Contents

Chapter 5

Choosing the Representation

This chapter introduces adaptive time-frequency approximation of signals and the matching pursuit algorithm.

Each of the dictionaries discussed in previous chapters efficiently represents some kind of structures. Spectrogram (Section 4.1) describes oscillations with chosen time resolution, wavelets (Section 4.3) zoom nicely on singularities, and so on. In general, it is very hard to guess which dictionary would provide an optimal representation for a given signal. On the other hand, efficient and informative decomposition can be achieved only in a dictionary containing functions reflecting the structure of the analyzed signal. This rule, illustrated by numerous examples from previous sections, is hardly a surprise: *The limits of my language mean the limits of my world.*[1]

So, why don't we extend the limits—by constructing a big dictionary, rich enough to fit all the structures, possibly occurring in any signal of interest?

5.1 GABOR DICTIONARY

The most common approach to the construction of time-frequency dictionaries relies on Gabor functions, that is Gaussian envelopes modulated by sine oscillations. By multiplying these two functions, we can obtain a wide variety of shapes, depending on their parameters (Figures 5.1 and 5.2).

1 Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, thesis 5.6.

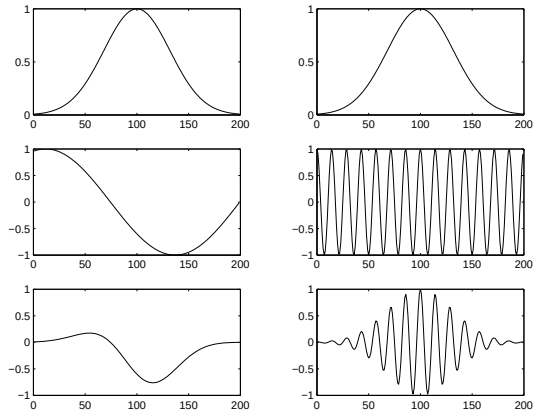


Figure 5.1 Gabor functions (bottom row) are constructed by multiplying Gaussian envelopes (upper row) with oscillations of different frequencies and phases (middle row).

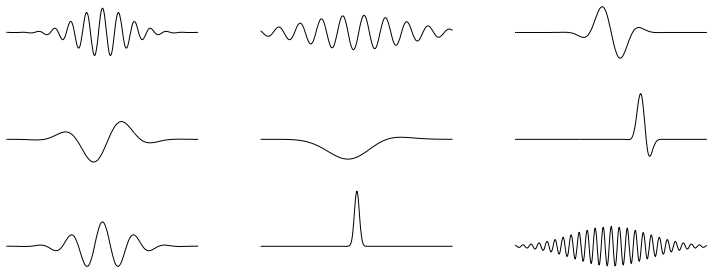


Figure 5.2 Examples of Gabor functions.

The advantage of this standardized approach is that all these different shapes, such as those presented in Figure 5.2 and many more,[2] can be described in terms of only four numbers per waveform: time width and position of the center of the Gaussian envelope, and frequency and phase of the modulating sine. Amplitudes are adjusted so that each function has equal (unit) energy, since the product of a waveform of unit energy with the signal will directly measure the contribution of that structure to the energy of the signal (product and energy were discussed in Section 2.1).

2 For example, pure sine waves and impulse functions can be treated as sines with very wide modulating Gauss and very narrow Gaussians, respectively.

Apart from the variety of shapes, there is a purely mathematical argument in favor of Gabor functions: they provide the best localization (lowest uncertainty, see Section 3.2) in the time-frequency plane [1]. However, this property is not crucial for understanding the procedure; the algorithm discussed in the next section can be used with any dense dictionary—that is, any dictionary containing at least enough functions to reproduce (efficiently or not) any signal.

5.2 ADAPTIVE APPROXIMATION

OK, we have a dictionary incorporating a variety of structures—how shall we use it? Following the classic approach described in the previous chapter for spectrogram and wavelets, we could try to use the products of all the functions from the dictionary with the signal. But in this case such representation would make no sense. If we use the whole dictionary of, say, millions of waveforms to describe 10 seconds of an EEG signal, we do not gain anything, either in terms of compression or understanding the signal structure. We must *choose* the representative waveforms. And if the choice will be adapted for each signal separately—unlike the a priori selections from Chapter 4—we may obtain a general, efficient, and adaptive procedure.

Matching pursuit is nothing but a procedure leading to such a choice. We will exemplify its operation on a signal composed of two Gabor functions and a little noise (Figure 5.3).

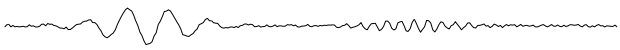


Figure 5.3 Signal constructed as a sum of two Gabor functions plus noise.

As the first candidate for the representation, we take the function, which, among all the functions from the dictionary, gives the largest product with the signal (Figure 5.4).

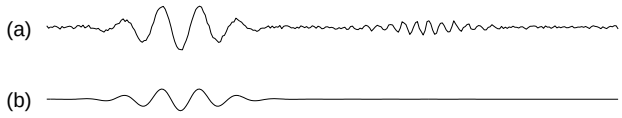


Figure 5.4 Function from the Gabor dictionary (b), giving the largest product with the signal (a).

However, a single function seldom explains exactly the whole signal. So, which function do we take as the second representative? The one giving the second largest product?

If the dictionary is rich, then it must contain many similar waveforms. One of them gives "the best fit" of the strongest structure. But the other, similar waveforms from the dictionary will most likely also fit the same structure—not "best," but still giving a large product. This product can be higher than the product of another, smaller structure, matched perfectly with a different waveform from the dictionary (Figure 5.5 (e)). Therefore, choosing *all* the waveforms that give a large product with the signal may result in a representation containing many similar waveforms, all approximating only the strongest structure of the signal—like (b), (c), and (d) from Figure 5.5—and completely omitting weaker structures, like (e).

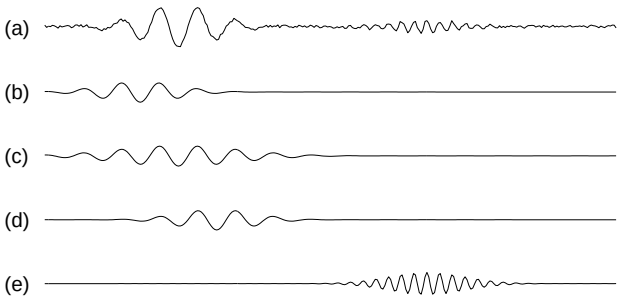


Figure 5.5 Similar functions from the dictionary (b–d) giving a large product with the signal (a). All these functions, more or less similar to the stronger structure, give larger products with the signal than the function exactly reflecting the weaker structure (e).

This is the price we pay for using a redundant dictionary. Smaller dictionaries, used in wavelet transform or STFT, are chosen so that their functions have little possible overlap—in the case of an orthogonal basis, the overlap is zero. If the dictionary were not so redundant, at most one of the functions (b), (c), and (d) in Figure 5.5 would be present. In such a case, we could use for the representation all the functions giving large products with the signal. But if the dictionary is so redundant, we must subtract the contribution of the first chosen function before fitting the next one.

5.3 MATCHING PURSUIT

Matching pursuit (MP) algorithm was first proposed in the context of signal analysis in 1993 by Mallat and Zhang in [2].[3] It is an iterative procedure, which can be described as follows:

1. Find (in the dictionary) the first function, that best fits the signal.
2. Substract its contribution from the signal.
3. Repeat these steps on the remaining residuals, until the representation of the signal in terms of chosen functions is satisfactory.

The first two iterations of this procedure, applied to the signal from Figure 5.3, are illustrated in Figure 5.6.

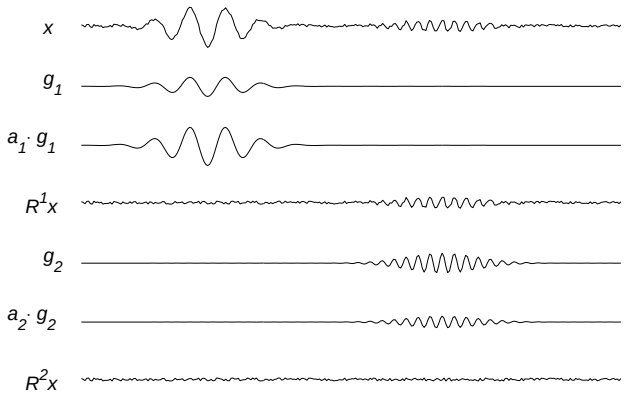
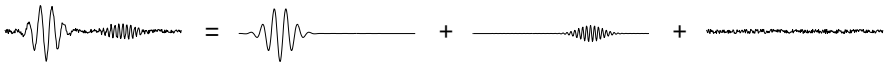


Figure 5.6 Matching pursuit algorithm. In the first step, we find the function g_1, which gives the largest product with the analyzed signal (x, upper trace). Then we adjust the amplitude of g_1 to the structure present in the signal and subtract it from the signal. The resulting (first) residual R^1x does not contain the contribution explained by the first fitted function. Therefore, the next function (g_2), found in the dictionary as giving the largest product with the residual, will fit the next structure present in the signal. If these two functions (g_1 and g_2) give a satisfactory representation of the signal x (e.g., explains the required percentage of energy), we can stop the procedure at this point and leave R^2x as the unexplained residual.

3 A similar approach to signals was proposed in [3]. The general idea was known previously from statistics as a regression pursuit.

As a result, we obtain an approximation of the signal x in terms of the functions g_1 and g_2 and their amplitudes a_1 and a_2, plus the unexplained residual:



or

$$x = a_1 \cdot g_1 + a_2 \cdot g_2 + R^2 x \ ,$$

where $R^2 x$ denotes the residual of signal x left after the second iteration.

5.4 TIME-FREQUENCY ENERGY DENSITY

So, we have a nice and compact description of the signal in terms of a sum of known functions. What about the picture of its time-frequency energy density?

First, let us recall the meaning of such a picture. As an image, it naturally has two dimensions. The horizontal dimension corresponds to time, and its extent reflects the length of the analyzed epoch. Vertical dimension represents frequency and may extend from zero to the half of the sampling frequency (i.e., the Nyquist frequency).[4]

Graphically we represent the energy density as shades of gray, color scale, or height in 3-D plots. If the signal has a high-energy activity of a given frequency in a given time epoch, we expect the corresponding area of the picture to exhibit high values of energy density. In the previous chapters we used a fixed set of functions (e.g., sines or wavelets), so we could a priori divide the time-frequency plane into boxes corresponding to these functions. Examples of such pictures are given in Figures 4.1 and 4.2 (spectrograms) and 4.7 and 4.8 (orthogonal wavelet transforms).

Using the matching pursuit expansion, we do not know a priori which functions will be chosen for the representation. The decomposition is adaptive, so we cannot draw a prior division of the time-frequency plane like in Figures 4.3 or 4.6. But for each of the functions, chosen for the representation of the signal, we can compute the corresponding time-frequency distribution of energy density, by means of the Wigner transform (Section 4.4). For the Gabor functions from the dictionary discussed in Section 5.1, we obtain blobs in the time-frequency plane extending in time approximately in the regions of the represented structures.[5] Their frequency

4 We recall from Section 1.2 that Nyquist frequency is the maximum frequency which can be reliably detected in a digital signal.

5 This ellipsoidal blob is actually a 2-D Gauss function.

extent is determined by the uncertainty principle (Section 3.2), which states that the shorter the structure is in time, the wider is its frequency content. An impulse can be considered an extremely short structure, so it will be represented as a vertical line. Infinite sine will be infinitely narrow in frequency, so it will make a horizontal line. By adding energy densities for all the functions from the decomposition, we obtain representations like in Figure 5.7.

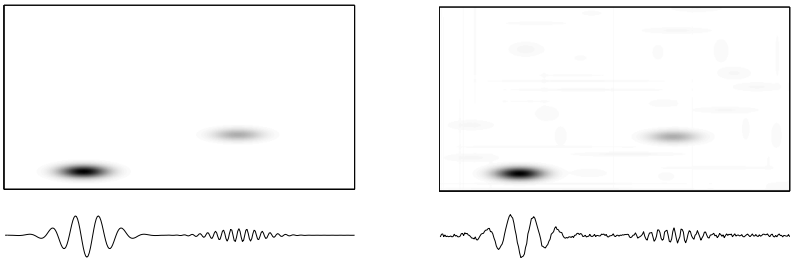


Figure 5.7 Left: sum of the first two functions (g_1 and g_2 from Figure 5.6) and its time-frequency energy density. Right: time-frequency representation of the signal from Figure 5.3, including noise. Apart from the same two structures representing the major components of the signal, we observe a lot of weaker blobs distributed uniformly across the time-frequency plane, representing the white noise added in simulation to the two structures. Vertical axis corresponds to frequency, increasing upwards.

But what do we actually gain from this approach, compared to the direct calculation of the Wigner transform of the signal, as in Section 4.4? From the matching pursuit decomposition, we know (or assume) that the signal is a sum $g_1 + g_2$. Using this information, we do not have to take the Wigner transform of the whole sum (i.e., the whole signal), which would give $g_1^2 + g_2^2 + 2g_1g_2$, but instead we take only $g_1^2 + g_2^2$, that is the sum of energy densities of the components, not including cross-terms. The problem of cross-terms is exemplified in Figure 4.9.

Figure 5.8 presents decompositions of a bit more complex signal, constructed from a continuous sine wave, one-point impulse (Dirac's delta), and three Gabor functions. Panel (b) gives time-frequency energy distribution obtained for this signal from MP decomposition. In the left, three-dimensional plots, energy is proportional to the height. In the right two-dimensional maps, energy is coded in shades of gray. Panels (c) and (d) present decompositions of the same signal with addition of a white noise of energy twice and four times the signal's energy. We observe basically retained representation of the major signal structures (slightly disturbed in the presence of the stronger noise) and a uniform distribution of the noise-related, weaker structures.

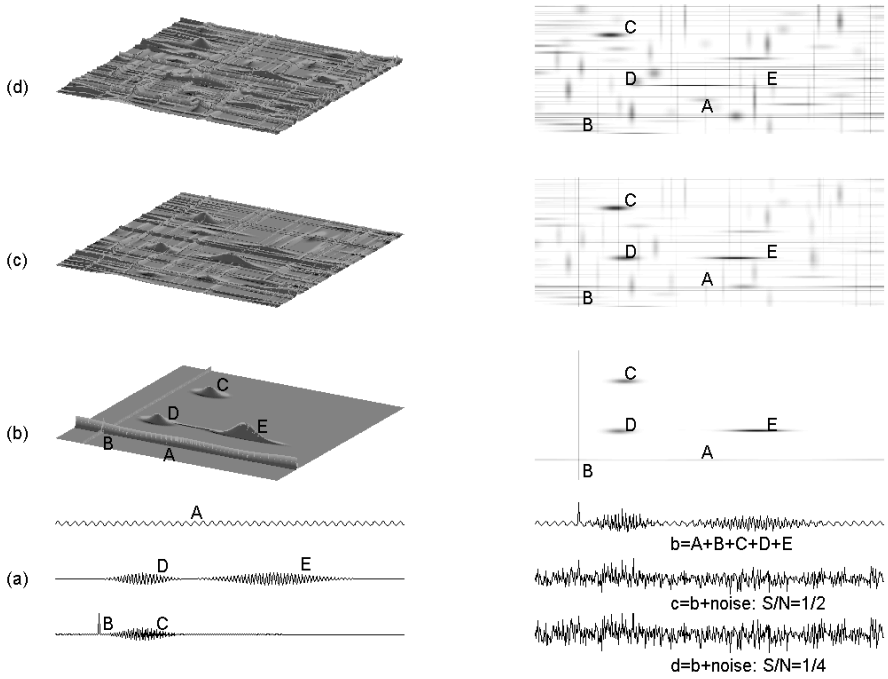


Figure 5.8 (a): Left—components of the simulated signal: sine A, Dirac's delta B, and Gabor functions C, D, and E. Right—signals, labelled b, c, and d, constructed as sum of structures A–E and white noise, and decomposed in corresponding panels (b), (c), and (d). (b): Time-frequency energy density obtained for sum of structures A–E; in 3-D representation on the left, energy is proportional to the height, while in the right panel it is proportional to the shades of gray. Panels (c) and (d): decompositions of signals with linear addition of noise, S/N = 1/2 (-3 dB) in (c) and -6 dB in (d). The same realization of white noise was used in both cases. (Reprinted from [4] with permission. © 2001 by PTBUN and IBD.)

References

[1] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed., New York: Academic Press, 1999.

[2] S. Mallat and Z. Zhang, "Matching Pursuit with Time-Frequency Dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, December 1993, pp. 3397–3415.

[3] S. Qian and D. Chen, "Signal Representation Using Adaptive Normalized Gaussian Functions," *Signal Processing*, vol. 36, March 1994, pp. 1–11.

[4] K. J. Blinowska and P. J. Durka, "Unbiased High Resolution Method of EEG Analysis in Time-Frequency Space," *Acta Neurobiologiae Experimentalis*, vol. 61, 2001, pp. 157–174.

About the Author

Piotr J. Durka received an M.Sc. and a Ph.D. and habilitation in physics from the Warsaw University. His views on the mission of a physicist in biomedical sciences are coherent with Richard Feynman's speech on cargo cult science, Jon Clearbout's notion of reproducible research, and the Occam's Razor.

In 1995 Dr. Durka introduced matching pursuit to biomedical signal analysis; these early studies were among the first real-world applications of adaptive time-frequency approximations. After a decade of continued work, subsequent applications reached the critical volume needed to support his thesis about possible unification in the field of EEG analysis. In 2004 he proposed SignalML (http://signalml.org), an elegant solution to the problem of incompatibility of digital formats used for storage of biomedical time series.

His other activities include the presidency of the award-winning software company Bitlab Ltd. (1994–1997), the design of the EEG.pl neuroinformatics portal (2003), lectures on signal processing, statistics, and computer science, and writing three books. For a full list of research papers, software, and other information, see http://durka.info.

Index