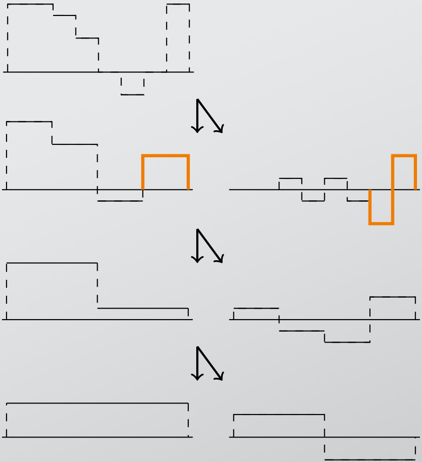
MRA wavelet frames and the UEP

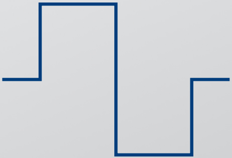
Applied harmonic analysis
YSC4206 Fall 2016

For the material of this lecture and more, see:
I.Daubechies, *Ten lectures on wavelets*, SIAM 1992
B.Dong, Z.Shen, *MRA-based wavelet frames and applications*, IAS/Park City Mathematics Series: The Mathematics of Image Processing 2010
S.Mallat, *A wavelet tour of signal processing: the sparse way*, Academic Press 2008

Recall the simplest MRA and wavelet idea



Approximation (MRA), Details (wavelets)



Haar wavelet

- MRA's and wavelets allow representation of signals at different levels of resolution

Wavelet systems

Give $\Psi := \{\psi_1, \ldots, \psi_r\} \subset L_2(\mathbb{R})$ consider the **wavelet system**

$$X(\Psi) := \{\psi_{\ell,n,k} \colon 1 \leq \ell \leq r; n, k \in \mathbb{Z}\},$$

where $\psi_{\ell,n,k} := \mathcal{D}^n T_k \psi_\ell = 2^{n/2}\psi_\ell(2^n \cdot -k)$ are **shifts** and **dilations**.

If $X(\Psi)$ is a tight frame for $L_2(\mathbb{R})$, then it is called a **tight wavelet frame** for $L_2(\mathbb{R})$ and the elements of Ψ are called **wavelets**.

Note: In all the following by tight frame we mean 1-tight, i.e. with frame bound 1.

The main **question** is:

- How to choose Ψ such that $X(\Psi)$ is a tight frame for $L_2(\mathbb{R})$?

Wavelet systems

Theorem: $X(\Psi)$ is a tight frame for $L_2(\mathbb{R})$ if and only if

$$\sum_{\psi\in\Psi}\sum_{k\in\mathbb{Z}}|\widehat{\psi}(2^k\xi)|^2 = 1; \quad \sum_{\psi\in\Psi}\sum_{k=0}^{\infty}\widehat{\psi}(2^k\xi)\overline{\widehat{\psi}(2^k(\xi+(2j+1)2\pi)} = 0 \quad j\in\mathbb{Z}, \quad (1)$$

holds for a.e. $\xi\in\mathbb{R}$.

- Beautiful result but provides little help in the actual construction of compactly supported wavelets. So we don't prove this.
- Instead, we turn to **wavelet systems connected to MRA's**. They are of great practical importance since for them **fast decomposition and reconstruction algorithms** exist. They can be constructed via extension principles.

A.Ron, Z.Shen: *Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator*, J Func Anal 1997.

Unitary extension principle

... where we prove a theorem on wavelets, leading to explicit examples.

Unitary extension principle

Let $\{V_n\}$ an MRA generated by refinable ϕ with mask h_0.

Goal: Find wavelets $\Psi := \{\psi_1, \ldots, \psi_r\} \subset V_1$ by finding masks $h_\ell \in \ell_2(\mathbb{Z})$ and letting

$$\psi_\ell(x) := 2 \sum_{k \in \mathbb{Z}} h_\ell[k] \phi(2x - k),$$

such that $X(\Psi)$ is a tight frame for $L_2(\mathbb{R})$.

- In Fourier domain this means finding 2π-periodic $\widehat{h}_\ell$ with

$$\widehat{\psi}_\ell(2\cdot) = \widehat{h}_\ell \widehat{\phi}.$$

- $h_1, \ldots, h_r$ are called **wavelet masks**.
- We will see that in numerical applications the masks govern "everything". One does not need to access the refinement function and wavelets. This is in contrast to Fourier analysis where time and again one needs to compute values $e^{i\omega}$.

Unitary extension principle

UEP-Theorem [Ron-Shen] Let $\phi \in L_2(\mathbb{R})$ be a compactly supported refinable function with finite mask h_0 and $\widehat{\phi}(0) = 1$. Suppose $\Psi = \{\psi_1, \ldots, \psi_r\}$ are defined by finite masks $h_1, \ldots, h_r$. Then $X(\Psi)$ is a tight frame for $L_2(\mathbb{R})$ provided for all $\xi \in \mathbb{R}$

$$\sum_{\ell=0}^{r} |\widehat{h}_\ell(\xi)|^2 = 1 \quad \text{and} \quad \sum_{\ell=0}^{r} \widehat{h}_\ell(\xi)\overline{\widehat{h}_\ell(\xi+\pi)} = 0 \tag{2}$$

If furthermore $r = 1$ and $\|\phi\| = 1$ then $X(\Psi)$ is an orthonormal basis of $L_2(\mathbb{R})$.

Unitary extension principle

- Assumptions on ϕ ensure that it generates an MRA and that $[\widehat{\phi}, \widehat{\phi}]$ is bounded (thus shifts of ϕ are Bessel system)
- UEP condition (2) means

$$\begin{pmatrix} \widehat{h}_0(\xi) & \widehat{h}_1(\xi) & \cdots & \widehat{h}_r(\xi) \\ \widehat{h}_0(\xi+\pi) & \widehat{h}_1(\xi+\pi) & \cdots & \widehat{h}_r(\xi+\pi) \end{pmatrix}$$

  has orthonormal rows. It can thus be extended to $(r+1) \times (r+1)$ unitary matrix.
- First column of extended matrix has norm 1. Thus necessarily

$$|\widehat{h}_0|^2 + |\widehat{h}_0(\cdot + \pi)|^2 \leq 1. \tag{3}$$

  Electronical engineers call this a *quadrature mirror filter* condition. If ϕ is orthonormal then this holds with equality.
- UEP condition rewritten in spatial domain becomes: For all $p \in \mathbb{Z}$:

$$\sum_{\ell=0}^{r} \sum_{k \in \mathbb{Z}} \overline{h_\ell[k]} h_\ell[k-p] = \delta_{p,0} \text{ and } \sum_{\ell=0}^{r} \sum_{k \in \mathbb{Z}} (-1)^{k-p} \overline{h_\ell[k]} h_\ell[k-p] = 0$$

Unitary extension principle

Proof of UEP: Let $\phi_{n,k} := \mathcal{D}^n T_k \phi$ and consider the operator

$$\mathcal{P}_n \colon f \mapsto \sum_{k \in \mathbb{Z}} \langle f, \phi_{n,k} \rangle \phi_{n,k}$$

- $[\widehat{\phi}, \widehat{\phi}] \leq C \implies (\phi_{n,k})_{k \in \mathbb{Z}}$ Bessel seq $\implies \mathcal{P}_n$ well-defined and $\|\mathcal{P}_n f\| \leq C \|f\|$
- Notice that $\mathcal{P}_n = \mathcal{D}^n \mathcal{P}_0 \mathcal{D}^{-n}$
- **Proposition:** If $f, \phi \in L_2(\mathbb{R})$ and $[\widehat{\phi}, \widehat{\phi}] < C$, then in L_2

$$\sum_{k \in \mathbb{Z}} \langle f, \phi(\cdot - k) \rangle e^{-ik\xi} = [\widehat{f}, \widehat{\phi}](\xi)$$

- **Lemma 1:** $\forall f \in L_2(\mathbb{R}) :$

$$\mathcal{P}_n f = \mathcal{P}_{n-1} f + \sum_{\ell=1}^{r} \sum_{k \in \mathbb{Z}} \langle f, \psi_{\ell,n-1,k} \rangle \psi_{\ell,n-1,k}$$

- **Lemma 2:** $\forall f \in L_2(\mathbb{R}) : \lim_{n \to -\infty} \mathcal{P}_n f = 0$
- **Lemma 3:** $\forall f \in L_2(\mathbb{R}) : \lim_{n \to \infty} \mathcal{P}_n f = f$

B-spline tight frames

... here are the promised examples.

B-spline tight frames

Recall mask of **B-spline of order m:** $\widehat{h}_0(\xi) := e^{-ij\frac{\xi}{2}} \cos^m(\xi/2)$, where $j = 0$ for even m, and $j = 1$ for odd m.

Idea: $|\widehat{h}_0(\xi)|^2$ **is first term in binomial expansion of**

$$(\cos^2(\xi/2) + \sin^2(\xi/2))^m = 1.$$

Define m **wavelet masks as the other terms.**

$$\widehat{h}_\ell(\xi) := -i^\ell e^{-ij\frac{\xi}{2}} \sqrt{\binom{m}{\ell}} \sin^\ell(\xi/2) \cos^{m-\ell}(\xi/2), \quad \ell = 1, \ldots, m.$$

Then $\sum_{\ell=0}^m |\widehat{h}_\ell(\xi)|^2 = 1$ and

$$\sum_{\ell=0}^m \widehat{h}_\ell(\xi) \overline{\widehat{h}_\ell(\xi + \pi)} = e^{ij\frac{\xi}{2}} (\cos(\xi/2)\sin(\xi/2))^m (1-1)^m = 0.$$

Thus, UEP implies that the wavelets

$$\widehat{\psi}_\ell(\xi) := -i^\ell e^{-ij\frac{\xi}{2}} \sqrt{\binom{m}{\ell}} \frac{\sin^{m+\ell}(\xi/4)\cos^{m-\ell}(\xi/4)}{(\xi/4)^m}, \quad \ell = 1, \ldots, m,$$

generate a tight frame for $L_2(\mathbb{R})$. ($(-i^\ell)$-factor to stay real in spatial domain)

Piecewise linear B-spline tight frames

Example: $B_2(x) = \max(1 - |x|, 0)$ has mask

$$\widehat{h}_0(\xi) = (\cos \xi/2)^2 = \frac{1}{4}\left(e^{i\xi/2} + e^{-i\xi/2}\right)^2 = \frac{1}{4}(e^{i\xi} + 2 + e^{-i\xi})$$

i.e. $h_0 = [\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$.

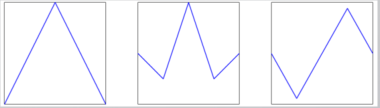
The wavelet masks are

$$\widehat{h}_1(\xi) = -i\sqrt{2}(\sin \xi/2)(\cos \xi/2) = \frac{\sqrt{2}}{4}(e^{-i\xi} - e^{i\xi})$$

and

$$\widehat{h}_2(\xi) = -(\sin \xi/2)^2 = -\frac{1}{4}(e^{-i\xi} - 2 + e^{i\xi})$$

i.e. $h_1 = [\frac{\sqrt{2}}{4}, 0, -\frac{\sqrt{2}}{4}]$ and $h_2 = [-\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}]$.



Piecewise cubic B-spline tight frames

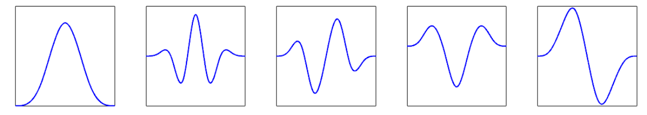
Example: Refinement mask

$$h_0 = \left[\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\right].$$

Wavelet masks

$$h_1 = \left[\frac{1}{16}, -\frac{1}{4}, \frac{3}{8}, -\frac{1}{4}, \frac{1}{16}\right], \quad h_2 = \left[-\frac{1}{8}, \frac{1}{4}, 0, -\frac{1}{4}, \frac{1}{8}\right],$$

$$h_3 = \left[\frac{\sqrt{6}}{16}, 0, -\frac{\sqrt{6}}{8}, 0, \frac{\sqrt{6}}{16}\right], \quad h_4 = \left[-\frac{1}{8}, -\frac{1}{4}, 0, \frac{1}{4}, \frac{1}{8}\right].$$



Fast wavelet transform

... where we turn to the practical implementations.

Fast wavelet transform

In practice, decompose signal down to certain level instead of negative infinity:

- If $\psi_1, \ldots, \psi_r$ constructed from UEP with refinable ϕ and $L \in \mathbb{Z}$, then
$$\{\phi_{L,k}, \psi_{\ell,n,k} \colon 1 \le \ell \le r, n \ge L, k \in \mathbb{Z}\}$$
is a tight frame for $L_2(\mathbb{R})$.
- Precisely, if $f \in L_2(\mathbb{R})$, then
$$f = \lim_{N \to \infty} \mathcal{P}_N f$$
where
$$\mathcal{P}_N f = \sum_{k \in \mathbb{Z}} \langle f, \phi_{L,k} \rangle \phi_{L,k} + \sum_{\ell=1}^{r} \sum_{n=L}^{N-1} \sum_{k \in \mathbb{Z}} \langle f, \psi_{\ell,n,k} \rangle \psi_{\ell,n,k}.$$

One also has to cut off at some fine scale and, instead of letting $N \to \infty$, consider the approximation $\mathcal{P}_N f$ of f for some fixed N. Since $\mathcal{P}_N = \mathcal{D}^N \mathcal{P}_0 \mathcal{D}^{-N}$ we may use $\mathcal{P}_0 f \in V_0$ to approximate f. If necessary, one can consider $f(2^{-N} \cdot)$ instead of f, since the approximation of f in V_N is the same as that of $f(2^{-N} \cdot)$ in V_0.

Coefficients in this tight frame expansion are inner products, i.e. integrals, and thus expensive to compute. Is there an efficient and fast way to compute all these coefficients?

Fast wavelet transform

Indeed, there is an hierarchical and fast scheme to iteratively compute the inner products. It's a consequence of the refinement equation and the way we defined the wavelets in the first place:

- For all $k \in \mathbb{Z}$:

$$\langle f, \phi_{n,k} \rangle = \sqrt{2} \sum_{j \in \mathbb{Z}} \overline{h_0[j-2k]} \langle f, \phi_{n+1,j} \rangle \quad \forall n = L, \dots, -1,$$

$$\langle f, \psi_{\ell,n,k} \rangle = \sqrt{2} \sum_{j \in \mathbb{Z}} \overline{h_\ell[j-2k]} \langle f, \phi_{n+1,j} \rangle \quad \forall n = L, \dots, -1; \, \ell = 1, \dots, r.$$

Thus f can be decomposed from the coefficients $(\langle f, \phi_k \rangle)_{k \in \mathbb{Z}}$ without explicitly computing any other inner products.

Fast wavelet transform

Fast wavelet **decomposition** can be viewed as computation of successively coarser approximations of f (or a given approximation $(\langle f, \phi_k\rangle)_{k\in\mathbb{Z}}$ of f [1]) together with difference in information between every two successive levels.

- Start from fine scale approximation $\mathcal{P}_0 f$ of f (think digital image) given by $(\langle f, \phi_{0,k}\rangle)_{k\in\mathbb{Z}}$ and compute information for the next coarser level $(\langle f, \phi_{-1,k}\rangle)_{k\in\mathbb{Z}}$ along with information $(\langle f, \psi_{\ell,-1,k}\rangle)_{k\in\mathbb{Z},\ell=1,\dots,r}$ for difference between the two approximation levels.
- Stop after finite number of levels.
- Then the information in $(\langle f, \phi_{0,k}\rangle)_{k\in\mathbb{Z}}$ is rewritten as levels of details $(\langle f, \psi_{\ell,n,k}\rangle)_{k\in\mathbb{Z},\ell=1,\dots,r;n=L,\dots,-1}$ plus a final coarse approximation $(\langle f, \phi_{L,k}\rangle)_{k\in\mathbb{Z}}$.

[1] Using these coeffis as input is called "a wavelet crime" by Strang/Nguyen, *Wavelet and Filter Banks*, Wellesley-Cambridge Press 1996, p232

Fast wavelet transform

In practice: Efficient implementation via convolution
In theory: Easier to think in terms of matrix multiplications.

- Let

$$v_{\ell,n} := \begin{pmatrix} \vdots \\ \langle f, \psi_{\ell,n,k} \rangle \\ \vdots \end{pmatrix}, \quad \ell = 0, \ldots, r;\, n = L, \ldots, 0,$$

  where $\psi_0 := \phi$. Think: Coefficients of channel ℓ and level n.

- Define infinite matrix

$$H_\ell := \left(\sqrt{2} \cdot \overline{h_\ell[k-2j]} \right)_{j,k \in \mathbb{Z}}.$$

- Refinement equation and wavelet definition via masks yield

$$v_{\ell,n} = H_\ell v_{0,n+1} \text{ for } n = L, \ldots, -1.$$

Fast wavelet transform

Spatial domain UEP condition implies (as we showed already)

$$2\sum_{\ell=0}^{r}\sum_{k\in\mathbb{Z}}\overline{h_\ell[j-2k]}h_\ell[j'-2k]=\delta_{j,j'} \quad \text{i.e.} \quad \sum_{\ell=0}^{r}H_\ell^* H_\ell = I.$$

Thus

$$v_{0,n+1}=\sum_{\ell=0}^{r}H_\ell^* H_\ell v_{0,n+1}=\sum_{\ell=0}^{r}H_\ell^* v_{\ell,n} \text{ for } n=L,\ldots,-1.$$

Fast wavelet transform

- H_ℓ is convolution and downsampling

$$(H_\ell v)[l] = \sqrt{2}\sum_{k\in\mathbb{Z}} \overline{h_\ell[k-2l]}v[k] \quad \text{i.e.} \quad H_\ell v = \downarrow(\sqrt{2}\cdot\overline{h_\ell[-\cdot]} * v)$$

  H_ℓ^* is upsampling and convolution

$$(H_\ell^* v)[k] = \sqrt{2}\sum_{l\in\mathbb{Z}} \overline{h_\ell[k-2l]}v[l] \quad \text{i.e.} \quad H_\ell^* v = \sqrt{2}h_\ell * (\uparrow v).$$

  Here $\downarrow v := (\dots, v_{-2}, v_0, v_2, \dots)$ and $\uparrow v := (\dots, v_1, 0, v_0, 0, v_1, \dots)$.
- f is **decomposed by convolution followed by downsampling**.
- The information $(\langle f, \phi_{n+1,k}\rangle)_{k\in\mathbb{Z}}$ can be **reconstruct**ed from the information $(\langle f, \phi_{n,k}\rangle)_{k\in\mathbb{Z}}$ on the next coarser level and the "difference" in information $(\langle f, \psi_{\ell,n,k}\rangle)_{k\in\mathbb{Z},\ell=1,\dots,r}$, **by up-sampling followed by convolution**.
- In practice **signals of finite length**. Thus **need to do some extension** of the signal to avoid artifacts caused by boundaries. After convolution (`conv2` in matlab) with the masks, restrict again to original support of the signal.
- To implement this, apparently we only need the masks. Masks are called filters in the EE-community where the design of **filter banks** is its own topic.

Fast decomposition and reconstruction

Let $v \in \mathbb{R}^N$ a univariate signal, N an integer multiple of 2^L and L the number of levels. Denote $v_{0,0} := v$.

- **Decomposition:** For each $j = 1, 2, \ldots, L$
  - Obtain low frequency approximation to v at level j

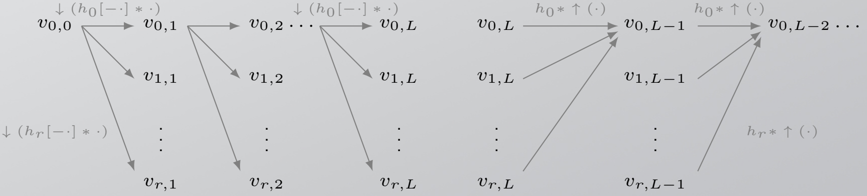
$$v_{0,j} = \downarrow (\sqrt{2} \cdot \overline{h_0[-\cdot]} * v_{0,j-1})$$

  - Obtain wavelet coefficients of v at level j

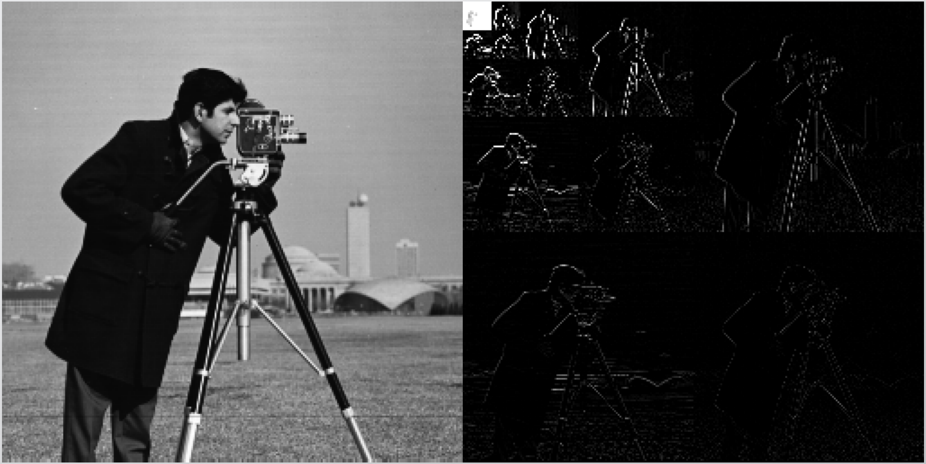
$$v_{\ell,j} = \downarrow (\sqrt{2} \cdot \overline{h_\ell[-\cdot]} * v_{0,j-1}) \quad \text{for } \ell = 1, 2, \ldots, r$$

- **Reconstruction:** For each $j = L, L-1, \ldots, 1$

$$v_{0,j-1} = \sum_{\ell=0}^{r} \sqrt{2} h_\ell * (\uparrow v_{\ell,j})$$



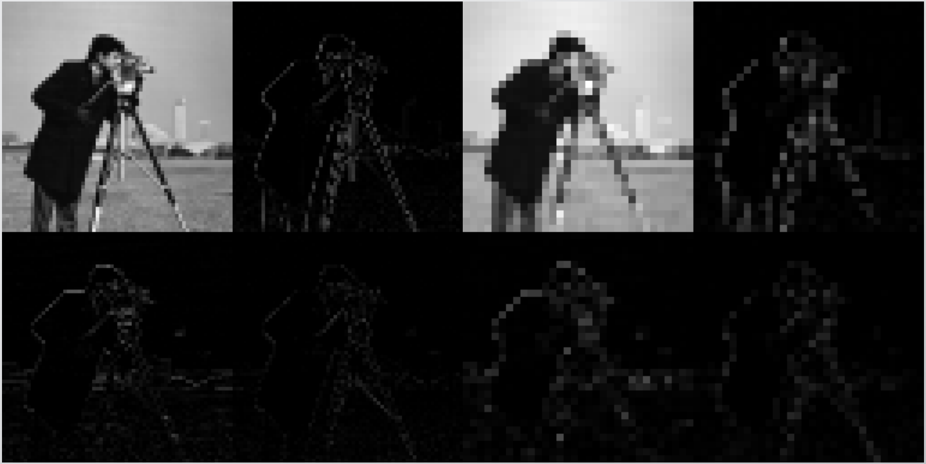
Fast decomposition and reconstruction



Fast decomposition and reconstruction

Fast decomposition and reconstruction

Fast decomposition and reconstruction



Shift-invariant systems

... where we just don't downsample but in return get shift-invariance.

Stationary wavelet transform

- In signal or image processing **shift-invariant** wavelet systems are preferred. Otherwise small shift in the signal can result in totally different wavelet coefficients. (This can be problematic e.g. in detection algorithms.)
- Wavelet system $X(\Psi) := \{\psi_{\ell,n,k} \colon 1 \leq \ell \leq r; n,k \in \mathbb{Z}\}$ is **not** shift-invariant.

  Indeed, $\psi_{\ell,n,k}(x-s) = 2^{n/2}\psi_\ell(2^n x - (2^n s + k))$ but if $n < 0$ then $2^n s + k$ is in general not an integer.

Oversample $X(\Psi)$ below level 0 to achieve **shift-invariance**: Let

$$\psi^q_{\ell,n,k} := \begin{cases} \mathcal{D}^n T_k \psi_\ell & n \geq 0 \\ 2^{n/2} T_k \mathcal{D}^n \psi_\ell & n < 0 \end{cases}$$

and consider

$$X^q(\Psi) := \{\psi^q_{\ell,n,k} \colon 1 \leq \ell \leq r; n,k \in \mathbb{Z}\}.$$

Stationary wavelet transform

- Corresponding masks are derived by **zero padding:** For $n < 0$ (sic)

$$\begin{aligned}
\psi^q_{\ell,n-1,k} &= 2^{n-1} T_k \psi_\ell(2^{n-1}\cdot) = 2^n T_k \sum_{k' \in \mathbb{Z}} h_\ell[k'] \phi(2^n \cdot - k') \\
&= \sum_{k' \in \mathbb{Z}} h_\ell[k'] 2^n \phi(2^n(\cdot - k - 2^{-n}k')) \\
&= \sum_{k' \in 2^{-n}\mathbb{Z}} h_\ell[2^n k'] 2^n \phi(2^n(\cdot - k - k')) \\
&= \sum_{k' \in \mathbb{Z}} h_{\ell,n}[k'] 2^n T_{k+k'} \phi(2^n \cdot)
\end{aligned}$$

where

$$h_{\ell,n}[k] := \begin{cases} h_\ell[2^n k] & \text{if } k \in 2^{-n}\mathbb{Z} \\ 0 & \text{else} \end{cases}$$

i.e. $h_{\ell,n}$ **is obtained from** $h_{\ell,n+1}$ **by inserting zeros between every two entries; and** $h_{\ell,0} = h_\ell$**.**

Stationary wavelet transform

Theorem: Let $\Psi = \{\psi_\ell \colon 1 \le \ell \le r\}$ be wavelets constructed from the UEP with refinable function ϕ. Then the **shift-invariant system**

$$X^q(\Psi) := \{\phi_{0,k}, \psi^q_{\ell,n,k} \colon 1 \le \ell \le r, n \ge 0, k \in \mathbb{Z}\}$$

is a tight frame for $L_2(\mathbb{R})$, i.e.

$$f = \sum_{k\in\mathbb{Z}} \langle f, \phi_{0,k}\rangle \phi_{0,k} + \sum_{\ell=1}^{r} \sum_{n=0}^{\infty} \sum_{k\in\mathbb{Z}} \langle f, \psi^q_{\ell,n,k}\rangle \psi^q_{\ell,n,k}.$$

- The **fast algorithm with the padded filters works exactly the same as before, except that there is NO down/up sampling**. In particuar, the signal can be of any length.
- We don't proof this result and refer to the literature (it's similar in spirit to the UEP result we did).

Tensor products for 2D

... where we move on to $\mathbb{R}^2$ so that we can work on images.

Higher dimensions

Most results discussed so far have been established in multivariate setting.

- For $f \in L_1(\mathbb{R}^d)$ and $a \in \ell_2(\mathbb{Z}^d)$ let

$$\widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-i\xi\cdot x}\, dx \quad \text{and} \quad \widehat{a}(\xi) := \sum_{k\in\mathbb{Z}^d} a[k] e^{-ik\cdot\xi}, \quad \xi \in \mathbb{R}^d.$$

- $\phi \in L_2(\mathbb{R}^d)$ is called refinable if

$$\phi = 2^d \sum_{k\in\mathbb{Z}^d} h_0[k]\phi(2\cdot -k)$$

  for some $h_0 \in \ell_2(\mathbb{Z}^d)$ or equivalently if $\widehat{\phi}(2\cdot) = \widehat{h}_0\widehat{\phi}$.

Replacing $\mathbb{Z}$ by $\mathbb{Z}^d$ the results on MRAs then carry over verbatim from univariate $L_2(\mathbb{R})$ to multivariate setting $L_2(\mathbb{R}^d)$.

Higher dimensions

- Adjust normalization factor in dilation operator, i.e.

$$\psi_{\ell,n,k} := \mathcal{D}^n T_k \psi_\ell := 2^{nd/2}\psi_\ell(2^n \cdot -k)$$

  Then the unitary extension principle generalizes as follows:

Theorem (multivariate UEP) Let $\phi \in L_2(\mathbb{R}^d)$ compactly supported and refinable with finite mask h_0 and $\widehat{\phi}(0) = 1$. Let $\Psi = \{\psi_1, \ldots, \psi_r\}$ be defined by the finitely supported masks $h_1, \ldots, h_r$. Then $X(\Psi)$, as well as its shift-invariant oversampling, are tight frames for $L_2(\mathbb{R}^d)$ provided

$$\sum_{\ell=0}^{r} |\widehat{h}_\ell(\xi)|^2 = 1 \quad \text{and} \quad \sum_{\ell=0}^{r} \widehat{h}_\ell(\xi)\overline{\widehat{h}_\ell(\xi+\nu)} = 0$$

holds for all $\xi \in [-\pi, \pi]^d$ and $\nu \in \{0, \pi\}^d$.
Furthermore, if $r = 2^d - 1$ and $\|\phi\| = 1$, then $X(\Psi)$ is an ONB of $L_2(\mathbb{R}^d)$.

Tight frames for 2D via tensor products

- Given univariate masks $\{h_\ell \colon \ell = 0, 1, \ldots, r\}$ define 2D masks

$$h_{i,j}[k_1, k_2] := h_i[k_1] h_j[k_2] \quad \text{for } 0 \le i, j \le r;\ (k_1, k_2) \in \mathbb{Z}^2.$$

- Then

$$\psi_{i,j}(x, y) := \psi_i(x) \psi_j(y) \quad \text{for } 0 \le i, j \le r;\ (x, y) \in \mathbb{R}^2.$$

  These are $(r+1)^2 - 1$ wavelets and the refinable function $\psi_{0,0}$.

For univariate masks constructed from UEP, this tensor product construction satisfies 2D UEP and thus yields tight frame for $L_2(\mathbb{R}^2)$.

Fast algorithm for 2D tensor products

Due to tensor product construction the two directions can be seperated

For the stationary case this means:

$$(D^x_{\ell,j}v)[\cdot,k_2] := \overline{h_{\ell,j}[-\cdot]} * v[\cdot,k_2]; \quad (D^{x*}_{\ell,j}v)[\cdot,k_2] := h_{\ell,j} * v[\cdot,k_2]$$
$$(D^y_{\ell,j}v)[k_1,\cdot] := \overline{h_{\ell,j}[-\cdot]} * v[k_1,\cdot]; \quad (D^{y*}_{\ell,j}v)[k_1,\cdot] := h_{\ell,j} * v[k_1,\cdot]$$

Let $v \in \mathbb{R}^{N_1 \times N_2}$ and L the number of levels. Denote $v_{0,0,0} := v$.

- **Decomposition:** For each $j = 1,2,\ldots,L$
  - Obtain low frequency approximation to v at level j

$$v_{0,0,j} = D^y_{0,1-j}(D^x_{0,1-j}v_{0,0,j-1})$$

  - Obtain wavelet coefficients of v at level j

$$v_{\ell_1,\ell_2,j} = D^y_{\ell_2,1-j}(D^x_{\ell_1,1-j}v_{0,0,j-1}) \quad \text{for } \ell_1,\ell_2 = 1,2,\ldots,r;\ (\ell_1,\ell_2) \neq (0,0)$$

- **Reconstruction:** For each $j = L, L-1, \ldots, 1$

$$v_{0,0,j-1} = \sum_{\ell_1=0}^{r}\sum_{\ell_2=0}^{r} D^{x*}_{\ell_1,1-j}(D^{y*}_{\ell_2,1-j}v_{\ell_1,\ell_2,j})$$

Wavelet wishlist

... one theoretical result on sparsity of images.

What to look for in a wavelet?

Competing attributes

- Smoothness
  - Errors less visible
- Vanishing moments
  - $\int x^m \psi(x)\, dx = 0$ for $m = 0, \ldots, M$
- Short support
  - Numerical performance
  - Detecting singularities
- Symmetry
  - Visual system more tolerant to symmetric errors
  - Boundary issues
- Directions in multivariate case
  - Tensor product emphasizes coordinate directions

Getting better in one usually means loosing in another.

Vanishing moments and sparsity

- Vanishing moments impact algorithms that involve manipulation of the wavelet coefficients.

Definition: Compactly supported ψ has **vanishing moments of order** m if

$$\int x^k \psi(x)\, dx = 0, \quad \forall k = 0, \ldots, m-1.$$

- This means polynomials up to order $m-1$ have zero wavelet coeffis (i.e. more complicating functions can be represented by refinable function alone).

Theorem: If $\psi \in L_2(\mathbb{R})$ compactly supported with vanishing moments of order m, then

$$\langle f, \psi_{n,k} \rangle = O(2^{-n(m+1/2)})$$

for all $f \in L_2(\mathbb{R}) \cap C^m(\mathbb{R})$ and n sufficiently large.

Proof: Use Taylor formula

Vanishing moments and sparsity

Remarks

- Thus majority of wavelet coefficients not near singularities can be small for piecewise smooth f.
- Thus if all wavelets have high order vanishing moments and short support, then the wavelet frame can provide sparse approximation for piecewise smooth functions. Many natural images are piecewise smooth.
- However, support size and vanishing moments are usually competing attributes and in applications one has to balance between them.