Systems for Sparse Representations: Fourier Analysis, Wavelets & Shearlets

Felix Voigtlaender

DEDALE Tutorial Day
November 16, 2016

Applied Functional Analysis Group
Technische Universität Berlin

Outline

Outline

1. Wavelets

2. Curvelets & Shearlets

Definition: Let $\phi, \psi \in L^2(\mathbb{R})$ be a scaling function and a wavelet (i.e., $\int_{\mathbb{R}} |\widehat{\psi}(\omega)|^2 \frac{d\omega}{|\omega|} < \infty$). Then the **inhomogeneous discrete** wavelet system generated by ϕ, ψ is

$$\Big(\phi(\bullet - m)\Big)_{m\in\mathbb{Z}} \cup \Big(2^{j/2} \cdot \psi(2^j \bullet - m)\Big)_{j\in\mathbb{N}_0, m\in\mathbb{Z}}.$$

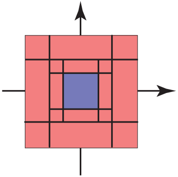
Wavelets in 1D and 2D

Definition: Let $\phi, \psi \in L^2(\mathbb{R})$ be a scaling function and a wavelet (i.e., $\int_{\mathbb{R}} |\widehat{\psi}(\omega)|^2 \frac{\mathrm{d}\omega}{|\omega|} < \infty$). Then the **inhomogeneous discrete** wavelet system generated by ϕ, ψ is

$$\Big(\phi(\bullet - m)\Big)_{m\in\mathbb{Z}} \cup \Big(2^{j/2} \cdot \psi(2^j \bullet - m)\Big)_{j\in\mathbb{N}_0, m\in\mathbb{Z}}.$$

Definition: Let ϕ, ψ as above. A 2D inhomogeneous discrete wavelet system is defined by

$$\Big(\phi^{(1)}(\bullet - m)\Big)_{m\in\mathbb{Z}^2} \cup \Big(2^{j} \cdot \psi^{(\ell)}(2^j \bullet - m)\Big)_{j\in\mathbb{N}_0, m\in\mathbb{Z}^2, \ell\in\{1,2,3\}}$$

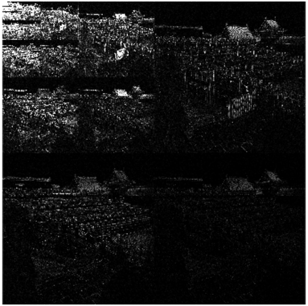


where

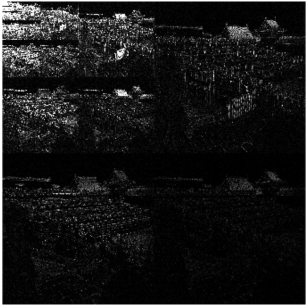
$$\phi^{(1)} := \phi \otimes \phi, \quad \psi^{(1)} := \phi \otimes \psi, \quad \psi^{(2)} := \psi \otimes \phi, \quad \psi^{(3)} := \psi \otimes \psi.$$

For good compression of a signal f, it would be desirable that f can be well approximated using only a few wavelets:

For good compression of a signal f, it would be desirable that f can be well approximated using only a few wavelets:



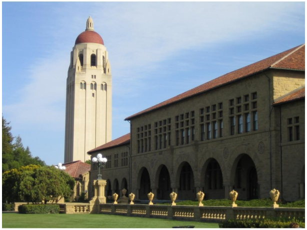
For good compression of a signal f, it would be desirable that f can be well approximated using only a few wavelets:

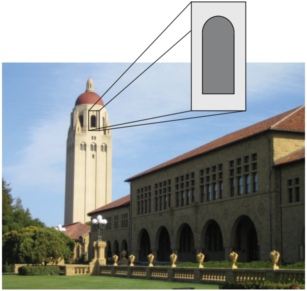
Theorem: Discrete wavelets provide optimal approximation rates for those $f \in L^2(\mathbb{R}^2)$ which are C^2 apart from **point singularities**:

$$\|f - f_N\|_{L^2} \lesssim N^{-1/2} \qquad (N \to \infty).$$

Living on the point

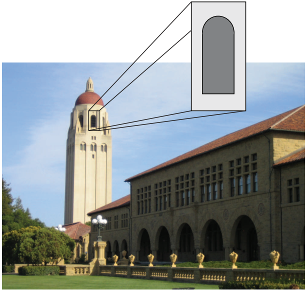


Living on the ~~point~~ edge

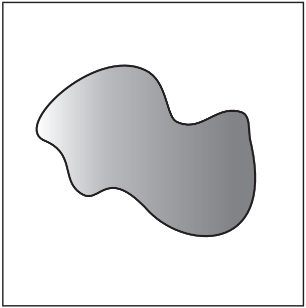


Natural images are governed by curved singularities, not point singularities!

Living on the ~~point~~ edge



$\rightsquigarrow$



Natural images are governed by curved singularities, not point singularities!

Definition (Donoho; 2001)

With $Q := (0,1)^2$, the set of cartoon-like functions is defined as

$$\mathcal{E}^2(\mathbb{R}^2) = \left\{ f_0 + f_1 \cdot \mathbb{1}_B \,\middle|\, \partial B \subset Q \text{ closed } C^2 \text{ curve and } f_0, f_1 \in C_c^2(Q) \right\}.$$

Approximation of cartoon-like functions

Theorem: For $f \in \mathcal{E}^2\left(\mathbb{R}^2\right)$, the best N-term approximation f_N

- using Fourier basis: $\|f - f_N\|_{L^2} \lesssim N^{-1/4}$,
- using Wavelets: $\|f - f_N\|_{L^2} \lesssim N^{-1/2}$.

Approximation of cartoon-like functions

Theorem: For $f \in \mathcal{E}^2(\mathbb{R}^2)$, the best N-term approximation f_N

- using Fourier basis: $\|f - f_N\|_{L^2} \lesssim N^{-1/4}$,
- using Wavelets: $\|f - f_N\|_{L^2} \lesssim N^{-1/2}$.

Theorem (Donoho; 2001): Let $(\psi_n)_{n \in \mathbb{N}}$ be **any** countable family in $L^2(\mathbb{R}^2)$. If

$$\|f - f_N\|_{L^2} \lesssim N^{-\theta} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2),$$

then $\theta \leq 1$. Here, we assume polynomial depth search for forming f_N.

Approximation of cartoon-like functions

Theorem: For $f \in \mathcal{E}^2(\mathbb{R}^2)$, the best N-term approximation f_N

- using Fourier basis: $\|f - f_N\|_{L^2} \lesssim N^{-1/4}$,
- using Wavelets: $\|f - f_N\|_{L^2} \lesssim N^{-1/2}$.

Theorem (Donoho; 2001): Let $(\psi_n)_{n\in\mathbb{N}}$ be **any** countable family in $L^2(\mathbb{R}^2)$. If

$$\|f - f_N\|_{L^2} \lesssim N^{-\theta} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2),$$

then $\theta \leq 1$. Here, we assume polynomial depth search for forming f_N.

Hence, Wavelets **cannot** approximate curvilinear singularities optimally.

Approximation of cartoon-like functions

Theorem: For $f \in \mathcal{E}^2(\mathbb{R}^2)$, the best N-term approximation f_N

- using Fourier basis: $\|f - f_N\|_{L^2} \lesssim N^{-1/4}$,
- using Wavelets: $\|f - f_N\|_{L^2} \lesssim N^{-1/2}$.

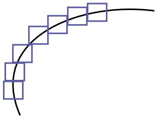
Theorem (Donoho; 2001): Let $(\psi_n)_{n\in\mathbb{N}}$ be **any** countable family in $L^2(\mathbb{R}^2)$. If

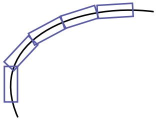
$$\|f - f_N\|_{L^2} \lesssim N^{-\theta} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2),$$

then $\theta \leq 1$. Here, we assume polynomial depth search for forming f_N.

Hence, Wavelets **cannot** approximate curvilinear singularities optimally.

Intuitive explanation: This is caused by the scalar dilations:

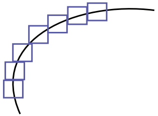


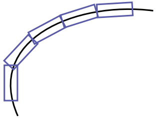


Outline

Towards a directional system

Recall: Isotropic dilations are suboptimal for approximating curves:

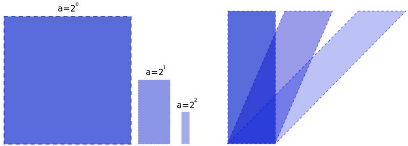




Recall: Isotropic dilations are suboptimal for approximating curves:

Idea for shearlets: Use

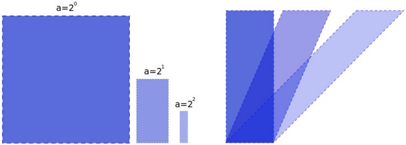
- Parabolic scaling
- Different orientations via shearing.



Recall: Isotropic dilations are suboptimal for approximating curves:

Idea for shearlets: Use

- Parabolic scaling
- Different orientations via shearing.



Advantages of shearing:

- The shearing matrices $\left(\begin{smallmatrix}1 & k\\ 0 & 1\end{smallmatrix}\right)$ leave the digital grid $\mathbb{Z}^2$ invariant.
- Uniform theory for the analog and digital situation.

Definition (Kutyniok, Labate; 2006): For $\psi \in L^2\left(\mathbb{R}^2\right)$, the associated discrete shearlet system is

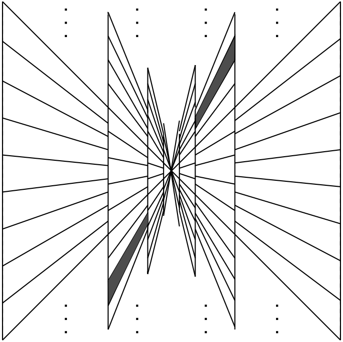
$$\left(2^{\frac{3}{4}j} \cdot \psi\left(S_k A_{2^j} \bullet - m\right)\right)_{j,k\in\mathbb{Z},m\in\mathbb{Z}^2}.$$

A first definition of shearlets

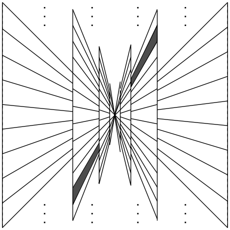
Definition (Kutyniok, Labate; 2006): For $\psi \in L^2(\mathbb{R}^2)$, the associated discrete shearlet system is

$$\left(2^{\frac{3}{4}j} \cdot \psi(S_k A_{2^j} \bullet - m)\right)_{j,k \in \mathbb{Z}, m \in \mathbb{Z}^2}.$$

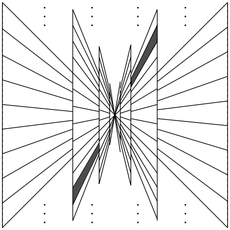
Induced frequency tiling:



The induced frequency tiling

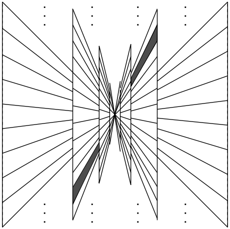


The induced frequency tiling



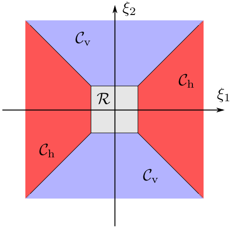
Very different treatment of x-direction and y-direction!

The induced frequency tiling

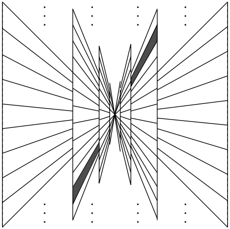


Very different treatment of *x*-direction and *y*-direction!

Solution: Use cone-adapted shearlets:

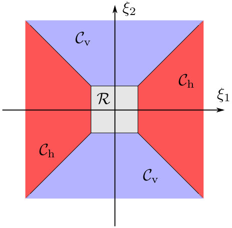


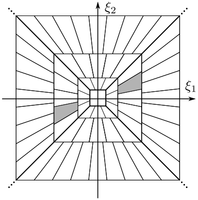
The induced frequency tiling



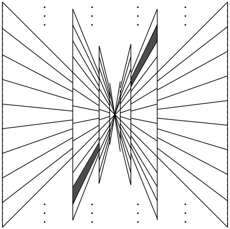
Very different treatment of x-direction and y-direction!

Solution: Use cone-adapted shearlets:



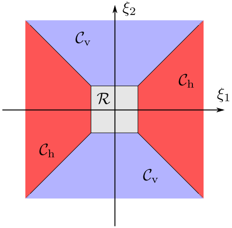


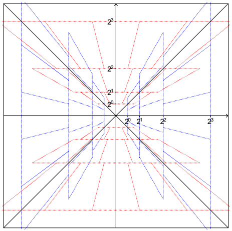
The induced frequency tiling



Very different treatment of *x*-direction and *y*-direction!

Solution: Use cone-adapted shearlets:





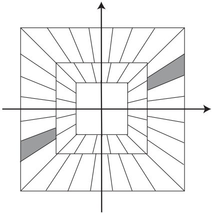
Cone-adapted Discrete Shearlet Systems

Definition (Kutyniok, Labate; 2006):
The cone-adapted discrete shearlet system $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ with sampling density $c > 0$ generated by $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is the union of

$$\left\{\phi(\bullet - c \cdot m) \,\middle|\, m \in \mathbb{Z}^2\right\},$$
$$\left\{\psi_{j,k,m,h} := \psi(S_k A_{2^j} \bullet - c \cdot m) \,\middle|\, j \in \mathbb{N}_0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\right\},$$
$$\left\{\psi_{j,k,m,v} := \tilde{\psi}\left(S_k^T \tilde{A}_{2^j} \bullet - c \cdot m\right) \,\middle|\, j \in \mathbb{N}_0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2\right\},$$

where

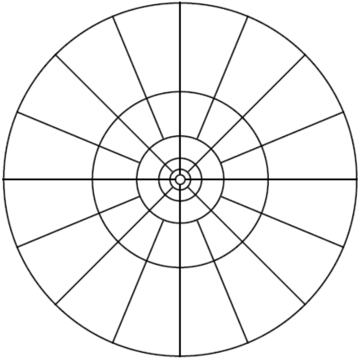
$$S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \qquad A_{2^j} = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \qquad \tilde{A}_{2^j} = \begin{pmatrix} 2^{j/2} & 0 \\ 0 & 2^j \end{pmatrix}.$$

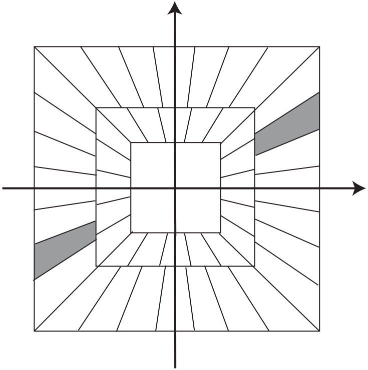


Curvelets induce a similar frequency tiling than shearlets, but using rotations instead of shearings.

Side remark: Curvelets

Curvelets induce a similar frequency tiling than shearlets, but using rotations instead of shearings.





Theorem (Kittipoom, Kutyniok, Lim; 2012):
Let $\phi, \psi, \tilde{\psi} \in L^2\left(\mathbb{R}^2\right)$ be compactly supported and assume that

- $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions,
- we have

$$|\widehat{\phi}(\xi)|^2 + \sum_{j,k} |\widehat{\psi}_{j,k}(\xi)|^2 + |\widehat{\tilde{\psi}_{j,k}}(\xi)|^2 \geq C > 0 \quad \text{a.e.} \qquad (\dagger)$$

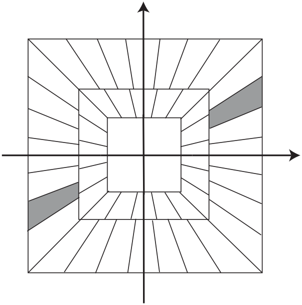
Then there is some $c > 0$ such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2\left(\mathbb{R}^2\right)$.

Theorem (Kittipoom, Kutyniok, Lim; 2012):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported and assume that

- $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions,
- we have

$$|\widehat{\phi}(\xi)|^2 + \sum_{j,k} |\widehat{\psi}_{j,k}(\xi)|^2 + |\widehat{\tilde{\psi}_{j,k}}(\xi)|^2 \geq C > 0 \quad \text{a.e.} \qquad (\dagger)$$

Then there is some $c > 0$ such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\mathbb{R}^2)$.



Theorem (Kittipoom, Kutyniok, Lim; 2012):
Let $\phi, \psi, \tilde{\psi} \in L^2\left(\mathbb{R}^2\right)$ be compactly supported and assume that

- $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions,
- we have

$$|\widehat{\phi}(\xi)|^2 + \sum_{j,k} |\widehat{\psi}_{j,k}(\xi)|^2 + |\widehat{\tilde{\psi}_{j,k}}(\xi)|^2 \geq C > 0 \quad \text{a.e.} \tag{$\dagger$}$$

Then there is some $c > 0$ such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2\left(\mathbb{R}^2\right)$.

Remarks

- For $(\dagger)$, it suffices to have $\tilde{\psi}((\xi_1, \xi_2)) = \psi((\xi_2, \xi_1))$ as well as

$$\begin{aligned} \left|\widehat{\phi}(\xi)\right| &\gtrsim 1 && \text{for } \xi \in [-1,1]^2, \\ |\widehat{\psi}(\xi)| &\gtrsim 1 && \text{for } \xi_1 \in [1/3,\, 3] \text{ and } |\xi_2| \leq |\xi_1|, \end{aligned}$$

- There are special examples with frame bounds $B/A \approx 4$.

Theorem (Kutyniok, Lim; 2011):
Let $\phi, \psi, \tilde{\psi} \in L^2\left(\mathbb{R}^2\right)$ be compactly supported and such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2\left(\mathbb{R}^2\right)$. Further, assume that $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ provides an optimally sparse approximation of all $f \in \mathcal{E}^2\left(\mathbb{R}^2\right)$, i.e.,

$$\|f - f_N\|_{L^2} \lesssim N^{-1} \cdot (\log N)^{3/2} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2).$$

Theorem (Kutyniok, Lim; 2011):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported and such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\mathbb{R}^2)$. Further, assume that $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ provides an optimally sparse approximation of all $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e.,

$$\|f - f_N\|_{L^2} \lesssim N^{-1} \cdot (\log N)^{3/2} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2).$$

Here, f_N is a linear combination of (at most) N elements of the dual frame of $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$.

Theorem (Kutyniok, Lim; 2011):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported and such that $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ is a frame for $L^2(\mathbb{R}^2)$. Further, assume that $\widehat{\psi}, \widehat{\tilde{\psi}}$ satisfy certain decay conditions. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$ provides an optimally sparse approximation of all $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e.,

$$\|f - f_N\|_{L^2} \lesssim N^{-1} \cdot (\log N)^{3/2} \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2).$$

Here, f_N is a linear combination of (at most) N elements of the dual frame of $\mathcal{SH}(\phi, \psi, \tilde{\psi}; c)$.

Remark

The proof shows

$$\sum_{n>N} |\theta(f)|_n^2 \lesssim N^{-2} \cdot (\log N)^3 \qquad \forall N \in \mathbb{N} \qquad \forall f \in \mathcal{E}^2(\mathbb{R}^2),$$

where $|\theta(f)|_n$ denotes the n-th largest shearlet coefficient. Hence, the shearlet coefficients are in ℓ^p for all $p > \frac{2}{3}$.

Thank you!

Thank you!

Questions, comments, counterexamples?

☺