**Abstract**

This paper presents a block-based algorithm designed to measure the local perceived sharpness in an image. Our method utilizes both spectral and spatial properties of the image: For each block, we measure the slope of the magnitude spectrum and the total spatial variation. These measures are then adjusted to account for visual perception, and then the adjusted measures are combined via a weighted geometric mean. The resulting measure, \(S_3\) (Spectral and Spatial Sharpness), yields a perceived sharpness map in which greater values denote perceptually sharper regions. This map can be collapsed into a single index which quantifies the overall perceived sharpness of the whole image. We demonstrate the utility of the \(S_3\) measure for within-image and across-image sharpness prediction, for global blur estimation, and for no-reference image quality assessment of blurred images.

**Index Terms**

Sharpness, blur, natural scenes, human visual system, total variation.

1. Introduction

Most consumer photographs contain particular regions which are perceived to be sharper than others. Although the term “sharpness” lacks a precise technical definition, any human can effortlessly point out the sharp regions in an image. Intuitively, a sharp region is one in which fine details are resolvable (high resolution) and in which edges and object boundaries appear to be of high contrast (high acutance). Indeed, most professional photographers attempt to maximize perceived sharpness by using a high-resolution camera and by employing digital retouching to increase acutance (e.g., via unsharp masking).

Yet, despite the ease with which perceived sharpness can be determined by eye, this task remains quite challenging for a computer. The ability to quantify the perceived sharpness of an image, and of regions within the image, can be useful for a variety of image-processing applications. For example, auto-enhancement algorithms can use this information to sharpen images in a spatially adaptive fashion [1]. Similarly, in main-subject detection, objects which are sharp are usually considered candidates of main subjects. Sharpness can also be a useful factor in no-reference image quality assessment (e.g., [2], [3]).

Previous work on sharpness estimation has largely been geared toward determining a single scalar value which quantifies the overall sharpness of an image. A common technique involves measuring the spread of edges. For example, Marziliano et al. [2] measure vertical edges in an image and then estimate overall sharpness based on the average edge width. Ferzli et al. [3] measure edge widths in 8x8 blocks, which are then weighted by a Mean Just-Noticeable Blur factor computed based on properties of the early human visual system (HVS). Shaked et al. [1] measure sharpness in the frequency domain using a technique based on the ratio of high-pass to low-pass energy of the spatial derivative of each line/column. Caviedes et al. [4] first build a block-based sharpness measure using the kurtosis of the DCT of each block, and then average the sharpness values for edge profiles to get the final sharpness metric. A good review of other sharpness/blur estimators can be found in [5].

In this paper, we propose a block-based algorithm which can measure the perceived sharpness of local image regions, and which does not require the presence of edges. Indeed, images commonly contain both edges and textures, and it is often the textures which appear sharper than the edges (e.g., a seashell in the sand). Our measure is based on two factors: (1) a spectral measure based on the slope of the local magnitude spectrum, (2) a spatial measure based on local maximum total variation (TV). Our work draws on the concepts proposed by Shaked et al. [1] and Field & Brady [6] (spectral aspect) and on the concept of total variation proposed in [7] (spatial aspect).

It is well known that the attenuation of high-frequency content can lead to an image which appears blurred. One way to measure this effect is examine the image’s magnitude spectrum \(M(f)\) which is known to fall inversely with frequency-i.e., \(M(f) \propto f^{-\alpha}\), where \(f\) is frequency and \(-\alpha\) is the slope of the line \(\log M \propto -\alpha \log f\). For natural scenes, \(\alpha\) is typically in the range 0.7 – 1.6 [8]. Researchers
have argued that the HVS is tuned to this natural-scene spectrum [9]. Our spectral measure of sharpness follows from the argument put forth by Field & Brady [6]: An image region whose spectrum exhibits a slope factor of $0 \leq \alpha \leq 1$ will appear sharp, whereas regions with $\alpha > 1$ will appear progressively blurred as $\alpha$ increases (see Figure 1).

In addition, we also employ a spatial measure based on the total variation [7]. The total variation of an image region effectively measures the sum of absolute differences between the region and a spatially shifted version of that region. Thus, a region which is smooth (e.g., sky) will exhibit a lesser total variation than a region, such as a texture, which demonstrates a greater variation across space. In a probabilistic framework, Blanchet et al. [7] have shown that TV can be used as a contrast-invariant measure of phase coherence, and thereby can be used as a measure of local sharpness. We argue that a non-probabilistic application of TV can be useful for measuring local sharpness due, in part, to its ability to take into account local contrast.

Here, we demonstrate that a weighted geometric combination of these two measures (spectral slope and total variation), when adjusted to account for visual perception, can lead to an effective measure of local perceived sharpness. We argue that a non-probabilistic application of TV can be useful for measuring local sharpness due, in part, to its ability to take into account local contrast.

Thus, a region which is smooth (e.g., sky) will exhibit a lesser total variation than a region, such as a texture, which demonstrates a greater variation across space. In a probabilistic framework, Blanchet et al. [7] have shown that TV can be used as a contrast-invariant measure of phase coherence, and thereby can be used as a measure of local sharpness. We argue that a non-probabilistic application of TV can be useful for measuring local sharpness due, in part, to its ability to take into account local contrast.

This paper is organized as follows: Section 2 provides the details of the $S_3$ measure. In Section 3 we present results of the $S_3$ measure for within-image sharpness prediction, across-image sharpness prediction, prediction monotonicity, and correlation with subjective ratings of image quality. General conclusions are presented in Section 4.

2. Algorithm

Let $X$ denote the $M_1 \times M_2$-pixel input grayscale image. We divide $X$ into blocks of size $m \times m$ pixels with $d$ pixels of overlap between neighboring blocks. Let $x$ denote a block of $X$.

2.1. Spectral Measure of Sharpness

Let $S_1(x)$ denote a measure of the perceived sharpness based on the slope of the magnitude spectrum of block $x$. Specifically, given $x$, we first compute the 2-D Discrete Fourier transform (DFT) denoted $y(f, \theta)$, where $f$ is the radial frequency and $\theta$ is the orientation. From $y(f, \theta)$, we next compute the total magnitude spectrum across all orientations $z(f) = \sum_{\theta} |y(f, \theta)|$. The slope of the spectrum of $x$ is then given by

$$\alpha^* = \arg \min_{\alpha} \| \beta f^{-\alpha} - z(f) \|_2,$$

where the $L_2$-norm is taken over all radial frequency $f$.

Finally, $S_1(x)$ is given by

$$S_1(x) = 1 - \frac{1}{1 + e^{\tau_1(\alpha^* - \tau_2)}},$$

where $\tau_1 = -3$ and $\tau_2 = 2$. As demonstrated in Figure 1, this sigmoid attempts to account for the HVS tuning to the spectrum of natural scenes; namely, regions with a slope factor of $0 \leq \alpha \leq 1$ appear sharp, whereas regions with $\alpha > 1$ appear blurred. Notice from the noise images in Figure 1 that perceived sharpness changes slowly for $\alpha = 0 - 1$, changes rapidly for $\alpha = 1 - 3$, and saturates for $\alpha > 3$.

In applying Equation (2), we use blocks of size $m = 32$, an overlap of $d = 24$, and a Hanning window to suppress edge effects. We have found that this relatively large block size of $m = 32$ provides a sufficient number of DFT coefficients to accurately estimate the slope of the local spectrum. The collection of $S_1(x)$ for all blocks are assembled to form our spectral-based sharpness map denoted by $S_1(X)$.

2.2. Spatial Measure of Sharpness

For our spatial-based measure of sharpness, we employ a modified version of the total variation proposed in [7].
total variation of block $\mathbf{x}$ is given by

$$v(\mathbf{x}) = \sum_{i,j} |x_i - x_j|$$  \hspace{1cm} (3)

where $x_i$ and $x_j$ are 8-neighbor pixels in $\mathbf{x}$.

Let $S_2(\mathbf{x})$ denote a measure of perceived sharpness based on the total variation of block $\mathbf{x}$. Instead of deriving $S_2(\mathbf{x})$ directly from the total variation of $\mathbf{x}$, we compute it as the maximum of the total variation of smaller blocks of $\mathbf{x}$:

$$S_2(\mathbf{x}) = \max_{\xi \in \mathbf{x}} v(\xi)$$  \hspace{1cm} (4)

where $\xi$ is a $2 \times 2$ block of $\mathbf{x}$.

The max operator used in Equation (4) attempts to account for visual summation across space [10]. Namely, the HVS combines perceptual responses across space in a non-linear fashion which has been shown to be well approximated via a max operator ($L_\infty$-norm). The overall perceived sharpness of a block can thus be approximated by the maximum perceived sharpness of its sub-blocks. For this local total variation measure, we use blocks of size $m = 8$ with no overlap between blocks. The collection of $S_2(\mathbf{x})$ for all blocks are assembled to form our spatial-based sharpness map denoted by $S_2(\mathbf{X})$.

Figure 2 shows $S_1$ and $S_2$ measurements for noise images with different spectral slopes. Figure 3 shows $S_1$ and $S_2$ measurements for an image containing a single annular edge blurred such that they exhibit varying spectral slopes. Following from the previous argument on visual summation, the $S_1$ and $S_2$ values in each figure denote the maximum of each corresponding map $S_1(\mathbf{X})$ and $S_2(\mathbf{X})$. Notice that these two measures are quite good at predicting the relative perceived sharpness between images.

### 2.3. Combining the Spectral and Spatial Measures

Given the the spectral-based sharpness map $S_1(\mathbf{X})$ and the spatial-based sharpness map $S_2(\mathbf{X})$, we combine them into an overall perceived sharpness map $S_3(\mathbf{X})$ via

$$S_3(\mathbf{X}) = S_1(\mathbf{X})^\gamma \times S_2(\mathbf{X})^{1-\gamma}$$  \hspace{1cm} (5)

where $0 \leq \gamma \leq 1$. We have found $\gamma = 0.5$ to generally yield the best results (see also Section 3.3) and have accordingly used this value for the results presented in Section 3.

When looking at two images to estimate which one is sharper, the image that appears sharper is not necessarily the one which contains more sharp regions. Instead, visual summation dictates that the overall sharpness is determined based on the sharpest region in each image. Accordingly, from the sharpness map $S_3(\mathbf{X})$, we propose that the overall sharpness of $\mathbf{X}$ be given by the maximum value of $S_3(\mathbf{X})$:

$$S_3 = \max_{\mathbf{x} \in \mathbf{X}} S_3(\mathbf{x})$$  \hspace{1cm} (6)

where $\mathbf{x}$ is a block in $\mathbf{X}$ and $S_3(\mathbf{x})$ denotes the intensity of the corresponding value in the $S_3(\mathbf{X})$ map.

We call the resulting scalar value given by Equation (6) the $S_3$ index.

### 3. Results and Analysis

In this section, we present results of our $S_3$ measure for within-image sharpness prediction and across-image sharpness prediction. We also demonstrate the utility of $S_3$ on two tasks: monotonic estimation of the standard deviation of the impulse response used in Gaussian blurring, and no-
reference quality assessment of blurred images.

3.1. Sharpness Maps

Figure 4 depicts $S_3$ maps and indices for a variety of images containing commonplace subject matter. In terms of across-image prediction, the $S_3$ indices of image *branches* and *pelicans* are the greatest, and the blurry *ball* image has the lowest $S_3$ index, which visually agrees with the relative perceived sharpness across these images. For example, two of the images, *beans* and *peak*, are not as sharp as images *pelicans* and *branches*, but clearly the former are much sharper than image *ball*. The $S_3$ indices for these images reflect this rank ordering of perceived sharpness. Image *petal*, on the other hand, is sharper than image *ball* but not as sharp as *beans* (since *beans* contains very sharp textures). Again, the $S_3$ index is able to capture this rank ordering of overall sharpness.

In terms of within-image sharpness prediction, all of the $S_3$ maps quite accurately capture the sharp regions in each image. For example, in image *ball*, the map is able to point out that the right vertical line is sharper than the left line. Although the ball is blurred, which is correctly shown as dark in the map, the map is also able to capture the sharp shadow under the ball.

In image *petal*, the flower’s stamens are the sharpest regions in the image, which is accurately noted in its corresponding $S_3$ map. The second sharpest region in this image occurs along the borders of the flower’s petals, particularly, the lower petals. This fact is also captured in the map: The borders of the upper petals receive a low sharpness value, the lower petals receive a higher sharpness value, and the stamens receive the greatest sharpness value.

The perceived sharpness of image *beans* arises from the visibility of fine details in the textured background and from the high-contrast edges of the beans themselves. The interior of the beans, on the other hand, appears quite smooth. The $S_3$ map is able to quantify the relative sharpness of each region; in particular, notice that the interior of the beans receives a low sharpness value, whereas the edges of beans receive the greatest sharpness value in this image.

In image *peak*, even though the peak is the main subject of this image, it is blurred. The sharpest region, instead, is the region containing the bushes in the foreground. Again, the $S_3$ map supports these observations. The darkest region in the map corresponds to the sky, since it is smooth; the peak itself receives a slightly higher sharpness value, particularly along its edges, though also in its textured areas. Also notice that some of the brush in the foreground is actually blurred, a fact which is also captured by the map.

In image *pelicans*, the sharpest region includes the pelicans and bushes in the middle of the image. The regions in the background, which include the building and the sky, are very blurred, and the water is increasingly blurred toward the bottom of the image. Accordingly, the $S_3$ map indicates the greatest sharpness toward the center of the image (corresponding to the pelicans and bushes) and decreasing sharpness away from the center.

Image *branches* includes two regions: the very sharp branches and very smooth sky. Again, as can be seen from the $S_3$ map, the branches receive a very high sharpness value while the sky receives a very low sharpness value.

Yet, as with any computational measure, $S_3$ does indeed fail on some images. One failure case is shown in Figure 5. Here, the people in the background are noticeably more blurred than the diver in the foreground. The $S_3$ map, however, overestimates the sharpness of the background,
ball  $S_3 = 0.447$

petal  $S_3 = 0.537$

beans  $S_3 = 0.702$

peak  $S_3 = 0.728$

pelicans  $S_3 = 0.804$

branches  $S_3 = 0.808$

Figure 4. Original images and corresponding $S_3$ maps and indices ($\gamma = 0.5$).

which we suspect is due to contrast. We are currently in the process of extending the proposed measure to better account for the perception of contrast and its influence on perceived sharpness.

3.2. Monotonic Prediction of Blur Parameter

A common technique of validating a sharpness measure is to examine the measure’s performance in monotonically predicting increasing amounts of blur (see, e.g., [2], [3]). To test this, we used the 29 reference images from the LIVE image database [11]. The 29 original 24-bits/pixel reference images were first converted to grayscale via a pixel-wise transformation of $I = 0.2989 R + 0.5870 G + 0.1140 B$, where $I$, $R$, $G$, and $B$ denote the 8-bit grayscale, red, green, and blue intensities, respectively. The grayscale images were then blurred using a Gaussian filter of size $15 \times 15$ pixels and standard deviations of $\sigma = \{0.4, 0.8, 1.6, 2.0, 2.4, 2.8\}$. We then computed $S_3$ indices for these blurred images.

The results of this test are shown in Figure 6, in which the $S_3$ indices for blurred versions of the same original image are represented by individual lines plotted against $\sigma$. The
Figure 6. Relationship between the standard deviation of the Gaussian blurring filter’s impulse response and the $S_3$ index (normalized to have a maximum value of 1.0 across the 29 test images).

average of these data are shown by the solid black line. Notice that we indeed obtain a monotonic trend of $S_3$ vs. $\sigma$.

Figure 7 shows $S_3$ maps and indices for a close-up of image parrots and for its sharpened and blurred versions. As can be seen from this figure, the sharpened image has the highest $S_3$ index, whereas the other two blurred versions have smaller $S_3$ values. Also notice that the maps accurately reflect the relative sharpness across these images: The maps indicate that sharpness progressively decreases from the sharpened image to the original image, followed by the image blurred using $\sigma = 0.8$, and then followed by the image blurred using $\sigma = 2.0$.

3.3. No-Reference Quality Assessment

Figure 7. $S_3$ maps and indices for the same image with different amount of sharpness and blurriness.

Finally, we analyze the performance of $S_3$ for no-reference quality assessment of blurred images. Here, we again employ the LIVE image database since subjective ratings of image quality are also available (differential mean-opinion scores, DMOS values). We computed $S_3$ indices for grayscale versions of the 145 blurred images provided in the LIVE database (varying amounts of Gaussian blur were applied to the 29 reference images). Table 1 shows the correlation coefficient and root mean-squared error (RMSE) of a cubic regression of the 145 subjective ratings on the corresponding 145 $S_3$ indices. Figure 8 shows a scatterplot between the subjective ratings and our $S_3$ index (linearized based on the cubic regression).

As listed in Table 1, the $S_3$ index achieves a maximum correlation coefficient and minimum RMSE using $\gamma = 0.5$. The resulting correlation of 0.914 is quite noteworthy given that $S_3$ does not have access to the original (non-blurred) image. As a comparison, full-reference quality-assessment methods (which require an original image) achieve correlations of 0.784 (PSNR), 0.945 (SSIM [12]), and 0.934 (VSNR [13]) on these images.
Table 1. Performance of $S_3$ on no-reference quality assessment of blurred images

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Correlation Coeff.</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/4</td>
<td>0.9081</td>
<td>7.843</td>
</tr>
<tr>
<td>2/3</td>
<td>0.9104</td>
<td>7.750</td>
</tr>
<tr>
<td>1/2</td>
<td><strong>0.9143</strong></td>
<td>7.585</td>
</tr>
<tr>
<td>1/3</td>
<td>0.9115</td>
<td>7.706</td>
</tr>
<tr>
<td>1/4</td>
<td>0.9111</td>
<td>7.721</td>
</tr>
</tbody>
</table>

Figure 8. Scatterplot of $S_3$ index ($\gamma = 0.5$) vs. subjective ratings for no-reference image quality assessment of blurred images from the LIVE image database.

4. Conclusion

In this paper, we presented a block-based algorithm to measure local perceived sharpness in an image which we call the $S_3$ measure. By using both spectral and spatial properties, the proposed $S_3$ measure is able to quantify local perceived sharpness within and across images. We also demonstrated that the resulting sharpness map can be collapsed via a max operator to obtain a scalar index which quantifies an image’s overall perceived sharpness. The utility of this index was demonstrated both for monotonic estimation of the standard deviation of the impulse response used in Gaussian blurring, and for no-reference quality assessment of blurred images. Additional research is needed to better account for the effect of contrast on perceived sharpness.

References