Image Restoration with Union of Directional Orthonormal DWTs

Shogo MURAMATSU, Natsuki AIZAWA and Masahiro YUKAWA
Dept. of Electrical and Electronic Eng., Niigata University
8050 2-no-cho, Ikarashi, Nishi-ku, 950-2181, Niigata, Japan
E-mail: {shogo,yukawa}@eng.niigata-u.ac.jp, natsu@telecom0.eng.niigata-u.ac.jp

***Abstract*—This work proposes to apply directional lapped orthogonal transforms to image restoration. A DirLOT is an orthonormal transform of which basis is allowed to be anisotropic with the symmetric, real-valued and compact-support property. In this work, DirLOTs are used to generate symmetric orthonormal discrete wavelet transforms and then a redundant dictionary as a union of unitary transforms. The multiple directional property is suitable for representing natural images which contain diagonal edges and textures. The performances of deblurring, super-resolution and inpainting are evaluated for several images with the iterative-shrinkage/thresholding algorithm. It is verified that the proposed dictionary yields comparable or superior restoration performance to the non-subsampled Haar transform.**

***Index Terms*—DirLOT, sparse representation, deblurring, super-resolution, inpainting, iterative-shrinkage/thresholding algorithm**

I. Introduction

In this study, we deal with a common problem of image restoration e.g. deblurring, super-resolution and inpainting [1]–[3]. Let $\mathbf{x} \in \mathbb{R}^N$ be an observed image which is represented by

$$\mathbf{x} = \mathbf{P}\mathbf{u}^* + \mathbf{w},$$

where $\mathbf{u}^* \in \mathbb{R}^M$ $(M \geq N)$ is an unknown original image, $\mathbf{P} \in \mathbb{R}^{N \times M}$ is a linear discrete operator which represents degradation and pixel loss through the measurement process, and $\mathbf{w} \in \mathbb{R}^N$ is a measurement noise modeled as a zero-mean additive white Gaussian noise (AWGN), respectively.

Image restoration is a problem of finding a good candidate image $\hat{\mathbf{u}} \in \mathbb{R}^M$ of the unknown high-resolution clean image $\mathbf{u}^*$ only from the observed image $\mathbf{x}$. Since the operator $\mathbf{P}$ is in general not invertible, the problem is ill-posed. In this situation, sparsity works well for the solution, and the iterative shrinkage approaches converges to an exact result [1]–[3]. The framework of the problem setting is shown in Fig. 1

In the sparse representation approach, the candidate $\hat{\mathbf{u}}$ is expressed by a linear-combination of image prototypes (atoms) in a dictionary $\mathbf{D} \in \mathbb{R}^{M \times L}$, i.e.

$$\hat{\mathbf{u}} = \mathbf{D}\hat{\mathbf{y}},$$

where $\hat{\mathbf{y}} \in \mathbb{R}^L$ is a candidate coefficient vector, and refers to the solution of the following form of optimization problem:

$$\hat{\mathbf{y}} = \arg\min_{\mathbf{y}} \|\mathbf{x} - \mathbf{P}\mathbf{D}\mathbf{y}\|_2^2 + \lambda\rho(\mathbf{y}), \tag{1}$$

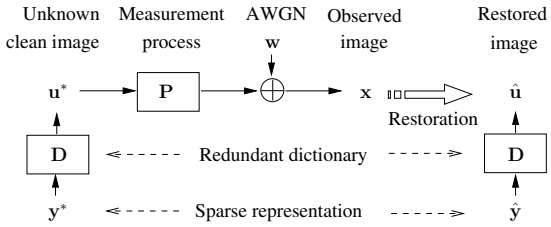

Fig. 1. Framework of problem setting

where $\|\cdot\|_2$ is the ℓ_2-norm of a vector, $\mathbf{y} \in \mathbb{R}^L$ is a coefficient vector, $\rho(\cdot)$ is a regularization term and $\lambda \in \mathbb{R}$ is a scalar parameter to control the trade-off between reconstruction fidelity and sparsity. When $\rho(\cdot)$ is a convex function, the proximal forward-backward algorithm can be used to solve Eq. (1). For $\rho(\mathbf{y}) = \|\mathbf{y}\|_1$, i.e. the ℓ_1-norm regularization is selected, the solver reduces to the iterative shrinkage/thresholding algorithm (ISTA), which guarantees the convergence to exact solution and applicable to large data such as images [4].

The selection of the dictionary $\mathbf{D}$, i.e. the modeling of the original signal, is a quite important task for solving the problem in Eq. (1) since it influences both of the restoration quality and the computational complexity. Recent development of image transforms involves non-separable ones for handling diagonal edges and textures since separable transforms are weak in representing such geometrical structures [3]. As a previous work, we have proposed 2-D non-separable directional lapped orthogonal transforms (DirLOTs) [5]. The bases are allowed to be anisotropic with the fixed-critically-subsampling, overlapping, orthogonal, symmetric, real-valued and compact-support property. The hierarchical tree construction yields a 2-D directional symmetric orthonormal discrete wavelet transform (SOWT). In the articles [5]–[7], we adopted a union of the directional SOWTs as a dictionary $\mathbf{D}$ so that we obtain efficient image denoising. In this work, we further extend the application of DirLOTs to image restoration with ISTA.

II. Union of Directional SOWT

In this section, let us review DirLOTs, its hierarchical construction to obtain directional symmetric orthonormal wavelet transforms (DirSOWTs) and a union of DirSOWTs.

A. Review of 2-D Directional LOT

DirLOTs are 2-D non-separable transforms and has a capability to be directional. Compared with existing transforms, DirLOTs have a special feature that the system simultaneously satisfies the fixed-critically-subsampling, overlapping, orthonormal, symmetric, real-valued and compact-support property with a non-separable basis. As well, it can hold the trend vanishing moments (TVMs) for any direction and has an appropriate boundary operation [5], [8]. The directional property works well for diagonal textures and edges[1].

Let $\mathbf{z} = (z_{\mathrm{y}}, z_{\mathrm{x}})^T \in \mathbb{C}^2$ be a variable vector in the 2-D Z-transform domain. Then, the polyphase matrix of order $[N_{\mathrm{y}}, N_{\mathrm{x}}]^T$ is represented by the following product form:

$$\mathbf{E}(\mathbf{z}) = \prod_{n_{\mathrm{y}}=1}^{N_{\mathrm{y}}} \left\{ \mathbf{R}_{n_{\mathrm{y}}}^{\{\mathrm{y}\}} \mathbf{Q}(z_{\mathrm{y}}) \right\} \cdot \prod_{n_{\mathrm{x}}=1}^{N_{\mathrm{x}}} \left\{ \mathbf{R}_{n_{\mathrm{x}}}^{\{\mathrm{x}\}} \mathbf{Q}(z_{\mathrm{x}}) \right\} \cdot \mathbf{R}_0 \mathbf{E}_0, \tag{2}$$

where $\mathbf{Q}(z) = \frac{1}{2} \left(\begin{smallmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{smallmatrix}\right) \left(\begin{smallmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & z^{-1}\mathbf{I} \end{smallmatrix}\right) \left(\begin{smallmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{smallmatrix}\right)$, $\mathbf{R}_0 = \left(\begin{smallmatrix} \mathbf{W}_0 & \mathbf{O} \\ \mathbf{O} & \mathbf{U}_0 \end{smallmatrix}\right)$, and $\mathbf{R}_n^{\{d\}} = \left(\begin{smallmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{U}_n^{\{d\}} \end{smallmatrix}\right)$. The product of sequential matrices is defined by $\prod_{n=1}^{N} \mathbf{A}_n = \mathbf{A}_N \mathbf{A}_{N-1} \cdots \mathbf{A}_2 \mathbf{A}_1$. $\mathbf{E}_0$ is an $M \times M$ symmetric orthonormal transform matrix given directly through the 2-D separable discrete cosine transform (DCT), where M is the number of channels. Symbols $\mathbf{W}_0$, $\mathbf{U}_0$ and $\mathbf{U}_{n_{\mathrm{d}}}^{\{d\}}$ denote orthonormal matrices of size $M/2 \times M/2$, which are freely controlled during the design process. The support region of each analysis (or synthesis) filter results in $2(N_{\mathrm{y}}+1) \times 2(N_{\mathrm{x}}+1)$.

B. Construction of Redundant Dictionary

A single DirLOT has a drawback to represent multiple directional structures in images. In this section, let us review constructing a union of tree-structured DirLOTs as a dictionary $\mathbf{D}$ [5]–[7].

A simple way to construct a redundant dictionary is to unite multiple unitary matrices. Shift invariant, i.e. non-subsampling, dictionary with orthonormal transforms is also another popular way. The existing techniques, however, do not take account of the directional property. In this work, we propose to construct a dictionary by using multiple DirSOWT so that diagonal textures and edges are sparsely represented. Our proposed dictionary $\mathbf{D}$ is represented by

$$\mathbf{D} = \begin{bmatrix} \mathbf{\Phi}_{0\cup\frac{\pi}{2}}^T & \mathbf{\Phi}_{\phi_1}^T & \mathbf{\Phi}_{\phi_2}^T & \mathbf{\Phi}_{\phi_3}^T & \cdots & \mathbf{\Phi}_{\phi_{K-1}}^T \end{bmatrix}, \tag{3}$$

where $\mathbf{\Phi}_{0\cup\frac{\pi}{2}}$ is a nondirectional symmetric orthonormal DWT with the classical two-order vanishing moments (VMs) [9], and $\mathbf{\Phi}_{\phi}$ is a DirSOWT constructed by a DirLOT with the two-order TVMs for the direction $\mathbf{u}_{\phi}$ [5]. K denotes the number of the DWTs, i.e. the redundancy of dictionary $\mathbf{D}$. Since $\mathbf{D}$ gives a tight frame with normalized atoms, $\mathbf{D}\mathbf{D}^T = K\mathbf{I}$ holds.

[1]DirLOT Toolbox is available from MATLAB Central (`http://www.mathworks.com/matlabcentral/`).

Data: Observed picture $\mathbf{x} \in \mathbb{R}^N$
Result: Restored picture $\hat{\mathbf{u}} \in \mathbb{R}^M$
Initialization;
$i \leftarrow 0$;
$\mathbf{y}^{(0)} \leftarrow \frac{1}{K}\mathbf{A}^T\mathbf{x}$;
Main iteration to find $\mathbf{y}$ that minimizes
$f(\mathbf{y}) = \|\mathbf{x} - \mathbf{A}\mathbf{y}\|_2^2 + \lambda\|\mathbf{y}\|_1$;
repeat
  $i \leftarrow i + 1$;
  $\mathbf{y}^{(i)} \leftarrow \mathcal{T}_{\frac{\lambda}{\alpha}}\left(\mathbf{y}^{(i-1)} - \frac{2}{\alpha}\mathbf{A}^T(\mathbf{A}\mathbf{y}^{(i-1)} - \mathbf{x})\right)$;
until $\|\mathbf{y}^{(i)} - \mathbf{y}^{(i-1)}\|_2^2 / \|\mathbf{y}^{(i)}\|_2^2 < \epsilon$;
$\hat{\mathbf{u}} \leftarrow \mathbf{D}\mathbf{y}^{(i)}$;

Algorithm 1: ISTA, where $\mathbf{A} = \mathbf{P}\mathbf{D}$ and $\alpha = 2\lambda_{\max}(\mathbf{A}^T\mathbf{A}) = 2K\lambda_{\max}(\mathbf{P}^T\mathbf{P})$, i.e. the Lipschitz constant of the gradient of $\|\mathbf{x} - \mathbf{A}\mathbf{y}\|_2^2$ [4].

III. Image Restoration with ISTA

We here propose to apply a union of DirSOWTs to image restoration problems as shown in Fig. 1. Due to the measurement process $\mathbf{P}$, the matrix $\mathbf{A} = \mathbf{P}\mathbf{D} \in \mathbb{R}^{N\times L}(N \le L)$ which relates the observed image $\mathbf{x}$ to the coefficient vector $\mathbf{y}^*$ is no longer a tight frame, much less a union of unitary matrices, and the norms of the column vectors are not guaranteed to be unit. There are several approaches to solve the problem. ISTA is an example of such solvers, which is computationally efficient and applicable to large data such as images [4].

A. ISTA with a tight frame

If we select the ℓ_1-norm as the sparsity measure $\rho(\cdot)$, we can use ISTA. ISTA for $\mathbf{D}$ is written as shown in Algorithm 1 [4], where $\mathcal{T}_{\lambda}(\cdot)$ is the vector function that performs the element-wise scalar soft-shrinkage operation

$$\mathcal{T}_{\lambda}(\mathbf{v}) = \mathrm{diag}(\mathrm{sign}(\mathbf{v})) \cdot (|\mathbf{v}| - \lambda\mathbf{1})_+,$$

where $\mathrm{sign}(\cdot)$ and $|\cdot|$ take the element-wise signs and absolute values, respectively, and $(\cdot)_+$ replaces negative elements to zeros and remains positive elements. The Lipschitz constant α is determined only by the degradation process since $\mathbf{D}$ constitutes a tight frame and then $\lambda_{\max}(\mathbf{A}^T\mathbf{A}) = \lambda_{\max}(\mathbf{A}\mathbf{A}^T) = K\lambda_{\max}(\mathbf{P}\mathbf{P}^T) = K\lambda_{\max}(\mathbf{P}^T\mathbf{P})$ holds, where $\lambda_{\max}(\cdot)$ denotes the maximum eigen value.

In Algorithm 1, we can use the fact that the dictionary $\mathbf{D}$ is decomposed into K orthonormal matrices and thus the block-wise implementation is available.

B. Examples of Measurement Process

The linear operation $\mathbf{P}$ includes blur, decimation and/or pixel loss. Let us summerize the image restoration of these problems.

1) Deblurring: Deblurring is a problem to restore a clear picture from blurred one, where AWGN is often assumed. In the framework shown in Fig. 1, $\mathbf{P}$ is modeled as a convolution matrix which consists of the impulse response, i.e. point-spread-function (PSF), with spatial shifts. The operation with

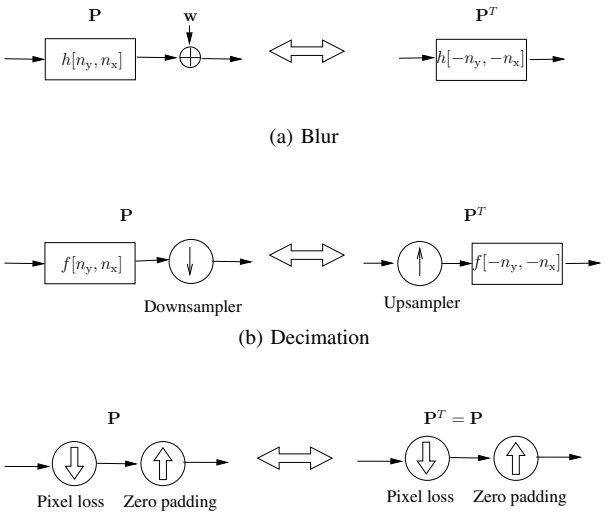

Fig. 2. Measurement processes with representative degradations and their transposed operations.

TABLE I
ADOPTED TRANSFORMS AND THE FEATURES.

Abrv.	Features
NSHT	Two-level non-subsampled Haar DWT, separable, tight, nondirectional
UDN4	Union of six-level SON4 and DirSOWTs with two TVMs of $[N_{\rm y}, N_{\rm x}]^T = [4,4]^T$, nonseparable, tight, multidirectional

$\mathbf{P}^T$ required ISTA is realized by the convolution with spatially reversal system, i.e. the 180-degree rotated filter as shown in Fig. 2 (a).

2) Super-resolution: Super-resolution is a problem to restore a clear high-resolution picture from decimated or low-resolution one. In Fig. 1, $\mathbf{P}$ is modeled as a convolution and downsampling matrix. The operation with $\mathbf{P}^T$ is composed of the upsampling and 180-degree-rotated convolution matrix. Fig. 2 (b) shows the pair of the operations.

3) Inpainting: Inpainting is a problem to restore a missing pixels from the other observed remaining pixels. $\mathbf{P}$ is simply modeled as a diagonal matrix of which elements are either of 0 or 1, which denote missing and remaining pixel position, respectively. Thus, the operation with $\mathbf{P}^T$ is exactly the same as $\mathbf{P}$ since $\mathbf{P}^T = \mathbf{P}$. Fig. 2 (c) shows the pair of the operations.

IV. SIMULATION RESULTS

This section shows some simulation results of deblurring, super-resolution and inpainting in order to verify the significance of DirSOWTs. To assess the performance, it is compared with the non-subsampled Haar transform (NSHT). The transforms adopted in this simulation are summarized in Tab. I, where UDN4 denotes the proposed dictionary $\mathbf{D}$ consisting of multiple DirSOWTs of $[N_{\rm y}, N_{\rm x}]^T = [4,4]^T$. We select four angles for ϕ_k in Eq. (3) as $\phi_k \in \{-\frac{\pi}{6}\frac{\pi}{6}, \frac{2\pi}{6}, \frac{4\pi}{6}\}$.

(a) *goldhill* (b) *lena* (c) *barbara* (d) *baboon*

Fig. 3. Original pictures $\mathbf{u}^*$ of size 512×512, 8-bit grayscale.

TABLE II
COMPARISON OF PSNRS AMONG THREE METHODS FOR VARIOUS PICTURES AND MEASUREMENT PROCESSES, WHERE PARAMETER λ, OF WHICH VALUE IS GIVEN IN THE PARENTHESIS, IS EXPERIMENTALLY GIVEN. 'CLASSICAL' MEANS WIENER, BICUBIC AND MEDIAN FILTER FOR DEBLURRING, SUPER-RESOLUTION AND INPAINTING, RESPECTIVELY.

Process	Picture	Classical	NSHT	UDN4
Deblurring	*goldhill*	24.41	28.45 (0.0010)	**28.48** (0.0033)
	lena	25.41	29.90 (0.0011)	**30.12** (0.0047)
	barbara	22.12	23.92 (0.0011)	**23.94** (0.0036)
	baboon	21.15	21.84 (0.0007)	21.84 (0.0000)
Super Resolution	*goldhill*	25.97	**29.11** (0.0003)	28.98 (0.0005)
	lena	26.98	**30.69** (0.0005)	30.61 (0.0012)
	barbara	22.91	**24.14** (0.0004)	24.10 (0.0004)
	baboon	20.84	**21.97** (0.0002)	21.93 (0.0003)
Inpainting	*goldhill*	21.70	21.34 (0.0259)	**33.73** (0.0332)
	lena	21.73	21.27 (0.0282)	**35.25** (0.0350)
	barbara	19.86	20.42 (0.0231)	**31.64** (0.0378)
	baboon	18.83	19.10 (0.0201)	**27.96** (0.0384)

Thus, the redundancy results in $K = 1+4 = 5$. The number of levels of each DWT is set to six. The basis termination method for the boundary operation is also applied to [8]. The non-subsampling Haar wavelet adopted here has two-level construction, where the redundancy results in $K = 1+3+3 = 7$ and identical to our dictionary.

Fig. 3 shows pictures used as unknown clean ones, $\mathbf{u}^*$, and Tabs. II and III summerize the performance evaluations in terms of the peak-signal to noise ratio (PSNR) and structural similarity (SSIM) indeces. The SSIM index measures a simi-

TABLE III
COMPARISON OF SSIM INDEXES AMONG THREE METHODS FOR VARIOUS PICTURES AND MEASUREMENT PROCESSES, WHERE PARAMETER λ, OF WHICH VALUE IS GIVEN IN THE PARENTHESIS, IS EXPERIMENTALLY GIVEN. 'CLASSICAL' MEANS WIENER, BICUBIC AND MEDIAN FILTER FOR DEBLURRING, SUPER-RESOLUTION AND INPAINTING, RESPECTIVELY.

Process	Picture	Classical	NSHT	UDN4
Deblurring	*goldhill*	0.633	0.720 (0.0007)	**0.724** (0.0026)
	lena	0.666	0.796 (0.0006)	**0.820** (0.0054)
	barbara	0.543	0.654 (0.0007)	**0.667** (0.0043)
	baboon	0.517	**0.529** (0.0005)	0.528 (0.0000)
Super Resolution	*goldhill*	0.982	**0.767** (0.0003)	0.759 (0.0004)
	lena	0.802	**0.859** (0.0003)	0.854 (0.0004)
	barbara	0.646	**0.701** (0.0004)	0.696 (0.0004)
	baboon	0.433	**0.551** (0.0002)	0.544 (0.0003)
Inpainting	*goldhill*	0.632	0.609 (0.0318)	**0.931** (0.0197)
	lena	0.655	0.560 (0.0385)	**0.945** (0.0206)
	barbara	0.602	0.647 (0.0300)	**0.944** (0.0220)
	baboon	0.522	0.638 (0.0324)	**0.907** (0.0251)

(a) Observed (b) Wiener (c) NSHT (d) UDN4

Fig. 4. Partial results of deblurring for "*lena*."

(a) Observed (b) Bicubic (c) NSHT (d) UDN4

Fig. 5. Partial results of super-resolution for "*lena*."

larity of two images, which approaches to one when the two images are perceptually close to each other[2] [10].

A. Deblurring

As the PSF $h[n_{\rm y}, n_{\rm x}]$, we used the 2-D Gaussian filter with standard deviation $\sigma_{\rm h} = 2$. AWGN is also assumed with standard deviation $\sigma_{\rm n} = 5$. Fig. 4 shows an observed picture of "*lena*" and three different restoration results. As a classical deblurring technique, we adopted the Wiener filter (MATLAB '`deconvwnr`' function). From Tabs. II and III, it is observed that UDN4 shows almost the best performance among three methods in terms of PSNR and SSIM index.

B. Super Resolution

In this simulation, we assumed the 2-D Gaussian filter with standard deviation $\sigma_{\rm h} = 2$ as a PSF $f[n_{\rm y}, n_{\rm x}]$ and the down-sampling with factor two. Any noise is not explicitly added. In Fig. 5 the super-resolution performances are compared among three methods for "*lena*." As a classical scaling-up technique, we used the bicubic interpolator (MATLAB '`imresize`' function). From Tabs. II and III, it is observed that the performances of NSHT and UDN4 show comparable to each other and significantly superior to the bicubic interpolation.

The performance of NSHT is slightly superior to that of UDN4. Our conjecture is that the Haar transform is suitable for the assumed degradation process with the isotropic Gaussian blur and the rectangular subsampling.

C. Inpainting

Fig. 6 compares the inpainting performances among three methods for "*lena*." The observed picture in Fig. 6 loses 20% pixels randomly. As a classical approach to fill the lost pixels, we used the median filter (MATLAB '`medfilt2`' function). Any noise is not explicitly assumed. From Tabs. II and III, UDN4 shows significant performance improvement of inpainting. This is because UDN4 has larger extent of basis images than those of NSHT.

[2]MATLAB function `ssim_index.m` from `http://www.cns.nyu.edu/~lcv/ssim/` was used.

(a) Observed (b) Median (c) NSHT (d) UDN4

Fig. 6. Partial results of inpainting for "*lena*."

V. Conclusions

A novel image restoration technique was proposed by introducing a union of DirLOTs. Through the application to the deblurring, super-resolution and inpainting with ISTA, it is verified that the proposed dictionary shows superior or comparable performance to the non-subsampled Haar transform.

Acknowledgment

This work was supported by JSPS KAKENHI (23560443). We would like to express our appreciation to Dr. Masao Yamagishi and Mr. Shunsuke Ono, Tokyo Institute of Technology, for their valuable comments.

References

[1] Michael Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, 2010.

[2] Stèphane Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, Academic Press, 2008.

[3] Jean-Luc Starck, Fionn Murtagh, and Jala M. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*, Cambridge University Press, 2010.

[4] Daniel P. Palomar and Yonina C. Eldar, Eds., *Convex Optimization in Signal Processing and Communications*, Cambridge University Press, 2009.

[5] Shogo Muramatsu, Dandan Han, Tomoya Kobayashi, and Hisakazu Kikuchi, "Directional lapped orthogonal transform: Theory and design," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2434–2448, May 2012.

[6] Shogo Muramatsu and Dandan Han, "Image denoising with union of directional orthonormal DWTs," in *IEEE Proc. of ICASSP*, Mar. 2012, pp. 1089–1092.

[7] Shogo Muramatsu, "SURE-LET image denoising with multiple directional LOTs," in *IEEE Proc. of PCS*, May 2012.

[8] Shogo Muramatsu, Tomoya Kobayashi, Minoru Hiki, and Hisakazu Kikuchi, "Boundary operation of 2-D non-separable linear-phase paraunitary filter banks," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2314–2318, Apr. 2012.

[9] Tomoya Kobayashi, Shogo Muramatsu, and Hisakazu Kikuchi, "Two-degree vanishing moments on 2-D non-separable GenLOT," in *IEEE Proc. of ISPACS*, Dec. 2009, pp. 248–251.

[10] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.