DATA-DRIVEN AUTOMATIC EDGE SHARPENING

BY

MYRA NAM

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Adviser:

Professor Narendra Ahuja
Sharpening is a powerful image transformation because sharp edges can bring out image details. Sharpness is achieved by increasing local contrast and reducing edge widths. We present a method that enhances sharpness of images and thereby their perceptual quality.

Most existing enhancement techniques require user input to improve the perception of the scene in a manner most pleasing to the particular user. Our goal of image enhancement is to improve the perception of sharpness in digital images for human viewers. We consider two parameters in order to exaggerate the differences between local intensities. The two parameters exploit local contrast and widths of edges. We start from the assumption that color, texture, or objects of focus such as faces affect the human perception of photographs.

When human raters are presented with a collection of images with different sharpness and asked to rank them according to perceived sharpness, the results have shown that there is a statistical consensus among the raters.

We introduce a ramp enhancement technique by modifying the optimal overshoot in the ramp for different region contrasts as well as the new ramp width. Optimal parameter values are searched to be applied to regions under the criteria mentioned above. In this way, we aim to enhance digital images automatically to create pleasing image output for common users.
Dedicated to my family.
I would like to thank all people who have helped and inspired me during my study in the master's program. I especially want to thank my adviser, Prof. Narendra Ahuja, for his guidance during my research and study at the University of Illinois at Urbana-Champaign. His enthusiasm in research drove me to pursue my studies despite obstacles. His advice and inspiration were always motivating for me.

I also thank all my lab mates in the Computer Vision and Robotics Lab for always being friendly and helpful. My deepest gratitude goes to my family for their unconditional love and support throughout my life; this thesis simply would have been impossible without them.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1</th>
<th>INTRODUCTION</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Overview of the Approach</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Thesis Overview</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2</th>
<th>BACKGROUND</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Motivations</td>
<td>4</td>
</tr>
<tr>
<td>2.2</td>
<td>Human Visual System and Perception</td>
<td>6</td>
</tr>
<tr>
<td>2.3</td>
<td>Techniques Preserving Image Structure</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>Region-Based Image Segmentation</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 3</th>
<th>RAMP ANALYSIS AND EDGE ENHANCEMENT</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Ramp Enhancement Models</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>Experiments</td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 4</th>
<th>EXPERIMENT DESIGN FOR OPTIMAL EDGE SHARPENING</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Experimental Procedure</td>
<td>30</td>
</tr>
<tr>
<td>4.2</td>
<td>Observers</td>
<td>34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 5</th>
<th>RESULTS AND RANK-ORDER CLASSIFICATION</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Notation and Problem Formulation</td>
<td>36</td>
</tr>
<tr>
<td>5.2</td>
<td>Data Analysis</td>
<td>39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 6</th>
<th>STATISTICAL ANALYSIS OF ORDINAL DATA</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Statistical Analysis by Analysis of Variance</td>
<td>47</td>
</tr>
<tr>
<td>6.2</td>
<td>Cumulative Logit Models for Ordinal Responses</td>
<td>51</td>
</tr>
</tbody>
</table>

| CHAPTER 7 | CONCLUSIONS AND FUTURE WORK | 61 |

| REFERENCES | | 63 |
1.1 Overview of the Approach

It is widely known that humans prefer photographs with sharp and crisp edges. Sharpening is a powerful image transformation because sharp edges can bring out image details. Sharpness is achieved by increasing local contrast and reducing edge widths. We present a method that enhances the sharpness of images and thereby their perceptual quality.

Most existing enhancement techniques require user input to improve the perception of the scene in a manner most pleasing to the particular user. Extreme sharpening can cause each pixel to stand out from neighboring pixels. Unnaturally pronounced edges make dark objects outlined with light halos and light objects outlined with dark halos. Too little sharpening makes no difference from the original image. Moreover, small noise in the image can appear as a texture in an otherwise smooth image after sharpening. Most existing image sharpening techniques are empirical or heuristic.

Our goal of image enhancement is to improve the perception of sharpness in digital images for human viewers. Instead of enhancing all the edges, we enhance the perceptual image structure that is composed of regions of homogeneous intensity. We define “overshoot” to be the excess of intensity over the homogeneous intensity inside the region due to the local contrast increase. A ramp is a transition from one region to an adjacent region. Region boundaries are located on ramps and are equivalent to ramp discontinuities. The ramp across region boundaries will be targeted for image sharpening. We consider two parameters in order to exaggerate the differences between local intensities. The two parameters exploit local contrast and widths of edges – overshoot and ramp width in the
vicinity of the region boundaries.

Unlike the existing methods, we attempt to make an automated sharpening technique. We start from the assumption that color, texture, or objects of focus such as faces affect the human perception of photographs. In other words, we seek the best enhancement for these different criteria by finding the best combination of parameter values. Texture is considered to adapt to the local spatial content in the image. The perceived color of an entire region can change depending on overshoot. In addition, among many objects, we have chosen faces for our experiment since faces tend to be at the center of digital photographs and people tend to locate them first in images. An object of focus as such can require different values of parameters from other ordinary objects. Or we might need parameter values to avoid the skin texture and sharpen the eyes of a portrait. When human raters are presented with a collection of images with different sharpness and asked to rank them according to sharpness, the results have shown that there is a statistical consensus among the raters.

We introduce a ramp enhancement technique by modifying the optimal overshoot in the ramp for different region contrasts as well as the new ramp width. Optimal parameter values are searched to be applied to regions under the criteria mentioned above. In this way, we aim to enhance digital images automatically to create pleasing image output for common users.

1.2 Thesis Overview

Backgrounds in human visual perception in edge sharpening will be presented in Chapter 2. In Chapter 3, we present an edge sharpening approach driven by an objective function that is controlled by two parameters. One parameter, overshoot, controls local contrast of edges, while the other, ramp, controls edge widths.

Optimal parameter values are driven by humans to enhance the sharpness of images. In other words, Chapter 4 describes the experiments performed to obtain the parameter values. The output controlled by the optimal parameters has the best perceived quality in sharpness irrespective of the observer. Our approach is that the two parameter values
are determined optimally according to texture sizes, color, and objects of focus, especially faces, in images. We produce a trained model that generates ratings of sharpness perception that are in accordance with the human ratings.

In Chapter 5, we describe the data analysis to compute relative ranking of parameters by supervised rank-order classification. In Chapter 6, the main effects of the two parameters will be discussed by a method, analysis of variance. In Chapter 7, the ordinal data will be analyzed to obtain marginal probabilities for a particular rank with different parameter values.
CHAPTER 2
BACKGROUND

2.1 Motivations

The goal of image enhancement is to process an image so that the output is more suitable than the original image for a specific application. One of the applications is a preprocessing procedure for high-level image structural analysis. Especially, our goal is to produce a pleasant output for viewers, with better sharpness and crispness of edges. The existing edge sharpening technique is a point-wise operation [1]. Let \( f(x) \) be a one-dimensional grayscale intensity profile, where \( x \) is the 1-D pixel coordinate. The first-order derivative is defined as the difference

\[
\frac{\partial f}{\partial x} = f(x + 1) - f(x).
\]

A second-order derivative is defined as the difference

\[
\frac{\partial^2 f}{\partial x^2} = f(x + 1) + f(x - 1) - 2f(x).
\]

First-order derivatives produce thicker edges in an image while second-order derivatives have a stronger response to thin lines and points.

The Laplacian operation is a two-dimensional operation defined as

\[
\nabla^2 f = [f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1)] - 4f(x, y).
\]

Adding the results of the Laplacian operation to the original image sharpens edges by highlighting the gray-level discontinuities. This is called unsharp masking.
Software with edge sharpening techniques is available for professionals and common users. One example is a graphics editing program, Adobe Photoshop CS3 extended, version 10.0.1. It allows users to determine values of several parameters as part of the unsharp masking filter. First, radius determines how many pixels on both sides of an edge will be enhanced. The radius setting is influenced by the resolution of the image. For the best outcome, users need to try out different values to create distinct borders of objects, and edges will have different thickness. Second, the amount parameter adjusts contrast levels of edges. High value of amount brings strong edges with high contrast, and a very high value could result in adjacent pixels of edges being black and white. The last setting is removal of Gaussian blur, motion blur, and lens blur. Motion blur is estimated in the direction that the user chooses and is removed.

The software first detects prominent edges and sharpens them. First, the picture is smoothed or blurred slightly (unsharpened), and the contrast is reduced just enough to reduce noise or dirt and to further blend areas of gradual tonal transition. The filter then looks for neighboring pixels with a certain level of contrast, and it increases the contrast of these pixels further, providing the sharpening effect. The drawback of Photoshop is that users need to try out different values of parameters until they obtain the results that meet their own standards, which is a heuristic method.

In order to achieve edge sharpening, we focus on two parts. First of all, we will make edge sharpening models. We will discuss how to locate edges and enhance edges by increasing local contrast without amplifying noise. Second, we will discuss how much we should sharpen images to produce the best image quality.

There is no existing study or model with the human visual perception in edge sharpness. Our goal is to draw on human perception in order to obtain optimal sharpness of any input digital images.

The following section deals with the human visual system, image formation and perception. And in the next section, we will discuss image segmentation that will be used in our algorithm to extract semantic object boundaries.
2.2 Human Visual System and Perception

As illustrated in Figure 2.1(a), light is transmitted through the lens to form images on the retina which are sensed by two classes of retinal photoreceptors, rods and cones [2], [3].

Rods are responsible for vision at low light levels, referred to as scotopic vision. They have high sensitivity and monochrome vision. Cones are active at higher light levels, referred to as photopic vision, are capable of color vision and are responsible for high spatial acuity. They are receptors that contain a visual pigment, called rhodopsin, or visual purple.

The fovea has a very high density of cones, and no rods. There are three types of cones: S-cone (the short-wavelength sensitive cones), M-cones (the middle-wavelength sensitive cones), and L-cones (the long-wavelength sensitive cones). The six to seven million cones can be divided into red cones (64%), green cones (32%) and blue cones (2%) based on measured response curves.

Understanding perception processing in simultaneous contrast is important for relating the effects of overshoot and perceptual fidelity in image enhancement. We will discuss how the human visual system interprets visual image data such as brightness and edges. Visual information is processed and transmitted by neurons. As illustrated in Figure 2.1(a), the optic nerve, running inside from the retina out to the back of the eyeball and into the brain, is a bundle containing about a million individual neurons. The cell bodies of the neurons that make up the optic nerve are called retinal ganglion cells. The ganglion cells exist on the surface of the retina. The light signal is transmitted from receptors, bipolar cells, to ganglion cells in Figure 2.1(b). When a retinal ganglion cell receives strong input signals, it is activated to generate an impulse, and the impulse is propagated down its axon that extends to the brain.

Synapses are junctions where receptors and neurons exchange signals. The receptors or neurons release chemical mediator substances, which act upon the receiving cell membrane. The action at the synapse is called excitation. Some of the input endings that interact with the retinal ganglion cells produce excitatory effects, while some of the endings on the same cell body may have an inhibitory effect. Lateral inhibition sharpens the spatial profile
Figure 2.1: Anatomy of human eye.
of excitation because neurons inhibit their neighbors.

The lateral inhibition is found in an optical illusion, Mach bands, invented by Ernst Mach in 1865. Each band is a solid color in Mach bands. However, the bands reflect different amounts of light with the darker bands reflecting less and lighter bands reflecting more. It appears as if the color bands curve inward or that each band is a gradient.

In Figure 2.2, the receptive fields are represented as a center disk and annulus. The center disk indicates an excitatory area (+), and the annulus an inhibitory area (-). The receptive fields in areas of uniform intensities receive the same stimulation in their excitatory centers and inhibitory surrounds. Therefore, the center excitations are in balance with the surround inhibitions [4].

The receptive field over the bright Mach band gives a stronger response in the center because part of the surround is in the darker area. Therefore, it receives less inhibition from the surround than the center at the left or right ends of bands. Since the excitatory response is less, the transition region is brighter on the right side and darker on the left side.

The Mach band effect from Figure 2.3 shows that perceived brightness is not uniform along the boundary of the bar although each bar has uniform intensity.

Brightness is not a simple function of intensity. The ability of the human eye to discrimi-
Figure 2.3: Perception and illusion: Mach band effect. The red curve in the graph indicates the perceived brightness of the series of homogeneous bars. It has overshoot effects near edges.

Distinguish between different intensity levels is an important consideration in subjective brightness [1]. The Weber-Fechner law implies a logarithmic relationship between physical luminance and subjectively perceived brightness. Suppose that the uniform illumination in the background is $L_B$ and the foreground luminance change is $\nabla L$. The ratio of $\frac{\nabla L}{L_B}$ is constant when the surround luminance $L_S$ and background luminance $L_B$ are the same. Refer to Figure 2.4(a). When an increment of illumination, $\nabla L$, is discriminative against the uniform background illuminance, $L_B$, the ratio $\frac{\nabla L}{L_B}$ is called the Weber ratio. The Weber ratio indicates the just noticeable visibility threshold. Although the ratio varies with intensity values, the Weber ratio, $c$, is approximately 0.02.

$$\nabla L = c L_B, \quad c = 0.01, \ldots 0.02$$

Our visual system is sensitive to luminance contrast rather than the absolute luminance. The Weber ratio and Mach band illustrate that human perception is sensitive to local contrast, not to actual illuminance. Perceived brightness is not the same as the true intensity. If local contrast across edges is increased enough not to be perceived as halo effects, it could be an aid in the design of image enhancement techniques. However, when the surround luminance $L_S$ is not the same as $L_B$, the Weber ratio changes. In image sharpening, it is important to find the optimal local contrast with different surround luminance in order to produce a pleasing output.
Figure 2.4: (a) Experiment of Weber-Fechner law. The Weber ratio is a just noticeable
visibility threshold and indicates the importance of contrast for perceived brightness of a
region. (b) Spatial contrast sensitivity.
In addition, spatial contrast sensitivity should be considered in edge sharpening. Lateral inhibition and excitation together lead to a bandpass characteristic of the contrast sensitivity function of the human visual system as shown in Figure 2.4(b).

The perceived brightness is different from absolute luminance and depends on the intensity contrast. A region is defined as a connected set of homogeneous pixels that allows smooth variations inside and is surrounded by steep discontinuities. By adding overshoot and undershoot in the vicinity of region boundaries, the local contrast can be enlarged in the ramp that transitions from the region to background. This ensures that the region is distinguishable from the background by human perception. Still, the human visual system perceives that the region preserves the homogeneity in intensity. Therefore, the structure of a region will be enhanced.

2.3 Techniques Preserving Image Structure

In [5], Tomasi and Manduchi introduced the bilateral filtering technique, an enhancement technique that smoothes images but preserves main discontinuities. It is a combination of two Gaussian filters of Euclidean distance in space and of the intensity difference.

With the Gaussian filter, \( g_\sigma(x) = e^{-x^2/\sigma^2} \), the filter is defined by

\[
h(I)_p = \frac{1}{W_p} \sum_{q \in I} g_\sigma_s(\| p - q \|) g_\sigma_r(\| I_p - I_q \|)I_q,
\]  

(2.1)

where \( W_p \) is the normalization factor of the weights

\[
W_p = \sum_{q \in I} G_\sigma_s(\| p - q \|)G_\sigma_r(\| I_p - I_q \|).
\]  

(2.2)

The output is a weighted mean of the neighborhood pixel intensities, whose weights are inversely proportional to the Euclidian distance of the neighbor pixel locations and the
Figure 2.5: Bilateral filter uses spatial and range Gaussian filters to smooth regions while keeping discontinuities. Thus, image structure is preserved.

intensity difference, which creates a piecewise constant image as an output, as shown in Figure 2.5.

The purpose of the bilateral filter is to preserve the global structure of the input image. We present an image enhancement technique that, like the bilateral filtering, preserves the image structure. However, our approach does not smooth the interior of regions; rather, it sharpens the edges that construct the image structure.

2.4 Region-Based Image Segmentation

We claim that the enhancement of image structure can be performed by ramp enhancement. A region is a low-level descriptor that represents an image structure, which contains homogeneous intensities with small variations allowed inside [6]. Ramp is a transition from one region to an adjacent region, and structure boundaries exist inside a ramp.

From the definition of the region model, ramp pixels, which are located in the vicinity of a region boundary, are variations of non-ramp homogeneous pixels of the region. Ideally, region boundaries should be sharp and exist in the same place as the exact boundaries of
Therefore, our hypothesis is as follows: If the ramp pixels are enhanced to the intensity level of the homogeneous non-ramp region, the enhancement effect by preserving the image structure is achieved.

We argue that human visual perception captures the image structure by grouping pixels together with similar distributions in intensity. This perception process corresponds to the region model. Increased contrast near structure boundaries makes the adjacent regions more distinguishable, while the perceptual luminance of the region remains the same. Therefore, instead of enhancing all the edges, we enhance the vicinities of region boundaries only.

In order to obtain regions, a segmentation algorithm in [7] is used prior to the enhancement. The image segmentation is built based on grayscale contrast and ramp discontinuities. Ramp estimation is performed by the proposed ramp model in [7]. In this way, the perceptual structure is constructed of regions by the segmentation.

The first step in edge sharpening is to detect the low-level structure in images. We detect structure boundaries as regions. A region consists of spatially adjacent pixels, and is measured by grayscale similarity. The region model has homogeneous intensity with slight variation, surrounded by ramp discontinuities. Ramp discontinuities exist while surrounding edges in Figure 2.6(a) because of blur during image acquisition process, lens defocus, and low resolution.

The segmentation algorithm partitions an image into a set of regions with contrast equal to or greater than a photometric scale, $\sigma_g$.

For each point $p$ in the image, define a function as

$$f_p(c) = \min ||p - q|| : |I(p) - I(q)| \leq c.$$

where $c$ is intensity contrast and $q$ is a neighboring pixel.

A region is divided into ramp pixels and non-ramp pixels using this function. It increases rapidly with $c$ for non-ramp pixels, while ramp pixels have small values of $f_p(c)$. Each ramp
discontinuity has two parameters, an associated width and contrast. The width in Figure 2.6(c) is a distance between a ramp pixel and another ramp pixel that is farthest along the ramp cross section. The cross section of the ramp is detected as the direction of the gradient. The contrast in Figure 2.6(b) is the intensity difference between these two pixels. The two functions decrease to local minimum at the center of the ramp, where an edge is located. All the pixels inside the ramp are assigned true contrast of ramp $\sigma_g$ and true width $w$.

After estimating ramp pixels with associated width and contrast, segmentation is performed by region growing method at a photometric scale, $\sigma_g$. Because contrast $c$ varies, different sizes of regions are formed according to $c > \sigma_g$ at different values of $\sigma_g$. Thus, the ramp discontinuities are estimated. The estimated ramp width and contrast are the input to the edge sharpening method.

2.4.1 Sharpening techniques

Before introducing our model, we now explore sharpening techniques in image enhancement and investigate how the proposed technique is different from the existing sharpening approaches in terms of overshoot and ramp width. First of all, the unsharp masking filter \cite{1} adds high-frequency components to the input image by isotropic Laplacian filter,

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \tag{2.3}$$
The second-order derivative of the ramp profile is defined as the following difference:

\[
\frac{\partial^2 f}{\partial x^2} = f(x + 1, y) + f(x - 1, y) - 2 f(x, y).
\] (2.4)

A ramp is formulated to have a constant slope in the middle of a ramp profile. Hence, the second-order derivative is non-zero only at the beginning and end of the ramp. The unsharp masking, \( g(x, y) = f(x, y) + \nabla^2 f \), increases the intensity of the bright side of the edge, while that of the dark side is lowered. As a result, it enhances the local contrast of the edge, while the ramp width and slope remain the same. Its emphasis on the high frequencies, however, amplifies noise. Therefore, Polesel et al. in [8] introduced a new method of unsharp masking to avoid the sharpening effects in smooth areas. An adaptive filter is employed to enhance mid- and high-frequency components while suppressing the enhancement in smooth areas. Nonetheless, local contrast by the Laplacian filter is added to the image, which causes artifacts if the contrast is high, especially near edges. Therefore, the overshoot effects of the unsharp masking could result in ringing effects. In summary, sharpening effects can be accomplished by spatial differentiation as shown in Figure 2.7(b).

Second, while the former method produces overshoot effect, Leu in [9] presented a different sharpening approach by reducing ramp width. For each pixel, its gradient direction and magnitude are estimated by searching for the maximum deviation of the intensity changes in the neighborhood. Then, it is determined to be a ramp pixel if the maximum value of the first-order derivative of the given ramp profile is higher than the given threshold. As shown in Figure 2.7(c), the ramp width is reduced while keeping the centerline of the edge. Compared to the first method, the ramp width reduction will not change the intensity levels by adding artificial overshoots along the edge. It will produce no effects on edges with small width, while reducing large ramp width significantly.

Yet, overshoots could be effective edge sharpening components if they were added around the edges to the extent that the human visual system did not perceive the existence of overshoots as artifacts. As discussed in Section 2.2, the perceived brightness is not always the same as the absolute luminance. As shown in Figure 2.7(d), the goal in this research is to
modify both the ramp width and overshoot in the vicinity of region boundaries. Therefore, we will have two factors, the desired ramp width and overshoot, as sharpening effects. As opposed to the edge-sharpening methods mentioned previously, our proposed method will sharpen only region boundaries by ramp modification through ramp width reduction and overshoot. Edges are formed by gray-level discontinuities at a point. A region boundary is a closed path that surrounds a finite region. Therefore, our approach will achieve the global enhancement of the image structure by enhancing respective regions locally.
Numerous techniques have been introduced in image enhancement for multiple purposes. First of all, contrast enhancement is a point operation that stretches the dynamic range of low-contrast image intensities. Low-contrast images occur because of poor lighting conditions or poor dynamic range of the imaging sensor. Contrast stretching is a global technique and could amplify noise. Second, smoothing techniques such as spatial averaging or low-pass filtering are used to extract image features as a blob-like form. Spatial averaging is performed by a convolution of a finite impulse filter, called spatial mask. Spatial averaging filters are commonly used for images that contain Gaussian noise. However, it introduces blurring across edges. Median filter, one of the order-statistics filters, arranges pixel values inside a window in ascending or descending order and chooses the middle value. It is a nonlinear filter, well known for removing salt-and-pepper noise. It performs poorly when the noise is Gaussian.

On the other hand, sharpening effects are performed with the emphasis on the mid- and high-frequency contents to enhance edges and detail in the input image. For example, unsharp masking is a neighborhood operation that crisps the edges. Sharpening effects are created by increasing local contrast across edges or reducing the width of an edge [8], [9].

In this chapter, we present a method that achieves sharpening effects with the combination of local contrast and the width of an edge. Moreover, instead of enhancing all the edges, we enhance the perceptual image structure that is composed of regions that are homogeneous in intensity. Sharp edges are important for the visual appearance of an image. Region boundaries are located on ramp regions and are equivalent to ramp discontinuities.
Therefore, in order to emphasize the structure of the region boundary, ramp enhancement will be performed for edge sharpening effects.

Region boundaries exist inside ramps. Ramps in an image are caused by degradations due to the following: circular aperture of a lens, blur due to camera being out-of-focus, edges of round objects, degradation due to sensor noise, and atmospheric turbulence, whose impulse response is formulated as a point spread function, a circularly shaped Gaussian function [10].

Therefore, ramps inside a region will be targeted for enhancement to reduce the effect of degradations in sharp region boundaries. Two parameters are considered for edge sharpening – overshoot and ramp width in the vicinity of the region boundaries.

In summary, we introduce a ramp enhancement technique by modulating the optimal overshoot in the ramp for different region contrasts as well as the new ramp width. The outline of the chapter is as follows. In this Section, we have presented the background of sharpening techniques and the motivation for the proposed enhancement techniques. In Section 3.1, two techniques in ramp enhancement are proposed and compared. Section 3.2 discusses implementation details and results, where we evaluate the outputs of synthetic and real images in terms of the two parameters, overshoot and ramp width. In conclusion, we discuss the strength and limitation of our proposed work, followed by tasks for future work.

### 3.1 Ramp Enhancement Models

In this section, two models are presented in order to achieve the following purpose: given a ramp profile as an input, a modified ramp profile is obtained with the desired overshoot and ramp width.

Without the existence of overshoot, if ramp pixels are updated by the replacement of the spatially close non-ramp homogeneous pixels as in Figure 3.1, the new ramp pixels will ensure the homogeneity of the region. Furthermore, if overshoot and undershoot are added on the ramp region, local contrast among regions will be enhanced while keeping the region
structure. An assumption is that we know in advance the location of region boundaries. The input to the models is the given ramp profile as a cross section of the ramp, and adjustable parameters are the desired ramp width and overshoot.

3.1.1 Heat equation method

From the assumption of a homogeneous region model, a ramp should be an extension of a non-ramp part of the region. However, a ramp exists due to atmospheric turbulence or lack of focus. The problem of enhancing a ramp can be formulated as the propagation of intensities at the contour of non-ramp parts of a region towards the region boundary. Again, Figure 3.1 explains this approach. It is analogous to a heat equation with homogeneous boundary conditions. Let us consider a cross section of a ramp, $f(x)$, as a function of a ramp pixel position, $0 < x < L$, lying on the cross section. Let the first non-ramp pixel intensity be located at the position $x = 0$. The value $x = L > 0$ indicates the location of the region boundary. As illustrated in Figure 3.2(a), the problem becomes

$$f_{xx} = f_t, t > 0$$
$$IC : f(x, 0) = f_o(x), 0 < x < L$$
$$BC : f(0, t) = f_o(0), f(L, t) = f_o(L)$$
Figure 3.2: Diffusion: (a) A model for homogeneous heat equation; f(t) indicates temperature in heat equation and intensity in the image enhancement model. (b) The proposed method by nonhomogenous heat equation for enhancing the ramp profile.

where \( f_o(x) \) is the initial ramp profile and \( t \) is time.

The physical interpretation of zero overshoot condition results in smoothing effects in the ramp. Analytically, when the initial condition is given, the fundamental solution of the homogeneous heat equation is a Gaussian function [11],

\[
G(x, t) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}.
\]

The solution of the heat equation, \( \Delta f = f_t \), is equivalent to the convolution of a Gaussian filter with the variance \( \sigma = \sqrt{2t} \).

\[
f(x, t) = G(x, t) * f_o(x)
\]

Therefore, the heat equation with homogeneous boundary condition has the smoothing effect on the ramp, ensuring the homogeneity in the region.

Accordingly, by adding overshoot and undershoot near the region boundary in Figure
3.2(b), the boundary condition becomes nonhomogeneous as follows:

\[ f_{xx} = f_t, \quad t > 0 \]  \hspace{1cm} (3.1)

\[ IC : f(x, 0) = f_o(x), \quad 0 < x < L \]  \hspace{1cm} (3.2)

\[ BC : f(0, t) = f_o(0), \quad f(L, t) = f_o(0) + I_{inc}, \]  \hspace{1cm} (3.3)

where \( f_x \) and \( f_{xx} \) indicate the first and second order derivatives of a function, \( f(x) \), respectively, with respect to \( x \). Likewise, \( f_t \) is the first-order derivative of the function with respect to \( t \). Without loss of generality, let a ramp profile be translated and positioned into the direction of \( f_x < 0 \). The initial condition at \( x = L \) in (3.2) is added by the overshoot, \( I_{inc} \). The initial condition and boundary condition guarantee uniqueness and existence of solution. Thus, the solution of non-homogeneous boundary condition becomes:

\[ f(x, t) = \left\{ I_{inc} \cdot \frac{x}{L} + f_o(0) \right\} + \sum_{n=1}^{\infty} C_n e^{-n^2 \pi^2 t / L^2} \sin\left( \frac{n \pi x}{L} \right). \]  \hspace{1cm} (3.4)

The first term in (3.4) is called the steady state solution, while the second term is the transient solution. The coefficients \( C_n \) in the second term are cosine series coefficients and can be determined by the initial condition, (3.2). Since we are not interested in the evolution of the ramp profile according to the heat equation, only the steady state solution is used for the implementation. In summary, as the ramp pixel is located nearer to the region boundary, higher intensity is assigned to the pixel, which allows the maximum local contrast across the region boundary.

3.1.2 Method by constrained linear programming

The second method is formulated by curve fitting such that the slope of the ramp profile at the region boundary is maximized, constrained to the condition of overshoot and the width of the ramp that allows overshoot. Let \( f(x) \) be a cubic equation. Again, suppose the ramp profile is oriented in such a way that \( f_x < 0 \), where the region boundary is located at \( x = 0 \). Refer to Figure 3.3. The position on which the maximum overshoot is placed is given as
Figure 3.3: A model of constrained optimization by cubic curve fitting.

\( x = \pm \alpha \), where \( \alpha > 0 \). Let \( I_1 \) and \( I_2 \) be the intensities of the two adjacent homogeneous non-ramp regions. Then, \( I_1 > I_2 \). The problem is formulated as follows: given \( L, I_1, \) and \( I_2 \), we want to find \( \alpha \) and \( a \) such that the enhanced intensity profile is best approximated to \( f(x) \):

\[
\begin{align*}
  f_x &= a(x - \alpha)(x + \alpha) = a(x^2 - \alpha^2) \\
  \Rightarrow f(x) &= \frac{a}{3}x^3 - a\alpha^2 x + I_m, \quad (3.5)
\end{align*}
\]

where \( I_m = \frac{(I_1 + I_2)}{2} \).

The objective of the problem is that, given overshoot, \( I_{inc} \), and the ramp width, \( L \), we want to maximize the slope of the ramp profile, \( f_x(0) \). It leads to:

\[
\begin{align*}
\min -f_x(0) &= a\alpha^2 \\
\text{s. t.} \quad f(\alpha) &= \frac{a}{3}\alpha^3 - a\alpha^3 + I_m \\
&= I_{min}, \\
I_{min} &= I_2 - I_{inc}.
\end{align*}
\]

This problem is infeasible since the constraints are not bounded. For the geometrical interpretation, as the ramp slope goes steeper, the cubic curve changes drastically with
respect to $x$. This phenomenon could produce strong overshoots inside the small portion of the ramp near the region boundary, called the ringing effect. Therefore, by introducing another constraint such that $\alpha$ lies further than half the length of the ramp width from the region boundary, the constrained optimization problem becomes feasible:

$$\begin{align*}
\min v \\
\text{s. t. } v &= -a\alpha^2 \\
a\alpha^3 &= \frac{3}{2} \left( \frac{I_1 - I_2}{2} + I_{inc} \right) \\
\alpha &\geq \frac{L}{2}
\end{align*}$$

The maximized ramp slope results in the sharpening effect.

### 3.1.3 Comparison of two methods

The two models presented above produce different shapes of a ramp profile by sharpening effects. The model inspired by the heat equation builds linearly increasing overshoot, which reaches the maximum value at the region boundary. As a result, the slope at the transition of one region to another will be very steep, where the highest contrast is allowed to cause the sharpening effect.

On the other hand, the model formulated by the constrained optimization allows gradual change of intensity between adjacent regions depending on the value of the desired ramp width parameter. Although achieving the same overshoot, $I_{inc}$, and the same ramp width, $L$, the two models will transfer different degrees of sharpness in a ramp.
3.2 Experiments

3.2.1 Implementation of the proposed models

We demonstrate our technique using two different types of structure enhancement models with two parameters – desired ramp width and overshoot. The proposed models have been illustrated in the 1-D continuous domain. It represents a cross section of ramp at a region boundary. For the implementation in 2-D discretized domain, pixel-wise update is performed for the pixels inside ramps only. Region boundary is given from the segmentation algorithm in [7], and ramp estimation is performed accordingly. Now, our models require the given ramp width from the input image and the true contrast between two adjacent regions for each ramp pixel.

According to our assumption, a region is formulated as a contiguous homogeneous set of pixels surrounded by ramp discontinuities. Therefore, a contour of the homogeneous non-ramp region can be estimated. For each ramp pixel, ramp width is estimated to be the summation of the closest distance from the non-ramp contour to the current pixel, and that from the region boundary.

The value of overshoot is determined by the true contrast between adjacent regions. True contrast is obtained as a difference between the best approximation of intensities of the two neighbor regions. Let the intensity of the region that the current ramp pixel belongs to be $I_1$. Let the intensity of the neighbor be $I_2$. Then, the true contrast is $|I_1 - I_2|$. For each ramp pixel, a kernel whose center is the ramp pixel itself is used to estimate $I_1$. The size of the kernel is chosen to be twice the distance from the non-ramp contour. Therefore, the size of the kernel is smaller as the ramp pixel is physically closer to the non-ramp contour. The size of the kernel ensures the existence of pixels that lie on the non-ramp region within the kernel. The value, $I_1$, is computed as the weighted average of the chosen pixels whose weights decrease as the distance from the current ramp pixel increases. We use Gaussian distribution to determine the weights. Intensities of spatially close pixels have higher similarities in intensity by the nature of images. Therefore, this approach of weighted average guarantees a good region intensity approximation. Similar computation
is performed for the estimation of $I_2$.

Two real positive numbers are received as input parameters to the proposed systems. One parameter is the desired ramp width ratio, $\mu_1$. The desired ramp width $\alpha$ is determined by the multiplication of the parameter and the given ramp width, $L$, in the input image: $\alpha = \mu_1 \cdot L$. In the diffusion method, the range of $x$ in (3.2) varies within $0 < x < \alpha$, and the boundary condition in (3.3) changes to $f(\alpha, t) = f_0(0) + I_{inc}$. In the curve fitting method, $\alpha$ is the pixel location where the maximum overshoot lies. Therefore, the output ramp width will vary depending on the given conditions of ramps.

The other input parameter $\mu_2$ is a ratio of the desired overshoot to the true contrast of neighboring regions. Therefore, the maximum overshoot, $I_{inc}$, along the cross section of a ramp is $I_{inc} = \mu_2 \cdot |I_1 - I_2|$. The input number is desired to be larger than the Weber’s ratio since it is the visibility threshold.

The proposed method is a non-iterative method like filtering. It is performed once for one pixel, and only ramp pixels are targeted for the enhancement. The computation of homogeneous region intensities is similar to the convolution with an adaptively selected kernel, whose size is approximately $5 \times 5$ matrix by experiments. However, the ramp width calculation could be expensive for large ramps since computation of distances from all ramp pixels inside the ramp is required to find the minimum distance from comparison. The computation time is approximately proportional to the size of the image. For a color image, each R, G, and B channel is exploited in the same manner.

We present the outputs of our proposed techniques. First, outputs of a synthetic image are shown in Figure 3.4. An image in Figure 3.4(i) is blurred by a $3 \times 3$ averaging filter. Here, the ramp width ratio is 1 so that the ramp width to be updated, $\alpha$, is chosen to be the given ramp width, $L$. Both the cubic curve fitting and the diffusion method sharpen the ramp width. However, the cubic curve fitting method has smooth transition across regions due to the characteristics of the curve polynomial equation. The overshoot for the cubic curve fitting is placed at the borderline between ramp and non-ramp parts inside the region. The diffusion method reduces the ramp width to be very sharp. The overshoot ratio parameter is chosen as 0.1. As shown from the ramp intensity profiles in 3.4(d) and 3.4(f),

25
Figure 3.4: Experiments on a synthetic image: (a) A synthetic image blurred by $3 \times 3$ averaging matrix. The bar indicates the location of the intensity profile. (b) Intensity profile across a ramp in (a). (c) Sharpened by cubic curve fitting. (d) Intensity profile across a ramp in (c). (e) Sharpened by diffusion equation. (f) Intensity profile across a ramp in (e). (g) Sharpened by adaptive unsharp masking. (h) Intensity profile across a ramp in (g). (i) The original image. (j) Intensity profile across a ramp in (i).

the ramp profile is sharpened and overshoot is added near the region boundary. Therefore, the manipulation of these two parameters will enhance the look of the image as in Figure 3.4(c) and Figure 3.4(e). In addition, the unsharp masking method is compared for the same input. Artifacts are introduced along edges while the ramp width is preserved. As mentioned previously, overshoot effects could become artifacts if the edge contrast is too high.

Next, examples of real images are demonstrated. Figure 3.5 shows an image of a dark circle with the input and outputs, the input ramp map, and the updated ramp map for different widths and overshoots. We can manipulate various ramp widths to be updated accordingly for the desired effects. For the cubic curve fitting method, region boundaries become smoother as the ramp width to be updated is chosen to be larger. In the diffusion method,
Figure 3.5: Experiments on a real image with a dark circle: (a) The original image. (b) Ramp map of the original image. (c) Output sharpened with no overshoot. (d), (e), and (f) Respectively, outputs by cubic curve fitting, diffusion method, and pixels updated with chosen ramp width with ramp width ratio 2 and the overshoot ratio 0.1. The same for (g), (h), and (i), with ramp width ratio 0.5 and the overshoot ratio 0.05.

overshoot increases more gradually for larger ramp width.

Other examples of real images are in Figure 3.6.

3.2.2 Discussions

We have presented an approach that manipulates different ramp widths and overshoots in the vicinity of region boundaries to bring edge sharpening effects to the image structure in digital photographs.

Our method is unique because adjustments of parameters in two proposed models result in various shapes of ramp profiles. It allows the edge sharpening effects along the region boundary for structure enhancement.

However, the performance depends heavily on the accuracy of segmentation outputs. In areas corrupted by noise or with too much texture, segmented region boundaries are not exact. As a solution, prior to applying our enhancement technique, we could perform
Figure 3.6: Experiments on a real image, Sculpture: (a) The original image with different ramp width ratio $\alpha$ for two proposed methods. (b) Output sharpened by the diffusion method with $\alpha = 1$. (c) Output sharpened by the curve fitting method with $\alpha = 1$. (d) Ramp map.
boundary corrections by minimizing the curvature along the boundary and maximizing the closeness factor of the pixel intensity to the homogeneous region to determine its membership to a region.

The future direction of this research goes to the automated enhancement for human visual perception. Other existing adaptive sharpening techniques require input parameters to obtain the desired output. For example, Wang et al. in [12] uses a modified gradient input in order to obtain the desired structural enhancement effects. Moreover, the adaptive unsharp masking method requires the mid- and high-frequency ranges to be determined as well as the emphasis degree of the sharpening effects on the selected frequency range.

By the comparison of different values in overshoot and undershoot, we need to find the maximum overshoot allowed so that human perception does not “perceivably” notice the difference as overshoot artifact. From the real image experiment, I have found that the range of the overshoot ratio parameter is approximately 0.02 to 0.2, where 0.02 corresponds to the Weber ratio. Perceivably allowable overshoot differs depending on the richness of textures near the region boundary. In addition, different ramp width ratios will result in different output ramp profiles. This parameter should be determined depending on the characteristics of the input image. For example, clouds tend to have large ramps, and small ramp width ratio could bring crispening effects along the region boundary of the clouds. For the automatic system, those parameters should be learned to determine the values of the desired ramp width and overshoot.

By learning the relationship between the blurriness of input image structure, the desired ramp widths and overshoot, we will be able to automatically enhance the image structure from the fluctuation created by autofocus or motion blur, which will reduce the user work in the art of photography.
In the previous chapter, we have created an edge sharpening technique with two input parameters, overshoot and ramp. There is no absolute scale to measure the sharpness quality. We are trying to make universal standards by learning the parameters from the human raters. Perception of sharpness for different parameter values is determined by human subjects. They rank-order images enhanced by different values of parameters according to how pleasing they find the images. We assume that all ratings are from the same distribution. In other words, individuals rate their preferred perceived sharpness to similar, universal standards. By learning from human raters, our goal is to implement an automatic data-driven edge sharpening engine.

4.1 Experimental Procedure

An image is enhanced by two parameters, an overshoot and ramp width. The objective in this section is to select the two parameter values that enhance an image so that the output is most pleasing to common people.

It is a ranking problem with a set of possible ratings, \(\{1, 2, \ldots, 12, 13\}\). We have a training sample of size \(m\), \(S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}\). \(x_i \in \mathcal{X}\) is a feature vector of a datapoint, and \(y_i \in \mathcal{Y}\) is its corresponding rank, \(1 \leq y_i \leq 13\), where 1 is the highest rank and 13 is the lowest rank. Given data under an image category, we learn a ranking function \(f : \mathcal{X} \rightarrow \mathcal{Y}\). After learning the ranking model, we obtain a ranker \(h\) and an ordered list of enhancement parameters under for each image category.

The training data is obtained from the survey. With 13 different enhanced output images,
15 image sets and \( N \) number of subjects participating in the survey, we have \( m = N \times 15 \times 13 = 273N \) responses under one image category.

From the classifier learning, we expect different ordered lists of enhancement parameters depending on the image categories – textures, color RGB, and faces. Test data are collected under the same categories but with different images from those in the training data categories. For each test image, we apply the learned classifier, and 13 enhanced images are ordered from rank 1 to rank 13 by using the ranking function \( f(\cdot) \).

Later, we verify the output of testing data from the human raters. We conduct the survey with the testing data in the same fashion as the one with the training data. If the two results reach consensus by agreeing on the same ordered list, then it confirms the robustness of our approach in the edge sharpening enhancement.

The sharpening effects are determined by two parameters, overshoot and ramp width. Overshoot controls the local intensity contrast in ramp area of regions. It controls the proportion of additional intensity to be added or subtracted by measuring the local intensity contrast. Therefore, the region boundaries become more pronounced. Four values \( \{0.2, 0.4, 0.6, 0.8\} \) were used as the overshoot parameter. Any value higher than 0.8 tends to strongly sharpen edges, so they look unnatural with halo effects. On the other hand, the ramp width ratio is a spatial component that allows the overshoot to be added around the ramp area. Three values \( \{1, 2, 4\} \) were used such that overshoot lies on the area that is the value times the original ramp widths from the region boundaries. The overshoot is controlled to be at its highest value near the region boundary, and its intensity gradually decreases as it goes farther from the boundary. The homogeneous non-ramp area is not touched. Therefore, we produce a total of 13 images: 12 differently enhanced images and the original image.

The two parameters are to be learned from the data in different categories of texture, color, and faces. The texture category is designed to have four classes according to the size of texture elements. Texture elements are defined as repeating objects that appear, regularly or nearly so, in the image. The sizes of texture elements (texel) are subjectively determined to be 8, 25, 64, and 144 texels in images of size 256 \( \times \) 256 for respective texture classes, as
Figure 4.1: Image database: four classes of texture levels. (a) Class 1 has approximately 9 texels within an image. (b) Class 2 has approximately 25 texels within an image. (c) Class 3 has approximately 64 texels within an image. (d) Class 4 has approximately 144 texels within an image.

shown in Figure 4.1. The class with 8 texels is the coarsest while the one with 144 texels is the most detailed. Each class contains 15 different images such as jelly beans, fruit, woolen sweater, flower, brick, etc. Detailed information could exist within the texture element, but the largest repeating object is considered a texel. The four texture classes are controlled to be zoomed differently from the same scenes. Each image is enhanced differently to form a set of 13 images as mentioned above.

The color categories have three classes – red, green, and blue. The image data is collected by modifying the histogram of a certain RGB channel to have a dominant color. The classes have the same 15 image sets with different colors. This experiment is based on the assumption that allowable overshoot could differ for different colors with the same image structure. In addition, we are seeking the maximum overshoot such that the color of a region perceived is not too different from the original color.

Lastly, the face class contains frontal faces with variation only in background. Faces are
focused and centered in photography or portraits. The desired output in this class would be to smooth skin and sharpen eyes. Here, we aim to study the tendency of parameters that can be applied to the entire face. The image categories of RGB color and faces are shown in Figure 4.2.

Now, the set of 13 images needs to be ordered from the highest to the lowest quality. We conduct a survey to rank-order the images. Each image has to be ranked by giving a score ranging 1 to 13, the best score being 1 with no ties allowed. The survey is designed to sort images by merge sort. The subjects are common people because they are the potential users of this method. From the collective data obtained by individuals, we attempt to establish the standardized version of optimal parameters for the most “pleasing” images under the image data categories.

Sorting is done according to the Mergesort algorithm. The survey is mainly conducted by pairwise comparison. Once the sublist is sorted by pairwise comparison, we show all the
images from the sublist in order. If the subject is satisfied with the order, he or she confirms it and goes to the next round of pairwise comparison in sublist ordering or merging of two ordered sublist. In the merge step, subjects are shown the sublist set of sorted images, and they can change the image order by drag-and-drop method if they are not satisfied with the order. The ordering is robust from the feedback in the merge steps. In sorting 13 images, merge sort has on average $13 \log 13$ steps of comparison.

Subjects are presented with images of the same objects. Some are higher in quality than others; i.e., they are crisper, sharper and overall better photographs. They are asked to rank-order the images - from rank 1, which the subject feels has the best quality, to rank 13, which is the worst. They have approximately 20-30 steps to complete one set of enhanced images in 5 minutes. If the presented images are hard to distinguish in quality, they are asked not to spend too much time selecting.

The sorting method is composed of two parts. One is pairwise comparison. This method is used for merging two subsets. The other is a feedback step to confirm the image order of merged sets by dragging and dropping images as shown in Figure 4.3. The website is designed with the PHP and Javascript, programming languages.

In the end, subjects confirm the image ranking. Image 1 has the highest quality while image 13 has the lowest.

4.2 Observers

We used Amazon Mechanical Turk (Mturk) to conduct the survey. It provides a global workforce for tasks requiring human intelligence. With this internet interface, we collected data from 800 subjects. Participants located all over the world allow a researcher to approach a universal standard. Each subject was given a randomly chosen set of enhanced images and a certain amount of time. A big population of participants represents a potential solution to the optimization problem of the best image enhancement.
Figure 4.3: (a) Pairwise comparison. (b) Drag-and-drop step.
CHAPTER 5

RESULTS AND RANK-ORDER CLASSIFICATION

We have designed a learning-based method of edge sharpening. The histograms in Figure 5.1 are raw data of class 1 of 8 texels, taken from the survey performed using Mturk. Each histogram shows the number of votes for the 13 ranks for each of 13 images, each with different parameter values.

From the training data, we will estimate the ranks of the parameters that best enhance a given test image. We will especially estimate relative rankings that can be similarity measures of preferred parameters by human sharpness perception. The learning involves a ranking problem, where subjects rank-order output images sharpened by different values of parameters. The ranking problem is a supervised multi-class classification [13].

5.1 Notation and Problem Formulation

Suppose that we have a training sample, \( S = \{ (x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m) \} \), where \( m \) is the number of data samples, \( x_i \in \mathcal{X} \) is an \( n \)-dimensional feature vector of a training datapoint, \( x_i \in \mathbb{R}^n \), and \( y_i \in \mathcal{Y} \) is the rank of \( x_i \). One is the highest rank and \( m \) is the lowest rank, and \( 1 \leq y_i \leq m \). Given \( S \in \{ (x_i, y_i) : x_i \in \mathcal{X}, y_i \in \mathcal{Y}, i = 1, 2, \ldots, m \} \), we learn a ranking function \( f : \mathcal{X} \to \mathcal{Y} \).

For linear estimation, we learn the linear projector \( h \) and thresholds, \( \theta_1 < \theta_2 < \ldots < \theta_{m-1} \). The thresholds are the relative rankings. The ranking function \( f(x) \) determines the
Figure 5.1: Raw data of class 1. Thirteen cases indicate 13 images, respectively: 12 enhanced images and 1 original image. For each image, the histogram of ranks is obtained by the survey.
ranks of a given feature vector

\[
f(x) = \begin{cases} 
1 & \text{if } h^T x < \theta_1, \\
i & \text{if } \theta_{i-1} < h^T x < \theta_i, \ i \neq 1, m, \\
m & \text{if } \theta_{m-1} < h^T x.
\end{cases}
\]

In learning the function \( f(x) \), we allow a margin of \( \epsilon \). For any \( x_i \) and \( y_i \), we have

\[
\theta_{y_j - 1} + \epsilon < h^T x_j, \ 1 < y_j \leq m.
\]

Therefore, the problem becomes

\[
\min \quad \frac{1}{2} ||h||^2 \\
\text{s. t. } h^T x_j < \theta_{y_j}, \ \forall 1 \leq y_j < m \\
h^T x_j > \theta_{y_j - 1} + \epsilon, \ 1 < y_j \leq m,
\]

which is formulated as a Lagrangian of minimizing \( L_d \)

\[
L_d = \frac{1}{2} ||h||^2 + \sum_{j=1}^{m-m_m} \gamma_j^+ (h^T x_j - \theta_{y_j}) + \sum_{j=m+1}^{m} \gamma_j^- (\theta_{y_j - 1} + \epsilon - h^T x_j) \\
- \sum_j \gamma_j^+ - \sum_j \gamma_j^-
\]

under the constraints that \( \gamma_j^+, \gamma_j^- > 0 \). The term \( m_i \) refers to the number of elements with rank \( i \).

For nonlinear data, a nonlinear mapping \( \phi \) is used with a kernel, \( K \), to embed the data points onto a linear subspace. A radial basis function is used for \( \phi \), and \( K(\cdot, \cdot) = \phi(\cdot)^T \phi(\cdot) \).

The dual of the Lagrangian above becomes

\[
L_d = -\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\gamma_i^--\gamma_i^+)(\gamma_j^- - \gamma_j^+)K(x_i, x_j) - \sum_j \gamma_j^+ - \sum_j \gamma_j^-
\]
This minimization problem has a numerical solution by either the perceptron or the Winnow method.

5.2 Data Analysis

The data set has three big categories: texture, color, and faces. The texture category is divided into four classes from the coarsest to finest, and the classes are designed to have the same patterns of 8, 25, 64, and 144 repeating objects, respectively. Let the texture class with 8 repeating objects be texture class 1, the one with 25 objects be texture class 2, the one with 64 objects be texture class 3, and the one with 144 objects be texture class 4. Edges are sharpened by two parameters: overshoot parameter with 4 values, \{0.2, 0.4, 0.6, 0.8\}, and ramp parameter with 3 values, \{1, 2, 4\}. Including the original image, 13 images are ranked in total.

The feature vector is \( x_i \in \mathbb{R}^4 \) and \( i \in \{1, \ldots, m\} \), where \( m \) is the number of data samples. \[
\begin{pmatrix}
\alpha \\
\beta \\
p_1 \\
p_2
\end{pmatrix}
\]

The value of \( \alpha = \sqrt{\frac{\sum_{k=1}^{3} \sum_{u=1}^{n_u} \sum_{v=1}^{n_v} (I_i(u,v,k) - I_o(u,v,k))^2}{3n_u n_v}} \) quantifies the difference between the enhanced image, \( I_i \), and the original image, \( I_o \), throughout the image coordinates \( u \) and \( v \) and three color RGB channels. The size of an image is \( n_u \times n_v \). \( \beta \) is the number of ramp pixels enhanced normalized by the total number of pixels. Parameter \( p_1 \) is an overshoot rate, descritized so that \( p_1 \in \{0.2, 0.4, 0.6, 0.8\} \). Parameter \( p_2 \) is a ramp width.
rate, $p_2 \in \{1, 2, 4\}$.

The ranking problem is a multi-class classification. We learn the linear projector $h$ and thresholds, $\theta_1 < \theta_2 < ... < \theta_{12}$. The thresholds indicate the relative ranking. Then, the ranking function $f(x)$ determines

$$f(x) = \begin{cases} 
1 & \text{if } h^T x < \theta_1, \\
 i & \text{if } \theta_{i-1} < h^T x < \theta_i, \ i \neq 1, 13, \\
13 & \text{if } \theta_{12} < h^T x.
\end{cases}$$

In learning, we allow a margin of $\epsilon$ such that for any $x_i$ and $y_i$ we have

$$\theta_{y_j - 1} + \epsilon < h^T x_j, \ 1 < y_j \leq 13.$$ 

Therefore, the problem is formulated as

$$\min \quad \frac{1}{2} \|h\|^2$$

s. t. $h^T x_j < \theta_{y_j}, \ \forall 1 \leq y_j < 13$

$$h^T x_j > \theta_{y_j - 1} + \epsilon, \ 1 < y_j \leq 13$$

The relative rankings, $\theta_i$, of images, $i = 1, 2, ..., 13$, in the data classes are shown in Figure 5.2 and Figure 5.3. The rankings are scaled to (0,1), where 0 is the lowest rank and 1 the highest. The original image, $x_{13}$, always has the lowest rank in all image classes. Other images, $x_j, j = 1, 2, ..., 13$, are located on average 0.5 away from $x_{13}$. The data indicate that enhanced images are preferred to the original image. Figures 5.4, 5.5, 5.6, and 5.7 show the estimated ranks of enhanced images. In the next chapter, we will analyze the effects of overshoot parameters and ramp parameters.
Figure 5.2: Scaled relative rankings of enhanced images in training data and test data in (a) class 1, (b) class 2, (c) class 3, (d) class 4.
Figure 5.3: Scaled relative rankings of enhanced images in (a) class 5, (b) class 6, (c) class 7, (d) class 8.
Figure 5.4: Estimated ranks in class 1. Rank 1 is the most preferred, and rank 13 is the least preferred.
Figure 5.5: Estimated ranks in class 4. Rank 1 is the most preferred, and rank 13 is the least preferred.
Figure 5.6: Estimated ranks in class 5. Rank 1 is the most preferred, and rank 13 is the least preferred.
Figure 5.7: Estimated ranks in class 8. Rank 1 is the most preferred, and rank 13 is the least preferred.
CHAPTER 6
STATISTICAL ANALYSIS OF ORDINAL DATA

6.1 Statistical Analysis by Analysis of Variance

6.1.1 Analysis of variance (ANOVA)

An important technique for analyzing the effect of categorical factors on a response is analysis of variance (ANOVA) [14]. The ANOVA tests whether there are any differences between the groups with a single probability distribution. The null hypothesis tested is that all groups are samples from populations having the same normal distribution. There are assumptions underlying analysis of variance. We must have independence between groups. The sampling distributions of sample means must be normally distributed. Lastly, the groups should come from populations with equal variances.

Testing means is related to variance. The assumptions that underly ANOVA follow from the F-ratio. Let $MS_{within}$ be the variance between samples. Let $MS_{bet}$ be the average of the sample variances. F-ratio is defined as

$$F = \frac{MS_{between}}{MS_{within}}.$$ 

$MS_{between}$ and $MS_{within}$ are weighted if the sample sizes are different. $MS_{between}$ is called variation due to treatment, and $MS_{within}$ is called variation due to error. F-ratio indicates the ratio of two independent estimates of the variance of a normal distribution. The ANOVA test depends on the fact that $MS_{between}$ can be influenced by population differences among means of the several groups. Since $MS_{within}$ compares values of each group to its own group mean, it does not depend on group means. $MS_{between}$ and $MS_{within}$ can be computed
by the sum of squares, degree of freedoms, and mean squares. The term, main effect, indicates that an input variable has an effect on the output. The number of main effects equals the number of input variables. The term, interaction, means the effect of an input variable depends on the particular level of the other input variable. The main effect of a variable is significant if the null hypothesis is not true. The p-value, computed from the F-ratio, determines the significance of the main effect or interaction.

6.1.2 Data analysis

To analyze effects of the two parameters, overshoot and ramp, two-way ANOVA is used to test whether the images enhanced by the two independent parameters have similar centroids.

Table 6.1 shows tests of within-subject effects of the significant parameters. Sig. (statistical significance) in the last column indicates p-value, which means the probability of an outcome is as extreme as the observed one provided that the null hypothesis is true. If the p-value is less than 0.05, a null hypothesis is rejected. In class 1, the F-test is 64.012 with degree of freedom (within) 3. Since the p-value of overshoot is 0.012 < 0.05, its main effect is significant. However, the main effect of ramp is not significant in class 1, which indicates that ramp does not affect the ranking in sharpness perception. In class 2, both overshoot and ramp are significant. In class 3, the main effect of overshoot is significant while that of ramp is not significant by p-value = 0.057 > 0.05. In class 4, both the main effect of overshoot and that of ramp are significant. In summary, ranking in classes of different texels (texture elements) is affected by both overshoot and ramp parameters.

In classes 5-7, main effect of overshoot is significant. However, interaction between overshoot and ramp happens by p-value = 0.006 < 0.05. Interaction indicates that overshoot and ramp may not be independent variables. In summary, in classes of RGB color, the overshoot parameter contributes to the ranking of the sharpness perception while the ramp parameter does not.

In class 8, the class of faces, both overshoot and ramp affect the ranking.
Table 6.2 shows the significant parameter values of overshoot and ramp in the tests of within-subject contrasts. In class 1, the p-value = 0.007 < 0.05 indicates that the change in overshoot from 0.2 to 0.8 is significant in determining ranks of images. In class 3, change in overshoot from 0.2 to 0.8 is significant in ranking, so is the change in ramp from 1 to 3. In class 4, both overshoot and ramp values have significant effects on ranking similar to class 3. Therefore, we observe that in texture classes, the overshoot parameter tends to be significant, and the ramp parameters show significance across the levels of ranking.

In classes 5-7, classes of RGB color, the change in overshoot parameter is significant in determining ranks. In addition, we observe joint effects of overshoot and ramp on ranking. In class 8, the faces class, the p-values < 0.05 for the contrasts of 0.2 vs. 0.8, 0.4 vs. 0.8, 0.6 vs. 0.8 indicate that the change from 0.2-0.6 to 0.8 in overshoot is significant in change of ranking, as is that of 2 vs. 3 in ramp parameter. Interaction between overshoot and ramp is observed. However, since p-value ≃ 0.05, it is still valid to assume that overshoot and ramp are independent variables.

In summary, we have analyzed the effects of overshoot and ramp parameters in classes. Both overshoot and ramp parameters significantly affect ranking in classes of different texels. In classes of RGB color, the overshoot parameter has the biggest effect on the
### Table 6.2: Tests of within-subject contrasts

<table>
<thead>
<tr>
<th>Source</th>
<th>Effect</th>
<th>Level</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>class 1</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1</td>
<td>7.696</td>
<td>.007</td>
</tr>
<tr>
<td>class 3</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1</td>
<td>4.700</td>
<td>.033</td>
</tr>
<tr>
<td>class 3</td>
<td>ramp</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>5.154</td>
<td>.026</td>
</tr>
<tr>
<td>class 4</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>9.358</td>
<td>.003</td>
</tr>
<tr>
<td>class 4</td>
<td>ramp</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>3.633</td>
<td>.060</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>29.216</td>
<td>.000</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>1 vs. 3</td>
<td>11.91</td>
<td>.001</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>2 vs. 3</td>
<td>8.997</td>
<td>.004</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>1 vs. 3</td>
<td>11.389</td>
<td>.001</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>18.462</td>
<td>.000</td>
</tr>
<tr>
<td>class 5</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>9.762</td>
<td>.002</td>
</tr>
<tr>
<td>class 6</td>
<td>Overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>27.645</td>
<td>.000</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>2 vs. 3</td>
<td>8.854</td>
<td>.004</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>1 vs. 3</td>
<td>6.060</td>
<td>.016</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>2 vs. 3</td>
<td>10.366</td>
<td>.002</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>1 vs. 3</td>
<td>38.311</td>
<td>.000</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>10.172</td>
<td>.002</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>15.027</td>
<td>.000</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>15.794</td>
<td>.000</td>
</tr>
<tr>
<td>class 6</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>8.315</td>
<td>.005</td>
</tr>
<tr>
<td>class 7</td>
<td>Overshoot</td>
<td>0.2 vs. 0.8</td>
<td>1 vs. 3</td>
<td>36.059</td>
<td>.000</td>
</tr>
<tr>
<td>class 7</td>
<td>overshoot</td>
<td>0.2 vs. 0.8</td>
<td>2 vs. 3</td>
<td>23.833</td>
<td>.000</td>
</tr>
<tr>
<td>class 7</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>1 vs. 3</td>
<td>9.504</td>
<td>.003</td>
</tr>
<tr>
<td>class 7</td>
<td>overshoot</td>
<td>0.4 vs. 0.8</td>
<td>2 vs. 3</td>
<td>7.918</td>
<td>.006</td>
</tr>
<tr>
<td>class 7</td>
<td>overshoot</td>
<td>0.6 vs. 0.8</td>
<td>2 vs. 3</td>
<td>4.009</td>
<td>.049</td>
</tr>
</tbody>
</table>
ranking, and interaction of overshoot and ramp is observed. In the class of faces, both overshoot and ramp parameters are significant. In the next section, we will estimate the marginal probability of getting a rank, $i$, out of 13 items in the two parameter combinations by cumulative logistic regression.

6.2 Cumulative Logit Models for Ordinal Responses

6.2.1 Cumulative logistic regression

Logistic regression is applied to model multivariate binary or categorical outcomes. The logit function, defined as $\text{logit}(p) = \log \frac{p}{1-p}$, is used in logistic regression. Alternatively, cumulative logit models are appropriate for ordinal responses and produce a marginal logistic structure for the individual outcomes [15]. The approach is to specify the probability model for the categorical outcomes in terms of underlying continuous variables. First of all, we consider a binary response variable with univariate logistic regression model. For a binary response variable $Y$, let $\pi(x)$ denote the probability of getting “success” (1) at value $x$. The logistic regression model has linear form for the logit of this probability,

$$\text{logit}\pi(x) = \log \left( \frac{\pi(x)}{1 - \pi(x)} \right) = \alpha + \beta x$$

where $\pi(x)$ is the parameter for the binomial distribution, $x$ is a predictor, and $\alpha$, $\beta$ are unknown regression coefficients.

Therefore,

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}.$$

Now, we model a cumulative logit model for multivariate categorical outcomes. Suppose that we have the $k$ ordered categories with probabilities $\pi_1(x), \pi_2(x), \ldots, \pi_k(x)$. $Y$ denotes the response with values in the range 1, ..., $k$ with the corresponding probabilities given
above. For a category $j$, the cumulative probability is

$$P(Y \leq j) = \pi_1 + ... \pi_j, \ j = 1, ..., J.$$  

We know that $P(Y \leq 1) \leq P(Y \leq 2) \leq ... \leq P(Y \leq J) = 1$.

Now, the cumulative logit function is defined as

$$\text{logit}[P(Y \leq j)] = \log\left[\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right] = \log\left[\frac{\pi_1 + ... + \pi_j}{\pi_{j+1} + ... + \pi_J}\right],$$

where $j = 1, ..., J - 1$.

A model for the cumulative logit of $j$ can be interpreted as a binary logistic regression model. Categories from 1 to $j$ form one category, $(Y \leq j)$, and categories from $j + 1$ to $J$ form the other category, $(Y > j)$. The model

$$\text{logit}[P(Y \leq j)] = \alpha_j + \beta x, j = 1, ..., J - 1$$

has parameter $\beta$ describing the effect of $x$ on the log odds of response in categories 1 to $j$. By using constant $\beta$, the model has an underlying assumption that the effect of $x$ is identical for all $J - 1$ cumulative logits. Therefore, the $J$ curves have shifted marginal probability distributions of the same shape. The category probabilities can be computed as $P(Y = j) = P(Y \leq j) - P(Y \leq j - 1)$.

### 6.2.2 Data analysis

We have two parameters, overshoot and ramp. The Overshoot values are taken from $\{0, 0.2, 0.4, 0.6, 0.8\}$, and ramp are taken from $\{0, 1, 2, 4\}$. We have designed models based on the significance from p-values. For classes of texels, a model is selected such that the overshoot variable is a continuous variable, $x_1$. The ramp parameter is discrete and represented as $x_2 \in R^3$ such that $x_2(0) = [1, 0, 0]^T$, $x_2(1) = [0, 1, 0]^T$, $x_2(2) = [0, 0, 1]^T$, 

52
\( x_2(4) = [-1, -1, -1]^T \). Therefore, the model is

\[
\text{logit}[P(Y \leq j)] = \alpha_j + \beta_1 x_1 + \beta_2^T x_2, j = 1, \ldots, 12.
\]

We solve the model for \( \alpha_j, \beta_1, \) and \( \beta_2 \). Figure 6.1 shows the probabilities that, given a ramp parameter, an image has rank \( Y = j, j = 1, 2, \ldots, J \) in texture class 4 as well as cumulative probabilities. By logistic regression, the model is fitted based on the data of overshoot range of \((0, 1)\). The model predicts that the probability of rank \( j \), and \( P(Y = j | \text{ramp}) \) is a function of overshoot. Figure 6.2 shows the probabilities that given an overshoot parameter, an image has rank \( Y = j, j = 1, 2, \ldots, J \) in class 5 of color red.

Figure 6.2 shows both cumulative probabilities \( P(Y \leq j | \text{overshoot}) \) and the probabilities \( P(Y = j | \text{overshoot}) \). They are functions of a continuous parameter, ramp.

Figure 6.3 shows the probability \( P(Y = 1 | \text{ramp}) \) for classes of texels (class 1, 2, 3, and 4). The overshoot parameter is chosen as a continuous parameter due to its significance by p-value < 0.05. For all ramp values, \( P(Y = 1 | \text{ramp}) \) increases as overshoot parameter increases. However, when \( \text{ramp} = 4 \), \( P(Y = 1 | \text{ramp}) \) reaches the maximum as overshoot increases. When overshoot > 0.7, the probability is almost at maximum. The data indicate that \( \text{ramp} = 4 \) and overshoot > 0.7 are preferred in images of class 1. In class 2, the maximum probability of rank 1 is at \( \text{ramp} = 2 \), and overshoot > 0.5. In class 3, the maximum probability of rank 1 is at \( \text{ramp} = 4 \), and overshoot > 0.3. In class 4, the maximum probability of rank 1 is at \( \text{ramp} = 4 \), and overshoot > 0.5. We have observed that preferred \( \text{ramp} \) and overshoot parameters vary in different texels.

Figure 6.4 shows the probability \( P(Y = 1 | \text{overshoot}) \) for classes of colors (classes 5-7) and class of faces (class 8). The ramp parameter is chosen as a continuous parameter due to its significance by p-value < 0.05.

Table 6.3 shows the most preferred parameters for each class based on the probabilities of rank 1. In class 5, overshoot = 0.6, 2 < ramp < 4 are preferred. In class 6, overshoot = 0.6, 1 < ramp < 4 are preferred. In class 7, 2 < ramp < 4, the maximum at overshoot 0.6. In class 8, at ramp = 4, overshoot = 0.6, with maximum probability of rank 1.
Figure 6.1: (a) Cumulative probability, $P(Y \leq j|ramp)$. (b) Probability $P(Y = j|ramp)$ in class 4.
Figure 6.2: (a) Cumulative probability, $P(Y \leq j | \text{overshoot})$. (b) Probability $P(Y = j | \text{overshoot})$ in class 5.
Figure 6.3: Scaled relative rankings of enhanced images in training data and test data in (a) class 1, (b) class 2, (c) class 3, (d) class 4.

Table 6.3: Tests of within-subject contrasts

<table>
<thead>
<tr>
<th>Class</th>
<th>overshoot</th>
<th>ramp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&gt; 0.7</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>&gt; 0.3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>&gt; 0.5</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.6</td>
<td>&lt; 2</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>7</td>
<td>0.6</td>
<td>&lt; 2</td>
</tr>
<tr>
<td>8</td>
<td>0.6</td>
<td>4</td>
</tr>
</tbody>
</table>
Figure 6.4: Scaled relative rankings of enhanced images in training data and test data in (a) class 1, (b) class 2, (c) class 3, (d) class 4.
6.2.3 Test of cumulative logistic regression

Based on Table 6.3, images were generated so that the best image quality agrees with the perceived sharpness in Figure 6.5 and Figure 6.6.
Figure 6.5: Scaled relative rankings of enhanced images in training data and test data in (a) class 1, (b) class 2, (c) class 3, (d) class 4.
Figure 6.6: Scaled relative rankings of enhanced images in training data and test data in (a) class 5, (b) class 6, (c) class 7, (d) class 8.
In this work, we have explored an edge enhancement approach by collecting human data to produce statistical consensus in sharpness perception. As a region-based approach, we sharpen the region boundaries, so we enhance the perceptual image structure that is composed of regions of homogeneous intensity. Ramp is estimated as the vicinity of the region boundaries. The edge sharpening technique is controlled by two parameters, overshoot and ramp width. The two parameters exploit local contrast and width of edges.

We attempt to make an automated sharpening technique. We have collected an image database based on texture granularity, color, and faces. By conducting experiments with human subjects, we have developed the best enhancement for these different criteria by finding the best combination of parameter values.

The experiment was conducted via Mechanical Turk provided by Amazon, so that the population sample is big enough. The experiment was designed as a web site, and subjects were asked to rank-order images according to how pleasing the images were, based on the sharpness perception.

We assume that individuals rate their preferred perceived sharpness according to similar, universal standards. By learning the parameters from the human raters, global standards for perception of sharpness can be established with the parameter values.

Then, we analyzed the data obtained from the experiments with different approaches. First, relative ranking was estimated based on rank-order classification. The ranking problem is a supervised multi-class classification. Second, the main effects of two parameters, overshoot and ramp, on the ranking were investigated by analysis of variance. Lastly, probabilities of getting ranks for different values of parameters were computed by cumulative
logistic regression.

Images with texture granularity show significant effects of overshoot on ranking in the sharpness perception. Images in color and face classes show significant effects of both overshoot and ramp on ranking. The dependency between overshoot and ramp is not significant. Therefore, it is valid to assume that overshoot and ramp are independent parameters.

For future work, we need a rigorous validation process. We will ask human subjects to rank test images, and compare the results with the learned parameters.

We have studied the perceived sharpness in real images. The images were categorized based on texture levels, color, and an object of focus, that is, faces. However, real images may contain other variables that could affect the human perception, such as noise or the setting in which the experiments take place. We will extend this study by removing the variables. The experiment will be designed from the psychophysical point of view. Especially, we will test the effects of color in determining overshoot values with simple synthetic images. The experiments will take place under fixed environments such as constant lighting and fixed distance between the computer monitor and the human subject. In this way, we will be able to implement a robust, automatic, data-driven, edge sharpening engine.
REFERENCES


