

Data-Driven Tight Frame for Multi-Channel Images and Its Application to Joint Color-Depth Image Reconstruction

Jin Wang*

Jian-Feng Cai[†]

Abstract

In image restoration, we usually assume that the underlying image has a good sparse approximation under a certain system. Wavelet tight frame system has been proven to be such an efficient system to sparsely approximate piecewise smooth images. Thus it has been widely used in many practical image restoration problems. However, images from different scenarios are so diverse that no static wavelet tight frame system can sparsely approximate all of them well. To overcome this, recently, Cai et. al. [7] proposed a method that derives a data-driven tight frame adapted to the specific input image, leading to a better sparse approximation. The data-driven tight frame has been applied successfully to image denoising and CT image reconstruction. In this paper, we extend this data-driven tight frame construction method to multi-channel images. We construct a discrete tight frame system for each channel and assume their sparse coefficients having a joint sparsity. The multi-channel data-driven tight frame construction scheme is applied to joint color and depth image reconstruction. Experimental results show that the proposed approach has a better performance than state-of-the-art joint color and depth image reconstruction approaches.

1 Introduction

Image restoration [12] is one of the most important areas in imaging science. Its major task is to estimate the clean original image from a corrupted/noisy image during the process of imaging, acquisition and communication. A popular class of approaches for image restoration are the sparsity-based regularization methods, assuming the underlying image has a good sparse approximation under some properly designed system. Actually, in the past decades, sparse approximation not only has been a powerful tool for image restoration but also has played important roles in many signal processing areas such as compression and data analysis.

The foundation of sparse approximation for image restoration is the design of systems under which a large class of images have a good sparse approximation. There are various such systems available. Some of them are bases, which use linearly independent atoms, and therefore the dimension of the coefficients is minimized. One representative class of bases for sparse approximation are orthonormal wavelet bases [13, 26], which can approximate piecewise smooth images very efficiently with only a few non-zero wavelet coefficients. This leads to the great success of orthonormal wavelet in image processing. However, linear independency of bases limits the number of atoms in the basis system. As a result, the sparse approximation ability by a basis is limited. Also, the linear independency constraint introduces inflexibility in the construction of bases for sparse approximation. Due to these reasons, over-complete systems, which sacrifice the linear independency of atoms in bases, have become more and more widely used in sparse approximation. When the linear independency in orthonormal wavelet bases are discarded while keeping Parseval's identity, one gets over-complete systems called wavelet tight frames derived from extension principles [14, 28]. While wavelet tight

*Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing, China, 100124. Email: reakwang@gmail.com

[†]Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA. Email: jianfeng-cai@uiowa.edu. This author is partially supported by NSF DMS-1418737.

frames inherit many good properties from orthonormal wavelet bases such as efficient decomposition and reconstruction algorithms, they possess a more powerful sparse approximation ability for piecewise smooth images than orthonormal wavelet bases. Wavelet tight frames are now widely used in image restoration for various tasks; see, e.g., [4, 6, 8, 9]. Some other over-complete systems that have a good sparse approximation for images are X-lets [10, 21, 22].

Images from real applications are diverse. Different applications give images with different characteristics. For example, images arising from medical imaging contain relatively fewer structures than images captured by a digital camera. Even for the same application, images behavior quite differently from instance to instance. The images of human faces captured by a digital camera are definitely different from those of street views. Consequently, systems of sparse approximation cannot always perform very well for practical problems. For example, it is well known that both orthonormal wavelet and wavelet tight frame cannot handle images with abundant texture structures very well. Motivated by this, [7] recently proposed a method that derives a discrete tight frame system, called data-driven tight frame, from the input image itself to provide a better sparse approximation to the input image. By relaxing the unitary extension principle [17, 28], an optimization involving a 0-norm and a unitary matrix constraint is used to obtain the filters in the data-driven tight frame. An alternating minimization scheme is employed for the resulting optimization. However, its convergent analysis is not provided in [7]. To overcome this issue, [2] proposed a theoretically convergent algorithm for data-driven tight frame construction. The advantage of data-driven tight frame over static tight frame has been demonstrated for image denoising [7], image inpainting [1], CT image reconstruction [34], and seismic data restoration [23]. In [1], the data-driven tight frame construction method is reformulated as an orthonormal dictionary learning in the image patch space.

In this paper, we extend the data-driven tight frame construction method to multi-channel images. We construct one discrete tight frame system for each channel and assume their sparse coefficients having a common sparsity pattern. This leads to an optimization problem involving a vector 0-norm and several orthonormal matrix constraints. The multi-channel data-driven tight frame is applied to joint color and depth image (RGB-D image) reconstruction, which arises from recent development on the emerging consumer RGB-D cameras. The RGB-D cameras can capture both the color image and the depth of the scene. However, the depth in an RGB-D image is often noisy and incomplete, especially around object boundaries. Furthermore, the RGB-D camera may produce a severely noised color image under some bad illumination conditions. Thus the enhancement of depth image and RGB-D image is receiving increasing research interest [18, 24, 30, 35]. We will employ the proposed multi-channel data-driven tight frame for the joint task of restoration of the depth image and noise removal of the color image. The experiments showed that the proposed approach has a better performance than state-of-the-art joint color and depth image reconstruction approaches in [24].

The rest of the paper is organized as follows. In Section 2, we give some preliminary backgrounds on wavelet tight frame, and the data-driven tight frame construction method in [7]. In Section 3, we present our multi-channel data-driven tight frame construction scheme, which will be applied to joint color and depth image reconstruction in Section 4. The paper is concluded in Section 5.

2 Tight Frames and Data-Driven Tight Frames

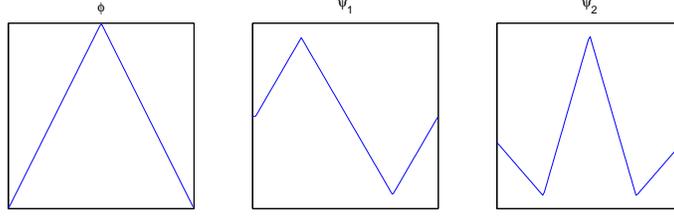
This section is devoted to some preliminaries of wavelet tight frame and data-driven tight frame construction. In Section 2.1, wavelet tight frame and its construction from unitary extension principles are introduced. In Section 2.2, we present the construction scheme of data-driven tight frames.

2.1 Tight Frame and Wavelet

In this section, we discuss briefly on tight frame, wavelet tight frame, and their construction from unitary extension principle; see [14, 28, 32] for more details. Let \mathcal{H} be a Hilbert space and $\{x_n\} \subset \mathcal{H}$ be a countable set. Then, we call $\{x_n\}$ a tight frame for \mathcal{H} if

$$\|x\|^2 = \sum_n |\langle x, x_n \rangle|^2, \quad \forall x \in \mathcal{H}. \quad (1)$$

Figure 1: The scaling function and framelets in piecewise linear B-spline wavelet tight frame



It is easy to see that any orthonormal basis is a tight frame. However, a tight frame is not necessarily an orthonormal basis. For example, the union of two distinct orthonormal bases divided by $\sqrt{2}$ is a tight frame but not an orthonormal basis. Therefore, tight frame is a generalization of orthonormal basis. While keeping Parseval's identity of orthonormal basis, tight frame discards linear dependency, and it is over-complete.

Wavelet tight frames are a widely used class of tight frames for sparse approximation in image restoration. We start with the univariate case where $\mathcal{H} = L_2(\mathbb{R})$. In wavelet tight frame systems, all the functions are shifts and dilations of some finite generators. In particular, let $\Psi = \{\psi_1, \psi_2, \psi_3, \dots, \psi_{m-1}\} \in L_2(\mathbb{R})$ be a set of finite generators. Then, we call the following system $X(\Psi)$ a wavelet tight frame if it forms a tight frame in $L_2(\mathbb{R})$:

$$X(\Psi) = \{\psi_{j,k}^i, 1 \leq i \leq m-1, j \in \mathbb{Z}, k \in \mathbb{Z}\},$$

where $\psi_{j,k}^i = 2^{j/2}\psi_i(2^j \cdot -k)$.

The construction of wavelet tight frames is then the construction of Ψ . One construction scheme is using multi-resolution analysis (MRA) and the unitary extension principle (UEP). We start from a compactly supported refinable function ϕ (often referred as a scaling function) with a refinement mask \mathbf{a}_0 satisfying

$$\widehat{\phi}(2\cdot) = \widehat{\mathbf{a}_0}\widehat{\phi},$$

where $\widehat{\phi}$ is the Fourier transform of ϕ , and $\widehat{\mathbf{a}_0}$ is a 2π -periodic trigonometric polynomial defined as $\widehat{\mathbf{a}_0}(\omega) := \sum_{k \in \mathbb{Z}} \mathbf{a}_0[k]e^{ik\omega}$ and satisfies $\widehat{\mathbf{a}_0}(0) = 1$. We then define the generators ψ_ℓ , $i = 1, \dots, m-1$, in Ψ in the Fourier domain by

$$\widehat{\psi}_i(2\cdot) = \widehat{\mathbf{a}_i}\widehat{\phi}, \quad i = 1, \dots, m-1,$$

where the associated masks $\{\widehat{\mathbf{a}_i}\}_{i=1}^{m-1}$ are 2π -periodic trigonometric polynomials $\widehat{\mathbf{a}_i}(\omega) = \sum_{k \in \mathbb{Z}} \mathbf{a}_i[k]e^{ik\omega}$. The unitary extension principle (UEP) [14, 28] tells that $X(\Psi)$ forms a tight frame provided that

$$\sum_{i=0}^{m-1} \widehat{\mathbf{a}_i}(\omega)\overline{\widehat{\mathbf{a}_i}(\omega + \gamma\pi)} = \delta_\gamma, \quad \forall \gamma \in \{0, 1\}, \quad (2)$$

where $\delta_0 = 1$ and $\delta_1 = 0$. The elements in Ψ are called framelets. The sequence \mathbf{a}_0 is called the low-pass filter, and $\mathbf{a}_1, \dots, \mathbf{a}_{m-1}$ are called high-pass filters.

One example class of wavelet tight frames derived from MRA and UEP are the B-spline wavelet tight frame systems. Figure 1 depicts the scaling function and the framelets in piecewise linear B-spline wavelet tight frames, and the associated low-pass and high-pass filters are

$$\mathbf{a}_0 = \frac{1}{4}[1, 2, 1], \quad \mathbf{a}_1 = \frac{\sqrt{2}}{4}[1, 0, -1], \quad \mathbf{a}_2 = \frac{1}{4}[-1, 2, -1].$$

Wavelet tight frames for $L_2(\mathbb{R}^2)$ can be obtained by tensor products: the scaling function, framelets, and filters for wavelet tight frames in $L_2(\mathbb{R}^2)$ are tensor products of their correspondences in $L_2(\mathbb{R})$.

Digital images are vectors in \mathbb{R}^{N^2} that are discretizations of functions in $L_2(\mathbb{R}^2)$. One can construct tight frame systems for digital images using filters of wavelet tight frames with proper boundary conditions. The

construction is derived from discrete wavelet tight frame transform. Let $\{\mathbf{a}_i\}_{i=0}^{m-1}$ be the low-pass filter and high-pass filter associated with the wavelet tight frame in $L_2(\mathbb{R}^2)$. Let $\mathbf{a} * \mathbf{b}$ be the filtering of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{N^2}$, viewed as 2D arrays of size $N \times N$, with proper boundary conditions. Then, the one level translation invariant discrete wavelet tight frame transform is a linear transform

$$W : \mathbf{x} \in \mathbb{R}^{N^2} \rightarrow \begin{bmatrix} \mathbf{a}_0 * \mathbf{x} \\ \vdots \\ \mathbf{a}_{m-1} * \mathbf{x} \end{bmatrix} \in \mathbb{R}^{mN^2}. \quad (3)$$

The unitary condition (1) and UEP (2) implies that $W^T W = I$. In other words, the rows of W forms a tight frame in \mathbb{R}^{N^2} . See [3, 11, 32] for more details.

2.2 Data-Driven Tight Frames

Though the tight frame introduced in the previous section is applied widely for image restoration [3, 4, 6, 8, 9, 11, 32], they are not good at reconstructing of images with rich textures. To broaden its applicability and improve its ability of sparse approximation, [7] proposed a scheme for constructing a discrete tight frame, called data-driven tight frame, adapted to a given image.

Since a wavelet tight frame is completely determined by its filters, the construction of data-driven tight frame is constructed by finding its filters. Instead of using static filters no matter what the underlying image is, the data-driven tight frame uses filters that are tailored to the specific input image. In particular, let $\{\mathbf{h}_i\}_{i=0}^{R-1}$ be a family of compact supported 2D filters that are to be determined later, and write

$$\mathbf{H} = [\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_{R-1}].$$

In the following, we will not distinct the set of filters $\{\mathbf{h}_i\}_{i=0}^{R-1}$ and the matrix of filters \mathbf{H} . Then, by (3), the associated one level tight frame transform is

$$W(\mathbf{H}) : \mathbf{x} \in \mathbb{R}^{N^2} \rightarrow \begin{bmatrix} \mathbf{h}_0 * \mathbf{x} \\ \vdots \\ \mathbf{h}_{R-1} * \mathbf{x} \end{bmatrix} \in \mathbb{R}^{RN^2}, \quad (4)$$

where $*$ is the filtering of two 2D arrays. In order rows of $W(\mathbf{H}) \in \mathbb{R}^{RN^2 \times N^2}$ forming a tight frame in \mathbb{R}^{N^2} , the perfect reconstruction property $W(\mathbf{H})^T W(\mathbf{H}) = I$ should be satisfied, which is finally reduced to certain conditions on the filters. By using the UEP, we have the following theorem from [7].

Theorem 1 ([7]). *Assume $\mathbf{H} \in \mathbb{R}^{R \times R}$ with $R = r^2$ for some integer r , and columns of \mathbf{H} are vectorization of 2D arrays of size $r \times r$. Then the rows of $W(\mathbf{H})$ forms a tight frame of \mathbb{R}^{N^2} if $\mathbf{H}^T \mathbf{H} = \mathbf{I}/R$.*

Let $\mathbf{g} \in \mathbb{R}^{N^2}$ be the input image. To get a tight frame, in the rest of the section, we will assume $\mathbf{H} \in \mathbb{R}^{R \times R}$ with $R = r^2$ for some integer r , and columns of \mathbf{H} are vectorization of 2D arrays of size $r \times r$. With the help of Theorem 1, one can formulate the data-driven tight frame construction as the following minimization problem

$$\min_{\mathbf{H}, \mathbf{v}} \|W(\mathbf{H})\mathbf{g} - \mathbf{v}\|_2^2 + \lambda^2 \|\mathbf{v}\|_0, \quad \text{subject to} \quad \mathbf{H}^T \mathbf{H} = \mathbf{I}/R, \quad (5)$$

where $\|\mathbf{v}\|_0$ is the 0-norm of \mathbf{v} . The first term of the cost function in (5) is to keep the coefficient $\mathbf{v} \in \mathbb{R}^{RN^2}$ close to $W(\mathbf{H})\mathbf{g}$. The second term is to enforce the sparsity of \mathbf{v} . The constraint ensures that $W(\mathbf{H})$ is a tight frame, according to Theorem 1. Therefore, through (5), we obtain a set of filters \mathbf{H} and a tight frame $W(\mathbf{H})$, under which the input image has a sparse coefficient. Eq. (5) is solved numerically by an alternating direction minimization scheme. Fixing \mathbf{H} , the minimization problem (5) becomes

$$\min_{\mathbf{v}} \|W(\mathbf{H})\mathbf{g} - \mathbf{v}\|_2^2 + \lambda^2 \|\mathbf{v}\|_0,$$

whose solution is

$$\mathbf{v} = \mathcal{T}_\lambda(W(\mathbf{H})\mathbf{g}). \quad (6)$$

Here \mathcal{T}_λ is the hard-thresholding operator defined by

$$\mathcal{T}_\lambda(\mathbf{x}) = \mathbf{x} \cdot \chi_{|\mathbf{x}| \geq \lambda},$$

where $\chi_{|\mathbf{x}| \geq \lambda}$ is the characteristic vector for the componentwise comparison $|\mathbf{x}| \geq \lambda$. Fixing \mathbf{v} , the minimization problem (5) is equivalent to

$$\min_{\mathbf{H}} \|W(\mathbf{H})\mathbf{g} - \mathbf{v}\|_2^2, \quad \text{subject to } \mathbf{H}^T \mathbf{H} = \mathbf{I}/R. \quad (7)$$

The closed-form solution of (7) is given by singular value decomposition (SVD), summarized in the following theorem.

Theorem 2 ([7]). *Let $\mathbf{G} \in \mathbb{R}^{R \times N^2}$ be the matrix by sampling all the $r \times r$ patches of \mathbf{g} , viewed as an $N \times N$ 2D array. Let $\mathbf{V} \in \mathbb{R}^{R \times N^2}$ be the reshape of \mathbf{v} such that the j -th column of \mathbf{V} corresponds to coefficients for pixel j . Let $\mathbf{G}\mathbf{V}^T = \mathbf{P}\mathbf{\Sigma}\mathbf{Q}^T$ be a singular value decomposition of $\mathbf{G}\mathbf{V}^T$. Then, the solution of (7) is given by*

$$\mathbf{H} = \mathbf{P}\mathbf{Q}^T/r. \quad (8)$$

Finally, the construction of data-driven tight frame from \mathbf{g} is to applying formulas (6) and (8) alternatively.

The improvement of data-driven tight frame approach over static tight frame approach has been demonstrated for image denoising [7], image inpainting [1], CT image reconstruction [34], and seismic data restoration [23]. In [1], the data-driven tight frame construction method is reformulated as an orthonormal dictionary learning in the image patch space.

3 Construction of Multi-Channel Data-Driven Tight Frame

In this section, we extend the construction scheme of data-driven tight frames to multi-channel images. We will derive one tight frame for each channel and assume their coefficients having a joint sparsity.

Let

$$\mathbf{G} = [\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \dots, \mathbf{g}^{(c)}]$$

be a multi-channel image, where $\mathbf{g}^{(i)} \in \mathbb{R}^{N^2}$ is the i -th channel. Since each channel may contain information of different features, we construct one tight frame for each channel. Let $\mathbf{H}^{(i)}$, $i = 1, \dots, c$, be the filters of the tight frame for $\mathbf{g}^{(i)}$, and $W(\mathbf{H}^{(i)})$ be its associated tight frame transform as in (4). We assume, for any $i = 1, \dots, c$, $\mathbf{H}^{(i)} \in \mathbb{R}^{R \times R}$ with $R = r^2$ for some integer r and columns of $\mathbf{H}^{(i)}$ are vectorization of 2D arrays of size $r \times r$. Denote by $\mathbf{v}^{(i)}$ the coefficient of $\mathbf{g}^{(i)}$ under the transform of $W(\mathbf{H}^{(i)})$, and write

$$\mathbf{V} = [\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(c)}].$$

As channels in \mathbf{G} are interacted, sparsity of columns of \mathbf{V} should be related. Here we assume that columns of \mathbf{V} have a common sparsity pattern. We assume that most of the rows in \mathbf{V} are zero vectors. Such a joint sparsity constraint is a widely used model for vector-valued data, and it has been successfully used in various applications [15, 16, 31, 33]. To promote such a joint sparsity, we employ $\|\mathbf{V}\|_{0, \text{row}}$ defined in the following as a regularization term

$$\|\mathbf{V}\|_{0, \text{row}} = \sum_{j=1}^{N^2} \chi_{\mathbf{v}_j \neq \mathbf{0}},$$

where \mathbf{v}_j is the j -th row of \mathbf{V} , and $\chi_{\mathbf{v}_j \neq \mathbf{0}}$ is the characteristic function of $\mathbf{v}_j \neq \mathbf{0}$. The function $\|\cdot\|_{0, \text{row}}$ is a generalization of 0-norm to vector-valued data.

Altogether, we construct the data-driven tight frame for multi-channel images by solving the following minimization problem

$$\begin{aligned} \min_{\mathbf{V}, \mathbf{H}^{(1)}, \dots, \mathbf{H}^{(c)}} \quad & \sum_{i=1}^c w_i \|W(\mathbf{H}^{(i)})\mathbf{g}^{(i)} - \mathbf{v}^{(i)}\|_2^2 + \lambda^2 \|\mathbf{V}\|_{0, \text{row}}, \\ \text{subject to} \quad & (\mathbf{H}^{(1)})^T \mathbf{H}^{(1)} = \dots = (\mathbf{H}^{(c)})^T \mathbf{H}^{(c)} = \mathbf{I}/R \end{aligned} \quad (9)$$

where w_i for $i = 1, \dots, c$ are positive constants to balance difference channels. The first term of the objective function in (9) is a data-fidelity term, which is to keep $\mathbf{v}^{(i)}$ the coefficient of $\mathbf{g}^{(i)}$ under the discret tight frame transform $W(\mathbf{H}^{(i)})$; the second term is a regularization term, which forces \mathbf{V} to satisfy the joint sparsity prior; the constraints, together with the additional assumptions on $\mathbf{H}^{(i)}$, ensure that $W(\mathbf{H}^{(i)})$, $i = 1, \dots, c$, are tight frame transforms, according to Theorem 1.

The alternating minimization method is applied to solve (9). Fixing either the filters $\mathbf{H}^{(i)}$, $i = 1, \dots, c$, or the coefficient \mathbf{V} , (9) is solved with respect to the other variable only. When the filters $\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(c)}$ are fixed, solving (9) with respect to \mathbf{V} can be recast as

$$\min_{\mathbf{V}} \sum_{i=1}^c w_i \|W(\mathbf{H}^{(i)})\mathbf{g}^{(i)} - \mathbf{v}^{(i)}\|_2^2 + \lambda^2 \|\mathbf{V}\|_{0, \text{row}}. \quad (10)$$

We have the following theorem on the closed form of the solution of (10).

Theorem 3. *The solution of (10) is*

$$\mathbf{V} = \mathcal{T}_{\lambda, \mathbf{w}, \text{row}} \left([W(\mathbf{H}^{(1)})\mathbf{g}^{(1)}, \dots, W(\mathbf{H}^{(c)})\mathbf{g}^{(c)}] \right). \quad (11)$$

where $\mathcal{T}_{\lambda, \mathbf{w}, \text{row}}$ is a row-vector hard-thresholding operator defined as follows: Let the j -th row of $\mathcal{T}_{\lambda, \mathbf{w}, \text{row}}(\mathbf{X})$ and \mathbf{X} be $[\mathcal{T}_{\lambda, \mathbf{w}, \text{row}}(\mathbf{X})]_j$ and \mathbf{x}_j respectively. Then

$$[\mathcal{T}_{\lambda, \mathbf{w}, \text{row}}(\mathbf{X})]_j = \begin{cases} \mathbf{x}_j, & \text{if } \|\mathbf{x}_j\|_{\mathbf{w}} \geq \lambda, \\ \mathbf{0}, & \text{otherwise,} \end{cases}$$

where $\|\mathbf{z}\|_{\mathbf{w}} = (\sum_{i=1}^c w_i z_i^2)^{1/2}$ is a weighted 2-norm.

Proof. Let $\mathbf{X} = [W(\mathbf{H}^{(1)})\mathbf{g}^{(1)}, \dots, W(\mathbf{H}^{(c)})\mathbf{g}^{(c)}]$. Let x_{ij} and \mathbf{x}_j be the (i, j) -th entry and j -th row of \mathbf{X} respectively. Then, (10) can be rewritten as

$$\min_{\mathbf{V}} \sum_{j=1}^{N^2} \left(\sum_{i=1}^c w_i (x_{ij} - v_j^{(i)})^2 + \lambda^2 \chi_{\mathbf{v}_j \neq \mathbf{0}} \right),$$

meaning that (10) can be separated as N^2 independent small-sized minimization problems

$$\min_{\mathbf{v}_j} \left(\sum_{i=1}^c w_i (x_{ij} - v_j^{(i)})^2 + \lambda^2 \chi_{\mathbf{v}_j \neq \mathbf{0}} \right) \quad (12)$$

for $j = 1, \dots, N^2$. Let us find the solution \mathbf{v}_j of (12). If $\mathbf{x}_j = \mathbf{0}$, then the solution of (12) is obviously $\mathbf{v}_j = \mathbf{0}$. Thus we assume $\mathbf{x}_j \neq \mathbf{0}$. By considering the cases where $\mathbf{v}_j \neq \mathbf{0}$ and $\mathbf{v}_j = \mathbf{0}$ respectively, we rearrange (12) as

$$\min_{\mathbf{v}_j} \begin{cases} \sum_{i=1}^c w_i (x_{ij} - v_j^{(i)})^2 + \lambda^2, & \text{if } \mathbf{v}_j \neq \mathbf{0}, \\ \|\mathbf{x}_j\|_{\mathbf{w}}^2, & \text{if } \mathbf{v}_j = \mathbf{0}. \end{cases}$$

Apparently, the minimizer in the first case is $\mathbf{v}_j = \mathbf{x}_j \neq \mathbf{0}$, and the associated minimum is λ^2 . Therefore, if $\|\mathbf{x}_j\|_{\mathbf{w}} \geq \lambda$, then the minimum is achieved by choosing the first case and the minimizer is $\mathbf{v}_j = \mathbf{x}_j$, otherwise the minimum is achieved by choosing the second case and the minimizer is $\mathbf{v}_j = \mathbf{0}$. \square

When the coefficient \mathbf{V} is fixed, solving (9) with respect to $\mathbf{H}^{(i)}$, $i = 1, \dots, c$, is equivalent to solving the following c independent minimization problems

$$\min_{\mathbf{H}^{(i)}} \|W(\mathbf{H}^{(i)})\mathbf{g}^{(i)} - \mathbf{v}^{(i)}\|_2^2 \quad \text{s.t.} \quad (\mathbf{H}^{(i)})^T \mathbf{H}^{(i)} = \mathbf{I}/R, \quad (13)$$

for $i = 1, \dots, c$. By Theorem 2, the solution of (13) has a closed form

$$\mathbf{H}^{(i)} = \mathbf{P}^{(i)}(\mathbf{Q}^{(i)})^T / r, \quad (14)$$

where $\mathbf{P}^{(i)}$ and $\mathbf{Q}^{(i)}$ are the matrices of singular vectors in the SVD $\mathbf{G}^{(i)}(\mathbf{V}^{(i)})^T = \mathbf{P}^{(i)}\mathbf{\Sigma}(\mathbf{Q}^{(i)})^T$. Here $\mathbf{G}^{(i)} \in \mathbb{R}^{R \times N^2}$ is the matrix by sampling all the $r \times r$ patches of $\mathbf{g}^{(i)}$, viewed as an $N \times N$ 2D array; $\mathbf{V}^{(i)} \in \mathbb{R}^{R \times N^2}$ is the reshape of $\mathbf{v}^{(i)}$ such that the j -th column of $\mathbf{V}^{(i)}$ corresponds to coefficients for pixel j .

Our algorithm for the construction of multi-channel data-driven tight frame is summarized in Algorithm 1.

Algorithm 1 Construction of Multi-Channel Data-Driven Tight Frame

Input: A c -channel image $\mathbf{G} = [\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(c)}]$, and a set of initial filters $\mathbf{A}^{(i)}$, $i = 1, \dots, c$.

Output: A set of filters of $\mathbf{H}^{(i)}$, $i = 1, \dots, c$ of learned multi-channel data-driven tight frame.

Procedures:

1. Initialize $\mathbf{H}_{(0)}^{(i)}$ by setting it to $\mathbf{A}^{(i)}$ for $i = 1, \dots, c$.
 2. Iterates $k = 1, 2, \dots$.
 - (a) Update the coefficient \mathbf{V} by (11) with $\mathbf{H}^{(i)} = \mathbf{H}_{(k-1)}^{(i)}$, $i = 1, \dots, c$.
 - (b) Update the filters $\mathbf{H}_{(k)}^{(i)}$ for $i = 1, \dots, c$ by:
 - i. Construct the matrices $\mathbf{G}^{(i)}$ and $\mathbf{V}^{(i)}$.
 - ii. Compute the SVD of $\mathbf{G}^{(i)}(\mathbf{V}^{(i)})^T = \mathbf{P}^{(i)}\mathbf{\Sigma}(\mathbf{Q}^{(i)})^T$.
 - iii. Obtain $\mathbf{H}_{(k)}^{(i)}$ by (14).
 - (c) If the stopping criteria is met, then set $K = k$ and quit the loop.
 3. Output the filters $\mathbf{H}^{(i)} = \mathbf{H}_{(K)}^{(i)}$ for $i = 1, \dots, c$.
-

4 Application to Joint Color and Depth Image Reconstruction

In this section, we focus on the application of proposed multi-channel data-driven tight frame construction scheme for joint color and depth image reconstruction. Calculating the distance of points in the scene relative to the position of the camera is one of the important tasks in computer vision [19]. Such depth information estimation problems are long-standing and used to be challenging. Recent development of RGB-D sensor is completely changing the world of computer vision, where the emerging consumer RGB-D sensor enables the easy acquisition of depth maps and corresponding color images in real time. Some representatives of commercial RGB-D cameras are Microsoft Kinect and ASUS Xtion Pro. The feature of capturing depth map at a low cost by the RGB-D cameras is gaining the attention of the research community, application developers and the game industry. Related computer vision applications [20, 25] become more and more widespread.

However, the depth images captured by depth cameras are usually noisy and incomplete, especially at the edge area of objects. Also, the color image might be severely noised when it is captured under a bad illumination condition. Thus the enhancement of both depth image and color image is receiving increasing research interest. Varieties of optimization recovery based schemes have been developed to solve this problem. [35] proposes to approximate piecewise smooth functions in frame domain with sparse representation, and a tight frame based scheme is developed to recover depth image from noise and outliers corrupted data. [30] presents a convex optimization framework to simultaneously estimate super-resolution of both depth images and color images, in which the problem is formulated into a single energy minimization with a convex function. Similarly, [18] proposes to utilize the depth cues from the induced optical-flow to enhance the depth images captured by a moving device. The problem is also addressed by minimizing an energy function. [24] assumes that similar RGB-D patches lie in a low dimensional subspace. Based on the low-rank subspace constraint, the depth image enhancement problem is formulated into a low-rank matrix completion problem. However, the low-rank completion model is still not novel enough to represent the diverse images well, especially for the complex texture regions.

4.1 Algorithm

A straightforward method for RGB-D image reconstruction is to apply the data-driven tight frame construction and image reconstruction for each channel independently. However, strong correlation exists among different channels in multi-channel images. See, e.g., [18, 24]. Good algorithms should take this into account to achieve good reconstructed RGB-D image quality. Actually, all of the aforementioned algorithms for depth image reconstruction are based on the fact that the color image (RGB channels) and its corresponding depth image (D channel) are highly related. Motivated by this, we apply the multi-channel data-driven tight frame (DDTF) for the joint task of depth image reconstruction and removal of possible noise in the color image, as channel correlations are considered significantly in the construction of the multi-channel DDTF.

A mathematical description of the joint color and depth image reconstruction is as follows. Let $\mathbf{g}^{(1)}$, $\mathbf{g}^{(2)}$, and $\mathbf{g}^{(3)}$ be the red (R), green (G), and blue (B) channels of the color image respectively, and $\mathbf{g}^{(4)}$ be the depth (D) channel. The observed images (noisy color image and noisy incomplete depth image) are

$$\mathbf{f}^{(i)} = \mathbf{g}^{(i)} + \boldsymbol{\eta}^{(i)}, \quad i = 1, 2, 3, \quad \mathbf{f}_{\Lambda}^{(4)} = \mathbf{g}_{\Lambda}^{(4)} + \boldsymbol{\eta}^{(4)},$$

where $\boldsymbol{\eta}^{(i)}$, $i = 1, 2, 3, 4$ are noises, Λ is the set of positions of observed pixels in the depth image, and \mathbf{x}_{Λ} is the restriction of \mathbf{x} on Λ . We would like to reconstruct both the RGB channels and the D channel.

The RGB-D channel image reconstruction problem can be seen as a problem of inpainting with missing pixels in the D channel only. Suppose that we have already constructed a multi-channel tight frame with filters $\mathbf{H}^{(i)}$, $i = 1, 2, 3, 4$. Inspired by the tight-frame inpainting algorithm in [5], a multi-channel tight frame inpainting algorithm is

$$\begin{cases} \mathbf{g}_{(k+1/2)}^{(4)} = \mathcal{P}_{\Lambda} \mathbf{f}^{(4)} + (\mathcal{I} - \mathcal{P}_{\Lambda}) \mathbf{g}_{(k)}^{(4)}, \\ \begin{bmatrix} \mathbf{v}_{(k+1)}^{(1)}, & \dots, & \mathbf{v}_{(k+1)}^{(4)} \end{bmatrix} = \mathcal{T}_{\lambda, w, \text{row}} \left(\begin{bmatrix} W(\mathbf{H}^{(1)}) \mathbf{f}^{(1)}, & W(\mathbf{H}^{(2)}) \mathbf{f}^{(2)}, & W(\mathbf{H}^{(3)}) \mathbf{f}^{(3)}, & W(\mathbf{H}^{(4)}) \mathbf{g}_{(k+1/2)}^{(4)} \end{bmatrix} \right), \\ \begin{bmatrix} \mathbf{g}_{(k+1)}^{(1)}, & \dots, & \mathbf{g}_{(k+1)}^{(4)} \end{bmatrix} = \begin{bmatrix} (W(\mathbf{H}^{(1)}))^T \mathbf{v}_{(k+1)}^{(1)}, & \dots, & (W(\mathbf{H}^{(4)}))^T \mathbf{v}_{(k+1)}^{(4)} \end{bmatrix}, \end{cases} \quad (15)$$

where \mathcal{P}_{Λ} is the diagonal matrix with diagonals 1 if the position is in Λ and 0 otherwise. The algorithm (15) is explained as follows. Let $\mathbf{g}_{(k)}^{(4)}$ be a guess of the D-channel. We replace its pixels on Λ by the noisy observation $\mathbf{f}_{\Lambda}^{(4)}$, which is done by the first step of (15). This estimated D-channel $\mathbf{g}_{(k+1/2)}^{(4)}$ is combined with the observed RGB-channels $\mathbf{f}^{(1)}$, $\mathbf{f}^{(2)}$, and $\mathbf{f}^{(3)}$ to get an estimated noisy RGB-D image, whom a multi-channel tight frame denoising scheme is applied to in Steps 2 and 3 in (15). By this way, we get a better guess of the depth image, as well as a denoised color image.

We finally obtain our joint color and depth image reconstruction method by combining the multi-channel tight frame inpainting algorithm (15) with the multi-channel data-driven tight frame construction Algorithm

2. In particular, the tight frame filters $\mathbf{H}^{(i)}$, $i = 1, 2, 3, 4$, in (15) is updated in each iteration by the output of one step of Algorithm 1. The algorithm is described in Algorithm 2.

Algorithm 2 Joint Color Image and Depth Image Reconstruction by Multi-Channel Data-Driven Tight Frame

Input: The noisy color image $\mathbf{f}^{(i)}$ for $i = 1, 2, 3$, the noisy incomplete depth image $\mathbf{f}_\Lambda^{(4)}$, and the set Λ of known pixels of the depth image.

Output: The denoised color image $\mathbf{g}^{(i)}$ for $i = 1, 2, 3$, and the reconstructed depth image $\mathbf{g}^{(4)}$.

Procedures:

1. Initialize $\mathbf{g}_{(0)}^{(4)}$ and the set of filters $\mathbf{H}_{(0)}^{(i)}$ for $i = 1, 2, 3, 4$.
2. Iterates $k = 1, 2, \dots$

(a) Set $\mathbf{G}_{(k-1/2)} = [\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \mathbf{f}^{(3)}, \mathcal{P}_\Lambda \mathbf{f}^{(4)} + (\mathcal{I} - \mathcal{P}_\Lambda) \mathbf{g}_{(k-1)}^{(4)}]$.

(b) Construct a multi-channel data-driven tight frame $\mathbf{H}_{(k)}^{(i)}$, $i = 1, 2, 3, 4$, by performing one iteration of Algorithm 1 with input image $\mathbf{G} = \mathbf{G}_{(k-1/2)}$ and filters $\mathbf{A}^{(i)} = \mathbf{H}_{(k-1)}^{(i)}$, $i = 1, 2, 3, 4$.

(c) Denoise $\mathbf{G}_{(k-1/2)}$ by thresholding the multi-channel tight frame coefficients

$$\begin{cases} [\mathbf{v}_{(k)}^{(1)}, \dots, \mathbf{v}_{(k)}^{(4)}] = \mathcal{T}_{\lambda, \mathbf{w}, \text{row}} \left(\left[W(\mathbf{H}_{(k)}^{(1)}) \mathbf{g}_{(k-1/2)}^{(1)}, \dots, W(\mathbf{H}_{(k)}^{(4)}) \mathbf{g}_{(k-1/2)}^{(4)} \right] \right), \\ [\mathbf{g}_{(k)}^{(1)}, \dots, \mathbf{g}_{(k)}^{(4)}] = \left[(W(\mathbf{H}_{(k)}^{(1)}))^T \mathbf{v}_{(k)}^{(1)}, \dots, (W(\mathbf{H}_{(k)}^{(4)}))^T \mathbf{v}_{(k)}^{(4)} \right], \end{cases}$$

and denote $\mathbf{G}_{(k)} = [\mathbf{g}_{(k)}^{(1)}, \dots, \mathbf{g}_{(k)}^{(4)}]$.

(d) If the stopping criteria is met, the set $K = k$ and quit the loop.

3. Output $\mathbf{G}_{(K)}$.
-

4.2 Experimental Results

In this section, we give the experimental validation and results of Algorithm 2, and compare it with state-of-the-art algorithm proposed in [24] for joint color and depth image reconstruction. To validate the efficiency of our scheme, a trivial solution that reconstructs each channel independently by applying the method in [7] is also compared.

The experiments are conducted in Matlab 2014a on a Windows PC with Intel i7 CPU and 8G memory. The public dataset Middlebury stereo dataset [29] is used as color images and depth images¹. For the purpose of comparison, the same experimental setting is utilized to obtain the color and depth images as in [24] (e.g., Gaussian white noise with stand deviation $\sigma = 25$ and $\sigma_d = 5$ is added into the RGB and depth channel respectively). To evaluate the quality of reconstruction, the enhancement result of [27] is used as ground truth, following the settings in [24]. The filters $\mathbf{H}^{(i)}$, $i = 1, \dots, 4$, of the multi-channel tight frame are of size 8×8 and initialized with 8×8 DCT transform. The missing values in the incomplete depth image are initialized by conventional interpolation methods. We set the weights $w_1 = w_2 = w_3 = 1$ in (10) for the RGB channels. As the depth channel is different from RGB channels, the weights w_4 should be different from the others. We choose w_4 manually according to the ratio of the 2-norms of the initial color image and the initial depth image. The range of w_4 is in [1, 5]. The thresholding value for the filter update and coefficient update is set as $\lambda^2 = 14\sigma^2$. In our experiment the number of iterations is fixed $K = 50$.

¹The dataset can be downloaded at <http://web.cecs.pdx.edu/~fliu/project/depth-enhance/Middlebury.htm>

Both the PSNR and visual reconstruction results are compared between our method Algorithm 2 and the low-rank matrix completion based method (LRMC) in [24]. The comparison of LRMC with other popular joint color-depth reconstruction and depth enhancement algorithms are presented in [24], and LRMC has shown to outperform its counterparts in the context of joint color and depth image reconstruction. Therefore, we compare our Algorithm 2 with the state-of-the-art depth enhancement method LRMC only. As a comparison, the result of a straightforward solution of applying the method in [7] in each channel independently is also provided. This is demonstrated in Table 1, Fig. 2 and Fig. 3. More results can be found in our project website. It is seen from Table 1 that the multi-channel data-driven tight frame approach Algorithm 2 produces better image quality in terms of PSNR than LRMC and the straightforward method. Also, Fig. 2 demonstrates the superior visual quality of the reconstructed images by Algorithm 2 over LRMC. For example, Figs. 2(g) and 2(o) produced by Algorithm 2 contain finer textures than Figs. 2(e) and 2(m) by LRMC. For some images LRMC has better results on the depth image, as demonstrated in Fig. 3. Note that in this case our algorithm still have better color image reconstruction results.

Table 1: PSNR comparison of reconstructed color and depth images with LRMC

IMAGE	01			02			03			04		
METHODS	LRMC	[7]	Alg. 2									
Color Image	26.29	26.01	27.32	30.72	29.18	30.95	29.46	29.12	30.24	30.07	29.77	30.91
Depth Image	35.26	35.54	35.63	33.10	31.53	32.29	39.00	38.33	38.92	42.71	42.41	42.99
IMAGE	05			06			07			08		
METHODS	LRMC	[7]	Alg. 2									
Color Image	30.94	30.49	31.54	30.19	29.22	31.05	32.29	31.84	32.58	31.33	30.95	32.14
Depth Image	42.66	43.29	43.33	41.13	41.29	42.12	38.29	37.59	38.00	39.78	37.94	38.18
IMAGE	09			10			11			12		
METHODS	LRMC	[7]	Alg. 2									
Color Image	25.61	25.38	27.08	28.71	28.10	29.20	28.82	27.91	29.09	26.46	26.23	27.57
Depth Image	46.23	48.08	46.86	42.39	41.66	42.33	42.62	41.68	42.16	37.72	37.29	37.57
IMAGE	13			14			15			16		
METHODS	LRMC	[7]	Alg. 2									
Color Image	27.59	26.45	28.23	29.52	28.07	29.77	33.35	33.28	34.50	32.11	31.48	32.17
Depth Image	39.03	38.41	38.56	41.86	41.75	42.13	43.01	42.67	43.77	39.13	38.37	39.37
IMAGE	17			18			19			20		
METHODS	LRMC	[7]	Alg. 2									
Color Image	34.15	33.39	34.29	31.31	29.52	31.57	32.16	30.69	32.83	32.62	31.33	33.31
Depth Image	38.88	38.19	39.32	38.59	37.40	37.91	39.91	38.56	39.06	39.10	38.21	38.52
IMAGE	21			22			23			24		
METHODS	LRMC	[7]	Alg. 2									
Color Image	31.13	29.84	31.25	31.23	29.86	32.40	34.48	33.88	35.11	31.08	29.73	31.12
Depth Image	41.40	41.24	41.67	39.85	38.29	40.02	42.35	42.11	43.34	36.13	35.64	35.90
IMAGE	25			26			27			28		
METHODS	LRMC	[7]	Alg. 2									
Color Image	29.90	28.85	30.25	29.38	28.48	29.85	28.41	27.44	29.28	33.56	33.18	34.07
Depth Image	45.26	45.41	45.49	46.13	46.60	46.58	40.77	39.94	40.80	39.96	39.60	40.80
IMAGE	29			30			AVERAGE					
METHODS	LRMC	[7]	Alg. 2	LRMC	[7]	Alg. 2	LRMC		[7]	Alg. 2		
Color Image	33.70	32.83	34.11	27.65	26.53	28.26	30.47		29.63	31.07		
Depth Image	40.32	38.36	41.04	41.26	39.43	41.18	40.46		39.89	40.53		

5 Conclusion

In this paper, we present a method for the construction of a data-driven tight frame from an input multi-channel image. We construct one discrete tight frame system for each channel and assume their sparse coefficients having a common sparsity pattern. Similar to [7], this leads to an optimization problem involving a row-vector 0-norm and several orthonormal matrix constraints, which is solved by an iterative algorithm that performs thresholding and SVDs alternatively. The multi-channel data-driven tight frame is applied to joint color and depth image (RGB-D image) reconstruction. The experiments showed that the

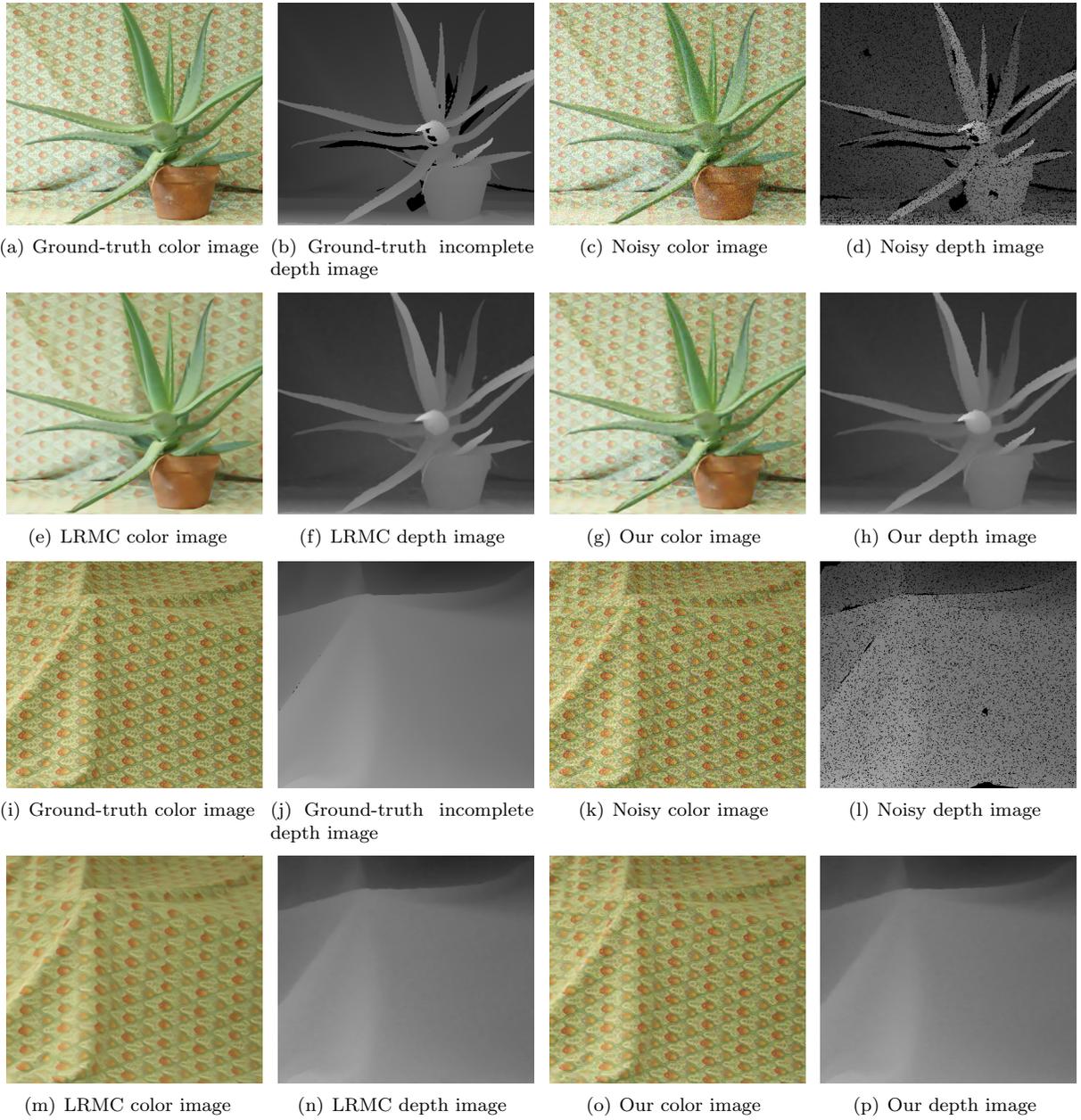


Figure 2: Visual quality comparison between Algorithm 2 and LRMC.



Figure 3: Visual quality comparison between Algorithm 2 and LRMC.

proposed approach has a better performance than state-of-the-art joint color and depth image reconstruction approaches in [24].

References

- [1] C. BAO, J.-F. CAI, AND H. JI, *Fast sparsity-based orthogonal dictionary learning for image restoration*, in 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, 2013, pp. 3384–3391.
- [2] C. BAO, H. JI, AND Z. SHEN, *Convergence analysis for iterative data-driven tight frame construction scheme*, Appl. Comput. Harmon. Anal., In Press (2014).
- [3] J.-F. CAI, R. H. CHAN, L. SHEN, AND Z. SHEN, *Restoration of chopped and noded images by framelets*, SIAM J. Sci. Comput., 30 (2008), pp. 1205–1227.
- [4] J.-F. CAI, R. H. CHAN, L. SHEN, AND Z. SHEN, *Convergence analysis of tight framelet approach for missing data recovery*, Adv. Comput. Math., 31 (2009), pp. 87–113.
- [5] J.-F. CAI, R. H. CHAN, AND Z. SHEN, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149.
- [6] J.-F. CAI, B. DONG, S. OSHER, AND Z. SHEN, *Image restoration: Total variation, wavelet frames, and beyond*, J. Amer. Math. Soc., 25 (2012), pp. 1033–1089.
- [7] J.-F. CAI, H. JI, Z. SHEN, AND G.-B. YE, *Data-driven tight frame construction and image denoising*, Appl. Comput. Harmon. Anal., 37 (2014), pp. 89–105.
- [8] J.-F. CAI, S. OSHER, AND Z. SHEN, *Linearized Bregman iterations for compressed sensing*, Math. Comp., 78 (2009), pp. 1515–1536.
- [9] J.-F. CAI, S. OSHER, AND Z. SHEN, *Linearized Bregman iterations for frame-based image deblurring*, SIAM J. Imaging Sci., 2 (2009), pp. 226–252.
- [10] E. J. CANDÈS AND D. L. DONOHO, *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities*, Comm. Pure Appl. Math., 57 (2004), pp. 219–266.
- [11] R. H. CHAN, T. F. CHAN, L. SHEN, AND Z. SHEN, *Wavelet algorithms for high-resolution image reconstruction*, SIAM J. Sci. Comput., 24 (2003), pp. 1408–1432 (electronic).
- [12] T. F. CHAN AND J. SHEN, *Image processing and analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. Variational, PDE, wavelet, and stochastic methods.
- [13] I. DAUBECHIES, *Ten lectures on wavelets*, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [14] I. DAUBECHIES, B. HAN, A. RON, AND Z. SHEN, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harmon. Anal., 14 (2003), pp. 1–46.
- [15] M. FORNASIER, R. RAMLAU, AND G. TESCHKE, *The application of joint sparsity and total variation minimization algorithms to a real-life art restoration problem*, Advances in Computational Mathematics, 31 (2009), pp. 157–184.
- [16] M. FORNASIER AND H. RAUHUT, *Recovery algorithms for vector-valued data with joint sparsity constraints*, SIAM Journal on Numerical Analysis, 46 (2008), pp. 577–613.
- [17] B. HAN, G. KUTYNIOK, AND Z. SHEN, *Adaptive multiresolution analysis structures and shearlet systems*, SIAM J. Numer. Anal., 49 (2011), pp. 1921–1946.

- [18] T.-W. HUI AND K. NGI NGAN, *Motion-depth: Rgb-d depth map enhancement with motion and depth in complement*, June 2014.
- [19] R. JAIN, R. KASTURI, AND B. G. SCHUNCK, *Machine vision*, vol. 5, McGraw-Hill New York, 1995.
- [20] C. KUSTER, T. POPA, C. ZACH, C. GOTSMAN, M. H. GROSS, P. EISERT, J. HORNEGGER, AND K. POLTHIER, *Freecam: A hybrid camera system for interactive free-viewpoint video.*, in Proceedings of VMV, 2011, pp. 17–24.
- [21] D. LABATE, W.-Q. LIM, G. KUTYNIOK, AND G. WEISS, *Sparse multidimensional representation using shearlets*, in Optics & Photonics 2005, International Society for Optics and Photonics, 2005, pp. 59140U–59140U.
- [22] E. LE PENNEC AND S. MALLAT, *Sparse geometric image representations with bandelets*, IEEE Trans. Image Process., 14 (2005), pp. 423–438.
- [23] J. LIANG, J. MA, AND X. ZHANG, *Seismic data restoration via data-driven tight frame*, Geophysics, 79 (2014), pp. V65–V74.
- [24] S. LU, X. REN, AND F. LIU, *Depth enhancement via low-rank matrix completion*, in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2014.
- [25] A. MAIMONE AND H. FUCHS, *Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras*, in Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on, IEEE, 2011, pp. 137–146.
- [26] S. MALLAT, *A wavelet tour of signal processing*, Elsevier/Academic Press, Amsterdam, third ed., 2009. The sparse way, With contributions from Gabriel Peyré.
- [27] C. RICHARDT, C. STOLL, N. A. DODGSON, H.-P. SEIDEL, AND C. THEOBALT, *Coherent spatio-temporal filtering, upsampling and rendering of RGBZ videos*, Computer Graphics Forum (Proceedings of Eurographics), 31 (2012).
- [28] A. RON AND Z. SHEN, *Affine systems in $L_2(\mathbf{R}^d)$: the analysis of the analysis operator*, J. Funct. Anal., 148 (1997), pp. 408–447.
- [29] D. SCHARSTEIN AND R. SZELISKI, *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*, International journal of computer vision, 47 (2002), pp. 7–42.
- [30] H. SEOK LEE AND K. MU LEE, *Simultaneous super-resolution of depth and images using a single camera*, in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2013.
- [31] S. SHEKHAR, V. M. PATEL, N. M. NASRABADI, AND R. CHELLAPPA, *Joint sparsity-based robust multimodal biometrics recognition*, in Computer Vision–ECCV 2012. Workshops and Demonstrations, Springer, 2012, pp. 365–374.
- [32] Z. SHEN, *Wavelet frames and image restorations*, in Proceedings of the International Congress of Mathematicians. Volume IV, New Delhi, 2010, Hindustan Book Agency, pp. 2834–2863.
- [33] X. YANG, S. KIM, AND E. P. XING, *Heterogeneous multitask learning with joint sparsity constraints*, in Advances in neural information processing systems (NIPS), 2009, pp. 2151–2159.
- [34] W. ZHOU, J.-F. CAI, AND H. GAO, *Adaptive tight frame based medical image reconstruction: a proof-of-concept study for computed tomography*, Inverse Problems, 29 (2013), pp. 125006, 18.
- [35] H. JI, Z. SHEN, AND Y. XU, *Wavelet frame based scene reconstruction from range data*, Journal of Computational Physics, 229 (2010), pp. 2093–2108, 6.