., corresponding to the LoG operator size, $w = 5$), noise is hardly removed. This is not a great disadvantage at all, because noise is considered as part of texture. For a detailed discussion about noise issues refer to Chapter 4. Figure 5.10 through Figure 5.19 depict blobs of the test images.

5.2.3 Determination of the Parameters of the Gabor Filters

This procedure includes steps 4, 5, and 6 which determine filter size, orientation, and spatial frequency of the Gabor filters. Appropriate parameters are determined for each image block. Partial results of the Marchetsreut image are listed in Table 5.1 (see also Figure 4.6 in Chapter 4).
The spatial frequency is computed by $2\pi / (\text{Blob size})$.

ND in orientation denotes no dominant local orientation.

Some portion of the data is shown.

Table 5.1 Parameters of the Gabor filters of the Marchesreut image.
Figure 5.10 Blobs of the Marchesreut image.
Figure 5.11 Blobs of the OSU campus image #193.
Figure 5.12 Blobs of the OSU campus image #195.
Figure 5.13 Blobs of the GPSVan Image.
Figure 5.14 Blobs of the tire image.
Figure 5.15 Blobs of the SAR image.
Figure 5.16 Blobs of the halftone image.
Figure 5.17 Blobs of the IR band of the multi-spectral image.
Figure 5.18 Blobs of the green band of the multi-spectral image.
Figure 5.19 Blobs of the blue band of the multi-spectral image.
5.2.4 Image Padding

Step 7 for mirror padding is not absolutely necessary. Its purpose is to avoid zagging around borders of the image because the filter size varies block by block. Another advantage of using padded image is that they preserve the original image size after processing. Figure 5.20 shows the mirror padded image "Marchetsreut." Padding was applied to all test images (The padding size is 64 pixels).

![Padding border]

Figure 5.20 The Marchetsreut image with mirror padding.
5.2.5 Adaptive Processing to Extract Texture Information and Segmentation

This section presents the results of texture segmentation obtained after the adaptive processing with the Gabor filters. The purpose of these experiments is to test and evaluate the proposed texture analysis scheme, as described in Chapter 4.

*ISODATA* clustering requires the number of the classes to be classified. The test images that have been evaluated were classified with three or four different classes. With a maximum number of six iterations, the convergence rate was 95% for all images. Before demonstrating the proposed strategy, an experiment is shown in Figure 5.21 which was obtained by processing the entire image with a pair of Gabor filters, chosen after visual inspection of the image (*i.e.*, $\sigma = 12$, $\theta = 45^\circ$, $T=32$, and $\varphi = 0^\circ$, $90^\circ$).

![Image of Marchetsreut segmentation](image)

$\sigma = 12$, $\theta = 45^\circ$, $T=32$, and $\varphi = 0^\circ$, $90^\circ$. *Segmentation with 4 classes.*

Figure 5.21 Segmentation with a pair of the Gabor filters for the *Marchetsreut* image.
This experiment demonstrates that Gabor filters with fixed parameters do not provide reasonable results. Although, the major road was well recognized because its orientation and width probably coincide with the parameters of the selected filters, the overall result is not acceptable.

5.2.5.1 *Marchtsreut Image*

Three different starting block sizes were used: 128 x 128, 64 x 64, and 32 x 32 pixels. The number of classes for segmentation was four. The results show that a smaller starting block size provide sharper boundaries (see Figures 5.22 and 5.24). The vertical boundaries shown in Figure 5.23 and two small squares on the road shown in Figure 5.24 were caused by normalization after processing block by block.

![Segmentation of the Marchtsreut image](image)

Figure 5.22 Segmentation of the *Marchtsreut* image with starting block size of 128x128 pixels.
Figure 5.23 Segmentation of the Marchtsreut image with starting block size of 64x64 pixels.

Figure 5.24 Segmentation of the Marchtsreut image with starting block size of 32x32 pixels.
5.2.5.2 OSU Campus Images

Three different initial starting block sizes were used, and the number of classes for segmentation was four, as for the Marchetsreut image. The results of the segmentation are similar to each other regardless of the starting block size (see Figure 5.25 through Figure 5.30). This result can be explained as: The entropy of each block is relatively higher compared with the Marchetsreut image. Therefore, a small \( \sigma \) for the Gabor filters was applied to each block. In consequence, details could be preserved during the process. As in the Marchetsreut image, some false segmentations (vertical and horizontal boundaries) can be seen in Figure 5.25 through Figure 5.30 due a problem in normalization.

![Figure 5.25 Segmentation of the OSU campus image #193 with starting block size of 128x128 pixels.](image)
Figure 5.26 Segmentation of the OSU campus image #193 with starting block size of 64x64 pixels.

Figure 5.27 Segmentation of the OSU campus image #193 with starting block size of 32x32 pixels.
Figure 5.28 Segmentation of the *OSU campus* image #195 with starting block size of 128x128 pixels.

Figure 5.29 Segmentation of the *OSU campus* image #195 with starting block size of 64x64 pixels.
The buildings which are located in the left area in image #193 and the middle area in image #195 make shadows on the ground. It is clearly shown in all figures (see Figures 5.25 through 5.30) that the shadow areas were segmented with different texture regions in the near right sides of the buildings. Shadow changes reflectance and is one of the significant factors in texture analysis. Texture primitives are illumination invariant; however, it is difficult to group the same primitives as one under the reflectance change. This is also difficult even for human vision. The same situation will be found in the experiment with the tire image discussed in Section 5.2.5.4.
5.2.5.3 GPSVan Image

The size of the starting image block was 48(H) × 30(V) pixels, which created 16 × 16 image blocks. The number of classes for segmentation was four, as for the previous test images. The purpose of the experiments with a close-range image is to evaluate the influence of the filter size.

The results show that proper selection of the σ value (i.e., filter size) is a crucial factor. One interesting point was that shadow of the power lines on the road (see the original image shown in Figure 5.4) did not influence the results. Different types of asphalt surface (fresh or worn asphalt) were classified as different textures. However, the result was different for each experiment. The left side of the road surface has two different classes in Figure 5.31, while the right side of the road surface has two different classes in Figure 5.32.

Figure 5.31 Segmentation of the GPSVan image with fixed σ = 6.
Figure 5.32 Segmentation of the *GPSVan* image with fixed $\sigma = 3$.

5.2.5.4 Tire Image

The experiments in this section are aimed at examining the influence of the fixed filter size and dynamic filter size (*i.e.*, proposed adaptive scheme). The same size of the starting block sizes, $128 \times 128$ pixels were used for two experiments.

Figure 5.33 results from the adaptive processing with fixed $\sigma$ (*i.e.*, all parameters were selected by the proposed scheme while $\sigma$ was kept constant), while Figure 5.34 results from full adaptive processing (*i.e.*, all parameters were selected by the proposed scheme including $\sigma$). The results suggest that to use a dynamic $\sigma$ leads to better texture segmentation. The shadowed area (*i.e.*, the upper part of the tire) was segmented as different class from the unshadowed lower area. The reflectance of the surface can influence the texture analysis.
Figure 5.33 Segmentation of the tire image with fixed $\sigma = 12$.

Figure 5.34 Segmentation of the tire image with dynamic $\sigma$. 
5.2.5.5 SAR Image

The SAR image was segmented with three classes. We demonstrate one experiment of $32 \times 32$ pixel starting block, since other larger starting block sizes did not provide better results.

The result shown in Figure 5.35 is not as good as we expected. Some parts of the area were not correctly segmented. Especially, the lower right part of the original image (see Figure 5.5) is covered with homogeneous area, however, the area was segmented with two different classes. Some parameters might be determined incorrectly due to reflectance change in that area. However, the upper left part of the image was properly segmented.

Figure 5.35 Segmentation of the SAR image
5.2.5.6 Halftone Image

We include the scanned image from the journal cover in order to evaluate the performance of our scheme with halftone noise. The number of classes for segmentation is three and the starting block size was 64 × 64 pixels. The other starting block sizes provided similar results which are not displayed.

Figure 5.36 Segmentation of the halftone image

The result displayed in Figure 5.36 shows that the proposed scheme works well in the presence of noise. We treated the halftone pattern as noise. The second experiment was to apply the proposed strategy to the image after removing the halftone pattern. The procedure we performed to remove the halftone noise is as follows (Russ, 1992):
1. The original image was transformed into frequency domain by the Fourier transform.

2. The power spectrum of the Fourier transform was computed.

3. The power spectrum was displayed and the unique pattern (i.e., systematic and periodic patterns with well-defined narrow peaks) of the halftone noise was removed by setting the magnitude of the power spectrum to zero with visual inspection (see Figure 5.37).

4. A new image (possibly without halftone noise) in the spatial domain was obtained by the inverse Fourier transform.

Figure 5.37 Power spectrum of the halftone pattern of the scanned image (see Figure 5.7).
Comparing Figure 5.38 with the original image (Figure 5.7), we realize that the overall quality of the new image is smooth and free of the halftone pattern. Figure 5.39 displays the extracted halftone pattern.
However, it might be difficult to differentiate two images displayed in Figure 5.7 and in Figure 5.38, because the resolution of the printer (600 dpi) is not high enough and another halftone pattern was introduced during printing.

The result of the texture segmentation after removing the halftone noise is shown in Figure 5.40. The resulting segmentation has smooth boundaries and lost details compared with Figure 5.36. The reason for this is that some texture primitives might also be removed during the process of removing halftone noise. Therefore, special attention should be paid when noise removing is applied to an image. As discussed in Chapter 4, it may be reasonable to consider noise as texture when we deal with "image texture."
5.2.5.7 Multi-Spectral Image

The number of classes for segmentation is four and the starting block size was 128x128 pixels. The segmentation result from each band was different (see Figures 5.41, 5.42, and 5.43). IR and green bands provided more detailed segmentation for the vegetation areas.

Figure 5.41 Segmentation of the IR band of the multi-spectral image.
Figure 5.42 Segmentation of the green band of the multi-spectral image.

Figure 5.43 Segmentation of the blue band of the multi-spectral image.
5.2.6 Physical Discontinuity versus Texture Boundary

Texture boundaries (after segmentation) can be detected by any edge operator. Physical discontinuity includes object boundaries which are represented by intensity edges. In this study, the LoG operators were used to detect both intensity edges and texture boundaries. A small σ value (σ = 1) was used for extracting texture boundaries while a relatively large σ value (σ = 5) was used to detect intensity edges. The reason why we choose different σ's for detecting texture boundaries and intensity edges is as follows:

1. Texture boundaries were obtained by processing with the Gabor filters which have smoothing capability. Therefore, it is fair to perform smoothing with larger σ for edge detection in order to compare texture boundaries. However, to find the relationship of the smoothing effect between the Gabor filter and the LoG operator is beyond the scope of the study.

2. If we use a small σ for edge detection, false edges are also detected since the LoG is noise sensitive.

3. Finally, we selected an appropriate σ (= 5) for edge detection by visual examination after applying the LoG with several different σ's.

Intensity edges and texture boundaries of the test images are alternatively displayed in the following pages for the purpose of the visual comparison (see Figure 5.44 through Figure 5.59).
Figure 5.44 Edges of the *Marchetsreut* image.

Figure 5.45 Texture boundaries of the *Marchetsreut* image.
Figure 5.46  Edges of the *OSU campus* image #193.

Figure 5.47  Texture boundaries of the *OSU campus* image #193.
Figure 5.48 Edges of the OSU campus image #195.

Figure 5.49 Texture boundaries of the OSU campus image #195.
Figure 5.50 Edges of the GPSVan image.

Figure 5.51 Texture boundaries of the GPSVan image.
Figure 5.52 Edges of the tire image.

Figure 5.53 Texture boundaries of the tire image.
Figure 5.54 Edges of the SAR image.

Figure 5.55 Texture boundaries of the SAR image.
Figure 5.56 Edges of the halftone image.

Figure 5.57 Texture boundaries of the halftone image.
Figure 5.58 Edges of the IR band of the multi-spectral image.

Figure 5.59 Texture boundaries of the IR band of the multi-spectral image.
Figure 5.60 Edges of the green band of the multi-spectral image.

Figure 5.61 Texture boundaries of the green band of the multi-spectral image.
Figure 5.62 Edges of the blue band of the multi-spectral image.

Figure 5.63 Texture boundaries of the blue band of the multi-spectral image.
We evaluated the best texture boundary image based on the texture segmentation results for each test image. The properties of the images are summarized in Table 5.2.

<table>
<thead>
<tr>
<th>Texture boundary image</th>
<th>Starting block size (pixel)</th>
<th>σ of Gabor filter</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marchetsreut</td>
<td>32 × 32</td>
<td>Adaptive</td>
<td>5.45</td>
</tr>
<tr>
<td>OSU campus #193</td>
<td>32 × 32</td>
<td>Adaptive</td>
<td>5.47</td>
</tr>
<tr>
<td>OSU campus #195</td>
<td>128 × 128</td>
<td>Adaptive</td>
<td>5.49</td>
</tr>
<tr>
<td>GPSVan</td>
<td>48(H) × 30(V)</td>
<td>3 fixed</td>
<td>5.51</td>
</tr>
<tr>
<td>Tire</td>
<td>128 × 128</td>
<td>Adaptive</td>
<td>5.53</td>
</tr>
<tr>
<td>SAR</td>
<td>32 × 32</td>
<td>Adaptive</td>
<td>5.55</td>
</tr>
<tr>
<td>Halftone</td>
<td>8 × 8</td>
<td>Adaptive</td>
<td>5.57</td>
</tr>
<tr>
<td>Multi-spectral: IR band</td>
<td>128 × 128</td>
<td>Adaptive</td>
<td>5.59</td>
</tr>
<tr>
<td>Multi-spectral: green band</td>
<td>128 × 128</td>
<td>Adaptive</td>
<td>5.61</td>
</tr>
<tr>
<td>Multi-spectral: blue band</td>
<td>128 × 128</td>
<td>Adaptive</td>
<td>5.63</td>
</tr>
</tbody>
</table>

Table 5.2 Summary of the texture boundary images.

The texture boundaries of each image were visually compared with corresponding intensity edges. In general, the major texture boundaries coincided with the edges. Therefore, we concluded that texture boundary is part of the physical discontinuity. Examples for this conclusion are displayed in Figures 5.65 and 5.66. We extracted the common texture boundaries and intensity edges from the Marchetsreut image. Two different σ values of the LoG operator were chosen to detect edges; σ = 1 (see Figure 5.64), and σ = 5 (see Figure 5.44). It is obvious that more common boundaries are found with smaller σ.
Figure 5.64 Edges of the Marchetsreut image with $\sigma = 1$. 
Figure 5.65 Texture boundaries coincide with intensity edges with $\sigma = 5$.

Figure 5.66 Texture boundaries coincide with intensity edges with $\sigma = 1$. 
5.3 Summary of the Analysis

In this chapter, we demonstrated the proposed strategy with various test images. We regard that the overall performance of the strategy was reasonably acceptable. This is mainly due to the excellent behavior of the Gabor filters in analyzing texture. The choice of the parameters is the key issue when we apply the Gabor filters. The analyses of the results are summarized in Table 5.3.

<table>
<thead>
<tr>
<th>Test image</th>
<th>Analysis</th>
</tr>
</thead>
</table>
| Marchetsreut     | • The starting block size influenced the segmentation, *i.e.*, the smaller size provides the more detail segmentation.  
                   • The homogeneous areas were reasonably segmented.  
                   • There was a problem in normalization, especially for small block size. |
| OSU campus images| • The starting block size did not affect in extracting texture information.  
                   • Shadowed areas were segmented as different region.  
                   • There were problems in normalization.                                                        |
| GPSVan image     | • The size of the Gabor filter affected the overall results.  
                   • Multiple size of the filter may be applied based on the scale change which occurs across the image. |
| Tire image       | • The experiment shows that the influence of fixed and dynamic size of the filter.  
                   • The adaptive scheme was proven to work reasonably well through this example.  
                   • Texture analysis is influenced by reflectance change.                                     |

Table 5.3  Summary of the analysis.
Table 5.3 (continued)

<table>
<thead>
<tr>
<th>Image Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAR image</td>
<td>• The strategy did not work in some parts of the image due to improper determination of the parameters.</td>
</tr>
<tr>
<td>Halftone image</td>
<td>• The proposed scheme worked well in the presence of noise.</td>
</tr>
<tr>
<td></td>
<td>• The experiment shows that removing noise may cause loss of important texture information (i.e., texture primitives).</td>
</tr>
<tr>
<td></td>
<td>• It is reasonable to consider noise is part of texture.</td>
</tr>
<tr>
<td>Multi-spectral image</td>
<td>• Each band provides a different segmentation result.</td>
</tr>
<tr>
<td></td>
<td>• Information from the IR and green bands is more detail for vegetation areas.</td>
</tr>
</tbody>
</table>
CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

An adaptive strategy for extracting texture information was developed and the performance was evaluated. We have reviewed the existing texture analysis methods for better understanding of the properties of the texture and identifying problems in texture analysis, and developing a robust and reliable method. The properties of the Gabor filters, which are conceived as hypothetical structures of retinal receptive fields in the human visual system, were identified and the advantages of using the Gabor filters were discussed. Human can successfully extract texture information and use it for image understanding, surface reconstruction and scene analysis without difficulty. Therefore, it is natural to be attracted by approaches which resemble the performance of human visual perception.

The purpose of this study was to develop an efficient strategy in applying Gabor filters for texture extraction and segmentation. An effort was made to propose an adaptive strategy that can select appropriate parameters of the Gabor filters. The motivation is to achieve automatic selection of the appropriate parameters without preliminary visual inspection and human intervention. The entire framework of
the proposed strategy is based on the theory of texture perception in human visual system.

Information from texture identification and segmentation can be utilized in digital photogrammetry, computer vision, remote sensing, and geographic information system (GIS) such as matching, surface reconstruction, object recognition, classification of the terrain, boundary detection, etc.

6.1 Conclusions

Experiments with various test images of natural scenes were carried out to examine and demonstrate the feasibility of the proposed adaptive texture analysis method, and the following conclusions are drawn:

- Two-dimensional Gabor filters should be used with appropriate parameters in order to extract texture information accurately.

- Improvement of the texture segmentation can be achieved after processing the image with the Gabor filters region by region which are progressively determined based on the characteristics of the image.

- The entire framework is to integrate multi-channel filter responses which is accompanied by multi-resolution texture analysis in an image. This is based on the human visual information processing.
• Adaptive processing for dynamic determination of the parameters reduces human intervention for visual examination and possibly the prior knowledge of the image.

• The optimal block size depends on the image contents. Therefore, the starting block size may influence the quality of the final segmentation, even though it is not significant as long as the size is not too big or not too small.

• The performance of the proposed method is relatively outstanding with the presence of noise. Removing noise may cause loss of the texture primitives, therefore, special attention should be paid. In addition, it might be reasonable to consider the noise as part of the texture.

• Multi-spectral imagery provides more information in texture analysis.

• More meaningful object boundaries can be obtained from texture segmentation.

6.2 Recommendations for Future Research

The following issues for extension to this study and applications in digital photogrammetry, GIS and computer vision are recognized:

1. Recommendations for extension of this work:

   • Development of a more reliable method to compute the parameters of the texture primitives in order to improve the performance of the Gabor filters.
• Proper normalization and integration of the multi-channel information are required to improve texture segmentation.

• Scale space approach with the proposed adaptive strategy has potential to improve the performance of the texture analysis.

2. Suggestions for applications in digital photogrammetry, GIS, and computer vision:

• Matching is the essential task to achieve automatic orientation in digital photogrammetry. The conventional entity for matching is a point or line feature. The texture information (i.e., properly identified texture primitives or texture parameters) can be used as a matching entity with the conventional entity. This has potential to improve the matching result. In addition, texture boundary can be used as initial global matching.

• In GIS, texture segmentation of the terrain provides classification of land type or use. For boundary detection, more accurate and physically meaningful result can be obtained by combining the classification technique used in remote sensing and segmentation from texture information.

• Identification of the texture parameters (e.g., orientation and spatial frequency) provides possibility to develop the symbolic description of texture which may be a gateway to high-level vision.
BIBLIOGRAPHY


127


