Qpro: An improved no-reference image content metric using locally adapted SVD

Xiang Zhu and Peyman Milanfar

Electrical Engineering Department
University of California at Santa Cruz, CA, 95064
xzhu@soe.ucsc.edu, milanfar@ucsc.edu

ABSTRACT

We present an improvement to our earlier Q metric,[1] a no-reference image content measure based on singular value decomposition (SVD) of local image gradient matrix. The new extension, Qpro, is capable of better measuring the amount of latent image content in the presence of both blur and random noise. As desired, its value drops monotonically when image becomes either blurry or noisy. Compared with our earlier metric Q which was computed using only anisotropic patches, Qpro implements SVD in transformed coordinates which are adapted to local estimated structure. In this way, it can measure a much wider variety of local image content, leading to significantly improved performance. Experiments demonstrate that this metric correlates with subjective quality evaluations even better than some full-reference quality metrics. It also outperforms other metrics in optimizing tuning parameters for image denoising filters.

Keywords: No-reference metric, singular value decomposition, denoising, sharpness, parameter optimization.

1. INTRODUCTION

Noise and blur are two major factors that affect image quality in both imaging and image processing systems. Inevitable noise can be generated by, for example, imaging sensor, scanning process or lossy compression. Blur may be introduced by defocusing, camera motion, image filtering, etc. Estimating the amount of image content in the presence of both blur and noise is an important problem, which is not trivial especially in a no-reference scenario. Most no-reference image sharpness metrics can hardly distinguish *true* signal from the high-frequency behavior due to noise.[2,3] Their values increase as their input images get either sharpened or noisy.

In article[1] we developed a no-reference image content metric Q based on the singular value decomposition (SVD) of local image gradients. This metric is capable of measuring true image content for many types of structured local patches (e.g. edge patch in Fig. 1 (a)). Its value drops monotonically when image/patch becomes either blurred or noisy. One problem for metric Q is that it fails in structured regions that do not contain a single dominant direction. One example is the *isotropic* patch in Fig. 1 (b), which has a circular shape and the energy in every orientation is basically the same. To solve this problem, a new metric Q_{pro} is proposed in this paper. Q_{pro} implements SVD in *transformed coordinates* which are adapted to local image content. Old Q becomes a special case of Q_{pro} that is valid for most structured regions, including isotropic ones.

(a) Edge (b) Circular (c) Corner

Figure 1. Different types of patches.

This work was supported in part by the US Air Force under Grant FA9550-07-1-0365, and by the National Science Foundation under Grant CCF-1016018.

2. REVIEW OF METRIC Q

Let us first review the definition of Q. Considering an image of interest g, we define the gradient matrix of an $N \times N$ local analysis window (w_i) centered at pixel i as:

$$\mathbf{G} = \begin{bmatrix} \vdots \\ \nabla \mathbf{g}_j^T \\ \vdots \end{bmatrix}, \quad j \in w_i \tag{1}$$

where $\nabla \mathbf{g}_j = [g_x(j), g_y(j)]^T$ denotes the gradient vector of the image at point j. The dominant orientation of the local window can be calculated by computing SVD of $\mathbf{G}$:

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^T \tag{2}$$

where both $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices. Singular vector $\mathbf{v}_1$ represents the dominant orientation of the local gradient field. The second singular vector $\mathbf{v}_2$ is the direction perpendicular to $\mathbf{v}_1$. Singular values $s_1 \geq s_2 \geq 0$ represent the energy in $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively.[4]

We define the image content metric Q as:

$$Q = s_1 \frac{s_1 - s_2}{s_1 + s_2}, \tag{3}$$

and it has been shown that for patches, where s_1 is significantly larger than s_2, Q can be viewed as a rough indicator of signal-to-noise-ratio (SNR), and its value drops when the patch becomes either blurry or noisy. Such patches are called *anisotropic*, such as the edge patches in Fig. 1 (a).

However, there still exist many *isotropic* patches that have $s_1 \approx s_2$ which nonetheless have strongly structured gradient fields (for example Fig. 1 (b), (c)). In these cases, metric Q cannot capture the major structure of the image content, and thus fails in measuring their quality change. In this paper, we propose a solution to this problem.

3. MEASURING ISOTROPIC PATCHES

Consider a typical isotropic patch example in Fig. 2, where in traditional Cartesian coordinate system we have $s_1 = s_2$. Our goal is to find a way to separate the gradient energy toward the center from the one in its perpendicular direction. One way is through rotation transform. For example, given a position j on the circle, we can rotate the coordinate by an angle θ_j using the following formula (see Fig. 2):

$$\nabla \mathbf{g}'_j = \begin{bmatrix} g'_x(j) \\ g'_y(j) \end{bmatrix} = \begin{bmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{bmatrix} \begin{bmatrix} g_x(j) \\ g_y(j) \end{bmatrix} \tag{4}$$

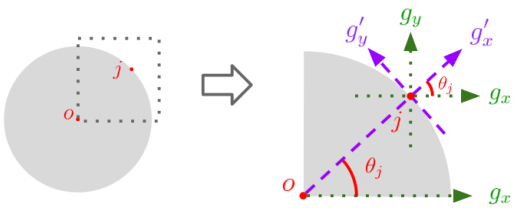
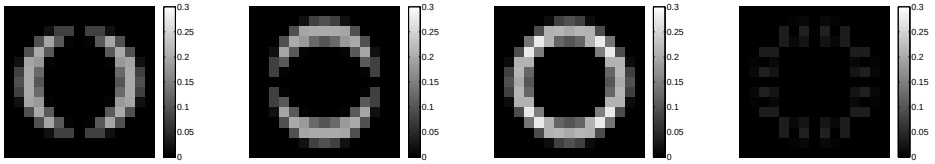


Figure 2. Rotational transformation example.

Note that the angle θ_j is determined by both the pixel position j and the center position (denoted as $o = [x_o, y_o]$). We implement this transform on the patch (b) from Fig. 1, and its corresponding gradients are given in Fig. 3, where we can see that after applying the transform, most gradient energy toward the center is



(a) $|g_x|$ (b) $|g_y|$ (c) $|g'_x|$ (d)$|g'_y|$

Figure 3. Horizontal and vertical derivative magnitudes of Fig. 1 (b) before (a, b) and after (c, d) the rotational transform.

transferred to g'_x. Though the transform does not change the l_2 norm of $\nabla \mathbf{g}$ at each pixel, it produces a dominant direction in the gradient field.

Because the rotation matrix is orthogonal, it does not change the correlation structure of random noise that is assumed to be IID. So all the statistical analysis on Q in article[1] is still valid. We define the new metric Q_{pro} as:

$$Q_{pro} = s'_1 \frac{s'_1 - s'_2}{s'_1 + s'_2} \tag{5}$$

where s'_1 and s'_2 are the singular values of the transformed gradient matrix $\mathbf{G}'$ in local window w_i

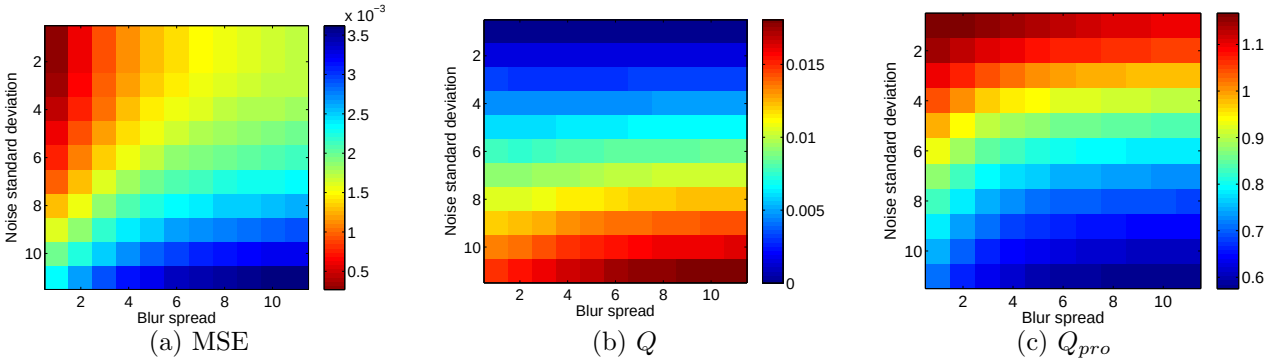
$$\mathbf{G}' = \begin{bmatrix} \vdots \\ \nabla \mathbf{g}_j'^T \\ \vdots \end{bmatrix}, \quad j \in w_i \tag{6}$$

To verify the validity of Q_{pro} with respect to both blur and noise, a simulated experiment is carried out. The test patch in Fig. 1 (b) is distorted through the following model:

$$\hat{g} = g \otimes h + n \tag{7}$$

where g stands for the clean patch, h represents a 9×9 Gaussian blur kernel, $\otimes$ denotes a convolution operator, and n is the WGN.

Both noise variance and blur kernel spread could be varied to alter the distortion level. The corresponding values of metric Q and Q_{pro} are given in Fig. 4 (b) and (c). Full-reference metric MSE is also plotted in (a) for comparison. Images get more and more blurred from left to right, and more and more noisy from top to bottom. It can be observed that, like MSE, the change of the metric Q_{pro} successfully reflects the change of the image quality: Q_{pro} monotonically drops as the distortion level rises, while the old Q failed.



(a) MSE (b) Q (c) Q_{pro}

Figure 4. Evaluations of MSE, Q and Q_{pro} on patch Fig. 1 (b), and its successively degraded versions. In (a) we inverted the color scale just to show the similarity between MSE and Q_{pro} in capturing the trend of image quality change.

One problem for Q_{pro} is that for each patch it is sensitive to the center position o, while the isotropic content is not always centered at the middle of the patch. In other words, before calculating Q_{pro} we need to detect

the proper center position first. On the other hand, in practice anisotropic patches appear more frequently than isotropic ones, which means given a patch we also need to decide which metric should be employed.

In article[1] we have discovered that the validity of Q depends on the relative difference between s_1 and s_2, which can be measured through its Coherence:

$$R(s_1, s_2) = \frac{s_1 - s_2}{s_1 + s_2} \tag{8}$$

The higher its Coherence is, the better the Q performs. Similarly, we can also employ Coherence $R(s_1', s_2')$ to determine the best center position o. Assume that we have a set of candidate positions $\{o_j\}$, the one with the highest $R(s_1', s_2')$ could be selected as a proper center position. Note that this position can be either inside or outside the patch. Specifically, if the center position is infinitely far from the patch ($x_o = -\infty$), the rotation matrix would become identity, and its corresponding $Q_{pro} = Q$. In other words, Q is nothing but a special case of Q_{pro}.

4. EXPERIMENTS

In article[1] a Q-based parameter optimization system for denoising filters is introduced (see Fig. 5). In this system, we vary the value of tuning parameter while observing the resulting metric. The best value of the algorithm parameter is then selected as that which maximizes the metric on the output. In this section, we employ both Q and Q_{pro} in the system to compare their performances. We simulate noisy data by adding white Gaussian noise to a clean test image (PSNR=23dB, 30dB). State of the art denoising method BM3D is employed as an example with its (variance) parameter σ_{est}^2 to be tuned.[5] We optimize σ_{est} in the broad range of $[1, 30]$. Because latent clean images are also available in this simulation, MSE is also calculated as comparison even though MSE is a full-reference metric that is not practical.

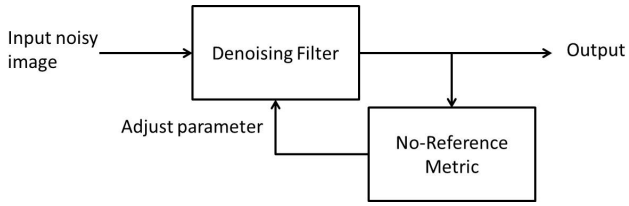


Figure 5. Selecting the tuning parameter using a no-reference image quality metric.

Plots of the experimental results using 23dB and 30dB inputs are given in Fig. 6 and Fig. 7, respectively. We can observe that in this experiment both Q and Q_{pro} successfully captured the changing trend of quality in the output as parameter were varied. However, probably due to its weak ability of structure capturing, Q prefers images that look slightly over-smoothed, while Q_{pro} yields the results which are visually good in both noise suppression and image detail preservation. Q_{pro} optimized results are also very close to the ones selected by full-reference MSE.

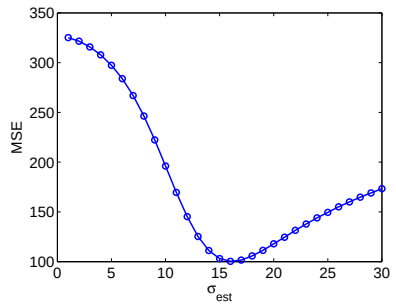
REFERENCES

[1] Zhu, X. and Milanfar, P., "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. on Image Processing* , 3116 (2010).

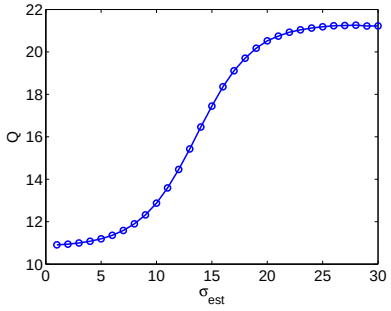
[2] Marziliano, P., Dufaux, F., Winkler, S., and Ebrahimi, T., "A no-reference perceptual blur metric," *Proceedings of the International Conference on Image Processing* **3**, 57–60 (2002).

[3] Ferzli, R. and Karam, L. J., "A no-reference objective sharpness metric using Riemannian tensor," *Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics* , 25–26 (January 2007).

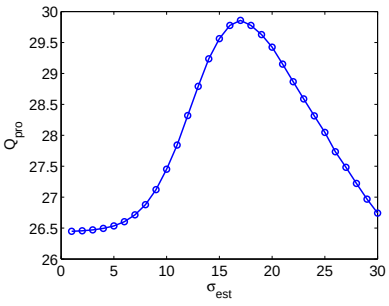
[4] Feng, X. and Milanfar, P., "Multiscale principal components analysis for image local orientation estimation," *Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers* **1**, 478–482 (November 2002).



(a) Plot of MSE



(b) Plot of Q



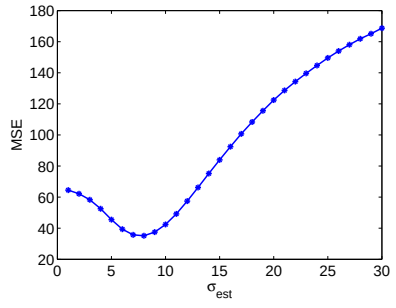
(c) Plot of Q_{pro}

(d) MSE optimized, $\sigma_{est} = 16$

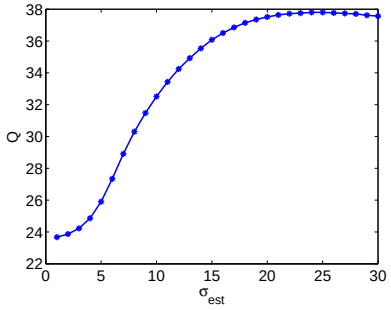
(e) Q optimized, $\sigma_{est} = 28$

(f) Q_{pro} optimized, $\sigma_{est} = 17$

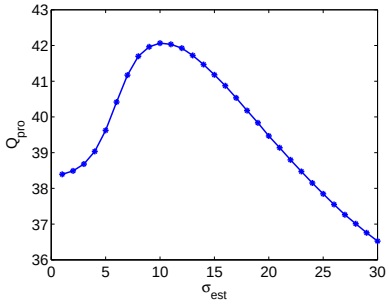
Figure 6. Plots of MSE, Q and Q_{pro} versus the tuning parameter in BM3D, and the corresponding optimized images. PSNR of the input image is 23dB.



(a) Plot of MSE



(b) Plot of Q



(c) Plot of Q_{pro}

(d) MSE optimized, $\sigma_{est} = 8$

(e) Q optimized, $\sigma_{est} = 25$

(f) Q_{pro} optimized, $\sigma_{est} = 10$

Figure 7. Plots of MSE, Q and Q_{pro} versus the tuning parameter in BM3D, and the corresponding optimized images. PSNR of the input image is 30dB.

[5] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing* **16**, 2080–2095 (August 2007).