

RESEARCH

Open Access

Multivariate matching pursuit in optimal Gabor dictionaries: theory and software with interface for EEG/MEG via Svarog

Rafał Kuś, Piotr Tadeusz Różański and Piotr Jerzy Durka*

*Correspondence:
durka@fuw.edu.pl
Faculty of Physics, University of
Warsaw, ul. Hoża 69, 00-681
Warszawa, Poland

Abstract

Background: Matching pursuit algorithm (MP), especially with recent multivariate extensions, offers unique advantages in analysis of EEG and MEG.

Methods: We propose a novel construction of an optimal Gabor dictionary, based upon the metrics introduced in this paper. We implement this construction in a freely available software for MP decomposition of multivariate time series, with a user friendly interface via the Svarog package (Signal Viewer, Analyzer and Recorder On GPL, <http://braintech.pl/svarog>), and provide a hands-on introduction to its application to EEG. Finally, we describe numerical and mathematical optimizations used in this implementation.

Results: Optimal Gabor dictionaries, based on the metric introduced in this paper, for the first time allowed for *a priori* assessment of maximum one-step error of the MP algorithm. Variants of multivariate MP, implemented in the accompanying software, are organized according to the mathematical properties of the algorithms, relevant in the light of EEG/MEG analysis. Some of these variants have been successfully applied to both multichannel and multitrial EEG and MEG in previous studies, improving preprocessing for EEG/MEG inverse solutions and parameterization of evoked potentials in single trials; we mention also ongoing work and possible novel applications.

Conclusions: Mathematical results presented in this paper improve our understanding of the basics of the MP algorithm. Simple introduction of its properties and advantages, together with the accompanying stable and user-friendly Open Source software package, pave the way for a widespread and reproducible analysis of multivariate EEG and MEG time series and novel applications, while retaining a high degree of compatibility with the traditional, visual analysis of EEG.

Keywords: Matching pursuit, EEG, MEG, Time-frequency, Gabor dictionary, Metrics

Background

Since the first application to EEG in 1995 [1], matching pursuit algorithm (MP) has been shown to significantly improve the EEG/MEG analysis in a variety of paradigms, including pharmaco-EEG [2,3], assessment of propagation [4], dynamics [5] and signal complexity [6-8] in epileptic seizures, detection of seizures [9,10], analysis of somatosensory evoked potentials in humans [11] and rats [12], detection of sleep spindles in Obstructive Sleep Apnea [13] and investigation of their chirping properties [14], studies of high gamma in

humans [15] and monkeys [16], investigation of brain's pain processing [17,18], parameterization of vibrotactile driving responses [19] and event-related desynchronization and synchronization [20,21].

New area of applications opened with the advent of multivariate MP (MMP) algorithms. MMP preprocessing was shown to significantly improve stability and sensitivity of EEG inverse solutions [22-27] and allowed for tracing evoked responses in single trials of EEG and MEG [28-31].

Finally, the algorithm offers also unique compatibility with the traditional, visual analysis of EEG. Specific mode of operation of MP, which is sequential focusing on locally strongest ("most visible") signal structures, resembles the working of an electroencephalographer who visually evaluates the EEG time series. Proper interpretation of the MP parameterization can provide a direct link to the results of visual analysis of EEG [32]—that is, we may find a direct correspondence between the waveforms fitted to the EEG time series, and the structures marked by visual scorer, including sleep spindles, slow waves or epileptic spikes [33-35]. This advantage should not be underestimated in the field, where most of our knowledge about behavioral and neurological correlates of EEG comes from visual analysis, which is still the only golden standard and point of reference: according to the Report of the American Academy of Neurology and the American Clinical Neurophysiology Society [36], quantitative EEG analysis *should be used only by physicians highly skilled in clinical EEG, and only as an adjunct to and in conjunction with traditional EEG interpretation*.

In spite of all these advantages, PubMed search of "matching pursuit and EEG" still returns only a two-digit number. This limited acceptance of such a promising method may be partly due to the lack of:

1. Well defined criteria for setting the most important parameters of the algorithm, which are the number and distribution of dictionary's functions.
2. Common framework for a variety of multivariate MP algorithms.
3. Robust and user friendly software based upon solid mathematical foundations.

Following sections address these issues from two different angles. Next section *An interactive tour of matching pursuit* provides plain English introduction of the major aspects and parameters of the algorithm, based on example computations. Following sections use equations to introduce optimal sampling of Gabor dictionaries and a common framework for multivariate MP (MMP), and Appendix discusses major numerical and mathematical optimizations used in the MMP implementation accompanying this paper.

An interactive tour of matching pursuit

[Additional file 1] is a video tutorial for downloading and configuration of the Svarog package, used for computations and visualization of results (Figures 1, 2, 3, 4, 5 and 6 are screenshots from Svarog). Screencast in [Additional file 2] shows the steps from loading the signal to displaying interactive map of time-frequency energy distribution. [Additional file 3] contains help of the MP module from Svarog. Complete software environment used for computations presented in this paper (including examples from the following sections) is freely available from <http://braintech.pl/svarog>, as described in section *Software availability and requirements*.

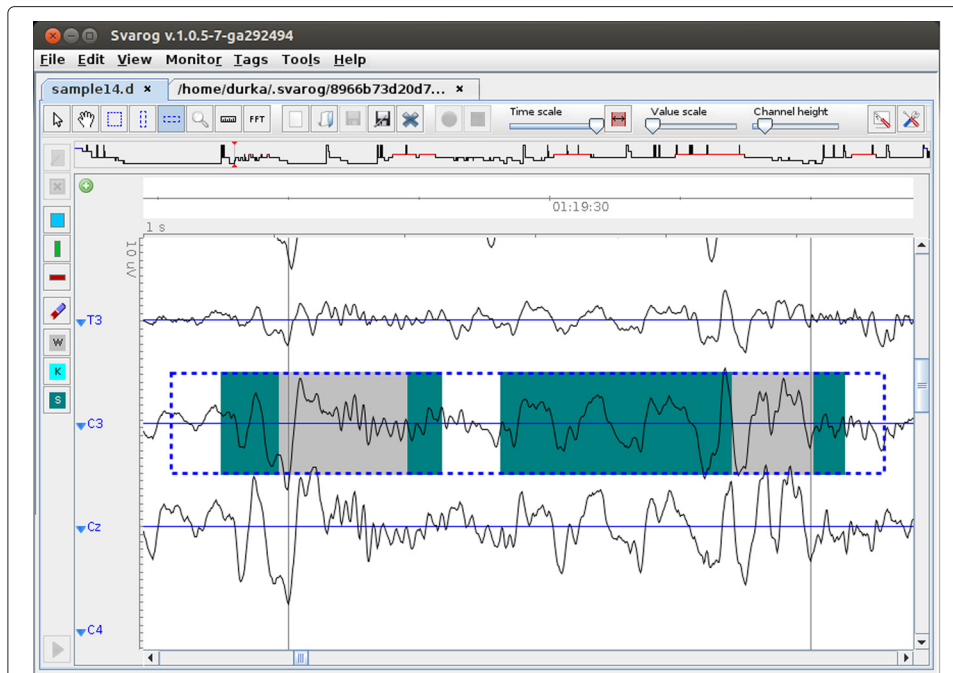


Figure 1 Sample epoch of sleep EEG. Screenshot displaying an epoch of sleep EEG recording. SWA and sleep spindles were marked by an electroencephalographer as green and gray rectangles, correspondingly (in this case both spindle tags fall inside the epochs marked as SWA). Blue dashed line outlines the epoch from C3 selected for MP decomposition, shown in Figure 2.

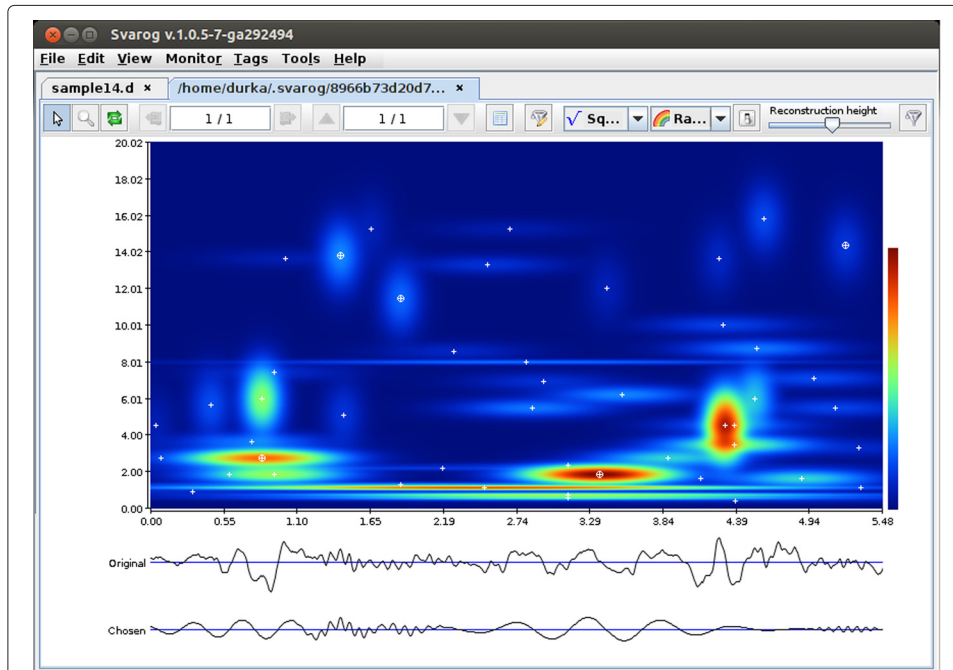


Figure 2 Time-frequency distribution of signal's energy. Results of MP decomposition displayed as an interactive time-frequency of map signal's energy density in Svarog. Clicking center of a blob (marked by white cross) adds the corresponding function to the reconstruction (bottom signal).

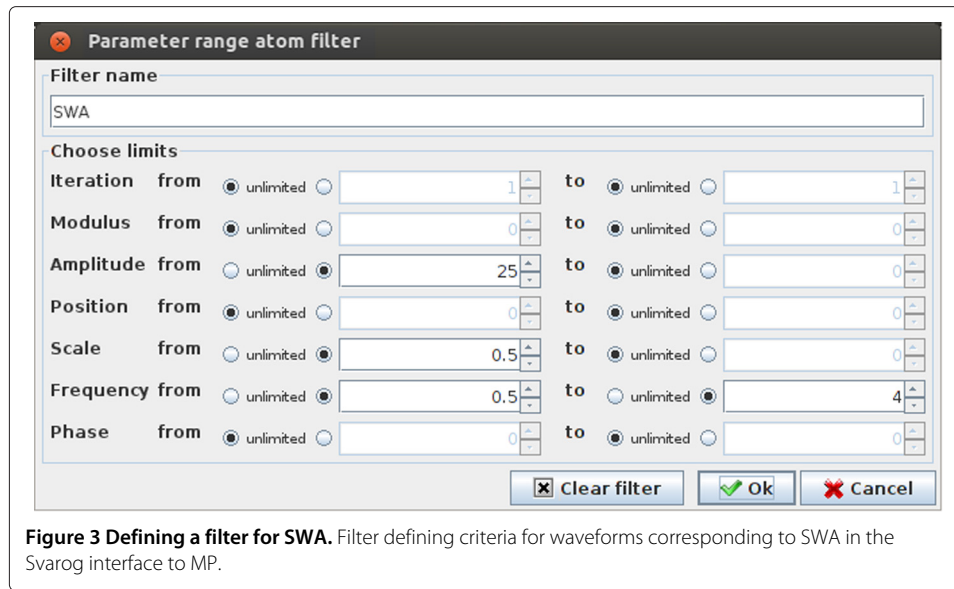


Figure 3 Defining a filter for SWA. Filter defining criteria for waveforms corresponding to SWA in the Svarog interface to MP.

The algorithm

The gist of the matching pursuit algorithm can be summarized as follows:

1. We start by creating a huge, redundant set (called a dictionary) of candidate waveforms for representation of structures possibly occurring in the signal. For the time-frequency analysis of signals we use dictionaries composed of sines with Gaussian envelopes, called Gabor functions, which reasonably represent waxing and waning of oscillations.

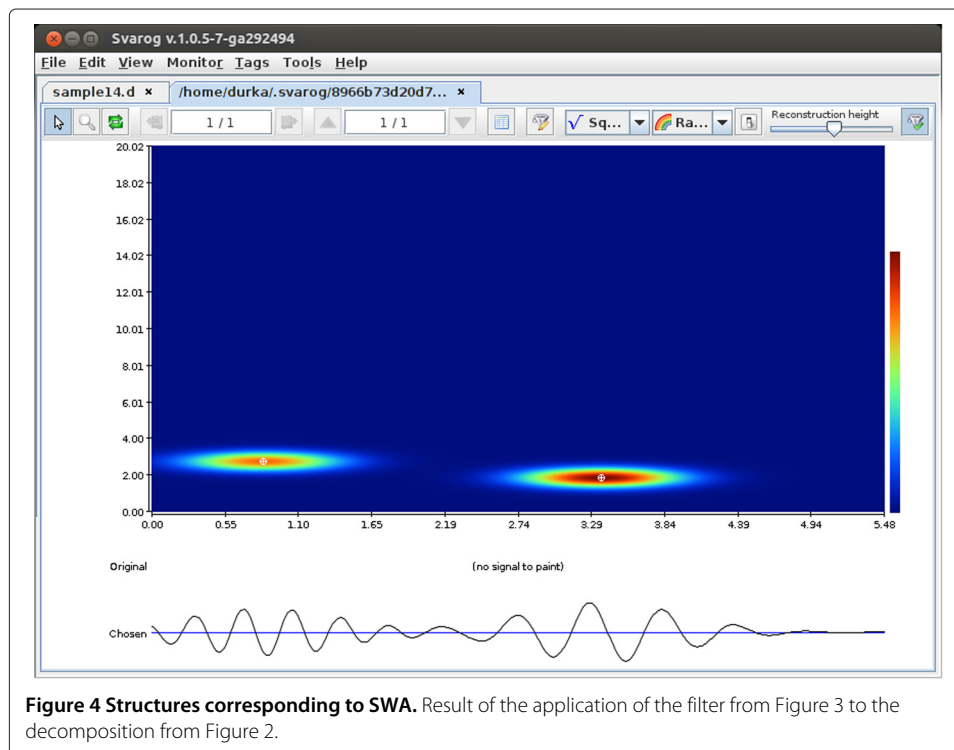
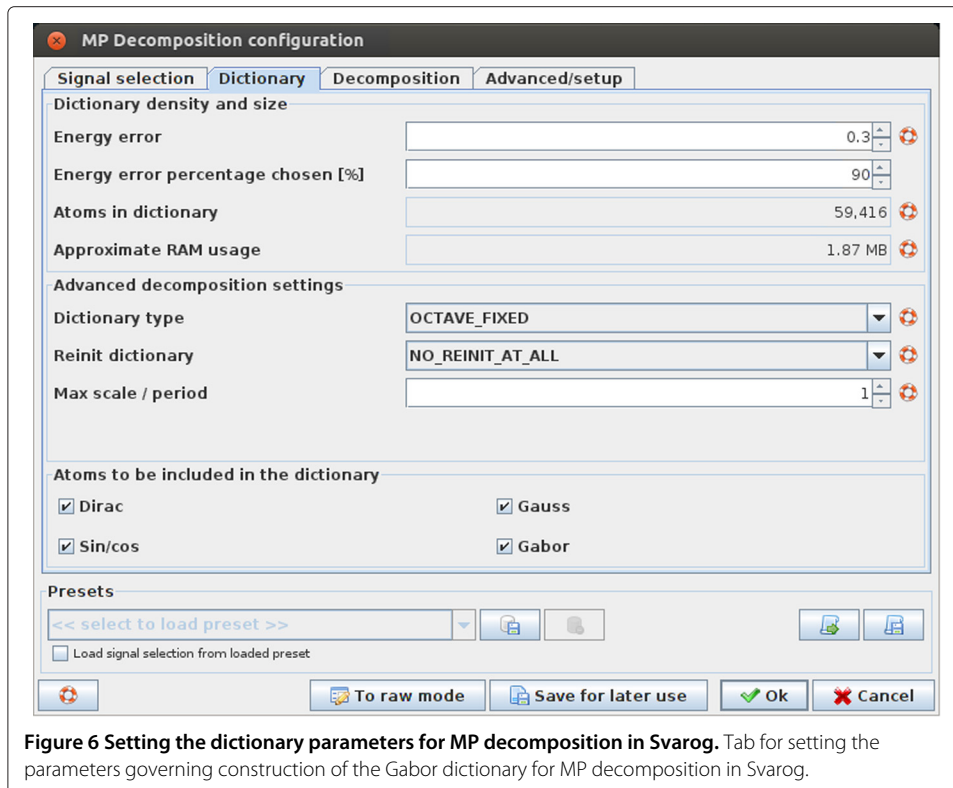
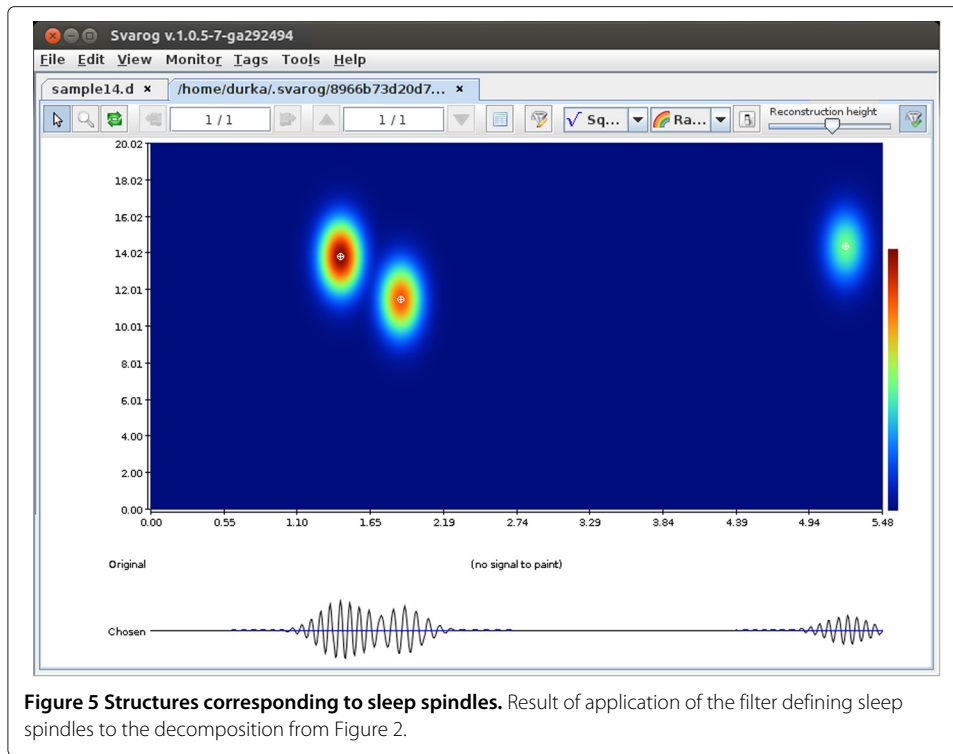


Figure 4 Structures corresponding to SWA. Result of the application of the filter from Figure 3 to the decomposition from Figure 2.



2. From this dictionary we choose only those functions, which fit the local signal structures. In such a way, the width of the analysis window is adjusted to the local properties of the signal. Local adaptivity of the procedure is somehow similar to the process of visual analysis, where an expert tends to separate local structure and assess their characteristics. Owing to this local adaptivity, MP is the only signal processing method returning explicit time span of detected structures.
3. The above idea is implemented in an iterative procedure: in each step we find one “best” function, and then subtract it from the signal being decomposed in the following steps.

MP advantages in EEG

We will discuss some of the advantages of MP in EEG analysis using an example from the field of sleep research. As mentioned in section *Background*, visual analysis of EEG is still the golden standard; in the area of sleep the basic reference [37] comes from 1968 (with later updates [38]). It defines criteria for division of sleep into stages, based mostly upon presence/prevalence of certain structures in the corresponding epochs of EEG recording. As for the definitions of these structures, formulated for standardization of their visual detection, let us take the example of sleep spindles:

*The presence of sleep spindle should not be defined unless it is at least 0.5 sec. duration, i.e. one should be able to count 6 or 7 distinct waves within the half-second period. (...)
The term should be used only to describe activity between 12 and 14 cps [37].*

Over the years, common definition drifted towards frequencies from 11 to 15 Hz, duration 0.5–2 seconds and amplitude usually above 15 μV . Reading this definition after 45 years, we are still surprised that:

1. Criteria are defined almost explicitly in time-frequency terms.
2. Before 1993 (that is before the introduction of the matching pursuit [39]), no signal processing method returning explicitly the frequency, amplitude and time width of the oscillations present in a signal was known.

In the following we present MP decomposition of an EEG epoch containing sleep spindles and slow wave activity (SWA). Figure 1 presents a sample epoch of sleep EEG loaded into Svarog (Signal Viewer, Analyzer and Recorder on GPL, see section *Software availability and requirements*). Green and gray rectangles represent SWA and sleep spindles, marked visually by an encephalographer.

Figure 2 presents results of MP decomposition of the epoch selected in Figure 1. Curves below the time-frequency map of energy density represent:

1. Original signal chosen for decomposition (marked in Figure 1).
2. Reconstruction (time course) of selected functions (marked by circled crosses).

Central panel holds the time-frequency map of signals energy density. Each function fitted to the signal by MP is presented as a Gaussian blob in the appropriate time and frequency coordinates, with time and frequency widths corresponding to its parameters. Formula for computing this representation is given by Eq. (4). One can argue about the superior properties of this estimate; indeed, there is a lot of possible ways to estimate signals energy density in the time-frequency plane (spectrogram, wavelets, Wigner-Ville etc.)

and none of them is perfect for all signals. Therefore, we shall concentrate on a unique feature of the MP decomposition, which is an explicit parameterization of the transients present in a signal.

Each of the blobs in Figure 2 represents a Gabor function of known time position and width, amplitude, frequency and phase, as in Eq. (3). These parameters contain the primary information about the signal's content (Eq. (2)), and they can be used directly for the investigation of the properties of the signal, as exemplified below.

Let us define the occurrence of SWA as a wave of amplitude above $50 \mu V$, width above half a second and frequency from 0.5 to 4 Hz. Definition of such filter in Svarog is presented in Figure 3. The amplitude is entered as $25 \mu V$, because the encephalographic convention relates to the peak-to-peak amplitude, which for the Gabor function is double of the mathematical amplitude. This convention is not the only difference between mathematics and visual perception of structures in the signal. For example, visual perception of both amplitude and time width depends on the context (mostly the variance of the signal). Therefore, if exact replication of the visual detection is the main goal, factors like S/N ratio can be incorporated into the post processing criteria.

Result of the application of the filter from Figure 3 to the decomposition from Figure 2 is given in Figure 4. Automagically, all that is left are structures conforming to the definition of SWA. Two major advantages of this result over previously available methods should be noted:

1. This selection takes into account all the features defining SWA, not only their frequency content as would be the case e.g. for a bandpass filter.
2. We have separate parameterization of all the conforming structures, including e.g. the duration of each of them.

The latter feature was explored in [40] for detection of sleep stages III and IV. These stages are defined as epochs occupied by SWA in 20–50% and over 50% of their time, correspondingly. As the previously employed signal processing methods did not parametrize explicitly the time span of relevant activity, it was the first explicit implementation of this rule, designed for standardization of visual detection, in a fully automatic system. We observe also a good concordance of these results with visual detection, represented by the two green marks (intersected by the gray marks for spindles) representing visually tagged SWA in Figure 1. This concordance was quantitatively evaluated in [34].

Similar operation can be performed for sleep spindles; application a filter reflecting their time-frequency parameters, quoted at the beginning of this section, gives us Figure 5. Again, structures detected by the algorithm fall within the borders marked by encephalographer for spindles (gray tags in Figure 1). We observe also a typical example, when two superimposed spindles (on the left) were marked by the human expert as one. It exemplifies the fact that MP-based methods are in most cases downward compatible with visual detection, yet, apart from repeatability and automatization offer also increased sensitivity and resolution. Indeed, in the noisy signal it is almost impossible to see the boundary, and human brain did not evolve for an online calculation of instantaneous frequency, which differentiates these two spindles.

Similarly to the detection of SWA, this scheme is not only more sensitive and selective compared to previous approaches, while retaining a high degree of concordance with

visual detection [33,41], but also allows us, for example, to explicitly count the number of spindles occurrences in any given epoch—a parameter also relevant in sleep analysis.

Finally, it is worth noting that the above proposed procedure is essentially free from any method-dependent settings, like the choice of the mother wavelet in wavelet transform, window width in spectrogram, or smoothing kernel in Wigner-Ville derived representations. All these parameters can significantly alter the results of analysis, and optimal settings can be different for each analyzed epoch. On the contrary, matching pursuit decomposition as such is “more or less” uniquely defined by Equation (1). However, subtle relations between dictionaries and results of MP decomposition were not fully understood so far—their explanation constitutes the main mathematical result presented in this paper. So for the rest of this section let us concentrate on the “more or less”.

Parameters of MP decomposition

As introduced in section *MP algorithm*, MP searches for functions fitting the signal in a large and redundant set called dictionary. The bigger (more redundant) this dictionary is, the better the chance of a perfect fit. Although it's a vague statement, nothing substantially more precise had been said on the relation of the dictionary size and quality of MP decomposition—until now.

This study introduces a novel construction of the dictionary, in which the inner products of the adjacent atoms are kept constant. In other words, distribution of dictionary's functions is uniform with respect to the metric related to their inner product. Using this construction gives us control over the maximum error of a single MP iteration, measured in terms of energy (product of the signal and a function from the optimal dictionary). Upper bound on this error is given by the worst case, where a structure present in the signal falls “in the middle” between dictionary's functions available for decomposition. This error is not equivalent to the resolution of approximation, because MP is an iterative, nonlinear procedure. Formal introduction of the mentioned metric and construction of the optimal dictionary are given in section *Optimal sampling of Gabor dictionaries*.

Figure 6 presents the “Dictionary” tab from the MP configuration window. The most important parameter (first from the top) is called “Energy error”. It corresponds to ϵ from Equation (7) in section *Optimal sampling of Gabor dictionaries*, and relates directly to the mentioned above maximum error of a single MP iteration.

On the practical side, this parameter regulates the price/performance tradeoff in MP decomposition. Smaller (closer to zero) values result in larger amount of functions in the dictionary and higher resolution of resulting MP decomposition, at a price of increased computation times and memory requirements. After each change of this parameter the window displays the amount of RAM that will be occupied by the corresponding dictionary. Keeping “Energy error” constant for analysis of epochs of different sizes will result in larger dictionaries for longer epochs, but the accuracy of decomposition, related to the density of the dictionary, will be the same except for the border effects.

Another issue related to the dictionary is a possible statistical bias, present in decompositions of many epochs in the same, relatively small dictionary. This problem was discussed in [42], where solution in terms of stochastic dictionaries was proposed. Stochastic dictionaries introduced in [42] were constructed by explicit random drawing of functions parameters from defined ranges (subsection *Uniform sampling*). This approach gave less control over the structure of the dictionary and excluded possibilities of several valuable

numerical optimizations. In this paper we propose a different approach, where randomization is achieved by random removal of a defined fraction of functions from a dense, structured dictionary. This procedure is applied when user chooses “OCTAVE_STOCH” as “Dictionary type”. In such case, a dictionary is first created according to the parameter ‘Energy error’ (ϵ from Equation (7)), and then selected fraction of functions is randomly removed from the dictionary. This fraction is governed by the parameter “Percentage chosen”, so that $(100 - \text{Percentage chosen})\%$ of functions is removed.

Finally, the last of the major parameters sets the number of matching pursuit iterations, which is equivalent to the number of functions chosen for the representation (compare Equation (2)). Owing to the convergence property of MP, they are ordered by decreasing energy. Increasing the number of iterations does not influence the parameters of functions chosen in earlier steps.

That is, if we make two separate decompositions of the same signal epoch and using the same dictionary—the first one with 50 iterations and the second one with 10 iterations—then the result of the first decomposition will be the same as taking the first 10 functions from the second decomposition (however, in case of stochastic dictionaries, these two decompositions will be performed using slightly different dictionaries). This difference is more pronounced for smaller dictionaries.

There are several mathematical criteria for stopping the decomposition; their influence was evaluated e.g. for MP-based descriptors of signal complexity [6] or optimal video coding [43]—general discussion of this parameter can be found in [44]. Software presented in this paper implements the two most basic options, based upon the maximum number of iterations (“Max iterations”) and percentage of signal’s energy explained by the whole decomposition (“Energy percent”). These options work in logical conjunction. Hence e.g. for the default settings (99% and 50 iterations), the procedure is completed after the 50th iteration—or before, if the representation (2) explains 99% of energy of the original signal for a lower number of functions M .

Detailed information on the parameters of MP decomposition employed in mp5 is available from [Additional file 3], which contains the help of the mp5 module from Svarog.

Optimal sampling of Gabor dictionaries

We start by formally reintroducing the MP following the notation from the seminal paper by Mallat and Zhang [39]. Denoting the function fitted to the signal x in the n -th iteration of MP as g_{γ_n} , and n -th residual as $R^n x$, we define the procedure as:

$$\begin{cases} R^0 x = x \\ R^n x = \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} x \\ g_{\gamma_n} = \arg \max_{g_{\gamma_i} \in D} |\langle R^n x, g_{\gamma_i} \rangle| \end{cases} \quad (1)$$

As a result we get an expansion

$$x \approx \sum_{n=0}^{M-1} \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n} \quad (2)$$

where M equals the number of iterations. For a time-frequency analysis of real-valued signals, dictionary D is usually composed from Gabor functions

$$g_{\gamma}(t) = K(\gamma) e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(\omega(t-u) + \phi) \quad (3)$$

where $\gamma = (u, \omega, s)$ and $K(\gamma)$ is such that $\|g_{\gamma}\| = 1$.

From expansion (2) we can derive a time-frequency distribution of the signal's energy (as shown in Figure 2), by summing Wigner distributions W of selected functions and explicitly omitting cross-terms:

$$Ex(t, \omega) = \sum_{n=0}^M |\langle R^n x, g_{\gamma_n} \rangle|^2 W_{g_{\gamma_n}}(t, \omega) \quad (4)$$

This magnitude is presented in Figures 2, 4 and 5.

Real world implementations of Gabor dictionaries

Parameters (u, ω, s) of functions g_γ from Equation (3) form a 3-dimensional continuous subspace of \mathbb{R}^3 —the infinite set D_∞ . Ranges of parameters delimiting this subspace correspond to the usual assumptions that the time center u and time spread s do not exceed the length of the analyzed epoch, and to the fact that ω makes no sense above the Nyquist frequency. The fourth parameter ϕ does not have to be taken into account in the construction of the dictionary D , since for any g_γ we can find the phase ϕ maximizing its product $\langle g_\gamma, x \rangle$ with given signal x in a single step—we recall this method in the Appendix (section *Optimal phase of a Gabor function*).

From this D_∞ , a finite dictionary $D \subset D_\infty$ must be chosen for any practical implementation of MP. This choice is equivalent to the choice of a discrete and finite set γ from the continuous ranges of u, ω , and s . However, up to now the major questions, crucial for real-world implementations of MP, were unanswered:

1. How should we choose the elements of the finite dictionary D from D_∞ and why?
2. How much do we gain by increasing the size of the dictionary?

It has been a well known fact, that using a larger dictionary of functions for decomposition of the same signal should lead to more accurate parametrization. However, no direct relation between the density of the dictionary and resolution of the resulting decomposition was derived so far. While an exact measure of the resolution of the highly nonlinear matching pursuit remains an open question, in the following sections we relate the maximum error of a single MP iteration to the density of an optimally sampled Gabor dictionary, filled with Gabor functions distributed uniformly with respect to the metric related to their inner product.

Dyadic sampling

The first, wavelet-like subsampling of D_∞ was proposed in the seminal paper [39]:

$$\gamma = (u, \omega, s) = \left(p a^j \Delta u, k \frac{\Delta \omega}{a^j}, a^j \right) \quad (5)$$

where $j, k, p \in \mathbb{Z}$ and the basic intervals for time and frequency (Δu and $\Delta \omega$) are chosen so that $\Delta u = \frac{\Delta \omega}{2\pi} < 1$. If a dictionary D is constructed from the functions g_γ with parameters $\gamma = (u, \omega, s)$ chosen according to (5), then for any signal $x \in L^2(\mathbb{R})$ there exists an optimality factor $\alpha > 0$ such that

$$\sup_{g_\gamma \in D} |\langle x, g_\gamma \rangle| \geq \alpha \sup_{g \in D_\infty} |\langle x, g \rangle| \quad (6)$$

However, there were no quantitative estimates for the performance of the MP in given dictionary, constructed for given $\Delta \omega$ and a , and it seems that nothing can be said about the optimality factor α except for that it exists and is greater than zero.

Uniform sampling

The quest for a justification of a particular scheme of subsampling D_∞ leads us to the considerations of a uniform sampling. If we know that the parameters (u, ω, s) are uniformly sampled with steps Δu , $\Delta \omega$, and Δs , then for any $g \in D_\infty$ there exists $g_\gamma \in D$, which has the parameters differing at most by halves of the sampling steps. Such sampling provides an estimate of one-step resolution of the highly nonlinear MP algorithm in the space of the parameters of Gabor functions. It was implemented in the mp4 software package [45] as a byproduct of the first approach to the issue of stochastic dictionaries [42], where parameters of the dictionary's functions were drawn from uniform probability distributions.

However, such a sampling scheme in some cases is far from optimal for the MP algorithm. For example, constant sampling of the time interval Δu for all the widths s will result in a dictionary containing a lot of strongly overlapping Gabors with large widths s (which will give large inner products with neighbors from the dictionary), and much more sparse coverage of positions of "shorter" Gabors with small s . Choice (1) of functions for the MP expansion (2) is based entirely on inner products, and the Cartesian metric in the space of the parameters of these functions is far from optimal in this respect. This leads us to the search for a metric that would correspond to the MP selection procedure (1).

In pursuit of a relevant metric

A metric that would reflect the distance between the dictionary elements as "seen" by the MP algorithm should be based on the inner product of dictionary's elements. We propose a metric $d(g_0, g_1)$ to ensure that the distance between nearest Gabor atoms in the dictionary does not exceed some *a priori* given threshold, ϵ . Similarly to dyadic dictionary (see Eq. 5), we assume that the scale parameter s varies by a factor a (dilation factor), as follows: $s_j = a^j$, where $j > 0$ and $a > 1$. However, we will propose a new method for determination of step values $\Delta \omega$ and Δu . We will show that this parametrization allows us to derive such step values, that for every (s, ω, u)

$$\begin{aligned} d(g_{(s,\omega,u,\phi)}, g_{(as,\omega,u,\phi)}) &\leq \epsilon \\ d(g_{(s,\omega,u,\phi)}, g_{(s,\omega+\Delta\omega,u,\phi)}) &\leq \epsilon \\ d(g_{(s,\omega,u,\phi)}, g_{(s,\omega,u+\Delta u,\phi)}) &\leq \epsilon. \end{aligned} \quad (7)$$

In order to construct a proper metric $d(g_0, g_1)$, we start by introducing a simple, intuitive measure for Gabor function similarity which satisfies all the properties of a metric:

$$d_0(g_0, g_1) = 2^{-\frac{1}{2}} \|g_0 - g_1\| = 2^{-\frac{1}{2}} \sqrt{\langle g_0 - g_1 | g_0 - g_1 \rangle} = \sqrt{1 - \langle g_0 | g_1 \rangle}. \quad (8)$$

Norm d_0 above is naturally defined based upon the inner product in the space of the real signals. The constant $2^{-\frac{1}{2}}$ has been introduced to set the distance between orthogonal Gabor atoms to 1, while greater distances (up to 2) can exist between Gabor atoms whose inner products are negative.

Inner product of two Gabor functions has the dimension of energy, thus the dimension of the metric (8) is amplitude. According to this fact, the conditions (7) imply the dimension of the square of the parameter ϵ as the quantity of energy. Also, $\langle G_0 | G_1 \rangle_{\max}$ will be nonnegative for any two Gabor functions. Therefore, since Gabors are normalized, the value of their inner product is limited to the range (0, 1) and thus the limit of the ϵ is

(0, 1). However, $\epsilon = 1$ allows the neighboring atoms to be orthogonal, so for a reasonable MP decomposition one should require $\epsilon \ll 1$.

To the best of our knowledge, the above measure has not been previously applied to the Gabor dictionary construction problem, although a similar one for packing in projection spaces was introduced by Tropp [46]. Interpretation and possible values of parameter ϵ will be discussed in the following section, after we construct a final metric based on $d_0(g_0, g_1)$.

Phase-related equivalence

According to the section *Optimal phase of a Gabor function* from the Appendix, one can easily calculate phase ϕ of the Gabor function that maximizes its inner product with a given signal. Therefore, each atom in MP dictionary can be replaced with any other representative from the set $\tilde{G}_{(s,\omega,u)}$, which is an equivalence class in respect to a relation \sim defined as

$$\mathcal{G}_{(s_0,\omega_0,u_0,\phi_0)} \sim \mathcal{G}_{(s_1,\omega_1,u_1,\phi_1)} \iff (s_0 = s_1) \wedge (\omega_0 = \omega_1) \wedge (u_0 = u_1). \quad (9)$$

This feature has to be taken into account by introducing a correct metric for Gabor atoms

$$d(g_0, g_1) = \tilde{d}(g_0/\sim, g_1/\sim), \quad (10)$$

where $\tilde{d}(G_0, G_1)$ is the distance function defined on equivalence classes

$$\tilde{d}(G_0, G_1) = \min_{\substack{g_0 \in G_0 \\ g_1 \in G_1}} d_0(g_0, g_1) = \sqrt{1 - \max_{\substack{g_0 \in G_0 \\ g_1 \in G_1}} \langle g_0 | g_1 \rangle} = \sqrt{1 - \langle G_0 | G_1 \rangle_{\max}}, \quad (11)$$

and we introduce the *maximal scalar product* $\langle G_0 | G_1 \rangle_{\max}$, which can be calculated as

$$\langle \mathcal{G}_{(s_0,\omega_0,u_0)} | \mathcal{G}_{(s_1,\omega_1,u_1)} \rangle_{\max} = \max_{\phi_0, \phi_1 \in [0; 2\pi]} \langle \mathcal{G}_{(s_0,\omega_0,u_0,\phi_0)} | \mathcal{G}_{(s_1,\omega_1,u_1,\phi_1)} \rangle. \quad (12)$$

Inner products of Gabor functions

The measure d_0 (8) is based on an inner product of Gabor functions, so it is necessary to calculate analytical formulae for this product. Such expressions can be found for real Gabor functions

$$g_\gamma(t) = K_\gamma e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(\omega(t-u) + \phi), \quad (13)$$

where $\gamma \equiv (s, \omega, u, \phi)$ is a set of the parameters, and

$$K_\gamma = \frac{2^{3/4}}{\sqrt{s}} \left[1 + \cos(2\phi) e^{-\frac{s^2 \omega^2}{2\pi}} \right]^{-\frac{1}{2}} \quad (14)$$

is the normalization factor.

To present an analytical formula for the inner product of two real Gabor functions (13) g_0 and g_1 with normalization factors K_0 and K_1 respectively, we introduce following constants:

$$A = \pi \left(\frac{1}{s_0^2} + \frac{1}{s_1^2} \right) \quad (15)$$

$$B = \frac{\pi}{A} \left(\frac{u_0}{s_0^2} + \frac{u_1}{s_1^2} \right) \quad (16)$$

$$C = -\pi \left(\frac{u_0^2}{s_0^2} + \frac{u_1^2}{s_1^2} \right) + AB^2 = -\pi \frac{(u_0 - u_1)^2}{s_0^2 + s_1^2}. \quad (17)$$

The complete formula for the product has been adapted from [47]:

$$\begin{aligned} \langle g_0, g_1 \rangle = K_0 K_1 \sqrt{\frac{\pi}{4A}} e^C & \left(\cos [(\omega_0 + \omega_1)B + (\phi_0 + \phi_1) - (\omega_0 u_0 + \omega_1 u_1)] e^{-\frac{(\omega_0 + \omega_1)^2}{4A}} + \right. \\ & \left. + \cos [(\omega_0 - \omega_1)B + (\phi_0 - \phi_1) - (\omega_0 u_0 - \omega_1 u_1)] e^{-\frac{(\omega_0 - \omega_1)^2}{4A}} \right). \end{aligned} \quad (18)$$

Products of adjacent atoms

To simplify the inner product (18) for special cases of adjacent dictionary atoms, we will calculate distances between two atoms which differ in only one of the three parameters s, ω, u . Let us discuss each case separately.

Scale

Because $u_0 = u_1$, the scalar product is invariant to a shift in position u , so we can safely choose $u = 0$. In this case $A = \frac{\pi(a^2+1)}{a^2s^2}$, $B = 0$, $C = 0$. Therefore,

$$\langle g_{(s,\omega,u=0,\phi_0)} | g_{(as,\omega,u=0,\phi_1)} \rangle = \sqrt{\frac{2a}{a^2+1}} \frac{\cos(\phi_0 - \phi_1) + \cos(\phi_0 + \phi_1) e^{-\frac{a^2s^2\omega^2}{(a^2+1)\pi}}}{\sqrt{1 + \cos(2\phi_0) e^{-\frac{s^2\omega^2}{2\pi}}} \sqrt{1 + \cos(2\phi_1) e^{-\frac{a^2s^2\omega^2}{2\pi}}}. \quad (19)$$

The global maximum of this product will be present for $\phi_0 = \phi_1 = 0$ or $\phi_0 = \phi_1 = \frac{\pi}{2}$, depending on parameter values. Therefore, one can calculate

$$\langle G_{(s,\omega,u)} | G_{(as,\omega,u)} \rangle_{\max} = \sqrt{\frac{2a}{a^2+1}} \max \left\{ \frac{1 - e^{-\frac{a^2s^2\omega^2}{(a^2+1)\pi}}}{\sqrt{1 - e^{-\frac{s^2\omega^2}{2\pi}}} \sqrt{1 - e^{-\frac{a^2s^2\omega^2}{2\pi}}}; \frac{1 + e^{-\frac{a^2s^2\omega^2}{(a^2+1)\pi}}}{\sqrt{1 + e^{-\frac{s^2\omega^2}{2\pi}}} \sqrt{1 + e^{-\frac{a^2s^2\omega^2}{2\pi}}} \right\}. \quad (20)$$

Frequency

The frequency parameter ω changes by $\Delta\omega$, $s_0 = s_1$ and $u_0 = u_1$. We can choose $u = 0$, analogically to the previous case. Here $A = \frac{2\pi}{s^2}$, $B = 0$ and $C = 0$, so the product is as follows:

$$\langle g_{(s,\omega,u=0,\phi_0)} | g_{(s,\omega+\Delta\omega,u=0,\phi_1)} \rangle = e^{-\frac{\Delta\omega^2s^2}{8\pi}} \frac{\cos(\phi_0 - \phi_1) + \cos(\phi_0 + \phi_1) e^{-\frac{s^2\omega(\omega+\Delta\omega)}{2\pi}}}{\sqrt{1 + \cos(2\phi_0) e^{-\frac{s^2\omega^2}{2\pi}}} \sqrt{1 + \cos(2\phi_1) e^{-\frac{s^2(\omega+\Delta\omega)^2}{2\pi}}}. \quad (21)$$

The global maximum of this product will be present for $\phi_0 = \phi_1 = \frac{\pi}{2}$. Therefore,

$$\langle G_{(s,\omega,u)} | G_{(s,\omega+\Delta\omega,u)} \rangle_{\max} = e^{-\frac{\Delta\omega^2s^2}{8\pi}} \frac{1 - e^{-\frac{s^2\omega(\omega+\Delta\omega)}{2\pi}}}{\sqrt{1 - e^{-\frac{s^2\omega^2}{2\pi}}} \sqrt{1 - e^{-\frac{s^2(\omega+\Delta\omega)^2}{2\pi}}}. \quad (22)$$

Position

In this case, $s_0 = s_1$, $\omega_0 = \omega_1$, and position parameter u differs by Δu . We can choose $u_0 = -\frac{\Delta u}{2}$ and $u_1 = \frac{\Delta u}{2}$. Applying $A = \frac{2\pi}{s^2}$, $B = 0$, $C = -\frac{\pi \Delta u^2}{2s^2}$ lead to

$$\langle g_{(s,\omega,u_0=-\frac{\Delta u}{2},\phi_0)} | g_{(s,\omega,u_1=\frac{\Delta u}{2},\phi_1)} \rangle = e^{-\frac{\pi \Delta u^2}{2s^2}} \frac{\cos(\phi_0 - \phi_1 + \omega \Delta u) + \cos(\phi_0 + \phi_1) e^{-\frac{s^2 \omega^2}{2\pi}}}{\sqrt{1 + \cos(2\phi_0) e^{-\frac{s^2 \omega^2}{2\pi}}} \sqrt{1 + \cos(2\phi_1) e^{-\frac{s^2 \omega^2}{2\pi}}}. \tag{23}$$

The global maximum of the product will appear for $\phi_0 = -\phi_1$. Therefore,

$$\langle G_{(s,\omega,u)} | G_{(s,\omega,u+\Delta u)} \rangle_{\max} = e^{-\frac{\pi \Delta u^2}{2s^2}} \max_{\phi \in [0;2\pi]} \frac{\cos(2\phi + \omega \Delta u) + e^{-\frac{s^2 \omega^2}{2\pi}}}{1 + \cos(2\phi) e^{-\frac{s^2 \omega^2}{2\pi}}}. \tag{24}$$

Construction of the optimally sampled dictionary

Formulae (20), (22) and (24) substituted into (7) could be used to construct an optimal dictionary. However, to improve computational performance of MP algorithms, one can use Fast Fourier Transform (FFT) as described in the Appendix. Therefore, it would be preferable to have values of $\Delta \omega$ and Δu independent of ω , yet still satisfying (7).

To construct such “uniform” dictionary, step values (a , $\Delta \omega$, Δu) must be selected for such ω that minimizes the value of the maximal scalar product (12). Exemplary step values obtained from (20), (22) and (24) are shown in Figure 7, as a function of frequency ω , for three different values of parameter ϵ . It can be observed that in order to obtain a , $\Delta \omega$, Δu fulfilling the condition (7) and independent of ω , one has to select step values corresponding to a large value of frequency ω .

Moreover, at frequency $\omega \approx \pi$, one can safely substitute in Equations (20), (22) and (24):

$$\begin{aligned} e^{-\frac{s^2 \omega^2}{2\pi}} &\approx 0 \\ e^{-\frac{a^2 s^2 \omega^2}{2\pi}} &\approx 0 \\ e^{-\frac{a^2 s^2 \omega^2}{(a^2+1)\pi}} &\approx 0 \\ e^{-\frac{s^2 \omega(\omega+\Delta \omega)^2}{2\pi}} &\approx 0. \end{aligned} \tag{25}$$

With these approximations, maximal scalar products can be calculated easily as

$$\begin{aligned} \langle G_{(s,\omega,u)} | G_{(as,\omega,u)} \rangle_{\max} &\approx \sqrt{\frac{2a}{a^2 + 1}} \\ \langle G_{(s,\omega,u)} | G_{(s,\omega+\Delta \omega,u)} \rangle_{\max} &\approx e^{-\frac{\Delta \omega^2 s^2}{8\pi}} \\ \langle G_{(s,\omega,u)} | G_{(s,\omega,u+\Delta u)} \rangle_{\max} &\approx e^{-\frac{\pi \Delta u^2}{2s^2}}, \end{aligned} \tag{26}$$

taking into account that $\max_{\phi \in [0;2\pi]} \cos(2\phi + \omega \Delta u) = 1$.

Therefore, the optimal step values independent of ω can be calculated from Equations (7) in the following steps:

1. For a given threshold ϵ (see Eq. 7) one can calculate the dilation factor a :

$$a = \frac{1 + \epsilon \sqrt{(2 - \epsilon^2)(\epsilon^4 - 2\epsilon^2 + 2)}}{(1 - \epsilon^2)^2} \tag{27}$$

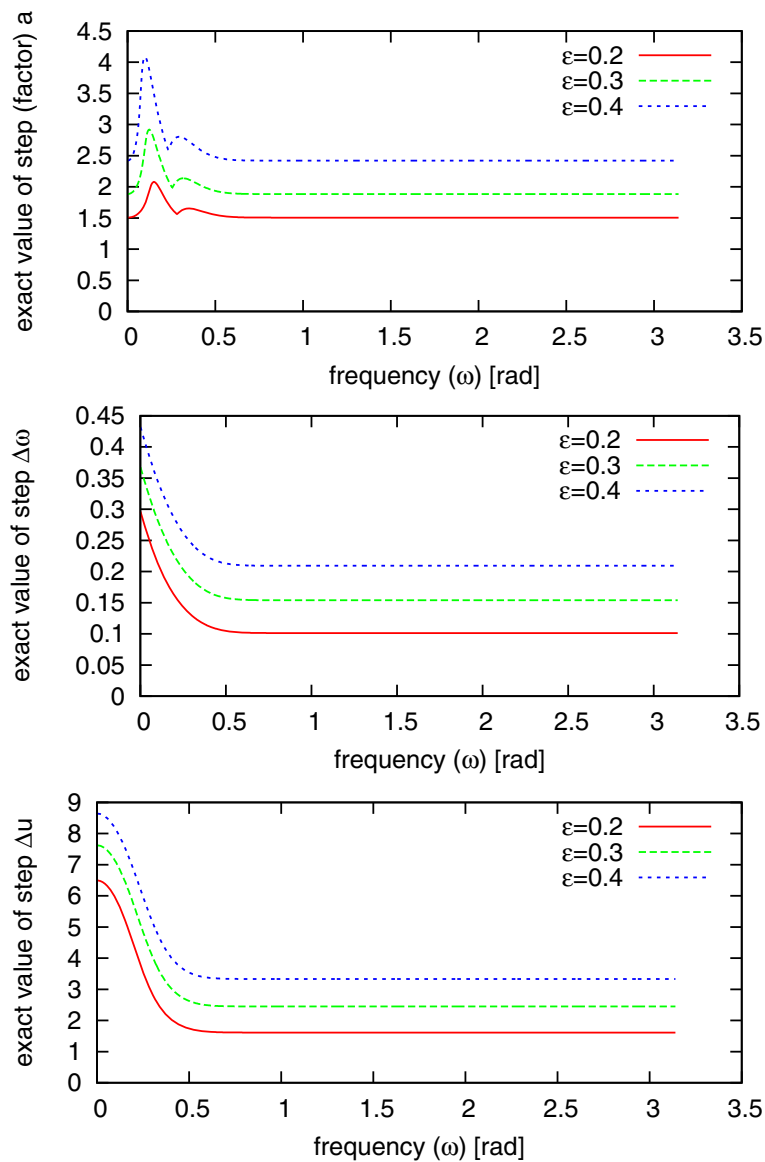


Figure 7 Optimal scale factor, frequency step and position step. Optimal scale factor a (top plot—see Eq. (20)), frequency step $\Delta\omega$ (middle plot—see Eq. (22)) and position step Δu (see Eq. (24)) as a function of ω for three values of ϵ and scale $s = 10$.

- The set of scales s_1, s_2, \dots, s_{N_s} of Gabors are determined as:

$$s_j = a^j, \tag{28}$$

where $j \in \mathbb{Z}_+, N_s = \lfloor \log_a N \rfloor$ is the number of different scales and N is the length of the analyzed signal.

- For each scale s obtained from Eq. (28) the step $\Delta\omega$ in frequency domain and Δu in time domain is calculated according to the formula:

$$\Delta\omega = \frac{1}{s} \sqrt{-8\pi \log(1 - \epsilon^2)} \tag{29}$$

$$\Delta u = s \sqrt{-\frac{2}{\pi} \log(1 - \epsilon^2)}. \tag{30}$$

Multivariate matching pursuit (MMP)

Need to analyze jointly more than one epoch usually arises in case of simultaneous measurements of more than one signal, or when subsequent epochs are treated as realizations of the same process (usually via some kind of ergodicity assumption). In EEG/MEG analysis these two cases usually refer to:

1. Simultaneous recording of EEG/MEG from more than one electrode/sensor, called hereafter “channels”.
2. EEG/MEG recording of subsequent time-locked responses to repetitive stimuli (event-related potentials or fields, ERPs/ERFs), called hereafter “trials”.

Several versions of the MMP algorithm were separately proposed for particular applications—we will briefly discuss this issue in section *Discussion*. In this paper we propose a general framework, where variants of the MMP are identified on the basis of mathematical formulation of the two crucial elements of the multivariate algorithm:

- A. Inter-channel/inter-trial constrains, that is parameters that we keep constant or allow to change across the channels/trials.
- B. Criteria for selection of the function from the dictionary in each MMP iteration—in MMP, contrary to MP, this function is fitted to many epochs at the same time, and optimality of fit to many signals at once can be expressed in different ways.

Such an universal approach allows for a direct implementation of the same code to a multitude of different paradigms encountered in recordings of any multivariate time series, not only EEG/MEG. Nevertheless, through this paper we will stick to “channels” and “epochs” in relation to the organization of the multivariate datasets. The following chapters define basic variants of MMP.

MMP1 (max sum of the moduli of products, constant phase)

The most straightforward multivariate extension of the MP—let’s call it MMP1—can be achieved as follows:

- A1. Only the amplitude varies across channels.
- B1. We maximize the sum of products in all the channels.

We maximize the sum of moduli of the products rather than their squares, as proposed in [48], due to the more efficient selection of the common phase for all the channels. Also, owing to the Gabor update formula (38), in all the iterations but the first one we compute the products of dictionary’s atoms with the function from the previous iteration, which has all the parameters fixed except for the amplitude. This saves a lot of computations comparing to the case of computing products with all the channels residua separately, as in MMP3.

Let us denote the multichannel signal as \mathbf{x} , and the signal in the i th channel as \mathbf{x}^i , with $i = 1 \dots N_c$, where N_c is the number of channels. We may express the condition (B1) for the choice of atom g_γ in n -th iteration as

$$\max_{g_\gamma \in D} \sum_{i=1}^{N_c} |\langle R^n \mathbf{x}^i, g_\gamma \rangle| \quad (31)$$

The whole procedure can be described as

$$\begin{cases} R^0 \mathbf{x} = \mathbf{x} \\ R^n \mathbf{x}^i = \langle R^n \mathbf{x}^i, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} \mathbf{x}^i \\ g_{\gamma_n} = \arg \max_{g_{\gamma} \in D} \sum_{i=1}^{N_c} |\langle R^n \mathbf{x}^i, g_{\gamma} \rangle| \end{cases} \quad (32)$$

Results of MMP1 are given in terms of functions g_{γ_n} , selected in consecutive iterations, and their weights in all the channels, determined for channel i by the real-valued products $\langle R^n \mathbf{x}^i, g_{\gamma_n} \rangle$. In each iteration, the multichannel residuum $R^{n+1} \mathbf{x}$ is computed by subtracting from the previous residua in each channel i the contribution of g_{γ_n} , weighted by $\langle R^n \mathbf{x}^i, g_{\gamma_n} \rangle$.

MMP2 (max modulus of the sum of products, constant phase)

Assumption of invariant phase in all the channels was explored in [22] to yield an efficient decomposition algorithm. If we modify the criterion of choice from the previous section to

$$\max_{g_{\gamma} \in D} \left| \sum_{i=1}^{N_c} \langle R^n \mathbf{x}^i, g_{\gamma} \rangle \right|, \quad (33)$$

we get the conditions:

- A2. Only the amplitude varies across the channels.
- B2. We maximize the absolute value of the sum of products across channels.

Due to the linearity of the residuum operator R [22], this choice allows for implementing a simple trick. Instead of finding in each step the product of each dictionary's waveform with all the channels separately, and then computing their sum (33), in each step we decompose the average signal $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = \frac{1}{N_c} \sum_{i=1}^{N_c} \mathbf{x}^i \quad (34)$$

$$\begin{cases} R^0 \bar{\mathbf{x}} = \bar{\mathbf{x}} \\ R^n \bar{\mathbf{x}} = \langle R^n \bar{\mathbf{x}}, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} \bar{\mathbf{x}} \\ \mathbf{g}_{\gamma_n} = \arg \max_{\mathbf{g}_{\gamma} \in D} |\langle R^n \bar{\mathbf{x}}, \mathbf{g}_{\gamma} \rangle| \\ R^n \mathbf{x}^i = \langle R^n \mathbf{x}^i, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} \mathbf{x}^i \end{cases} \quad (35)$$

This procedure yields a computational complexity close to the monochannel MP—compared to MMP1, reduced by the factor N_c (that is, number of channels).

Convergence of this procedure may be relatively slower for waveforms appearing in different channels with exactly opposite phases.

Due to operating on the average of channels, this version of the algorithm cannot be directly applied to the data presented in the average reference (montage). These problems are absent in MMP1 as well as in the next implementation, allowing for arbitrary phases across the channels.

MMP3 (variable phase)

- A3. Phase and amplitude vary across the channels.
- B3. We maximize the sum of squared products (energies) across channels.

Again, as in (31), we maximize

$$\max_{g_\gamma \in D} \sum_{i=1}^{N_c} |\langle R^n \mathbf{x}^i, g_\gamma^i \rangle|^2 \quad (36)$$

but this time $g_{\gamma_n}^i$ are not the same g_{γ_n} for all channels i —they can have different phases.

$$\begin{cases} R^0 \mathbf{x} = \mathbf{x} \\ R^n \mathbf{x}^i = \langle R^n \mathbf{x}^i, g_{\gamma_n}^i \rangle g_{\gamma_n}^i + R^{n+1} \mathbf{x}^i \\ g_{\gamma_n}^i = \arg \max_{g_\gamma \in D} \sum_{i=1}^{N_c} |\langle R^n \mathbf{x}^i, g_\gamma^i \rangle|^2 \end{cases} \quad (37)$$

As presented in the Appendix (section *Optimal phase of a Gabor function*), computing an optimal phase of Gabor function g_γ , maximizing absolute value of the product $\langle R^n \mathbf{x}^i, g_\gamma \rangle$, can be implemented very efficiently. Value of (31) for phases optimized separately will never be smaller than in the case of the phase common to all the channels, so this freedom should improve the convergence.

MMPXY

In case of multichannel recordings of event-related potentials (ERPs), we are dealing with a slightly more complicated structure of the data. Instead of a vector of signals \mathbf{x}^i , where i indexes channels/sensors, we get a matrix of signals \mathbf{X}_k^i , with additional index k reflecting subsequent repetitions (single trials). On such data, MMP can be performed across either of the two dimensions i and k . For the sake of simplicity, we will call these indices “channels” and “trials”, although for the algorithm they can represent any direction of a multidimensional dataset. The algorithm does not contain any problem-specific optimizations and as such preserves generality.

Usually, the first index in multiplexed multichannel data is the channel (sensor) i and the second is the time index t . Number of trial k comes next, and indexes the the set of whole epochs and all channels i , related to the k -th repetitions of an event (usually ERP). MP operates on epochs indexed by discrete time t . MMP will operate on channel and trial indices i and k , allowing for different constrains across sensors or repetitions.

For the mp5 package we assumed the naming convention in the form MMPXY, where X denotes the version of MMP algorithm used in each iteration across the channel dimension, and Y—across the repetitions:

MMP11: For each channel and trial fit the optimal g_γ with $\phi = \text{const}$.

MMP12: Average trials k in each channel i and fit optimal g_γ with $\phi = \text{const}$ to these averages.

MMP21: For each trial k find the average across channels i and fit optimal g_γ with $\phi = \text{const}$ to these averages.

MMP23: For each trial k find the average across channels i and fit to these averages optimal g_γ with potentially different phase ϕ in each channel.

MMP32: For each channel i find the average across trials k and fit to these averages optimal g_γ with potentially different phase ϕ in each trial.

MMP33: For each channel and trial fit the optimal g_γ with potentially different phase ϕ in each channel and trial.

For example, MMP12 denotes a setting where in each iteration the choice of the atom fitting best all the channels will be effectuated according to MMP1, while across the trials

MMP2 will be used. That is, in each iteration the residua of single trials are averaged separately in each channel, and to these averages the best g_j is fitted according to MMP1. Naming of the algorithms (MMPX) corresponds to the three above subsections. MMP13 and MMP31 were not implemented in mp5.

Discussion

Resolution of MP

One of the problems, faced by everyone applying MP to exploratory analysis of signals, is “how big should be the dictionary so that I do not miss some important structures?” Of course “the bigger the better”, but increasing the size of the dictionary increases significantly the computational cost, and the exact gain in resolution of MP representation was not known so far. Optimal Gabor dictionaries, introduced in this paper, for the first time allow to relate the density of the dictionary to the maximum error of a single iteration of the algorithm.

Details of implementation

This paper describes mathematical foundations and numerical optimizations, implemented in the freely available software package mp5, developed at the University of Warsaw for decomposition of signals using multivariate matching pursuit. This software made possible several published works [24,28,29] plus some studies in progress. It builds on another decade of experience in using the previous, monovariate version of the algorithm which we made freely available as mp4 at <http://eeg.pl/mp>, and used in over a dozen published studies. Although all these cited works would not be possible without this software, they were published as “classical scientific papers” where author is supposed to concentrate only on the scientific novelty, not tools. There was no space for all the important details of the implementations, which make a big difference in the results and hence constitute the core of the Reproducible Research. Therefore, although the source code of both these packages have been freely available for years, a complete description of employed mathematical and numerical optimizations and tricks, as presented in Appendix Implementation and optimizations, was not published up to now.

Applications of MMP to EEG/MEG

Probably the most promising field of future applications is related to the multivariate matching pursuit (MMP). Real world applications were made possible owing to the progress in computer hardware in recent decades. In section *Multivariate matching pursuit (MMP)* we propose, in concordance with [44], a simple classification of MMP variants according to the combinations of the constraints on parameters and criteria of choice. However, there is a potential continuum of other approaches outside of this simple scheme, tailored very specifically for particular applications. Most of them suffer from the need of arbitrary setting some parameter that may significantly change the results of decomposition. In this light we believe that in most cases the simple approach proposed in section *Multivariate matching pursuit (MMP)*, which is free from arbitrary parameters, is optimal and most elegant, at least for starting. A procedure that is free of task-specific settings has also obvious advantages stemming directly from its generality.

However, specific approaches of course may offer addition advantages in particular cases. For example, MMP tailored for the analysis of stereo recordings of sound

in [49] allows for different time positions of the time-frequency atoms present in the two channels. Together with different amplitudes in each channel, it relates to modeling the microphones as gain-delay filters in the anechoic case. Unfortunately, a model explaining relations between channels of EEG/MEG recordings is far more complicated, even in the case of known distribution of sources (so-called forward EEG problem).

An attempt to incorporate constraints, reflecting the generation of multichannel EEG, into the MMP procedure, was presented in [26]. To the purely energetic criterion of MMP1 (31), a second term was added to favor those g_γ which give smooth distribution of amplitudes across the channels. Spatial smoothness (quantified by Laplacian operators) means basically that the values of $\langle R^n \mathbf{x}^i, g_\gamma \rangle$ should be similar for i corresponding to the neighboring channels. However, a choice combining two completely different criteria requires some setting of their relative weights. For example, if we attribute too much importance to the spatial criterion, in favor of the energetic one, we may obtain atoms giving very smooth scalp distributions across electrodes. But in such a case the convergence of the MMP procedure, measured in the rate-distortion sense, relating to the amount of explained energy, may be severely impaired. Up to now, no objective or optimal settings for regulating the influence of such extra criteria on the MMP algorithms was proposed.

Finally, we should also point out some of the possible novel and promptly implementable applications of MMP. Variants of the multivariate algorithms, described in section *Multivariate matching pursuit (MMP)*, can be related to several models of multichannel recordings of repetitive trials of evoked brains activity. Algorithms developed for parameterization of EEG structures in subsequent channels [22,24] have been applied to decomposing subsequent trials of event-related potentials [28,29].

Ongoing works include application of MMP3 to evoked potentials, where variable phase accounts for the variable latencies, and instantaneous decomposition of both repetitions and channels of event-related potentials, with some of the discussed constraints applied separately to the relevant dimensions.

Apart from that, MMP3 may be also used to compute estimates of the phase locking factor [50] (also called inter-trial coherence, [51]). Simultaneous decomposition of all the repetitions will be crucial in this case: in separate MP decompositions of subsequent trials, atoms representing possibly the same structures can have slightly different frequencies, which makes their relative phase insignificant. By estimating the phase coherence only in those are of the time-frequency plane, where indeed an oscillatory activity appears, we may get rid of a lot of noise blurring previously applied estimates.

Software availability and requirements

Software package described in this article is freely available from <http://braintech.pl/svarog/>. It can be run on a computer with a reasonably recent version of MS Windows, Mac OS or GNU/Linux with Java runtime environment. Screencasts (video files), showing (1) downloading and configuration of the package and (2) MP decomposition process of a sample EEG epoch via Svarog, are included as [Additional file 1] and [Additional file 2]. These videos can be also viewed in a variety of formats embedded in HTML5 at <http://braintech.pl/svarog>. [Additional file 3] contains a snapshot of Svarog's help related to mp5.

Complete source code for the MMP engine written in C is available from <http://git.braintech.pl/matching-pursuit.git>. GUI is implemented in Java within the Svarog system,

for which the source code is available at <http://git.braintech.pl/svarog.git>. Both projects are released on terms of the General Public License (GPL).

Svarog is a loose acronym for Signal Viewer, Analyzer and Recorder on GPL, and constitutes the core of the world's first professional EEG recording and analysis system based entirely on GPL software (<http://braintech.pl/Manifesto.html>). This multiplatform system is developed primarily for GNU/Linux. Current versions of the system, including the above discussed software plus the OpenBCI framework for brain-computer interfaces and a modified version of the PsychoPy framework for experiments design, are available as packages for Ubuntu and Debian (see <http://deb.braintech.pl>).

Appendix

Implementation and optimizations

Product update formula

This optimization was proposed in [39].

Let us recall from (1) the formula for the n th residuum:

$$R^n x = \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} x$$

Taking the product of both sides with g_{γ_i} , which is the candidate for selection in the next iteration, we get

$$\langle R^{n+1} x, g_{\gamma_i} \rangle = \langle R^n x, g_{\gamma_i} \rangle - \langle R^n x, g_{\gamma_n} \rangle \langle g_{\gamma_n}, g_{\gamma_i} \rangle \quad (38)$$

This equation expresses the product of a dictionary function g_{γ_i} with the residuum in step $n + 1$ using two products, which were already calculated in the previous iteration— $\langle R^n x, g_{\gamma_i} \rangle$ and $\langle R^n x, g_{\gamma_n} \rangle$ —and a product of two functions from the dictionary— $\langle g_{\gamma_n}, g_{\gamma_i} \rangle$. Therefore, the only thing that remains to be computed is a product of two known functions.

Inner product of continuous Gabor functions can be expressed in terms of elementary functions (see [47,52]). Unfortunately, it does not reflect with enough accuracy the numerical value of the product of two discrete vectors, representing sampled versions of the same Gabor functions. Exact formula for the product of the latter involves theta functions, which can be approximated by relatively fast converging series [47].

Sin, Cos, and Exp: fast calculations and tables

In spite of the trick from the previous section, still—at least in the first iteration—we need to compute the “plain” inner products of the signal with Gabor functions. Using the result from section *Optimal phase of a Gabor function*, of all the phases ϕ we calculate products only for $\phi = 0$ and $\phi = \frac{\pi}{2}$.

Computationally, the most expensive part is not the actual calculation of the inner products, but the generation of discrete vectors of samples from Equation (3), which contains cosine and exponent. Compilers usually approximate these functions by high-order polynomials. Although contemporary CPUs may implement directly some special functions, they will still be much more expensive to compute than basic additions or multiplications. Therefore, avoiding explicit calls of these functions may result in significant acceleration—together with tabularization, it accelerated the MP implementation [45] by over an order of magnitude. In the following we show (after [53]) how to fill in a vector with values of sines and cosines for equally spaced arguments using only one call of these functions.

Since the time t in (3) in the actual computations is discrete, the trick is to compute $\sin(\omega(t + 1))$ knowing $\sin(\omega t)$. Using the trigonometric identity (59) with its corresponding form for the sine function, we get

$$\sin(\omega(t + 1)) = \sin(\omega t + \omega) = \cos(\omega t) \sin(\omega) + \sin(\omega t) \cos(\omega) \quad (39)$$

$$\cos(\omega(t + 1)) = \cos(\omega t + \omega) = \cos(\omega t) \cos(\omega) - \sin(\omega t) \sin(\omega) \quad (40)$$

We start with $t = 0$, setting $\cos(0) = 1$ and $\sin(0) = 0$, and computing constants $\cos(\omega)$ and $\sin(\omega)$. Values of (39) and (40) for subsequent t can be filled in a recursive way, using the computed $\cos(\omega)$ and $\sin(\omega)$ and taking as $\sin(\omega t)$ and $\cos(\omega t)$ values from the previous steps.

A similar approach can accelerate computation of the factors $e^{-\alpha t^2}$ present in (3):

$$e^{-\alpha(t+1)^2} = e^{-\alpha t^2 - 2\alpha t - \alpha} = e^{-\alpha t^2} e^{-2\alpha t} e^{-\alpha} \quad (41)$$

To compute (41) we need $e^{-\alpha t^2}$ from the previous iteration, constant $e^{-\alpha}$ independent of t , and $e^{-2\alpha t}$. The last factor can be updated in each iteration at a cost of one multiplication: to get $e^{-2\alpha(t+1)}$ from $e^{-2\alpha t}$ we multiply it by a precomputed constant $e^{-2\alpha}$.

In all these cases we also take into account the symmetries $\sin(-x) = -\sin(x)$, $\cos(-x) = \cos(x)$, and $e^{-(-x)^2} = e^{-x^2}$ to double the savings. Values of these vectors can be stored in memory for subsequent calculations of Gabor vectors (3) with different combinations of sin/cos and exp, but only if we restrict the discretization of parameters to some integer grid, for example:

$$u = 1 \dots N, \quad \omega = (1 \dots N)\pi/N, \quad s = 1 \dots N \quad (42)$$

Apart from fast Sin, Cos and Exp function generation, the optimal dictionary allows for saving in the computer memory the tables with values of these functions. The number N_s of different scales in an optimal dictionary is (see Equation (27)):

$$N_s = \lfloor \log_a N \rfloor \quad (43)$$

where N is the number of samples in signal and a is the parameter expressed by Equation (27). A typical epoch of EEG/MEG contains some thousands samples, so it is possible to store all EXP functions in computer memory. Due to the fact, that Gabors, for given scale, are arranged in frequency domain in increments of $\Delta\omega$ (29) one can save also in computer's memory one period of Sine/Cosine signal of the lowest frequency. The sine and cosine signal of higher frequencies, for example $k \times \Delta\omega$, where k is the natural number, can be generated by means of selecting every k -th sample from Cos/Sine signal of frequency $\Delta\omega$ stored in the computer's memory.

Fast detection of two orthogonal Gabors

Update formula (section *Product update formula*) in combination with optimal dictionary allows for uses the next numerical optimization—fast assessment of orthogonality of two Gabors functions. Let us analyze the analytical formula for an inner product of two

Gabors. After substituting to the Equation (18) the exact expression for normalization factor K_γ and constant C , and introducing the following factors:

$$\begin{aligned}
 X &= \left(\cos [(\omega_0 + \omega_1)B + (\phi_0 + \phi_1) - (\omega_0 u_0 + \omega_1 u_1)] e^{-\pi \frac{(\omega_0 + \omega_1)^2}{4A}} \right. \\
 &\quad \left. + \cos [(\omega_0 - \omega_1)B + (\phi_0 - \phi_1) - (\omega_0 u_0 - \omega_1 u_1)] e^{-\pi \frac{(\omega_0 - \omega_1)^2}{4A}} \right) \\
 Y &= K_0 K_1 \sqrt{s_1 s_2} = \frac{2^{6/4}}{\sqrt{\left(1 + \cos(2\phi_0) e^{-\frac{s_0^2 \omega_0^2}{2\pi}}\right) \left(1 + \cos(2\phi_1) e^{-\frac{s_1^2 \omega_1^2}{2\pi}}\right)}} \\
 Z &= \sqrt{\frac{\pi}{4A}} (s_1 s_2)^{-\frac{1}{2}} e^C = \sqrt{\frac{s_0 s_1}{4(s_0^2 + s_1^2)}} e^{-\pi \frac{(u_0 - u_1)^2}{s_0^2 + s_1^2}}
 \end{aligned} \tag{44}$$

one can obtain:

$$\langle g_0, g_1 \rangle = X \cdot Y \cdot Z \tag{45}$$

It is straightforward to estimate that value of factor X fulfils the condition

$$|X| < 2 \tag{46}$$

for every possible set of parameters $\{s_0, s_1, \omega_0, \omega_1, \phi_0, \phi_1, u_0, u_1\}$. The maximal value of parameter Y is limited by the lowest values of scale s and frequency ω , which, based on (29), are:

$$\begin{aligned}
 s &= s_{\min} > 0 \\
 \omega &= \Delta\omega = \frac{2\sqrt{\pi}}{s_{\min}} \sqrt{-\log(1 - \epsilon^2)}
 \end{aligned} \tag{47}$$

and phases $\phi_0 = \phi_1 = \pm \frac{\pi}{2}$. Substituting above formulae into factor Y results in following expression:

$$|Y| \leq \frac{2^{3/2}}{\sqrt{\left(1 + \cos(2\phi_0) (1 - \epsilon^2)^2\right) \left(1 + \cos(2\phi_1) (1 - \epsilon^2)^2\right)}} \leq \frac{2^{3/2}}{\epsilon^2(2 - \epsilon^2)} \approx \frac{\sqrt{2}}{\epsilon^2} \tag{48}$$

Based on this observation, it is possible to introduce a numerical threshold, defining an approximate orthogonality of two Gabor functions. The factor Z depends on the relative position and the width of two Gabors. Atoms which differ mostly in these parameters will give a small factor Z . Therefore

$$\frac{2\sqrt{2}}{\epsilon^2} Z < \eta \implies |XYZ| < \eta \implies \langle g_0, g_1 \rangle \approx 0 \tag{49}$$

where η can be set for example at the accuracy of a double precision number (10^{-16}). Condition (49) allows for efficient detection of orthogonal atoms in dictionary and replacing their inner product by zero in Equation (38). Moreover, in case of a dictionary with uniform step $\Delta\omega$ at a given scale, it is possible to determine the set of Gabor functions, characterized by the same position u , for which inner product with the Gabor selected in previous iteration will be zero.

Limiting domain of the product

When Equation (49) is not fulfilled, the two Gabor functions cannot be treated as orthogonal, and their product has to be determined. Full scalar product of two Gabor functions g_1 and g_2 can be written as

$$\langle g_1, g_2 \rangle = K_1 K_2 \int_{-\infty}^{+\infty} e^{-A(t-B)^2} e^C \cos(\omega_1(t - u_1) + \phi_1) \cos(\omega_2(t - u_2) + \phi_2) dt \quad (50)$$

where A, B and C are defined in Equations (15–17).

To perform numerical integration, one can replace the improper integral above with a definite integral on a sufficiently large interval $[a; b]$, so that

$$\left| \int_{-\infty}^{+\infty} g_1(t)g_2(t) dt - \int_a^b g_1(t)g_2(t) dt \right| < \eta \quad (51)$$

for given error bound η . Such interval can be constructed as $[B - \Delta; B + \Delta]$ to fulfil

$$\begin{aligned} \left| \int_{B+\Delta}^{+\infty} g_1(t)g_2(t) dt \right| &< \frac{1}{2}\eta \\ \left| \int_{-\infty}^{B-\Delta} g_1(t)g_2(t) dt \right| &< \frac{1}{2}\eta. \end{aligned} \quad (52)$$

Therefore, Δ can be calculated as

$$\Delta = \frac{1}{\sqrt{A}} \operatorname{erf}^{-1} \left(1 - \frac{\eta \epsilon^2}{\sqrt{2}} e^{-C} \sqrt{\frac{s_1}{s_2} + \frac{s_2}{s_1}} \right), \quad (53)$$

where erf^{-1} is an inverse of the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (54)$$

The value of Δ in (53) is well-defined unless inequality (49) is fulfilled, in which case the atoms are orthogonal to the given precision and there is no need to define integration interval.

To simplify formula (53), one can use inequality

$$\operatorname{erf}^{-1}(1 - x) \leq \sqrt{-\log x}, \quad (55)$$

which is fulfilled for all $x \in (0; 1]$. Therefore,

$$\Delta' = \frac{1}{\sqrt{A}} \sqrt{C - \log \left(\frac{\eta \epsilon^2}{\sqrt{2}} \sqrt{\frac{s_1}{s_2} + \frac{s_2}{s_1}} \right)} \quad (56)$$

is guaranteed to fulfil $\Delta' \geq \Delta$.

Optimal phase of a Gabor function

In the following we find an explicit formula for the phase ϕ_{\max} , that maximizes the product of signal x with Gabor function of given time position u , frequency ω , and scale s . Let us recall the formula (3) of a real Gabor function:

$$g_\gamma(t) = K(\gamma) e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(\omega(t - u) + \phi).$$

γ denotes the set of parameters $\gamma = \{u, s, \omega, \phi\}$ and $K(\gamma)$ is such that $\|g_\gamma\| = 1$. Writing $K(\gamma)$ explicitly gives

$$g_{(\gamma)}(t) = \frac{e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(\omega(t - u) + \phi)}{\|e^{-\pi \left(\frac{t-u}{s}\right)^2} \cos(\omega(t - u) + \phi)\|} \quad (57)$$

Phase shift ϕ can be also expressed as a superposition of two orthogonal oscillations $\cos(\omega(t-u) + \phi)$ and $\sin(\omega(t-u) + \phi)$. We define

$$\begin{aligned} C &= e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos(\omega(t-u)) \\ S &= e^{-\pi\left(\frac{t-u}{s}\right)^2} \sin(\omega(t-u)) \end{aligned} \quad (58)$$

and, using the trigonometric identity

$$\cos(\alpha + \phi) = \cos \alpha \cos \phi - \sin \alpha \sin \phi, \quad (59)$$

we write the Gabor function (57) as

$$g_\gamma(t) = \frac{C \cos \phi - S \sin \phi}{\|C \cos \phi - S \sin \phi\|} \quad (60)$$

Using $\|x\|^2 = \langle x, x \rangle$, and the orthogonality of C and S defined in (58) ($\langle C, S \rangle = 0$), we write the product of Gabor function defined as (60) with the signal x as

$$\langle x, g_\gamma \rangle = \frac{\langle x, C \rangle \cos \phi - \langle x, S \rangle \sin \phi}{\|C \cos \phi - S \sin \phi\|} = \frac{\langle x, C \rangle - \langle x, S \rangle \tan \phi}{\sqrt{\langle C, C \rangle + \langle S, S \rangle \tan^2 \phi}} \quad (61)$$

We are looking for the maximum absolute value of this product. For the sake of simplicity we will maximize $\langle x, g_\gamma \rangle^2$ instead of $|\langle x, g_\gamma \rangle|$. Denoting $v = \tan \phi$

$$\langle x, g_\gamma \rangle^2 = \frac{(\langle x, C \rangle - \langle x, S \rangle v)^2}{\langle C, C \rangle + \langle S, S \rangle v^2} \quad (62)$$

To find v which maximizes (62), we look for zeros of the derivative $\frac{\partial}{\partial v} \langle x, g_\gamma \rangle^2$ and find two roots:

$$v_1 = \frac{\langle x, C \rangle}{\langle x, S \rangle}, \quad v_2 = -\frac{\langle x, S \rangle \langle C, C \rangle}{\langle x, C \rangle \langle S, S \rangle}$$

Substituting these values for $\tan \phi$ in (61) we get

$$\begin{aligned} \langle x, g_\gamma \rangle|_{v=v_1} &= 0 \\ \langle x, g_\gamma \rangle|_{v=v_2} &= \frac{\langle x, C \rangle + \frac{\langle x, S \rangle^2 \langle C, C \rangle}{\langle x, C \rangle \langle S, S \rangle}}{\sqrt{\langle C, C \rangle + \frac{\langle x, S \rangle^2 \langle C, C \rangle^2}{\langle x, C \rangle^2 \langle S, S \rangle}}} \end{aligned}$$

Since for v_1 the square of the product is minimum (zero), the other extremum is a maximum in v_2 . Therefore, the phase ϕ that maximizes $\langle x, g_\gamma \rangle^2$ is given by

$$\phi_{\max} = \arctan \frac{\langle x, S \rangle / \langle S, S \rangle}{\langle x, C \rangle / \langle C, C \rangle} \quad (63)$$

and the maximum value of the product is

$$\langle x, g_\gamma \rangle_{\max} = \frac{\langle x, C \rangle^2 / \langle C, C \rangle + \langle x, S \rangle^2 / \langle S, S \rangle}{\sqrt{\langle x, C \rangle^2 / \langle C, C \rangle + \langle x, S \rangle^2 / \langle S, S \rangle}} \quad (64)$$

Applying Fast Fourier Transform

Estimation of the product of a Gabor function with signal x (61) requires computing of the inner product of signal x with functions C and S defined in Equation (58). Let us rewrite the formulae for $\langle x, C \rangle$ and $\langle x, S \rangle$:

$$\langle x, C \rangle = \int_{-\infty}^{\infty} x e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos(\omega(t-u)) \quad (65a)$$

$$\langle x, S \rangle = \int_{-\infty}^{\infty} x e^{-\pi\left(\frac{t-u}{s}\right)^2} \sin(\omega(t-u)) \quad (65b)$$

Multiplying (65a) by $i = \sqrt{-1}$ and subtracting it from (65b), one can obtain the following equation:

$$\langle x, C \rangle - i \langle x, S \rangle = \int_{-\infty}^{\infty} x e^{-\pi \left(\frac{t-u}{s}\right)^2} e^{-i\omega(t-u)} \quad (66)$$

The right side of Equation (66) is the Fourier Transform of signal x windowed by Gauss functions $e^{-\pi \left(\frac{t-u}{s}\right)^2}$. This formula allows for fast computation of inner products $\langle x, C \rangle$ and $\langle x, S \rangle$, since in an optimal dictionary the atoms with the given scale s are arranged in frequency domain with uniform step $\Delta\omega$ (see Equation (29)).

Additional structures in the dictionary

Apart from the Gabor functions, Gabor dictionary implemented in mp5 contains also the following functions:

- “Pure” harmonic waves

$$S(t) = K_s \cos(\omega t + \phi) \quad (67)$$

where K_s is normalization factor such that $\langle S(t), S(t) \rangle = 1$ on the analysed signal length. The phase ϕ of the signal $S(t)$ is estimated according to Equation (63).

Harmonic functions are distributed in frequency domain with step $\Delta\omega$ determined by Equation (29);

- Kronecker delta functions

$$\Delta(t - u) = \begin{cases} 1, & \text{for } t = u \\ 0, & \text{for } t \neq u \end{cases} \quad (68)$$

In this work, the delta functions are distributed across the whole time domain, that is for each point of the time series.

- “pure” Gaussians

$$G(t) = K_g e^{-\pi \frac{(t-u)^2}{s^2}} \quad (69)$$

Where K_g is normalization factor such that $\langle G(t), G(t) \rangle = 1$ on the analysed signal length. These functions are distributed in the scale domain with step a (see Equation (27)) and with step Δu (30) in the time domain.

Additional files

Additional file 1: A video tutorial for downloading and configuration of the Svarog package, used for computations and visualization of results.

Additional file 2: Shows the steps from loading the signal into Svarog to displaying interactive map of time-frequency energy distribution.

Additional file 3: Contains help of the MP module from Svarog.

Abbreviations

EEG: Electroencephalogram; FFT: Fast Fourier Transform; MEG: Magnetoencephalogram; MP: Matching pursuit; MMP: Multivariate matching pursuit; S/N: Signal to noise; SWA: Slow wave activity.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RK wrote from the scratch the the mp5 implementation in C, and contributed most of the text in Appendix Implementation and optimizations. PTR derived the formulae for product-related metric and resulting dictionary construction, adapted Svarog and its GUI to the modifications in the algorithm, and wrote most of the sections introducing optimal sampling of Gabor dictionaries. PJD proposed the idea of optimal sampling of Gabor dictionaries

based upon the product-related metrics in 2003, over the next decade supervised and coordinated projects which led to this article and the accompanying software, and wrote the remaining text. All authors read and approved the final manuscript.

Acknowledgements

This work was supported from Polish funds for science, including grant from the Polish Ministry of Science and Higher Education (Decision 644/N-COST/2010/0). Authors are grateful for constructive remarks of prof. Kenneth Foster on the manuscript organization and style of presentation. Finally, we thank our colleagues Dobiesław Ircha, Marek Barwiński and Artur Matysiak for years of collaboration on related issues and inspiring discussions.

Received: 3 August 2013 Accepted: 2 September 2013

Published: 23 September 2013

References

1. Durka PJ, Blinowska KJ: **Analysis of EEG transients by means of matching pursuit.** *Ann Biomed Eng* 1995, **23**:608–611.
2. Durka PJ, Szelenberger W, Blinowska K, Androsiuk W, Myszka M: **Adaptive time-frequency parametrization in pharmaco EEG.** *J Neurosci Methods* 2002, **117**:65–71.
3. Lelic D, Olesen AE, Brock C, Staahl C, Drewes AM: **Advanced pharmaco-EEG reveals morphine induced changes in the brain's pain network.** *J Clin Neurophysiol* 2012, **29**(3):219–225.
4. Koubeissi MZ, Jouny CC, Blakeley JO, Bergey GK: **Analysis of dynamics and propagation of parietal cingulate seizures with secondary mesial temporal involvement.** *Epilepsy Behav* 2009, **14**:108–112. [<http://www.sciencedirect.com/science/article/pii/S1525505008002746>]
5. Jouny CC, Adamolekun B, Franaszczuk PJ, Bergey GK: **Intrinsic ictal dynamics at the seizure focus: effects of secondary generalization revealed by complexity measures.** *Epilepsia* 2007, **48**(2):297–304. [<http://dx.doi.org/10.1111/j.1528-1167.2006.00963.x>]
6. Jouny CC, Franaszczuk PJ, Bergey GK: **Characterization of epileptic seizure dynamics using Gabor atom density.** *Clin Neurophysiol* 2003, **114**:426–437.
7. Jouny CC, Franaszczuk PJ, Bergey GK: **Signal complexity and synchrony of epileptic seizures: is there an identifiable preictal period?** *Clinph* 2005, **116**:552–558.
8. Bergey GK, Franaszczuk PJ: **Epileptic seizures are characterized by changing signal complexity.** *Clin Neurophysiol* 2001, **112**:241–249.
9. Wilson SB, Scheuer ML, Emerson RG, Gabor AJ: **Seizure detection: evaluation of the reveal algorithm.** *Clin Neurophysiol* 2004, **115**(10):2280–2291.
10. Zhang ZG, Yang JL, Chan SC, Luk K, Hu Y: **Time-frequency component analysis of somatosensory evoked potentials in rats.** *BioMed Eng OnLine* 2009, **8**:4. [<http://www.biomedical-engineering-online.com/content/8/1/4>]
11. Zhang Z, Luk KDK, Hu Y: **Identification of detailed time-frequency components in somatosensory evoked potentials.** *Neural Syst Rehabil Eng, IEEE Trans* 2010, **18**(3):245–254.
12. Zhang ZG, Yang JL, Chan SC, Luk K, Hu Y: **Time-frequency component analysis of somatosensory evoked potentials in rats.** *BioMed Eng OnLine* 2009, **8**:4. [<http://www.biomedical-engineering-online.com/content/8/1/4>]
13. Schönwald S, Carvalho D, de Santa-Helena E, Lemke N, L Gerhardt G: **Topography-specific spindle frequency changes in obstructive sleep apnea.** *BMC Neurosci* 2012, **13**:89. [<http://www.biomedcentral.com/1471-2202/13/89>]
14. Schönwald SV, Carvalho DZ, Dellagustin G, de Santa-Helena EL, Gerhardt GJ: **Quantifying chirp in sleep spindles.** *J Neurosci Methods* 2011, **197**:158–164. [<http://www.sciencedirect.com/science/article/pii/S0165027011000525>]
15. Cervenka MC, Franaszczuk PJ, Crone NE, Hong B, Caffo BS, Bhatt P, Lenz FA, Boatman-Reich D: **Reliability of early cortical auditory gamma-band responses.** *Clin Neurophysiology* 2013, **124**:70–82.
16. Ray S, Maunsell JHR: **Different origins of gamma rhythm and high-gamma activity in macaque visual cortex.** *PLoS Biol* 2011, **9**(4):e1000610.
17. Lelic D, Olesen SS, Valeriani M, Drewes AM: **Brain source connectivity reveals the visceral pain network.** *NeuroImage* 2012, **60**:37–46. [<http://www.sciencedirect.com/science/article/pii/S1053811911013991>]
18. Drewes AM, Gratkowski M, Sami SAK, Dimceviski G, Funch-Jensen P, Arendt-Nielsen L: **Is the pain in chronic pancreatitis of neuropathic origin? Support from EEG studies during experimental pain.** *World J Gastroenterol* 2008, **14**(25):4020–4027. [<http://www.biomedsearch.com/nih/pain-in-chronic-pancreatitis-neuropathic/18609686.html>]
19. Żygierewicz J, Kelly EF, Blinowska KJ, Durka PJ, Folger S: **Time-frequency analysis of vibrotactile driving responses by matching pursuit.** *J Neurosci Methods* 1998, **81**:121–129.
20. Durka PJ, Ircha D, Neuper C, Pfurtscheller G: **Time-frequency microstructure of event-related desynchronization and synchronization.** *Med Biol Eng Comput* 2001, **39**(3):315–321.
21. Durka PJ: **Time-frequency microstructure and statistical significance of ERD and ERS.** In *Progress in Brain Research. Volume 159*. Edited by Neuper C, Klimesch W: Elsevier BV; 2006:121–133.
22. Durka PJ, Matysiak A, Montes EM, Sosa PV, Blinowska KJ: **Multichannel matching pursuit and EEG inverse solutions.** *J Neurosci Methods* 2005, **148**:49–59.
23. Lelic D, Gratkowski M, Valeriani M, Arendt-Nielsen L, Drewes AM: **Inverse modeling on decomposed electroencephalographic data: a way forward?** *J Clin Neurophysiol* 2009, **26**(4):227–235. [<http://www.biomedsearch.com/nih/Inverse-modeling-decomposed-electroencephalographic-data/19584750.html>]
24. Zwoliński P, Roszkowski M, Żygierewicz J, Haufe S, Nolte G, Durka P: **Open database of epileptic EEG with MRI and postoperational assessment of foci—real world verification for the EEG inverse solutions.** *Neuroinformatics* 2010, **8**:285–299.

25. Bénar C, Papadopoulo T, Clerc M: **Topography time-frequency atomic decomposition for event related M/EEG signals.** In *Proceedings of 29th Annual International IEEE EMBS Conference*; 2007:5461–5464. [ftp://ftp-sop.inria.fr/odyssey/Publications/2007/benar-papadopoulo-etal:07.pdf]
26. Studer D, Hoffmann U, Koenig T: **From EEG dependency multichannel matching pursuit to sparse topographic decomposition.** *J Neurosci Methods* 2006, **153**(2):261–275.
27. Xu P, Yao D: **A novel method based on realistic head model for EEG denoising.** *Comput Methods Programs Biomed* 2006, **83**(2):104–110.
28. Sielużycki C, Kus R, Matysiak A, Durka P, Koenig R: **Multivariate matching pursuit in the analysis of single-trial latency of the auditory M100 acquired with MEG.** *Int J Bioelectromagnetism* 2009, **11**(4):155–160.
29. Sielużycki C, König R, Matysiak A, Kuś R, Ircha D, Durka P: **Single-trial evoked brain responses modeled by multivariate matching pursuit.** *IEEE Trans Biomed Eng* 2009, **56**:74–82.
30. Bénar C, Papadopoulo T, Torrèsani B, Clerc M: **Consensus matching pursuit for multi-trial EEG signals.** *J Neurosci Methods* 2009, **180**:161–170. [http://www.sciencedirect.com/science/article/B6T04-4VWHVX5-2/2/e6ebdc581a60cde843503fe30f9940d1]
31. Jörn M, Sielużycki C, Matysiak M, Żygierewicz J, Scheich H, Durka P, König R: **Single-trial reconstruction of auditory evoked magnetic fields by means of template matching pursuit.** *J Neurosci Methods* 2011, **199**:119–128. [http://www.sciencedirect.com/science/article/pii/S0165027011002238]
32. Durka PJ: **On the methodological unification in electroencephalography.** *BioMed Eng OnLine* 2005, **4**(15).
33. Żygierewicz J, Blinowska KJ, Durka PJ, Szelenberger W, Niemcewicz S, Androsiuk W: **High resolution study of sleep spindles.** *Clin Neurophysiol* 1999, **110**(12):2136–2147.
34. Durka PJ, Malinowska U, Szelenberger W, Wakarow A, Blinowska KJ: **High resolution parametric description of slow wave sleep.** *J Neurosci Methods* 2005, **147**:15–21.
35. Durka PJ: **Adaptive time-frequency parametrization of epileptic EEG spikes.** *Phys Rev E* 2004, **69**(051914). [http://pre.aps.org/abstract/PRE/v69/i5/e051914]
36. Nuwer M: **Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society.** *Neurology* 1997, **49**:277–292.
37. Rechtschaffen A, Kales A (Eds): *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages in Human Subjects. No. 204 in National Institutes of Health Publications.* Washington DC: US Government Printing Office; 1968.
38. Iber C, Ancoli-Israel S, Chesson A, Quan S: *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification.* American Academy of Sleep Medicine; 2007.
39. Mallat S, Zhang Z: **Matching Pursuit with time-frequency dictionaries.** *IEEE Trans Signal Process* 1993, **41**:3397–3415.
40. Malinowska U, Klekowicz H, Wakarow A, Niemcewicz S, Durka P: **Fully parametric sleep staging compatible with the classical criteria.** *Neuroinformatics* 2009, **7**(4):245–253.
41. Schonwald S, Desantahelena E, Rossatto R, Chaves M, Gerhardt G: **Benchmarking matching pursuit to find sleep spindles.** *J Neurosci Methods* 2006, **156**(1–2):314–321. [http://dx.doi.org/10.1016/j.jneumeth.2006.01.026]
42. Durka PJ, Ircha D, Blinowska KJ: **Stochastic time-frequency dictionaries for matching pursuit.** *IEEE Trans Signal Process* 2001, **49**(3):507–510.
43. Vleeschouwer CD, Zakhor A: **In-loop atom modulus quantization for matching pursuit and its application to video coding.** *IEEE Trans Image Process* 2003, **12**(10):1226–1242.
44. Durka PJ: *Matching Pursuit and Unification in EEG Analysis.* Artech House; 2007. [Engineering in Medicine and Biology], [ISBN 978-1-58053-304-1]
45. Ircha D: **MP4—software for matching pursuit with stochastic Gabor dictionaries.** [http://eeg.pl/mp20002009]
46. Tropp JA: **Constructing packings in projective spaces and Grassmannian spaces via alternating projection.** ICES Report 04-23, UT-Austin 2004.
47. Ferrando SE, Doolittle EJ, Bernal AJ, Bernal LJ: **Probabilistic matching pursuit with Gabor dictionaries.** *Signal Process* 2000, **80**(10):2099–2120.
48. Gribonval R: **Piecewise linear source separation.** In *Proc. SPIE 03, Volume 5207 Wavelets: Applications in Signal and Image Processing.* San Diego; 2003. [http://spiedigitallibrary.org/volume.aspx?volumeid=2241]
49. Gribonval R: **Sparse decomposition of stereo signals with Matching Pursuit and application to blind separation of more than two sources from a stereo mixture.** *Acoustics, Speech, Signal Process, Proc ICASSP'02, Orlando, Florida, USA* 2002, **3**:3057–3060.
50. Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J: **Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human.** *J Neurosci* 1996, **16**(13):4240–4249.
51. Delorme A, Makeig S: **EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis.** *J Neurosci Methods* 2004, **134**:9–21.
52. Barwiński M: **Product-based metric for Gabor functions and its implications for the matching pursuit algorithm.** *Master's thesis.* Warsaw University, Institute of Experimental Physics 2004. http://eeg.pl/Members/mbarwinski/m.sc.-on-matching-pursuit-theory
53. Ircha D: **Reprezentacje sygnałów w redundantnych zbiorach funkcji.** *Master's thesis.* University of Warsaw, Faculty of Physics 1997.

doi:10.1186/1475-925X-12-94

Cite this article as: Kuś et al.: Multivariate matching pursuit in optimal Gabor dictionaries: theory and software with interface for EEG/MEG via Svarog. *BioMedical Engineering OnLine* 2013 **12**:94.