

BOSTON UNIVERSITY
COLLEGE OF ENGINEERING

Dissertation

**TEXTURE- AND STRUCTURE-BASED IMAGE REPRESENTATION
WITH APPLICATIONS TO IMAGE RETRIEVAL AND
COMPRESSION**

by

ZHIHUA HE

B.S., University of Electronic Science and Technology of China, 1999
M.S., Drexel University, 2002

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2007

PREVIEW

UMI Number: 3246606

INFORMATION TO USERS

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 bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send 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.

UMI[®]

UMI Microform 3246606

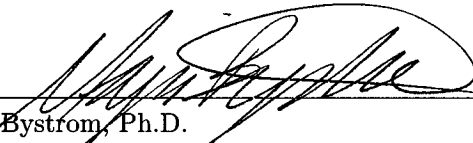
Copyright 2007 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

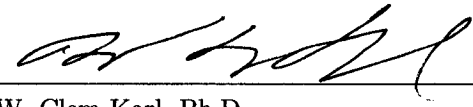
Approved by

First Reader



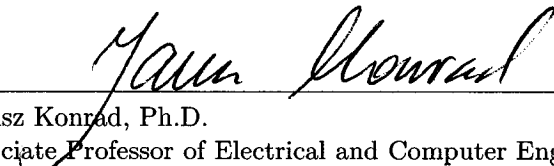
Maja Bystrom, Ph.D.
Associate Professor of Electrical and Computer Engineering

Second Reader



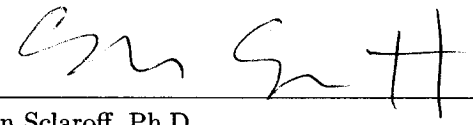
W. Clem Karl, Ph.D.
Professor of Electrical and Computer Engineering and Professor of
Biomedical Engineering

Third Reader



Janusz Konrad, Ph.D.
Associate Professor of Electrical and Computer Engineering

Fourth Reader



Stan Sclaroff, Ph.D.
Associate Professor of Computer Science

PREVIEW

Acknowledgments

I am usually not good at writing acknowledgment, one reason is that I am never good at writing, and another reason is that I believe action speaks more for words. However, when I look back over all these years I spent in schools, I do realize that I owe lots of thanks to people along the way.

First, I would like to thank my parents, ShuiBao He and Yong Xu. They have provided me with a stable home, without which I can never dream of getting advanced education; especially thanks for my father, who gave me a dream of being a scientist from my early childhood. I definitely want to thank my brother, Zhipeng He, who takes the responsibility to take care of the family while I am away from home.

Then, a special thank goes to my PhD adviser, Dr. Maja Bystrom. Without her continuous support, I will not be able to come here for my PhD research. Without her guidance and patience, I can hardly imagine how I could get through the PhD research alone. I will not forget the warmth brought by her in those cold winters. I also want to thank Dr. Arkady Kopansky, my first office-mate here, who has been always willing to help and answer the questions.

Here, I want to thank my committee members: Dr. Stan Sclaroff, Dr. Janusz Konrad and Dr. Clem Karl. Thank you for your careful reviews and valuable comments.

In the mean time I would also like to express my gratitude to all my teachers from primary school to college. You have led me step by step into the science world, and have shown me how to approach problems from different aspects.

Finally, I want to express my gratitude to all my friends, life will be so lonely and so boring without you guys! It is really hard for me to name all of you here, but I want to let you all know that I really appreciate your love and care, and I will always be willing to share mine with you. Below I want share my special thanks to Fang, Doreen, Violet, Shuang and Ling. Thank you for being there with me always.

**TEXTURE- AND STRUCTURE-BASED IMAGE REPRESENTATION
WITH APPLICATIONS TO IMAGE RETRIEVAL AND
COMPRESSION**

(Order No.)

ZHIHUA HE

Boston University, College of Engineering, 2007

Major Professor: Maja Bystrom, Ph.D.,
Department of Electrical and Computer Engineering

ABSTRACT

The design of an efficient image representation methods using small numbers of features can facilitate image processing tasks such as compression of images and content-based retrieval of images from databases. In this dissertation, three methods for capturing and concisely representing two distinguishing characteristics of images, namely texture and structure, are developed. Applications of these compact representations of image characteristics to image compression as well as retrieval of images and hand-sketches of images from databases are given and performance is compared with other compression and retrieval methods.

The first method to be introduced is a directional, hidden-Markov-model-based method for succinctly describing image texture using a small number of features. This method employs the well known, multi-scale contourlet and steerable-pyramid transforms to isolate in different subbands the edges that comprise the image texture. Statistical inter- and intra-subband dependencies are captured via hidden Markov models, and model parameters are used to represent texture in small feature sets. Application of this method to content-based retrieval of images with homogeneous textures from database is shown. At the similar computation cost, about 10% higher retrieval rates over comparable methods are

demonstrated; when approximately one third fewer features are used, similar retrieval rates can be obtained using the proposed method.

A method for concisely describing large image structures, that is, significant image edges, is then proposed. This method decomposes an image using the contourlet transform into directional subbands which contain edges of different orientations. Each subband is then projected onto its associated primary and orthogonal directions and the resulting projections are filtered and then modeled using piece-wise linear approximations or Gaussian mixture models. The model parameters then form the concise feature sets used to represent the image's structure. An application of this image-representation method to retrieval of images from databases based on users' sketches of the images is shown. An retrieval rate increase of 13% using the proposed method is demonstrated over a current spatial-histogram-based method.

Finally, a new multi-scale curve representation framework, the chordlet, is constructed for succinct curve-based image structure representation. This framework can be viewed as an extension to curves of the well known beamlet transform, a multi-scale line representation system. In this dissertation, the representation efficiency, in terms of bits versus distortion, of the chordlet transform is compared with that of the beamlet transform. An algorithm for performing a fast chordlet transform has been developed. A chordlet-based coding system is constructed for application of the chordlet transform to compression of image shapes. By using the proposed method increased compression is obtained at lower distortion when compared with two well known methods.

PREVIEW

Contents

1	Introduction	1
2	An Overview of Multi-scale Representation	9
2.1	Introduction	9
2.2	Classical Transform	10
2.3	Gaussian Pyramid and Laplacian Pyramid	12
2.4	Wavelet Transform and Wavelet Packets	16
2.5	Directional Multi-Scale Transform	20
2.5.1	Steerable Pyramid Transform	23
2.5.2	Ridgelet	24
2.5.3	Curvelet	27
2.5.4	Contourlet	29
2.6	Multi-scale Line and Curve Representation	33
3	Texture Representation	35
3.1	Background on Hidden Markov Models	37
3.2	Texture Representation using Directional HMMs in the Contourlet Domain	40
3.2.1	HMM Exploiting Luminance In-band Directional Dependencies in Contourlet Domain	42
3.2.2	HMM Exploiting Luminance Cross-scale and Directional In-band De- pendencies in the Contourlet Domain	45
3.2.3	HMM Exploiting Color In-Band Directional Dependencies in the Contourlet Domain	48

3.2.4	HMM Exploiting Color Cross-Scale Dependencies in the Contourlet Domain	50
3.3	Rotation-Invariant Texture Representation Using an HMM and Exploiting In-band and Cross-band Dependencies in Steerable Pyramid	52
3.4	Applications to Texture-Based Retrieval	57
3.4.1	Non-Rotation-Invariant Texture-Based Retrieval	57
3.4.2	Rotation-Invariant Texture-Based Retrieval	64
3.5	Conclusion	66
4	Structure Representation	69
4.1	Structure Representation Using Directional Projection	70
4.1.1	DP-SR Framework	70
4.1.2	Directional Filtering	71
4.1.3	Directional Projection	72
4.1.4	Low-Pass Filtering of 1D Profiles	74
4.2	Application to Sketch-Based Image Retrieval	74
4.2.1	DP-SR-I: Piece-wise Linear Profile Modeling and Similarity Measurement	74
4.2.2	DP-SR-II: Non-linear Gaussian Mixture Profile Modeling and Similarity Measurement	76
4.2.3	Results	80
4.3	Conclusion	83
5	Curve Representation Via the Chordlet	86
5.1	Introduction	86
5.1.1	Beamlet Construction	87
5.2	Chordlet Construction	91
5.3	Comparison Between Beamlet and Chordlet for the Representation Efficiency	93
5.3.1	Formulation	94

5.3.2	Curve Representation Via Chordlet and Beamlet Elements . . .	95
5.3.3	Comparison and Discussion	98
5.4	Implementation of the Chordlet Transform	100
5.4.1	Construction of FFT-based Chordlet	102
5.4.2	Direct Chordlet Versus FFT-Based Chordlet	108
5.4.3	Partial Transform	111
5.4.4	Savings Due to Combination of the Direct Chordlet and the FFT- based Chordlet Transform	112
5.4.5	Is the Estimated Computational Cost for the Direct Chordlet and FFT-based Chordlet Transforms Consistent with the True Value? . .	115
5.5	Chordlet Coefficient Significance Model	116
5.5.1	Coefficient Significance Model	117
5.5.2	Simulation	122
5.5.3	Discussion	124
5.6	A Thresholding-based Chordlet Representation System	125
5.6.1	An Image Structure Representation System	125
5.6.2	Simulation and Discussion	127
5.7	Rate-Distortion Based Multi-Scale Chordlet Representation Systems	133
5.7.1	Introduction	133
5.7.2	Fundamentals and Definitions	134
5.7.3	Mono-Scale Chordlet Representation	137
5.7.4	Multi-Scale Chordlet Representation	141
5.8	Applications	146
5.8.1	Chordlet-Based Representation Coder	146
5.8.2	Simulation	149
5.9	Conclusions and Discussion	156
5.10	Appendix	157
5.10.1	Beamlet Dictionary Size	157

5.10.2	Chordlet Dictionary Size	158
5.10.3	The Definition of Curvature	163
5.10.4	The Proof of Lemma 1	163
5.10.5	The Optimal Line Approximation for a SmoothCurve	164
165		
5.10.7	A Lower-Bound for the Average Number of Beamlets Required for a SmoothCurveSegment C_s in an $N \times N$ Lattice Grid	168
5.10.8	The Proof of Lemma 3	171
5.10.9	The Proof of Lemma 4	172
5.10.10	The Proof of Lemma 5	174
5.10.11	The Proof of Lemma 6	177
6	Contributions and Future Work	180
6.1	Contributions	180
6.2	Future Work	182
	References	184
	Curriculum Vitae	190

PREVIEW

List of Figures

1·1	Example of texture images. (a) Image <i>Tile</i> . (b) Image <i>Bark</i>	2
1·2	An example of sketches and corresponding images. (a) <i>Flower 1</i> . (b)-(c) Two sketches of <i>Flower 1</i> . (d) <i>Flower 2</i> . (e)-(f) Two sketches of <i>Flower 2</i> . .	6
2·1	A block diagram of a three-level Gaussian pyramid.	13
2·2	An example of a three-level Gaussian pyramid decomposition of image <i>Lena</i>	13
2·3	A block diagram of a two-level Laplacian pyramid.	15
2·4	An example of a two-level Laplacian pyramid decomposition of image <i>Lena</i>	15
2·5	An illustration of the analysis and synthesis filter bank. (a) Analysis filter bank. (b) Synthesis filter bank.	17
2·6	A block diagram of a two-level wavelet pyramid.	18
2·7	An example of a two-level wavelet decomposition of image <i>Lena</i>	19
2·8	A block diagram of a two-level wavelet packet pyramid.	21
2·9	An example of a two-level wavelet packet decomposition of image <i>Lena</i>	22
2·10	A block diagram of a two-level steerable decomposition.	24
2·11	An example of a two-level steerable pyramid decomposition of image <i>Lena</i>	25
2·12	A block diagram of a two-level contourlet decomposition.	30
2·13	An example of a two-level contourlet decomposition of image <i>Lena</i> with subbands labeled according to the block diagram in Fig. 2·12.	31
2·14	An illustration of the representation of a curve using wavelet, curvelet/contourlet, beamlet and chordlet. (a) Representation using wavelet. (b) Representation using curvelet/contourlet. (c) Representation using beamlet. (d) Represen- tation using chordlet.	32

3-1	An illustration of a two-level contourlet decomposition of image <i>Peppers</i> . . .	36
3-2	Three classes of Markov dependencies in a wavelet pyramid. (a) Cross-scale dependencies between subband pairs at scale L and $L + 1$ (b) In-band spatial directional dependencies at scale L for horizontal, vertical and diagonal subbands along the horizontal (H), vertical (V) and diagonal (D) directions, respectively. (c) Cross-band dependencies within scales L and $L + 1$ among horizontal, vertical and diagonal subbands.	38
3-3	An illustration of cross-scale dependencies, indicated by ellipsoids, in contourlet pyramid of Fig. 3-1. (a) The vertical subband at scale 2, labeled 21 in Fig. 3-1. (b) The child subband of (a), which is the near-vertical subband at scale 1 labeled 12 in Fig. 3-1.	39
3-4	An illustration of directional spatial in-band dependencies in the contourlet pyramid of Fig. 3-1. The regions of dependency are indicated by ellipsoids. (a) Vertical in-band dependencies in vertical subband at scale 2, labeled 21 in Fig. 3-1. (b) Horizontal in-band dependencies in horizontal subband at scale 2, labeled 23 in Fig. 3-1.	40
3-5	An illustration of cross-band dependencies, indicated by the ellipsoids, in the contourlet pyramid of Fig. 3-1. (a) Near vertical subband at scale 1, labeled 11 in Fig. 3-1. (b) Near vertical subband at scale 1, labeled 12 in Fig. 3-1.	41
3-6	An illustration of the luminance HMM exploiting directional in-band Markov dependencies for the vertical and horizontal subbands.	44
3-7	An illustration of the luminance HMM exploiting directional in-band dependencies and cross-scale dependencies in a vertical subband.	47
3-8	Two levels of contourlet decomposition for three color components of test color image. (a) Original color image. (b) Red component decomposition. (c) Green component decomposition. (d) Blue component decomposition. .	49

3.9	An illustration of a color HMM exploiting directional in-band dependencies in vertical triplets.	51
3.10	An illustration of a color HMM exploiting cross-scale dependencies in vertical triplets.	52
3.11	An illustration of a two-level steerable pyramid decomposition of a texture image. The image labeled H is the high-pass subband image, images labeled $B1$ and $B2$ are two directional subband images at the first level, the images labeled $LB1$ and $LB2$ are the two directional subband images at the second level, and the image labeled LL is the low-pass subband image.	54
3.12	An illustration of a directional HMM exploiting vertical dependencies of the hidden states of cross-band vectors.	56
3.13	An illustration of a texture retrieval system exploiting contourlet domain-based HMMs. (a) A general retrieval system; $N = 1$ when only luminance component is adopted for feature extraction, and $N = 3$ when color components are adopted for feature extraction. (b) The feature extraction subsystem using the proposed texture representations.	58
3.14	Sample texture images from the MIT VisTex database.	59
3.15	Top six retrieval results for texture image <i>Bark</i> using color-based and intensity-based HMMs and the method in (Po, 2003). (a) Query image <i>Bark</i> with three other bark images generated from the same original image. (b) Retrieval results using the proposed color HMM exploiting cross-scale dependencies. (c) Retrieval results using the proposed color HMM exploiting in-band dependencies. (d) Retrieval results using the proposed HMM exploiting the luminance in-band and cross scale dependencies. (e) Retrieval results using the method of (Po, 2003). In each case the query image is shown on the left and top six retrieved images are shown on the right.	62

3.16	Top six retrieval results using the color in-band model, the luminance in-band model and the luminance cross-scale model of (Po, 2003). (a) Query image <i>Flower</i> with three other flower images generated from the same original image. (b) Retrieval results using the luminance in-band model. (c) Retrieval results using the luminance cross-scale model of (Po, 2003). (d) Retrieval results using color in-band model. In each case the query image is shown on the left and top six retrieved images are shown on the right. . . .	63
3.17	Sample texture images from the Brodatz texture database (University of Southern California, Brodatz texture image database, 2004).	64
3.18	An illustration of rotated texture images. From top left to bottom right, the images with rotation angles: 0° (original image), 30° , 60° , 90° , 120° , and 150°	65
4.1	DP-SR framework.	71
4.2	A retrieval system based on DP-SR.	71
4.3	Sample image on edge map, sketch and representative subband projections. (a) Original image. (b) Horizontal subband image with principal horizontal projection and orthogonal vertical projection. (c) Image edge map. (d) Sample sketch of the image.	72
4.4	Image profiles of Fig. 4-3(b). (a) Profile 1: projecting along the principal direction (0°) of the vertical subband. (b) Profile 2: projection along the orthogonal direction (90°) of vertical subband. (c) Low-pass filtered version of (a). (d) Low-pass filtered version of (b). (e) Peak-valley approximation of (a). (f) Peak-valley approximation of (b).	73
4.5	Sample images from MPEG-7 <i>Shape</i> database (Latecki et al., 2000) and their corresponding edge-maps and sketches. (a) Images from <i>Shape</i> database (Latecki et al., 2000). (b) Edge-map of images in (a). (c) Sketches of images in (a).	81
4.6	Sample images from the <i>Flower</i> database.	83

4.7	Sample sketches of images from <i>Flower</i> database. (a) Original image and sketches of <i>Flower</i> 33. (b) Original image and sketches of <i>Flower</i> 91. (c) Original image and sketches of <i>Flower</i> 95.	83
5.1	An illustration of beamlet elements at different scales. (a) Scale 1. (b) Scale 2. (c) Scale 3. (d) Scale 4.	89
5.2	An example of representation of a curve, shown in red, by beamlet elements, shown in green, in a beamlet pyramid.	90
5.3	Construction of a set of chordlet elements with different curve heights but common end points and chord length $2r$. The solid curves are chordlet elements constructed from dotted circles with curve height k_i	93
5.4	An illustration of the approximation of SmoothCurveSegment $C_{ij,t}$ with chordlet $c_{ij,k}^+$	96
5.5	An illustration of the approximation of a curve by line segments.	97
5.6	Average number of beamlet elements required for approximation of C_s in an $N \times N$ grid with approximation error ϵ_0	101
5.7	Illustration of the computation of circle radii.	104
5.8	Illustration of approximation of diagonal line segments to determine the expanded image size.	106
5.9	Illustration of circular convolution aliasing for a 1D signal. The blue region represents the periodic signal y , and the red region represents the signal x . Different periods of y are identified by T_i . Note that signal x overlaps two neighboring periods of y , and this introduces aliasing.	107
5.10	Illustration of suppression of circular convolution aliasing by padding the signal with zeros. The blue region denotes the valid y signal, while the blank blue region denotes the zeros-padded region. Note that no aliasing is introduced when the blank blue region is large enough.	108

5-11	Illustration of circular convolution aliasing for expanded image I_e and circular kernel K	109
5-12	Illustration of suppression of circular convolution aliasing by padding the signal with adequate zeros to form expanded image I_e	110
5-13	An illustration of a curve segment Γ_2 in R_J^+ , and the curve Γ_1 , which returns maximum curve approximation error $\tilde{\epsilon}(m, n)$. Bounding curves for Γ_1 and Γ_2 are given by dashed lines.	119
5-14	An illustration of the areas S_1 and S_2 for Γ_1 and Γ_2 of Fig. 5-13, respectively.	120
5-15	The area S_{12L} between Γ_1 and Γ_2 in Fig. 5-13.	121
5-16	An illustration for computing the overlapping area, S_{ol} , of Γ_1 and Γ_2 of Fig. 5-13.	122
5-17	Probabilities for curves to be captured at different scales as a function of tolerable approximation errors.	123
5-18	Thresholding-based chordlet structure representation of test image <i>Four-Petal</i> . (a) Original image. (b) Low-resolution edge map. (c) Structure at scale 1. (d) Structure at scale 2. (e) Structure at scale 3. (f) Structure at scale 4. (g) Structure at scale 5. (h) Overall representation.	128
5-19	Chordlet structure representation of test image <i>Lily</i> . (a) Original image. (b) Low-resolution edge map. (c) Structure at scale 1. (d) Structure at scale 2. (e) Structure at scale 3. (f) Structure at scale 4. (g) Structure at scale 5. (h) Overall representation.	129
5-20	Chordlet structure representation of test image <i>Lily</i> when neither intra-scale inhibition nor cross-scale inhibition are utilized. (a) Original image. (b) Low-resolution edge map. (c) Structure at scale 1. (d) Structure at scale 2. (e) Structure at scale 3. (f) Structure at scale 4. (g) Structure at scale 5. (h) Overall representation.	131

5·21	An illustration of the first three levels of a fat quadtree representation of an image with two curves. The curve segments indicated by solid red lines are actual curves, and the curve segments indicated by dash-dotted blue lines are curves to be represented at finer scales. (a) Original image. (b),(c) Representation of each curve of (a) at level 1. Here two nodelets are generated, the curves indicated by blue dash-dotted line indicates the curves will be represented through chordlet split, corresponding to nodelet label Q , at finer scales. (d),(e) Representation of each curve of (a) at level 2, corresponding to (b),(c), respectively. (f),(g) Representation of each curve of (a) at level 3, corresponding to (b),(c), respectively.	147
5·22	The map images used in simulations for shape representation and compression. From top left to bottom right: Belgium, Brazil, China, Cuba, Germany, and Switzerland.	150
5·23	Sample shape images used in simulations for shape representation and compression.	151
5·24	An illustration of shape image <i>Deer</i> using different representation schemes. (a) Shape image <i>Deer</i> . (b) JBEAM representation, with distortion 182 and bit count 7008. (c) Quadtree-based chordlet representation, with distortion 169 and bit count 6424. (d) Fat quadtree-based chordlet representation, with distortion 150 and bit count 6184.	154
5·25	An illustration of very-low-bit-rate representation of shape image <i>Apple</i> using different representation schemes. (a) Shape image <i>Apple</i> . (b) JBEAM representation, with bit count 312. (c) Quadtree-based chordlet representation, with bit count 312. (d) Fat quadtree-based chordlet representation, with bit count 184.	155
5·26	An illustration of the line-segment approximation of a SmoothCurve.	165

Chapter 1

Introduction

Image representation serves as an integral part of image processing algorithms. For different applications, designing or selecting an appropriate representation for image data plays an important role. While image representation is a broad research field, in this dissertation, we are interested in efficient representation of primary image features. More specifically, texture representation and structure representation are studied. All of the work included in this dissertation is based on multi-scale frameworks. To help the readers have a better understanding of our work, Chapter 2 gives an overview of multi-scale representations. In particular, the directional multi-scale representations are discussed in more details, which are used throughout this dissertation.

Examples of texture images are given in Figs. 1.1 (a)-(b). In texture representation, the key is to find an accurate and efficient description of the texture information in an image. This representation could be used for texture-related applications such as texture classification and texture retrieval. For example, for an image with checker-board patterns, one texture representation could be simply text, e.g., '*checker-board texture*'. Another texture representation could be the histogram of the image, which characterizes the statistical distribution of the colors/intensities. Furthermore, different models can be adopted for the characterization of color/intensity distributions. For this example, if the user would like to find the image with the similar textures as this checker-board image in an image database, which is known as texture-based image retrieval, then the text-based representation might be applicable for retrieval in small database. However, when large image database is considered, text annotation of every image in the database becomes very expensive, thus, the histogram-based representation might be more appropriate, since the representation can be

automatically generated for all the images. On the other hand, the histogram-based representation requires the storage of the entire histogram, and the histogram only characterizes the frequencies of colors/intensities in its corresponding image and discards geometry information, therefore, it might be a poor descriptor for images with similar textures. Thus, more sophisticated statistical modeling of the color/intensity distribution may offer a more flexible and robust representation for texture-based retrieval applications.

A strict definition of texture is debatable because of its observer-dependent nature for different applications (Wang et al., 1998). In general, two classes of methods, namely, the structural-based and statistical-based methods, are used for texture representation (Long et al., 2003). Structural-based representation is typically used for textures with strong and regular patterns. On the other hand, statistical-based methods are used more widely. Assuming textures are generated from certain statistical sources, statistical models, for which rich mathematical frameworks are available, can be constructed to capture texture information in a small number of parameters. However, depending on the construction of statistical models, structural information of textures might be lost or reduced to certain extent.

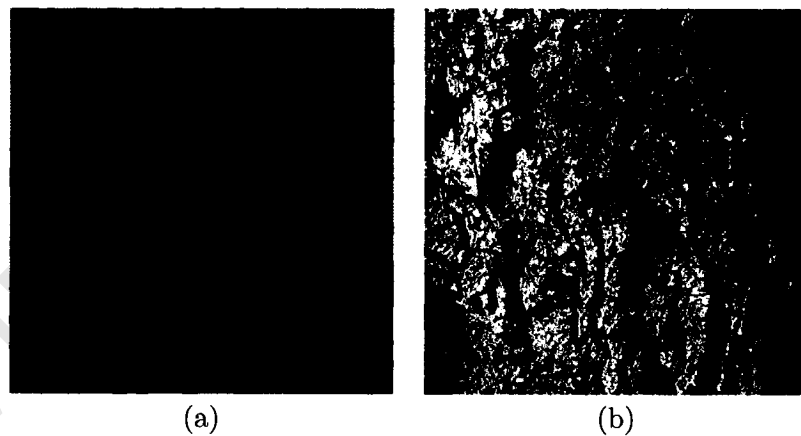


Figure 1-1: Example of texture images. (a) Image *Tile*. (b) Image *Bark*.

Many texture feature models are derived from an early work of Tamura, et al. (Tamura

et al., 1978), which models texture features based on human texture perceptions using coarseness, contrast, directionality, linelikeness, regularity, and roughness. In recent years, with the wide applications of multi-scale analysis in image processing, active studies (Randen and Husoy, 1999; Manjunath et al., 2000; Wouwer et al., 1999; Crouse et al., 1998) have been carried out on texture models in the multi-scale spatial-frequency domains, in particular, the wavelet domain. The popularity of multi-scale wavelet decomposition in image analysis is partly supported by the physiological studies of the human visual system (HVS) (Valois and Valois, 1988), and facilitated by the success of wavelets in still image compression in JPEG-2000 (Joint Photographic Experts Group Homepage, 2005). In (Randen and Husoy, 1999), Randen and John gave a comparative study of texture feature extractions using different filter banks, with local energy from each filter bank used as features. The role of different filter banks is compared by using the same similarity measure for corresponding texture features in retrieval in (Randen and Husoy, 1999). The first- and second-order statistics of the wavelet coefficients are employed by Wouwer, et al. in (Wouwer et al., 1999) for texture modeling. A texture descriptor based on the Gabor decomposition is described in (Manjunath et al., 2000), where the structuredness of the textures are captured by a perceptual browsing component and the statistics of the texture is captured by a similarity retrieval component.

While texture models based on local energy and basic statistics in the wavelet domain yield reasonable retrieval performance in texture-based retrieval applications, more sophisticated statistical models in the wavelet domain are developed to capture both statistical and structural properties of textures. In particular, wavelet-based hidden Markov models (HMMs) are widely used in texture modeling (Crouse et al., 1998; Do and Vetterli, 2002a; Fan and Xia, 2003; Po, 2003). In (Crouse et al., 1998), Crouse, et al. exploited cross-scale hidden Markov dependencies in the estimation of statistical model parameters for feature extraction in the wavelet domain, and suggested possible further extensions utilizing cross-band and in-band (spatial) dependencies. In (Do and Vetterli, 2002a), Do and Vetterli adopted a vector HMM model for characterization of textures in the steerable pyramid

wavelet domain, investigating both cross-scale and cross-band dependencies. Fan and Xia (Fan and Xia, 2003) have proposed a HMM model for texture retrieval and synthesis also considering cross-band correlation. Recently, Po and Do (Po, 2003) have studied statistical modeling in the contourlet domain (Do and Vetterli, 2003), and exploited cross-scale dependencies for their HMM approach to the texture retrieval problem. In Chapter 3, we investigate the importance of different dependencies in constructions of contourlet-based HMMs, along with possible combination of color information for texture representation. The application to the texture-based image retrieval is studied.

The high-level visual contents of an image, including objects' shapes and their spatial relationships, can be characterized in shape and spatial features. Shape and spatial feature extraction is often object-based, i.e., an image is usually first classified/segmented into regions/objects before feature extraction. Two classes of methods have been used for shape extraction, namely, boundary-based approaches and region-based approaches (Long et al., 2003). Boundary-based approaches extract features based on object contours; this includes approaches such as Fourier-based shape descriptors (Persoon and Fu, 1977; Kauppinen et al., 1995), and curvature-based shape descriptors (Abbasi et al., 1999). The classical region-based approaches focus on the statistical moments of the object region (Hu, 1962; Yang and Albrechtsen, 1994). Spatial features representing the spatial relationships between objects can be modeled using 2-D strings (Chang et al., 1987; Lee et al., 1992) or symbolic images (Gudivada and Raghavan, 1995). Most of the methods for shape and spatial feature extraction rely on image segmentation; however image segmentation is another active research field, and accurate segmentation is in general difficult to achieve. Therefore, applications relying on robust shape features and spatial feature extraction are limited as well. On the other hand, structure representation, which includes shape and spatial information, can be adopted with more flexibility and robustness in different image applications.

Many of approaches to structure representation focus on closed shape description (Latecki et al., 2000; Grigorescu and Petkov, 2003; Abbasi et al., 1999; Persoon and Fu, 1977).

While these results are important for many applications, the work presented in this dissertation focuses on representation of the edges/boundaries in an image without restriction to closed shape. The edges/boundaries in an image are usually approximated by line segments and curve segments. Accordingly, two structure representation frameworks are constructed in this work, namely, directional-projection-based structure representation (DP-SR) and curve representation via a chordlet framework. The DP-SR is constructed based on the multi-scale contourlet pyramid, it captures line-segment information through projection of the edges along different orientations, and summarizes the line-approximated edges in a set of projection profiles through profile modelings. Projection-based methods such as the Hough transform, can efficiently capture line features (Tipwai and Madarasmi, 2002; Franti et al., 2000; Fung et al., 1996). However, the Hough transform treats points on integration lines equally, and thus reduces the spatial information of the shapes. Alternatively, Kadyrov and Petrou (Kadyrov and Petrou, 2001) proposed a trace transform as the generalization of the Hough transform, where selected functionals along the straight lines are utilized and some invariance to translation, rotation and scaling is achieved. In (Abdel-Mottaleb, 2000), histograms of edge segments in four orientations of a block of pixels are used as the shape representation. Tang and Wang (Tang and Wang, 2004) approached the face sketch recognition problem by first transforming face photos into pseudo-sketches, and then comparing the query sketches with pseudo-sketches in a database using an eigenface representation method. In (Sclaroff, 1997), Sclaroff has described a method deformable shape-based image retrieval, where the modal matching technical is used. In DP-SR, through a contourlet transform, edges with different orientations are captured in directional subbands at various scales. Following the projection of each subband into a pair of orthogonal profiles, the shape and spatial information is compactly captured by identifying the primary structures of each 1-D projection profile. In Chapter 4, we first introduce the proposed directional projection based structure representation DP-SR, based on which a piece-wise peak-valley approximation-based model DP-SR-I is developed. Then we describe a non-linear Gaussian mixture approximation-based model DP-SR-II. The advantages and disadvantages of the

proposed models are discussed, and a potential application of DP-SR to image retrieval based on sketches is described with simulations performed. Examples of sketches and their original images are given in Figs. 1-2 (a)-(f).

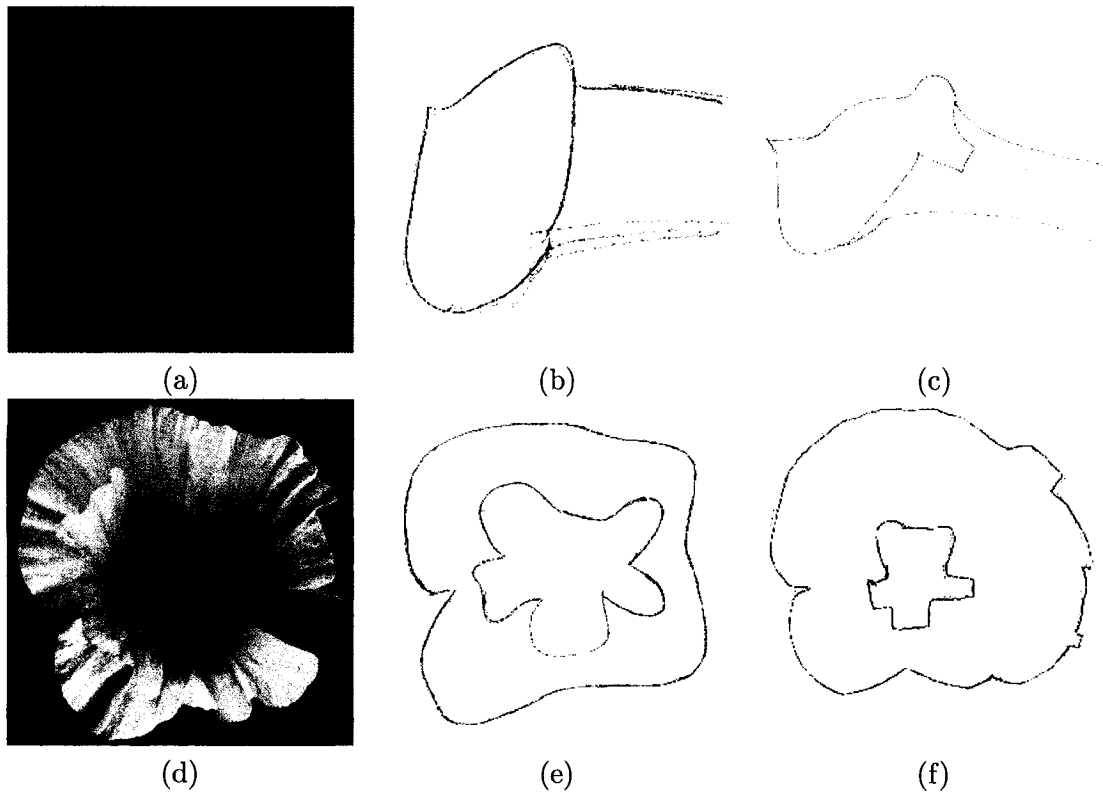


Figure 1-2: An example of sketches and corresponding images. (a) *Flower 1*. (b)-(c) Two sketches of *Flower 1*. (d) *Flower 2*. (e)-(f) Two sketches of *Flower 2*.

Among the multi-scale representations, it has been shown in (Donoho et al., 1998) that wavelets are optimal in capturing point singularities, thus, they provide good representations for data with point discontinuities. To provide a good representations for line segments, which are crucial in 2D image representations, the curvelet (Candes and Donoho, 1999a), contourlet (Do and Vetterli, 2003) and beamlet (Donoho and Huo, 2002) have been constructed. Chapter 2 will give introduction to these transforms, and a brief review of them is given below. In the curvelet (Candes and Donoho, 1999a), line singularities, are

captured through a double transform, i.e., the wavelet transform and ridgelet transform, where the wavelet transform captures the point singularities, whose accumulation, line singularities are captured through a ridgelet transform. The construction of contourlet (Do and Vetterli, 2003) can be viewed as a discrete approximation of the curvelet. A double filter bank, Laplacian pyramid and directional filter banks, framework is adopted. Besides the curvelet and contourlet, the beamlet (Donoho and Huo, 2002) offers a straight forward and flexible framework for line segment representation. In the beamlet framework, a dictionary of line segments at different scales and different locations is defined; the elements in the dictionary are used to represent arbitrary line segments in the image. In a summary, the curvelet, contourlet and beamlet all focus on the line-segment representation as approximation to curves.

In this dissertation, structure representation using curve segments is studied through the construction of the chordlet framework, a hierarchical multi-scale curve representation. The chordlet can be viewed as the curve extension of beamlet. Through a representation efficiency model constructed in Section 5.3, the chordlet-based representation is first shown to be theoretically efficient as compared to beamlet-based representation. A fast chordlet transform algorithm, using the combination of direct and FFT-based chordlet transform, is proposed in Section 5.4 and its performance has been evaluated. For the application to thresholding-based representation, a chordlet coefficient significance model is proposed in Section 5.5 to help the user determine the significance level for the chordlet coefficients in representation system. A simple thresholding-based chordlet representation system is then described based on this significance model. To resolve the in-band and cross-scale correlation between the chordlet elements in chordlet-based representation, two chordlet coding systems are constructed in Section 5.7 using a rate-distortion model. The application of the proposed chordlet coding systems to shape structure representation/compression is studied and compared with beamlet-based coding system. The derivation and discussion of coding efficiency, fast implementation, and coefficient significance in Chapter 5 suggest that the proposed chordlet provides a good curve representation for image analysis. The simulation