THESIS

AN ANALYSIS OF GRATING CELL FEATURES FOR TEXTURE

DISCRIMINATION

Submitted by

John Stevens

Department of Computer Science

COLORADO STATE UNIVERSITY

May 13, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY JOHN STEVENS ENTITLED AN ANALYSIS OF GRATING CELL FEATURES FOR TEXTURE DISCRIMINATION BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

_____
Lucy Troup

_____
J. Ross Beveridge

_____
Advisor: Bruce A. Draper

_____
Department Chair: Darrell Whitley

ii

ABSTRACT OF THESIS

AN ANALYSIS OF GRATING CELL FEATURES FOR TEXTURE DISCRIMINATION

The design of artificial vision systems has been influenced by knowledge of the early stages of processing in the human vision system. The discovery of directionally sensitive cells in the human visual cortex lead to the theory of edge detection in computer vision, and the discovery that simple cell receptive fields can be modeled as Gabor filters has lead to the development and use of Gabor jets.

In this thesis, we evaluate a low-level image feature inspired by "grating" cells found in the human visual cortex. These cells, and the features based on them, detect spatial gratings–repeated patterns of light and dark bars–in their receptive fields. We evaluate the utility of grating cell model features to distinguish different textures using Fisher's linear discriminant. It will be shown that the grating cell features contain significantly more distinguishing information than another standard Gabor-filter-based image feature.

John Stevens
Department of Computer Science
Colorado State University
Fort Collins, CO 80523
Summer 2010

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

The design of artificial vision systems has been directly influenced by our knowledge of the human visual system, and in particular our knowledge of the early stages of visual processing in cortical areas V1 and V2. For example, Hubel and Wiesel's early work with single-cell [HW62] recordings lead to the discovery of directionally sensitive cells in V1, which lead to the theory of edge detection in computer vision [MH80]. More recently, the discovery in the 1980's that simple cell receptive fields can be modeled as Gabor filters [JP87, PGJ89] has lead to the development and use of Gabor jets [FA91].

This thesis aims to evaluate another model derived from single-cell recordings taken from neurons in the visual cortex. Von der Heydt observed cells that respond to grating patterns, but not to single bars or edges [vdHPD92]. Petkov and Kruizinga developed a non-linear computational model that mimics the observed properties of these grating cells [PK97]. This thesis tests the hypothesis that grating cell responses, being non-linear, contribute useful information above and beyond simple linear filters. The problem of texture discrimination is analyzed as an example of problems where grating cell responses might be expected to improve performance.

Since Hubel and Wiesel's cell-recording work [HW62], neuroscientists have performed electrode readings on thousands of cells found in the primary visual cortex (V1) and other areas of the mammalian visual cortex. These experiments are generally conducted as follows: visual stimuli are presented to an animal while electrodes measure the firing rates of individual cells or small groups of neurons in its visual cortex. The goal is to model neurons by finding functions that map the input in the visual field to their firing rates.

One family of cells observed by Hubel and Wiesel, called simple cells, have receptive fields that consist of adjacent excitatory (light in this region causes the cell's firing rate to increase) and inhibitory (light in this region causes the cell's firing rate to decrease) bands. It was later found [Mar80, JP87] that the Gabor function [Gab46] closely modeled the receptive fields of simple cells.

Since this discovery, Gabor filters have been used as feature extractors in computational vision, in such tasks as texture [JF90], object recognition [RB97], and face recognition [SB06]. At a Face Authentication Test in 2004, the top two performing algorithms used features extracted with Gabor filters [MKS$^+$04]. Gabor filters can also be viewed as an example of steerable filters [FA91].

Not all cells in the primary visual cortex are simple cells. Another class of cells, called complex cells, includes a broader range of receptive field behaviors and do not have a single mathematical model. One type of complex cell, first found by Von der Heydt [vdHPD92], responds to gratings, alternating patterns of light and darkness at a particular orientation and spatial frequency, but not to single bars or edges.

Petkov and Kruizinga developed a computational model of these grating cells [PK97]. This model was developed to emulate the behavior of the grating

cells observed by von Der Heydt by responding to grating patterns at particular orientations and spatial frequencies, but not responding to single bars or edges, even if the bars or edges occur at the selected orientation and frequency. This is achieved through a non-linear combination of linear simple cell (Gabor) responses.

Petkov and Kruizinga provide a limited analysis of their grating cell operator in the context of the problem of texture discrimination [KP99]. They concluded that the grating cell operator extracts meaningful information about oriented texture. This thesis expands that work, providing an independent evaluation of the grating cell operator on the texture discrimination problem. Our work differs from that of Petkov and Kruizinga in that we perform the evaluation on larger dataset developed by a third party.

We use the Fisher criterion to evaluate the performance of grating cell model features and another Gabor-based feature–Gabor energy–for discriminating texture. Significant improvement with grating features would support the hypothesis that grating cell responses are useful texture measures. Alternatively, a lack of improvement does not mean that grating cells are useless; they may simply serve another purpose than texture discrimination. Nonetheless, since texture discrimination has been suggested as a major role for grating cells [KP99], a negative result would suggest that the role of grating cells needs to be rethought.

With grating cell model features, pairs of images will be shown to have much higher Fisher criteria scores than with Gabor-energy features. In addition, we show that the grating cell features are a much sparser representation of texture in an image. Research suggests that V1 may employ a sparse coding to represent visual input [OF97].

Chapter 2 reviews the biological literature on Gabor and grating cells, and

3

discusses their computational models. Chapter 3 details the grating cell model under evaluation. Chapter 4 describes the evaluation of grating cell models for texture discrimination. Chapter 5 details the experimental results, and Chapter 6 presents final conclusions and future work.

Chapter 2

Literature Review

2.1 Biology

Hubel and Wiesel [HW62] observed different cell types in the mammalian visual cortex, which they termed simple cells and complex cells. Simple cell receptive fields consist of excitatory and inhibitory regions, with the response of the cell determined by linear spatial summation. Complex cells encompass a broader range of cell behaviors that lack a single model.

2.1.1 Simple Cells

Hubel and Wiesel defined cells whose receptive fields had four certain properties as simple cells. These properties are: 1) distinct excitatory regions, in which light stimuli increase the firing rate of the cell, and inhibitory areas, in which stimuli decrease the firing rate of the cell 2) linear summation within each excitatory or inhibitory area, so that the firing rate of the cell is excited or inhibited linearly with the strength of the stimuli in the corresponding area 3) antagonism between excitatory and inhibitory areas, so that stimuli in inhibitory regions cancels out stimuli in excitatory regions, and 4) the ability to predict the response to a stimulus

5

from a map of the excitatory and inhibitory areas.

Simple cell receptive fields consist of excitatory and inhibitory bands. Some observed receptive fields were bipartite, consisting of two adjacent regions, one excitatory and one inhibitory, and some had three bands, with an excitatory or inhibitory band between two bands of the opposite type.

It was hypothesized [Mar80] that simple cell receptive fields are modeled by Gabor filters [Gab46]. Gabor filters had been used in one-dimensional signal processing because they minimize the joint error in the spatial and spatial-frequency domains. Daugman proved [Dau85] that a two-dimensional Gabor filter minimizes an analogous two-dimensional error, paving the way for the use of Gabor filters in computational vision.

2.1.2 Complex Cells

The class of complex cells represents a variety of cell behaviors and is not described by a single mathematical model, as the class of simple cells is. Complex cells do not generally have the summation and mutual antagonism properties of simple cells.

Hubel and Wiesel observed several different behaviors in the cells they classified as complex cells [HW62]. Some responded to bars of a specific width at a specific orientation located anywhere in the cell's receptive field. Another cell responded to a vertically oriented edge anywhere in its receptive field.

Von der Heydt [vdHPD92] observed cells in V1 that responded only to spatial grating patterns, alternating light and dark bars at a specific orientation. These grating cells respond weakly or not at all to single bars or edges. The finding of cells with such non-linear behavior challenged the idea that V1 performed only

linear spatial-frequency filtering and that non-linear processing only occurs in later stages.

2.2 Petkov and Kruizinga's Grating Cell Model

Petkov and Kruizinga [PK97] developed a computational model of the grating cells that were observed by von der Heydt. This model is designed to mimic the properties of the observed grating cells and respond only to repeated spatial gratings and not respond to single bars, edges, or partial gratings. A spatial grating is a repeated pattern of oriented light and dark bars containing at least three repetitions.

The grating cell model uses features computed by a simple cell model based on Gabor filters, which is described in the next section. The simple cell responses are combined non-linearly to produce the grating cell model outputs.

The following sections discuss two image operators based on Petkov's simple cell and grating cell models. Another Gabor filter based operator used in the literature, the Gabor energy operator, is also discussed. These operators act on images and produce feature maps. A feature map for an operator on an input image consists of the responses of the operator at each pixel in the input image.

2.2.1 The Gabor Model of Simple Cell Receptive Fields

Simple cell responses are used to compute Petkov and Kruizinga's grating cell model. The model of simple cells first convolves an image with a Gabor kernel. The Gabor kernel responses are then normalized for contrast, then a hyperbolic ratio function is applied and negative values are culled.

The Gabor function is:

$$g_{\xi,\eta,\theta,\lambda,\phi}(x,y) = e^{-\dfrac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \cos\left(2\pi\frac{x'}{\lambda} + \phi\right) \tag{2.1}$$

where

$$x' = (x - \xi)\cos\theta - (y - \eta)\sin\theta$$

$$y' = (x - \xi)\sin\theta - (y - \eta)\cos\theta$$

The Gabor function is a Gaussian function convolved with a sinusoidal function. The parameters ξ and η position of the center of the receptive field of the Gabor function in an image, and the parameter θ specifies its orientation. The substitutions x' and y' accomplish this rotation and translation. The parameter γ is the eccentricity of the Gaussian function. γ is fixed at $\frac{1}{2}$ in all Gabor functions used in this thesis. This value has been found to be consistent with observations of the responses of real simple cells [PK97, JP87].

The standard deviation of the Gaussian, σ, is set to .56 λ, and is therefore a dependent parameter to the Gabor function. The constant .56 was chosen based on the spatial-frequency bandwidths of the observed receptive fields of real simple cells, per Petkov and Kruizinga [PK97].

The Gabor function is convolved with an image I:

$$r_{\xi,\eta,\lambda,\theta,\phi} = \int\int_I g_{\xi,\eta,\theta,\lambda,\phi}(x,y)I(x,y)dxdy \tag{2.2}$$

The value $r_{\xi,\eta,\lambda,\theta,\phi}$ is normalized by dividing the response by the average gray level within its receptive field. The result of this normalization is that the response is based on the contrast within the receptive field, not on the intensity. The average Gray level of the image in the operator's receptive field, $\alpha_{\xi,\eta,\lambda}$, is computed using the Gaussian component of $g_{\xi,\eta,\theta,\lambda,\phi}$:

$$\alpha_{\xi,\eta,\lambda} = \int \int_I e^{-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}} I(x,y) dx dy \tag{2.3}$$

where σ is set to .56λ, as in the Gabor function.

In order to obtain a response function similar to those of real cells, a hyperbolic ratio of the contrast-normalized Gabor filter responses is taken. The hyperbolic ratio function is:

$$h(x) = \frac{Rx}{x + C} \tag{2.4}$$

where R is the maximum response (as x goes to ∞) and C is the semi-saturation constant (value of x for which $h(x)$ is $\frac{R}{2}$).

The final response of the simple cell operator is as follows:

$$s = \begin{cases} 0 & \text{if } \alpha = 0, \\ \chi\left(h(\frac{r_{\xi,\eta,\lambda,\theta,\phi}}{\alpha_{\xi,\eta,\lambda}})\right) & \text{otherwise.} \end{cases} \tag{2.5}$$

χ is the Heaviside step function, which is defined as:

$$\chi(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \tag{2.6}$$

In the grating cell model (described in the next section), even symmetric ($\phi = 0$ and $\phi = \pi$) simple cell operators are used. An even symmetric simple cell operator with $\phi = 0$ will respond strongly to a bar that has a width of half of the operator's wavelength. Even symmetric cell operators with $\phi = 0$ are referred to as "on-center", since the central peak of the Gabor function is positive. An operator with a ϕ of π is referred to as "off-center" as its Gabor function has a negative central peak.

2.2.2 Grating Cell Model

The grating operator is designed to respond to oriented spatial gratings. As with the simple cell operator, orientation (θ) and wavelength (λ) are parameters to the grating cell operator. The operator responds to a pattern of three or more bars, and not have any response to two bars or a single bar, or to edges.

In broad terms, a grating operator with wavelength λ works by detecting bars of width $\frac{\lambda}{2}$ along an interval of length 3λ centered at a location. Even symmetric simple cell operators are used to detect bars. Non-linear post processing of the simple cell operator feature maps results in a grating subunit value for each pixel, which takes on a value of 0 or 1. A value of 1 indicates that that pixel is part of a spatial grating of the specified wavelength and orientation. There are two grating subunits for each pixel, one indicating whether the pixel is on a "bar", the other indicating whether it is in the space between bars. The final grating operator response at a pixel location (ξ,η) is a Gaussian weighted summation of the grating subunit values in a neighborhood around the pixel.

To determine the value of a grating subunit, the quantities $M_{\xi,\eta,n}$ and $\mathcal{M}_{\xi,\eta}$ are computed as follows:

$$M_{\xi,\eta,n} = \max\{s_{\xi',\eta',\lambda,\theta,\phi_n}|\eta',\xi' :$$
$$n\frac{\lambda}{2}\sin\theta \leq (\xi'-\xi) < (n+1)\frac{\lambda}{2}\sin\theta,$$
$$n\frac{\lambda}{2}\cos\theta \leq (\eta'-\eta) < (n+1)\frac{\lambda}{2}\cos\theta,$$
$$\phi_n = \begin{cases} 0 & n = -3,-1,1 \\ \pi & n = -2,0,2 \end{cases}\}$$
$$\mathcal{M}_{\xi,\eta} = \max\{M_{\xi,\eta,n}|n = -3\ldots2\}$$

Each grating subunit's value is based on the simple cell responses along a line segment with orientation θ and length 3λ centered at (ξ, η). This line segment is further broken up into six line segments of length $\frac{\lambda}{2}$, indexed by integers n from -3 to 2. The value $M_{\xi,\eta,n}$ for the subsegment n is the maximum simple cell response in an axis-aligned box with corners at the points $(\xi + n\frac{\lambda}{2}\cos\theta, \eta + n\frac{\lambda}{2}\sin\theta)$ and $(\xi + (n+1)\frac{\lambda}{2}\cos\theta, \eta + (n+1)\frac{\lambda}{2}\sin\theta)$. For odd values of n, $\phi_n = 0$, so the value of $M_{\xi,\eta,n}$ is the maximum on-center simple cell response in that box. For even values of n, $\phi_n = \pi$, so the value is the maximum off-center simple cell response. See figure 2.1 for an example of how the boxes are arranged along such a line segment.

The activity of a grating subunit, q, is computed as follows:

$$q_{\xi,\eta,\theta,\lambda} = \begin{cases} 1 & \text{if } \forall n, n \in \{-3\ldots2\} M_{\xi,\eta,\theta,\lambda,n} \geq \rho\mathcal{M}_{\xi,\eta,\theta,\lambda} \\ 0 & \text{if } \exists n, n \in \{-3\ldots2\} M_{\xi,\eta,\theta,\lambda,n} < \rho\mathcal{M}_{\xi,\eta,\theta,\lambda} \end{cases} \quad (2.7)$$

The value of a grating subunit is one if all values $M_{\xi,\eta,\theta,\lambda,n}$ are withing a proportion ρ of the largest of the values, $\mathcal{M}_{\xi,\eta,\theta,\lambda}$. Petkov and Kruizinga use a value of .9 for ρ.

From the subunit activations, the value of the grating operator at a point (ξ, η) is computed as follows:

$$w_{\xi,\eta,\theta,\lambda} = \int e^{\frac{(\xi-\xi')^2-(\eta-\eta')^2}{2(\beta\sigma)^2}} (q_{\xi',\eta',\theta,\lambda} + q_{\xi',\eta',\theta+\pi,\lambda}) d\xi' d\eta' \quad (2.8)$$

The grating response w is the Gaussian weighted sum of subunit responses. The parameter β determines the size of the receptive field of the grating cell operator relative to the size of the receptive field of the simple cell operators on which it is based. The sum of two subunit responses, $q_{\xi',\eta',\theta,\lambda}$ and $q_{\xi',\eta',\theta+\pi,\lambda}$, is taken in equation 2.8. $q_{\xi',\eta',\theta+\pi,\lambda}$ is activated by a grating that is 180° degrees (or π radians) out of phase with the grating that activates $q_{\xi',\eta',\theta,\lambda}$. Such a grating has the same

11
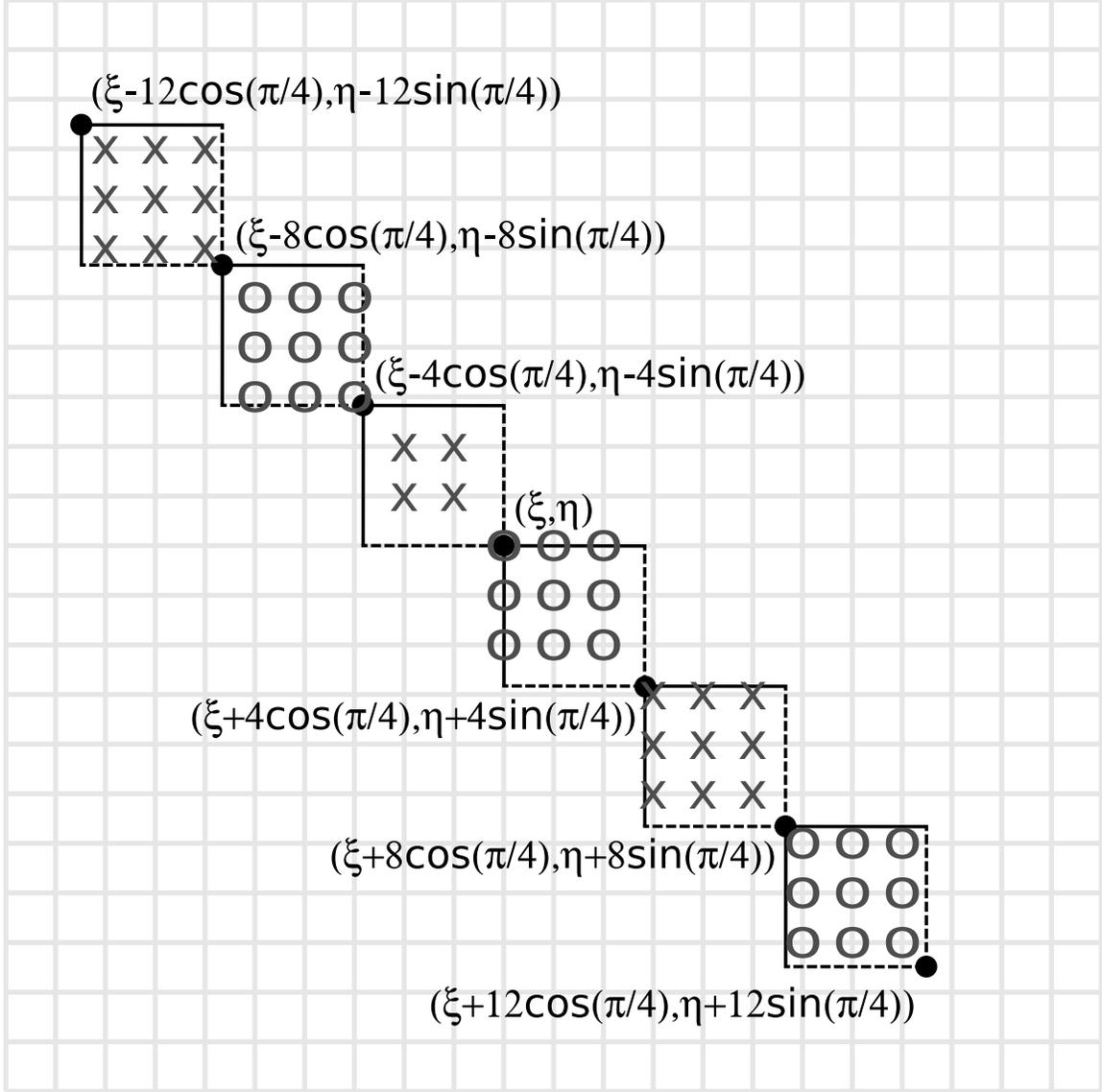
Figure 2.1: Figure showing how the line segment for a grating subunit at location (ξ, η) along which the simple cell responses are used to determine the activation of grating subunit $q_{\xi', \eta', \theta, \lambda}$. $\theta = \frac{\pi}{4}$, $\lambda = 8$ for this grating cell operator. The bounds of the box for each subinterval are marked. The pixels are located at the intersections of the gray lines. An 'O' indicates that the on-center simple cell response at a pixel is use. An 'X' indicates that the off-center response is used. Note that in the box for $n = -1$, there are only four pixels, whereas there are nine pixels in the other boxes.

12

orientation and scale, and image location as the one that activates $q_{\xi',\eta',\theta,\lambda}$— only the intensity of the bars are reversed.

The grating cell operator differs from linear filters, such as Gabor filters, in that it does not give a partial response to partially matching stimuli. For example, an on-center even-symmetric simple cell operator has three excitatory bands. The response of the simple cell operator will be highest if the image intensity is high in those excitatory bands and low in its inhibitory bands, but there will still be a non-zero response if the intensity is high in only one of the excitatory bands. The grating cell operator's response is zero unless there are three or more bars present.

2.2.3   Gabor Energy Operator

Gabor-based features known as Gabor energy features [KP99] or phase-invariant Gabor features [FS89] have been used in texture discrimination. The value of a Gabor energy feature at a pixel is:

$$e_{\theta,\lambda,\xi,\eta} = \sqrt{r^2_{\xi,\eta,\lambda,\theta,0} + r^2_{\xi,\eta,\lambda,\theta,\frac{\pi}{2}}} \tag{2.9}$$

where $r_{\xi,\eta,\lambda,\theta,\phi}$ is the result of the convolution of a Gabor function with the image, from equation 2.2. The value of the Gabor energy feature at a pixel is computed by taking the square root of the sum of the squares of two Gabor responses with phases 90° apart. As long as they are 90° apart, the value of the operator is invariant to the particular phases used. We use $\theta = 0^o$ and $\theta = 90^o$ for all Gabor energy operators in this thesis.

2.3 Uses of the Grating Cell Model

Kruizinga and Petkov [KP99] evaluate their grating operator in the context of texture discrimination and segmentation. They compare features derived from the grating cell operator with Gabor energy and co-occurence matrix [Pec91] features. Fisher linear discriminant analysis is used to analyze the separability of nine different textures using these three sets of features. Grating features achieved the highest distance between different textures, and different textures were always linearly separable using grating features. In addition, Kruizinga and Petkov show results of a K-means segmentation algorithm on mosaics of these textures using each set of pixel-wise features. The qualitatively best segmentation was achieved using the grating features.

Weinman et. al. make use of a modification of Petkov's grating cell model for detecting signs in natural images [WHM04]. This was chosen because sign text consists of a series of mostly vertically oriented strokes that resemble a spatial grating.

2.4 Other Grating Cell Models

Alternative models of grating cells exist. One such model is due to Lourens et. al. [LBOT05], and another due to Du Buf [dB07]. Both models are similar to Petkov's model in that they post-process pixel-wise Gabor features.

Lourens et. al. [LBOT05] created one such model. Like Petkov's model, Lourens' grating operator is based on a Gabor-based operator's response along a straight line. The length of the line along which the Gabor operator responses are examined varies based on those Gabor operator responses, unlike Petkov's op-

14

erator. The Gabor responses along are combined nonlinearly, not just with the maximum operator, but with the difference between the average and the minimum and maximum along the interval. The authors claim that this leads to a better match with the response profiles of actual grating cells than does Petkov's operator.

Du Buf's grating cell operator [dB07], in contrast to Petkov's operator, is aimed at providing precise boundary localization between regions that resemble spatial gratings and regions that do not. Du Buf's grating cell operator labels each pixel in an image as either "grating" or "non-grating".

Chapter 3

Grating Cell Model

This chapter gives a more detailed explanation of the grating cell model. Implementation issues are discussed first, followed by a qualitative analysis of the behavior of the simple and complex cell models. The Gabor energy operator against which the grating cell is compared in the evaluation is also discussed.

3.1   Implementation Details

The integral in equation 2.2 is implemented with the following summation:

$$r_{\xi,\eta,\lambda,\theta,\phi} = \sum_{i=\xi-\frac{s}{2}}^{\xi+\frac{s}{2}} \sum_{j=\eta-\frac{s}{2}}^{\eta+\frac{s}{2}} g_{\xi,\eta,\lambda,\theta,\phi}(i,j)I(i,j) \tag{3.1}$$

where $I(x,y)$ is the image intensity at the point (x,y). The mask size, s is chosen to be the smallest odd integer greater than or equal to 5σ. When ξ or η is close enough to the edge that the limits of the summation are out of the boundaries of the image, the indices are reflected, e.g. $I(-3,2) = I(3,2)$.

For each chosen set of parameters (orientation and wavelength) for a simple cell operator, a single Gabor kernel is created and convolved with the image, generating a feature map of raw Gabor responses to which contrast normalization, the

hyperbolic ratio function, and the Heaviside step function are applied. Thus, the location of the simple cell operator's receptive field is not truly a parameter, since the response at every pixel in the image is computed.

3.1.1 Contrast Normalization

Equation 2.3 is the calculation of the contrast normalization term as described in [PK97]. As described, the contrast normalization is performed with an unrotated Gaussian while the Gaussian used in calculating the Gabor kernel is rotated. This leads to artifacts in the feature maps. For the implementation used in this evaluation, the same oriented Gaussian as used in the Gabor function is used in contrast normalization:

$$\alpha_{\xi,\eta,\lambda,\theta,\gamma} = \int \int_I e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} I(x,y)dxdy \tag{3.2}$$

where

$x' = (x - \xi)\cos\theta - (y - \eta)\sin\theta$

$y' = (x - \xi)\sin\theta - (y - \eta)\cos\theta$

This integral is implemented as a convolution with a mask size of 5σ, as in equation 3.1.

3.1.2 Hyperbolic Ratio Parameters

The values of the semi-saturation constant (C) and maximum response (R) of the hyperbolic ratio function (equation 2.4) used in the simple cell operator are not defined in [PK97].

A hyperbolic ratio function like the one described in equation 2.4 has a horizontal asymptote at $h(x) = R$ and a vertical asymptote at $x = -C$. The value of
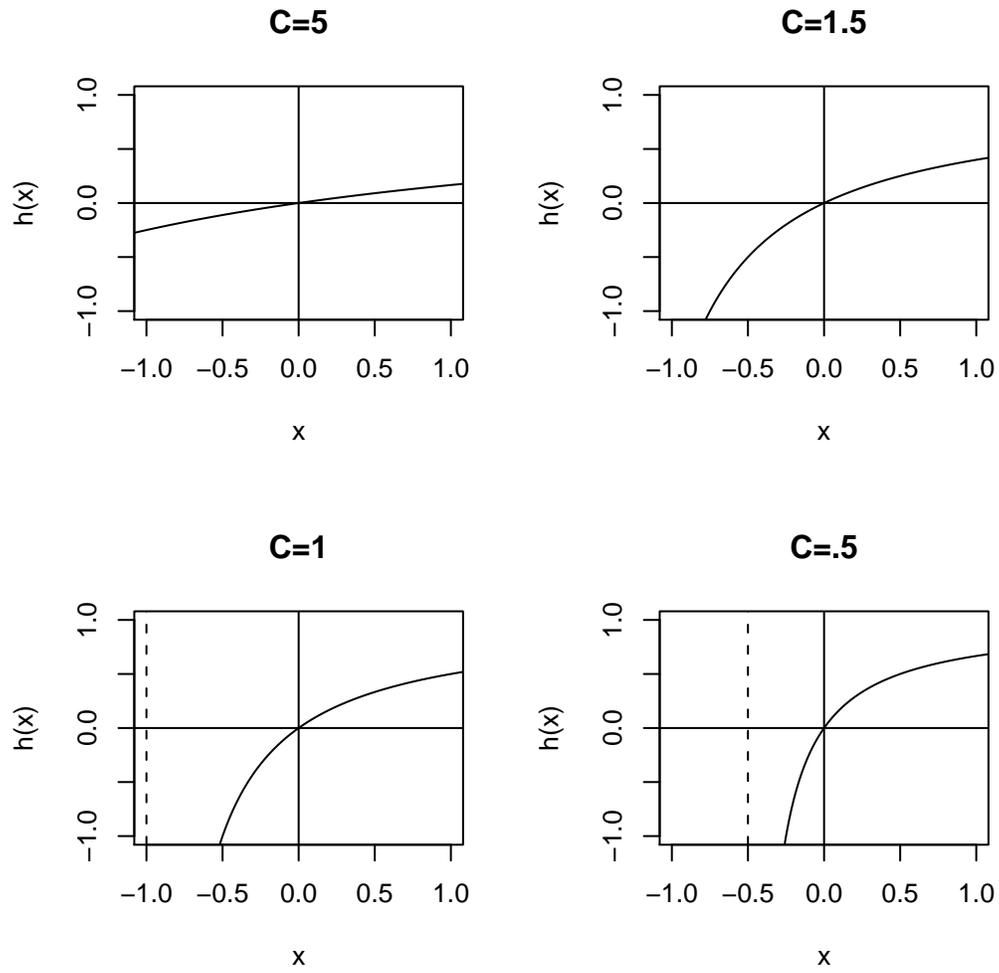
Figure 3.1: Plots of hyperbolic ratio h(x) with different C values, R=1.

$h(x)$ approaches $-\infty$ as x goes to ∞, and approaches $-\infty$ as x goes to $-C$ from the right. Figure 3.1 shows hyperbolic ratio functions with different values of C.

For x values greater than $-C$, the sign of $h(x)$ is the same as the sign of x. When C is less than or equal to one, the vertical asymptote is inside the range $[-1, 1]$ of values produced after the contrast normalization step to which the hyperbolic ratio function is applied. For x values slightly less than -C, h(x) is a large positive value. If C is less than one, negative values in the feature map after the contrast normalization step that would otherwise be set to zero by the Heaviside step function instead become large positive values. This produces artifacts in the feature maps generated by the simple cell operator. For this reason, a value of 1.5 is chosen for C in all simple cell operators used in this evaluation.

As the first step in the grating operator is to produce feature maps using the simple cell operator, the choice of the hyperbolic ratio parameters R and C needs to be evaluated in the context of the grating cell operator.

In the grating cell operator, the simple cell responses are used to determine grating subunit activations. A subunit at a location activates if all of the maximum simple cell responses within intervals along a line segment centered at that location are within a percentage of the maximum response along that line segment. The R parameter scales all responses linearly and equally, so it does not effect the subunit activations and thus does not effect the output of the grating operator.

Varying C in a range from 1.5 to 5 did not affect the output of the grating cell operator. A value of 1.5 is used for C in all grating operators used in this evaluation.
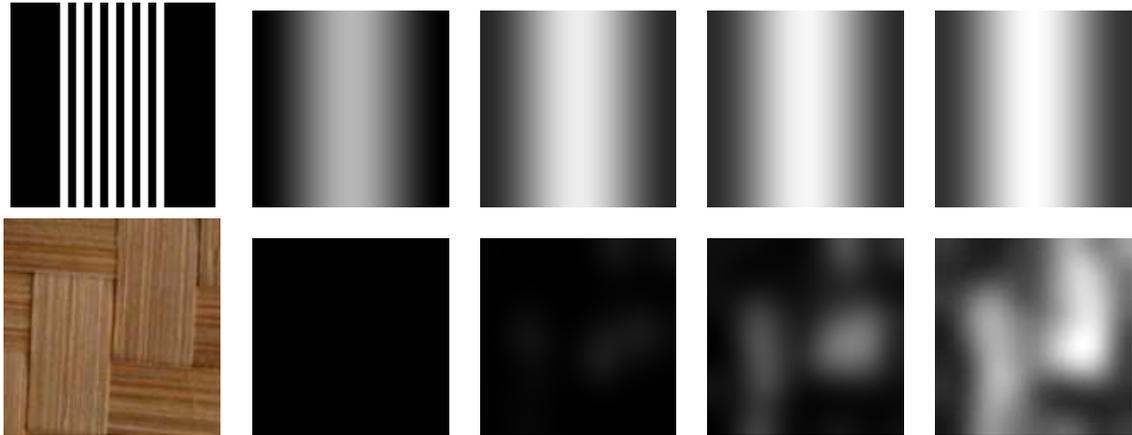
Figure 3.2: Grating operator feature maps on a synthetic grating (first row) and a natural image from the Prague texture segmentation benchmark (second row). First column: input image. Second column: $\rho = .9$. Third column: $\rho = .6$. Fourth column: $\rho = .4$. Fifth column: $\rho = .2$.

3.1.3 Choice of ρ in Grating Cell Model

Petkov use a constant .9 for ρ. The parameter ρ is the threshold for grating subunit activation. In order for a grating subunit to activate, the maximum simple cell responses from each subinterval of a grating cell's receptive field must be within a factor of ρ of the maximum response. Lower values make it easier for grating subunits to activate and result in stronger grating cell operator responses.

Grating operators with a ρ value of .9 responded well to synthetic images of gratings, such as the one in Figures 3.5 and 3.6. On natural images, however, a value of .9 was found to be too strict. Figure 3.2 for an example. The grating operator detects the synthetic grating for all four values of ρ, but does not detect the grating in the natural image for $\rho = .9$ or $\rho = .6$. The grating shown in the natural image in figure 3.2 is one of the most pronounced spatial gratings in any image in the dataset. Failing to detect such a grating would make for an

uninteresting comparison. For this reason, we use a value of .2 instead of .9 for ρ.

3.2 Qualitative Analysis

Figure 3.3 shows the stages of simple cell operators with different orientations on a synthetic image of a white bar on a black background. The bar is five pixels wide, and the wavelength parameter of all operators in the figure is 10 pixels. In this case, the width of the central peak of the Gabor mask used by the simple cell operator is five pixels, the same as the width of the bar in the input image. When the operator's orientation matches the bar's orientation, there is a peak in the simple cell response at the center of the bar. The response drops off as the center of the simple operator moves away from the center of the bar.

The response drops off as the orientation of the operator changes away from the orientation of the bar. The response is smaller at $\theta = 22.5^o$, and smaller still at $\theta = 45^o$. There is virtually no response at $\theta = 90^o$, at which point the bar is perpendicular to the excitatory and inhibitory bands of the simple cell operator's receptive field.

Figure 3.4 shows the stages of the simple cell operator as the wavelength is varied on the same synthetic bar image. The response at the center of the bar drops off as the wavelength decreases. Note that, when the wavelength is 10 pixels, there are three peaks of equal magnitude, one on the center of the bar, and two others to the left and right. These peaks in the feature map occur when the smaller magnitude excitatory bands of the Gabor mask align with the bar. The raw response is much lower than when the central peak overlaps the bar, but average Gaussian-weighted image intensity is also lower, resulting in a peak similar in magnitude to the central peak after contrast normalization.

21
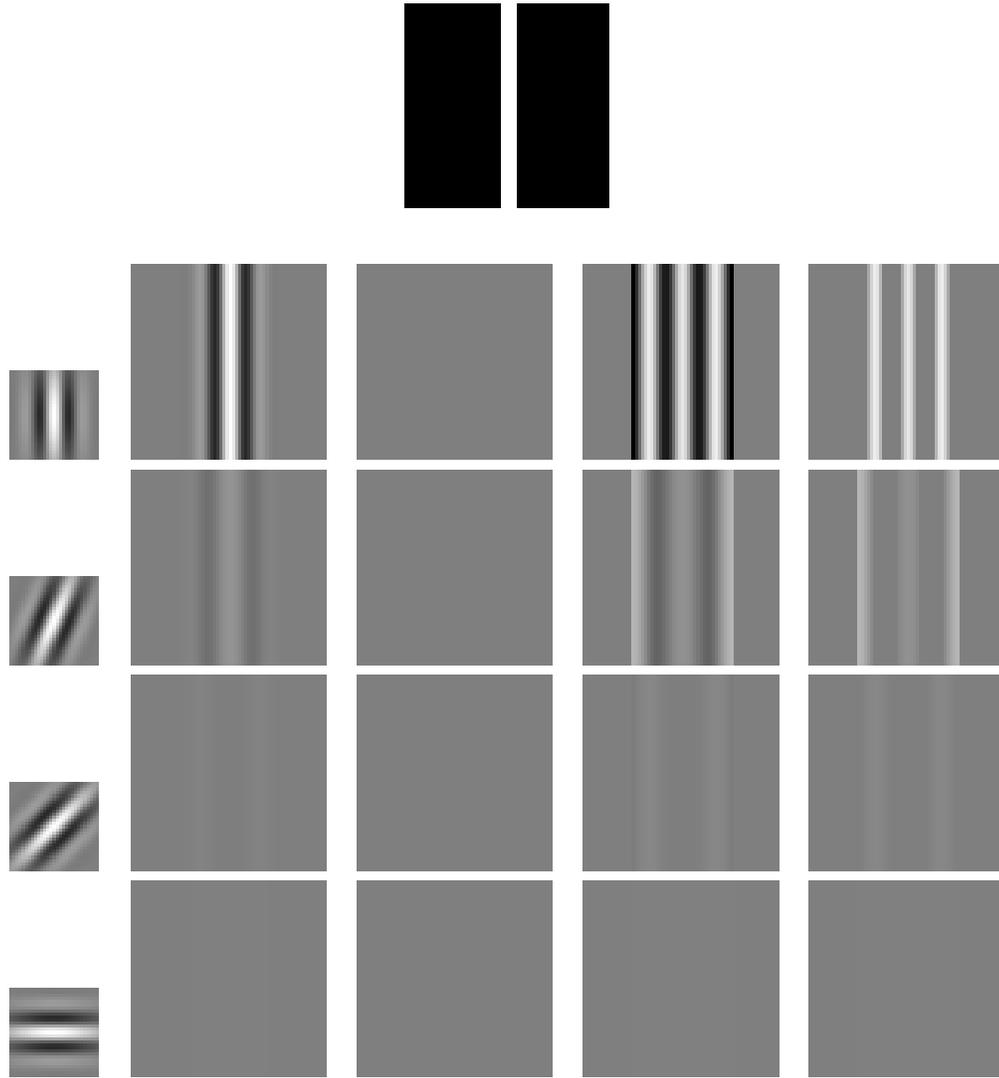
Figure 3.3: Stages of simple cell operators with different orientation parameters applied to a synthetic bar image. Top: Input image, a synthetic image of a five pixel wide white bar on a black background. First column: Gabor mask. Second column: convolution of Gabor mask with input image. Third column: contrast normalization of image from column 2. Fourth column: application of hyperbolic ratio function to image from column 3 (C=1.5). Fifth column: application of Heaviside step function to image from column 4. Row 1: $\theta = 0^o$, Row 2: $\theta = 22.5^o$, Row 3: $\theta = 45^o$, Row 4: $\theta = 90^o$. $\lambda = 10$ for all operators.
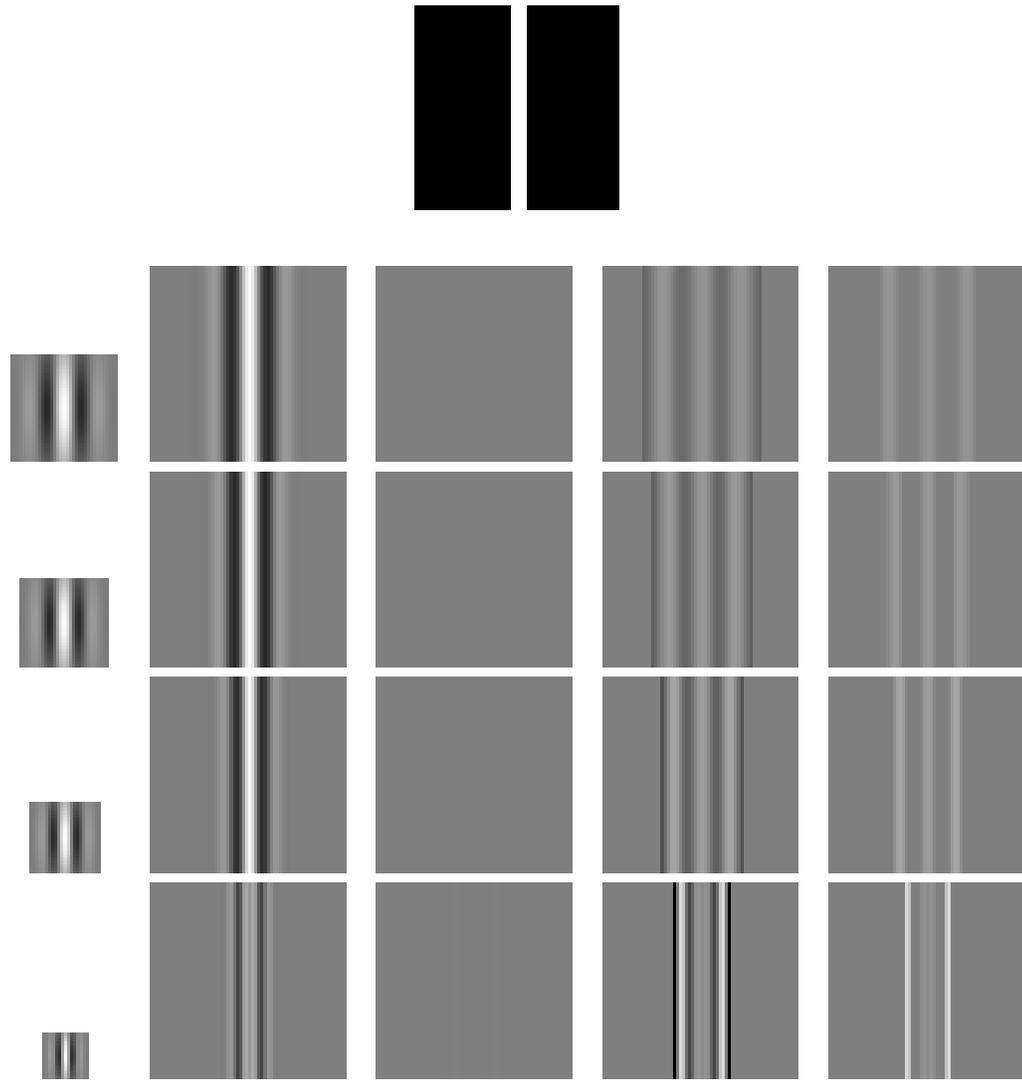
Figure 3.4: Stages of simple cell operators with different wavelength parameters applied to a synthetic bar image. Top: Input image, a synthetic image of a five pixel wide white bar on a black background. First column: Gabor mask. Second column: convolution of Gabor mask with input image. Third column: contrast normalization of image from column 2. Fourth column: application of hyperbolic ratio function to image from column 3 (C=1.5). Fifth column: application of Heaviside step function to image from column 4. Row 1: $\lambda = 12$, Row 2: $\lambda = 10$, Row 3: $\lambda = 8$, Row 4: $\lambda = 5$. $\theta = 0^o$ for all operators.
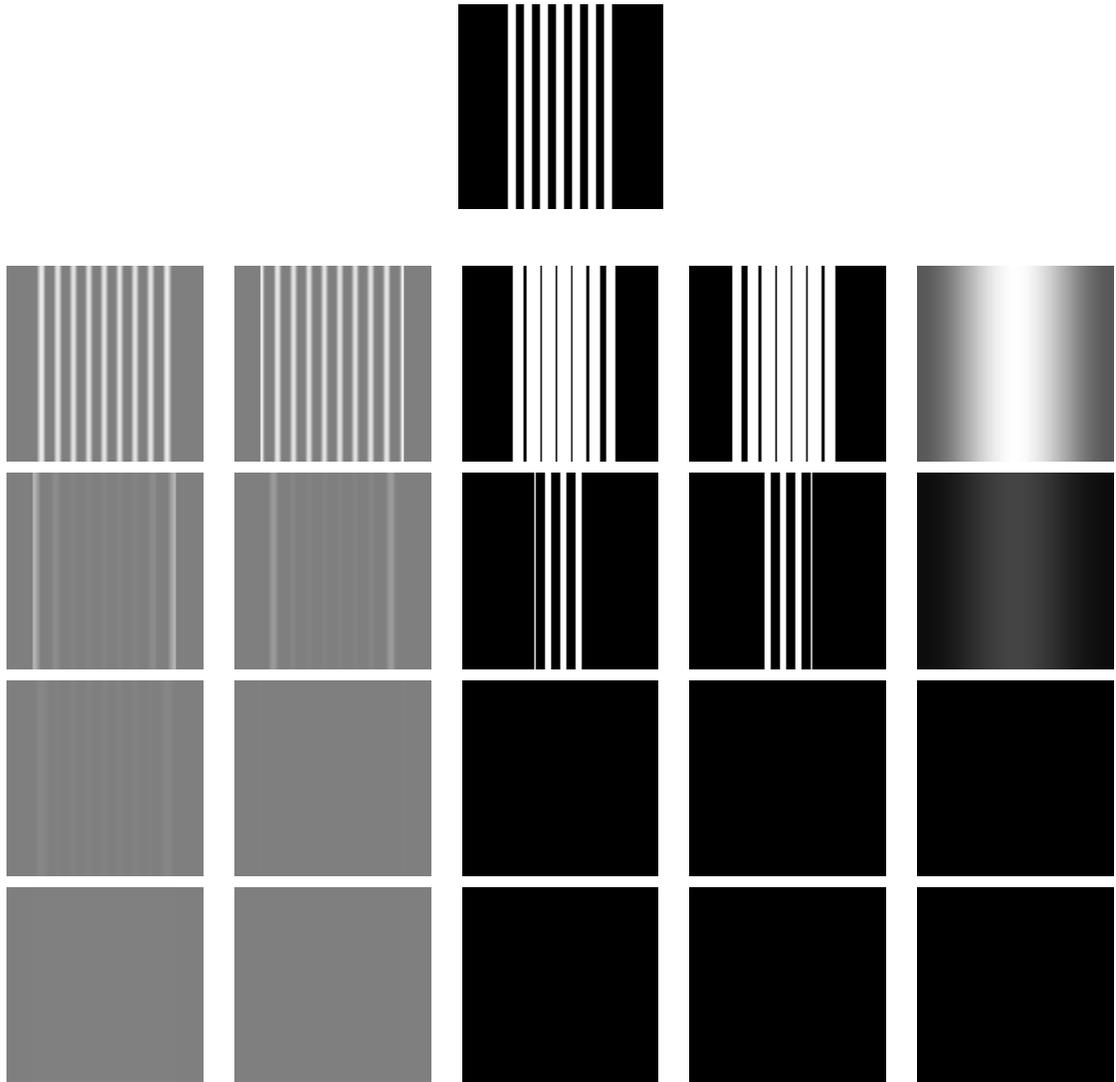
Figure 3.5: Shows stages of grating operators with different orientation parameters on a synthetic grating image. Top: input image, a grating of white five pixel wide bars on a black background. First column (from the left): on-center simple cell responses. Second column: off-center simple cell responses. Third column: $\theta = 0$ subunits. Fourth column: $\theta = \pi$ subunits. Fifth column: grating operator response. Row 1: $\theta = 0^o$, Row 2: $\theta = 22.5^o$, Row 3: $\theta = 45^o$, Row 4: $\theta = 90^o$. $\lambda = 10$ and $\rho = .6$ for all operators.

Figure 3.5 shows the stages of grating operators with different orientations on a synthetic image of a grating of five-pixel-wide white bars on a black background. This grating has a wavelength of ten pixels. The first two columns respectively show the $\phi = 0^o$ and $\phi = 180^o$ simple cell responses used in computing the grating cell response. The $\phi = 0^o$ simple cell response is highest on the white bars, and the $\phi = 180^o$ simple cell response is highest on the black bars. The third and fourth columns of the figure show the grating subunits. The fifth column shows the grating response feature map.

The response of the grating operator is maximized when the orientation of the grating in the images matches the orientation of the grating operator. The response drops off the farther the orientation of the operator is away from the orientation of the image grating. At $\theta = 22.5^o$, there is a much lower response from the grating operator. At $\theta = 45^o$ and $\theta = 90^o$, there is no response from the grating operator.

Figure 3.6 shows the stages of grating operators with different wavelengths on the same synthetic grating image. The first two columns respectively show the $\phi = 0^o$ and $\phi = 180^o$ simple cell responses used in computing the grating cell response. The grating operator response is highest when the wavelength parameter of the operator matches the wavelength of the the image grating.

3.3   Gabor energy

To provide a baseline for comparison, we use Gabor energy features in segmentation.

Figure 3.7 shows, for three synthetic images, the Gabor energy feature maps generated by Gabor energy operators with different orientation parameters.

Figure 3.8 shows, for three synthetic images, the Gabor energy feature maps
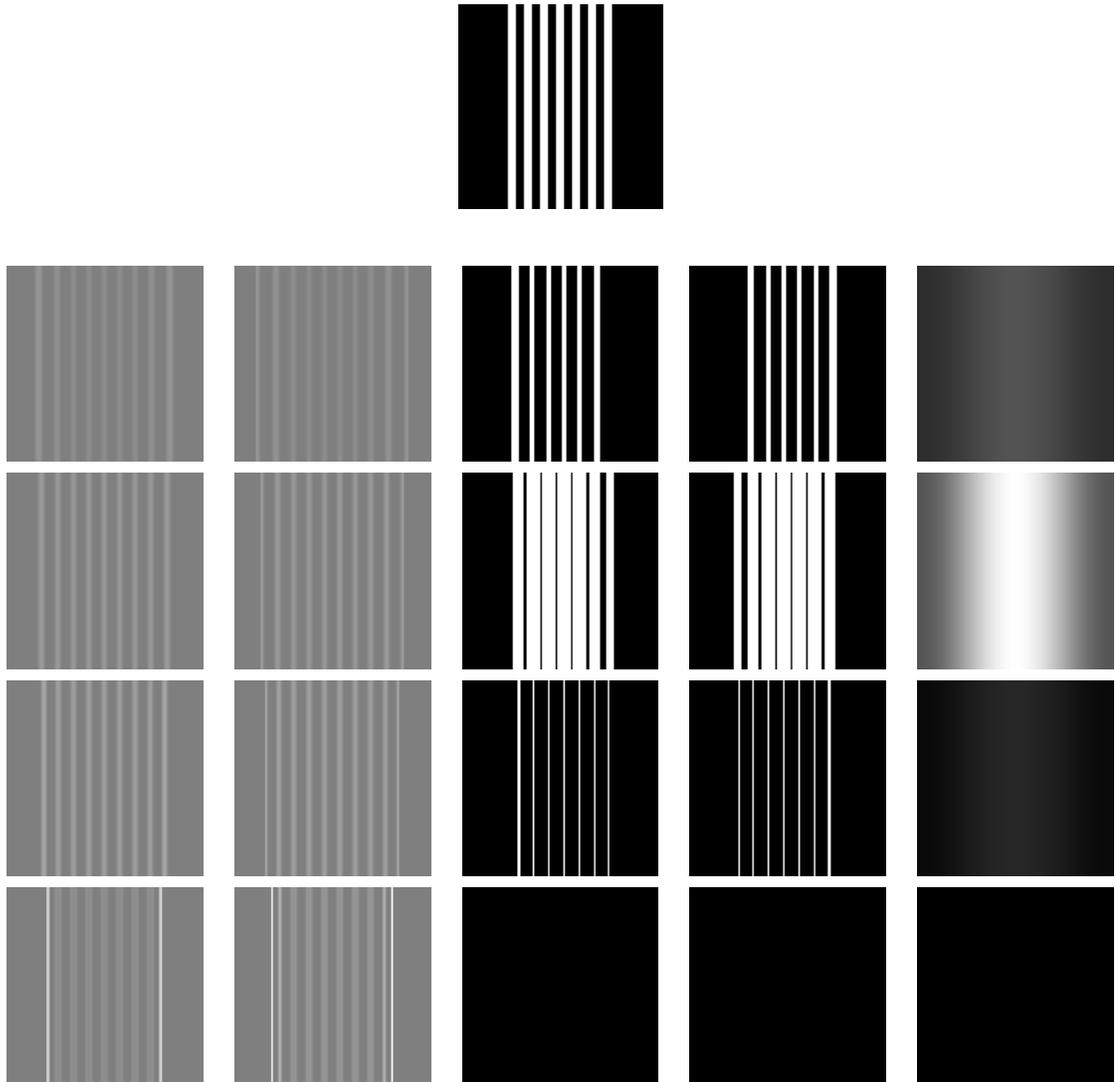
Figure 3.6: Shows stages of grating operators with different wavelength parameters on a synthetic grating image. Top: input image, a grating of white five pixel wide bars on a black background. First column (from the left): on-center simple cell responses. Second column: off-center simple cell responses. Third column: $\phi = 0$ subunits. Fourth column: $\phi = \pi$ subunits. Fifth column: grating operator response. Row 1: $\lambda = 12$, Row 2: $\lambda = 10$, Row 3: $\lambda = 8$, Row 4: $\lambda = 5$. $\theta = 0^o$ and $\rho = .6$ for all operators.

Symbol	Parameter
λ	Wavelength
θ	Orientation
ϕ	Phase
γ	Eccentricity
σ	Standard deviation
C	Semi-saturation constant of hyperbolic ratio function
R	Maximum value of hyperbolic ratio function
ρ	Threshold for grating subunit activation

Table 3.1: Table of parameter symbols.

Parameter	Petkov and Kruizinga	This Thesis
λ	Free parameter	
θ	Free parameter	
ϕ	Not a parameter, $\phi = 0$ and $\phi = \pi$ used in grating model	
γ	Fixed at .5	
σ	Set to .56λ	
C		Fixed at 1.5
R		Fixed at 1
ρ	Fixed at .9	Fixed at .2

Table 3.2: Table of parameters in both models.
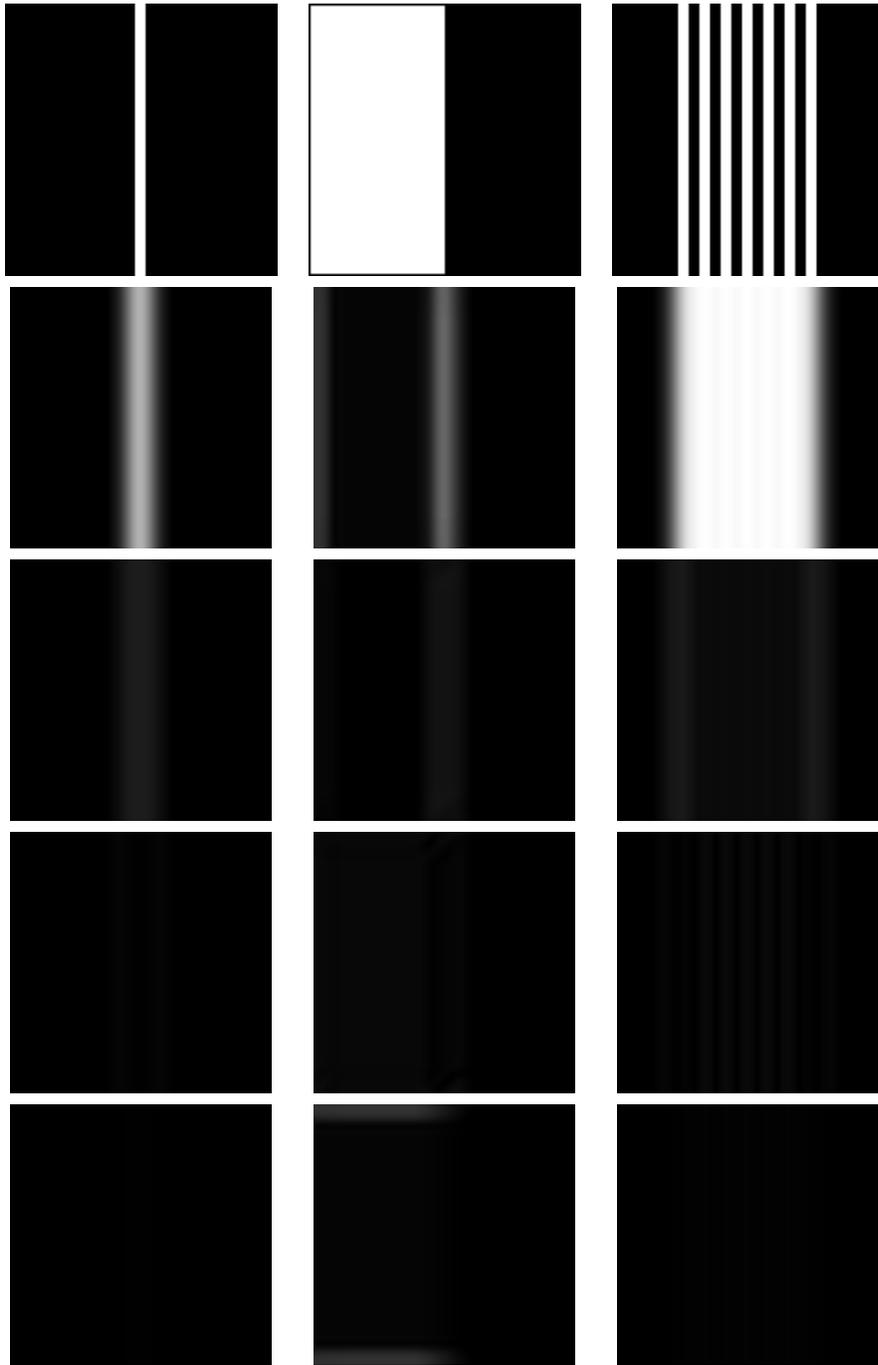
Figure 3.7: Gabor energy feature maps for different synthetic images and different orientations. Top row: synthetic bar, edge, and grating images. Second row: $\theta = 0^o$. Third row: $\theta = 22.5^o$. Fourth row: $\theta = 45^o$. Fifth row: $\theta = 90^o$. $\lambda = 10$ for all operators.
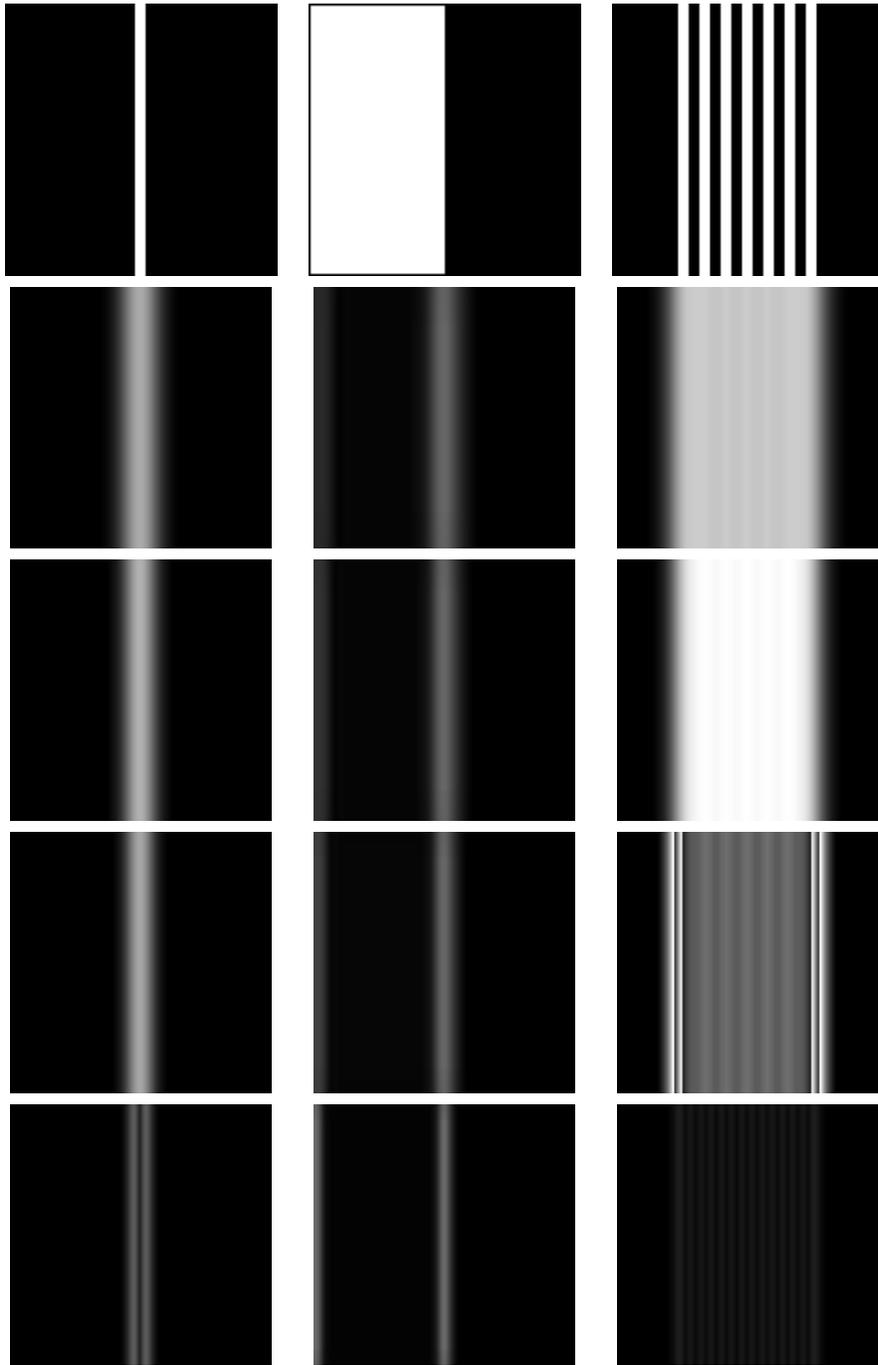
Figure 3.8: Gabor energy feature maps for different synthetic images and different wavelengths. Top row: synthetic bar, edge, and grating images. Second row: $\lambda = 12$. Third row: $\lambda = 10^o$. Fourth row: $\lambda = 8$. Fifth row: $\lambda = 5$. $\theta = 0^o$ for all operators.

generate by Gabor energy operators with different wavelength parameters.

3.4 Choice of θ Values

Figure 3.9 shows the response of the Gabor energy operator plotted against the orientation of the operator for individual pixels within three synthetic images. An orientation of zero degrees for the operator matches the orientation of the stimuli in the images and maximizes the operator response. The response drops off as the orientation changes away from zero degrees, becoming close to zero when the operator's orientation is 22.5 degrees off from the stimuli in the image. We therefore choose θ values in 22.5 degree increments from zero degrees up to 167.5 degrees for the Gabor energy operators generating the features used in the evaluation.

The same orientations are used in the grating operators. This is consistent with Petkov and Kruizinga's own use of their grating cell model [PK97, KP99].

3.5 Comparison

Figure 3.10 shows a side-by-side comparison of the Gabor bar, Gabor energy, and grating operators on four synthetic images: a single bar, an edge, a "near grating" consisting of two bars, and a true grating.

The Gabor bar operator has a high response to the bars in the bar, near-grating, and true grating images. Of note is that there is also a band with a strong response in the dark region of the edge image. This is another haloing artifact caused when the secondary excitatory band of the Gabor filter falls on the edge of the light region. The filter response is small, but so is the average gray level in the receptive field, so the response after contrast normalization is high.
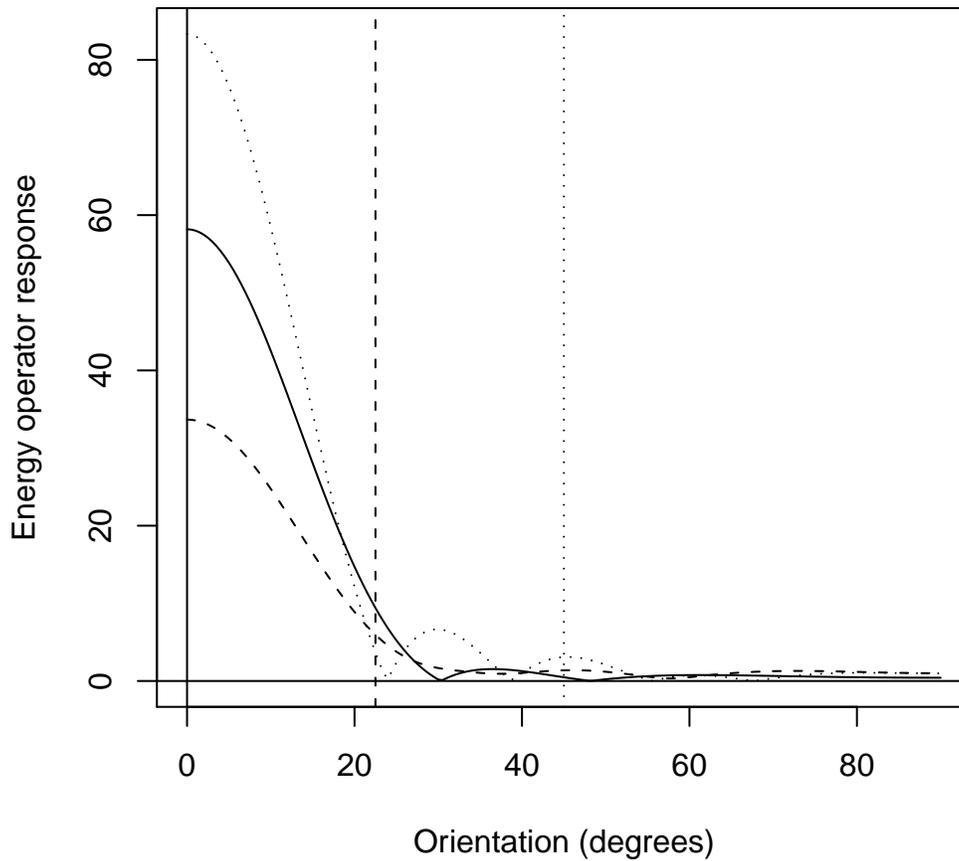
Figure 3.9: Response of the Gabor energy operator plotted against the θ parameter for three synthetic images. Solid line: synthetic bar image (sampled from the center pixel of the bar). Dashed line: synthetic edge image (sampled from on the edge). Dotted line: synthetic grating image (sampled from the center pixel of the central bar of the grating). The vertical dashed line is at 22.5^o, the vertical dotted line is at 45°.
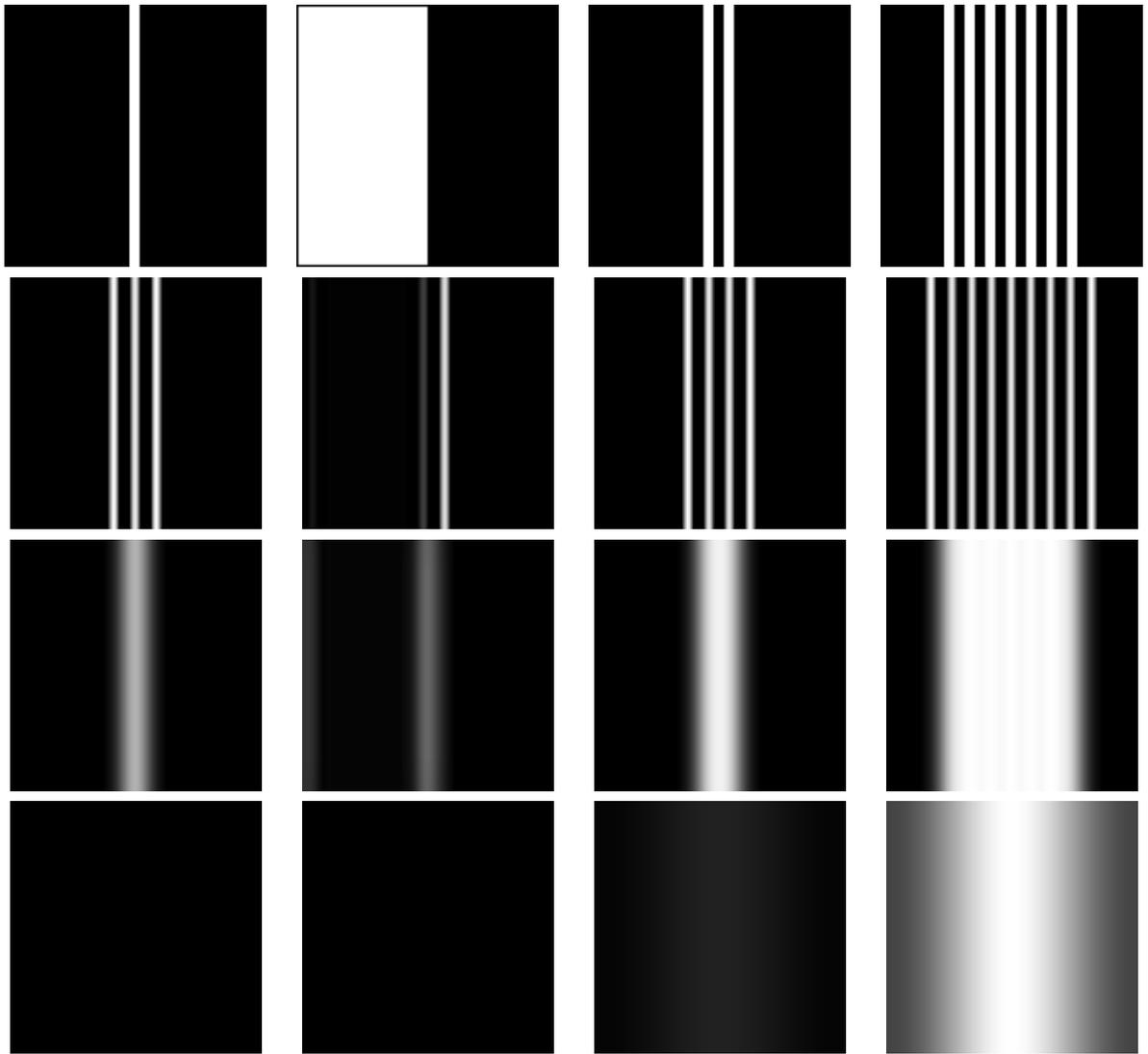
Figure 3.10: Comparison of different operators on different input images. Top row: input image. Second row: simple cell operator. Third row: Gabor energy operator. Fourth row: grating cell operator.

The Gabor energy operator responds strongly to all four stimuli. The grating operator has no response to the single bar and edge images, and only a faint response to the near-grating image. That the grating operator has any response at all to the near grating is a result of the haloing in the Gabor operator. As seen in Figure 3.10, there are four bars with a high response in the on-center Gabor operator feature map. An off-center Gabor operator would have three bars of high response, one between the two bars and one to either side. Since the grating operator operates on Gabor operator feature maps and not on the image directly, there are three or more bars and the operator has a response. The response is small compared to the response to a true grating, and so the grating operator can still be said to only respond strongly to gratings of three or more bars.

3.6   Image Pyramids and Scale Spaces

We want our image features to detect texture at different image scales. To accomplish this, we use a Gaussian image pyramid. An image pyramid is a collection of representations of an image at different scales [FP02].

Each level in an image pyramid is a smoothed and subsampled image of the previous level. The lowest level of the pyramid consists of the original image convolved with a Gaussian mask with $\sigma = 1$. Each subsequent level is generated by convolving the previous level with a $\sigma = \sqrt{3}$ Gaussian mask and downsampling by selecting every other pixel. The lower levels of the pyramid contain larger images and finer scales, while the higher levels contain smaller images and coarser scales.

Features are generated by running the same operators on all levels of the image pyramid, generating a feature map for each level. The feature maps for the

coarser scales are upsampled back to the size of the original image using bilinear interpolation. An image pyramid with six levels is computed for each image used in the experiments.

Chapter 4

Methods

The goal of this thesis is to experimentally determine whether grating cell responses improve texture discrimination. Our method is to evaluate the results of a standard texture discrimination algorithm on Gabor responses and grating cell responses.

The Prague texture segmentation benchmark [HM08] is a method and system for evaluating texture segmentation. The method creates synthetic texture mosaics with known ground-truth segmentations. Texture segmentation algorithms can be applied to these images, and the output compared to the ground truth. The system includes a web interface for specifying dataset parameters, downloading datasets, and uploading segmented images for evaluation.

Although the Czech system was intended to evaluate segmentation algorithms, we will use it for another purpose. Our goal is to evaluate the utility of grating cell responses in texture discrimination. To accomplish this, we use textures from the Prague benchmark, and use Fisher linear discriminant analysis to assess the separability of these textures using grating and Gabor energy features.

4.1 The Prague Texture Segmentation Benchmark

The Prague system generates two sets of images: textures and mosaics. The textures are natural images. Each texture is a 512 by 512 color image. There are 10 categories of textures, described in more detail in the next section.

The textures in the benchmark dataset are images of natural and man-made texture. The images are real images, as opposed to artificially synthesized textures, were chosen by the authors of the benchmark because they are harder to segment. This dataset provides a challenge to the texture discrimination task evaluated in this thesis. This evaluation does not aim for state-of-the-art performance on the Prague benchmark, but to measure the effect that grating cell responses have on performance.

The Prague benchmark is used via a web form found at http://mosaic.utia.cas.cz. This web form provides an interface for generating and downloading datasets and uploading and evaluating segmentations.

4.2 Dataset

The dataset used for the evaluation was obtained from the Prague texture segmentation benchmark system. The dataset consists of 108 textures in the following categories:

- bark (12 textures)

- flowers (10 textures)

- glass (7 textures)

- man-made (20 textures)

- nature (10 textures)

- plants (10 textures)

- rock (10 textures)

- stone (10 textures)

- textile (10 textures)

- wood (9 textures)

The dataset was generated from the Prague texture segmentation benchmark website using the following parameters:

- "Colour" button selected for a Colour benchmark.

- "dataset size" set to "large (4x)"

- "supervised" checkbox checked

- "noise type" set to "no"

4.3   Fisher Linear Discriminant Analysis

We wish to analyze the effectiveness of grating features in texture discrimination. We pose the problem of discriminating two textures as a classification problem, with the pixels from the two texture images as the samples to be classified. Pixels in the texture images are mapped to a high dimensional space using Gabor energy or grating operators, with each dimension corresponding to the response of one

image operator at that pixel. We use a linear classifier based on Fisher's linear discriminant. This description and implementation is based on [B+06].

Discriminant analysis for two-class classification works by projecting each data point to a single dimension. Linear discriminant analysis uses a linear projection, defined by a vector w. An input vector x is projected to one dimension by the following equation:

$$y = w^T x \qquad (4.1)$$

A threshold is then applied to classify the projected values. Fisher's linear discriminant is a way of finding the linear projection that maximizes the separation of the two classes. The input to the Fisher linear discriminant classifier is N_1 samples of class C_1, and N_2 samples of class C_2.

The Fisher criterion is defined as the ratio of the variance between the two classes to the variance within each class, and is given by the following equation:

$$J(w) = \frac{(m_2 - m_1)^2}{s_1^2 + s_2^2} \qquad (4.2)$$

where m_1 and m_2 are the means of the projected values of class 1 and class 2, and s_1^2 and s_2^2 are the variances. This equation can be rewritten in terms of matrix and vector operations like so:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \qquad (4.3)$$

where S_B, the between-class covariance matrix, is:

$$S_B = (m_2 - m_1)(m_2 - m_1)^T \qquad (4.4)$$

38

and S_W, the total within-class covariance matrix, is:

$$S_W = \frac{1}{N_1} \sum_{n \in C_1} (x_n - m_1)(x_n - m_1)^T + \frac{1}{N_2} \sum_{n \in C_2} (x_n - m_2)(x_n - m_2)^T \qquad (4.5)$$

The w vector that maximizes this criterion is a unit vector in the direction of:

$$S_W^{-1}(m_2 - m_1). \qquad (4.6)$$

The Fisher criterion increases as the projected points in different classes move farther away from each other, and the projected points for points within the same class move closer together. Higher values of the Fisher criterion indicate better class separability. If the projected values of the two classes do not overlap, the classes are said to be linearly separable.

In the experiments in this thesis, 10% of the pixels from a texture image are selected randomly as the samples on which to perform classification, and samples from two images are used as input samples to the classifier. The input samples from each class are randomly partitioned into a training set, comprising 80% of the samples, and a test set comprising the other 20%. The weight vector w and the threshold are calculated using the training set. We report the Fisher criterion for each comparison in Chapter 5. We also use receiver operating characteristic curves to evaluate the separability of the classes in each comparison.

4.4 Classifier Evaluation Using Receiver Operating Characteristic Curves

A classifier can be constructed from the projected Fisher values by applying a threshold to them, anc classifying all samples with a projected value above the

threshold as one class, and all samples with a projected value below the threshold as the other class. A receiver operating characteristic curve is a non-parametric way of analyzing the performance of such threshold based classifiers across all possible thresholds that could be chosen.

Let C_1 be the "positive" class, and C_2 be the "negative" class. A receiver operating characteristic curve, or ROC curve, is a plot of the true positive rate, that is the number of samples correctly classified as class C_1, against the false positive rate, the number of samples incorrectly classified as C_1, as the classification threshold is varied. Such a curve will always pass through the point (0,0) (for low thresholds, there are no true or false positives) and (1,1) (for high thresholds, all samples are classified as "positive"). If the classes are separable, the ROC cure will pass through the point (0,1), which indicates that there is a threshold that allows for a 100% true positive rate with no false positives.

The area under the ROC curve, or AUC, is a non-parametric measure of the classifier's effectiveness. Separable classes will have an AUC of 1. Two classes with little or no separability will have an AUC close to .5. For a more thorough discussion of ROC curves, see [Faw04].

In addition to calculating the AUC for each Fisher comparison, we also use the AUC value to evaluate the amount of distinguishing information in each individual feature band. For each comparison between two images, we take each individual feature band in turn and use only the values of each sample in that feature band to classify them. A classifier could be constructed by applying a threshold directly to the feature values in each individual band. We report on the AUC for such classifiers in Chapter 5.

The R package ROCR [SSBL05] was used to calculate the AUC and generate

the ROC curves in this thesis.

4.5 Feature Bands Used in Comparison

For each of the Gabor energy and grating operators, 48 feature bands were generated for each image: eight orientations at each of six levels of scale. The orientation parameters were 0°, 22.5^o, ... 157.5^o. These are the same orientations used by Petkov and Kruizinga for grating operator orientations in their analysis [PK97, KP99].

Each orientation is generated at six levels of scale. The first level of scale is the original image size, and each level of scale represents a doubling of the size of the receptive field of the operator. The wavelength parameter used for all operators was 4 pixels, so at the effective scales of the operators used to generate features are 4, 8, 16, 32, 64, and 128 pixels.

The set of 48 grating feature bands generated from an input image at the previously described eight orientations and six scales will be referred to as the "grating features" for an image. The set of Gabor energy bands will be referred to as the "Gabor energy features" of an image.

4.6 Experiment Design

A total of 48 grating and Gabor energy features are generated for each image: 8 orientations (0°,22.5^o,45°...157.5^o) at each of 6 scales. A feature map for each of these feature operators is generated for each 512x512 pixel texture image. This projects each pixel into a 48 dimensional space, with each dimension being the value of one of the feature maps at that pixel.

For each image, 10% of the pixels are selected randomly to be the samples, resulting in a set of 26214 samples for each image. The comparisons are performed on this subset of the image.

A Fisher analysis is performed on the 48-dimensional samples taken from each pair of texture images in the dataset, resulting in 5778 comparisons. 80% of the samples in each class are selected to be the training samples, the other 20% are the test samples. The Fisher projection vector and Fisher criteria is calculated using the training samples. The AUC statistic is calculated using the test samples.

Chapter 5

Results

In this chapter, we present the results of our evaluation. First, we present the Fisher criteria scores and the area under the ROC curve (AUC) for the texture pairs. Then, we provide a more detailed analysis of the performance of the grating cell operator. This analysis includes a subjective evaluation of the operator's performance across the different texture categories present in the Prague texture segmentation benchmark.

5.1 Fisher Results

We set out to show that grating cell features contain useful texture information. The data from the texture discrimination experiments supports this contention. Table 5.1 gives aggregate Fisher criteria statistics for the 5778 pairs of textures. The mean Fisher score for all pairs of textures using grating features is over 890, about 168 times higher than the mean score for all pairs of textures with Gabor energy features.

The higher Fisher criterion values for grating features are also reflected in the AUC values, although the difference is not as pronounced. There is a ceiling of 1

	Min.	Max.	Mean
Fisher Criterion (energy)	1.18772	30.24766	5.29668
Fisher Criterion (grating)	1.72117	6506.523	890.55103

Table 5.1: Minimum, maximum, and mean Fisher criteria for all pairs of textures.

	Min.	Max.	Mean
AUC (energy)	.8836155	1	.997503
AUC (grating)	.9999981	1	1

Table 5.2: Minimum, maximum, and mean AUC for all pairs of textures.

for the AUC value. Out of the 5778 pairs of textures, 2113 pairs are separable using Gabor energy features. This indicates that the Gabor energy features generally contain enough information to distinguish two textures. When grating features are used, all but two pairs of textures are separable. Grating features contain more distinguishing texture information than Gabor energy.

Figure 5.1 shows histograms of the Fisher projected sample values using both grating and Gabor energy features for the textures "bark17" and "glass8". The two textures are highly separable using grating features, but the two projected clusters overlap when using Gabor energy features. Figure 5.2 shows the ROC curves for Gabor energy and gratings for these two textures. As the two classes are separable using grating features, the ROC curve for the grating features passes through the point (0,1), indicating a 100% true positive rate with no false positives, and the area under the curve is 1.0. The two textures are not separable using Gabor energy features; the area under the ROC curve for Gabor energy features is .884.

Figure 5.3 shows histograms of the Fisher projected sample values using both grating and Gabor energy features for the "panel3" and "wood" textures. These two textures are separable using both Gabor energy and grating features. Using grating features, the samples from each class project to tighter clusters than with

Projected bark17 and glass8 points, energy features

Projected bark17 and glass8 points, grating features
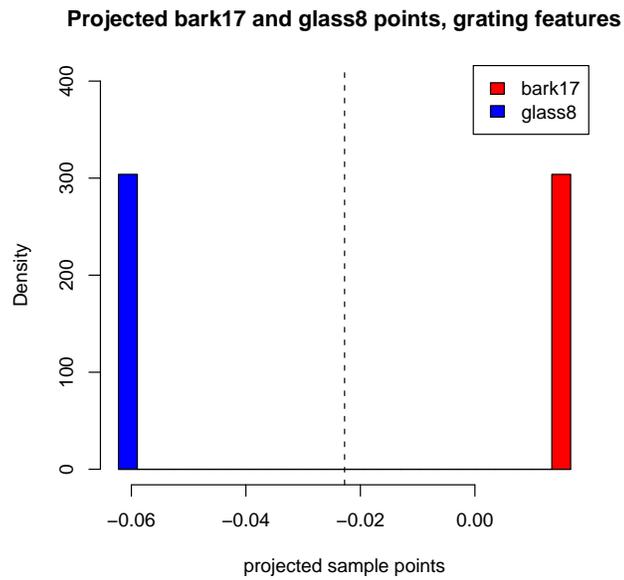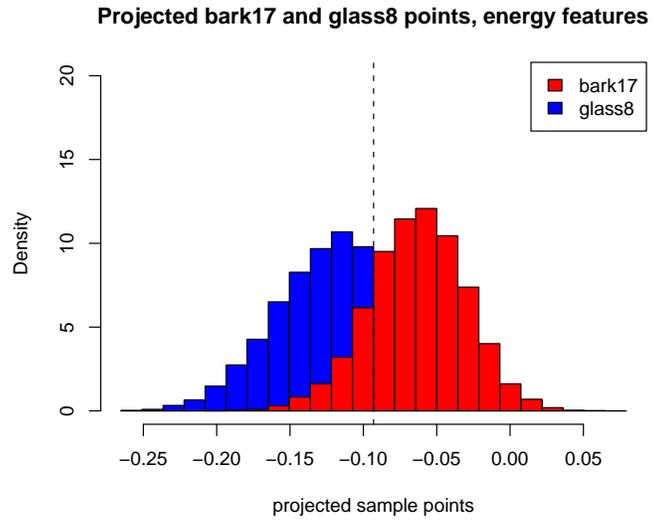
Figure 5.1: Histograms of projected samples for bark17 and glass8 textures.

45
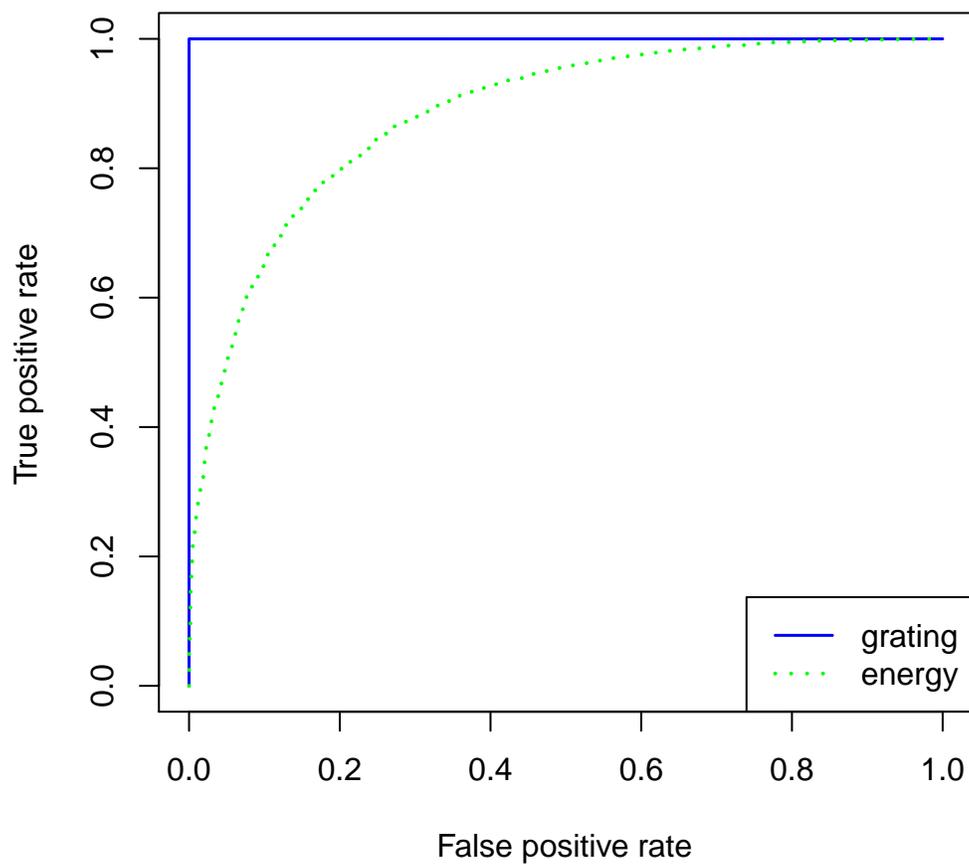
Figure 5.2: ROC curves for bark17 and glass8 textures.

Projected panel3 and wood points, energy features

Projected panel3 and wood points, grating features
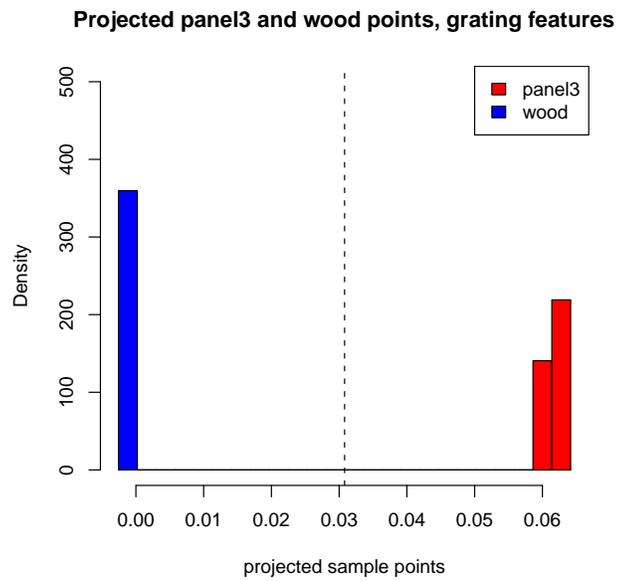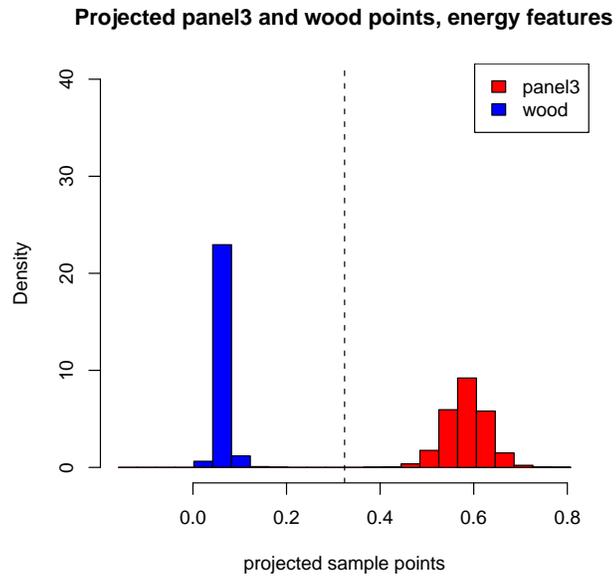
Figure 5.3: Histograms of projected samples for agave2 and bark10 textures.

First Texture	Second Texture	J (energy)	AUC (energy)	J (grating)	AUC (grating)
panel3	wood4	30.03542	1	1.721174	1
panel3	woodpanel3	14.96378	1	2.967591	1
wood4	woodpanel3	18.30076	1	3.190641	1
kaolin5	stone21	4.126295	0.9999425	3.855945	0.9999981
glass6	marble4	6.103584	1	4.217152	1
roofTiles4	woodpanel3	12.33496	1	4.372276	0.9999982
kaolin1	marble4	1.587033	0.9473344	4.563894	1
glassWall	oak	6.202118	1	4.737781	1
glass5	wood4	19.26053	1	4.793367	1
glass5	glass6	3.773492	0.999931	4.877406	1

Table 5.3: Lowest performing grating pairss.

Gabor energy features.

5.1.1 Worst Grating Performance

Though grating cell features generally were better at separating two textures than Gabor energy features, there are pairs of images in the dataset for which they did not perform well. In this section, we examine the texture pairs with the lowest Fisher criteria for grating cell features.

Table 5.3 shows the bottom 10 pairs ranked by Fisher criterion using grating features. Of the worst 10 pairs for grating features, the panel3 and wood4 textures are present in four.

Of note here is that for eight of these texture pairs, the Fisher criteria for energy is greater than that for grating. The mean Fisher criterion value for the pairs using energy features is 11.6687927, higher than the mean of the Fisher criteria for all pairs using energy features.

For the ten pairs with the lowest grating Fisher criteria, the AUC value when Gabor energy features are used is above the mean nine out of ten times. On the

	Minimum	Maximum	Mean
All Energy Bands	.50000003	1	.7297038
All Grating Bands	.5	1	.7997665

Table 5.4: Minimum, maximum, and mean AUC statistic for single band classifiers.

pairs that had low Fisher criteria with grating features, the Fisher criteria with Gabor energy features was high. We explore this relationship more in section 5.1.4.

Figures 5.4 and 5.5 shows the projected histograms for both energy and grating features for the worst performing grating pairs. What is interesting here is that the grating feature histograms do not appear to be Gaussian, which is the assumption made by Fisher's linear discriminant. For "roofTiles4" and "woodpanel3", the right tail of the "woodpanel3" class is significantly longer than the left, and overlaps the other class. For "panel3" and "wood4", the "panel3" class has an almost uniform distribution over its range, rather than a Gaussian.

5.1.2   Individual Feature Bands

For each pair of textures, we computed the area under the ROC curve for each individual feature band. Overall, 277344 pairs of feature bands were compared. Table 5.4 shows the aggregate statistics for the AUCs for all bands.

On average, each individual pair of grating feature bands is easier to distinguish than each pair of Gabor energy feature bands. As Figures 5.6 and 5.7 show, the distributions of AUC scores for grating and Gabor energy features is significantly different.

The distribution for Gabor energy features is non-uniform, with fewer band pair at higher AUC values. Since the AUC caps at 1, there is a ceiling effect, and therefore a spike at 1. Compare this with the distribution for grating feature band

49

Projected panel3 and wood4 points, energy features

Projected panel3 and wood4 points, grating features
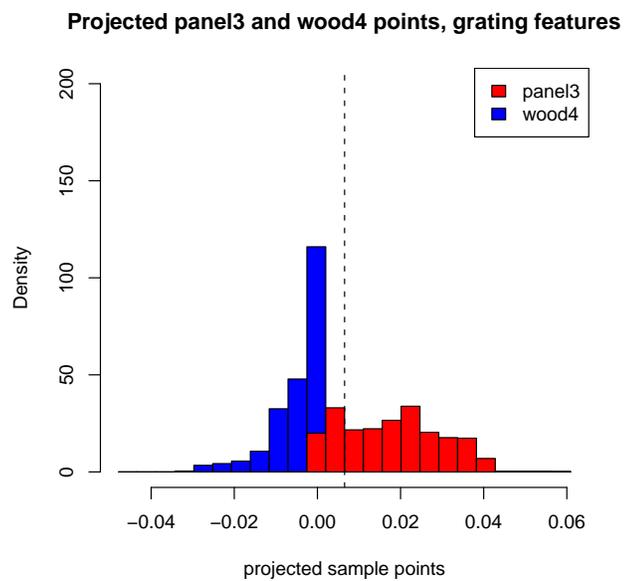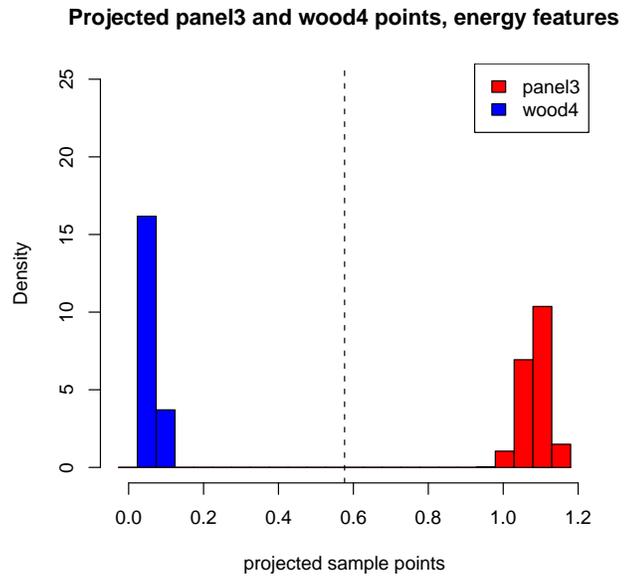
Figure 5.4: Histograms of projected samples for panel3 and wood4 textures.

Projected roofTiles3 and woodpanel3 points, energy featur
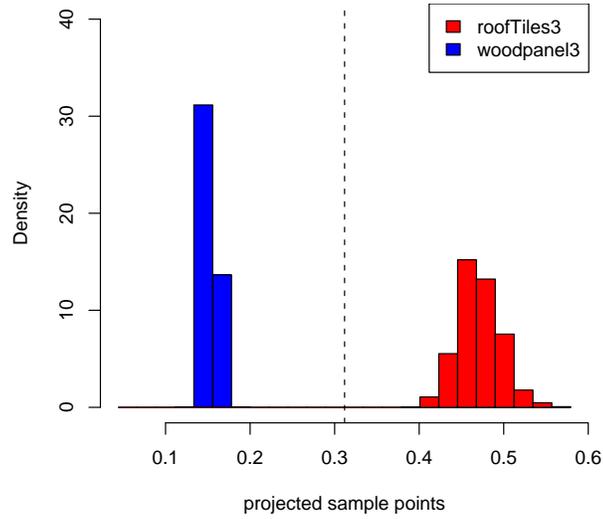
Projected roofTiles3 and woodpanel3 points, grating featur
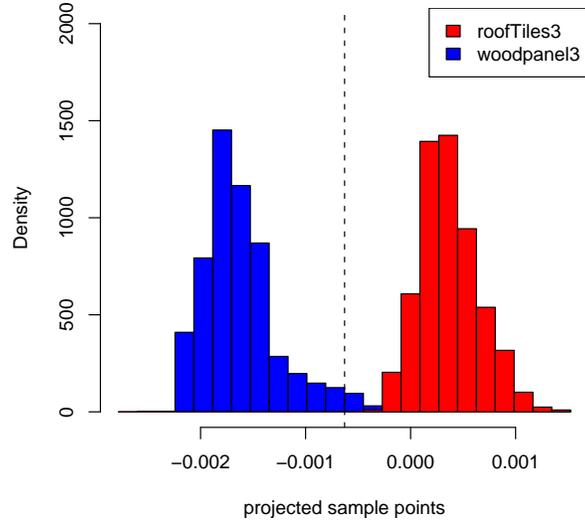
Figure 5.5: Histograms of projected samples for roofTiles4 and woodpanel3 textures.

AUC for Gabor energy feature bands
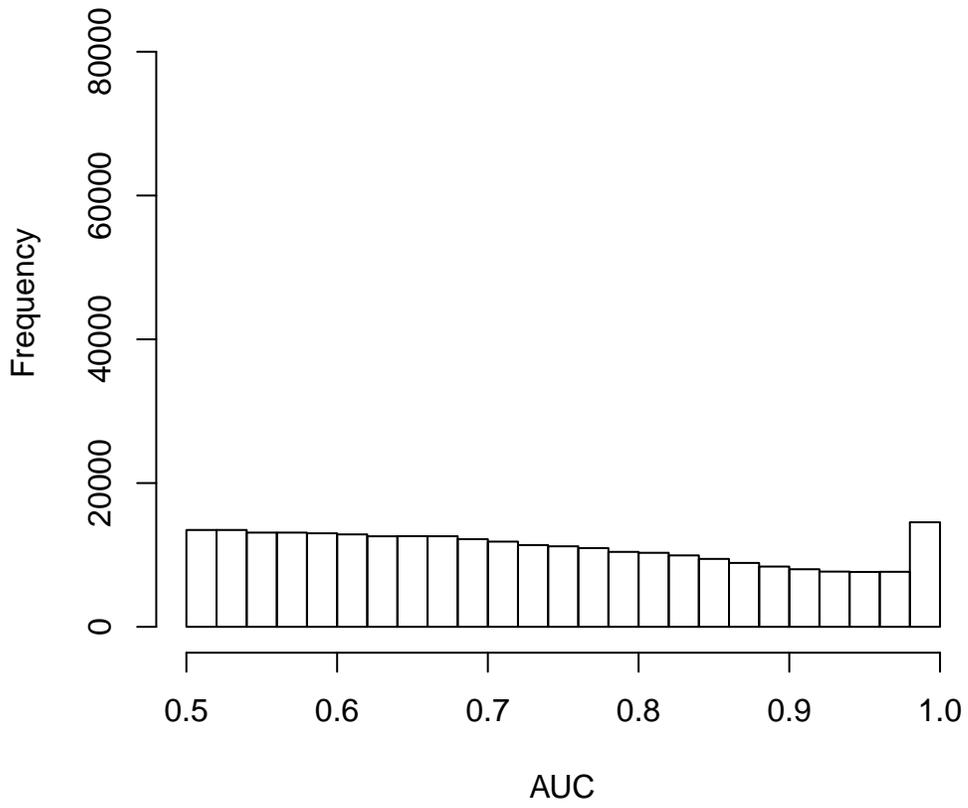
Figure 5.6: Histogram of AUC values for Gabor energy feature band pairs.

AUC for grating feature bands
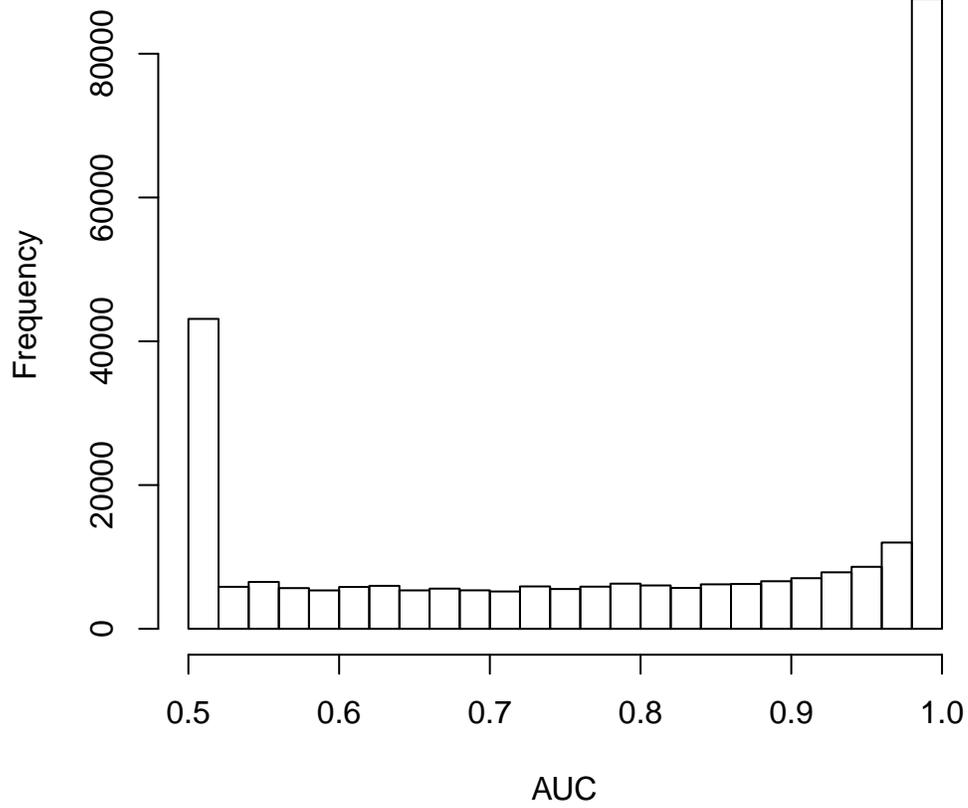
Figure 5.7: Histogram of AUC values for grating feature band pairs.

pairs. The distribution is nearly uniform between AUC values of .5 to 1, with two large concentrations around .5 and around 1. The distribution of the AUC values for pairs of grating feature bands is highly bimodal.

A single grating feature band is far more likely than a single Gabor energy feature band to contain enough information to perfectly distinguish textures in this dataset than a single energy band. 61321 pairs of grating feature bands had an AUC value of 1. This is around 22.1%. If two images and one grating feature band are selected at random from the dataset, there is a 22.1% chance that that one band contains enough information to distinguish the pixels of the two images. In contrast, only 1758, or .6%, of Gabor energy feature band pairs had an AUC value of 1.

It is also true that each individual grating feature band is more likely than a single Gabor energy feature band to contain little or no distinguishing information. 35618, or 12.8%, of the individual grating band pairs had an AUC value of exactly .5. No individual Gabor energy band pair had an AUC value of exactly .5, and only 6781 (2.4%) had AUC values less than .51.

The presence of AUC values of exactly .5 for grating band pairs is explained by the presence of grating bands that are identically 0. Of the 5184 grating feature bands generated, 1404, or 27.1%, are identically 0. When a grating feature band is entirely 0 for two textures, the AUC for those two textures along that band is exactly .5. Any threshold less than 0 results in 0% true and false positive rates, and any threshold greater than 0 results in 100% true and false positive rates.

None of the Gabor energy feature bands are entirely 0. When no grating of a particular orientation and scale is present in an image, the corresponding grating cell operator will produce a response of exactly 0. This is in contrast to the Gabor

54

energy operator, which will produce a small magnitude response due to noise even when there is no actual energy at the operator's orientation and scale.

5.1.3 Grating Cell Features as a Sparse Image Representation

The grating operator's final response is a Gaussian smoothing of a discrete intermediate step, the grating subunit responses. This suppresses all responses to stimuli that do not closely resemble the target spatial grating. This intermediate discretization is reflected in the final outputs, which show a tendency to be either exactly 0 or to high across much of an image.

102 out of the 108 images in the dataset have at least one identically 0 grating feature band. The mean number of identically 0 grating feature bands per image is 13. The maximum number of identically zero grating feature bands in one image is 44. Grating feature bands are therefore a much sparser representation of an image than Gabor energy feature bands, none of which were identically 0. This is true both in terms of the number of bands which contained any information, and in terms of the number of bands necessary to have enough information to distinguish images.

5.1.4 Correlation Between Gabor Energy and Grating Results

The correlation, ρ, between two sets, X and Y, is defined as follows:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \tag{5.1}$$

with $Cov(X,Y)$ being the covariance between X and Y, and σ_X and σ_Y being the standard deviation. Correlation ranges from -1 to 1. A value close to 1 indicates

Category	Textures	Mean J, grating	Mean J, energy	Correlation
bark	12	786.2892	3.3978	0.04928685
flowers	10	1206.411	3.1235	0.1493595
glass	7	350.2044	4.0329	0.1479511
man-made	20	1091.002	7.6148	0.4690026
nature	10	777.6871	4.1392	0.3095019
plants	10	804.3594	3.0183	0.4056697
rock	10	428.7988	2.6138	-0.1190436
stone	10	631.0043	3.6923	0.6033728
textile	10	413.2916	8.7002	-0.173837
wood	9	504.7391	8.6570	-0.1491143

Table 5.5: By category statistics.

that X and Y vary linearly together. Larger magnitude values indicate a stronger degree of dependence. We do not perform any statistical tests to determine the significance of the correlation values presented in this thesis. We merely report them and make anecdotal observations.

We compute the correlation between the Fisher criteria for pairss with Gabor energy features and the Fisher criteria for the same pairs with grating features.

The correlation between the Fisher criteria values for pairs of textures using Gabor energy features and the Fisher criteria values for pairs of textures using grating features is 0.1695534. We found this value to be surprisingly low. This indicates that grating cells are not just exploiting the same information in an image more effectively than the Gabor energy features, but suggests that grating cell features are fundamentally different from, not a refinement of, Gabor energy features.

5.2 By Category Study

Table 5.5 shows a summary of statistics for pairs of images from the same category. The correlation between the Fisher criteria for grating and Gabor energy features is highest for the "stone", "man-made", and "plants" categories. The best performing categories for grating features were "flowers", "man-made", and "plants".

Figure 5.8 shows the images in the "stone" category. The textures in this category appear to have little oriented texture. No image has an obvious spatial grating. This is reflected in the grating feature performance: the in-class pair mean Fisher criteria is 631.0043, which is below the overall mean.

Figure 5.9 shows the images from the "rock" category. Comparisons on textures within this category also performed worse than average with grating features, with an in-class pair mean Fisher criteria of 428.7988. Some of the textures in this category exhibit obvious spatial gratings, but many do not.

Figure 5.11 shows a selection of textures from the "man-made" category. Most textures in this category contain obvious oriented texture. Many of these textures are repeating simple geometric patterns. The Fisher criteria values for pairs with grating features is suitably high: 1091.002.

Figure 5.10 shows the textures in the "flowers" category. The pairs within this category with grating features did very well: the mean Fisher criteria was 1206.411. While there appears to be oriented texture in the flower textures, most images do not contain an obvious spatial grating.

Categories with textures containing more obvious spatial gratings tended to perform better using grating features than categories with textures that did not contain spatial gratings, although this does not appear to be a hard-and-fast rule.
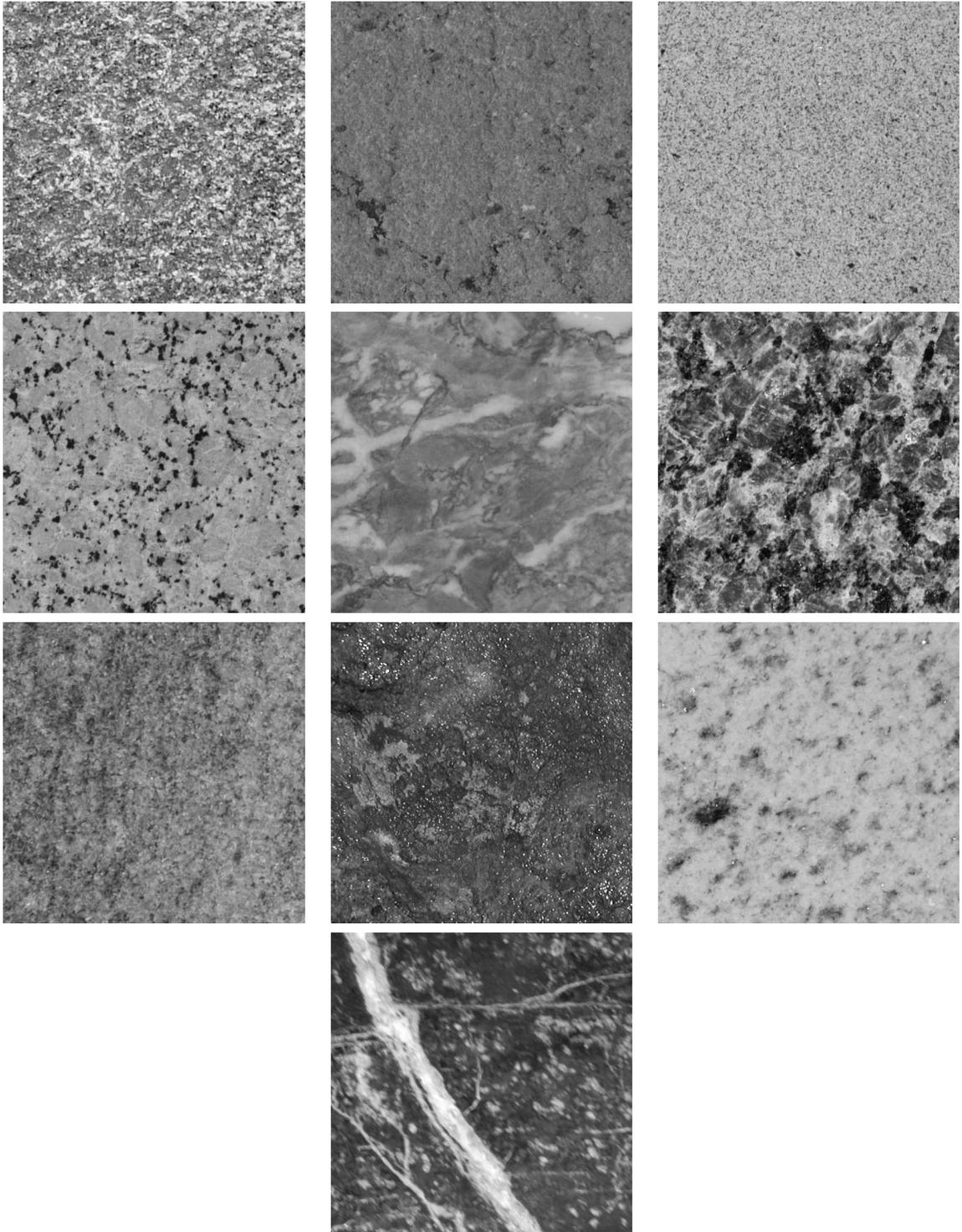
Figure 5.8: Textures in the "stone" category.
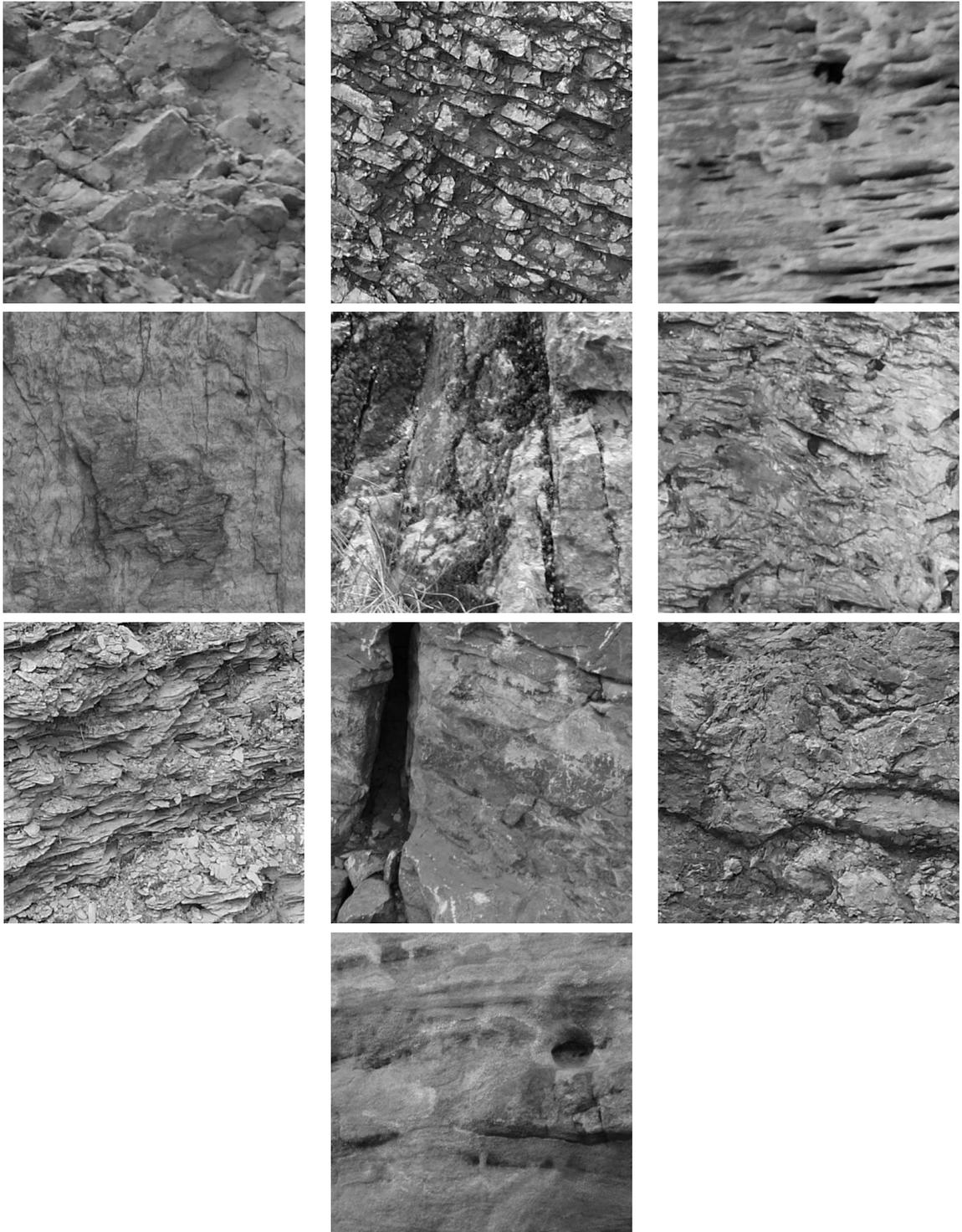
Figure 5.9: Textures in the "rock" category.
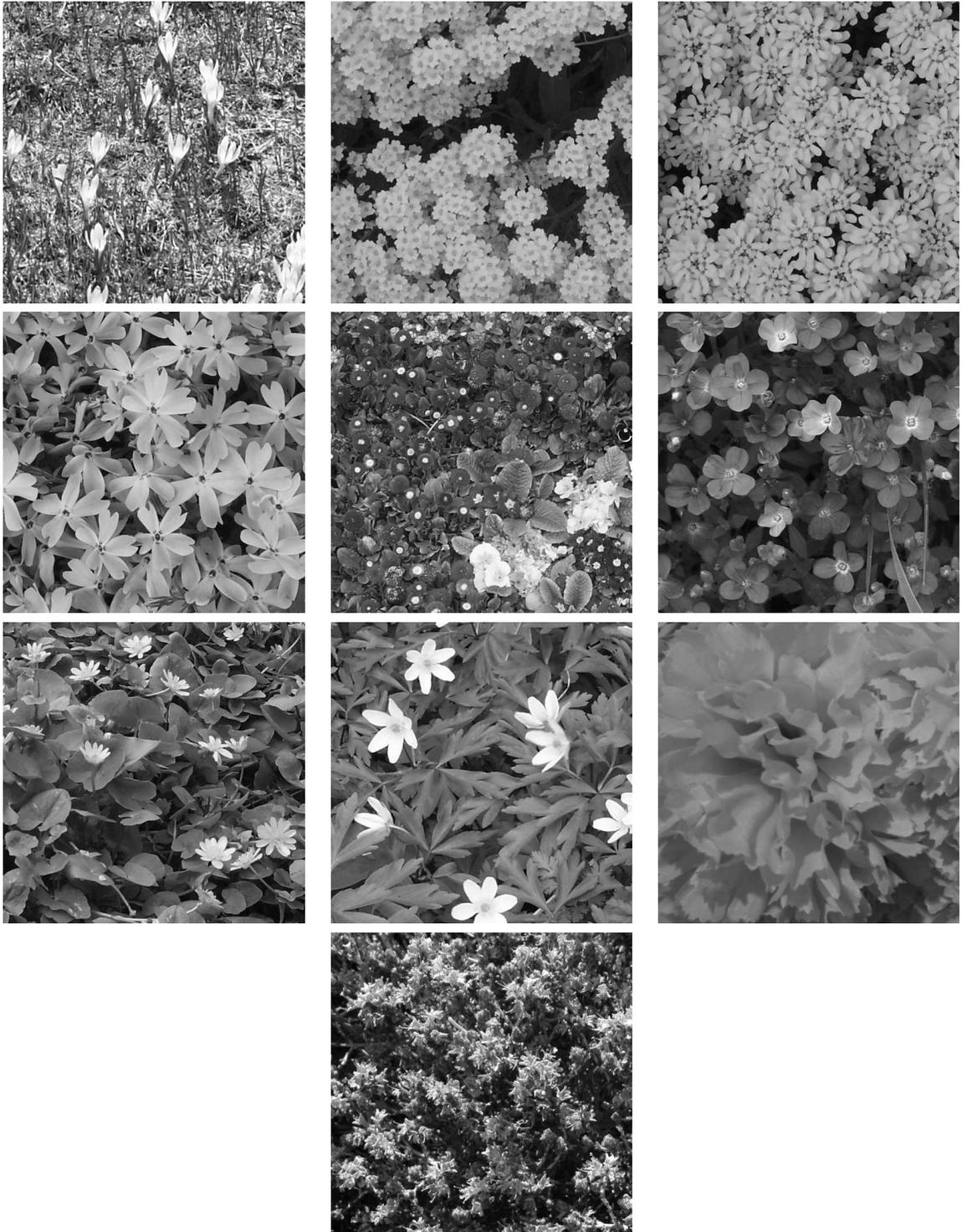
Figure 5.10: Textures in the "flowers" category.
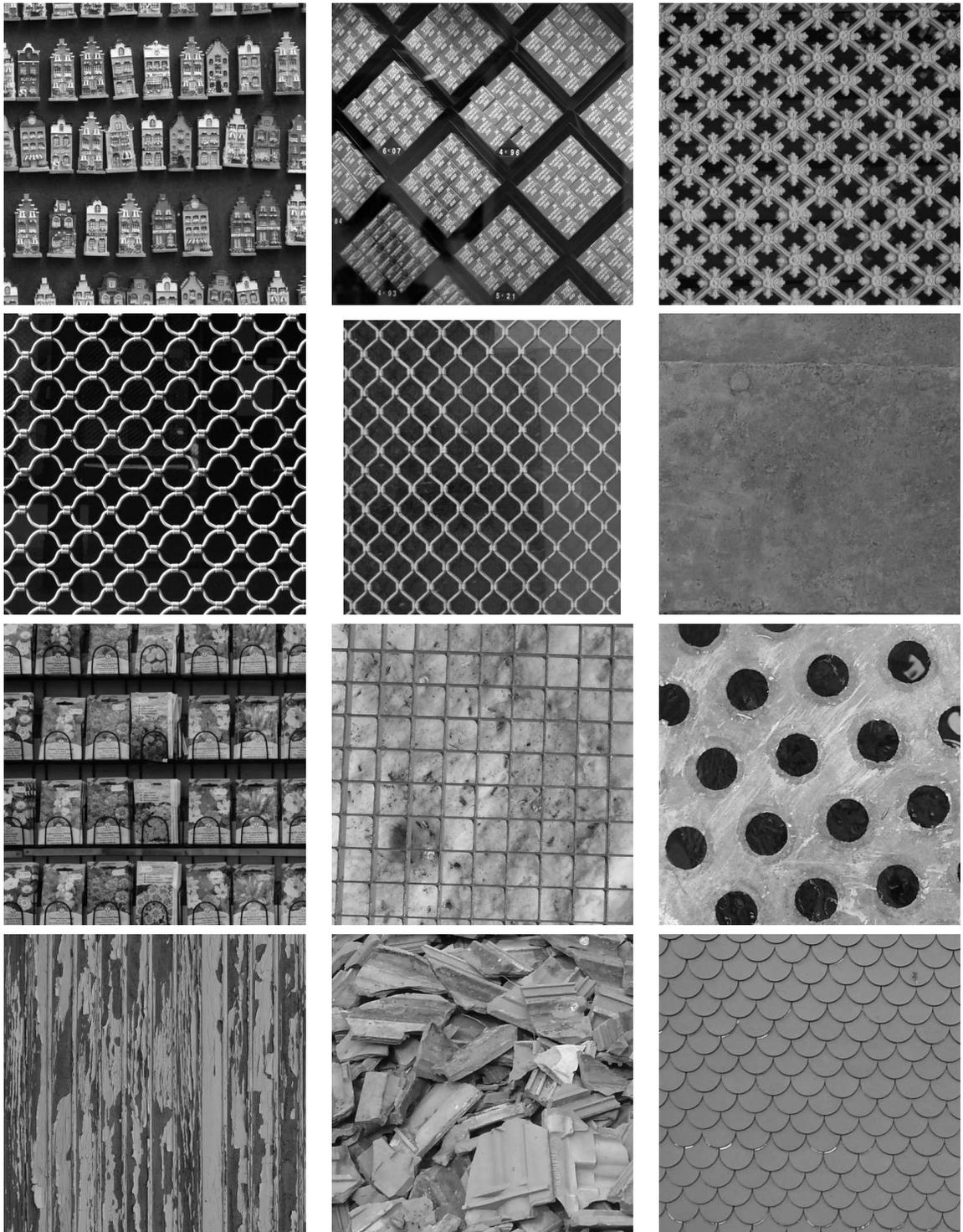
Figure 5.11: Twelve textures in the "man-made" category. The category contains twenty textures in total.

We leave a more thorough analysis of how the behavior of the grating operator differs across categories of images to future work.

Chapter 6

Conclusions

In this thesis, we compared features derived from a model of biological grating cells to another Gabor-based image feature. The performance task for comparison was texture discrimination.

The grating cell features were found to be more effective than Gabor-energy features overall at distinguishing textures in a set of difficult, natural textures. Image pairs had, on average, Fisher criterion scores 168 times higher in grating cell feature space than in Gabor-energy feature space.

This result suggests the usefulness of the grating operator as a means to extract low level information from natural images. Petkov and Kruizinga evaluated the operator only in the context of "oriented texture". We have broadened the analysis to the general class of texture images in general, not just those visually identified as being "oriented".

The correlation between the Fisher criteria for pairs of images using grating cell features and the Fisher criteria using Gabor-energy features is 0.1695534. This low correlation suggests that grating cell features are not tapping into the same image information as the Gabor energy operator.

Grating cell features are much sparser than Gabor energy features. A significant

fraction of the grating feature bands contain no image information. However, a significant fraction of pairs of grating feature bands from different images contain enough information alone to distinguish those two images. Grating cell features provide a sparser, but more information rich, representation of texture in an image. Sparse coding has been suggested as a strategy employed in the visual cortex to represent visual information [OF97]. The sparsity and effectiveness of the grating cell ceatures for texture discrimination suggests that grating cells may play a role in sparse representations in biological vision.

6.1   Future Work

The effectiveness of grating cell features in texture discrimination suggests their use in other computer vision applications. The most direct next step is to use grating features in the related problem of texture image segmentation. This is the problem of taking an image consisting of regions containing different textures, and determining the boundaries between the regions.

Kevin Burnett and Jason Remington used the grating cell model implementation developed for this thesis to test the effectiveness of grating cell features for the texture segmentation task [BR].

The application of grating cells to other problems in Computer Vision is of interest, given the performance of the grating cell features in the texture discrimination task.

Other grating cell models can be analyzed and compared to Petkov's. The reproducible success of Petkov's model has shown that grating cell features are useful. Next, it could be established which existing grating cell model is the most useful.

REFERENCES

[B+06]   C.M. Bishop et al. *Pattern recognition and machine learning.* Springer New York:, 2006.

[BR]   Kevin Burnett and Jason Remington. personal communication.

[Dau85]   J.G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 2(7):1160–1169, 1985.

[dB07]   JMH du Buf. Improved grating and bar cell models in cortical area V1 and texture coding. *Image and Vision Computing*, 25(6):873–882, 2007.

[FA91]   W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

[Faw04]   T. Fawcett. ROC graphs: Notes and practical considerations for researchers. *Machine Learning*, 31, 2004.

[FP02]   D.A. Forsyth and J. Ponce. *Computer vision: a modern approach.* Prentice Hall Professional Technical Reference, 2002.

[FS89]   I. Fogel and D. Sagi. Gabor filters as texture discriminator. *Biological cybernetics*, 61(2):103–113, 1989.

[Gab46]   D. Gabor. *Theory of Communication.* Institution of Electrical Engineering, 1946.

[HM08]   Michael Haindl and Stanislav Mikes. Texture segmentation benchmark. In *Proceedings of ICPR 2008*, 2008.

[HW62]   DH Hubel and TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106, 1962.

[JF90]   A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on*, pages 14–19, Nov 1990.

[JP87]   JP Jones and LA Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1233–1258, 1987.

[KP99]   P. Kruizinga and N. Petkov. Nonlinear operator for oriented texture. *Image Processing, IEEE Transactions on*, 8(10):1395–1407, 1999.

[LBOT05]   T. Lourens, E. Barakova, H.G. Okuno, and H. Tsujino. A computational model of monkey cortical grating cells. *Biological Cybernetics*, 92(1):61–70, 2005.

[Mar80]   S. Marcelja. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am*, 70:70–11, 1980.

[MH80]   D. Marr and E. Hildreth. Theory of Edge Detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences (1934-1990)*, 207(1167):187–217, 1980.

[MKS+04]   K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostin, F. Cardinaux, S. Marcel, S. Bengio, C. Sanderson, N. Poh, et al. Face authentication test on the BANCA database. In *Proc. of International Conference on Pattern Recognition*, pages 523–532, 2004.

[OF97]   B.A. Olshausen and D.J. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision research*, 37(23):3311–3325, 1997.

[Pec91]   S. Peckingpaugh. An improved method for computing gray-level coocurrence matrix based texture measures. *CVGIP: Graphical Models and Image Processing*, 53(6):574–580, 1991.

[PGJ89]   DA Pollen, JP Gaska, and LD Jacobson. Physiological constraints on models of visual cortical function. *Models of Brain Functions*, pages 115–135, 1989.

[PK97]   N. Petkov and P. Kruizinga. Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biological Cybernetics*, 76(2):83–96, 1997.

[RB97]   R.P.N. Rao and D.H. Ballard. Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation*, 9(4):721–763, 1997.

[SB06]   L. Shen and L. Bai. A review on Gabor wavelets for face recognition. *Pattern Analysis & Applications*, 9(2):273–292, 2006.

[SSBL05]   T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer. ROCR: visualizing classifier performance in R. *BIOINFORMATICS-OXFORD-*, 21(20):3940, 2005.

[vdHPD92]   R. von der Heydt, E. Peterhans, and MR Dursteler. Periodic-pattern-selective cells in monkey visual cortex. *Journal of Neuroscience*, 12(4):1416–1434, 1992.

[WHM04]   J. Weinman, A. Hanson, and A. McCallum. Sign detection in natural images with conditional random fields. In *Proc. of IEEE International Workshop on Machine Learning for Signal Processing*, 2004.