

Matching pursuits with time-frequency dictionaries

A.M. Alvarez-Meza, Ph.D.

A.A. Orozco-Gutierrez, Ph.D.

`andres.alvarez1;aaog@{utp.edu.co}`

Grupo de automática
Universidad Tecnológica de Pereira



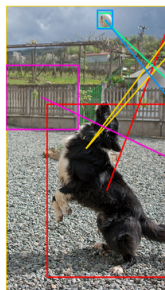
Outline

- 1 Signal representation
- 2 Time-frequency atomic decomposition
- 3 Matching pursuit in Hilbert spaces
- 4 Matching pursuit with time-frequency dictionaries
- 5 Matching pursuit with Gabor atoms
- 6 Matlab toolkits and exercises

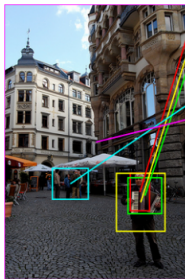
Signal representation

- Natural languages have large vocabularies (words with close meanings).
- Low level signal representations provide explicit information.
- Numerical parameters should offer compact characterizations of the elements we are looking for.
- A suitable representation gives simple cues to differentiate patterns!

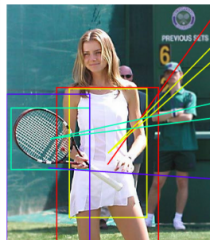
Signal representation



1.31 dog
0.31 plays
0.45 catch
-0.02 with
0.25 white
1.62 ball
-0.10 near
-0.07 wooden
0.22 fence



0.26 man
0.31 playing
1.51 accordion
-0.07 among
-0.08 in
0.42 public
0.30 area



1.12 woman
-0.28 in
1.23 white
1.45 dress
0.06 standing
-0.13 with
3.58 tennis
1.81 racket
0.06 two
0.05 people
-0.14 in
0.30 green
-0.09 behind
-0.14 her

source: Computer vision example - Stanford vision Lab

The generalized Fourier series

Let Ψ be a set of square-integrable functions in \mathcal{F}

$$\Psi = \{\phi_n : [a, b] \rightarrow \mathcal{F}\}_{n=0}^{\infty},$$

which are pairwise orthogonal for the inner product

$$\langle f, g \rangle = \int_a^b f(x)g^*(x)w(x)dx$$

$w(x)$: weight function.

The generalized Fourier series with respect to Ψ , is then

$$f(x) \sim \sum_{n=0}^{\infty} c_n \phi_n(x) \quad (1)$$

where the coefficients are given by

$$c_n = \frac{\langle f, \phi_n \rangle}{\|\phi_n\|_w^2}$$

If Ψ is a complete set, i.e., an orthonormal basis on $[a, b]$, \sim becomes $=$ in the L^2 sense.

The Fourier transform

Let $f(t) \in L^2(\mathbb{R})$ be a function in the Hilbert space $L^2(\mathbb{R})$, the Fourier transform $\hat{f}(\omega)$ is defined by:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt. \quad (2)$$

The inverse Fourier transform is defined as:

$$\hat{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega.$$

The Short Time Fourier transform

- Employs a window function to localize the complex sinusoid as:

$$F(\omega, u) = \int_{-\infty}^{\infty} f(t)h(t-u)e^{-i\omega t}dt \quad (3)$$

- $u \in \mathbb{R}$ is a translation parameter and $h(t-u)$ is the window function which confines the complex sinusoid $e^{-i\omega t}$.
- Some Window shapes: Hanning, hamming, cosine, Gaussian, etc.
- Gaussian windowed STFT (Gabor transform):

$$h_{\xi,u}(t) = h(t-u)e^{i2\pi\xi t}, \quad (4)$$

$$h(t) = \frac{1}{\sqrt{\sigma}\pi^{1/4}}e^{-\frac{1}{2}(t^2/\sigma^2)}, \sigma \in \mathbb{R}^+, \text{ and } \xi \in \mathbb{R}: \text{ modulating factor.}$$

- $h_{\xi,u}(t)$ is known as a windowed Fourier atom.

The Short Time Fourier transform

- Convolution of the atom with the signal:

$$\begin{aligned} F(\xi, u) &= \int_{-\infty}^{\infty} f(t) h_{\xi, u}^*(t) dt \\ &= \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{\sigma} \pi^{1/4}} e^{\frac{-1}{2}[(t-u)^2/\sigma^2]} e^{-i2\pi\xi t} dt \end{aligned} \quad (5)$$

The Wavelet transform

- To match a signal by moving and stretching/squeezing a wavelet function.

$$T(s, u) = \int_{-\infty}^{\infty} f(t) \psi_{s,u}^*(t) dt \quad (6)$$

- Example: Morlet wavelet $\psi(t) = \frac{1}{\pi^{1/4}} e^{i2\pi\xi_0 t} e^{-t^2/2}$, ξ_0 : central frequency.
- The Morlet wavelet transform gives:

$$T(s, u) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}\pi^{1/4}} e^{\frac{-1}{2}[(t-u)^2/s^2]} e^{-i2\pi(\xi_0/s)(t-u)} dt \quad (7)$$

The Time-frequency atomic decomposition

- To **extract information from complex signals**, we need to adapt the time-frequency decomposition to the particular signal structures.
- A family of time-frequency atoms can be generated by scaling, translating, and modulating a single window function $g(t) \in L^2(R)$.
- $g(t)$ is continuously differentiable and $\|g(t)\|=1$, $g(0) \neq 0$.
- For any scale $s > 0$, frequency modulating ξ , and translation u , we denote $\gamma=(s, u, \xi)$ and define:

$$g_{\gamma}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t} \quad (8)$$

where $\frac{1}{\sqrt{s}}$ normalizes $\|g_{\gamma}(t)\| = 1$.

The Time-frequency atomic decomposition

- The Fourier transform of $g(t)$ yields:

$$g_{\gamma}(\omega) = \sqrt{s} g(s(\omega - \xi)) e^{-i(\omega - \xi)u}, \quad (9)$$

since $|g(\omega)|$ is even, $g_{\gamma}(\omega)$ is centered at frequency $\omega = \xi$.

- The family $\mathcal{D} = \{g_{\gamma}(t)\}_{\gamma \in \Gamma}$ is redundant ($\Gamma = \mathbb{R}^+ \times \mathbb{R}^2$).
- To represent properly $f(t)$ we must select a subset of atoms $\{g_{\gamma_n}(t)\}_{n \in N}$, with $\gamma_n = (s_n, u_n, \xi_n)$, so that:

$$f(t) = \sum_{n=-\infty}^{\infty} a_n g_{\gamma_n}(t) \quad (10)$$

The Fourier and wavelet transforms correspond to different families of time-frequency atoms

The Fourier transform's atoms

$$F(\xi_n, u_n) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{\sigma} \pi^{1/4}} e^{\frac{-1}{2}[(t-u_n)^2/\sigma^2]} e^{-i2\pi \xi_n t} dt \quad (11)$$

- All the atoms have a constant scale $s_n = s_0$ ($s_0 \sim \sigma$ for Gabor window).
- If the main signal structures are localized over a time-scale of the order of s_0 and a_n gives important insights on their localization and frequency content.
- Window Fourier transform is not well adapted to describe structures that are much smaller or much larger than s_0 .
- To analyze components of varying sizes, it is necessary to use time-frequency atoms of different scales.

The wavelet transform's atoms

$$T(s_n, u_n) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s_n} \pi^{1/4}} e^{\frac{-1}{2}[(t-u_n)^2/s_n^2]} e^{-i2\pi(\xi_0/s_n)(t-u_n)} dt \quad (12)$$

- Decomposes signals over time-frequency atoms of varying scales, called wavelets.
- The resulting family is composed of dilations and translations of a single function, multiplied by complex phase parameter.
- The expansion coefficients a_n characterize the scaling behavior of signal structures.
- However, wavelet coefficients do not provide precise estimates of the frequency content of waveforms whose Fourier transforms is well localized (restriction on the frequency parameter $\xi_n = \xi_0/s_n$).

Matching pursuit in Hilbert spaces

- We want to compute a linear expansion of $f \in L^2(R)$ over a set of vectors selected from $\mathcal{D} = (g_\gamma)_{\gamma \in \Gamma}$ to best match its inner structures.
- Procedure: successive approximations of f with orthogonal projections on elements of \mathcal{D} .
- Let $g_{\gamma_0} \in \mathcal{D}$, then:

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + Rf, \quad (13)$$

where Rf is the residual vector.

- Since g_{γ_0} is orthogonal to Rf (Exercise: proof it):

$$\|f\|^2 = \|\langle f, g_{\gamma_0} \rangle\|^2 + \|Rf\|^2. \quad (14)$$

Matching pursuit in Hilbert spaces

- To minimize $\|Rf\|$ we must choose $g_{\gamma_0} \in \mathcal{D}$ such that $|\langle f, g_{\gamma_0} \rangle|$ is maximum.
- It is possible to find a vector that is almost the best:

$$|\langle f, g_{\gamma_0} \rangle| \geq \alpha \sup_{\gamma \in \Gamma} |\langle f, g_{\gamma} \rangle|, \quad (15)$$

where $0 < \alpha \leq 1$.

- Matching pursuit is an iterative algorithm that subdecomposes the residue Rf by projecting it on a vector of \mathcal{D} that matches Rf almost at best.

Matching pursuit in Hilbert spaces

- Energy conservation equation:

$$\|f\|^2 = \sum_{n=0}^{m-1} |\langle R^n f, g_{\gamma_n} \rangle|^2 + \|R^m f\|^2. \quad (16)$$

- In practice, the algorithm is terminated either when the residual energy below a preset cut-off level:

$$\|R^n f\|^2 < \epsilon \|f\|^2; \quad \forall \epsilon \in \mathbb{R}^+. \quad (17)$$

- The original f is decomposed into a sum of dictionary elements, that are chosen to best match its residues.
- The decomposition is nonlinear, however, we maintain an energy conservation as if it was a linear orthogonal decomposition.
- Matching pursuit is a greedy algorithm that chooses at each iteration a waveform that is best adapted to approximate part of the signal.

Matching pursuit with time-frequency dictionaries

- For dictionaries of time-frequency atoms a matching yields an adaptive time-frequency transform.
- The function f is decomposed into a sum of complex time-frequency atoms that best match its residues.
- The matching pursuit decomposes f as:

$$f = \sum_{n=0}^{+\infty} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n}, \quad (18)$$

where $\gamma_n = (s_n, u_n, \xi_n)$ and

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{s_n}} g\left(\frac{t - u_n}{s_n}\right) e^{i\xi_n t}. \quad (19)$$

Matching pursuit with time-frequency dictionaries

- At each iteration the MP algorithm selects a vector g_{γ_n} that satisfies:

$$|\langle R^n f, g_{\gamma_n} \rangle| \geq \alpha \sup_{\gamma \in \Gamma} |\langle R^n f, g_{\gamma} \rangle|. \quad (20)$$

- From the decomposition of any $f(t)$ within a time-frequency dictionary a new time-frequency energy distribution is obtained by adding the Wigner distribution of each selected atom.
- The cross Wigner distribution of two functions $f(t)$ and $h(t)$ is defined by:

$$W[f, h](t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f\left(t + \frac{\tau}{2}\right) h^*\left(t - \frac{\tau}{2}\right) e^{-i\omega\tau} d\tau. \quad (21)$$

where $Wf(t, \omega) = W[f, f](t, \omega)$.

Matching pursuit with time-frequency dictionaries

- A picture of the energy distribution of $f(t)$ in the time-frequency plane can be defined as ([Exercise: Proof it](#)):

$$Ef(t, \omega) = \sum_{n=0}^{\infty} |\langle R^n f, g_{\gamma_n} \rangle|^2 W g_{\gamma_n}(t, \omega). \quad (22)$$

Matching pursuit with Gabor atoms

- The signal approximation reconstructed from N MP expansion coefficients is given by:

$$f_N(t) = \sum_{n=0}^{N-1} M_n h_{s_n, u_n, \xi_n}(t). \quad (23)$$

- The Gabor atom for the MP method is defined as:

$$h_{s_n, u_n, \xi_n}(t) = \frac{1}{\sqrt{s_n}} h\left(\frac{t - u_n}{s_n}\right) \quad (24)$$

Matching pursuit with Gabor atoms

- A Gaussian window is defined for the MP as:

$$h(t) = 2^{1/4} e^{-\pi t^2} \quad (25)$$

- The Gabor atom for MP can be written as:

$$h_{s_n, u_n, \xi_n, \phi_n}(t) = K_n \frac{2^{1/4}}{\sqrt{s_n}} e^{-\pi[(t-u_n)/s_n]^2} \cos(2\pi\xi_n t + \phi_n) \quad (26)$$

s_n : scale, u_n : localization factor for Gaussian envelope, ξ_n : frequency, ϕ_n : phase of the real sinusoid, K_n : normalization factor to preserve unit energy.

- The expansion coefficients are determined as follows:

$$M(s_n, u_n, \xi_n, \phi_n) = \int_{-\infty}^{\infty} f(t) K_n \frac{2^{1/4}}{\sqrt{s_n}} e^{-\pi[(t-u_n)/s_n]^2} \cos(-2\pi\xi_n t + \phi_n) dt \quad (27)$$

Summary: playing with signals (Gabor atoms example)

- Fourier:

$$F(\xi_n) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi \xi_n t} dt. \quad (28)$$

- Short-time Fourier transform:

$$F(\xi_n, u_n) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{\sigma} \pi^{1/4}} e^{\frac{-1}{2}[(t-u_n)^2/\sigma^2]} e^{-i2\pi \xi_n t} dt \quad (29)$$

- Wavelet transform:

$$T(s_n, u_n) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s_n} \pi^{1/4}} e^{\frac{-1}{2}[(t-u_n)^2/s_n^2]} e^{-i2\pi(\xi_0/s_n)(t-u_n)} dt \quad (30)$$

- Matching pursuit transform:

$$M(s_n, u_n, \xi_n, \phi_n) = \int_{-\infty}^{\infty} f(t) K_n \frac{2^{1/4}}{\sqrt{s_n}} e^{-\pi[(t-u_n)/s_n]^2} \cos(-2\pi \xi_n t + \phi_n) dt \quad (31)$$

- Fourier: `fft.m`
- Short-time Fourier transform: `spectrogram.m`
- Wavelet transform
 - Continuous wavelet transform: `cwt.m`
 - Discrete wavelet transform: `dwt.m`
- Matching pursuit: `wmpalg.m`

- Summarize the main properties of the following time-frequency decompositions: Fourier, Short-time Fourier, Wavelet, and MP.
- Proof equation (14) (orthogonal property for MP in Hilbert spaces).
- Proof equation (22) (energy estimation in MP using the Wigner distribution).
- Test the Fourier, STFT, wavelet, and MP transforms over a non-stationary signal (see `wavemenu` Matlab function).
- Discuss and conclude about the obtain results.

References I



Addison, P. S. (2002).

The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance.

CRC press.



Mallat, S. G. and Zhang, Z. (1993).

Matching pursuits with time-frequency dictionaries.

IEEE Transactions on signal processing, 41(12):3397–3415.