Ridgelets

1. The continuous transform

In this section we study a version of the continuous ridgelet ransform which is also closely related to a neural network with one hidden layer.

If now ψ is a given function, then we define

$$\psi_{\mathbf{u},a,b}(\mathbf{x}) = \frac{1}{\sqrt{a}}\psi\left(\frac{\mathbf{u}\cdot\mathbf{x}-b}{a}\right), \quad |\mathbf{u}| = 1, \quad a > 0, \quad b \in \mathbb{R}.$$

We have the following result.

Theorem 9.1. *Let $d \geq 1$ and let ψ and $\varphi \in L^1(\mathbb{R})$ be such that*

$$\int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)||\hat{\varphi}(\omega)|}{|\omega|^d}\,\mathrm{d}\omega < \infty \quad \textit{and} \quad K_{\psi,\varphi} \stackrel{\text{def}}{=} \int_{\mathbb{R}} \frac{\overline{\hat{\psi}(\omega)}\hat{\varphi}(\omega)}{|\omega|^d}\,\mathrm{d}\omega \neq 0.$$

If $f \in L^1(\mathbb{R}^d)$ is such that $\hat{f} \in L^1(\mathbb{R}^d)$, then

$$f(\mathbf{x}) = \frac{1}{K_{\psi,\varphi}} \int_{\mathbb{S}^{d-1}} \int_0^\infty \int_{\mathbb{R}} \langle f, \psi_{\mathbf{u},a,b}\rangle\, \varphi_{\mathbf{u},a,b}(\mathbf{x})\,\mathrm{d}b\,\mathrm{d}a\,\mathrm{d}\mathbf{u}.$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in $L^2(\mathbb{R}^d)$.

Observe that if ψ and φ are real-valued functions, then $K_{\psi,\varphi}$ is real-valued as well. From the proof we see that we have

$$\int_{\mathbb{S}^{d-1}} \int_0^\infty \left| \int_{\mathbb{R}} \langle f, \psi_{\mathbf{u},a,b}\rangle\, \varphi_{\mathbf{u},a,b}(\mathbf{x})\,\mathrm{d}b \right| \mathrm{d}a\,\mathrm{d}\mathbf{u} < \infty,$$

and that the integral $\int_{\mathbb{R}} \langle f, \psi_{\mathbf{u},a,b}\rangle\, \varphi_{\mathbf{u},a,b}(\mathbf{x})\,\mathrm{d}b$ is the convolution of L^1-functions, and hence well-defined. If $\psi = \varphi$, then it is not difficult to show that the triple integral converges absolutely as well.

Proof. Let $\mathbf{u} \in \mathbb{R}^d$ be such that $|\mathbf{u}| = 1$. We define the Radon-transform $P_{\mathbf{u}} f$ as follows:

$$(P_{\mathbf{u}} f)(t) = \int_{\mathbb{R}^{d-1}} f(t\mathbf{u} + U^{\perp}\mathbf{s})\,\mathrm{d}\mathbf{s},$$

where $U^{\perp}$ is a $d \times (d-1)$ matrix with columns that form an orthonormal basis for the subspace of vectors in $\mathbb{R}^d$ orthogonal to $\mathbf{u}$. It is not difficult to show that $P_{\mathbf{u}} f \in L^1(\mathbb{R})$ and that

$$\widehat{P_{\mathbf{u}} f}(\underline{\omega}) = \hat{f}(\underline{\omega}\mathbf{u}). \tag{9.1}$$

Furthermore, we let, abusing our notation somewhat,

$$\psi_a(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t}{a}\right) \quad \text{and} \quad \tilde{\psi}_a(t) = \overline{\psi_a(-t)}, \quad a > 0 \quad t \in \mathbb{R}.$$

We observe that

$$\langle f, \psi_{\mathbf{u},a,b}\rangle = (\tilde{\psi}_a * P_{\mathbf{u}} f)(b). \tag{9.2}$$

We let

$$\phi(\underline{\omega}) = \overline{\hat{\psi}(\underline{\omega})}\hat{\varphi}(\underline{\omega}) + \overline{\hat{\psi}(-\underline{\omega})}\hat{\varphi}(-\underline{\omega}),$$

and observe that

$$\begin{aligned}\int_0^{\infty} \phi(a\omega)\frac{1}{a^d}\,\mathrm{d}a &= \omega^{d-1}\int_0^{\infty} \phi(a)\frac{1}{a^d}\,\mathrm{d}a \\ &= \omega^{d-1}\int_{\mathbb{R}} \frac{\overline{\hat{\psi}(\eta)}\hat{\varphi}(\eta)}{|\eta|^d}\,\mathrm{d}\eta = \omega^{d-1}K_{\psi,\varphi}, \quad \omega > 0.\end{aligned}$$

The same calculation shows, of course, that there is a constant C such that

$$\int_0^{\infty} |\phi(a\omega)|\,\frac{1}{a^d}\,\mathrm{d}a \le C\omega^{d-1}, \quad \omega > 0. \tag{9.3}$$

If we now let $\mathbf{x} \in \mathbb{R}^d$ be arbitrary and define

$$g(\mathbf{x}) \stackrel{\text{def}}{=} \int_{\mathbb{S}^{d-1}}\int_0^{\infty}\int_0^{\infty} \mathrm{e}^{\mathrm{i}2\pi\omega\mathbf{u}\cdot\mathbf{x}}\phi(a\omega)\hat{f}(\omega\mathbf{u})\frac{1}{a^d}\,\mathrm{d}a\,\mathrm{d}\omega\,\mathrm{d}\mathbf{u}, \tag{9.4}$$

then it follows from (9.3) and our assumptions on f that this integral converges absolutely, and we have in fact

$$\begin{aligned}g(\mathbf{x}) &= K_{\psi,\varphi}\int_{\mathbb{S}^{d-1}}\int_0^{\infty} \mathrm{e}^{\mathrm{i}2\pi\omega\mathbf{u}\cdot\mathbf{x}}\hat{f}(\omega\mathbf{u})\omega^{d-1}\,\mathrm{d}\omega\,\mathrm{d}\mathbf{u} \\ &= K_{\psi,\varphi}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}2\pi\mathbf{y}\cdot\mathbf{x}}\hat{f}(\mathbf{y})\,\mathrm{d}\mathbf{y} = K_{\psi,\varphi}f(\mathbf{x}).\end{aligned} \tag{9.5}$$

By Fubini's theorem and the fact that $\mathbb{S}^{d-1}$ is invariant under the mapping $\mathbf{u} \mapsto -\mathbf{u}$ we get

$$(9.6) \quad g(\mathbf{x}) = \int_{\mathbb{S}^{d-1}} \int_0^\infty \int_0^\infty \Big(\mathrm{e}^{\mathrm{i}2\pi\omega\mathbf{u}\cdot\mathbf{x}} \overline{\hat{\psi}(a\omega)} \hat{\varphi}(a\omega) \hat{f}(\omega\mathbf{u}) + \mathrm{e}^{-\mathrm{i}2\pi\omega\mathbf{u}\cdot\mathbf{x}} \overline{\hat{\psi}(-a\omega)} \hat{\varphi}(-a\omega) \hat{f}(\omega\mathbf{u}) \Big) \frac{1}{a^d} \,\mathrm{d}\omega \,\mathrm{d}a \,\mathrm{d}\mathbf{u}$$
$$= \int_{\mathbb{S}^{d-1}} \int_0^\infty \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}2\pi\omega\mathbf{u}\cdot\mathbf{x}} \overline{\hat{\psi}(a\omega)} \hat{\varphi}(a\omega) \hat{f}(\omega\mathbf{u}) \frac{1}{a^d} \,\mathrm{d}\omega \,\mathrm{d}a \,\mathrm{d}\mathbf{u}.$$

Next we note from (9.1) that the Fourier transform of the function $\tilde{\psi}_a * P_{\mathbf{u}} f * \varphi_a$ is $a\overline{\hat{\psi}(a\underline{\omega})}\hat{\varphi}(a\underline{\omega})\hat{f}(\underline{\omega}\mathbf{u})$, and therefore we get by the Fourier inversion formula

$$(9.7) \qquad g(\mathbf{x}) = \int_{\mathbb{S}^{d-1}} \int_0^\infty (\tilde{\psi}_a * P_{\mathbf{u}} f * \varphi_a)(\mathbf{u}\cdot\mathbf{x}) \frac{1}{a^{d+1}} \,\mathrm{d}a \,\mathrm{d}\mathbf{u}.$$

(By the results above we know that $\int_{\mathbb{S}^{d-1}} \int_0^\infty |(\tilde{\psi}_a * P_{\mathbf{u}} f * \varphi_a)(\mathbf{u}\cdot\mathbf{x})| \frac{1}{a^{d+1}} \,\mathrm{d}a \,\mathrm{d}\mathbf{u} < \infty$.) Now by (9.2)

$$(\tilde{\psi}_a * P_{\mathbf{u}} f * \varphi_a)(\mathbf{u}\cdot\mathbf{x}) = \int_{\mathbb{R}} (\tilde{\psi}_a * P_{\mathbf{u}} f)(b) \varphi_a(\mathbf{u}\cdot\mathbf{x} - b) \,\mathrm{d}b$$
$$= \int_{\mathbb{R}} \langle f, \psi_{\mathbf{u},a,b} \rangle \, \varphi_{\mathbf{u},a,b}(\mathbf{x}) \,\mathrm{d}b.$$

When this result is combined with (9.5) and (9.7) we get the claim of the theorem. □

2. An orthonormal basis of almost-ridgelets

Let us consider the space $L^2(\mathbb{R}^2)$ and construct a special kind of orthonormal basis. Let φ be the scaling function and ψ be the corresponding wavelet function for an orthonormal multiresolution of $L^2(\mathbb{R})$ and we assume that φ and ψ have compact support. For $n_0 \le 0$ we define

$$w_{n_0,p,0}(\underline{x}) = \sum_{j\in\mathbb{Z}} 2^{-\frac{n_0}{2}} \varphi(2^{-n_0}(\underline{x}+j)-p), \quad p = 0,\dots,2^{n_0-1},$$
$$w_{n,p,1} = \sum_{j\in\mathbb{Z}} 2^{-\frac{n}{2}} \psi(2^{-n}(\underline{x}+j)-p), \quad n \le n_0, p = 0,\dots,2^{n-1}.$$

We leave it as an exercise to show that these functions form an orthonormal basis for $L^2(\mathbb{T})$.

Next we choose a wavelet function Ψ such that $(\Psi_{m,k})_{n,k\in\mathbb{Z}}$ is an orthonormal basis in $L^2(\mathbb{R})$ where $\Psi_{m,k}(\underline{x}) = 2^{-\frac{m}{2}}\Psi(2^{-m}\underline{x}-k)$. But in addition we require that

$$\Psi_{m,k}(-\underline{x}) = \Psi_{m,1-k}(\underline{x}).$$

(Other variants of this condition could be used as well). In order to construct such a wavelet one can start by taking an even function Ψ_* such that $\sum_{m\in\mathbb{Z}}\big|\hat{\Psi}_*(2^m\underline{\omega})\big|^2 \overset{\text{a.e.}}{=} 1$ and $\hat{\psi}(\underline{\omega})\overline{\hat{\psi}\big((\underline{\omega}+k)\big)} \overset{\text{a.e.}}{=} \sum_{p=1}^{\infty}\hat{\psi}(2^p\underline{\omega})\overline{\hat{\psi}\big(2^p(\underline{\omega}+k)\big)}$ for all odd integers k, and then one takes $\Psi(\underline{x}) = \Psi_*(\underline{x}+\frac{1}{2})$ or equivalently $\widehat{\Psi}(\underline{\omega}) = e^{i\pi\underline{\omega}}\widehat{\Psi_*}(\underline{\omega})$. One possible choice, which does not give very fast decay for Ψ is to take $\Psi_*(\omega) = 1$ when $\frac{1}{2} < \omega < 1$ and 0 otherwise.

Now we choose $n_0 \le -1$ and we define the functions ρ_λ where $\lambda = (m,k,n,p,\sigma)$ with $m \in \mathbb{Z}$, $k \ge 1$ $\sigma \in \{0,1\}$, $n = n_0$ if $\sigma = 0$ and $n \le n_0$ if $\sigma = 1$, and $p = 0,1\ldots,2^{-n}-1$. We do this by defining the Fourier transform

$$\widehat{\rho_\lambda}\big(r\cos(2\pi\theta), r\sin(2\pi\theta)\big) = \frac{1}{\sqrt{2\pi r}}\Big(\widehat{\Psi_{m,k}}(r)w_{n,p,\sigma}(\theta) + \widehat{\Psi_{m,k}}(-r)w_{n,p,\sigma}(\theta+\tfrac{1}{2})\Big).$$

Now we have

$$\begin{aligned}
\langle\rho_\lambda,\rho_{\lambda'}\rangle &= \langle\widehat{\rho_\lambda},\widehat{\rho_{\lambda'}}\rangle \\
&= \int_0^\infty\int_0^1 \frac{2\pi r}{2\pi r}\Big(\widehat{\Psi_{m,k}}(r)w_{n,p,\sigma}(\theta) + \widehat{\Psi_{m,k}}(-r)w_{n,p,\sigma}(\theta+\tfrac{1}{2})\Big) \\
&\quad\times\overline{\Big(\widehat{\Psi_{m',k'}}(r)w_{n',p',\sigma'}(\theta) + \widehat{\Psi_{m',k'}}(-r)w_{n',p',\sigma'}(\theta+\tfrac{1}{2})\Big)}\,\mathrm{d}\theta\,\mathrm{d}r \\
&= \int_0^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(r)}\,\mathrm{d}r\int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta \\
&\quad+\int_0^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r\int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta \\
&\quad+\int_0^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r\int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta \\
&\quad+\int_0^\infty \widehat{\Psi_{m,k}}(-r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r\int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta.
\end{aligned}$$

Observe that $w_{n,p,\sigma}(\underline{\theta}+\frac{1}{2}) = w_{n,\tilde{p},\sigma}(\underline{\theta})$ where $|p-\tilde{p}| = 2^{-n-1}$. If now $\lambda = \lambda'$ then we have

$$\int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta = \int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta = 1,$$
$$\int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta = \int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta = 0,$$

and hence we get

$$\begin{aligned}\langle \rho_\lambda, \rho_{\lambda'}\rangle &= \int_0^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(r)}\,\mathrm{d}r + \int_0^\infty \widehat{\Psi_{m,k}}(-r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r \\ &= \int_{-\infty}^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(r)}\,\mathrm{d}r = 1.\end{aligned}$$

If $\sigma \neq \sigma'$ or $n \neq n'$ or neither $p = p'$ nor $|p-p'| = 2^{-n-1}$ then

$$\int_0^1 w_{n,p,\sigma}(\theta + z\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta + z'\tfrac{1}{2})}\,\mathrm{d}\theta = 0$$

for all $z, z' \in \{0,1\}$. Thus when $\lambda \neq \lambda'$ it remains to consider the cases where $\sigma = \sigma'$, $n = n'$ and either $p = p'$ or $|p - p'| = 2^{-n-1}$. Assume first that $p = p'$. Then we get by the same argument as above that

$$\langle \rho_\lambda, \rho_{\lambda'}\rangle = \int_{-\infty}^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(r)}\,\mathrm{d}r = 0,$$

because we have $(m,k) \neq (m',k')$. Next we assume that $|p-p'| = 2^{-n-1}$. Then we have

$$\begin{aligned}\int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta &= \int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta = 0,\\ \int_0^1 w_{n,p,\sigma}(\theta+\tfrac{1}{2})\overline{w_{n',p',\sigma'}(\theta)}\,\mathrm{d}\theta &= \int_0^1 w_{n,p,\sigma}(\theta)\overline{w_{n',p',\sigma'}(\theta+\tfrac{1}{2})}\,\mathrm{d}\theta = 1,\end{aligned}$$

and then we get

$$\begin{aligned}\langle \rho_\lambda, \rho_{\lambda'}\rangle &= \int_0^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r + \int_0^\infty \widehat{\Psi_{m,k}}(-r)\overline{\widehat{\Psi_{m',k'}}(r)}\,\mathrm{d}r \\ &= \int_{-\infty}^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',k'}}(-r)}\,\mathrm{d}r = \int_{-\infty}^\infty \widehat{\Psi_{m,k}}(r)\overline{\widehat{\Psi_{m',1-k'}}(r)}\,\mathrm{d}r = 0,\end{aligned}$$

because $k \neq 1 - k'$ since $k, k' \geq 1$.

It remains to show that the functions ρ_λ span $L^2(\mathbb{R}^2)$, but we leave this as an exercise.

Another way to look at these functions, and which is the reason why they could be called ridgelets is to define

$$\Psi^+_{m,k}(\underline{t}) = \int_{-\infty}^\infty \sqrt{|\omega|}\,\widehat{\Psi_{m,k}}(\omega)\mathrm{e}^{\mathrm{i}2\pi\omega\underline{t}}\,\mathrm{d}\omega.$$

Then we get by the Fourier inversion formula (without worrying about convergence questions for the moment)

$$
\begin{aligned}
\rho_\lambda((x_1,x_2)) &= \int_0^\infty \int_0^1 2\pi r \mathrm{e}^{\mathrm{i}2\pi(rx_1\cos(2\pi\theta)+rx_2\sin(2\pi\theta))} \\
&\quad\times \frac{1}{\sqrt{2\pi r}}\Big(\widehat{\Psi_{m,k}}(r)w_{n,p,\sigma}(\theta) + \widehat{\Psi_{m,k}}(-r)w_{n,p,\sigma}(\theta+\tfrac{1}{2})\Big)\,\mathrm{d}\theta\,\mathrm{d}r \\
&= \int_0^1\int_0^\infty \sqrt{2\pi r}\mathrm{e}^{\mathrm{i}2\pi(rx_1\cos(2\pi\theta)+rx_2\sin(2\pi\theta))}\widehat{\Psi_{m,k}}(r)w_{n,p,\sigma}(\theta)\mathrm{d}r\,\mathrm{d}\theta \\
&\quad + \int_0^1\int_0^\infty \sqrt{2\pi|-r|}\mathrm{e}^{\mathrm{i}2\pi((-r)x_1\cos(2\pi(\theta+\frac{1}{2}))+(-r)x_2\sin(2\pi(\theta+\frac{1}{2})))} \\
&\qquad\qquad \times \widehat{\Psi_{m,k}}(r)w_{n,p,\sigma}(\theta+\tfrac{1}{2})\mathrm{d}r\,\mathrm{d}\theta \\
&= \int_0^1 \left(\int_{-\infty}^\infty \sqrt{2\pi|r|}\mathrm{e}^{\mathrm{i}2\pi(rx_1\cos(2\pi\theta)+rx_2\sin(2\pi\theta))}\widehat{\Psi_{m,k}}(r)\mathrm{d}r\right)\, w_{n,p,\sigma}(\theta)\,\mathrm{d}\theta \\
&= \int_0^1 \Psi^+_{m,k}\big(rx_1\cos(2\pi\theta)+rx_2\sin(2\pi\theta)\big)w_{n,p,\sigma}(\theta)\,\mathrm{d}\theta.
\end{aligned}
$$