Sharpness measure: towards automatic image enhancement

2 authors:

Doron Shaked
HP Inc.
64 PUBLICATIONS   **1,817** CITATIONS

SEE PROFILE

Ingeborg Tastl
HP Inc.
54 PUBLICATIONS   **460** CITATIONS

SEE PROFILE

Sharpness Measure: Towards Automatic Image Enhancement

Doron Shaked, Ingeborg Tastl
HP Laboratories Israel
HPL-2004-84(R.2)
June 10, 2005*

image
enhancement,
image analysis,
image sharpening,
feature extraction

We propose a measure for image sharpness, which facilitates automatic image sharpness enhancement. This way blurry images will be sharpened more whereas sufficiently sharp images will not be sharpened at all. The measure employs localized frequency content analysis in a feature-based context. Thereby it avoids many of the pitfalls of alternative methods: Frequency domain methods provide excellent sharpness measures for images of similar scenes, however they fail when the scene changes. Feature-based methods concentrate on features, however assumptions required for good performance are too restrictive for general purposes. The proposed sharpness measure correlates well with perceived sharpness, and is to a large degree invariant to image content. Furthermore, we show that the proposed image sharpness measure can be used to drive an enhancement algorithm, which will sharpen an input image to a nominal measure. Last but not least, the proposed sharpness measure is computationally efficient, and requires fewer computations than a 3x3 convolution.

Approved for External Publication

Sharpness Measure:
Towards Automatic Image Enhancement

Doron Shaked and Ingeborg Tastl

Abstract

We propose a measure for image sharpness, which facilitates automatic image sharpness enhancement. This way blurry images will be sharpened more whereas sufficiently sharp images will not be sharpened at all. The measure employs localized frequency content analysis in a feature-based context. Thereby it avoids many of the pitfalls of alternative methods: Frequency domain methods provide excellent sharpness measures for images of similar scenes, however they fail when the scene changes. Feature-based methods concentrate on features, however assumptions required for good performance are too restrictive for general purposes. The proposed sharpness measure correlates well with perceived sharpness, and is to a large degree invariant to image content. Furthermore, we show that the proposed image sharpness measure can be used to drive an enhancement algorithm, which will sharpen an input image to a nominal measure. Last but not least, the proposed sharpness measure is computationally efficient, and requires fewer computations than a 3x3 convolution.

1. Introduction

The data current imaging applications face can involve sources as diverse as:
- Disposable camera images, developed by a low quality developer, scanned at home, and compressed.
- Manually enhanced high-resolution images from a skilled professional photographer.

A printer or any other imaging device for that matter cannot perform well in such diverse conditions, unless it can estimate the quality of its input images and process them accordingly.

Thus, the notion of adaptive image processing has recently crystallized (in HP and elsewhere) to the notion of convergent image processing algorithms: We have to make sure that we converge to the same image quality (IQ) – no matter what the input IQ is. This notion of convergence implies an IQ space, a mapping from images into that space, and a mapping of imaging algorithms to displacements in that space.

While the ideal objective described above is still far ahead, this paper presents a significant step towards that goal. This work presents a measure for image sharpness (an important IQ feature). The results we have to date indicate a good fit between the proposed measure and perceptual sharpness. Furthermore, the measure may be used to drive an image enhancement algorithm placing a given image on a sharpness target. The enhancement algorithm [4] was chosen for its quality, and the flexibility it provides to enhance noisy images.

Adaptive image enhancement and specifically adaptive sharpening have been widely investigated. A large body of research appears under Blind Deconvolution. In a recent survey [3] blind deconvolution methods are classified into the following categories:
- Feature-based models, e.g. in astronomy images where one can assume a spot to represent the true point spread function.
- Model based zero analysis methods, which are impractical in our case since they

assume a known parametric blur model and no noise.

- ARMA filter models, are computationally intensive and occasionally unstable.
- Single stage algorithms are computationally intensive iterative algorithms estimating the blur and the reconstructed image simultaneously.

More practical algorithms include one by Eschbach and Fuss [2] who locate a sharp region (maximal local gradient), and tune a sharpening algorithm such that the gradient-based sharpness in the region will reach a nominal target. A recent sharpness measure (DSS) by Zhang et al. [8] averages the maximal local absolute difference over all edge pixels. Both measures mix sharpness and contrast, which are perceptually similar though technically different IQ elements.

Tretter [7] assumes a fractal image model, which translates to a *1/f* model in the frequency domain. This implies a nominal ratio between the low and high frequency bands, which can be tuned with the appropriate amount of sharpening.

The sharpness measure presented here has a little from each of the above methods, which makes it, so we argue, better than the alternatives. An important differentiator of the proposed measure is that it is implemented in a highly efficient form.

In the next section we motivate and derive the outline of the proposed measure. Section 3 details its implementation and focuses on efficiency considerations. Results are presented in Section 4. Section 5 is a summary.

2. Derivation

According to the fractal image model, natural images exhibit fractal behavior, which could be modeled in the frequency domain. Let us denote an image by $m(x, y)$, and its Fourier transform by $F(m) = M(\xi_x, \xi_y)$ or in short $M(\bar{\xi})$, where $\bar{\xi} = (\xi_x, \xi_y)$ are the Cartesian frequency coordinates. According to the fractal model [6]

$$\left| M\left(\bar{\xi}\right) \right| = \alpha \Big/ \left\| \bar{\xi} \right\|_2^{2H+2} \tag{1}$$

It follows, that if we knew the Hurst parameter H, the image spectrum would be well defined and any deviations may be attributed to linear acquisition degradations. We could then, in theory, reconstruct the original image by inverting the degradation. This however, would be taking the model too far. The frequency distribution could nevertheless be useful for estimating single parameter model based degradations such as in [7], where the chosen model parameter was the enhancement algorithm's sharpening parameter. Our initial single parameter sharpness model was based on the high to band pass frequency content ratio.

$$Sh_m = \frac{HP_m}{BP_m} \tag{2}$$

where

$$HP_m = \int\limits_{\bar{\xi} \in H} \left| M\left(\bar{\xi}\right) \right|^2 d\bar{\xi}$$

$$BP_m = \int\limits_{\bar{\xi} \in B} \left| M\left(\bar{\xi}\right) \right|^2 d\bar{\xi} \tag{3}$$

and *H, B* are respectively high and low-band pass frequency ranges. Note that by substituting (3) into (2), α of (1) is canceled out.

As may be expected from (1), we found out that Sh_m performs very well comparing sharpness of images of similar scenes. However it fails comparing the sharpness of images of different scene types.

What seems to be missing is a prevalent feature to replace the common scene, which would also serve to reduce the variation of the Hurst parameter *H*. Indeed we can safely assume that all images have some sharp edges, which we can use as features.

Whereas the notion of feature based sharpness estimation is hardly new, feature-based estimators will either assume the existence of a particular feature in the image, such as a singular point in astronomical images, or else work in the spatial domain, such as in the case of some gradient-based sharpness estimators. Gradient based estimators will usually not work in our case because gradients are affected by contrast changes as much as by

sharpness, and they are less sensitive to the small blur function support in which we are interested.

We propose to use a localized measure similar to (2), localized on feature locations $F \subset R^2$. For that we will need to define a local Fourier description. For now it will suffice to denote it by $M(\xi_x, \xi_y, x, y)$, or in short $M(\overline{\xi}, \overline{x})$. The proposed sharpness measure is thus

$$Sh_m^F = \frac{HP_m^F}{BP_m^F} \qquad (4)$$

where

$$HP_m^F = \int\limits_{\overline{x} \in F} \int\limits_{\overline{\xi} \in H} \left| M(\overline{\xi}, \overline{x}) \right|^2 d\overline{\xi} d\overline{x}$$

$$BP_m^F = \int\limits_{\overline{x} \in F} \int\limits_{\overline{\xi} \in B} \left| M(\overline{\xi}, \overline{x}) \right|^2 d\overline{\xi} d\overline{x} \qquad (5)$$

and

$$F = \left\{ \overline{x} \mid \int\limits_{\overline{\xi} \in B} \left| M(\overline{\xi}, \overline{x}) \right|^2 d\overline{\xi} > T \right\} \qquad (6)$$

for some threshold T.

As evident from (6), although we propose to use edges as an indicator of a common scene, we do not explicitly assume knowledge of the original image at the feature location. Consequently we do not need to fit the local image to a template. In addition the use of the local frequency band ratio (4) as opposed to spatial features makes sure that the contrast of the features (modeled by α in (1)) does not overshadow its sharpness.

A different motivation for Sh_m^F can be found in our ultimate goal, namely, to have the measure drive an enhancement algorithm to a consistent target IQ, and more specifically a consistent target sharpness measure Sh_m^F. The selective algorithm we chose to drive [4] is based on unsharp masking, which modifies the high frequency band in a predictable way, while keeping the low frequency band relatively constant. However in contrast to unsharp masking, it does not affect (and even smoothes) noisy regions, and thus, unless the measure is

limited to strong edges it might behave unpredictably.

Sh_m^F is thus a frequency band ratio measured on strong image features.

3. Implementation details

The following are design considerations that turned out to be critical in the development of the sharpness estimation module. Most of those were driven by computational efficiency considerations.

1. As was already implied by the move to local frequency analysis in (5), we use spatial filters rather than a full 2D Fourier transform. In the appendix we show that

$$HP_m^F \approx \int\limits_{\overline{x} \in F} HP_m^2(\overline{x}) d\overline{x}$$

$$\qquad (7)$$

$$BP_m^F \approx \int\limits_{\overline{x} \in F} BP_m^2(\overline{x}) d\overline{x}$$

where

$$HP_m(\overline{x}) = (hp * m)(\overline{x})$$

$$\qquad (8)$$

$$BP_m(\overline{x}) = (bp * m)(\overline{x})$$

are ideal bandpass *bp(x)*, and highpass *hp(x)* filters. In practice we require sharp frequency profiles for *hp(x)*, and *bp(x)*.

2. All the above could be repeated for 1D rather than 2D, with a minor difference in (1), whose 1D equivalent is

$$\left| M(\overline{\xi}) \right| = \alpha \big/ \left\| \overline{\xi} \right\|_2^{2H+1} \qquad (9)$$

This difference does however not propagate further in the derivation.

The input *m* in (8) is thus a collection of image rows and columns. Furthermore, *bp*, and *hp* are 1D filters.

A possible down side of working in 1D is that diagonal edges will appear more blurry than similar horizontal or vertical edges.

3. The rapid decay of the expected frequency content (9) would drive all ratios to zero. In order to get a reasonably scaled response we have to work with a

more balanced signal. One alternative is the 1D derivative of (9) (multiplication by ξ in the transform domain). According to (9) it should result in a uniform frequency distribution. Figure 1 depicts an average frequency distribution of image rows in red, and the distribution of the corresponding 1D row derivatives in blue. Considering that the DC component for the derivative signal is, by definition zero, it aligns fairly well with (9).
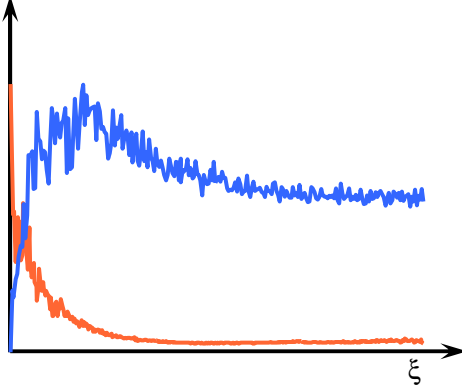

<u>Figure 1</u>: Average frequency distribution of image rows (red), and frequency distribution of row value derivatives (blue).

4. In order for the formulation (4-8) to work well the approximating spatial filters should be narrow and have a sharp frequency profile, which in turn requires large spatial support. Straightforward convolution kernels will necessarily be overly large. The obvious solution is IIR filtering. IIR filters are commonly used in signal processing [5], and much less so in image processing [1]. Let n denote the output of a linear filter. Equation (10) details the generic formulation of kernel filters (Finite Impulse response – FIR), and recursive filters (Infinite Impulse Response – IIR).

$$n^{FIR}(x) = \sum_{z=-N}^{N} K_z \cdot m(x-z) \qquad (10a)$$

$$n^{IIR}(x) = \sum_{z=1}^{N_\alpha} \alpha_z \cdot n(x-z) + \\ \sum_{z=0}^{N_\beta} \beta_z \cdot m(x-z) \qquad (10b)$$

The infinite support of IIR filters provides for the ability to design filters with a sharp frequency profile for a small number of taps (multiply and add operations). Figure 2 depicts the frequency response for the concatenation of the derivative and the high-pass and band-pass filters we propose. Each of the filters is implemented using 5 multiplication operations per pixel.
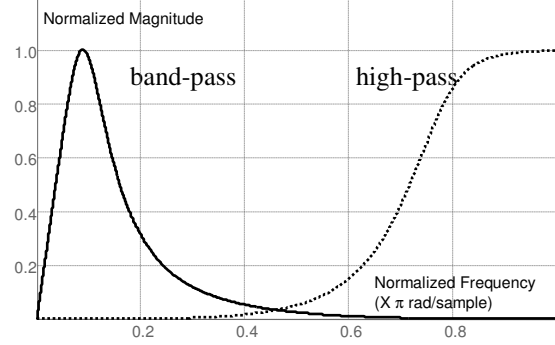

<u>Figure 2</u>: Frequency response magnitude of the derivative followed by the two filters. Both filters are normalized to a maximal magnitude of one.

Figure 3 depicts the frequency distributions of a signal derivative as in Figure 1, filtered by $BP_m(x)$ and $HP_m(x)$ (8) implemented by the 5 tap IIR filters depicted in Figure 2.
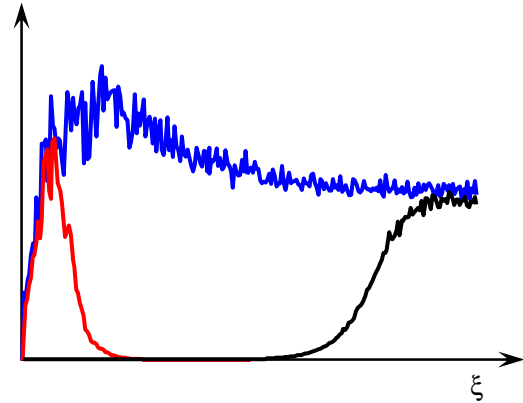

<u>Figure 3</u>: Frequency response magnitude of the derivative (blue), and the corresponding $BP_m(x)$ (red), and $HP_m(x)$ (black).

5. There are two problems with IIR filters. Some of them are unstable (which is not the case here), and they have nonlinear phase shifts. A partial correction for phase shifts can simply be implemented by

shifting the corresponding filter result according to the shift of the filter's dominant frequency. Another alternative is to concatenate the same filter, forwards and backwards, namely the output of a filter will be filtered again, by the same IIR filter but this time in the inverse 1D direction. This would result in a zero phase shift for all filters.

6. Substituting (7-8) into (4-5) measure Sh_m^F amounts to

$$Sh_m^F = \frac{\int\limits_{x \in F} HP_m^2(x)dx}{\int\limits_{x \in F} BP_m^2(x)dx} \quad (11)$$

Necessarily, F in (6) contains a variety of edges some of which do not represent the sharpness of the image. Namely, edges that are: blurry in the original scene, outside the image's depth of focus, or diagonal edges. The two averages in (11) weight edges according to their contrast, whereas we would prefer to focus on sharper edges. In order to reduce the relative weight of blurry edges we can average a local sharpness measure across the feature range F. Thus, we propose to use

$$Sh_m^{WF} = \int\limits_{x \in F} \left(\frac{HP_m(x)}{BP_m(x)} \right)^2 dx \quad (12)$$

7. According to (12) the ratio of high and band pass content is a local estimation of the sharpness, which is averaged over the feature region F. In most images the feature region is large enough and a good estimation is obtained even if only a small part of F is considered. We therefore redefine m as a sample collection of image rows and columns. For the purpose of estimating sharpness in photographic images sampling one in every 10 image rows or columns is usually sufficient.

To summarize, we propose to implement sharpness estimation Sh_m^{WF} (12), with $HP_m(x)$, and $BP_m(x)$ (8) implemented as

IIR filters (10b)[1], and substituting (7) into (6) to obtain

$$F = \left\{ x \mid BP_m(x) > T \right\} \quad (13)$$

4. Results

The proposed sharpness estimation is designed for efficiency and accuracy. The results reported below were obtained for measures using 5 multiplications per pixel per filter using one of every 10 rows and columns of the measured images. On the average less than 3 multiplications per pixel had to be performed.

To evaluate the proposed sharpness measure we applied it to a database of 115 images scanned from 5 different rolls of film. Three rolls (marked as Disposable 1, P&S 1, and SLR 1) were taken by Mike McGuire and Eric Montgomery of HP-Labs using a disposable, a Point&Shoot, and an SLR camera respectively. These images contain a variety of scenes, which are often similar (and sometimes identical) across the camera types. Two rolls (marked as Disposable 2, P&S 2) were taken by Doron Shaked using a disposable, and a Point&Shoot camera. These images contain pairs of identical scenes.

Images from the first three rolls contain scenes with highly variable scene texture (single person against a mosaic wall as opposed to a blank wall). One of the major sanity checks regarding the proposed sharpness measure was to make sure mosaic background images do not cluster separately (as they did in some of the earlier versions of the measure).

Figure 4 visualizes the sharpness measure for the 115 images. The horizontal axis is the log of the sharpness measure. Each dot represents an image. The 5 films are color-coded and separated on the vertical axis. Small vertical lines mark the average of a specific roll.

---

[1] For $BP_m(x)$ in (12): α_1=-2.3741, α_2=1.9294, α_3=-0.5321, β_0=β_3=0.0029, β_1=β_2=0.0087. For $HP_m(x)$ in (12): α_3= 1.4590, α_2= 0.9104, α_1=0.1978, β_3=-β_0=0.0317, β_1=-β_2=0.0951.

Our expectation was that better camera optics would result in sharper images (namely, SLR > P&S > Disposable). Evidently, the proposed sharpness measure agrees with our expectations, with an exception for Disposable-1 especially as compared to P&S-2. This mismatch is easily explained by looking at the image sets. Many images in Disposable-1 are taken from scenes at or close to the fixed focus distance of the disposable camera. Those are indeed perceptually sharper than most of the P&S-2 images.
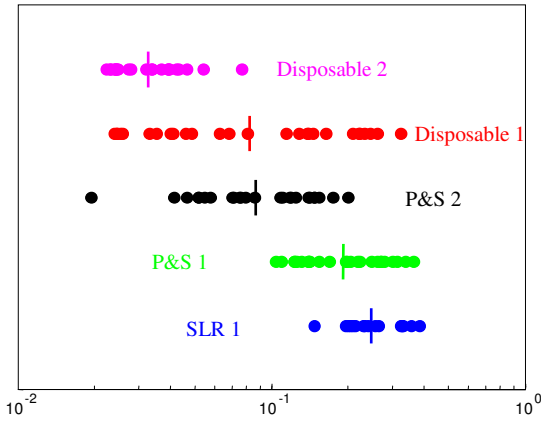

Figure 4: Log sharpness measure of the original scanned images (on the horizontal axis). Images are separated vertically and color-coded according to source.

A similar apparent paradox is depicted in Figure 5, which presents two image parts (same crop size). Oddly enough we noted that although the two images were of a similar scene, the Disposable-1 image (left) rated higher than the corresponding SLR-1 image (right). The reader would hopefully agree that the image on the left is sharper than the image on the right. Most other IQ parameters are, as one would expect, in favor of the SLR image, however, perceived sharpness goes the other way, as is predicted by the sharpness measure.

The ultimate test for the proposed measure is to fulfill its goal as a means towards automatic image enhancement. A minimal requirement is the ability to use the measure to drive a sharpness enhancement algorithm and align the measure of the enhanced images on a predetermined target. Results in this spirit are reported on the next page.
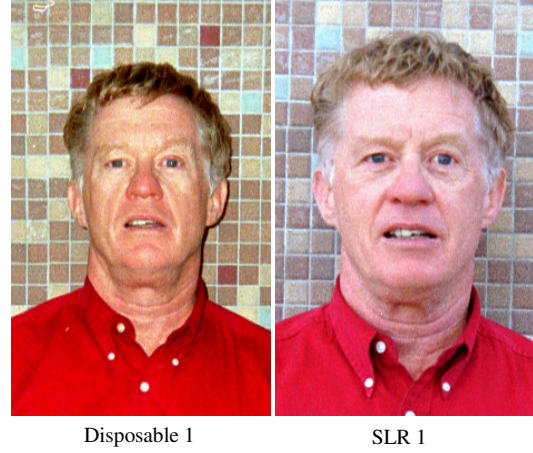

Disposable 1   SLR 1

Figure 5: A portrait taken by different cameras in a similar scene

We compared the sharpness measure for images that were sharpened using the method of [4] with a variable sharpening parameter and a fixed denoising threshold (T=3). We found empirically that the sharpness measure ratio of enhanced images with respect to originals is approximately linearly proportional to the respective sharpness parameter. As a consequence, we determined an empiric sharpness parameter map λ depicted in Figure 6, as a function of the required sharpness parameter ratio. The sharpness parameter, λ^*, to drive an image to a target sharpness is thus

$$\lambda^* = \lambda\left(\frac{\text{Target}}{Sh_{\text{orig}}^{WF}}\right) \qquad (14)$$

Notice that for images whose sharpness measure is higher that the target measure, the ratio is smaller than unity, and λ^* is zero, which means that the images are not enhanced (and naturally their sharpness measure remains higher than the target). On the other hand, if the required sharpness boost is higher than 7.35, λ^* is clipped at 1.5. The reason for the latter was that images for which λ^* exceeded 1.5 were often unacceptably sharp and their sharpness increase was less predictable.
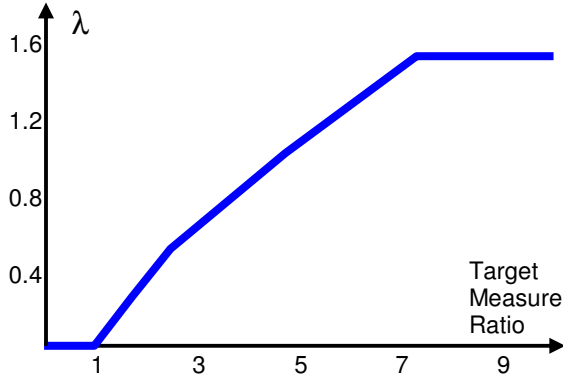
Figure 6: Empirically determined sharpness parameter map λ as a function of required sharpness measure enhancement ratio.

Figure 7 visualizes the sharpness measure for the 115 enhanced images. In the three lower sub figures images were enhanced using λ^* as described above, and considering target sharpness of 0.05, 0.1, and 0.3 respectively. For each of the sub figures the target sharpness is marked by a solid line. The original measure as in Figure 4 appears at the top. Note the way images below the target measure converge to the target. Most notable exceptions occur for higher target sharpness where λ^* is clipped to 1.5.

It should be noted that the sharpness convergence demonstrated in Figure 7 is empiric. We do not claim analytic convergence. Indeed the distribution of lagging images (in the lower subfigure of Figure 7) is not a shifted replica of the sharpness distribution in the top subfigure, as it should have been if the convergence were analytic.
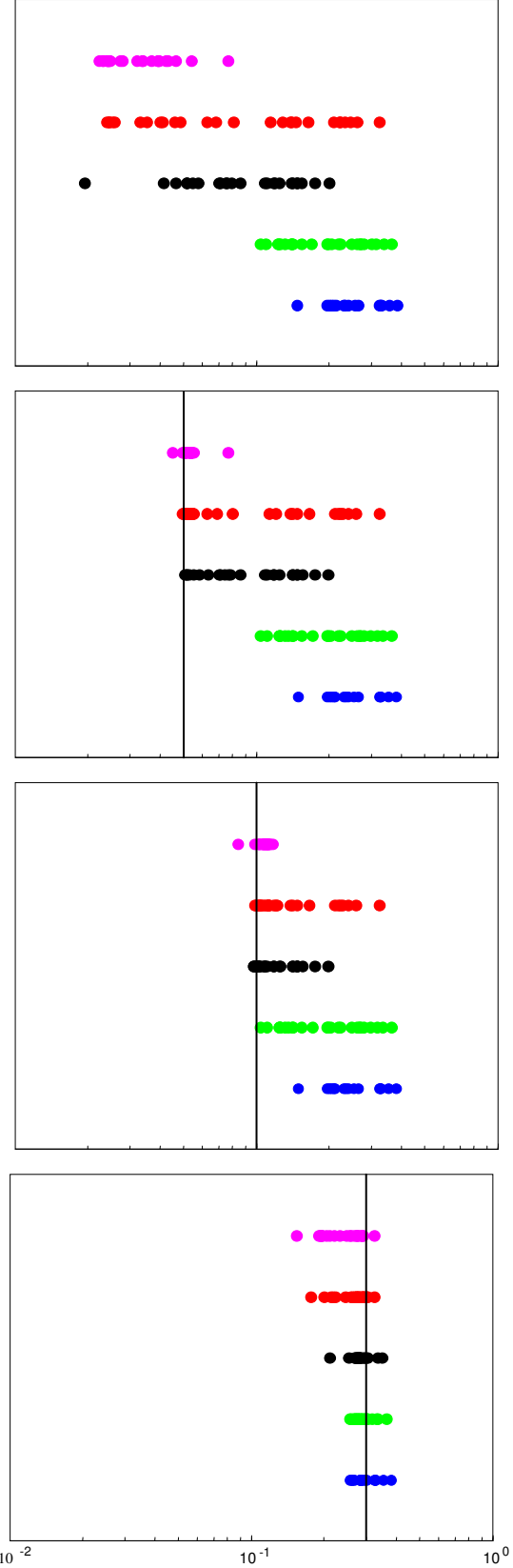

Figure 7: Log sharpness measure of adaptively sharpened images (on the horizontal axis). Measure of original appears in top subfigure. Target sharpness marked as solid line.

5. Summary

In this paper we proposed an image sharpness measure as a step towards automatic image enhancement. The proposed measure is based on localized frequency analysis. Its low computational requirement makes it suitable for commercial applications. We have shown empirically that the measure is consistent with the selected enhancement algorithm and provides convergent sharpening.

In order to provide for automatic image enhancement we still need to show that it can also provide for perceptual convergence. In a sequel we combine this measure with image noise estimation and other measures into an adaptive automatic image enhancement algorithm.

6. Acknowledgements

7. References

[1] R. Deriche, "Fast Algorithms for Low Level Vision", *IEEE Trans. PAMI*, Vol. 12, pp. 78-86, 1990.

[2] R. Eschbach, and W. A. Fuss, "Image Dependent Sharpness Enhancement", US patent 5,363,209.

[3] D. Kundur, and D. Hatzinakos, "Blind Image Deconvolution", *IEEE Signal Proc. Mag.*, pp. 43-64, May 1996.

[4] R. Maurer, **"**Selective Smoothing and Sharpening of Images by Generalized Unsharp Masking", US patent 6,665,448.

[5] Oppenheim, and Schafer, *Digital Signal Processing*, Prentice Hall, 1975.

[6] B. Pesquet-Popescu, and J. Levi Vehel, "Stochastic Fractal Models for Image Processing", *IEEE Signal Proc. Mag.*, pp. 48-62, May 2002.

[7] D. Tretter, "Apparatus and Method for Determining the Appropriate Amount of Sharpening for an Image", US patent 5,867,606.

[8] B. Zhang, J. P. Allebach, and Z. Pizlo, "An Investigation of Perceived Sharpness and Sharpness Metrics," *Proc. SPIE*, Vol. 5668, Image Quality and System Performance II, pp.98-110, January 2005.

Appendix A

Here we analyze the proportion factors in (7), arguing that it is a reasonable approximation.

Suppose that the local Fourier descriptor $M(\overline{\xi}, \overline{x})$ is a Fourier descriptor of a local window function applied to the signal.

$$M(\overline{\xi}, \overline{x}) = F_{\overline{z}}(w(\overline{z}) \cdot m(\overline{x} - \overline{z})) \quad (15)$$

For a weight w function such that

$$\int |w(\overline{x})|^2 d\overline{x} = 1 \quad (16)$$

Then according to Parseval

$$\int |M(\overline{\xi}, \overline{x})|^2 d\overline{\xi} = \int |w(\overline{z}) \cdot m(\overline{x} - \overline{z})|^2 d\overline{z} \quad (17)$$

and similarly

$$\int_{\overline{\xi} \in H} |M(\overline{\xi}, \overline{x})|^2 d\overline{\xi} = \int |w(\overline{z}) \cdot HP_m(\overline{x} - \overline{z})|^2 d\overline{z} \quad (18)$$

with HP_m the output of an ideal high pass filter as in (8).

Therefore for an unrestricted feature location F, we have

$$HP_m^{\infty} = \int \int_{\overline{\xi} \in H} |M(\overline{\xi}, \overline{x})|^2 d\overline{\xi} d\overline{x}$$

$$= \int \int |w(\overline{z})|^2 \cdot |HP_m(\overline{x} - \overline{z})|^2 d\overline{z} d\overline{x}$$

$$= \int |w(\overline{z})|^2 \cdot \int |HP_m(\overline{x} - \overline{z})|^2 d\overline{x} d\overline{z} \quad (19)$$

$$= \int |w(\overline{z})|^2 d\overline{z} \cdot \int |HP_m(\overline{x})|^2 d\overline{x}$$

$$= \int HP_m^2(\overline{x}) d\overline{x}$$

Note however, that (19) is true only if the integration over x is not restricted. In case it is restricted by F as in (5), we resort to an approximation. Let us replace (15) with

$$M(\overline{\xi}, \overline{x}) = F_{\overline{z}}(m(\overline{z}) \cdot w(\overline{x} - \overline{z})) \quad (20)$$

the equivalents for (17) and (18) follow immediately. Instead of (19) we get

$$HP_m^F = \int_{\overline{x} \in F} \int_{\overline{\xi} \in H} |M(\overline{\xi}, \overline{x})|^2 d\overline{\xi} d\overline{x}$$

$$= \int_{\overline{x} \in F} \int |HP_m(\overline{z})|^2 \cdot |w(\overline{x} - \overline{z})|^2 d\overline{z} d\overline{x}$$

$$= \int |HP_m(\overline{z})|^2 \cdot \int_{\overline{x} \in F} |w(\overline{x} - \overline{z})|^2 d\overline{x} d\overline{z} \quad (21)$$

$$= \int HP_m^2(\overline{z}) \cdot \int I_F(\overline{x}) |w(\overline{x} - \overline{z})|^2 d\overline{x} d\overline{z}$$

Where I is the indicator function

$$I_F(\overline{x}) = \begin{cases} 1 & \overline{x} \in F \\ 0 & \overline{x} \notin F \end{cases} \quad (22)$$

The inner integral in (21) is a convolution between the indicator function I and a square profile of the window function w. It results in a smoothed indicator function that resembles I_F more for narrower window functions. We get

$$HP_m^F = \int HP_m^2(\overline{x}) \cdot I_F^{Smooth}(\overline{x}) d\overline{x}$$

$$\approx \int_{\overline{x} \in F} HP_m^2(\overline{x}) \cdot d\overline{x} \quad (23)$$

All the above could be repeated for a bandpass filter to result in the second approximation of (7).