Image Denoising With Edge-Preserving and Segmentation Based on Mask NHA

Fumitaka Hosotani, Yuya Inuzuka, Masaya Hasegawa, Shigeki Hirobayashi, *Member, IEEE*, and Tadanobu Misawa

***Abstract*— In this paper, we propose a zero-mean white Gaussian noise removal method using a high-resolution frequency analysis. It is difficult to separate an original image component from a noise component when using discrete Fourier transform or discrete cosine transform for analysis because sidelobes occur in the results. The 2D non-harmonic analysis (2D NHA) is a high-resolution frequency analysis technique that improves noise removal accuracy because of its sidelobe reduction feature. However, spectra generated by NHA are distorted, because of which the signal of the image is non-stationary. In this paper, we analyze each region with a homogeneous texture in the noisy image. Non-uniform regions that occur due to segmentation are analyzed by an extended 2D NHA method called Mask NHA. We conducted an experiment using a simulation image, and found that Mask NHA denoising attains a higher peak signal-to-noise ratio (PSNR) value than the state-of-the-art methods if a suitable segmentation result can be obtained from the input image, even though parameter optimization was incomplete. This experimental result exhibits the upper limit on the value of PSNR in our Mask NHA denoising method. The performance of Mask NHA denoising is expected to approach the limit of PSNR by improving the segmentation method.**

***Index Terms*— Image denoising, image segmentation, image representation, edge detection, non-harmonic analysis (NHA).**

I. Introduction

THE NOISE reduction problem for images involves eliminating noise from an image to recover a cleaner one. In this problem, the image in question is often defined as the sum of the values of the relevant pixels of the clean (noise-free) image and those of the corresponding pixels in the noisy (random values) image. Denoising methods are used for noise reduction and have a wide range of applications. For example, Mairal et al. applied denoising methods to deblurring, inpainting, and demosaicing [1], [2].

Many cases can be obtained with regard to the distribution and strength of noise in an image. For example, Tian et al. closely analyzed noise at the circuit level [3]. Moreover, Hirakawa and Parks proposed a novel method to remove practical noise from real-world applications. In a closed approach, Xiao et al. proposed the restoration of images distorted by mixed Gaussian noise and impulse noise. In this paper, we propose a method for the removal of zero-mean white Gaussian noise and compare it with other prevalent techniques.

Numerous denoising methods have been developed in diverse fields of research. Katkovnik et al. in their review suggested a classification of denoising algorithms according to two features: local/non-local and pointwise/multipoint algorithms [6]. In local pointwise modeling, the size of the support is restricted by the distance between the estimation point and the observation point, and the estimator provides an estimate only for a single point. Local polynomial approximation (LPA) involves typical local pointwise models that use fitting polynomials [7]. Local multipoint estimates are calculated for all observation points used by the estimator and can constitute an image block. In practice, k-means singular value decomposition (K-SVD) [8] and Multi-scale K-SVD (MS K-SVD) [9] are commonly used algorithms for dictionary learning. The shape-adaptive discrete cosine transform (SA-DCT), which is applicable to non-uniform blocks, is used to analyze homogeneous regions [10]. Zhang et al. proposed a local pixel grouping principal component analysis (LPG-PCA) [11] that is classified into local and multipoint methods. While both these methods focus on the neighborhood of the observation point, a non-local (NL) LPG-PCA method by contrast samples points from the entire image. NL-means are further classed into non-local and pointwise methods, and both involve obtaining weights from the entire image [12]. Block matching and 3D filtering (BM3D) algorithms are known to be efficient and are classified into nonlocal and multipoint methods [13]. This method achieves high sparsity in the transform domain by grouping similar 2D image blocks into 3D data arrays. A combination of non-local and local methods has also been proposed in recent years. Yang et al. proposed such a method, which also incorporated dictionary learning and a non-local approach [14].

The multipoint methods reviewed by Katkovnik et al. are typically based on thresholding in the transform domain using orthogonal transforms. DFT and DCT are commonly used as orthogonal transforms [15]. Their accuracy depends on the size of the analysis window, and a signal cannot be analyzed if its period is longer than that of the analysis window.

Manuscript received December 22, 2014; revised October 7, 2015; accepted October 15, 2015. Date of publication October 26, 2015; date of current version November 11, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Peter Tay. *(Corresponding author: Shigeki Hirobayashi.)*

F. Hosotani, Y. Inuzuka, and M. Hasegawa are with the Department of Intellectual Information Engineering, Graduate School of Science and Engineering for Education, University of Toyama, Toyama 9308555, Japan (e-mail: fum0622@gmail.com; m1471104@ems.u-toyama.ac.jp; d1471002@ems.u-toyama.ac.jp).

S. Hirobayashi and T. Misawa are with the Department of Science and Engineering, University of Toyama, Toyama 9308555, Japan (e-mail: hirobays@eng.u-toyama.ac.jp; misawa@eng.u-toyama.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2015.2494461

Therefore, sidelobes are observed in the results of DFT or DCT analysis when the input signal in the analysis window does not have an integer value period.

To address the above-mentioned problem, Hirobayashi et al. designed a high-resolution frequency analysis method called non-harmonic analysis (NHA) [16]. NHA has been used in numerous applications and delivers satisfactory results [17]–[20]. It improves thresholding accuracy because it expresses frequency using a line spectrum that can suppress sidelobes. Using a large analysis window increases the frequency resolution of DFT and DCT but reduces their spatial resolution. NHA can obtain a frequency of satisfactory resolution with a relatively small window because it does not depend on the size of the analysis window.

NHA assumes a stationary signal in the analysis window. Therefore, non-stationarity in the analysis window produces a sidelobe during frequency analysis.

A ringing artifact occurs because of sidelobe deletion during thresholding. Chatterjee and Milanfar Chatterjee and Milanfar proposed a cluster-based denoising method that focuses on the structural similarity of an image. Regions with different textures separated by segmentation are all considered stationary signals. Thus, the occurrence of a ringing artifact can prevent us from securing stationary signals through segmentation.

The choice of threshold is an important problem in denoising. Our aim in this paper is to increase the denoising quality of an image through preprocessing. We will thus postpone a discussion of the choice of threshold until future research, and will only consider the possibility of improving image denoising using 2D NHA and a segmented image in this paper.

Our proposed denoising method consists of two steps. First, we expand the 2D NHA to a non-uniform shape to correspond to the segmented image in order to conduct a more accurate frequency analysis. Second, we consider an image segmentation method in order to separate regions with spatial stationarity due to noisy images.

The remainder of this paper is organized as follows: in Section 2, we consider an open problem in frequency analysis and present our high-resolution frequency analysis method using 2D NHA. Moreover, we propose an edge-preserving segmentation method that divides edges and homogeneous regions in an image. We then indicate a limitation of our approach through a denoising experiment that uses a simulation image. In Section 3, we test the performance of our proposed denoising method using natural images. Section 4 contains our conclusions.

II. Proposed Method

The analysis of images using DFT and DCT leads to sidelobes when the period of the signal in the analysis window is not an integer value. Hence, sidelobes and noise cannot be completely distinguished from the amplitude of DFT and DCT. Sidelobe reduction occurs often in thresholding denoising. A method for the analysis of sidelobe suppression is thus needed to solve this problem. Non-harmonic analysis, proposed by Hirobayashi et al., can distinguish the components of the original signal from those of the noise signal in a noisy image because it estimates frequency as a line spectrum. NHA is more effective than DCT in this context and thus is more suitable for thresholding. However, analysis using NHA leads to distorted results in case of non-stationary signals because like DFT and DCT, NHA assumes the stationarity of the signals of the analysis window. Further, sidelobes occur if an image containing edges is analyzed. To address this problem, we separate the homogeneous region in an image using image segmentation. However, every region contains a non-uniform shape as a result of segmentation because like DCT, 2D NHA uses a rectangular window. For the analysis of segmented images, 2D NHA needs to be adapted to nonrectangular windows. Our proposed method proceeds as follows:

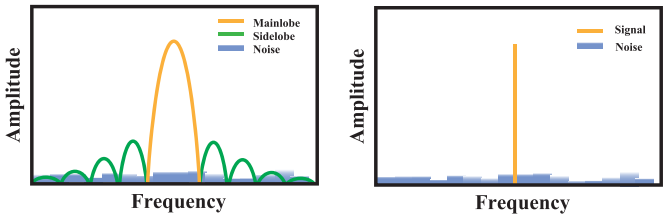


Fig. 1. The difference in frequency expression in a noisy environment between (a) traditional frequency analysis, and (b) high-frequency resolution analysis.

1) Extract edges from a noisy image.
2) Define the edge regions through binary dilation of edges.
3) Remove the edge regions from the noisy image.
4) Apply texture segmentation to the image with the edges removed.
    a) For each layer obtained from segmentation, do the following:
        i) Define the masking matrices from the segmentation results. The focused layer is assigned the value 1 and other layers are assigned 0.
        ii) Extract patches from the layer of interest.
        iii) Calculate spectra for each patch using mask NHA.
        iv) Apply thresholding to obtain spectra for each patch.
        v) Obtain the restored patches.
        vi) Obtain the restored layer by a synthesis of the patches.
5) Obtain restored image by combining all layers.

A. 2D NHA

The 2D DFT in the spatial frequency domain is expressed as follows:

$$X(k,l) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} I(n_1, n_2) e^{-j2\pi\left(\frac{kn_1}{N} + \frac{ln_2}{M}\right)}. \quad (1)$$

In this equation, N and M represent the size of the data, and I is the image signal. The short Fourier transform assumes a completely periodic signal in an $N \times M$ analysis window.

The frequency resolution of DFT is limited by the integer period of $\frac{k}{N}$. If the period of the signal in the analysis window has a non-integer value, sidelobes are observed in the results. In an analysis of DFT, the frequency resolution is chosen according to the change in analysis window size. However, NHA estimates the Fourier coefficients by using sinusoidal

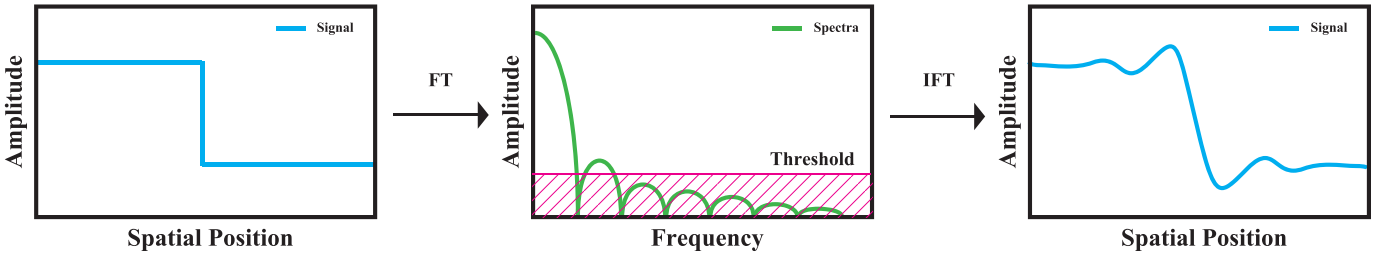


Fig. 2. Ringing artifact due to thresholding.

wave fitting [16]. The 2D sinusoidal model used by NHA is as follows:

$$I(n_1, n_2) = \hat{A}\cos\left(2\pi \frac{\hat{f}_x}{f_{x_s}} n_1 + \frac{\hat{f}_y}{f_{y_s}} n_2 + \hat{\phi}\right). \tag{2}$$

Here, $\hat{A}$ is the amplitude of the sinusoidal model, $\hat{f}_x$ and $\hat{f}_y$ represent spatial frequencies, and $\hat{\phi}$ denotes the phase parameter of the model. f_{x_s} and f_{y_s} are sampling frequencies that are defined as $f_{x_s} = \frac{1}{\partial x}$ and $f_{y_s} = \frac{1}{\partial y}$, respectively. Furthermore, the two axes of the image domain correspond to x and y. The 2D NHA estimates frequency by minimizing the mean squared error (MSE) between the target signal and the model signal, and is determined by the method of steepest descent. The frequency parameters are calculated as follows:

$$F(\hat{A}, \hat{f}_x, \hat{f}_y, \hat{\phi}) = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \{I(n_1, n_2) - \hat{I}(n_1, n_2)\}^2. \tag{3}$$

The above expression is solved as an optimization problem using the coefficient of 2D DCT as the initial value. In frequency analysis using DCT in general, the frequency resolution depends on the window size. The original signal can be distorted by the removal of the sidelobe through thresholding. This is the obstacle to denoising. NHA can suppress sidelobes using high-frequency resolution. Fig. 1 shows the difference in frequency expressions between a traditional analysis and a high-frequency analysis. As can be seen, sidelobes occur in the results of traditional analysis (left). High-frequency resolution analysis can distinguish the original signal from the noise signal, as shown in Fig. 1 (right).

B. Mask NHA

Image denoising using domain transformation often employs a block unit called a "patch." The smoothing effectiveness of denoising increases with the patch size. However, the likelihood of the edge mixing with the patches increases when using large patches. The edge is transformed into the sinc function by DFT. When part of the sinc function in the frequency domain is removed by thresholding, the restored image includes a ringing artifact. The edge details are hence lost by thresholding because sidelobes always occur as a result of the analysis of edges. The detailed deletion in the thresholding process is shown in Fig. 2. Edges in the image represent non-stationary signals. Hence, denoising needs to be applied to stationary signals to prevent detail deletion.

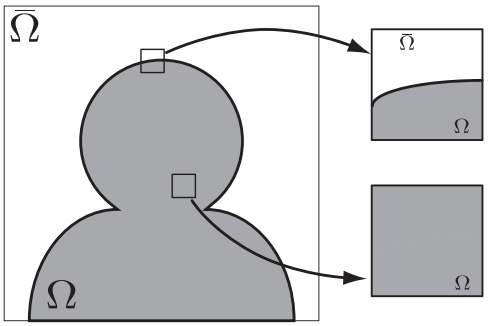


Fig. 3. Non-uniform analysis window occurred for the block contained two uniform region.

If the analysis window contains regions of different textures adjacent to each other, the corresponding signal in the analysis window is non-stationary. To minimize the influence of signal non-stationarity, it is necessary to exclude edges and different textures from the analysis window. Hence, the input signal in the analysis window needs to be masked. The masking area is obtained from image segmentation results.

In segmentation, the noisy image is classified into two regions, the edge region and the texture region. Moreover, texture region is classified into multiple regions using spatial similarity. Note that similar textures are classified into the same cluster and can be considered nearly stationary signals. Therefore, patches extracted from arbitrary regions can be considered approximate stationary signals.

However, the patch located on the segment boundary can contain different regions. Following segmentation, image regions contain non-uniform shapes, and thus 2DNHA needs to be adapted to a window of flexible shape. Non-uniform regions occur near the boundary of the region, as shown in Fig. 3. The target region is represented by Ω, and the remaining region is represented by $\bar{\Omega}$.

Therefore, we need to improve part of the 2D NHA algorithm. In order to differentiate between Ω and $\bar{\Omega}$, our proposed method employs a weighting factor of $w(n_1, n_2)$ constructed using binary information. This weighting factor is the masking matrix. When the outer region is represented in an image I, $w(n_1, n_2)$ is 0. When the target region is represented in an image I, $w(n_1, n_2)$ is 1. Moreover, in order to calculate a cost function that excludes $\bar{\Omega}$, we improve the cost function given in Eq. (3) as follows:

$$F(\hat{A}, \hat{f}_x, \hat{f}_y, \hat{\phi}) = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} w(n_1, n_2) \times \{I(n_1, n_2) - \hat{I}(n_1, n_2)\}^2. \tag{4}$$

(a) (b)

Fig. 4. Segmentation result by Edge-preserving segmentation, on the lenna image. (a) Lenna image, (b) segmentation result of Lenna image.

Using this cost function, we can extract the spectrum from Ω excluding $\bar{\Omega}$. Furthermore, Ω must be given a temporary value when initial values are assigned by the fast Fourier transform (FFT). The spectral structure must retain its shape in order to enable convergence to the global minimum in our proposed method. The algorithm of our proposed method is as follows. The median value or mean value of the known region $\bar{\Omega}$ is first applied to the missing region Ω. Initial values $\hat{A}$, f_x, f_y, and $\hat{\phi}$ are then assigned using FFT. By considering Eq. (4) as the cost function, this nonlinear problem is converted into a minimization problem, and $\hat{f}_x$, $\hat{f}_y$, and $\hat{\phi}$ are determined by repeatedly using the method of steepest descent.

C. Edge-Preserving Segmentation

An image needs to be divided into clusters without edges in order to prevent ringing artifacts. However, the boundary definition of each region with uniform texture is difficult to provide using existing segmentation methods. Thus, it is difficult to determine the boundary of a region within a noisy image. In particular, the segmentation of a noisy image often causes errors. Therefore, the segmentation of noisy images contains fuzziness in the choice of the region boundary. We believe that these fuzzy boundaries should be defined as edge regions. Specifically, the noisy image is separated into homogeneous textures and edge regions. Therefore, in our method, we first extract the edges from a noisy image and then define the edge regions. Boundary distortion due to segmentation is reduced by defining the texture boundaries. Fig. 4 shows the segmentation results of edge-preserving segmentation on the "Lenna" image, where the white line indicates the boundary of each region.

D. Edge Detection

We use Canny edge detection [22], which is commonly used to identify edge position. It uses a smoothing process prior to edge detection. The smoothing method often uses a Gaussian filter, but this distorts the edge position if it is used in Canny edge detection. An edge-preserving smoothing method is thus needed. Liu et al. improved the accuracy of Canny edge detection by using bilateral filtering [23], [24]. A bilateral filtering is an edge-preserving filter that uses weights defined according to the difference of pixel luminance between the pixel of interest and its neighboring pixels. The bilateral filter can be expressed as follows:

$$\boldsymbol{F}(\boldsymbol{x}) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{I}(\boldsymbol{\xi}) w(\boldsymbol{\xi}, \boldsymbol{x}) d\boldsymbol{\xi}}{\int_{-\infty}^{\infty} w(\boldsymbol{\xi}, \boldsymbol{x}) d\boldsymbol{\xi}} \tag{5}$$

where

$$w(\boldsymbol{\xi}, \mathbf{x}) = \exp\left(\frac{-(\boldsymbol{\xi} - \boldsymbol{x})^2}{2{\sigma_d}^2}\right) \exp\left(\frac{-(\boldsymbol{I}(\boldsymbol{\xi}) - \boldsymbol{I}(\boldsymbol{x}))^2}{2{\sigma_r}^2}\right). \tag{6}$$

$\boldsymbol{I}$ and $\boldsymbol{F}$ above denote the input image and the output image, respectively, $\boldsymbol{x}$ and $\boldsymbol{\xi}$ represent spatial variables, and σ_d and σ_r, represented by a standard deviation, characterize the filter intensity and edge sensibility, respectively. This Canny algorithm in our edge detection remains unchanged in the other step except for its bilateral filtering stage.

E. Segmentation

Two well-known strategies for image segmentation are region growing and region merging. Region growing is a method to iteratively expand the initial region, and region merging combines the over-segmented results. These methods cannot be used without prior knowledge of the image in question. We thus use the mean shift-based image segmentation proposed by Comaniciu and Meer [25]. The mean shift algorithm searches for local maximum density points in a feature space [26]. Christoudias et al. proposed a method to segment images by applying mean shift to the image feature space. Moreover, mean shift does not require information regarding the number of segments.

Mean shift is used in the kernel density function as follows:

$$\hat{f}_{h,K}(\boldsymbol{x}) = \frac{c_{k,d}}{nh^d} \sum_{n=1}^{n} k\left(\left\|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{h}\right\|^2\right) \tag{7}$$

$\boldsymbol{x}$ in the above is a variable vector, $\boldsymbol{x}_i$ represents sample vectors, k is the kernel function, h the band parameter, and $c_{k,d}$ is the normalization constant. The kernel density function constructs a curved surface and its local maximum is derived through iteration. The derived function for $k(x)$ is defined as follows:

$$g(x) = -k'(x).$$

Therefore, the recursion for convergence to the local maximum can be written as:

$$\boldsymbol{y}_{j+1} = \frac{\sum_{i=1}^{n} \boldsymbol{x}_i g\left(\left\|\frac{\boldsymbol{y}_j - \boldsymbol{x}_i}{h}\right\|^2\right)}{g\left(\left\|\frac{\boldsymbol{y}_j - \boldsymbol{x}_i}{h}\right\|^2\right)}. \quad j = 1, 2, \ldots \tag{8}$$

Image segmentation using mean shift is obtained through grouping based on the distance between each convergence point $\boldsymbol{y}_{i,1} = \boldsymbol{x}_i$ and its neighboring points. Image segmentation through the mean shift algorithm is available in the Edge Detection and Image Segmentation (EDISON) system developed by the Robust Image Understanding Laboratory (RIUL) at Rutgers University [27]. The summary of image segmentation through EDISON is as follows:

1) Mean shift segmentation
   a) Obtain N convergence points $\{z_i\}_{i=1,\ldots,n}$ by mean shift.
   b) Group each point z_i with its neighboring points comparing z_i^s, z_i^r and h_s, h_r, respectively.
   c) Obtain P clusters $\{\boldsymbol{C}_p\}$ through weighted averaging. The weight is obtained from the edge intensity.
   d) Assign label $L_i = \{p | z_i \in z_p\}$ to each input vector z_i to use indices of clusters p.

In the above, z_i and $\{z_i\}_{i=1,\ldots,n}$ denote the vectorized input image and the convergent points obtained by the mean shift algorithm, respectively. The subscripts s and r denote spatial and range components of a vector, respectively. and in the above denote spatial and range bandwidths, respectively.

F. Advantage of the Proposed Method

Our method performs satisfactorily if image segmentation is ideally carried out. The test images used in our experiment consisted of textures of two patterns, each of which separates vertically at the center of the images. Each pattern consists of three sinusoidal waves of different frequencies. Fig. 5*a* shows the tested image.

For our experiment, we assumed that texture segmentation was ideally performed. Therefore, masking matrices were properly formed for each pattern.

Moreover, the universal threshold proposed by Donoho and Johnstone was unavailable for the results of Mask NHA. Thus, we did not use this thresholding and instead employed an ideal choice of threshold in the experiment.

The ideal thresholding is implemented using spectral extraction that minimizes the difference between the thresholding result of the restored image and that of the clean image. Specifically, the results of this experiment reveal the limitation of our method. We plan to consider the choice of threshold in Mask NHA in future research.

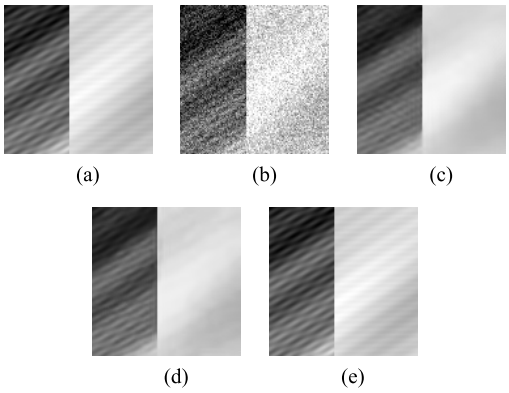


Fig. 5. Denoising results of the proposed method and two state-of-the-art denoising methods. (a) Clean Image. (b) Noisy $\sigma = 30$. (c) SA-DCT. (d) BM3D. (e) Proposed.

TABLE I
DENOISING PERFORMANCE OF EACH METHOD

σ	SA-DCT	BM3D	Proposed
$\sigma = 5$	45.24	46.04	**48.18**
$\sigma = 10$	40.73	41.32	**45.20**
$\sigma = 15$	37.56	38.45	**43.32**
$\sigma = 20$	35.20	36.15	**41.32**
$\sigma = 25$	33.31	33.99	**40.76**
$\sigma = 30$	31.75	32.56	**40.57**

We tested the performance of our method against that of SA-DCT and BM3D, two state-of-the-art methods. Figures 5*b*-5*e* show the denoising results for SA-DCT, BM3D, and our proposed method. Table *I* lists the quantitative evaluation results according to peak signal-to- noise ratio (PSNR) as a qualitative measure. We can see that the proposed method is superior to the two state-of-the-art methods at every noise intensity value.

III. EXPERIMENTAL RESULTS

In this section, we present and discuss the experimental results obtained from natural image denoising using our proposed method. We selected target images with varying features, such as a landscape and a portrait. The target images were selected from a traditional image database and the Laboratory for Image and Video Engineering (LIVE) image database at the University of Texas at Austin. The target images were assumed to be noise-free, and the noisy images consisted of noise-free images and those with additive white Gaussian noise (AWGN).

The proposed method needs to determine the values of several parameters according to the target images. Its performance is thus dependent on the values of these parameters. The values of the parameters should be chosen from the input images because the values of optimal parameters vary according to the nature of the input signal.

However, we used constant parameters throughout this experiment except in the spectral thresholding process. Each parameter was calibrated for the "Cameraman" image.

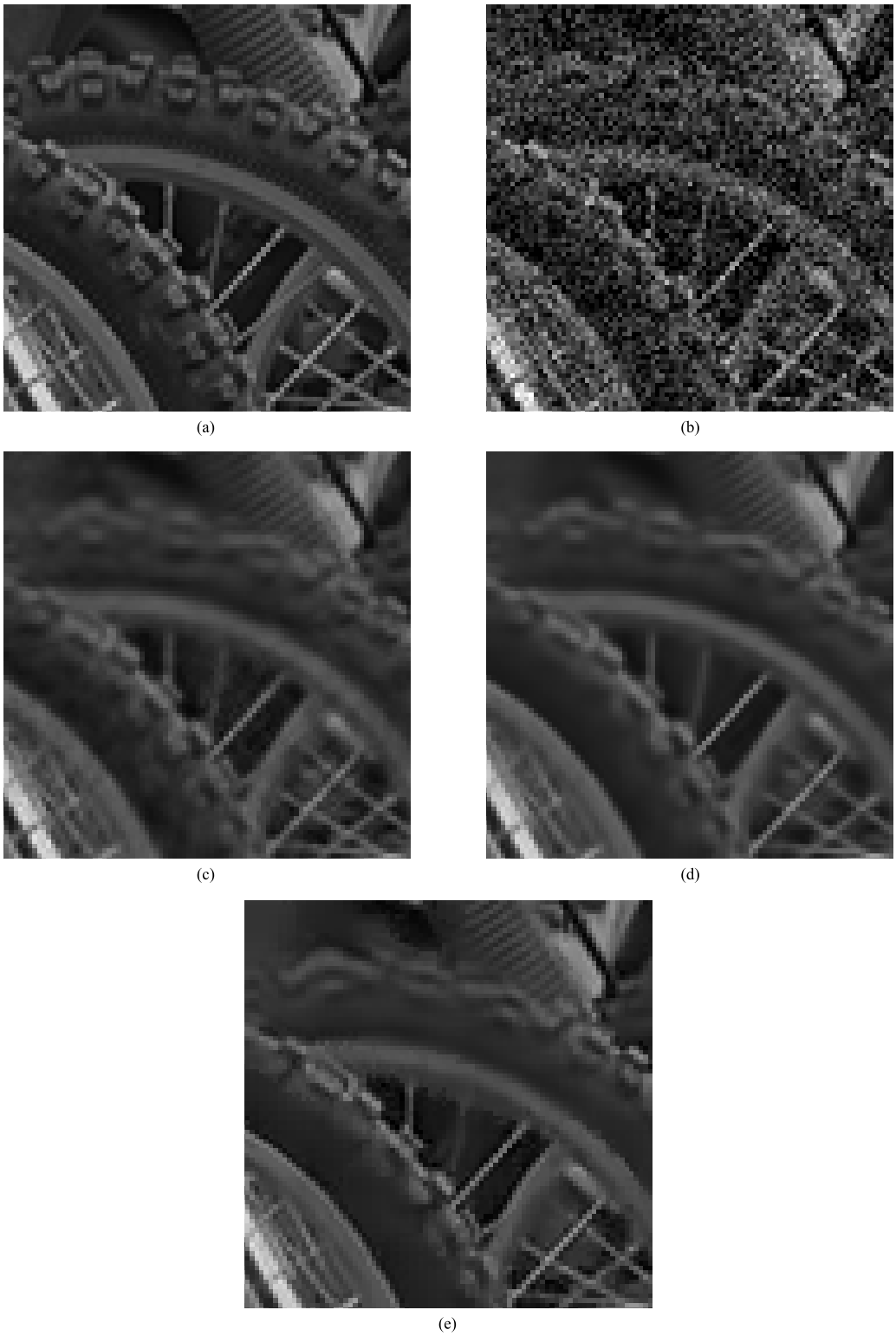


Fig. 6. Restored images of *Bikes* using each method. (a) Original Image. (b) Noisy $\sigma = 25$. (c) SA-DCT. (d) BM3D. (e) Proposed Method.

The smoothing process using a bilateral filter described in Section is controlled by two standard deviations and the window size. We chose values of 5 and 0.08 for σ_d and σ_r, respectively, and set the window size to 5×5. These constants were empirically determined to enable us to extract objects in the "Cameraman" image.

We used two threshold values in Canny edge detection. Each threshold value was automatically calculated by using the histogram of a differential image [28].

We used the EDISON system implemented in C++ for mean shift segmentation. Each parameter was set according to the initial settings of the EDISON system; therefore, no other

TABLE II
SUMMARY OF ALL DENOISING PSNR RESULTS

σ	Cameraman			Bikes			Ocean			Statue		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	40.83	**40.97**	38.50	40.36	**40.53**	39.35	40.78	**40.87**	40.81	40.89	**40.98**	40.14
$\sigma = 10$	**36.25**	36.15	36.24	35.04	35.20	**35.56**	35.96	36.16	**36.66**	36.58	36.81	**36.84**
$\sigma = 15$	33.37	33.29	**33.88**	31.94	32.14	**32.88**	33.42	33.81	**34.12**	34.20	34.52	**34.70**
$\sigma = 20$	31.62	31.50	**31.98**	29.85	30.09	**30.91**	31.76	32.32	**32.40**	32.55	32.91	**33.05**
$\sigma = 25$	30.41	30.26	**30.53**	28.28	28.58	**29.45**	30.63	**31.29**	31.12	31.33	31.59	**31.65**
$\sigma = 30$	**29.54**	29.37	29.47	27.07	27.41	**28.24**	29.79	**30.47**	30.18	30.34	30.51	**30.53**
σ	Womanhat			Barbara			Buildings			Buildings2		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	40.21	**40.35**	40.21	40.56	**40.76**	40.63	**39.65**	39.53	37.85	**39.42**	39.29	36.52
$\sigma = 10$	35.89	36.16	**36.41**	36.47	**37.09**	36.93	34.66	**34.67**	34.61	**33.88**	33.49	33.58
$\sigma = 15$	33.73	**34.22**	34.18	34.05	**34.98**	34.80	31.93	31.98	**32.28**	30.70	30.12	**31.02**
$\sigma = 20$	32.34	**32.99**	32.72	32.33	**33.47**	33.22	30.05	30.14	**30.56**	28.52	27.84	**29.05**
$\sigma = 25$	31.37	**32.12**	31.64	30.96	**32.31**	31.97	28.60	28.78	**29.23**	26.85	26.12	**27.49**
$\sigma = 30$	30.65	**31.43**	30.86	29.86	**31.29**	30.90	27.47	27.71	**28.09**	25.55	24.85	**26.25**
σ	Coinsfountain			Lenna			Lighthouse			Stream		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	**40.25**	40.08	39.63	**40.34**	40.29	40.30	**40.23**	40.22	39.41	**39.21**	38.97	35.15
$\sigma = 10$	34.99	34.68	**35.57**	37.32	**37.44**	37.34	35.31	35.32	**35.86**	**33.25**	32.74	32.75
$\sigma = 15$	32.12	31.84	**32.86**	35.47	**35.79**	35.52	32.79	33.03	**33.27**	30.00	29.48	**30.30**
$\sigma = 10$	30.26	30.03	**31.00**	34.10	**34.61**	34.26	31.30	**31.73**	31.72	27.93	27.45	**28.43**
$\sigma = 25$	29.05	28.80	**29.64**	33.04	**33.64**	33.21	30.28	**30.85**	30.68	26.44	26.03	**26.98**
$\sigma = 30$	28.07	27.79	**28.62**	32.07	**32.81**	32.40	29.48	**30.13**	29.93	25.37	25.01	**25.93**

TABLE III
SUMMARY OF ALL DENOISING MAE RESULTS

σ	Cameraman			Bikes			Ocean			Statue		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	2.26	**2.23**	2.46	2.94	**2.77**	2.85	2.29	**2.24**	2.25	2.29	**2.21**	2.27
$\sigma = 10$	3.34	**3.38**	3.43	4.86	4.46	**4.46**	3.49	3.40	**3.34**	3.37	**3.23**	3.27
$\sigma = 15$	4.34	4.26	**4.23**	6.35	5.81	**5.79**	4.35	4.24	**4.20**	4.19	**4.04**	4.13
$\sigma = 20$	**5.00**	5.12	5.15	7.61	**6.97**	7.00	5.05	**4.92**	4.95	4.86	**4.79**	5.00
$\sigma = 25$	**5.59**	5.83	6.13	8.64	**7.98**	8.06	5.63	**5.48**	5.65	**5.40**	5.48	5.87
$\sigma = 30$	**6.06**	6.55	6.98	9.68	**9.04**	9.09	6.12	**5.98**	6.24	**5.92**	6.23	6.70
σ	Womanhat			Barbara			Buildings			Buildings2		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	2.36	**2.28**	2.32	2.61	2.39	**2.38**	3.01	**2.89**	31.2	3.45	**3.38**	3.81
$\sigma = 10$	3.45	**3.31**	3.34	3.94	**3.42**	3.54	4.86	**4.53**	4.72	6.00	**5.83**	5.91
$\sigma = 15$	4.18	**4.00**	4.12	4.89	**4.15**	4.41	6.27	**5.77**	6.01	7.95	7.79	**7.68**
$\sigma = 20$	4.78	**4.57**	4.75	5.65	**4.82**	5.20	7.42	**6.83**	7.14	9.55	9.49	**9.21**
$\sigma = 25$	5.25	**5.05**	5.37	6.34	**5.43**	5.99	8.46	**7.80**	8.15	10.95	11.05	**10.63**
$\sigma = 30$	5.66	**5.45**	5.90	6.99	**6.06**	6.73	9.45	**8.73**	9.20	12.17	12.48	**11.96**
σ	Coinsfountain			Lenna			Lighthouse			Stream		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	2.78	**2.71**	2.78	2.32	2.27	**2.26**	2.55	**2.49**	2.57	3.47	**3.47**	4.07
$\sigma = 10$	4.39	4.34	**4.25**	3.12	**3.03**	3.12	3.90	3.75	**3.69**	**6.00**	6.01	6.02
$\sigma = 15$	5.60	5.62	**5.45**	3.73	**3.63**	3.76	4.85	4.61	**4.62**	7.94	7.99	**7.61**
$\sigma = 20$	6.59	6.70	**6.51**	4.21	**4.09**	4.33	5.55	**5.25**	5.39	9.48	9.56	**9.06**
$\sigma = 25$	**7.37**	7.67	7.47	4.67	**4.55**	4.87	6.14	**5.80**	6.03	10.73	10.85	**10.32**
$\sigma = 30$	**8.12**	8.57	8.40	5.10	**4.99**	5.34	6.63	**6.28**	6.57	11.80	11.95	**11.39**

parameters required setting. However, the minimum region size used was 8×8 and the edge strength threshold was 0.1. These values were also determined empirically. The same as section II-F, the threshold of the spectral extraction was determined using the original image.

Table *II*, Table *III*, and Table *IV* show the denoising results. Fig. 6 shows the denoised images obtained through the proposed method as well as the two aforementioned state-of-the-art methods used for performance comparison.

In Fig. 6, our method appears to restore the detail of the tire more clearly than other methods. Table *II* shows that our method performs better in terms of the PSNR comparison than both SA-DCT and BM3D for eight of the 12 images used in our experiment. However, the comparison results based on structural similarity (SSIM) index values listed in table IV yield a different perspective. SSIM may be much closer to the MSE than some might claim [29]. In this paper, we used SSIM as an objective evaluation indicator and found that our proposed method outperforms the other methods for only three images. With regard to noise intensity, the proposed method is effectively between $\sigma = 10$ and $\sigma = 25$. However, BM3D is better than our method when the lowest intensity or highest intensity. Table III shows that our method yields better mean absolute error (MAE) results than the other methods for three images. With regard to noise intensity, the proposed method is robust between $\sigma = 10$ and $\sigma = 20$. These results show that our method can satisfactorily reproduce the details of images.

TABLE IV
SUMMARY OF ALL DENOISING SSIM RESULTS

σ	Cameraman			Bikes			Ocean			Statue		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	0.970	**0.970**	0.965	0.985	**0.985**	0.984	0.970	0.970	**0.971**	0.966	0.966	**0.966**
$\sigma = 10$	0.943	0.942	**0.946**	0.954	0.954	**0.958**	0.924	0.927	**0.930**	0.929	0.932	**0.932**
$\sigma = 15$	0.915	0.913	**0.917**	0.915	0.917	**0.923**	0.876	**0.886**	0.883	0.896	**0.903**	0.899
$\sigma = 20$	**0.887**	0.886	0.881	0.871	0.877	**0.885**	0.832	**0.852**	0.838	0.868	**0.878**	0.864
$\sigma = 25$	**0.865**	0.862	0.838	0.828	0.838	**0.847**	0.791	**0.818**	0.797	0.843	**0.856**	0.823
$\sigma = 30$	**0.846**	0.841	0.801	0.788	0.803	**0.809**	0.760	**0.793**	0.763	0.821	**0.833**	0.784
σ	Womanhat			Barbara			Buildings			Buildings2		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	0.955	0.956	**0.958**	**0.973**	0.972	0.973	**0.978**	0.976	0.977	**0.990**	0.989	0.986
$\sigma = 10$	0.900	0.905	**0.907**	0.951	**0.954**	0.950	0.944	**0.944**	0.944	**0.965**	0.959	0.965
$\sigma = 15$	0.854	**0.867**	0.861	0.932	**0.939**	0.933	0.912	**0.916**	0.913	0.931	0.918	**0.932**
$\sigma = 20$	0.819	**0.840**	0.824	0.911	**0.924**	0.915	0.883	**0.891**	0.885	0.890	0.872	**0.894**
$\sigma = 25$	0.794	**0.819**	0.794	0.889	**0.908**	0.895	0.855	**0.867**	0.858	0.847	0.824	**0.852**
$\sigma = 30$	0.773	**0.800**	0.772	0.866	**0.892**	0.873	0.829	**0.846**	0.832	0.807	0.780	**0.813**
σ	Coinsfountain			Lenna			Lighthouse			Stream		
	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed	SA-DCT	BM3D	Proposed
$\sigma = 5$	**0.977**	0.975	0.976	**0.953**	0.952	0.952	0.966	**0.966**	0.964	**0.984**	0.982	0.976
$\sigma = 10$	0.935	0.927	**0.939**	0.926	**0.927**	0.923	0.917	0.917	**0.926**	0.935	0.925	**0.941**
$\sigma = 15$	0.886	0.876	**0.892**	0.906	**0.910**	0.901	0.867	0.873	**0.873**	0.863	0.848	**0.876**
$\sigma = 20$	0.841	0.831	**0.844**	0.888	**0.896**	0.883	0.827	**0.841**	0.830	0.788	0.776	**0.802**
$\sigma = 25$	**0.804**	0.794	0.803	0.872	**0.883**	0.866	0.800	**0.819**	0.801	0.725	0.718	**0.732**
$\sigma = 30$	**0.771**	0.761	0.769	0.858	**0.871**	0.853	0.779	**0.800**	0.781	0.669	0.670	**0.677**

Our method seems to be able to obtain better results when the image in question contains an explicit edge.

In this paper, we aimed for accurate image analysis using homogeneous texture segmentation. The performance of our method depends on preprocessing, each function of which is controlled by several parameters.

This experiment showed that our method can obtain denoising results comparable with those of state-of-the-art methods in spite of constraints. In this paper, edge-preserving segmentation was provisionally implemented. However, the implementation described above is incomplete. Image segmentation was used to secure the stationarity of the frequency, but the input feature vectors used for segmentation represented features of the spatial domain. We think that frequency domain features should be used in image segmentation instead. For instance, we can use DFT or DCT coefficients to search for similar frequency patches as arbitrary position patches. These coefficients are thus a suitable frequency feature for our purposes.

IV. CONCLUSION

In this paper, we proposed a novel noise-reduction method for images using image segmentation. Our proposed denoising method uses high-resolution analysis, called 2D NHA. However, 2D NHA is unable to accurately analyze non-stationary signals because part of the spectrum required to restore the image is lost while determining a threshold for the non-stationary patch. We thus proposed an edge-preserving segmentation method implemented by a compounded method consisting of Canny edge detection and a mean shift algorithm.

However, non-uniform patch analysis needs to be handled in order to process the segmented image. We thus adapted the 2D NHA method to non-uniform patches. In a pretest, our proposed method was compared with state-of-the-art methods for an artificial image. The result showed that our method can obtain results comparable with those of the state-of-the-art methods by using image segmentation.

Finally, we compared the denoising results of the proposed method with the state-of-the-art methods using natural images. The proposed method outperformed the other methods with regard to PSNR comparison, but was inferior to them with respect to SSIM values. In addition, a comparison of the MAE showed that pixel differences yield a similar result to the PSNR value. This indicates that denoising can be facilitated by image segmentation and edge preservation.

The main limitation of our work is that the performance of the proposed method is dependent on the parameters in the partial processes. In future research, we plan to replace edge-preserving segmentation with a superior method. For instance, the edge regions and homogeneous regions should be defined using frequency domain parameters. Further, we need to consider a method to determine the threshold values. Thus, there is room for improvement in subsequent research.

REFERENCES

[1] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[2] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[3] H. Tian, B. Fowler, and A. El Gamal, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, no. 1, pp. 92–101, Jan. 2001.

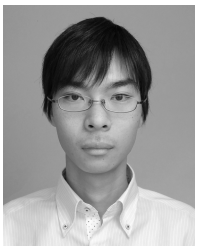
[4] K. Hirakawa and T. W. Parks, "Image denoising using total least squares," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2730–2742, Sep. 2006.

[5] Y. Xiao, T. Zeng, J. Yu, and M. K. Ng, "Restoration of images corrupted by mixed Gaussian-impulse noise via l_1–l_0 minimization," *Pattern Recognit.*, vol. 44, no. 8, pp. 1708–1720, 2011.

[6] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *Int. J. Comput. Vis.*, vol. 86, no. 1, pp. 1–32, 2010.

[7] C. Loader, *Local Regression and Likelihood*. New York, NY, USA: Springer-Verlag, 1999, vol. 47.

[8] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
[9] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representations for image and video restoration," *Multiscale Model. Simul.*, vol. 7, no. 1, pp. 214–241, 2008.
[10] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1395–1411, May 2007.
[11] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.*, vol. 43, no. 4, pp. 1531–1549, 2010.
[12] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
[13] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[14] S. Yang, L. Zhao, M. Wang, Y. Zhang, and L. Jiao, "Dictionary learning and similarity regularization based image noise reduction," *J. Vis. Commun. Image Represent.*, vol. 24, no. 2, pp. 181–186, 2013.
[15] O. G. Guleryuz, "Weighted averaging for denoising with overcomplete dictionaries," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 3020–3034, Dec. 2007.
[16] M. Hasegawa, T. Kako, S. H. Hirobayashi, T. Misawa, T. Yoshizawa, and Y. Inazumi, "Image inpainting on the basis of spectral structure from 2-D nonharmonic analysis," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3008–3017, Aug. 2013.
[17] T. Yoshizawa, S. Hirobayashi, and T. Misawa, "Noise reduction for periodic signals using high-resolution frequency analysis," *EURASIP J. Audio, Speech, Music Process.*, vol. 5, no. 5, pp. 1–19, 2011.
[18] T. Ichinose, S. Hirobayashi, T. Misawa, and T. Yoshizawa, "Forecast of stock market based on nonharmonic analysis used on NASDAQ since 1985," *Appl. Financial Econ.*, vol. 22, no. 3, pp. 197–208, 2012.
[19] X. Cao *et al.*, "Non-harmonic analysis applied to optical coherence tomography imaging," *Jpn. J. Appl. Phys.*, vol. 51, no. 2R, p. 022503, 2012.
[20] T. Ueda, K. Fujii, S. Hirobayashi, T. Yoshizawa, and T. Misawa, "Motion analysis using 3D high-resolution frequency analysis," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 2946–2959, Aug. 2013.
[21] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1438–1451, Jul. 2009.
[22] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.
[23] H. Liu, N. Klomp, and I. Heynderickx, "A perceptually relevant approach to ringing region detection," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1414–1426, Jun. 2010.
[24] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
[25] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
[26] K. Fukunaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 32–40, Jan. 1975.
[27] C. M. Christoudias, B. Georgescu, and P. Meer, "Synergism in low level vision," in *Proc. 16th Int. Conf. Pattern Recognit.*, vol. 4. 2002, pp. 150–155.
[28] H.-Q. Li, Q.-C. Yu, and M. Fang, "Application of otsu thresholding method on canny operator," *Comput. Eng. Design*, vol. 9, p. 046, 2008.
[29] R. Dosselmann and X. D. Yang, "A comprehensive assessment of the structural similarity index," *Signal Image Video Process.*, vol. 5, no. 1, pp. 81–91, 2011.



Fumitaka Hosotani received the B.E. degree in intellectual information engineering from the University of Toyama, Japan, in 2013, and the M.E. degree from the Department of Intellectual Information Engineering, Graduate School of Science and Engineering for Education, University of Toyama, in 2015. His research interests are in image processing (denoising and segmentation) and frequency analysis.

Yuya Inuzuka received the B.E. degree in intellectual information engineering from the University of Toyama, Toyama, Japan, in 2014, where he is currently pursuing the degree with the Department of Intellectual Information Engineering, Graduate School of Science and Engineering for Education. His research interests include digital signal processing and image inpainting.

Masaya Hasegawa received the M.E. degree with the Department of Intellectual Information Engineering, Graduate School of Science and Engineering for Education, University of Toyama, Toyama, Japan, in 2014. He is currently a Doctoral Program Student with the Graduate School of Science and Engineering for Education, University of Toyama. His research interests include digital signal processing and high accuracy analysis to image.

Shigeki Hirobayashi (M'03) received the D. Eng. degree. He is currently a Professor with the University of Toyama. He has been engaged in research on acoustic signal processing, acoustic field control, and image processing. Furthermore, he is a member of the Institute of Electronics Information and Communication Engineers, the Institute of Electrical Engineers of Japan, the Acoustic Society of Japan, and the Acoustic Society of America.

Tadanobu Misawa received the B.E., M.E., and Ph.D. degrees from Kanazawa University, in 1999, 2001, and 2004, respectively. He was a Doctoral Fellow with the Kanazawa Institute of Technology in 2004. He was an Assistant Professor with the School of Management, Tokyo University of Science, in 2005. He is currently an Associate Professor with the Graduate School of Science and Engineering, University of Toyama. His interests are artificial intelligence and brain–computer interface.