Image Sharpening with Reduced Sensitivity to Noise: A Perceptually Based Approach*

David J. Lieberman and Jan P. Allebach
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907-1285
{dlieberm,allebach}@ecn.purdue.edu

Abstract

Sharpening images using conventional highpass linear filters can result in two serious problems: An increase in the level of image noise and the addition of error which results from clipping pixel values to within the range [0,1]. We have found that these two problems are closely related. By modifying the neighborhood of each clipped pixel, we reduce its perceptibility and reduce image noise.

1 Introduction

It is widely accepted that most observers prefer the appearence of images that have been enhanced via sharpening. Unfortunately, using a highpass linear filter to sharpen an image can dramatically increase image noise. Clearly, high frequency noise power is expected to increase. The subject of this paper is an additional source of image noise: clipping the pixel values of the sharpened image.

Any filter which only uses positive filter coefficients is inherently lowpass. Thus, every highpass linear filter must have negative coefficients. In addition, if the filtering operation is expected to preserve average local tone, then the coefficients must sum to unity. Therefore, the sum of the positive filter coefficients must exceed unity. Now, consider a filter whose positive and negative coefficients sum to P and $-N$ respectively. If it is applied to a graylevel image with support limited to the range $[0,1]$, then the output image may have support throughout the range $[-N,P]$. In a typical implementation, all pixel values below 0 and above 1 are *clipped* to 0 and 1 respectively. Our goal is to reduce the perceptibility of this clipping by adjusting the pixel values in the neighborhood of each clipped pixel. One possible approach is to diffuse [1] the magnitude of the clipped pixel into the neighborhood. However, we choose to employ a model based approach. This technique is reminiscent of the algorithm proposed by Zakhor et al [2]. First, we introduce a visual model and error measure. Next, we show how our error measure is minimized by optimizing our neighborhood adjustment to reduce the perceptibility of the effects of clipping. Then, some implementation issues are discussed. In addition, an interpretation of how this algorithm is expected to impact the image quality is given. Lastly, we provide some experimental results and concluding remarks. Throughout this paper, we use $[\mathbf{m}] = [m,n]^T$ to represent discrete spatial coordinates.

1.1 Visual Model

We model the lowpass characteristics of the human visual system (HVS) with a linear shift-invariant impulse response. This perceptual filter kernel $\tilde{p}[\mathbf{m}]$ is based on the contrast sensitivity function (CSF) of the HVS [3]. Let $f[\mathbf{m}]$ represent the sharpened image and $\tilde{f}[\mathbf{m}]$ represent the perceived sharpened image. Note that the pixels of $f[\mathbf{m}]$ may extend beyond the range $[0,1]$. The relation between $f[\mathbf{m}]$ and $\tilde{f}[\mathbf{m}]$ is given by

$$\tilde{f}[\mathbf{m}] = \sum_{\mathbf{n}} f[\mathbf{n}]\tilde{p}[\mathbf{m}-\mathbf{n}]. \quad (1)$$

1.2 Error Metric

Define $f_r[\mathbf{m}]$ as another version of the sharpened image whose pixels are *restricted* to the range $[0,1]$. Our goal is to compute $f_r[\mathbf{m}]$ to minimize a measure of a perceptually filtered version of the error image $e[\mathbf{m}] = f[\mathbf{m}] - f_r[\mathbf{m}]$. We model the perceived error as

$$\tilde{e}[\mathbf{m}] = \sum_{\mathbf{n}} e[\mathbf{n}]\tilde{p}[\mathbf{m}-\mathbf{n}]. \quad (2)$$

The global error measure to be minimized is the total squared perceived error given by

$$E = \sum_{\mathbf{m}} (\tilde{e}[\mathbf{m}])^2. \quad (3)$$

2 Minimizing Perceived Clipping Artifacts

Under optimal conditions, no two clipped pixels are in close proximity to each other. In addition, an attempt to obscure the visibility of a clipping artifact doesn't require

*Research supported by the Hewlett-Packard Company.

saturating a neighboring pixel value to 0 or 1. Therefore, reducing the visibility of each clipped pixel can be considered individually; we refer to this as unconstrained optimization. When these conditions do not apply, constrained optimization is necessary.

2.1 Unconstrained Optimization

Consider a pixel of a sharpened image whose value exceeds 1. We assume this pixel is neither on nor near the image boundry; but the pixel's index is referred to as the origin for convenience. First, the neighborhood size is selected. For a given size, a unique optimal solution in the minimum-mean-square sense can be computed. Let N be the number of pixels in the neighborhood, and $\hat{\delta}^{(N)}[\mathbf{m}]$ be the kernel whose values represent the neighborhood adjustments which compensate for a unit level of clipping. Our use of this notation is explained in Sec. 4.

Using the unconstrained solution, the effect of each clipped pixel is considered individually; so assume that the only pixel that required clipping is the pixel at the origin. After clipping this pixel and compensating for its effect, the adjusted image is given by

$$f_r[\mathbf{m}] = f[\mathbf{m}] - \triangle\delta[\mathbf{m}] + \triangle\hat{\delta}^{(N)}[\mathbf{m}], \tag{4}$$

where $\triangle = f[\mathbf{0}] - 1$ is the difference between the initial and clipped pixel value, and $\delta[\mathbf{m}]$ equals 1 where $\mathbf{m} = \mathbf{0}$ and 0 otherwise. Notice that $\triangle$ is used to scale $\hat{\delta}^{(N)}[\mathbf{m}]$ because this kernel is assumed to compensate for the effects of a unit level of clipping. The error image $e[\mathbf{m}] = \triangle\delta[\mathbf{m}]$ - $\hat{\delta}^{(N)}[\mathbf{m}]$; so $\tilde{e}[\mathbf{m}]$ is given by

$$\tilde{e}[\mathbf{m}] = \triangle\tilde{p}[\mathbf{m}] - \sum_{\mathbf{n}} \hat{\delta}^{(N)}[\mathbf{m}]\tilde{p}[\mathbf{m} - \mathbf{n}]. \tag{5}$$

Given this correlate of perceived error, the procedure for minimizing E is straightforward: Begin by zero-padding $\tilde{p}[\mathbf{m}]$ to increase its width and height by at least the width and height of the region of support for $\hat{\delta}^{(N)}[\mathbf{m}]$. Then, concatenate the columns of this array into a column vector denoted as $\vec{\mathbf{b}}$. Note that for each $\mathbf{m}$, the term $\hat{\delta}^{(N)}[\mathbf{m}]\tilde{p}[\mathbf{m} - \mathbf{n}]$ represents a shifted and scaled version of this array. So corresponding to each of these terms, shift the non-zero region of support of the array so as to represent that term, and concatenate each of these shifted arrays into a column vector $\vec{a}_i$ where i=1,2,...,N. Here N is the number of pixels in the neighborhood. These vectors are used to construct matrix $\overline{A} = [\vec{a}_1\ \vec{a}_2\ ...\ \vec{a}_n]$. The system of equations $\overline{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is used to compute the coefficients of $\hat{\delta}^{(N)}[\mathbf{m}]$, which are the components of the column vector $\vec{\mathbf{x}}$. This is accomplished by computing the left-inverse, $\vec{\mathbf{x}}_{mmse} = (\overline{A}^T\overline{A})^{-1}\overline{A}^T\vec{\mathbf{b}}$.

For a series of different neighborhoods sizes, the optimal values of $\hat{\delta}^{(N)}[\mathbf{m}]$ are given in Fig. 1. Notice that all four coefficients of Fig. 1(b) have the same value. Similarly, only two distinct coefficient values are used in the 8-nearest-neighborhood (8NN) of Fig. 1(c) and five distinct

	X	.871

(a)

	.277	
.277	X	.277
	.277	

(b)

-.305	.553	-.305
.553	X	.553
-.305	.553	-.305

(c)

-.08	.207	-.287	.207	-.08
.207	-.521	.722	-.521	.207
-.287	.722	X	.722	-.287
.207	-.521	.722	-.521	.207
-.08	.207	-.287	.207	-.08

(d)

Figure 1: Optimal adjustments to be made to the neighborhood of a clipped pixel. The size of the neighborhoods considered include 1(a), 4(b), 8(c), and 24(d) pixels.

values in the 24-nearest-neighborhood of Fig. 1(d). These symmetries are due to the symmetry in $\tilde{p}[\mathbf{m}]$. This perceptual kernel is symmetric about the vertical, horizontal, and both diagonal axes. In addition, the coefficients of each kernel do not sum to unity. By minimizing mean-squared error, one doesn't necessarily guarantee correct average tone. However, the changes in tone are small and tend to decrease as a function of increasing neighborhood size; in Figs. 1(a) through (d), the coefficients sum to 0.871, 1.108, 0.992 and 0.994 respectively. The mean-square perceived clipping error as a function of neighborhood size is shown in Fig. 2. The error measure using 0 coefficients corresponds to the mean-square perceived error which results from a unit level of clipping whose neighborhood is not adjusted. The plotted mean-square perceived errors using 1, 4, 8, and 24 coefficients correspond to the kernels described in Fig. 1(a) through (d).

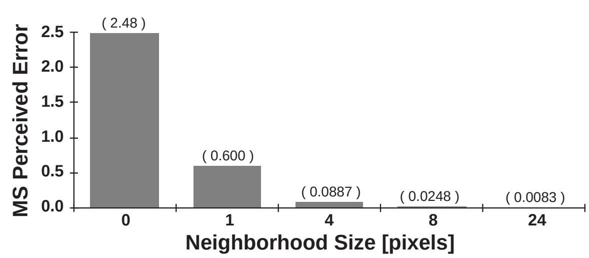


Figure 2: Effect of applying the unconstrained solution on the mean-square perceptual error using neighborhoods of 1, 4, 8, and 24 pixels.

Clearly, complexity increases with the size of the neighborhood whereas the resulting error decreases. Another tradeoff involving neighborhood size is associated with the likelihood that the unconstrained solution is feasible. As the chosen neighborhood size increases, a saturated pixel is more likely to be encountered. This decreases the probability that the unconstrained solution is feasible.

2.2 Constrained Optimization

When the unconstrained solution is not feasible, the constrained solution should be used. Consider the example illustrated in Fig. 3. In Fig. 3(a), the intended output of a sharpening filter is shown. The center pixel's magnitude exceeds one. Clipping its magnitude to one and attempting to apply the unconstrained solution $\hat{\hat{\delta}}^{(8)}[\mathbf{m}]$ results in (b). However, the magnitude of a neighboring pixel exceeds one, so this solution is not feasible. The optimal constrained solution is shown in (c). These solutions can be determined using a procedure similar to the procedure described in Sec. 2.1.

(a)

0.5	0.5	0.5
0.5	1.3	0.81
0.5	0.96	0.5

(b)

.409	.666	.409
.666	1.0	.976
.409	1.126	.409

(c)

.387	.711	.370
.659	1.0	1.0
.481	1.0	.464

Figure 3: The intended sharpened output (a), effect of applying the unconstrained solution (b), and the optimal constrained solution (c).

3 Implementation Issues

In general, reducing the effects of clipping following a sharpening operation may be performed as follows. First, sharpen the image and store each instance of clipping in the error image $e_{clip}[\mathbf{m}]$ defined as

$$e_{clip}[\mathbf{m}] = \begin{cases} 1 - f[\mathbf{m}]\,; & \text{if } f[\mathbf{m}] > 1 \\ 0\,; & \text{if } 0 \leq f[\mathbf{m}] \leq 1. \\ f[\mathbf{m}]\,; & \text{if } f[\mathbf{m}] < 0 \end{cases} \quad (6)$$

When pixels exceeding this range are simply clipped to 0 and 1, the resulting clipped image $f_{clip}[\mathbf{m}]$ is given by

$$f_{clip}[\mathbf{m}] = f[\mathbf{m}] - e_{clip}[\mathbf{m}]. \quad (7)$$

If an unconstrained solution can be applied to the neighborhood of each clipped pixel, then the resulting image $f_r[\mathbf{m}]$ may be computed as

$$f_r[\mathbf{m}] = (f[\mathbf{m}] - e_{clip}[\mathbf{m}]) + (\hat{\hat{\delta}}^{(N)}[\mathbf{m}] * e_{clip}[\mathbf{m}]), \quad (8)$$

where $*$ represents convolution, the first term equals $f_{clip}[\mathbf{m}]$, and the second term accounts for the superposition of all adjustments made upon the neighborhoods of clipped pixels. In general, some of these adjustments will not be feasible. At the position of each unfeasible adjustment, the constrained optimal solution is used to replace the unconstrained solution. Switching between the use of the unconstrained and constrained solutions is inherently nonlinear. This suggests that the order in which the neighborhoods of clipped pixels are adjusted may impact the resulting mean-square function cost. Certainly, an iterative approach may be used to *search* for an ordering whose use lowers function cost compared to other orderings. For simplicity, in this paper we assume that constrained solutions are applied to pixels in raster-scan order, from left to right and from top to bottom.

4 Maneuvering the Clipping Effects into the Perceptual Black Space: An Interpretation

In Sec. 2.1, $\hat{\hat{\delta}}^{(N)}[\mathbf{m}]$ is described as a kernel which is applied to the neighborhood of a clipped pixel to minimize the perceptibility of a unit level of negative clipping error. For this reason, it is convenient to consider the combined impact of the clipping and $\hat{\hat{\delta}}^{(N)}[\mathbf{m}]$ by defining $\hat{K}^{(N)}[\mathbf{m}]$ as

$$\hat{K}^{(N)}[\mathbf{m}] = \hat{\hat{\delta}}^{(N)}[\mathbf{m}] - \delta[\mathbf{m}]. \quad (9)$$

Reducing the perceptibility of clipping suggests the following relation:

$$\tilde{p}[\mathbf{m}] * \hat{K}^{(N)}[\mathbf{m}] \approx 0, \quad (10)$$

for all $\mathbf{m}$. These relations may be combined to show that

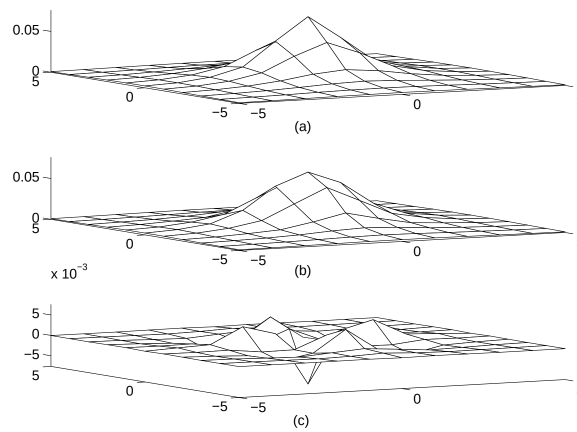


Figure 4: Comparison between $\tilde{p}[\mathbf{m}]$ (a), $\tilde{p}[\mathbf{m}] * \hat{\hat{\delta}}^{(8)}[\mathbf{m}]$ (b), and the perceived difference $\tilde{p}[\mathbf{m}] * \hat{K}^{(8)}[\mathbf{m}]$ (c).

$\tilde{p}[\mathbf{m}] \approx \tilde{p}[\mathbf{m}] * \hat{\hat{\delta}}^{(N)}[\mathbf{m}]$ as illustrated in Fig. 4, for n=8. Thus, we may interperet $\hat{\hat{\delta}}^{(N)}[\mathbf{m}]$ as the kernel which is identically zero at the origin, but which nevertheless perceptually reproduces the impulse response. This explains our choice of notation for the $\hat{\hat{\delta}}^{(N)}[\mathbf{m}]$ kernels. Similarly, we may interpret the clipping and neighborhood adjustment combination given by $\hat{K}^{(N)}[\mathbf{m}]$ as an attempt to maneuver these combined effects into the perceptual black space. Clearly, (10) and Fig. 4(c) suggest that $\hat{K}^{(N)}[\mathbf{m}]$ is perceived as being virtually invisible.

5 Experimental Results

A sample image with added noise appears in Fig. 5(a). The sharpened and clipped image $f_{clip}[\mathbf{m}]$ given by (7) is shown in Fig. 5(b). The effects of clipping introduce localized shifts in tone. These fluctuations are perceived as added noise and change the image's overall average tone. In particular, pixels within darker areas tend to be clipped up to 0 and within lighter areas tend to be clipped down to 1. As a result, clipping tends to increase and decrease the average graylevel in these areas respectively. This reduces contrast and results in a faded appearence. This effect is most evident in the tone shifts throughout face and jaw of Fig. 5(b). In Fig. 5(c), the image $f_r[\mathbf{m}]$ is shown. The effects of clipping are perceptually minimized in this image using the procedure outlined in Sec. 2. This sharpened image largely preserves average local tone and exhibits less noise. Due to the compession used in format conversion, the printed images will not appear as intended. These images in tiff format and the postscript version of this paper are availiable at http://www.ecn.purdue.edu/EISL/pubs/NIP/dlieberm1.

6 Conclusions

Sharpening an image by applying a highpass linear filter is a well known classical approach. Clearly, the level of high frequency noise is expected to increase. However, one shouldn't assume that high frequency noise gain alone is responsible for the poor quality of a sharpened image. In this paper, we have shown that the effects of clipping may be responsible in part for the decrease in image quality. Most of this decrease is due to shifts in average local tone. This is particularly so if the image is noisy and requires strong levels of sharpening. Fortunately, the shifts in tone and the perceptibility of the clipping effects can be largely avoided. This may improve image quality dramatically.

7 Biography

David Lieberman received his B.S. in Electrical Engineering from the University of Wisconsin at Madison in 1986 and his M.S in Electrical Engineering from Purdue University in 1996. He is a Ph.D. candidate in Electrical Engineering at Purdue where he is employed as a research assistant and expects to graduate in Dec. 98. His research interests include model based image enhancement, halftoning, and screen design.

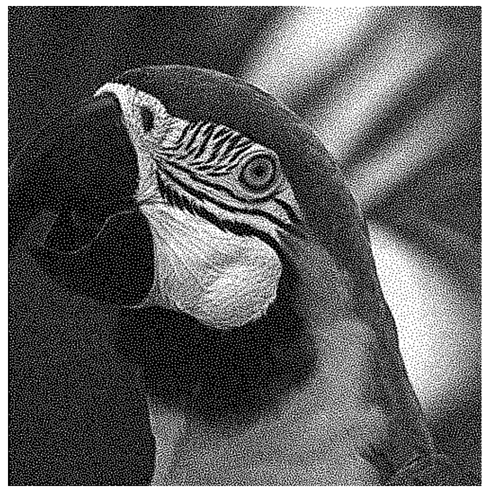
References

[1] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," *Proc. SID*, vol. 17, No. 2, pp. 75 – 77, 1976.

[2] A. Zakhor, S. Lin, and F. Eskafi, "A new class of b/w and color halftoning algorithms," *IEEE Trans. on Image Processing*, vol. 2, pp. 499 – 509, October 1993.

[3] D. Lieberman and J. Allebach, "Model based direct binary search halftone optimization with a dual interpretation," *Proceedings of the 1998 IEEE International Conference on Image Processing*, October 4-7 1998, Chicago, IL.

(a) Cont.-tone sample image.



(b) Sharpened image with clipped output

(c) Sharpened image with reduced perceived clipping

Figure 5: Illustrating the effect of image sharpening followed by ordinary clipping vs. clipping that is perceptually reduced.