Atomic Signal Models Based on Recursive Filter Banks

Michael Goodwin
University of California at Berkeley
michaelg@eecs.berkeley.edu

Martin Vetterli
École Polytechnique Fédérale de Lausanne
Martin.Vetterli@de.epfl.ch

Abstract

Time-frequency atomic models are useful for signal analysis, modification, and coding, especially when the time-frequency behavior of the atoms matches the behavior of the signal. Such adaptive representations can be derived using the matching pursuit algorithm with an overcomplete dictionary of time-frequency atoms. In this paper, we consider matching pursuit with atoms constructed by coupling causal and anticausal damped sinusoids. These provide advantages over symmetric Gabor atoms for modeling signals with transient behavior, such as music. Furthermore, the matching pursuit computation is simplified by the structure of the atoms; expansions based on these atoms can be derived using simple recursive filter banks.

1 Atomic Decompositions

Time-frequency atomic signal representations have been of growing interest since their explication by Gabor several decades ago [1, 2]. The fundamental principles of atomic modeling are that a signal can be decomposed into elementary functions that are localized in time-frequency and that such decompositions are useful for applications such as signal analysis, modification, and coding. Familiar examples of atomic models include basis expansions such as Fourier and wavelet representations. Such basis expansions, however, exhibit substantial limitations for representing arbitrary signals; for instance, a Fourier basis provides a poor representation of a time-localized signal. These limitations result from the attempt to represent arbitrary signal behavior in terms of a limited set of expansion functions. This problem can be resolved by using overcomplete expansions, wherein the expansion functions for a signal decomposition are chosen from a highly redundant set whose elements exhibit a wide range of time-frequency behaviors. Overcomplete expansions based on time-frequency dictionaries can be derived using the matching pursuit algorithm [3].

2 Matching Pursuit

Matching pursuit is a greedy iterative algorithm for deriving signal decompositions of the form

$$x[n] \approx \sum_i \alpha_i g_i[n], \tag{1}$$

where the functions $g_i[n]$ are chosen from a dictionary of atoms [3]. The pursuit starts by choosing the atom that best approximates the signal in a two-norm sense. The contribution of this atom is subtracted from the signal and the process is iterated on the residual. For a dictionary D, the task at the i-th stage of the algorithm is thus to find the atom $d_{m(i)}[n] \in D$ that minimizes the two-norm of the residual

$$r_{i+1}[n] = r_i[n] - \alpha_i d_{m(i)}[n], \tag{2}$$

where $m(i)$ is a dictionary index and α_i is the coefficient of the atom in the model being derived; note that $r_1[n] = x[n]$. The atom chosen at the i-th stage will hereafter be denoted by $g_i[n] = d_{m(i)}[n]$, where i refers to the iteration when $g_i[n]$ was chosen and $m(i)$ is the index of the atom in D.

Treating the signals as column vectors, the optimal atom to choose at the i-th stage can be expressed as

$$g_i = \arg\min_{g_i \in D} ||r_{i+1}||^2 = \arg\min_{g_i \in D} ||r_i - \alpha_i g_i||^2. \tag{3}$$

For unit-norm atoms, Eq. (3) can be reformulated as

$$g_i = \arg\max_{g_i \in D} |\langle g_i, r_i \rangle|. \tag{4}$$

The coefficient for the model is $\alpha_i = \langle g_i, r_i \rangle = g_i^H r_i$.

To enable representation of a wide range of signal features, a large dictionary of atoms is used in the matching pursuit algorithm. The computation of the correlations $\langle g, r_i \rangle$ for all $g \in D$ is thus expensive. The computational cost can be reduced by using an update formula based on Eq. (2); the correlations at stage $i+1$ are given by

$$\langle g, r_{i+1} \rangle = \langle g, r_i \rangle - \alpha_i \langle g, g_i \rangle, \tag{5}$$

where the dictionary cross-correlation term $\langle g, g_i \rangle$ can be precomputed and stored [3].

1058-6393/98 $10.00

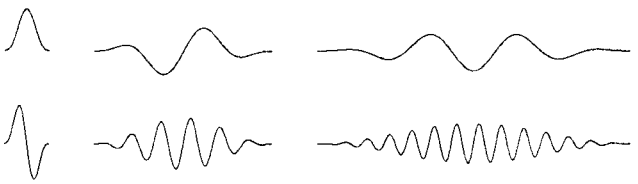

Figure 1. Gabor atoms derived from a symmetric window by scaling, modulation, and translation.

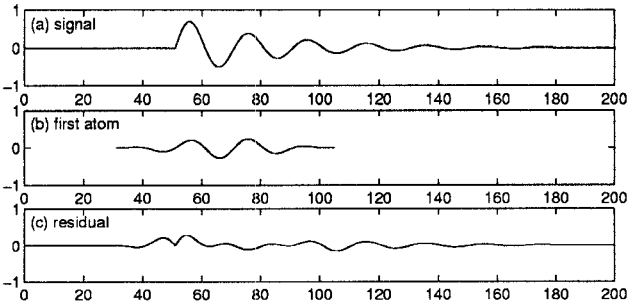

Figure 2. Pre-echo in atomic models: (a) the onset of a damped sinusoid, (b) the first atom chosen from a symmetric Gabor dictionary by matching pursuit, and (c) the residual.

3 Time-Frequency Dictionaries

In an atomic decomposition derived by matching pursuit, the atoms correspond to basic signal features. This is especially useful for analysis and coding if the atoms can be described by meaningful parameters such as time location, frequency modulation, and scale; then, the basic signal features can be identified and parameterized. Matching pursuit using a parametric overcomplete dictionary consisting of a wide variety of time-frequency atoms provides a compact, adaptive, and parametric signal model [3, 4].

3.1 Gabor Atoms

In matching pursuit, dictionaries of Gabor atoms have been of primary interest since these are appropriate functions for general time-frequency signal expansions [1, 3]. In continuous time, such atoms are derived from a unit-norm window $g(t)$ by scaling, modulation, and translation:

$$g_{\{s,\omega,\tau\}}(t) = \frac{1}{\sqrt{s}} g\left(\frac{n-\tau}{s}\right) e^{j\omega(n-\tau)}. \quad (6)$$

A sampling argument indicates that Gabor atoms can be expressed in discrete time as

$$g_{\{s,\omega,\tau\}}[n] = f_s[n-\tau_0]e^{j\omega(n-\tau)}, \quad (7)$$

where $f_s[n]$ is a unit-norm window of scale s [3]. Examples of Gabor atoms are given in Fig. 1. Note that these atoms can be indexed in a dictionary by the parameter set $\{s, \omega, \tau\}$; this parametric structure allows for a simple description of a specific dictionary, which is useful for compression. Also, note that the modulation of an atom can be defined independently of the time shift, or *dereferenced*, which will be useful in later considerations:

$$\tilde{g}_{\{s,\omega,\tau\}}[n] = f_s[n-\tau_0]e^{j\omega n} = e^{j\omega\tau} g_{\{s,\omega,\tau\}}[n]. \quad (8)$$

In applications of Gabor functions, $g[n]$ is typically an even-symmetric window. The associated dictionaries thus consist of atoms that exhibit symmetric time-domain behavior. This is problematic for modeling asymmetric features such as transients, which occur frequently in natural signals. Fig. 2 shows the pre-echo artifact that is introduced in the reconstruction if a simple transient is modeled with a symmetric Gabor dictionary [5]. The residual has a discontinuity at the onset time; later stages of the pursuit must incorporate small-scale atoms into the decomposition to model this discontinuity. One approach to this problem is the high-resolution pursuit proposed in [6, 7], where symmetric atoms are still used but the selection metric is modified so that atoms that introduce drastic artifacts are not chosen for the decomposition. Given that symmetric functions are intrinsically not well suited for modeling asymmetric events, another approach of interest for modeling signals with transient behavior is the use of a dictionary of asymmetric atoms, *e.g.* damped sinusoids.

3.2 Damped Sinusoids

In [4, 5, 8], time-frequency dictionaries of damped sinusoids are used to improve the representation of transients; such approaches are motivated by the common occurrence of damped oscillations in natural signals. Like the atoms in a general Gabor dictionary, damped sinusoidal atoms can be indexed by characteristic parameters, namely the damping factor a, modulation frequency ω, and start time τ:

$$g_{\{a,\omega,\tau\}}[n] = S_a\, a^{(n-\tau)} e^{j\omega(n-\tau)} u[n-\tau], \quad (9)$$

or, if the modulation is dereferenced,

$$\tilde{g}_{\{a,\omega,\tau\}}[n] = S_a\, a^{(n-\tau)} e^{j\omega n} u[n-\tau], \quad (10)$$

where the factor S_a accounts for unit-norm scaling. Examples are given in Fig. 3. Note that these atoms correspond to Gabor functions derived from a one-sided exponential window; the atomic structure is more readily indicated by a damping factor than a scale parameter, however, so the index set $\{a, \omega, \tau\}$ is used instead of the set $\{s, \omega, \tau\}$.

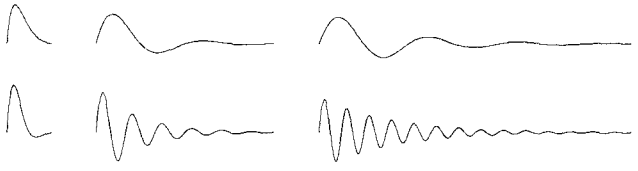

Figure 3. Damped sinusoids: Gabor atoms based on a one-sided exponential window.

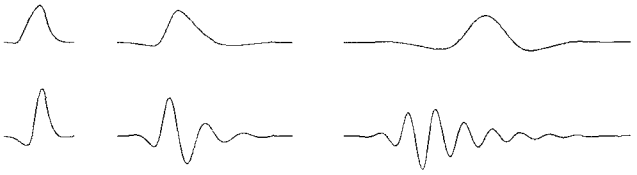

Figure 4. Composite atoms constructed by coupling causal and anticausal damped sinusoids and using low-order smoothing.

3.3 Composite Atoms

The simple example of Fig. 2 shows that symmetric atoms are inappropriate for modeling some signals. It is simple to construct examples for which asymmetric atoms prove similarly ineffective. Given the task of modeling arbitrary signals, it can thus be argued that a wide range of both symmetric and asymmetric atoms should be present in the dictionary. Such *composite* dictionaries are considered here. Due to computational advantages that will be discussed later, atoms constructed by coupling causal and anticausal damped sinusoids are of particular interest. These atoms take the form

$$g_{\{a,b,J,\omega,\tau\}}[n] = f_{\{a,b,J\}}[n-\tau]e^{j\omega(n-\tau)}, \quad (11)$$

or, if the modulation is dereferenced,

$$\tilde{g}_{\{a,b,J,\omega,\tau\}}[n] = f_{\{a,b,J\}}[n-\tau]e^{j\omega n}, \quad (12)$$

where the amplitude envelope is a unit-norm function constructed using a causal and an anticausal exponential:

$$f_{\{a,b,J\}}[n] = \quad (13)$$
$$S_{\{a,b,J\}}\left(a^n u[n] + b^{-n}u[-n] - \delta[n]\right) * h_J[n],$$

where $\delta[n]$ is subtracted because the causal and anticausal components overlap at $n = 0$. The function $h_J[n]$ is a smoothing window of length J. A variety of composite atoms are depicted in Fig. 4; these can exhibit a wide range of time-frequency behaviors and can be described by the simple parameter set $\{a, b, J, \omega, \tau\}$.

The unit-norm scaling factor for a composite atom is

$$S_{\{a,b,J\}} = \frac{1}{\sqrt{\Upsilon(a,b,J)}}, \quad (14)$$

where $\Upsilon(a,b,J)$ is given by

$$\sum_n \left|\left(a^n u[n] + b^{-n}u[-n] - \delta[n]\right) * h_J[n]\right|^2 = \quad (15)$$

$$\sum_{l=0}^{J-1}\sum_{k=0}^{J-1} \frac{a^{|l-k|}}{1-a^2} + \frac{b^{|l-k|}}{1-b^2} + \frac{a^{|l-k|}b - ab^{|l-k|}}{a-b}. \quad (16)$$

Precomputation of this scale factor is required for pursuit based on these atoms.

The composite atoms described above can be written in terms of unit-norm constituent atoms:

$$\tilde{g}_{\{a,b,J,\omega,\tau\}}[n] = \quad (17)$$
$$S_{\{a,b,J\}}\left(\frac{\tilde{g}^+_{\{a,\omega,\tau\}}[n]}{S_a} + \frac{\tilde{g}^-_{\{b,\omega,\tau\}}[n]}{S_b} - \delta[n]\right) * h_J[n],$$

where $\tilde{g}^+_{\{a,\omega,\tau\}}[n]$ is a causal atom and $\tilde{g}^-_{\{b,\omega,\tau\}}[n]$ is an anticausal atom defined as

$$\tilde{g}^-_{\{b,\omega,\tau\}}[n] = S_b\, b^{-(n-\tau)}e^{j\omega n}u[-(n-\tau)]. \quad (18)$$

For a rectangular $h_J[n]$, the convolution in Eq. (17) can be rewritten as

$$\sum_{\Delta=0}^{J-1} \frac{\tilde{g}^+_{\{a,\omega,\tau+\Delta\}}[n]}{S_a} + \frac{\tilde{g}^-_{\{b,\omega,\tau+\Delta\}}[n]}{S_b} - \delta[n+\Delta]. \quad (19)$$

Note that dereferenced atoms are used in the construction so that their modulations add coherently in the sum over the time lags Δ; otherwise, a phase shift of $\omega\Delta$ would be required for each atom to achieve coherent modulation of the composite atom. As will be seen in Section 4.2, this construction framework leads to a simple relationship between the correlations of the signal with the composite atom and with the underlying damped sinusoids.

The special case of symmetric atoms ($a = b$), one example of which is shown in Fig. 4, suggests the use of this approach to construct atoms similar to symmetric Gabor atoms based on common windows. Given a unit-norm window $w[n]$, the issue is to choose a damping factor a and a smoothing order J such that the resultant $f_{\{a,a,J\}}[n]$ is similar to $w[n]$. Using the two-norm as an accuracy metric, the objective is to minimize the error

$$\epsilon(a,J) = ||f_{\{a,a,J\}}[n] - w[n]||^2 \quad (20)$$

by optimizing a and J. Since $f_{\{a,a,J\}}[n]$ and $w[n]$ are both

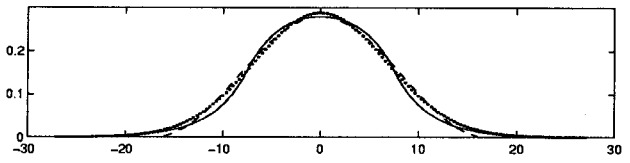

Figure 5. The envelope for a symmetric composite atom (solid) can be constructed to match a Hanning window (dashed) and a Gaussian window (dotted).

unit-norm, this expression can be simplified to:

$$\epsilon(a, J) = 2\left(1 - \sum_n f_{\{a,a,J\}}[n]w[n]\right). \quad (21)$$

The overall objective of the optimization is thus to maximize the correlation of $f_{\{a,a,J\}}[n]$ and $w[n]$,

$$\hat{\epsilon}(a, J) = \sum_n f_{\{a,a,J\}}[n]w[n]. \quad (22)$$

In an implementation, this would not be an on-line operation but rather a precomputation indicating values of a and J to be used in the parameter set of the composite dictionary; interestingly, this precomputation itself resembles a matching pursuit. Note that the values of a and J for the $f_{\{a,a,J\}}[n]$ in the composite dictionary are based on the scales of symmetric behavior to be included in the dictionary. Presumably, closed form solutions for a and J can be found for some particular windows; such solutions are of course limited by the requirement that J be an integer. The intent of this discussion, however, is not to investigate a window matching algorithm but instead to provide an existence proof that symmetric atoms constructed with one-sided exponentials can reasonably mimic Gabor atoms based on standard symmetric windows. Fig. 5 shows an example of an unmodulated composite atom that roughly matches a Hanning window and a Gaussian window.

It has been demonstrated that a composite dictionary containing a wide range of symmetric and asymmetric atoms can be constructed from uniform dictionaries of causal and anticausal damped sinusoids. Atoms resembling common symmetric Gabor atoms can readily be generated, meaning that this approach can be tailored to include standard symmetric atoms as a dictionary subset; there is no generality lost by constructing atoms in this fashion. As will be shown in Section 4.1, the pursuit computation for dictionaries of damped sinusoids is inexpensive; this leads to the algorithm of Section 4.2, namely a low cost implementation of matching pursuit based on a composite dictionary.

4 Computation Using Recursive Filter Banks

For arbitrary dictionaries, the computational cost of the matching pursuit iteration can be reduced using the update relation in Eq. (5). For dictionaries consisting of damped sinusoids or composite atoms constructed as described in Section 3.3, the correlation computation can be carried out with simple recursive filter banks. This computational framework is developed in the following two sections.

4.1 Pursuit of Damped Sinusoidal Atoms

In matching pursuit using a dictionary of complex damped sinusoids, correlations must be computed for every combination of damping factor, modulation frequency, and time shift. The correlation of a signal $x[n]$ with a causal atom $g^+_{\{a,\omega,\tau\}}[n]$ is given by

$$\eta_+(a, \omega, \tau) = S_a \sum_{n=\tau}^{\tau+L-1} x[n]\, a^{(n-\tau)} e^{-j\omega(n-\tau)} \quad (23)$$

where L is a truncation length that can be imposed on the atom. Truncation will not be considered further here; more details can be found in [4, 5]. A further simplification is that correlations with unnormalized atoms will be considered:

$$\rho_+(a, \omega, \tau) = \sum_{n=\tau}^{\tau+L-1} x[n]\, a^{(n-\tau)} e^{-j\omega(n-\tau)} \quad (24)$$

$$= \frac{\eta_+(a, \omega, \tau)}{S_a}. \quad (25)$$

Formulating the algorithm in terms of unnormalized atoms will reduce the computational cost of the algorithm developed in Section 4.2 for pursuing composite atoms.

The exponential structure of the atoms can be used to reduce the cost of the correlation computation over the time index; correlations at neighboring times are related by a simple recursion:

$$\rho_+(a, \omega, \tau - 1) = ae^{-j\omega}\rho_+(a, \omega, \tau) + x[\tau - 1]. \quad (26)$$

This is just a one-pole filter operated in reversed time to make the recursion stable for causal damped sinusoids; the similar forward recursion is unstable for $a < 1$. For anticausal atoms, the correlations are given by the recursion

$$\rho_-(b, \omega, \tau + 1) = be^{j\omega}\rho_-(b, \omega, \tau) + x[\tau + 1], \quad (27)$$

which is operated in forward time for the sake of stability. This formulation shows that the structure of damped sinusoidal atoms can be exploited to simplify the correlation computation irrespective of the update formula in Eq. (5). Similar simplifications based on recursive computation have been reported in the literature for short-time

Fourier transforms using exponential windows [9] as well as more general cases [10].

The results given above hold for atoms whose modulation is referenced to the atomic time origin as in Eqs. (6), (9), and (11). This local time reference leads to an interpretation of the pursuit computation in terms of a recursive filter bank. For the construction and pursuit of composite atoms, however, the dereferenced atoms defined in Eqs (8), (10), and (12) are of importance. The correlation formulae for dereferenced damped sinusoids can be derived by combining Eq. (8) with Eq. (24) to arrive at:

$$\tilde{\rho}_+(a,\omega,\tau) \;=\; e^{-j\omega\tau}\rho_+(a,\omega,\tau). \tag{28}$$

Eqs. (26) and (27) can then be reformulated as

$$\tilde{\rho}_+(a,\omega,\tau-1) \;= \tag{29}$$

$$a\tilde{\rho}_+(a,\omega,\tau) + e^{-j\omega(\tau-1)}x[\tau-1]$$

$$\tilde{\rho}_-(b,\omega,\tau+1) \;= \tag{30}$$

$$b\tilde{\rho}_+(b,\omega,\tau) + e^{-j\omega(\tau+1)}x[\tau+1].$$

For dereferenced modulation, the correlation computation thus corresponds to modulating the signal and using baseband filters [4]. As will be seen in the next section, dereferencing the modulation leads to a simple relationship between the signal correlations with composite atoms and the correlations with underlying damped sinusoids.

4.2 Pursuit of Composite Atoms

Using matching pursuit to derive a signal model based on composite atoms requires computation of the correlations of the signal with these atoms. Recalling the form of the composite atoms given in Eqs. (17) and (19), these correlations have, by construction, a simple relationship to the correlations with the underlying one-sided atoms:

$$\tilde{\rho}(a,b,J,\omega,\tau) \;=\; S_{\{a,b,J\}} \times \tag{31}$$

$$\sum_{\Delta=0}^{J-1}\left[\frac{\tilde{\eta}_+(a,\omega,\tau+\Delta)}{S_a} \;+\; \frac{\tilde{\eta}_-(b,\omega,\tau+\Delta)}{S_b} \;-\; x[\tau+\Delta]\right]$$

The sum can be rewritten in terms of correlations with unnormalized atoms:

$$\sum_{\Delta=0}^{J-1}\left[\tilde{\rho}_+(a,\omega,\tau+\Delta) \;+\; \tilde{\rho}_-(b,\omega,\tau+\Delta) \;-\; x[\tau+\Delta]\right]. \tag{32}$$

The correlation with any composite atom can thus be computed based on the correlations derived by the recursive filter banks discussed earlier; this computation is most straightforward if the underlying atoms are dereferenced and unnormalized. Once the underlying correlations have been derived, any atom constructed according to Eq. (19) can be added to the modeling dictionary at the cost of one multiply per atom to account for scaling. For a detailed computational consideration, including a comparison with the update approach of Eq. (5), the reader is referred to [4].

Finally, it should be noted that the computations given herein for atoms with complex modulation can be readily extended for deriving real models of real signals [5].

5 Conclusion

Atoms based on underlying exponential structures can exhibit a wide range of symmetric and asymmetric time-frequency behaviors. Dictionaries consisting of such atoms are useful for matching pursuit since the required correlation computations can be carried out with simple recursive filter banks. Future work involves application of such atomic models for audio coding and modification.

References

[1] D. Gabor. Theory of communication. *Journal of the Institute of Electrical Engineer*, 93:429–457, 1946.

[2] D. Gabor. Acoustical quanta and the theory of hearing. *Nature*, 159(4044):591–594, May 1947.

[3] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, December 1993.

[4] M. Goodwin. *Adaptive Signal Models: Theory, Algorithms, and Audio Applications*. PhD thesis, University of California at Berkeley, Fall 1997.

[5] M. Goodwin. Matching pursuit with damped sinusoids. *ICASSP-1997*, 3:2037–2040.

[6] S. Jaggi et al. High resolution pursuit for feature extraction. Technical Report LIDS-P-2371, MIT, November 1996.

[7] R. Gribonval et al. Analysis of sound signals with high resolution matching pursuit. *Proceedings of the TFTS Symposium*, pages 125–128, 1996.

[8] B. Friedlander and A. Zeira. Oversampled Gabor representation for transient signals. *IEEE Transactions on Signal Processing*, 43(9):2088–2094, Sept. 1995.

[9] S. Tomažic. On short-time Fourier transform with single-sided exponential window. *Signal Processing*, 55(2):141–148, December 1996.

[10] M. Unser. Recursion in short-time signal analysis. *Signal Processing*, 5(3):229–240, May 1983.