PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Affine shear tight frames with two-layer structure

Zhihua Che
Xiaosheng Zhuang

Affine Shear Tight Frames with Two-Layer Structure

Zhihua Che[a] and Xiaosheng Zhuang[a]

[a] Department of Mathematics, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong

ABSTRACT

Affine shear tight frames with 2-layer structure in arbitrary dimensions are constructed. Affine shear filter banks with 2-layer structure associated with affine shear tight frames with 2-layer structure are designed and show to be with the perfect reconstruction property. The redundancy rate and computational complexity of affine shear filter banks are discussed. Applications to video denoising are conducted to demonstrate the effectiveness of the affine shear filter banks with 2-layer structure.

Keywords: directional multiscale representation systems, affine shear tight frames, digital affine shear transforms, 2-layer filter banks, directional filter banks, image/video processing, high-dimensional data analysis, denoising/inpainting, tensor product complex tight framelets, shearlets, curvelets, contourlets, surfacelets

1. INTRODUCTION

Directional multiscale representation systems have been shown to be superb over many other multiscale representation systems in both theory (sparse approximation) and applications, e.g., see [2, 5, 8–11, 16, 20–23] and many references therein. Based on the framework of frequency-based affine systems [13], smooth affine shear tight frames, one of the recent developed directional multiscale representation systems, have been study systematically in dimension two [16] and in arbitrary dimension $d \geqslant 2$ [23]. In [16,23], it has been shown that an affine shear tight frames can be deduced from a directional affine wavelet tight frame via appropriate downsampling, and thus it can be regarded as a subsampled system from a directional affine wavelet tight frame thereby associating an affine shear tight frame with a underlying directional filter bank. More importantly, digital affine shear transforms can be efficiently implemented using their underlying filter banks and are very similar to the standard fast wavelet transforms. Such digital affine shear filter banks enjoy many nice properties including low-redundancy rate, arbitrary number of directional filters, shear structure, and so on. Applications of digital affine shear filter banks associated with affine shear tight frames in image/video processing demonstrate the advantages of such types of filter banks in comparison with many other frame-based directional filter banks [23]. On the other hand, in the recent papers [14,15], another type of directional systems, called tensor product complex tight framelets, has been successfully applied in image/vedio processing as well. Their associated filter banks (TP-$\mathbb{C}$TF$_6$ and TP-$\mathbb{C}$TF$_6^{\downarrow}$ filter banks) have many nice properties such as simple tensor product structure, easy frequency-domain design, and more importantly, 2-layer directional filter banks. The 2-layer structure brings many important features in a filter bank system. It results in the combinations of an inner layer of high-pass filters that are mainly for capturing edge-like features, and an outer layer of high-pass filters that are highly oscillating for texture-like features. However, due to the limitation of tensor-product structure, such type of filter banks can only have a fixed number of directional filters, which is not desirable in practice, especially when the resolutions of images getting higher and higher.

Motivated by the successful applications of the affine shear tight frames ([16,23]) and the tensor product complex tight framelets for image/video processing ([14, 15]), in [4], affine shear filter banks with 2-layer structure (DAS-2 filter bank) in dimension two are proposed and applied to the tasks of image denoising and image inpainting. Though the idea is simple, it brings significant performance improvement in image denoising/inpainting using the DAS-2 filter banks. In this paper, we further investigate d-dimensional affine shear tight frames with 2-layer structure in any dimension $d \geqslant 2$ and present the characterizations, construction, and applications of affine shear tight frames with 2-layer structure.

Send correspondence to Email: xzhuang7@city.edu.hk, Phone: +852 3442 5942. Further contact information: Zhihua Che, zhihuache2-c@my.cityu.edu.hk.

Wavelets and Sparsity XVII, edited by Yue M. Lu, Dimitri Van De Ville, Manos Papadakis, Proc. of SPIE Vol. 10394, 1039422 · © 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2272398

In the remaining part of this paper, we first introduce affine shear systems with 2-layer structure that have generators splitting the frequency region at each scale into inner and outer layers, where the inner layer generators are used to capture edge-like features while the outer layer texture-like features. After that, we provide the characterization and construction of a sequence of affine shear systems with 2-layer structure to be a sequence of affine shear tight frames with 2-layer structure, which naturally induce affine shear filter banks with 2-layer structure (DAS-2 filter banks). Last but not least, we show that digital affine shear transforms can be implemented with low redundancy rate and with near-linear computational complexity ($\mathcal{O}(N\log N)$). Numerical experiments are conducted to demonstrate the advantages of our digital affine shear filter banks with 2-layer structure over many other state-of-the-art frame-based methods in video processing.

2. AFFINE SHEAR TIGHT FRAMES WITH 2-LAYER STRUCTURE IN $\mathbb{R}^d$

In this section, we extend the definition of affine shear tight frames with 2-layer structure in [4] in dimension two to any dimension $d\geqslant 2$. We first introduce the notation of a sequence of d-dimensional affine shear systems with 2-layer structure and then provide the characterization for an affine shear system with 2-layer structure to be an affine shear tight frame with 2-layer structure in $L_2(\mathbb{R}^d)$ as well as the construction of d-dimensional affine shear tight frames with 2-layer structure.

2.1 Characterizaiton

Throughout the paper, we assume $d\geqslant 2$ and use the compact notation $f_{U;\mathsf{k},\mathsf{n}}(x):=|\det U|^{1/2}f(Ux-\mathsf{k})e^{-i\mathsf{n}\cdot Ux}$, $x\in\mathbb{R}^d$ to encode dilation by a $d\times d$ invertible matrix U, translation by $\mathsf{k}\in\mathbb{R}^d$, and modulation by $\mathsf{n}\in\mathbb{R}^d$ for a function f defined in $\mathbb{R}^d$. Such a notation is consistent with the classical wavelet notation $\psi_{j;k}:=2^{j/2}\psi(2^j\cdot-k)$. We denote the shear operator $S^{\vec{\tau}}$ with $\vec{\tau}=(\tau_2,\dots,\tau_d)\in\mathbb{R}^{d-1}$, anisotropic $d\times d$ dilation matrix A_λ, and isotropic $d\times d$ dilation matrix M_λ with $\lambda>1$ by

$$S^{\vec{\tau}}=\begin{bmatrix}1&\tau_2&\dots&\tau_d\\0&1&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&1\end{bmatrix},\quad \mathsf{A}_\lambda=\begin{bmatrix}\lambda^2&0&\dots&0\\0&\lambda&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&\lambda\end{bmatrix},\quad\text{and}\quad \mathsf{M}_\lambda=\begin{bmatrix}\lambda^2&0&\dots&0\\0&\lambda^2&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&\lambda^2\end{bmatrix}.$$

We use $\mathsf{N}_\lambda:=\mathsf{M}_\lambda^{-\mathsf{T}}$ and $\mathsf{B}_\lambda:=\mathsf{A}_\lambda^{-\mathsf{T}}$ to denote the transpose of the inverse of M_λ and A_λ, respectively. Note that $\mathsf{M}_\lambda=\mathsf{A}_\lambda\mathsf{D}_\lambda$ with $\mathsf{D}_\lambda:=\operatorname{diag}(1,\lambda\mathsf{I}_{d-1})$, where I_n denotes the $n\times n$ identity matrix. Define $S_{\vec{\tau}}:=(S^{\vec{\tau}})^{\mathsf{T}}$ and denote $d\times d$ matrix E_n to be the elementary matrix corresponding to the coordinate exchange between the first axis and the nth one. That is, $\mathsf{E}_1=\mathsf{I}_d$, $\mathsf{E}_2=\operatorname{diag}(\begin{bmatrix}0&1\\1&0\end{bmatrix},\mathsf{I}_{d-2})$, $\mathsf{E}_3=\operatorname{diag}(\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix},\mathsf{I}_{d-3})$, and so on.

Let $\varphi\in L_2(\mathbb{R}^d)$ be a scaling function for the low frequency part at scale j and Ψ_j^{in},Ψ_j^{out} be a set of generators for the high frequency part in the inner layer) and in the outer layer at scale j, respectively, be given by

$$\Psi_j^{in}:=\{\psi^{j,\vec{\ell},in}:|\vec{\ell}|\leqslant\vec{r}_{j,in}\},\quad \Psi_j^{out}:=\{\psi^{j,\vec{\ell},out}:|\vec{\ell}|\leqslant\vec{r}_{j,out}\}$$

with $\vec{\ell}:=(\ell_2,\dots,\ell_d)\in\mathbb{Z}^{d-1}$, $\vec{r}_{j,\iota}:=(r^\iota_{j,2},\dots,r^\iota_{j,d})\in\mathbb{Z}^{d-1}$, and $\psi^{j,\vec{\ell},\iota}$ being functions in $L_2(\mathbb{R}^d)$ for $\iota\in\{in,out\}$. Here and after we shall use the compact notation $\sum_{\vec{\ell}=-\vec{r}_{j,\iota}}^{\vec{r}_{j,\iota}}$ to mean $\sum_{\ell_2=-r^\iota_{j,2}}^{r^\iota_{j,2}}\cdots\sum_{\ell_d=-r^\iota_{j,d}}^{r^\iota_{j,d}}$ and $|\vec{\ell}|\leqslant\vec{r}_{j,\iota}$ to mean all $\vec{\ell}=(\ell_2,\dots,\ell_d)$ such that $|\ell_2|\leqslant r^\iota_{j,2}$, ..., $|\ell_d|\leqslant r^\iota_{j,d}$. The summation $\sum_n$ and $\sum_\iota$ are short for $\sum_{n=1}^d$ and $\sum_{\iota=in,out}$, respectively. A *d-dimensional affine shear system* (with 2-layer structure and starting at scale J) is then defined to be

$$\mathsf{AS}_J(\varphi;\{\Psi_j^{in},\Psi_j^{out}\}_{j=J}^\infty)=\{\varphi_{\mathsf{M}_{\lambda_J^{out}};\mathsf{k}}:\mathsf{k}\in\mathbb{Z}^d\}\cup\{\psi^{j,\vec{\ell},\iota}_{S^{-\vec{\ell}}\mathsf{A}_{\lambda_j^\iota}\mathsf{E}_n;\mathsf{k}}:\mathsf{k}\in\mathbb{Z}^d,n=1,\dots,d,\ |\vec{\ell}|\leqslant\vec{r}_{j,\iota},\iota=in,out\}_{j=J}^\infty. \tag{1}$$

The Fourier transform $\widehat{f}$ of a function $f\in L_1(\mathbb{R}^d)$ is defined to be $\widehat{f}(\xi):=\int_{\mathbb{R}^d}f(x)e^{-ix\cdot\xi}dx$ for $\xi\in\mathbb{R}^d$ and can be naturally extended to functions in $L_2(\mathbb{R}^d)$ or tempered distributions. Following the same lines of proof in Theorem 2 of [16]. We have the following simple characterization for a sequence of affine shear systems with 2-layer structure to be a sequence of affine shear tight frames with 2-layer structure for $L_2(\mathbb{R}^d)$ when all generators are nonnegative in the frequency domain (also see [13, Corollary 18]):

THEOREM 2.1. *Let J_0 be an integer and $\mathsf{AS}_J(\varphi;\{\Psi_j^{in},\Psi_j^{out}\}_{j=J}^\infty)$ be defined as in (1). Suppose that $\widehat{h}\geqslant 0$ for all $h\in\{\{\varphi\}\cup\Psi_j^{in}\cup\Psi_j^{out}\}_{j=J}^\infty$. Then, for all integers $J\geqslant J_0$, $\mathsf{AS}_J(\varphi;\{\Psi_j^{in},\Psi_j^{out}\}_{j=J}^\infty)$ is an affine shear tight frame for $L_2(\mathbb{R}^d)$; that is, all generators are from $L_2(\mathbb{R}^d)$ and*

$$\|f\|_2^2=\sum_{\mathsf{k}\in\mathbb{Z}^d}|\langle f,\varphi_{\mathsf{M}_{\lambda_j^{out}};\mathsf{k}}\rangle|^2+\sum_{j=J}^\infty\sum_{n,\iota}\sum_{\vec{\ell}=-\vec{r}_{j,\iota}}^{\vec{r}_{j,\iota}}\sum_{\mathsf{k}\in\mathbb{Z}^d}|\langle f,\psi^{j,\vec{\ell},\iota}_{S^{-\vec{\ell}}\mathsf{A}_{\lambda_j^\iota}\mathsf{E}_n;\mathsf{k}}\rangle|^2,\quad\forall f\in L_2(\mathbb{R}^d),\tag{2}$$

if and only if the following holds:

$$\begin{aligned}\widehat{h}(\xi)\widehat{h}(\xi+2\pi\mathsf{k})=&0,\quad a.e.,\xi\in\mathbb{R}^d,\ \mathsf{k}\in\mathbb{Z}^d\backslash\{0\},\ \forall h\in\{\{\varphi\}\cup\Psi_j^{in}\cup\Psi_j^{out}\}_{j=J}^\infty,\\ |\widehat{\varphi}(\mathsf{N}_{\lambda_{j+1}^{out}}\xi)|^2=&|\widehat{\varphi}(\mathsf{N}_{\lambda_j^{out}}\xi)|^2+\sum_{n,\iota}\sum_{\vec{\ell}=-\vec{r}_{j,\iota}}^{\vec{r}_{j,\iota}}|\widehat{\psi^{j,\vec{\ell},\iota}}(S_{\vec{\ell}}\mathsf{B}_{\lambda_j^\iota}\mathsf{E}_n\xi)|^2,\quad a.e.,\xi\in\mathbb{R}^d,\ j\geqslant J_0,\\ \lim_{j\to\infty}\langle|\widehat{\varphi}(\mathsf{N}_{\lambda_j^{out}}\cdot)|^2,\widehat{h}\rangle=&\langle 1,\widehat{h}\rangle\quad\forall\widehat{h}\in C_c^\infty(\mathbb{R}^d).\end{aligned}\tag{3}$$

2.2 Construction

Based on Theorem 2.1, we next provide the construction of affine shear tight frames with 2-layer structure in arbitrary dimension $d\geqslant 2$. We use the frequency domain approach to achieve (3), which heavily relies on a building block function $\boldsymbol{\nu}_{[c,\epsilon]}$.

Let ν be a function such that $\nu(x)=0$ for $x\leqslant -1$, $\nu(x)=1$ for $x\geqslant 1$, and $|\nu(x)|^2+|\nu(-x)|^2=1$ for all $x\in\mathbb{R}$. Such a function can be constructed to be smooth in $C^\infty(\mathbb{R})$ or differentiable in $C^k(\mathbb{R})$; see [16,23]. Define for $0<\epsilon\leqslant c$, the funciton $\boldsymbol{\nu}_{[c,\epsilon]}$ to be

$$\boldsymbol{\nu}_{[c,\epsilon]}(x):=\begin{cases}\nu(\frac{x+c}{\epsilon}) & \text{if } x<-c+\epsilon,\\ 1 & \text{if } -c+\epsilon\leqslant x\leqslant c-\epsilon,\\ \nu(\frac{-x+c}{\epsilon}) & \text{if } x>c-\epsilon.\end{cases}$$

Then the function $\boldsymbol{\nu}_{[c,\epsilon]}$ is a "bump" function supported on $[-c-\epsilon,c+\epsilon]$ and satisfies $\sum_{k\in\mathbb{Z}}|\boldsymbol{\nu}_{[c,\epsilon]}(\cdot-2kc)|^2\equiv 1$.

Define $\boldsymbol{\gamma}_\varepsilon:=\boldsymbol{\nu}_{[1/2,\varepsilon]}$ for $0<\varepsilon\leqslant 1/2$ to be the splitting function, and $\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi):=\boldsymbol{\nu}_{[\lambda^{-2}(1-t/2)\rho\pi,\lambda^{-2}t\rho\pi/2]}(\xi)$, $\boldsymbol{\beta}_{\lambda,t,\rho}(\xi):=(|\boldsymbol{\alpha}_{\lambda,t,\rho}(\lambda^{-2}\xi)|^2-|\boldsymbol{\alpha}_{\lambda,t,\rho}(\xi)|^2)^{1/2}$ to be the 1D Meyer-type scaling and wavelet functions with $\lambda>1$, $0<t\leqslant 1$, and $0<\rho\leqslant\lambda^2$. Then $\operatorname{supp}\boldsymbol{\gamma}_\varepsilon=[-1/2-\varepsilon,1/2+\varepsilon]$, $\operatorname{supp}\boldsymbol{\alpha}_{\lambda,t,\rho}=[-\lambda^{-2}\rho\pi,\lambda^{-2}\rho\pi]$ and $\operatorname{supp}\boldsymbol{\beta}_{\lambda,t,\rho}=[-\rho\pi,-\lambda^{-2}(1-t)\rho\pi]\cup[\lambda^{-2}(1-t)\rho\pi,\rho\pi]$.

The functions $\boldsymbol{\alpha}_{\lambda,t,\rho}$ and $\boldsymbol{\beta}_{\lambda,t,\rho}$ are used for the ξ_1-axis while the function $\boldsymbol{\gamma}_\varepsilon$ is for splitting pieces along the other axes. Roughly speaking, the core generator for our affine shear systems in the frequency domain looks like $\boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\prod_{n=2}^d\boldsymbol{\gamma}_\varepsilon(\xi_n/\xi_1)$, which is a pyramid shape generator. Application of parabolic scaling, shear, and translation operations to such a generator induces our affine shear systems. Further technical treatments are then applied on such systems to achieve tightness.

For $\lambda>1$, define $\ell_\lambda:=\lfloor\lambda-(1/2+\varepsilon)\rfloor+1=\lfloor\lambda+(1/2-\varepsilon)\rfloor$, $\lambda_j^{out}:=\lambda^j$, and $\lambda_j^{in}:=\lambda^{j-1/2}$. We next define $\boldsymbol{\Gamma}_j^\iota$ for $\iota=in,out$, which will be used for normalization of frequency splitting along the shear directions. Define d-dimensional splitting function $\boldsymbol{\gamma}(\xi):=\prod_{n=2}^d\boldsymbol{\gamma}_\varepsilon(\xi_n/\xi_1)$ for $\xi=(\xi_1,\ldots,\xi_d)\in\mathbb{R}^d$,

$$\boldsymbol{\gamma}^{j,\vec{\ell},\iota}(\xi):=\boldsymbol{\gamma}(S_{\vec{\ell}}\mathsf{B}_{\lambda_j^\iota}\xi)=\prod_{n=2}^d\boldsymbol{\gamma}_\varepsilon(\lambda_j^\iota\xi_n/\xi_1+\ell_n)\quad\text{and}\quad\boldsymbol{\Gamma}_j^\iota(\xi):=\sum_{n=1}^d\sum_{\vec{\ell}=-\vec{\ell}_{\lambda_j^\iota}}^{\vec{\ell}_{\lambda_j^\iota}}\boldsymbol{\gamma}^{j,\vec{\ell},\iota}(\mathsf{E}_n\xi),\tag{4}$$

where $\vec{\ell}:=(\ell_2,\ldots,\ell_d)\in\mathbb{Z}^{d-1}$ and $\vec{\ell}_{\lambda_j^\iota}:=(\ell_{\lambda_j^\iota},\ldots,\ell_{\lambda_j^\iota})\in\mathbb{Z}^{d-1}$. Then $\boldsymbol{\Gamma}_j^\iota$ has the following properties:

(i) $0 < \boldsymbol{\Gamma}_j^\iota(\xi) \leqslant 2$, $\boldsymbol{\Gamma}_j^\iota(\mathsf{E}_n\xi) = \boldsymbol{\Gamma}_j^\iota(\xi)$ for all $n = 1, \dots, d$, and $\boldsymbol{\Gamma}_j^\iota(t\xi) = \boldsymbol{\Gamma}_j(\xi)$ for all $t \neq 0$ and $\xi \neq 0$.

(ii) $\boldsymbol{\Gamma}_j^\iota(\xi) \equiv 1$ for $\xi \in \left\{\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d : \max\{|\xi_m/\xi_n| : m \neq n; m, n = 1, \dots, d\} \leqslant \frac{\lambda_j^\iota}{\ell_{\lambda_j^\iota} + 1/2 + \varepsilon}\right\}$.

Next, we define the d-dimensional scaling function φ and generators $\psi^{j,\vec{\ell},\iota}$. Let

$$\begin{aligned} \widehat{\varphi}(\xi) :=& [\otimes \boldsymbol{\alpha}_{\lambda,t,\rho}](\xi) = \prod_{n=1}^{d} \boldsymbol{\alpha}_{\lambda,t,\rho}(\xi_n), \\ \boldsymbol{\omega}^{out}(\xi) :=& \sqrt{|\widehat{\varphi}(\lambda^{-2}\xi)|^2 - |\widehat{\varphi}(\lambda^{-1}\xi)|^2}, \\ \boldsymbol{\omega}^{in}(\xi) :=& \sqrt{|\widehat{\varphi}(\lambda^{-1}\xi)|^2 - |\widehat{\varphi}(\xi)|^2}. \end{aligned}$$

Note that for simplicity of presentation, we omit the dependency of $\varphi, \psi^{j,\vec{\ell}}, \boldsymbol{\gamma}^{j,\vec{\ell},\iota}, \widehat{a}, \widehat{b}, \boldsymbol{\Gamma}_j^\iota$, etc., on the parameters $\lambda, t, \rho, \varepsilon$. Now define $\psi^{j,\vec{\ell},\iota}$ by

$$\widehat{\psi^{j,\vec{\ell},\iota}}(\xi) := \boldsymbol{\omega}^\iota(\lambda^{-2j}(S_{\vec{\ell}}\,\mathsf{B}_{\lambda_j^\iota})^{-1}\xi) \frac{\Pi_{n=2}^d \boldsymbol{\gamma}_\varepsilon(\xi_n/\xi_1)}{\sqrt{\boldsymbol{\Gamma}_j^\iota((S_{\vec{\ell}}\,\mathsf{B}_{\lambda_j^\iota})^{-1}\xi)}}, \quad \xi \in \mathbb{R}^d\backslash\{0\},\ \iota \in \{in, out\}. \tag{5}$$

and $\widehat{\psi^{j,\vec{\ell},\iota}}(0) := 0$, which gives $\widehat{\psi^{j,\vec{\ell},\iota}}(S_{\vec{\ell}}\,\mathsf{B}_{\lambda_j^\iota}\xi) = \boldsymbol{\omega}^\iota(\lambda^{-2j}\xi)\frac{\boldsymbol{\gamma}^{j,\vec{\ell},\iota}(\xi)}{\sqrt{\boldsymbol{\Gamma}_j^\iota(\xi)}}$. We have the following (quasi-stationary) *d-dimensional affine shear system with 2-layer structure*:

$$\mathsf{AS}_J(\varphi; \{\Psi_j^{in}, \Psi_j^{out}\}_{j=J}^\infty) := \{\varphi_{\mathsf{M}_{\lambda_J^{out}};\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^d\} \cup \{\psi^{j,\vec{\ell},\iota}_{S^{-\vec{\ell}}\mathsf{A}_{\lambda_j^\iota}\mathsf{E}_n;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^d, n = 1, \dots, d, |\vec{\ell}| \leqslant \vec{\ell}_{\lambda_j^\iota}, \iota = in, out\}_{j=J}^\infty. \tag{6}$$

The affine system with 2-layer structure define above is indeed an affine shear tight frame with 2-layer structure:

THEOREM 2.2. *Let $\lambda > 1$, $0 < t \leqslant 1$, $0 < \varepsilon \leqslant 1/2$, and $0 < \rho \leqslant \frac{1}{1+2\varepsilon}$. Let $\mathsf{AS}_J(\varphi; \{\Psi_j^{in}, \Psi_j^{out}\}_{j=J}^\infty)$ be defined as in (6) with $\widehat{\varphi} = \otimes^d \boldsymbol{\alpha}_{\lambda,t,\rho}$ and $\psi^{j,\vec{\ell},\iota}$ being given by (5). Then $\mathsf{AS}_J(\varphi; \{\Psi_j^{in}, \Psi_j^{out}\}_{j=J}^\infty)$ is an affine shear tight frame with 2-layer structure for $L_2(\mathbb{R}^d)$ for all $J \geqslant 0$. Moreover, $\mathsf{AS}_J(\varphi; \{\Psi_j^{in}, \Psi_j^{out}\}_{j=J}^\infty)$ contains a subsystem generated by a single generator ψ; that is*

$$\{\psi_{S^{-\vec{\ell}}\mathsf{A}_{\lambda_j^\iota}\mathsf{E}_n;\mathsf{k}} : \mathsf{k} \in \mathbb{Z}^d, n = 1, \dots, d, |\vec{\ell}| \leqslant \vec{r}_j, \iota = in, out\}_{j=J}^\infty \subseteq \mathsf{AS}_J(\varphi; \{\Psi_j^{in}, \Psi_j^{out}\}_{j=J}^\infty),$$

where $\vec{r}_j := (r_j, \dots, r_j) \in \mathbb{Z}^{d-1}$ with $r_j := \lfloor \lambda^{j-2}(1-t)\rho - (1/2+\varepsilon) \rfloor$ and $\widehat{\psi}(\xi) := \boldsymbol{\beta}_{\lambda,t,\rho}(\xi_1)\prod_{n=2}^d \boldsymbol{\gamma}_\varepsilon(\xi_n/\xi_1)$, $\xi \in \mathbb{R}^d$.

3. DIGITAL AFFINE SHEAR TRANSFORMS

As demonstrated in [16,23], an affine shear tight frame is associated with a underlying affine shear filter bank. In this section, we briefly discuss the construction of digital affine shear filter banks with 2-layer structure (DAS-2 filter banks) and the implementation of digital affine shear transforms based on the DAS-2 filter banks.

We can define inner, middle, outer functions $\widehat{a}, \widehat{a}_1, \widehat{a}_2 \in C(\mathbb{R}^d)$ by

$$\widehat{a}(\xi) := [\otimes^d \boldsymbol{\nu}_{[c_0,\epsilon_0]}](\xi), \quad \widehat{a_1}(\xi) := [\otimes^d \boldsymbol{\nu}_{[c_1,\epsilon_1]}](\xi), \quad \widehat{a_2}(\xi) := [\otimes^d \boldsymbol{\nu}_{[c_2,\epsilon_2]}](\xi), \quad \xi \in \mathbb{R}^d \tag{7}$$

for some parameters $0 < c_0 < c_1 < c_2 = \pi$ and $\varepsilon, \epsilon_0, \epsilon_1, \epsilon_2 > 0$ satisfying $c_0 + \epsilon_0 \leqslant \pi/2$ (for downsampling by 2), $(c_1 + \epsilon_1) - \frac{c_0 - \epsilon_0}{3/2+\varepsilon} \leqslant \pi/2$ (for downsampling by 4), and $(c_2 + \epsilon_2) - \frac{c_1 - \epsilon_1}{3/2+\varepsilon} \leqslant \pi/2$ (for downsampling by 4). Here, the parameter ε is the parameter in γ_ε. We identify the function $\widehat{a}$ as a function in $C(\mathbb{T}^2)$ in frequency domain. In the time domain, it then serves as our low-pass filter a. The filter $\widehat{a}$ is supported inside $[-\pi/2, \pi/2]^d$. The other two functions $\widehat{a}_1, \widehat{a}_2$ are auxiliary functions for building the inner and outer layer high-pass filters. Define the inner and outer functions b^{in}, b^{out} by

$$\widehat{b^{out}}(\xi) := \sqrt{|\widehat{a}_2(\xi)|^2 - |\widehat{a}_1(\xi)|^2}, \quad \widehat{b^{in}}(\xi) \quad := \sqrt{|\widehat{a}_1(\xi)|^2 - |\widehat{a}(\xi)|^2}. \tag{8}$$

Now, we apply the splitting technique to $\widehat{b^\iota}$ for the construction of high-pass filters $b^{j,\vec{\ell},\iota}$. At scale $j \geqslant 0$ and a nonnegative integers $k_j^\iota \in \mathbb{N}_0$, define $\vec{r}_{j,\iota} := (2^{k_j^\iota}, \dots, 2^{k_j^\iota}) \in \mathbb{Z}^{d-1}$. The number k_j^ι controls the total number of shear directions at scale j. Similar to the definition of normalization function $\mathbf{\Gamma}_j^\iota$ in (4), we define

$$\mathbf{\Gamma}_{k_j^\iota}(\xi) = \sum_{n=1}^{d} \sum_{\vec{\ell}=-\vec{r}_{j,\iota}}^{\vec{r}_{j,\iota}} |\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell}}(\mathsf{E}_n\xi)|^2, \quad \xi \neq 0 \quad \text{and} \quad \mathbf{\Gamma}_{k_j^\iota}(0) := 0, \tag{9}$$

where $\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell}}(\xi) := \prod_{n=2}^d \gamma_\varepsilon(2^{k_j^\iota}\xi_n/\xi_1 + \ell_n)$. To guarantee smoothness of boundary, we need to further split $\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell}}(\xi)$ to positive part and negative part of ξ_1-axis. Define

$$\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell},+}(\xi) := \boldsymbol{\gamma}^{k_j^\iota,\vec{\ell}}(\xi)\chi_{\{\xi_1>0\}} \quad \text{and} \quad \boldsymbol{\gamma}^{k_j^\iota,\vec{\ell},-}(\xi) := \boldsymbol{\gamma}^{k_j^\iota,\vec{\ell}}(\xi)\chi_{\{\xi_1<0\}}.$$

Note that $\widehat{b^\iota}(\xi)\frac{\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell},\pm}(\xi)}{\sqrt{\mathbf{\Gamma}_{k_j^\iota}(\xi)}}$ are not $2\pi\mathbb{Z}^d$-periodic functions. We define $b^{j,\vec{\ell},\iota,\pm}$ to be the $2\pi\mathbb{Z}^d$-periodization of $\widehat{b^\iota}(\xi)\frac{\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell},\pm}(\xi))}{\sqrt{\mathbf{\Gamma}_{k_j^\iota}(\xi)}}$ as follows.

$$\widehat{b^{j,\vec{\ell},\iota,\pm}}(\xi) := \sum_{\mathsf{k}\in\mathbb{Z}^d} \widehat{b^\iota}(\xi + 2\pi\mathsf{k})\frac{\boldsymbol{\gamma}^{k_j^\iota,\vec{\ell},\pm}(\xi + 2\pi\mathsf{k})}{\sqrt{\mathbf{\Gamma}_{k_j^\iota}(\xi + 2\pi\mathsf{k})}}, \quad \xi \in \mathbb{T}^d. \tag{10}$$

The total number of high-pass filters $b^{j,\vec{\ell},\iota,+}$ and $b^{j,\vec{\ell},\iota,-}$ at this scale j for $|\vec{\ell}| \leqslant 2^{k_j^\iota}$ is $2(2^{k_j^\iota+1}+1)^{d-1}$. Each filter of $\widehat{b^{j,\vec{\ell},\iota,\pm}}$ is $2\pi\mathbb{Z}^d$-periodic function on $\mathbb{T}^d$. The set of filters $\{b^{j,\vec{\ell},\iota,\tau} : |\vec{\ell}| \leqslant 2^{k_j^\iota}, \iota \in \{in, out\}, \tau \in \{+,-\}\}$ consists high-pass filters for the ξ_1-cone $\mathcal{C}_1 := \{\xi \in \mathbb{R}^d\backslash\{0\} : |\xi_n/\xi_1| \leqslant 1, n = 2, \dots, d\}$ along the ξ_1-axis. The cone $\mathcal{C}_k$ along the ξ_k-axis is then given by $\mathcal{C}_k := \{\xi \in \mathbb{R}^d\backslash\{0\} : |\xi_n/\xi_k| \leqslant 1, n = 1, \dots, d, n \neq k\}$. The high pass filters $b_n^{j,\vec{\ell},\iota,\tau} := b^{j,\vec{\ell},\iota,\tau}(\mathsf{E}_n\cdot)$ are then the high-pass filters for the cone $\mathcal{C}_n$. The *digital affine shear filter bank with 2-layer structure* (DAS-2 filter bank) at scale j is then defined to be

$$\{a; b_n^{j,\vec{\ell},\iota,\tau} : |\vec{\ell}| \leqslant 2^{k_j^\iota}, \iota \in \{in, out\}, \tau \in \{+,-\}, n = 1, \dots, d\}. \tag{11}$$

We have the following result.

THEOREM 3.1. *Retaining notations in this section and assuming $c_k, \epsilon_k, \varepsilon$ for $k = 0, 1, 2$ satisfying $0 < c_0 < c_1 < c_2 = \pi$, $c_0 + \epsilon_0 \leqslant \pi/2$, $(c_1 + \epsilon_1) - \frac{c_0-\epsilon_0}{3/2+\varepsilon} \leqslant \pi/2$, $(c_2 + \epsilon_2) - \frac{c_1-\epsilon_1}{3/2+\varepsilon} \leqslant \pi/2$, and $0 < \varepsilon \leqslant \frac{\pi}{c_2+\epsilon_2} - \frac{1}{2}$, then the filter bank defined in* (11) *forms a DAS-2 filter bank with the perfect reconstruction (PR) property as follows:*

$$|\widehat{a}(\xi)|^2 + \sum_{\tau,\iota,n} \sum_{\vec{\ell}=-\vec{r}_{j,\iota}}^{\vec{r}_{j,\iota}} \left|\widehat{b_n^{j,\vec{\ell},\iota,\tau}}(\xi)\right|^2 = 1, \tag{12}$$

$$\widehat{a}(\xi)\widehat{a}(\xi + 2\pi\omega) = 0, \tag{13}$$

$$\widehat{b_n^{j,\vec{\ell},\iota,\tau}}(\xi)\widehat{b_n^{j,\vec{\ell},\iota,\tau}}(\xi + 2\pi\omega_n) = 0, \tag{14}$$

for all $\xi \in \mathbb{T}^d$, $|\vec{\ell}| \leqslant (2^{k_j^\iota}, \dots, 2^{k_j^\iota})$, $\omega \in [\frac{1}{2}\mathbb{Z}^d] \cap [0,1)^d\backslash\{0\}$, and $\omega_n \in [(\mathsf{A}_n^{j,\iota})^{-\mathsf{T}}\mathbb{Z}^d] \cap [0,1)^d\backslash\{0\}$, where $\iota \in \{in, out\}$, $\tau \in \{+,-\}$, $n = 1, \dots, d$, and the $d \times d$ diagonal matrix $\mathsf{A}_n^{j,\iota} := \mathsf{E}_n \operatorname{diag}(4, 2^{k_j^\iota}, \dots, 2^{k_j^\iota})\mathsf{E}_n$.

Given a sequence of nonnegative integers $k_j^\iota : j = 0,\dots,J-1$ for some fixed integer $J \geqslant 0$ with respect to the finest scale. Let $\mathsf{M} := 2\mathsf{I}_d$ and $\mathsf{A}_n^{j,\iota} := \mathsf{E}_n \operatorname{diag}(4, 2^{k_j^\iota}\mathsf{I}_{d-1})\mathsf{E}_n$ for $n = 1,\dots,d$. We can then obtain a sequence of PR filter banks

$$\{a;\mathcal{B}_j^{in},\mathcal{B}_j^{out}\} := \{a \downarrow \mathsf{M}, b_n^{j,\vec{\ell},\iota,\pm} \downarrow \mathsf{A}_n^{j,\iota} : |\vec{\ell}| \leqslant \vec{r}_{j,\iota}, n = 1,\dots,d, \iota = in, out\} \tag{15}$$

for $j = 0,\dots,J-1$. Here M in $a \downarrow \mathsf{M}$ indicates downsampling matrix for filtered coefficients with respect to the low-pass filter a and $\mathsf{A}_n^{j,\iota}$ in $b_n^{j,\vec{\ell},\iota,\pm} \downarrow \mathsf{A}_n^{j,\iota}$ indicates downsampling matrix for filtered coefficients with respect to the high-pass filter $b^{j,\vec{\ell}}(\mathsf{E}_n\cdot)$. We call such a sequence of PR filter banks as a sequence of *d-dimensional digital affine shear filter banks with 2-layer structure* (DAS-2 filter banks) and denote it as $\mathcal{DAS}_J(\{a;\mathcal{B}_j^{in},\mathcal{B}_j^{out}\}_{j=0}^{J-1})$. Note that the total number of high-pass filters in $\{\mathcal{B}_j^{in},\mathcal{B}_j^{out}\}$ is $2d(2^{k_j^{in}+1}+1)^{d-1} + 2d(2^{k_j^{out}+1}+1)^{d-1}$.

Using such a sequence of DAS-2 filter banks, we can perform decomposition (forward transform) of d-dimensional data to a sequence of filtered coefficients as well as reconstruction (backward transform) of the d-dimensional data from the filtered coefficients. The forward and backward transform algorithms are depicted in Algorithms 1 and 2.

Algorithm 1. Forward Digital Affine Shear Transform

Input: Data v^J and DAS-2 filter banks $\mathcal{DAS}_J\left(\{a;\mathcal{B}_j^{in},\mathcal{B}_j^{out}\}_{j=0}^{J-1}\right)$.
Output: filtered coefficients $\{v^0\} \cup \{w_n^{j,\vec{\ell},\iota,\tau} : |\vec{\ell}| \leqslant \vec{r}_{j,\iota}, \iota, \tau, n\}_{j=0}^{J-1}$
Main steps:
for $j = J-1$ **to** 0 **do**
    $v^j \leftarrow [v^{j+1} \circledast a] \downarrow \mathsf{M}$.
    for each $b_n^{j,\vec{\ell},\iota,\tau}$ in $\mathcal{B}_j^{in} \cup \mathcal{B}_j^{out}$ **do**
        $w_n^{j,\vec{\ell},\iota,\tau} \leftarrow [v^{j+1} \circledast b_n^{j,\vec{\ell},\iota,\tau}] \downarrow \mathsf{A}_n^{j,\iota}$.
    end for
end for

Algorithm 2. Backward Digital Affine Shear Transform

Input: Coefficients $\{v^0\} \cup \{w_n^{j,\vec{\ell},\iota,\tau} : |\vec{\ell}| \leqslant \vec{r}_{j,\iota}, \iota, \tau, n\}_{j=0}^{J-1}$ and DAS-2 filter banks $\mathcal{DAS}_J\left(\{a;\mathcal{B}_j^{in},\mathcal{B}_j^{out}\}_{j=0}^{J-1}\right)$.
Output: Data v^J.
Main steps:
for $j = 0$ **to** $J-1$ **do**
    $v^{j+1} \leftarrow [v^j \uparrow \mathsf{M}] \circledast a^\star$.
    for each $b_n^{j,\vec{\ell},\iota,\tau}$ in $\mathcal{B}_j^{in} \cup \mathcal{B}_j^{out}$ **do**
        $v^{j+1} \leftarrow v^{j+1} + [w_n^{j,\vec{\ell},\iota,\tau} \uparrow \mathsf{A}_n^{j,\iota}] \circledast (b_n^{j,\vec{\ell},\iota,\tau})^\star$.
    end for
end for

The redundancy rate measures the storage complexity of a filter bank transform, which is usually given by the ratio of size of the output coefficients and the size of the input data. Similar to [23], we can obtain the redundancy rate r of the digital affine shear transform based on a sequence of DAS-2 filter banks, which is given by

$$r = \left(\sum_{j=0}^{J-1}\sum_{\iota=in,out}\frac{d(2^{-k^\iota_{J-1-j}}+2)^{d-1}}{2^{dj+1}} + \frac{1}{2^{dJ}}\right) \leqslant d(2^{-k_{min}}+2)^{d-1}\frac{2^d}{2^d-1}.$$

where $k_{min} := \min\{k_j^\iota : j = 0,\dots,J-1, \iota = in, out\}$. Moreover, the implementation of the forward and backward digital transforms can be based on the FFT algorithm, which results in the computational complexity proportional to $\mathcal{O}(N\log N)$ for N the size of input data.

4. NUMERICAL EXPERIMENTS ON VIDEO DENOISING

In [4], we have demonstrated the applications of digital affine shear filter banks with 2-layer structure in the tasks of image processing. In this section, we apply our high-dimensional DAS-2 filter banks in the tasks of video denoising. We compare the performance of our systems to several other stat-of-art directional multiscale representation systems. The usual peak-signal-noise-ratio (PSNR) index is used to measure the performance of different systems, which is defined to be $\mathrm{PSNR}(u, \widetilde{u}) = 10 \log_{10} \frac{255^2}{\mathrm{MSE}(u,\widetilde{u})}$, where $u : \Lambda \to \mathbb{C}$ is the original data defined on a lattice Λ, $\widetilde{u}$ is the denoised data of u, and $\mathrm{MSE}(u, \widetilde{u})$ is the mean square error $\frac{1}{|\Lambda|} \sum_{\mathsf{k} \in \Lambda} |u(\mathsf{k}) - \widetilde{u}(\mathsf{k})|^2$ with $|\Lambda|$ the cardinality of the lattice Λ. The unit of PSNR is dB. The local-soft (LS) thresholding technique is employed in the processing of filtered coefficients (see [23]) for more details).

The parameters $c_k, \epsilon_k, \epsilon_k$ of a, a_1, a_2 for $k = 0, 1, 2$ are given by $c_0 = 0.879$, $c_1 = 1.98$, $c_2 = \pi$ $\epsilon_0 = 0.268$, $\epsilon_1 = 0.14$, $\epsilon_2 = 0.1$ and $\varepsilon = 0.2$ for both inner layer and outer layer. We choose $J = 4$ for $\mathcal{DAS}_J(\{a; \mathcal{B}_j^{in}, \mathcal{B}_j^{out}\}_{j=0}^{J-1})$ as in (15); that is, we decompose to 4 scales. The shear parameters (k_0, k_1, k_2, k_3) is set to be $(1, 1, 1, 1)$ for inner layer and $(2, 2, 1, 1)$ for outer layer. That is, for the finest scale $j = 0$, we use totally $3((2^{2+1} + 1)^2 + (2^{1+1} + 1)^2) = 318$ shear directions (106 for each cone). The redundancy of our system $\mathcal{DAS}_J(\{a; \mathcal{B}_j^{in}, \mathcal{B}_j^{out}\}_{j=0}^{J-1})$ is 11.38. The convolution window size to compute local coefficient variance is set to be 9, i.e., we are using $9 \times 9 \times 9$ window.

We test three videos: `Mobile` and `Coastguard` which can be downloaded from `http://www.shearlab.org`. Both videos are of size $192 \times 192 \times 192$. We first employ symmetric boundary extension (with 32 pixels) on the noisy image to avoid boundary effect. We then apply our forward transform to obtained the coefficients. After performing the local-soft threshold procedure, we then apply the backward transform to the thresholded coefficients and throw away the extended boundary to obtained the final denoised image.

We compare our denoising performance to DAS-1 filter banks in [23], 3D dual-tree complex wavelets [21], 3D tensor product complex tight framelets [14, 15] , surfacelets [20], 3D DNST in [19]. The DAS-1 filter banks in [23] has redundancy rate 17.88. The 3D dual-tree complex wavelet transform (DT-$\mathbb{C}$WT) in [21] has redundancy rate 8. The number of directional filters of DT-$\mathbb{C}$WT at each scale is 56. The number of scales is 5. The 3D TP-$\mathbb{C}\mathrm{TF}_6$ and TP-$\mathbb{C}\mathrm{TF}_6^{\downarrow}$ are detailed in [15], which have redundancy rate 29.71 and 3.71, respectively. The number of scales is 4. Bivariate shrinkage thresholding technique is employed for DT-$\mathbb{C}$WT, TP-$\mathbb{C}\mathrm{TF}_6$, and TP-$\mathbb{C}\mathrm{TF}_6^{\downarrow}$. For 3D DNST from the ShearLab package, we choose the one with redundancy rate 154 (3 scales). The surfacelet transform (SURF) from SurfBox at `http://minhdo.ece.illinois.edu/software` has redundancy rate 6.4. The 3D DNST and sufacelet transform use hard thresholding for denoising.

We compare the denoising performance over different noise level $\sigma \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$. The comparison results are presented in Table 1. The numbers inside brackets are the difference of between the compared method and our method in terms of PSNR. From Table 1, we see that our method outperforms all other methods except a few case for DNST.

	$192 \times 192 \times 192$ Mobile						
σ	DAS-2 (11.38)	DAS-1 (17.88)	DT-$\mathbb{C}$WT (8)	TP-$\mathbb{C}\mathrm{TF}_6^{\downarrow}$ (3.71)	TP-$\mathbb{C}\mathrm{TF}_6$ (29.71)	SURF (6.4)	DNST (154)
10	35.69	34.99(0.70)	34.72(0.98)	35.15(0.55)	35.52(0.18)	32.79(2.91)	35.91(-0.21)
20	31.99	31.50(0.49)	30.86(1.13)	31.48(0.51)	31.77(0.22)	29.95(2.04)	32.18(-0.20)
30	29.90	29.57(0.33)	28.67(1.23)	29.44(0.46)	29.66(0.24)	28.26(1.64)	29.99(-0.09)
40	28.44	28.26(0.18)	27.14(1.30)	28.05(0.39)	28.20(0.25)	27.05(1.39)	28.42(0.02)
50	27.33	27.26(0.07)	26.06(1.27)	26.99(0.34)	27.08(0.25)	26.11(1.22)	27.22(0.11)
60	26.46	26.45(0.00)	25.21(1.25)	26.14(0.32)	26.18(0.27)	25.38(1.08)	26.25(0.21)
70	25.74	25.78(-0.04)	24.55(1.19)	25.44(0.30)	25.45(0.29)	24.77(0.96)	25.44(0.30)
80	25.14	25.20(-0.06)	24.00(1.14)	24.84(0.30)	24.82(0.31)	24.25(0.89)	24.75(0.39)
90	24.60	24.68(-0.09)	23.57(1.03)	24.33(0.26)	24.29(0.31)	23.80(0.80)	24.15(0.45)
100	24.15	24.23(-0.07)	23.17(0.98)	23.89(0.26)	23.82(0.33)	23.40(0.75)	23.62(0.54)
	$192 \times 192 \times 192$ Coastguard						
σ	DAS-2	DAS-1	DT-$\mathbb{C}$WT	TP-$\mathbb{C}\mathrm{TF}_6^{\downarrow}$	TP-$\mathbb{C}\mathrm{TF}_6$	SURF	DNST
10	34.18	33.70(0.48)	33.21(0.98)	33.86(0.32)	34.15(0.04)	30.86(3.32)	33.81(0.37)
20	30.71	30.27(0.44)	29.61(1.10)	30.26(0.44)	30.62(0.09)	28.26(2.45)	30.28(0.43)
30	28.85	28.47(0.38)	27.71(1.13)	28.39(0.46)	28.73(0.12)	26.87(1.98)	28.40(0.45)
40	27.60	27.27(0.33)	26.47(1.13)	27.14(0.46)	27.45(0.15)	25.91(1.69)	27.13(0.47)
50	26.66	26.40(0.27)	25.56(1.11)	26.21(0.45)	26.48(0.18)	25.17(1.49)	26.17(0.49)
60	25.91	25.70(0.21)	24.86(1.06)	25.47(0.44)	25.71(0.20)	24.57(1.34)	25.39(0.52)
70	25.28	25.14(0.14)	24.29(0.99)	24.86(0.42)	25.07(0.21)	24.06(1.22)	24.74(0.54)
80	24.76	24.65(0.10)	23.83(0.93)	24.34(0.41)	24.53(0.23)	23.61(1.15)	24.17(0.59)
90	24.31	24.23(0.07)	23.41(0.90)	23.90(0.41)	24.06(0.24)	23.22(1.09)	23.67(0.63)
100	23.88	23.86(0.02)	23.08(0.80)	23.51(0.37)	23.65(0.23)	22.87(1.01)	23.22(0.66)

Table 1. PSNR of denoised `Mobile` and `Coastguard`.

ACKNOWLEDGMENTS

The research of this paper was supported in part by the Research Grants Council of Hong Kong (Project No. CityU 11304414) and City University of Hong Kong (Project No.: 7200462 and 7004445).

REFERENCES

[1] B. G. Bodmann, G. Kutyniok, and X. Zhuang, Gabor shearlets, *Appl. Comput. Harmon. Anal.* **38** (1) (2015), 87–114.

[2] 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** (2) (2004), 219–266.

[3] E. J. Candès, L. Demannet, D. Donoho, and L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.* **5** (3) (2006), 861–899.

[4] Z. Che and X. Zhuang, digital affine shear filter banks with 2-layer structure and their applications in image Processing, submitted, 2017.

[5] C. K. Chui, H. N. Mhaskar, and X. Zhuang, Representation of functions on big data associated with directed graphs. *Appl. Comput. Harmon. Anal.*, doi:10.1016/j.acha.2016.12.005.

[6] A. L. Cunha, J. Zhou, and M. N. Do, The nonsubsampled contourlet transform: theory, design, and applications, *IEEE Trans. Image Proc.* **15** (10) (2006), 3089–3101.

[7] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, **61**, SIAM, Philadelphia, PA, 1992.

[8] K. Guo, D. Labate, W. Lim, G. Weiss, and E. Wilson, Wavelets with composite dilations, *Electr. Res. Ann. AMS.* **10** (2004), 78–87.

[9] K. Guo, G. Kutyniok, and D. Labate, Sparse multidimensional representations using anisotropic dilation and shear operators, Wavelets and Splines (Athens, GA, 2005), Nashboro Press, Nashville, TN (2006), 189-201.

[10] K. Guo, D. Labate, W. Lim, G. Weiss, and E. Wilson, Wavelets with composite dilations and their MRA properties, *Appl. Comput. Harmon. Anal.* **20** (2006) 231–249.

[11] K. Guo and D. Labate, Optimal sparse multidimensional representation using shearlets, *SIAM J. Math. Anal.* **9** (2007), 298–318.

[12] K. Guo and D. Labate, The construction of smooth Parseval frames of shearlets, *Math. Model. Nat. Phenom.* **8** (2013), 82–105.

[13] B. Han, Nonhomogeneous wavelet systems in high dimensions. *Appl. Comput. Harmon. Anal.* **32** (2012), 169–196.

[14] B. Han and Z. Zhao, Tensor product complex tight framelets with increasing directionality, *SIAM J. Imaging Sci.* **7** (2014), 997–1034.

[15] B. Han, Z. Zhao, and X. Zhuang, Directional tensor product complex tight framelets with low redundancy, *Appl. Comput. Harmon. Anal.*, to appear.

[16] B. Han and X. Zhuang, Smooth affine shear tight frames with MRA structures, *Appl. Comput. Harmon. Anal.* **39** (2) (2015), 300–338.

[17] E. J. King, G. Kutyniok, and X. Zhuang, Analysis of data separation and recovery problems using clustered sparsity, *J. Math. Imaging Vis.* **48** (2014), 205–234.

[18] G. Kutyniok, M. Sharam, and X. Zhuang, ShearLab: A rational design of a digital parabolic scaling algorithm, *SIAM J. Imaging Sci.* **5** (4) (2012), 1291–1332.

[19] W.-Q. Lim, Nonseparable shearlet transform. *IEEE Trans. Image Proc.* **22** (2013), 2056 - 2065.

[20] Y. M. Lu and M. N. Do, Multidimensional directional filter banks and surfacelets, *IEEE Trans. Image Process.* **16** (2007), 918–931.

[21] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, The dual-tree complex wavelet transform, *IEEE Signal Process. Mag.* **22** (6) (2005), 123–151.

[22] R. Yin, T. Gao, Y. M. Lu, and I. Daubechies, A tale of two bases: Local-nonlocal regularization on image patches with convolution framelets, SIAM J. Imaging Sci. **10** (2017), 711–750.

[23] X. Zhuang, Digital affine shear transforms: fast realization and applications in image/video processing, *SIAM J. Imaging Sci.* **9** (3) (2016), 1437–1466.