Noname manuscript No.
(will be inserted by the editor)

Sparsity Based No-Reference Image Quality Assessment for Automatic Denoising

Meisam Rakhshanfar · Maria A. Amer

Received: date / Accepted: date

Abstract In image and video denoising, a quantitative measure of genuine image content, noise, and blur is required to facilitate quality assessment, when the ground-truth is not available. In this paper, we present a no-reference image quality assessment for denoising applications, that examines local image structure using orientation dominancy and patch sparsity. We propose a fast method to find the dominant orientation of image patches, which is used to decompose them into singular values. Combining singular values with the sparsity of the patch in the transform domain, we measure the possible image content and noise of the patches and of the whole image. To measure the effect of noise accurately, our method takes both low and high textured patches into account. Before analyzing the patches, we apply a shrinkage in transform domain to increase the contrast of genuine image structure. We show that the proposed method is useful to select parameters of denoising algorithms automatically in different noise scenarios such as white Gaussian and processed noise. Our objective and subjective results confirm the correspondence between the measured quality and the ground-truth. We show that the proposed method rivals related state-of-the-art no-reference quality assessment approaches.

Keywords No-reference quality assessment · Denoising · Sparse representation

1 Introduction

Image and video quality suffers from noise, blur, and compression artifacts. During the capturing process, noise from different sources is added to image and video content. It is essential to reduce the noise for enhancing the quality, reducing the bit-rate, or improving the performance of subsequent image processing tasks. Blur may be introduced to an image either during capture or due to processing such as denoising. In order to evaluate the performance of a denoiser or a deblurrer, a quantitative measure of quality is required. In many practical cases where the reference image is not available, the role of quality measurement techniques is more highlighted. During the development process of image enhancement algorithms, the importance of no-reference image-quality assessment (NR-IQA) becomes clear. The effect of changing the parameters, or the algorithm (e.g., in an optimizing process), should be studied and the output quality needs to be verified. Subjective evaluation is tedious, especially when the test dataset is large. Thus, an automated quality measurement is necessary. Although it seems not possible to design an image-quality assessment that replaces the human visual system in all situations, many assessment methods are designed and successful at achieving reliable results at least in confined conditions (e.g., limited range of noise and image structure). NR-IQA techniques aim to distinguish image structure (e.g., salient geometric features) from distortion (e.g., noise and blur) and quantify the overall image distortion without a ground-truth according to visual perception [27], sharpness, and noise. As a consequence, parameters of image or video processing methods, such as noise estimation, noise removal, and deblurring can be optimized based on overall quality. By measuring the quality of the final output, with a recursive procedure, quality of current output is compared with previous outputs to find the optimal point. In addition of parameter selection, NR-IQA methods can be used to classify images based on

M. Rakhshanfar · M.A. Amer
Electrical and Computer Engineering Department,
Concordia University, Montreal, QC, Canada

their quality. As an example, among several captured images, such as in the burst mode, the one with the highest quality can be selected as the image of interest. Unlike *image fidelity* assessment that evaluates the similarity of processed image to original one, here we are looking for *image quality* assessment for automatic denoising, i.e., automatically determining what parameter leads to the highest output quality in denoising process.

Many NR-IQA methods have been introduced for limited types of distortions (e.g., white noise, blur, and compression artifact); recent methods are designed to cover more distortion types. In reference based methods, many types of distortion can be detected. However, lack of reference in NR-IQA methods makes it difficult to detect certain types of distortion. For the distortions related to denoising such as different types of noise and blur, NR-IQA methods can be helpful in automatic denoising, i.e., selecting optimal parameters of a denoiser. In this paper, we propose a sparsity and dominant orientation based (SDQI) method that can be used 1) to optimize parameters of image denoising algorithms and 2) to verify the quality of these algorithms. We assume noise may have different types such as Gaussian or processed (non-white). To do this, we quantify the genuine image content based on the sparsity of local gradients using singular value decomposition (SVD) and discrete Fourier transform (DFT). SVD is applied to find the orientation dominancy of the image gradient patches. For a more accurate estimation of orientation, a shrinkage (i.e., suppressing the small coefficient) in the transform domain is first applied on the gradient image. To address multi-orientation patches, where one orientation is not dominant, we employ DFT to detect image structure, which increases the reliability of signal detection. To compute the SVD, instead of recursive matrix operations, we propose a faster method simpler to implement.

The remainder of this paper is as follows: section 2 discusses related work; section 3 describes our image content model and NR-IQA method; section 4 presents objective and subjective results, and section 5 concludes the paper.

2 Related work

The diversity of NR-IQA methods using various image processing principles makes them difficult to be categorized. Based on their applications some focus on specific types of distortion (e.g., white Gaussian noise) and others consider different potential distortions. In [23], the unbiased risk estimate is proposed to calculate the distortion cost (e.g., mean squared error MSE) for the enhancement application. It assumes the noise is additive white Gaussian (AWGN), and accurate estimation of the noise variance is available. The technique proposed in [12] detects both noise and blur in one step based on the image anisotropy and measures the visual quality based on the variance of the entropy. The optimal performance of this method is achieved when the degradation is globally uniform and in the case of non-uniform noise or blur its performance decreases. [3] presents two separate pipelines for estimation of noise and blur. The noise is assumed to be AWGN implying the high-frequency part of the noise exists, which is not accurate under real (e.g., processed non-white) noise. [11,1,9,19,17,6] are developed to substitute human visual system to classify images (e.g., detection of blurred versus non-blurred in digital photography). Based on their applications, these methods are designed to be more sensitive to blur and less to noise. Thus, their performance decreases in the presence of the noise because the detection of edge versus high-frequency noise becomes challenging. Just noticeable blur (JNB) [11] is introduced to express the presence of blur around an edge using Sobel operator on local patches. [19] computes the cumulative probability of blur detection (CPBD) by classifying the blocks into edges and smooth areas. [18] developed a blind image quality assessment (BIQI) method, which utilizes support vector machine classifier to define the quality index (QI) based on subband coefficients of wavelet transform. The image is first ranked in each category of the degradations: JPEG and JPEG2000 compression, white noise, Gaussian blur, and fast fading; the final QI is estimated by combining all ranks. Blind/referenceless spatial quality evaluator (BRISQUE) [17] is a less computational complex method compared to [18] where, instead of wavelet, it employs Gaussian filter to extract low and high-frequency image components. Local phase coherence (LPC) of the wavelet image coefficients is employed in [14] to evaluate the image sharpness. The authors assume blur affects the LPC relationship near sharp image features and the degradation of LPC strength is employed to compute the image sharpness. The authors of [2] assess the image blur using a combination of natural scene statistics, multi-resolution decomposition, and machine learning. Based on training a probabilistic support vector machine, blur is measured using gradient histogram features. In [26], the method S3 evaluates the sharpness by combining both spectral (Fourier domain) and spatial (pixel domain) sharpness measurements. Spectral measure is based on the slope of the local magnitude spectrum and spatial measure is based on local maximum total variation. In [22], pixel scattering is used to determine the image content, re-

lying upon the fact that the noisier or blurrier the image is, the less entropy change is made by scattering the pixels. Its performance decreases as the noise properties deviate from AWGN since entropy becomes inaccurate for measuring the quality. NFERM [13] predicts the image structural degradation based on three groups of features; free energy model, structural information such as gradient magnitude, and naturalness. Then, extracted features are fed into a support vector regressor that has been already trained to obtain the quality score. GMLOG [29] proposes a learning based NR-IQA model that uses the joint statistics of simple gradient magnitude and Laplacian of Gaussian features. It uses joint adaptive normalization to boost the performance. MetricQ [30] is a local method that made a wider noise assumption. SVD of the local gradient is employed to exploit the sharpness and noise of patches. SVD is used to estimate the dominant direction and its perpendicular direction and energy of both are considered to estimate the quality of each patch based on estimated signal to noise ratio. The average of quality values of patches that contain relative high quality is considered as output quality. Although MetricQ addresses other types of noise in addition of AWGN, such as processed noise, the impact of noise in quality measurement is less emphasized by excluding noisy patches from the process. Our method also uses SVD of the local gradient to detect image content, however, dissimilar to MetricQ 1) we apply a Fourier shrinkage prior to orientation detection, 2) we take the Fourier sparsity of the patches into account, and 3) we address the blur by considering the absolute power of image signal.

3 Proposed method

3.1 Proposed Image Content Model

Assuming a transform such as SVD decorrelates a signal into orthogonal coefficients; generally most of the signal energy is represented in few coefficients creating a sparse representation of the image [8]. We employ this feature to differentiate between signal and noise to detect the true image content. Since the nature of the noise is random, it is unlikely to represent the noise in sparse form especially when the noise is represented in all orthonormal coefficients (e.g., white noise). Thus, the signal is more likely to be image content when it can be represented in a sparser form. We use SVD and DFT to maximize the chance of detecting content by benefiting the advantages of both. SVD is useful to detect single orientation signal, but cannot detect multi-orientation signals. On the other hand, DFT is beneficial in finding multi-orientation signals, but mistakes spatially correlated noise with signal.

High-frequency image components carries the edge information and the goal of a denoiser is to preserve them while removing the noise. We use the image gradient to extract edge information. For an image of interest I the gradients G_x, G_y, and the complex gradient image G can be calculated as

$$G_x = (H_s^T H_d) * I,\ G_y = (H_d^T H_s) * I,\ G = G_x + jG_y, \quad (1)$$

where $j = \sqrt{-1}$, $H_d = [-1\ 0\ 1]$, and $*$ denotes the two-dimensional convolution. Examples for H_s are $H_s = \frac{1}{2}[0\ 1\ 0]$ or the Sobel operator $H_s = \frac{1}{8}[1\ 2\ 1]$. Gradient orientations of the pixels on the edges are similar, whereas on the noisy pixels are random. The similarity of gradient orientations can be utilized to distinguish edge pixels and hence image content. We utilize SVD and DFT to find this similarity as follows.

3.1.1 Sparse DFT

Let x be a patch (block) of the gradient image G of size $N \times N$ which rearranged into a column vector with size of $N^2 \times 1$; thus x can be represented as a linear combination of orthonormal DFT dictionary (basis) matrices as

$$x = \mathbf{D}\alpha, \quad (2)$$

where $\mathbf{D}$ is a $N^2 \times N^2$ DFT dictionary and α is a $N^2 \times 1$ vector of complex numbers. The sparsity of DFT means that most of the power of α is concentrated in few elements. If we sort the α according to their magnitude from high to low to create $\hat{\alpha}$, we define l as the set of numbers that meets,

$$l = \left\{ n \mid n \in \mathbb{N}\ ,\ \frac{\sum_{k=1}^{n} |\hat{\alpha}_k|^2}{\sum_{k=1}^{N^2} |\hat{\alpha}_k|^2} \geq \delta \right\}, \quad (3)$$

where $0 \leq \delta \leq 1$ is a constant, $\hat{\alpha}_k$ means the k^{th} elements of $\hat{\alpha}$ and thus $\sum_{k=1}^{N^2} |\hat{\alpha}_k|^2$ is the whole energy of the patch. We define l_{min} as the minimum number that contains δ energy of the signal, $l_{min} = \min(l)$. For instance, when $\delta = 0.9$, at least 90% of the energy of α is assigned to the l_{min} element. We then define the inverse sparsity degree of DFT as follows,

$$\xi^{-1} = \frac{l_{min} \cdot \delta \sum_{k=1}^{N^2} |\hat{\alpha}_k|^2}{N^2 \sum_{k=1}^{l_{min}} |\hat{\alpha}_k|^2}, \quad (4)$$

The more sparse α is, the higher the ξ becomes. Fig. 1(b) shows an example of ξ for the image *Barbara* with $\delta = 0.75$.

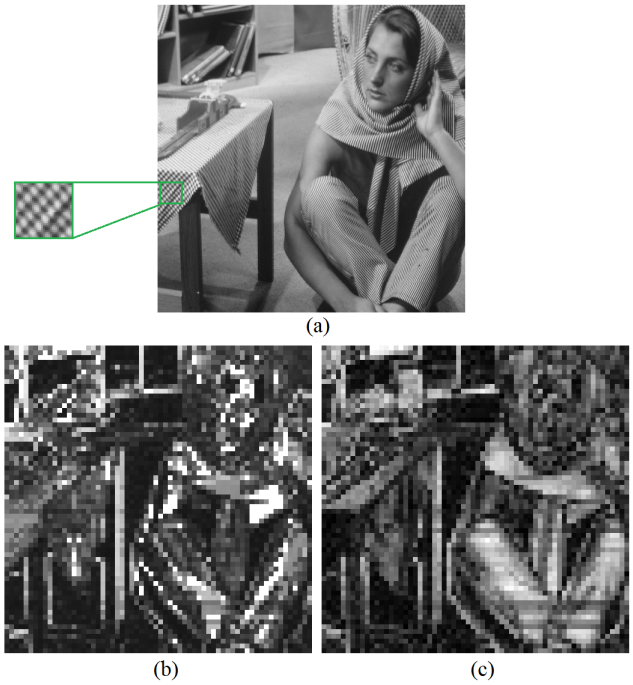

Fig. 1 Sparsity of gradients in 8×8 patches for the image *Barbara* (a). (b) shows the DFT sparsity ξ with $\delta = 0.75$ and (c) shows SVD sparsity β. Brighter patches are sparser. SVD sparsity cannot detect multi-orientation textures while DFT can.

3.1.2 Sparse SVD

Consider x a patch of gradient image G with size $N \times N$ is separated into real and imaginary part forming a $N^2 \times 2$ matrix $x_v = [x_r\ x_i]$. x_v can be decomposed into two singular values s_1 and s_2 [10,24] as

$$x_v = \begin{bmatrix} U_r\ U_i \end{bmatrix} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{5}$$

where $[U_r\ U_i]$ is an orthonormal matrix, meaning $U_r U_i^T = 0$, $s_1 \geq s_2 \geq 0$, and θ is a constant that represents the dominant orientation of G. When the signal energy is concentrated in one direction $s_1 \gg s_2$, *orientation dominancy* or *SVD sparsity* happens. Consider the SVD sparsity factor β as

$$\beta = \frac{s_1}{s_2} \geq 1. \tag{6}$$

The more the pixels of x are aligned in a single direction the higher the β becomes. On the other hand, if x contains pixels with random directions, β becomes small. Fig. 1(c) shows an example of β for the image *Barbara*. SVD sparsity locates single direction edges accurately but not image content with multi-direction repeated textures, as highlighted in Fig. 1(a), while DFT can.

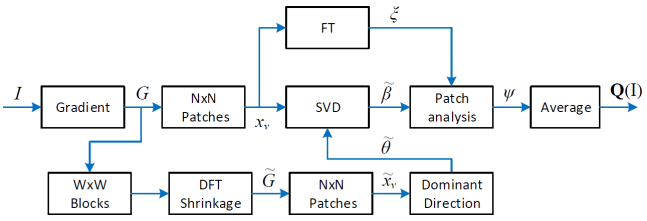

Fig. 2 Block diagram of the proposed algorithm.

3.2 Overview of proposed SDQI

The proposed method consists of three main steps. In the first step we compute the image gradient and apply a local Fourier shrinkage on the gradient image to generate a better approximation of the image signal. For speed consideration, we divide the gradient image into $W \times W$ overlapping blocks and we use two dimensional DFT to apply the shrinkage. In the second step, we divide the shrunk gradient image into non-overlapping patches of $N \times N$ and compute the dominant gradient orientation for each patch using the SVD analysis. In the final step, we divide the original gradient image into non-overlapping patches and we compute both SVD and DFT sparsity of each patch. We combine the sparsity information to measure the local signal noise power and finally image quality. The proposed NR-IQA method can be summarized as in Algorithm 1, and as in the block diagram of Fig. 2.

Algorithm 1: Proposed SDQI

i) **Compute** the complex gradient map G from the input image I using (1).
ii) **Divide** the gradient G into overlapping blocks of $W \times W$ and apply a Fourier shrinkage via (7).
iii) **Divide** the shrunk gradient $\tilde{G}$ into patches and compute the dominant direction $\tilde{\theta}$ for each patch via (11).
iv) **Divide** the gradient G into patches of $N \times N$ and compute the local sparsities using (4) & (6).
v) **Calculate** quality value for each patch via (13).
vi) **Output** the QI by averaging local values via (16).

3.3 Fourier shrinkage

The objective of this step is to increase the contrast of edge signals by suppressing the noisy Fourier coefficients. Assuming we divide the gradient image G into overlapping blocks of $W \times W$ and the DFT coefficients α are computed, a shrinkage procedure suppresses the noisy α and increases the contrast between signal and noise. We use the term *patch* to indicate an $N \times N$ image region and *block* for an $W \times W$ one. We assume α coefficients with a relative small magnitude are more likely to be noise, thus they should be suppressed. We

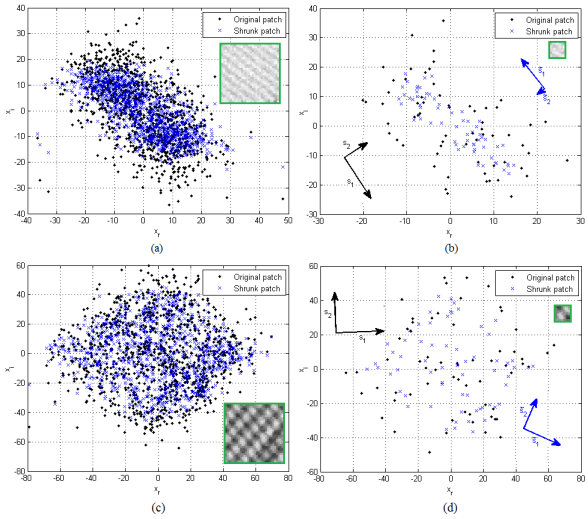

Fig. 3 Shrinkage with $W = 16$ changes the dominant orientation (angle of arrows). For a sample 32×32 part of *Barbara* (a) and (c), the shrinkage changes the dominant orientations of 8×8 patches (b) and (d).

use the median of $|\alpha|$ as a reference point. Let W be an even number and $\alpha_{med} = \frac{1}{2}(|\hat{\alpha}[\frac{1}{2}W^2 - 1]| + |\hat{\alpha}[\frac{1}{2}W^2]|)$ is the median of $|\alpha|$. In order to suppress small DFT coefficients we propose modifying each DFT coefficient as

$$\tilde{\alpha} = \alpha \cdot \exp\left(-\frac{c_\alpha \cdot \alpha_{med}^2}{|\alpha|^2}\right), \tag{7}$$

where c_α is a constant. (7) suppresses (or shrinks) the small coefficient relative to α_{med}. If α_{med} is small, i.e., α is sparse, the shrinkage has no effect. For each block, first α is computed and suppressed to $\tilde{\alpha}$ and then an inverse DFT is applied to obtain the modified (or the shrunk) gradient map in pixel domain. Since we are using sliding windows with overlapping, the results of the individual blocks are averaged to create the whole shrunk gradient map $\tilde{G}$. Let O be the size of overlapping in pixels; $\tilde{G}$ at each position is calculated from the average of $W^2/(W-O)^2$ blocks. For example, when $O = \frac{W}{2}$ the average of 4 blocks is required to calculate each pixels. Our idea is to set $W > N$ so a more global shrinkage on a larger block affects a local small patch. Fig. 3 shows how a global gradient shrinkage affects the dominant orientation, especially for smaller patches. Since the noise is suppressed, dominant orientation can be estimated more accurately after shrinkage.

3.4 Dominant orientation

When the pixels are locally aligned in a single orientation, there is a high chance of image content presence. To exploit this property, we divide the shrunk gradient map $\tilde{G}$ into non-overlapping patches of $N \times N$ and for each patch we form $\tilde{x}_v = [\tilde{x_r}\ \tilde{x_i}]$ by rearranging the real and imaginary part of patch values to form a $N^2 \times 2$ matrix. The dominant orientation of the shrunk patch $\tilde{\theta}$ is computed to meet (5). Generally, algorithms for computing singular values are related to eigenvalue computing of symmetric matrices. The QR algorithm [21] reduces rectangular matrix to *bidiagonal* using a *Householder* reduction. Although these iterative matrix computations give accurate results [7], their complexity makes them hard to implement. To compute the dominant orientation $\tilde{\theta}$, instead of iterative approaches, we propose a simpler solution. To meet the condition $U_r U_i^T = 0$ or equally $\tilde{s}_1 U_r \cdot \tilde{s}_2 U_i^T = 0$ in (5) we should have,

$$\sum_{k=1}^{N^2} (\tilde{x}_{r,k}\cos\tilde{\theta} + \tilde{x}_{i,k}\sin\tilde{\theta})(-\tilde{x}_{r,k}\sin\tilde{\theta} + \tilde{x}_{i,k}\cos\tilde{\theta}) = 0, \tag{8}$$

where $\tilde{x}_{r,k}$ and $\tilde{x}_{i,k}$ are the k^{th} element of $\tilde{x}_r$ and $\tilde{x}_i$, respectively; it follows,

$$\cos\tilde{\theta}\sin\tilde{\theta}\sum_{k=1}^{N^2}(\tilde{x}_{i,k}^2 - \tilde{x}_{r,k}^2) = (\cos^2\tilde{\theta} - \sin^2\tilde{\theta})\sum_{k=1}^{N^2}(\tilde{x}_{r,k}\tilde{x}_{i,k}). \tag{9}$$

With $cos(2\tilde{\theta}) = (\cos^2\tilde{\theta} - \sin^2\tilde{\theta})$ and $sin(2\tilde{\theta}) = \frac{1}{2}(\cos\tilde{\theta}\sin\tilde{\theta})$,

$$\tan(2\tilde{\theta}) = \frac{2\sum_{k=1}^{N^2}(\tilde{x}_{r,k}\tilde{x}_{i,k})}{\sum_{k=1}^{N^2}(\tilde{x}_{i,k}^2 - \tilde{x}_{r,k}^2)}, \tag{10}$$

so, the dominant orientation $\tilde{\theta}$ in $\tilde{x}_v$ can be computed as

$$\tilde{\theta} = \frac{1}{2}\tan^{-1}\left(\frac{2\sum_{k=1}^{N^2}(\tilde{x}_{r,k}\tilde{x}_{i,k})}{\sum_{k=1}^{N^2}(\tilde{x}_{i,k}^2 - \tilde{x}_{r,k}^2)}\right), \tag{11}$$

which is simpler to implement and faster compared to general iterative SVD computations. We use dominant orientation $\tilde{\theta}$ to compute SVD sparsity in (12).

3.5 Patch sparsity analysis

We developed a solution using two criteria; orientation dominancy (SVD sparsity) and DFT sparsity; we used them to analyze the likelihood of image signal presence in a patch. To exploit orientation dominancy, we divide the original gradient map G into patches of $N \times N$ and we use $\tilde{\theta}$ estimated from shrunk gradient map $\tilde{G}$ to compute the orientation dominancy of G. Assuming $\tilde{x}_v$ is a $N^2 \times 2$ of a patch of shrunk gradient $\tilde{G}$ at a certain

location and x_v is the corresponding patch of G at the same location; First, we compute the dominant orientation of shrunk patch $\tilde{\theta}$ according to (11) for $\tilde{x}_v$ and then we use it to calculate the modified singular values of x_v, i.e., $\tilde{s}_1$ and $\tilde{s}_2$ as

$$\tilde{s}_1 = \sqrt{\sum_{k=1}^{N^2} \left| x_{r,k} cos\tilde{\theta} + x_{i,k} sin\tilde{\theta} \right|^2}$$
$$\tilde{s}_2 = \sqrt{\sum_{k=1}^{N^2} \left| x_{i,k} cos\tilde{\theta} - x_{r,k} sin\tilde{\theta} \right|^2}. \quad (12)$$

$\tilde{s}_1$ and $\tilde{s}_2$ are different from s_1 and s_2 the singular values of x_v (see Fig. 3). In (12) we use $\tilde{\theta}$ the dominant orientation of shrunk patch instead of θ the dominant orientation of x_v. $\tilde{s}_1$ is the energy of signal along the $\tilde{\theta}$ and $\tilde{s}_2$ is the energy of along the perpendicular direction $(\pi/2 - \tilde{\theta})$. Fig. 4 illustrates how these energies are computed according to (12). Using (6), we define the sparsity of SVD after shrinkage as $\tilde{\beta} = \frac{\tilde{s}_1}{\tilde{s}_2}$. When the singular values are sparse, i.e., $\tilde{s}_1 \gg \tilde{s}_2$ (or $\tilde{\beta} \gg 1$), the probability of image content presence is higher. Theoretically $\beta \geq 1$; however, it is not guaranteed that $\tilde{\beta} \geq 1$. $\tilde{\beta} \leq 1$ implies that the probability that the patch contains image signal is low. We propose a likelihood function that maps this property to the local quality of the patch as

$$\psi = \frac{(\tilde{\beta} - 1 - \epsilon)}{\tilde{\beta} + \tilde{\beta}_0}, \quad (13)$$

where $\epsilon \geq 0$ and $\tilde{\beta}_0 \geq 0$ are computed in (14) and (15) based on $\tilde{s}_1$ and ξ. ψ is a value indicating the relative quality of the patch. In case $\tilde{\beta} \leq 1$, ψ becomes negative, implying that the patch contains no useful signal. In related work [30], patches with small signal to noise ratio are rejected and the effect of noise in noisy patches is not considered. Our idea is that ψ can be negative to highlight the impact of noise in overall image quality. We can consider $|\psi|$ as a probability that shows signal presence in the case of $\psi > 0$ and noise existence in case of $\psi < 0$. Therefore, the effect of noise is more considered in our algorithm. We propose to use the DFT sparsity of x_v, i.e., ξ to compute the ϵ,

$$\epsilon = \max(\xi^{-1} - \xi_{max}^{-1}, 0), \quad (14)$$

where ξ_{max} is a constant representing a relatively large value for ξ. When $\xi \geq \xi_{max}$, i.e., the DFT is very sparse, $\epsilon = 0$. On the other hand, when ξ is relatively small, ϵ becomes non zero. In this case when $\tilde{\beta}$ is also relatively small ($\tilde{\beta} < \epsilon$), ψ becomes negative. In fact, ϵ is an adjustment to increase the reliability by taking the DFT sparsity into account. Fig. 5 shows ψ with $\tilde{\beta}_0 = 0$ for different values of ϵ. In (13), only relative values of decomposed signal, i.e., ratios of high power to low-power coefficients are considered. In a weak-textured patch, it is possible that the absolute values $\tilde{s}_1$ and $\tilde{s}_2$ are small but their relative value $\tilde{\beta}$ is large. In order to detect blur and compression artifacts, weak-textured patches should be addressed by ranking lower the smaller s_1. We define $\tilde{\beta}_0$ to adjust the quality of patch as

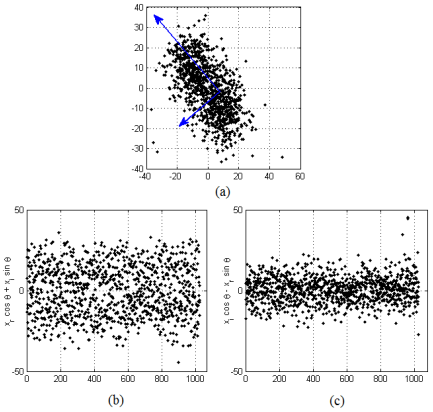

Fig. 4 Example of computing the singular values for (a) a gradient patch. Samples are rotated to be aligned in the dominant (b) and perpendicular (c) orientation. The energy $s_1 = 505.8$ along the dominant orientation is greater than the energy $s_2 = 267.9$ along the perpendicular orientation.

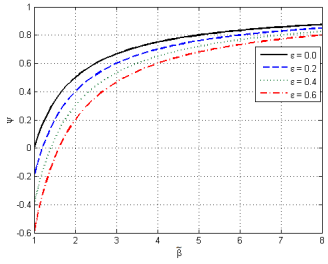

Fig. 5 Quality of a patch ψ according to the SVD sparsity $\tilde{\beta}$ and DFT based adjustment ϵ.

$$\tilde{\beta}_0 = \frac{c_\beta^2}{c_\beta^2 + \tilde{s}_1^2}, \quad (15)$$

where c_β is a constant. Smaller value of $\tilde{s}_1$ compared to c_β leads to larger value of $\tilde{\beta}_0$ and thus smaller ψ. Increasing c_β makes ψ more sensitive to blur, but decreases the sensitivity to noise. Thus, to define c_β, a suitable trade-off between noise and blur should be considered (see section 4).

3.6 Quality index

In (13) we have defined ψ as a measure to quantify the probability of the signal presence in each patch.

Assuming the directional energy $\tilde{s}_1$ is the signal of interest, we consider its expected value as a measure for genuine image content. The expected energy of signal is computed by multiplying $\tilde{s}_1$ by its presence probability ψ. By aggregating all of genuine energies (i.e., expected values), we compute the overall genuine energy for the entire image to quantify the quality of the image $\mathbf{Q}(I)$, assuming the input image contains K patches,

$$\mathbf{Q}(I) = \frac{1}{K}\sum_{k=1}^{K} \tilde{s}_{1,k}\psi_k, \tag{16}$$

where $\tilde{s}_{1,k}$ and ψ_k are the $\tilde{s}_1$ and ψ_k of the k^{th} patch. Negative value of ψ_k implies presence of noise without any signal. Thus, when $\psi < 0$, $|\psi|$ is the probability of noise presence with no genuine signal which leads to negative $\tilde{s}_{1,k}\psi_k$. In theory, $\mathbf{Q}(I)$ can be negative which means the power of noise is more than signal. In practice, only relative result of $\mathbf{Q}(I)$ is informative, therefore, the sign of $\mathbf{Q}(I)$ is important and a negative QI shows lower quality than a positive QI.

4 Experimental results

To evaluate the performance of the proposed SDQI, seven state-of-the-art methods BRISQUE [17], CPBD [19], JNB [11], LPC [14], S3 [26], BIQI [18], and MetricQ [30] have been compared objectively and subjectively. All analyses are performed on the gray-level image, however, there is no restriction for performing the algorithm on other channels.

We have run extensive simulations to set the algorithm's parameters: N SVD patch size, W shrinkage block size, O shrinkage overlapping size, c_α of (7), δ of (3), ξ_{max} of (14), and c_β of (15). N should be small enough to contain a distinct orientation. Since the proposed algorithm detects one dominant orientation, a large patch may contain many different orientations which cannot be accurately detected. However, for an accurate estimation of orientation, sufficient number of pixels are required and a very small patch does not satisfy this condition. Considering that the DFT operation is faster when N is power of 2, our experiments show that $N = 8$ is optimal for the performance. We set the shrinkage window size $W = 2N$ to process the image details more globally before analyzing the patch. We set $O = \frac{W}{2}$ and our experiments show that by increasing O (e.g., $O = \frac{3W}{4}$) the performance slightly improves; however, it does not justify the higher computational complexity (e.g., 4x). We have analyzed the effect of Fourier shrinkage on the performance of the algorithm by altering c_α to change the shrinkage strength. For this, we used the denoising methods BM3D [5] and bilateral filter [25] and we changed the noise removal force, i.e., input standard deviation of noise σ_n, and measured the output QI. Fig. 6 shows the QI of the proposed method by changing both σ_n and c_α. By increasing c_α, the ability of detecting noise increases by providing lower QI for noisy image ($\sigma_n = 0$); however, the capability of detecting blur decreases since the QI peak shifts towards higher σ_n. Thus, we set $c_\alpha = 4$ as a balanced trade-off to detect both noise and blur and leads to better results in most of our experiments. By conducting extensive simulations we determined other algorithm's parameters $\delta = 0.75$, $\xi_{max} = 8$, and $c_\beta = 20$ that give the highest correspondence with ground-truth quality metrics PSNR and MSSIM.

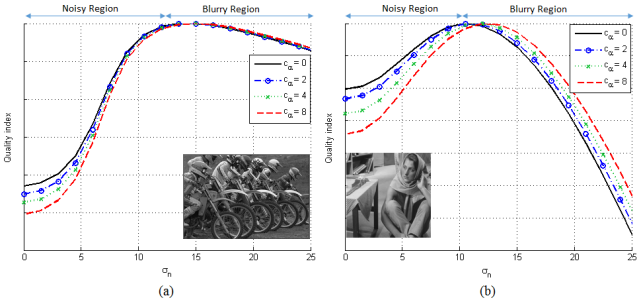

Fig. 6 The effect of Fourier shrinkage on the performance of the proposed algorithm. The noisy images with added $\sigma_a = 10$ (28dB) are denoised by BM3D [5] (a) and (b) bilateral filter [25] using different input σ_n. Increasing c_α increases the capability of detecting noise, but shifts the maximum QI to a blurrier result, i.e., higher σ_n. $c_\alpha = 0$ means no shrinkage.

4.1 Optimizing denoiser parameters using NR-IQA

In image or video denoising, the goal is to remove noise without degrading the sharpness (i.e., introducing blur). Thus, finding an optimal point between noise and blur is the key to achieve the highest quality. If the key denoising parameter $\mathbf{p}$ is not well selected, the output will be degraded with either noise or blur. A NR-IQA which detects the genuine image content, such as edges of physical objects, local sharpness, and textures, can be used to select such key parameter. Assuming I is the observed noisy image and a filtering process outputs the filtered image $I_{\mathbf{p}}$ using the input parameter $\mathbf{p}$. As proposed by [30] by changing $\mathbf{p}$ and measuring the output quality using a NR-IQA, the denoiser output can be optimized. In order to evaluate the performance of NR-IQA, we consider that the ground-truth quality metric such as mean opinion scores (MOS), PSNR, or MSSIM is available and denoted by $\Phi(\cdot)$, which measures the quality of the NR-IQA based denoiser output. Assuming $\mathbf{QI}(I_{\mathbf{p}})$ is the quality index measured by a NR-IQA, the NR-IQA based denoiser output leads to the highest

$\mathbf{QI}(I_\mathbf{p})$ as

$$I_\mathbf{opt} = \underset{I_\mathbf{p}}{\text{argmax}} \left[\mathbf{QI}(I_\mathbf{p})\right], \quad \Phi_{gtm} = \Phi(I_\mathbf{opt}, I_\mathbf{ref}), \qquad (17)$$

where $I_\mathbf{opt}$ is the output of NR-IQA based denoiser at highest QI, $I_\mathbf{ref}$ is the reference image, and Φ_{gtm} is its quality according to the reference and considered ground-truth. Due to imperfection of NR-IQA, the output quality may deviate from maximum achievable quality. Fig. 7 (top) is an example of computing Φ_{gtm} for NR-IQA methods BIQI, MetricQ, and SDQI for *Peppers* contaminated with AWGN with $\sigma_a = 10$ (PSNR = 28dB) and *I05* contaminated with spatially correlated noise. Fig. 7 (bottom) shows the NR-IQA vs. PSNR and NR-IQA vs. MSSIM scatter plots where QIs are normalized. The methods that are more positively correlated with the ground-truth and reach higher quality are more desirable to be used as denoising parameter selector. Scatter plots shows a positive correlation between NR-IQA and both PSNR and MSSIM for MetricQ and ours with higher values for ours. For BIQI the correlation is not consistent. Negative correlation happens when the reference quality (MOS) and QI behave inversely. For example, by adding more noise, MOS shows less quality but QI shows more quality. If the source of distortion (e.g., noise or blur) was known or the negative correlation was consistent for all distortions, we could use the negative correlation and adopt it inversely which leads to positive correlation. In Tables 2, 4, and 6 the methods CPBD, JNB and S3 give consistent negative correlation. Thus, to have a fair comparison negative signs of correlation for these methods were removed. For inconsistent negative values we did not change the sign.

Fig. 8 shows a block diagram for optimizing the parameters of denoising (NR-IQA based denoiser). We use Φ_{gtm} to evaluate the NR-IQA according to ground-truth metric (see Tables 3, 5 and 7). In Fig. 8 only the maximum value of QI is taken into account, however, the behaviour of all QIs with respect to ground-truth should also be examined. We consider two well-known correlation factors, Spearman and Kendall rank order correlation coefficient (SROCC and KROCC) to evaluate the performance of NR-IQA. We change the denoiser parameter and compute QI of denoiser output. We then compute SROCC and KROCC using the set (one value for each parameter) of QI and the corresponding ground-truth. Since SROCC, KROCC, and Φ_{gtm} do not necessarily match (compare Table 2 and 3) we consider all of them for our evaluations. When the reference is available and the degradation is noise or blur, many studies show a high correspondence between subjective and objective measures such as PSNR

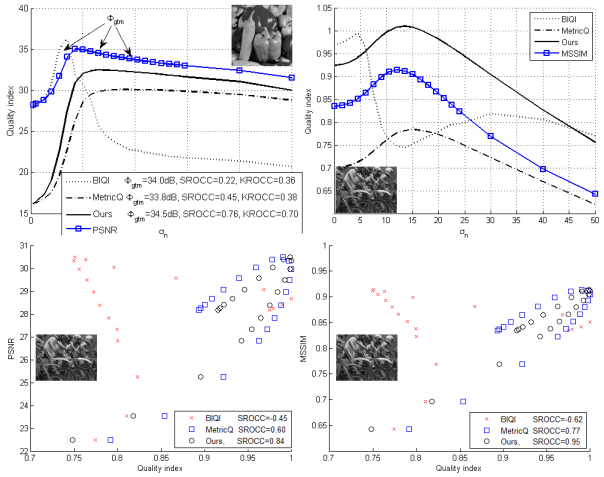

Fig. 7 Top: evaluation of NR-IQA methods in selecting the denoiser BM3D parameter for *Peppers* (left) and *I05* (right) corrupted with AWGN and spatially correlated noise considering PSNR and MSSIM as the ground-truth quality metric. Bottom: scatter plot of NR-IQA vs. PSNR and NR-IQA vs. MSSIM for *I05*.

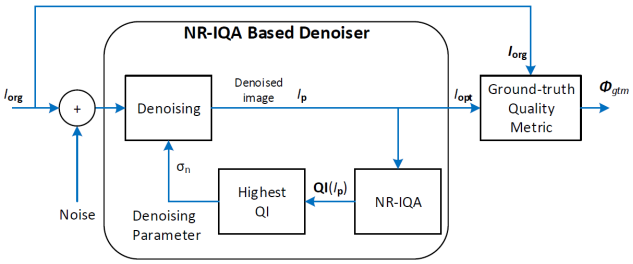

Fig. 8 Optimizing the parameters of a denoiser using a NR-IQA method and finding the quality of output based on the ground-truth quality metric.

and MSSIM (see [20]). Depending on the availability of the reference image, in our simulations we have used subjective metric MOS, and objective metrics, PSNR and MSSIM, as the ground-truth. Subjective metrics are more reliable, however, a fair MOS requires a large dataset and manpower.

4.2 Real Noise

To analyze the performance of the NR-IQA methods under real noise we have selected 25 real noisy images and video frames (see Fig. 9). We applied BM3D on each image and created sets of denoised images with two levels; medium (i.e., most of the noise is removed but image details are kept) and strong (i.e., noise as well as some image details are removed). For video frames where the temporal data was available we used VBM3D [4] and extracted one frame and added to the dataset. We have conducted a human subject study to obtain MOS. Two images, one noisy one denoised or both de-

Fig. 9 Samples of our image dataset corrupted with real noise.

noised were displayed on a full HD monitor and observers were asked to select a better image between the two and score them. We used five levels as recommended by ITU to compare the quality of right image compared to left one. The scores were "Bad", "Poor", "Fair", "Good", and "Excellent" equivalent to scores 1 to 5 for right image and 5 to 1 for left image. If both images have the same quality each image gets the score of 3. Since each image has scored twice, the minimum and maximum of MOS is 2 and 10, however, the results show no average MOS values equal to 2 or 10. The display order of right or left image for being noisy or denoised was based on a predefined random pattern. All observers were told to have a normally preferred distance to the monitor and not take more than 15 seconds for each pair-wise test. Total of 75 comparison for each subject took place and for each subject the time of the test was between 10 to 20 minutes. The total number of scores for each image were averaged as the final MOS for that each image. A total of 30 human subjects between the age of 20 and 55 participated in our experiments. We have obtained the MOS for all 75 images in the dataset. In average, MOS of original noisy, denoised by medium level, and denoised by strong level were 4.48, 8.01, and 5.51, respectively.

We have calculated the SROCC and KROCC for the real noise dataset. Table 1 shows the SROCC and KROCC result of each NR-IQA method. Our methods shows higher positive correlation with MOS. In CPBD, JNB, and S3 noise is less taken into account and for all cases the noisier images shows higher QI compared to denoised ones. BIQI and BRISQUE also tend to select noisier images over the denoised ones. One the other hand, MetricQ and LPC tend to select blurrier images with destructed details. Fig. 10 shows part of original and denoised images by different levels from our real noisy dataset. BIQI and BRISQUE select the noisiest (left column) as the highest quality. MetricQ selects the blurriest (third column) as the highest quality and ours selects the ones with highest MOS.

Table 1 Correlation factors between MOS and NR-IQA methods for denoising of real noise.

	BRISQUE [17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	MetricQ[30]	Ours
SROCC	-0.32	0.35	0.32	0.40	0.31	-0.15	0.37	**0.84**
KROCC	-0.19	0.23	0.20	0.31	0.19	-0.07	0.26	**0.73**

4.3 Synthetic noise

We have evaluated the performance the NR-IQA methods by analyzing the behavior of each method in finding the best balance between noise and blur in image denoising under synthetic AWGN and synthetic processed noise. We use both correlation factor (SROCC and KROCC) and Φ_{gtm} (quality at maximum QI) to objectively evaluate the performance of NR-IQA methods (see Fig. 8). We consider two well-known reference-based quality metrics PSNR and MSSIM [28] as the ground-truth quality metric. We have considered the TID2013 [20] database as the ground-truth and added synthetic noise, then, we have varied the main parameter of the denoiser (here σ_n and σ_{RF3D}) using 14 steps. We have computed the correlation between ground-truth quality metric and computed QI. We have considered two high-performance denoising methods BM3D and PID [15] in these experiments. The statistical significance of the correlation factors is related to the number of data points used in the correlation computation. For synthetic experiments, we used p-value to measure the significance of the correlation factors. For SROCC, p-value is lower than 10^{-3} when the absolute of correlation is less than 0.18. For the KROCC, this value is 0.12. We also considered the quality of NR-IQA output in these experiments. The PSNR and MSSIM values at the maximum QI are measured and compared (see Tables 3, 5, and 7). The NR-IQA method that gives the highest Φ_{gtm}, is more suitable to be used in a NR-IQA based denoiser design (see also [30]).

4.3.1 Synthetic AWGN

In this experiment we added zero-mean AWGN to the ground-truth images from TID2013. The noisy images were generated by adding AWGN with standard deviation σ_a of 10 (PSNR = 28dB). For all synthetic noise tests using TID2013, we consider the standard deviation of noise σ_n as the main parameter of the denoiser and we have varied that using 15 different levels of denoising from relatively small (which leads to noisy results) to large values (which leading to blurry results). Table 2 compares the SROCC and KROCC between different NR-IQA methods. In case of BM3D as the denoiser, BRISQUE clearly outperforms other methods followed by BIQI and proposed SDQI. In case

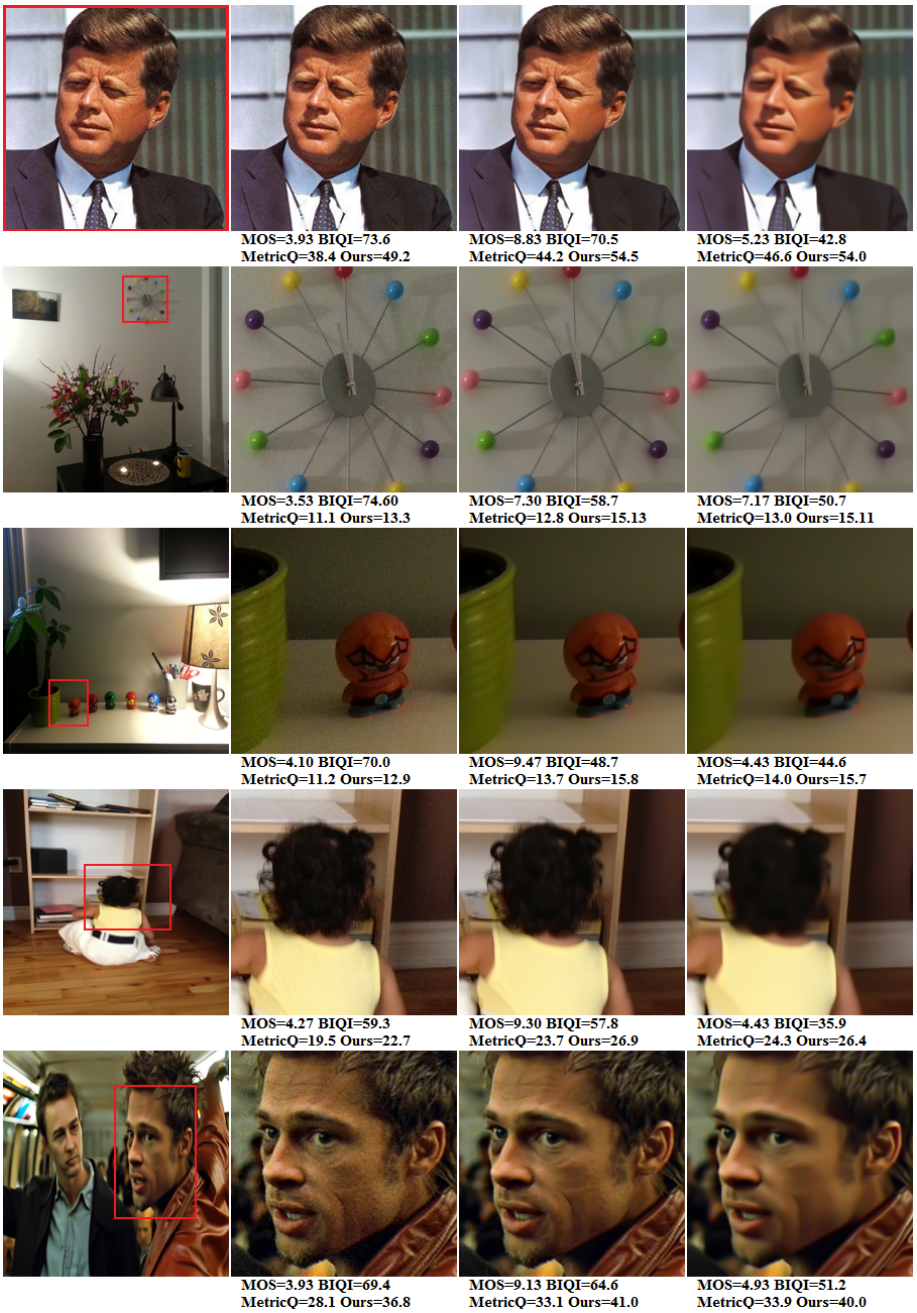

Fig. 10 Original, part of original, and two denoised with BM3D using two σ. SDQI selects the output with the highest MOS.

of PID as the denoiser, proposed method outperforms other methods followed by MetricQ and BRISQUE. CPBD, JNB, and S3 are more sensitive to blur and less to noise which yield negative correlations. Table 3 compares the Φ_{gtm} for NR-IQA methods averaged over TID2013 database. BRISQUE achieves slightly higher PSNR and MSSIM as proposed method. The performance of BRISQUE and BIQI is relatively higher when the noise is white, however, according to Table 1 it degrades when the noise is non-white. It is worth noting that we have also tested the performance of NR-IQA methods under signal-dependent noise. In this case the variance of noise is a function of image intensity. We selected a "close to reality" noise level function, i.e., variance of noise at each intensity. We computed the average Φ_{gtm} and the results show QI values relatively similar to Table 3 for all eight methods.

We have considered the case that the adjustable parameter of the denoiser is not the standard deviation of the noise. Fig. 11 shows the PSNR result of NR-IQA based denoiser using RF3D [16] as the video denoiser. Key input parameter of RF3D is the power spectral

Table 2 AWGN: Correlation factor for TID2013 database.

	BRISQUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	MetricQ[30]	Ours
BM3D denoiser, correlation with PSNR								
SROCC	**0.78**	0.28	0.28	0.47	0.28	0.58	0.50	0.57
KROCC	**0.66**	0.17	0.18	0.38	0.17	0.51	0.43	0.50
BM3D denoiser correlation with MSSIM								
SROCC	**0.73**	0.35	0.35	0.49	0.35	0.51	0.54	0.61
KROCC	**0.64**	0.19	0.20	0.39	0.19	0.47	0.46	**0.69**
PID denoiser, correlation with PSNR								
SROCC	0.75	0.33	0.33	0.42	0.33	0.60	0.73	**0.76**
KROCC	0.65	0.22	0.22	0.31	0.23	0.51	0.65	**0.69**
PID denoiser, correlation with MSSIM								
SROCC	0.72	0.37	0.37	0.42	0.37	0.55	0.73	**0.76**
KROCC	0.53	0.63	0.23	0.23	0.31	0.24	0.49	0.65

Table 3 AWGN: Quality of NR-IQA based denoiser, Φ_{gtm}, averaged over TID2013 database.

	BM3D							
	BRISQUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	MetricQ[30]	Ours
PSNR	**33.47**	28.24	28.24	31.47	28.22	32.93	32.06	32.46
MSSIM	**0.90**	0.69	0.69	0.84	0.68	0.89	0.86	0.87
	PID							
PSNR	**33.35**	28.26	28.26	30.68	28.23	33.11	32.96	33.06
MSSIM	**0.90**	0.69	0.69	0.81	0.68	0.89	0.88	0.88

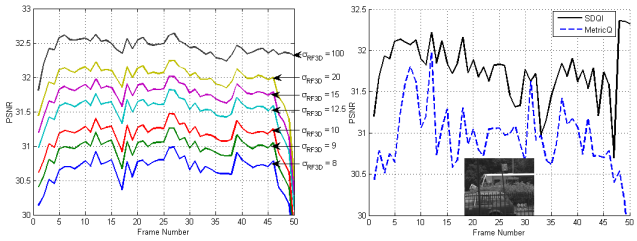

Fig. 11 Selecting the parameter of RF3D [16] video denoiser. (a) PSNR with different parameters for the 25dB *Bus* video. (b) PSNR of denoiser output when σ_{RF3D} is selected using MetricQ and proposed method. Since noise is white, higher σ_{RF3D} (flat PSD) leads to higher quality.

density (PSD) which is defined by a 2D Gaussian low-pass filter with different sigma σ_{RF3D}. Proposed method can better, i.e., leads to outputs with higher PSNR, select the parameter of RF3D σ_{RF3D} compared to MetricQ.

4.3.2 Spatially correlated noise

Camera noise usually becomes manipulated due to processing such as filtering, lossy compression, or demosaicing. Thus, in order to evaluate our method under this real conditions, we assume that the noise is spatially correlated (similar to noise after demosaicing, upscaling, or filtering) and we generated noisy images by adding filtered AWGN to the ground-truth images from TID2013. Noisy images were denoised using BM3D and PID with 15 levels of denoising. We used 5×5 Gaussian filter with sigma of 0.6 and $\sigma_a = 20$. Table 4 compares the SROCC and KROCC between selected NR-IQA methods using PSNR and MSSIM as the ground-truth. For both BM3D and PID, proposed SDQI outperforms other methods followed by MetricQ and LPC.

Table 4 Spatially correlated noise: Correlation factor for TID2013 database.

	BRISQUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	MetricQ[30]	Ours
BM3D denoiser, correlation with PSNR								
SROCC	-0.01	0.24	0.26	0.55	0.27	0.49	0.53	**0.63**
KROCC	0.10	0.20	0.21	0.48	0.24	0.41	0.47	**0.57**
BM3D denoiser correlation with MSSIM								
SROCC	-0.15	0.37	0.40	0.52	0.41	0.35	0.63	**0.70**
KROCC	0.01	0.29	0.31	0.46	0.33	0.33	0.56	**0.65**
PID denoiser, correlation with PSNR								
SROCC	-0.25	0.29	0.33	0.51	0.34	0.43	0.73	**0.76**
KROCC	-0.23	0.26	0.27	0.45	0.31	0.38	0.66	**0.70**
PID denoiser, correlation with MSSIM								
SROCC	-0.32	0.39	0.42	0.48	0.43	0.33	0.73	**0.75**
KROCC	-0.28	0.32	0.34	0.42	0.37	0.32	**0.67**	**0.67**

Table 5 Spatially correlated noise: Quality of NR-IQA based denoiser, Φ_{gtm}, averaged over images of TID2013 database.

	BM3D							
	BRISQUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	MetricQ[30]	Ours
PSNR	29.89	28.60	28.41	30.69	28.21	31.16	30.77	**31.23**
MSSIM	0.78	0.72	0.71	0.83	0.70	0.84	0.84	**0.85**
	PID							
PSNR	28.88	28.63	28.40	29.95	28.21	30.99	31.57	**31.65**
MSSIM	0.74	0.72	0.71	0.80	0.70	0.83	**0.86**	**0.86**

Table 4 results corresponds with real noise results in Table 1. The performance of BRISQUE and BIQI degrades as the noise deviates from whiteness and in some cases yield negative correlations. Similar to AWGN, CPBD, JNB, and S3 give negative correlations. Table 5 compares the average of Φ_{gtm} for considered NR-IQA. Our method achieves more accurate results followed by BIQI in case BM3D denoiser and MetricQ in case PID denoiser. Comparing Table 4 and Table 2 gives an idea about the sensitivity of methods to the high-frequency components of the noise. Performance of BIQI and BRISQUE significantly decreases in this situation while SDQI shows stable performance. Fig. 12 shows visual quality comparison, applying BM3D where the filter parameter σ_n is selected using BIQI, MetricQ, LPC and proposed. BIQI leads to noisy results, however, MetricQ and LPC yield blurry results (which correspond to the results in Table 1).

4.3.3 Lossy compressed noise

Images are often lossy compressed. Thus, we repeated the above experiments by applying a lossy compression on the noisy images. Noisy images were generated by adding AWGN with $\sigma_a = 10$ to TID2013 database, then we compressed them using standard JPEG with quality factor (QF) of 75, finally we denoised them using BM3D and PID with 15 levels of denoising. Table 6 compares the SROCC and KROCC between selected NR-IQA methods using PSNR and MSSIM as the ground-truth. For both BM3D and PID, proposed SDQI outperforms other methods followed by MetricQ and LPC. Similar to spatially correlated noise, the per-

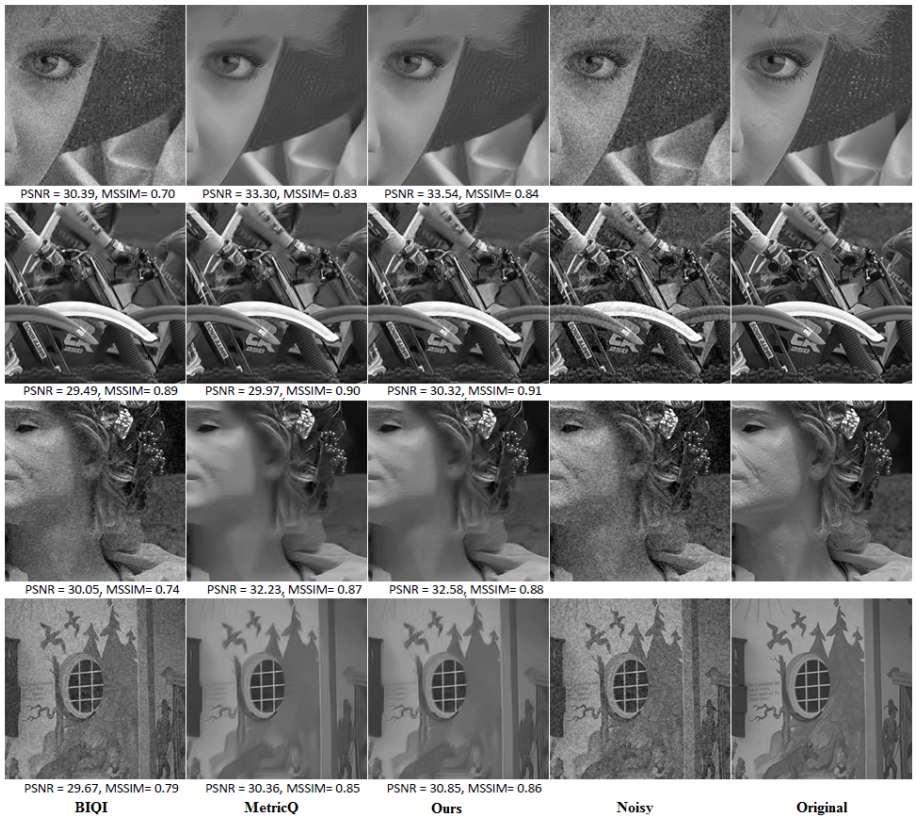

Fig. 12 Visual comparison: selecting BM3D parameter for denoising images from TID2013 corrupted with spatially correlated noise.

Table 6 Lossy compressed noise: Correlation factor for TID2013 database.

	BRIS-QUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	Metri-cQ[30]	Ours
BM3D denoiser, correlation with PSNR								
SROCC	0.30	0.39	0.39	0.50	0.39	0.33	0.58	**0.65**
KROCC	0.30	0.27	0.28	0.42	0.28	0.31	0.49	**0.57**
BM3D denoiser correlation with MSSIM								
SROCC	0.27	0.41	0.42	0.49	0.41	0.29	0.59	**0.64**
KROCC	0.30	0.27	0.27	0.40	0.27	0.30	0.49	**0.56**
PID denoiser, correlation with PSNR								
SROCC	0.32	0.44	0.44	0.38	0.44	0.34	0.79	**0.82**
KROCC	0.26	0.34	0.33	0.31	0.34	0.32	0.71	**0.75**
PID denoiser, correlation with MSSIM								
SROCC	0.32	0.44	0.44	0.37	0.44	0.32	0.76	**0.78**
KROCC	0.27	0.31	0.30	0.29	0.31	0.32	0.68	**0.71**

formance of BRISQUE and BIQI degrades as the noise becomes lossy compressed and CPBD, JNB, and S3 give negative correlations. Table 7 compares the average of Φ_{gtm} for considered NR-IQA methods using PSNR and MSSIM as the ground-truth. The proposed SDQI is able to select a more accurate σ_n compared to other methods, suggesting it being more suitable for denoising applications.

Table 7 Lossy compressed noise: Quality of NR-IQA based denoiser, Φ_{gtm}, averaged over images of TID2013 database.

	BM3D							
	BRIS-QUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	Metri-cQ[30]	Ours
PSNR	31.64	27.91	27.92	31.01	27.90	31.70	31.66	**31.92**
MSSIM	**0.86**	0.69	0.69	0.83	0.69	**0.86**	0.85	**0.86**
PID								
PSNR	31.73	27.91	27.95	30.23	27.90	31.85	32.39	**32.42**
MSSIM	0.86	0.69	0.69	0.79	0.69	0.86	**0.88**	**0.88**

4.4 General quality assessment

We have tested our algorithm, independent of denoising, in general degradation conditions. We have used all distortion provided in TID2013 database. Different types of noise such as AWGN, AWGN in color components (AWGN-C), spatially correlated (SPCN), masked (MSKN), high frequency (HFN), impulse (IMPN), quantization (QTNN) are considered. We considered also different types of blur such as Gaussian blur (G-Blur), denoising of a noise-free image (Blur-DEN) which is a type of anisotropic blur. Table 8 compares the SROCC values, considering MOS as the ground-truth where the proposed method gives higher values on average. The distortion *change of color saturation* (color-SAT) is not applicable since the NR-IQA are applied on the luminance component which remains unchanged in this type

Table 8 SROCC values for NR-IQA for all distortions in the TID2013 database using the MOS as the ground-truth.

	BRIS-QUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	Metri-cQ[30]	Ours
AWGN	0.85	-0.78	-0.73	0.56	-0.74	0.86	0.86	0.87
AWGN-C	0.84	-0.74	-0.62	0.41	-0.48	0.77	0.84	0.85
SPCN	-0.55	0.50	0.01	0.62	-0.57	0.81	0.83	0.69
MSKN	0.83	-0.60	-0.57	0.83	-0.59	0.79	0.83	0.85
HFN	0.73	-0.78	-0.73	0.80	-0.75	0.86	0.85	0.86
IMPN	0.84	-0.47	0.17	0.73	-0.53	0.80	0.79	0.86
QTNN	0.84	-0.70	-0.66	0.23	-0.61	0.51	0.13	0.62
G-Blur	0.87	0.86	0.87	0.83	0.83	0.85	0.82	0.85
DEN-Blur	0.79	0.84	0.78	0.70	0.85	0.34	0.81	0.82
JPEG	0.85	0.44	0.72	0.41	0.57	0.84	-0.08	0.47
JPEG2K	0.85	0.86	0.86	0.48	0.82	0.78	0.84	0.79
JPEG-Err	0.76	0.48	0.36	0.74	-0.29	0.54	0.70	0.72
JPEG2K-Err	0.54	0.52	0.47	0.72	-0.24	0.67	0.44	0.45
NEPN	0.69	-0.13	0.01	0.08	-0.46	-0.08	0.73	0.77
LBWD	0.65	0.11	0.17	0.42	0.72	0.41	0.66	0.70
Mean-Sh	0.39	-0.04	-0.11	0.36	-0.08	0.29	0.62	0.63
Contrast	-0.18	-0.28	-0.54	0.85	0.84	0.41	0.82	0.83
Color-SAT	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
MGN	0.85	-0.67	-0.57	0.73	-0.64	0.85	0.86	0.87
COMFN	0.32	0.85	0.78	0.59	0.64	0.58	0.78	0.77
LCNI	0.83	-0.19	0.47	0.49	0.56	0.71	0.86	0.87
ICQWD	0.84	-0.64	-0.60	0.73	0.16	0.78	0.84	0.85
CH-ABER	0.80	0.81	0.77	0.78	0.84	0.76	0.78	0.78
SS-REC	0.84	0.84	0.81	0.69	0.85	0.80	0.76	0.77
Average	0.66	0.05	0.09	0.60	0.07	0.65	0.71	0.76

of distortion. The proposed method provides higher correlation in average and in most of *actual* noise and blur distortions.

4.5 Implementation issues

The source codes of algorithms [18,17,19,11,14,30,26, 13] were obtained from the author's websites. In BIQI and BRISQUE the maximum and minimum quality happens at 0 and 100 respectively. To obtain the consistency with other methods which give higher QI for higher quality, we subtracted the QI of these methods from 100. We implemented our method using *MATLAB* (MEX). We have tested the acceleration factor of proposed SVD computation approach by comparing it to standard MATLAB *'svd'* function. For the patch size of $N = 32$ for instance, it shows a speedup by a factor of 1.9. Our approach saves, for example, 0.2 seconds for calculating the SVD of a full HD image with resolution of 1920×1080 with patch size of 8. We took the advantage of block-based operation to utilize the ability of parallel processing and accelerated our implementation using capability of GPU in parallel processing. We used OpenCL programming language to implement our method. Table 9 compared the speed of different methods to compute the QI of full HD image (1920×1080). For all tests we used Intel 3.07 GHz, i7 CPU and NVIDIA GTX 970 GPU. Table 9 cannot fully reflect the speed of algorithms since the implementations of the different algorithms may not be optimized. However, computation time gives an idea about the lower bound of the speed. We placed our *MATLAB* code, test datasets, and other supplementary materials on our project website `http://users.encs.concordia.ca/~amer/SDQI/`.

Table 9 Elapsed time in seconds to process a full HD image.

BRIS-QUE[17]	CPBD [19]	JNB [11]	LPC [14]	S3 [26]	BIQI [18]	Metri-cQ[30]	Ours	Ours GPU
0.67	4.47	7.29	6.94	188.8	1.03	9.31	0.75	0.018

5 Conclusion

In this paper, we presented a new no-reference image quality assessment approach for denoising based on single value decomposition and Fourier transform which are used to estimate the dominant orientation and coefficient sparsity. We propose a fast and easy to implement method to calculate the SVD avoiding recursive operations. Based on SVD analysis and Fourier sparsity, we measure local image structure, noise, and blur, and integrate them to compute overall quality. Our method takes both noisy and structured patches into account, therefore it measures the effect of noise more precisely compared to state-of-the-art methods. To reach more accurate results, especially under heavy noise, we use a Fourier shrinkage to increase the contrast of image structure before analyzing the patches. We have performed ample simulation using synthetic and real data to validate the performance of the proposed method. We used white Gaussian and processed noise in our simulations to support our claims. The proposed approach is fast and able to provide a more reliable estimation of image quality under different degradations compared to state-of-the-art NR-IQA approaches. For future work we plan to extend our method with learning based scheme such as support vector machine to make it more effective under general distortions. For example, learning and decision can be made based on extracted sparsity and orientation dominancy values of all patches.

References

1. Cao, G., Zhao, Y., Ni, R.: Edge-based blur metric for tamper detection. Journal of Information Hiding and Multimedia Signal Processing **1**(1), 20–27 (2010)
2. Chen, M., Bovik, A.: No-reference image blur assessment using multiscale gradient. In: Quality of Multimedia Experience. International Workshop on, pp. 70–74 (2009)
3. Cohen, E., Yitzhaky, Y.: No-reference assessment of blur and noise impacts on image quality. Signal, image and video processing **4**(3), 289–302 (2010)
4. Dabov, K., Foi, A., Egiazarian, K.: Video denoising by sparse 3D transform-domain collaborative filtering. In: Proc. 15th European Signal Processing Conference, vol. 1, p. 7 (2007)

5. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-D transform-domain collaborative filtering. Image Processing, IEEE Trans. on **16**(8), 2080–2095 (2007)
6. Dash, R., Sa, P., Majhi, B., et al.: Blur parameter identification using support vector machine. ACEEE International Journal on Control System and Instrumentation **3**(2) (2012)
7. Demmel, J., Kahan, W.: Accurate singular values of bidiagonal matrices. SIAM Journal of Sci. Stat. Comput **11**, 873–912 (1990)
8. Elad, M., Aharon, M.: Image denoising via learned dictionaries and sparse representation. In: Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, vol. 1, pp. 895–900 (2006)
9. Feichtenhofer, C., Fassold, H., Schallauer, P.: A perceptual image sharpness metric based on local edge gradient analysis. Signal Processing Letters, IEEE **20**(4), 379–382 (2013)
10. Feng, X., Milanfar, P.: Multiscale principal components analysis for image local orientation estimation. In: Signals, Systems and Computers, Conference Record of the Thirty-Sixth Asilomar Conference on, vol. 1, pp. 478–482 (2002)
11. Ferzli, R., Karam, L.: A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). Image Processing, IEEE Trans. on **18**(4), 717–728 (2009)
12. Gabarda, S., Cristóbal, G.: Blind image quality assessment through anisotropy. Journal of Opt. Soc. Am. A **24**(12), B42–B51 (2007)
13. Gu, K., Zhai, G., Yang, X., Zhang, W.: Using free energy principle for blind image quality assessment. IEEE Transactions on Multimedia **17**(1), 50–63 (2015)
14. Hassen, R., Wang, Z., Salama, M.: Image sharpness assessment based on local phase coherence. Image Processing, IEEE Trans. on **22**(7), 2798–2810 (2013)
15. Knaus, C., Zwicker, M.: Progressive image denoising. Image Processing, IEEE Trans. on **23**(7), 3114–3125 (2014)
16. Maggioni, M., Sanchez-Monge, E., Foi, A.: Joint removal of random and fixed-pattern noise through spatiotemporal video filtering. Image Processing, IEEE Trans. on **23**(10), 4282–4296 (2014)
17. Mittal, A., Moorthy, A., Bovik, A.: No-reference image quality assessment in the spatial domain. Image Processing, IEEE Trans. on **21**(12), 4695–4708 (2012)
18. Moorthy, A., Bovik, A.: A two-step framework for constructing blind image quality indices. Signal Processing Letters, IEEE **17**(5), 513–516 (2010)
19. Narvekar, N., Karam, L.: A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). Image Processing, IEEE Trans. on **20**(9), 2678–2683 (2011)
20. Ponomarenko, N., Jin, L., Ieremeiev, O., Lukin, V., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., et al.: Image database TID2013: Peculiarities, results and perspectives. Signal Processing: Image Communication **30**, 57–77 (2015)
21. Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3 edn. (2007)
22. Rakhshanfar, M., Amer, A.: No-reference image quality assessment for removal of processed and unprocessed noise. In: Image Processing (ICIP), IEEE Int. Conf. on (2015)
23. Ramani, S., Blu, T., Unser, M.: Monte-carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms. Image Processing, IEEE Trans. on **17**(9), 1540–1554 (2008)
24. Takeda, H., Farsiu, S., Milanfar, P.: Kernel regression for image processing and reconstruction. Image Processing, IEEE Trans. on **16**(2), 349–366 (2007)
25. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Computer Vision, Sixth International Conference on, pp. 839–846 (1998)
26. Vu, C., Phan, T., Chandler, D.: S3: A spectral and spatial measure of local perceived sharpness in natural images. Image Processing, IEEE Trans. on **21**(3), 934–945 (2012)
27. Wang, Z.: Applications of objective image quality assessment methods [applications corner]. Signal Processing Magazine, IEEE **28**(6), 137–142 (2011)
28. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. Image Processing, IEEE Trans. on **13**(4), 600–612 (2004)
29. Xue, W., Mou, X., Zhang, L., Bovik, A.C., Feng, X.: Blind image quality assessment using joint statistics of gradient magnitude and laplacian features. IEEE Transactions on Image Processing **23**(11), 4850–4862 (2014)
30. Zh., X., Milanfar, P.: Automatic parameter selection for denoising algorithms using a no-reference measure of image content. Image Processing, IEEE Trans. on **19**(12), 3116–3132 (2010)