# Adaptive Time-Frequency Decompositions with Matching Pursuit

Geoff Davis, Stéphane Mallat [1] and Zhifeng Zhang

New York University, Courant Institute
251 Mercer Street, New York, NY 10012

**Abstract**

To compute the optimal expansion of signals in redundant dictionary of waveforms is an NP complete problem. We introduce a greedy algorithm, called matching pursuit, that performs a sub-optimal expansion. The waveforms are chosen iteratively in order to best match the signal structures. Matching pursuits are general procedures to compute adaptive signal representations. With a dictionary of Gabor functions, a matching pursuit defines an adaptive time-frequency transform. We derive a signal energy distribution in the time-frequency plane, which does not include interference terms, unlike Wigner and Cohen class distributions. A matching pursuit is a chaotic map, whose attractor defines a generic noise with respect to the dictionary. We derive an algorithm that isolates the coherent structures of a signal and an application to pattern extraction from noisy signals is described.

## 1 Introduction

Flexible decompositions are particularly important for representing signal components whose localizations in time and frequency vary widely. The signal must be expanded into waveforms whose time-frequency properties are adapted to its local structures. Such waveforms are called time-frequency atoms. For example, impulses need to be decomposed over functions well concentrated in time, while spectral lines are better represented by waveforms which have a narrow frequency support. When the signal includes both of these elements, the time-frequency atoms must be adapted accordingly. One must therefore introduce a procedure that chooses the waveforms that are best adapted to decompose the signal structures, among all the time-frequency atoms of a large dictionary.

Computing the optimal approximation of a signal in a redundant dictionary is an NP complete problem. We introduce a sub-optimal greedy algorithm called matching pursuit, that decomposes any signal into a linear expansion of waveforms that belong to a redundant dictionary of functions. These waveforms are selected in order to best match the signal structures. Although a matching pursuit is non-linear, like an orthogonal expansion, it maintains an energy conservation which guaranties its convergence.

The application of matching pursuits to adaptive time-frequency decompositions is described in section 5. The signal is decomposed into waveforms selected among a dictionary of time-frequency atoms, that are the the dilations, translations and modulations of a single window function. We derive a time-frequency energy distribution, by adding the Wigner distribution of the selected time-frequency atoms. Contrarily to the Wigner distribution or Cohen's class distributions, this energy distribution does not include interference terms and thus provides a clear picture in the time-frequency plane. A matching pursuit defines a choatic map whose properties are studied. The residue converges to an attractor that corresponds to a class of signals that do not include any coherent structures, with respect to the dictionary. These properties are used to isolate the coherent patterns of noisy signals.

[1] This work was supported by the AFOSR grant F49620-93-1-0102, ONR grant N00014-91-J-1967 and the Alfred Sloan Foundation

 0-8194-1546-4/94/$6.00

**Notations**

The space $\mathbf{L}^2(\mathbf{R})$ is the Hilbert space of complex valued functions such that

$$\|f\|^2 = \int_{-\infty}^{+\infty} |f(t)|^2 dt < +\infty. \tag{1}$$

The inner product of $(f, g) \in \mathbf{L}^2(\mathbf{R})^2$ is defined by

$$< f, g >= \int_{-\infty}^{+\infty} f(t)\bar{g}(t)dt, \tag{2}$$

where $\bar{g}(t)$ is the complex conjugate of $g(t)$. The Fourier transform of $f(t) \in \mathbf{L}^2(\mathbf{R})$ is written $\hat{f}(\omega)$ and defined by

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}dt. \tag{3}$$

# 2 Time-Frequency Atoms

A general family of time-frequency atoms can be generated by scaling, translating and modulating a single window function $g(t) \in \mathbf{L}^2(\mathbf{R})$. We suppose that $g(t)$ is real and centered at 0. We also impose that $\|g\| = 1$, that the integral of $g(t)$ is non-zero and that $g(0) \neq 0$. For any scale $s > 0$, frequency modulation $\xi$ and translation $u$, we denote $\gamma = (s, u, \xi)$ and define

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g(\frac{t-u}{s}) e^{i\xi t}. \tag{4}$$

The index $\gamma$ is an element of the set $\mathbf{\Gamma} = \mathbf{R}^+ \times \mathbf{R}^2$. The factor $\frac{1}{\sqrt{s}}$ normalizes to 1 the norm of $g_\gamma(t)$. The function $g_\gamma(t)$ is centered at the abscissa $u$ and its energy is concentrated in a neighborhood of $u$, whose size is proportional to $s$. Its Fourier transform is centered at the frequency $\omega = \xi$ and has an energy mostly concentrated in a neighborhood of $\xi$, whose size is proportional to $1/s$.

The dictionary of time-frequency atoms $\mathcal{D} = (g_\gamma(t))_{\gamma \in \mathbf{\Gamma}}$ is a very redundant set of functions that includes window Fourier frames and wavelet frames.[2] When the signals include time-frequency structures of very different types, one can not choose a priori a frame that is well adapted to perform the expansion. For any given signal, we rather need to find the best atoms of the dictionary that match the signal structures, in order to perform a compact decomposition. Next section studies such adaptive decompositions in redundant dictionaries.

# 3 Matching Pursuit

Let $\mathbf{H}$ be a signal space. We define a dictionary as a family $\mathcal{D} = (g_\gamma)_{\gamma \in \mathbf{\Gamma}}$ of vectors in $\mathbf{H}$, such that $\|g_\gamma\| = 1$. We impose that linear expansion of vectors in $\mathcal{D}$ are dense in $\mathbf{H}$. In general $\mathcal{D}$ is a very redundant family of vectors. A signal does not have a unique representation as a sum of elements of a redundant dictionary. Unlike the case of a basis, we have some degrees of freedom in choosing a signal's particular representation. This freedom allows us to choose a subset of the dictionary that is tailored to the signal in questions and which provides the most compact representation. We can choose a subset of the dictionary for which the signal energy is concentrated in as few terms as possibles. The chosen vectors highlight the predominant signal features.

Let $\mathcal{D}$ be dictionary of vectors in $\mathbf{R}^\mathbf{N}$. For any given $\alpha \in ]0, 1[$, we define an optimal approximation of $f \in \mathbf{R}^\mathbf{N}$

to be an expansion

$$\tilde{f} = \sum_{n=1}^{\alpha N} a_n g_{\gamma_n},$$

where $a_n$ and $g_{\gamma_n} \in \mathcal{D}$ are chosen in order to minimize

$$||f - \tilde{f}||.$$

When the dictionary is redundant, computing the optimal solution requires to much computations. More precisely, if we restrict the number of bits of $a_n$ and the number of vectors in $\mathcal{D}$ to $\Theta(N^k)$, for a fixed $k$, then one can prove[3] that computing the optimal solution is an $NP$ complete problem.

Because of the impossibility to compute numerically an optimal solution, we develop a greedy algorithm that computes a good sub-optimal approximation. Let $f \in \mathbf{H}$. We want to compute a linear expansion of $f$ over a set of vectors selected from $\mathcal{D}$, in order to best match its inner structures. This is done by successive approximations of $f$ with orthogonal projections on elements of $\mathcal{D}$. Let $g_{\gamma_0} \in \mathcal{D}$. The vector $f$ can be decomposed into

$$f = < f, g_{\gamma_0} > g_{\gamma_0} + Rf, \tag{5}$$

where $Rf$ is the residual vector after approximating $f$ in the direction of $g_{\gamma_0}$. Clearly $g_{\gamma_0}$ is orthogonal to $Rf$, hence

$$\|f\|^2 = |< f, g_{\gamma_0} >|^2 + \|Rf\|^2. \tag{6}$$

To minimize $\|Rf\|$, we must choose $g_{\gamma_0} \in \mathcal{D}$ such that $|< f, g_{\gamma_0} >|$ is maximum. In some cases, it is only possible to find a vector $g_{\gamma_0}$ that is almost the best in the sense that

$$|< f, g_{\gamma_0} >| \geq \alpha \sup_{\gamma \in \mathbf{\Gamma}} |< f, g_\gamma >|, \tag{7}$$

where $\alpha$ is an optimality factor that satisfies $0 < \alpha \leq 1$.

A matching pursuit is an iterative algorithm that sub-decomposes the residue $Rf$ by projecting it on a vector of $\mathcal{D}$ that matches $Rf$ almost at best, as it was done for $f$. This procedure is repeated each time on the following residue that is obtained.

Let us explain by induction, how the matching pursuit is carried further. Let $R^0 f = f$. We suppose that we have computed the $n^{th}$ order residue $R^n f$, for $n \geq 0$. We choose, with the choice function $C$, an element $g_{\gamma_n} \in \mathcal{D}$ which closely matches the residue $R^n f$

$$|< R^n f, g_{\gamma_n} >| \geq \alpha \sup_{\gamma \in \mathbf{\Gamma}} |< R^n f, g_\gamma >|. \tag{8}$$

The residue $R^n f$ is sub-decomposed into

$$R^n f = < R^n f, g_{\gamma_n} > g_{\gamma_n} + R^{n+1} f, \tag{9}$$

which defines the residue at the order n+1. Since $R^{n+1} f$ is orthogonal to $g_{\gamma_n}$

$$\|R^n f\|^2 = |< R^n f, g_{\gamma_n} >|^2 + \|R^{n+1} f\|^2. \tag{10}$$

Let us carry this decomposition up to the order $m$. We decompose $f$ into the concatenated sum

$$f = \sum_{n=0}^{m-1} \left( R^n f - R^{n+1} f \right) + R^m f. \tag{11}$$

Equation (9) yields

$$f = \sum_{n=0}^{m-1} < R^n f, g_{\gamma_n} > g_{\gamma_n} + R^m f. \tag{12}$$

Similarly, $\|f\|^2$ is decomposed in a concatenated sum

$$\|f\|^2 = \sum_{n=0}^{m-1} \left(\|R^n f\|^2 - \|R^{n+1} f\|^2\right) + \|R^m f\|^2. \tag{13}$$

Equation (10) yields an energy conservation equation

$$\|f\|^2 = \sum_{n=0}^{m-1} |< R^n f, g_{\gamma_n} >|^2 + \|R^m f\|^2. \tag{14}$$

The original vector $f$ is decomposed into a sum of dictionary elements, that are chosen to best match its residues. Although this decomposition is non-linear, we maintain an energy conservation as if it was a linear orthogonal decomposition. A major issue is to understand the behavior of the residue $R^m f$ when $m$ increases. By transposing a result proved by Jones[5] for projection pursuit algorithms,[4] one can prove that the matching pursuit algorithm does converge, even in infinite dimensional spaces.

THEOREM 3.1. *Let $f \in \mathbf{H}$. The residue $R^m f$ defined by the induction equation (9) satisfies*

$$\lim_{m \to +\infty} \|R^m f\| = 0. \tag{15}$$

*Hence*

$$f = \sum_{n=0}^{+\infty} < R^n f, g_{\gamma_n} > g_{\gamma_n}, \tag{16}$$

*and*

$$\|f\|^2 = \sum_{n=0}^{+\infty} |< R^n f, g_{\gamma_n} >|^2. \tag{17}$$

*When $\mathbf{H}$ has a finite dimension, $||R^m f||$ decays exponentially to zero.*

# 4 Implementations of a Matching Pursuit

When the dictionary is very redundant, the search for the vectors that match best the signal residues can mostly be limited to a sub-dictionary $\mathcal{D}_\alpha = (g_\gamma)_{\gamma \in \mathbf{\Gamma}_\alpha} \subset \mathcal{D}$. We suppose that $\mathbf{\Gamma}_\alpha$ is a finite index set included in $\mathbf{\Gamma}$ such that for any $f \in \mathbf{H}$

$$\sup_{\gamma \in \mathbf{\Gamma}_\alpha} |< f, g_\gamma >| \geq \alpha \sup_{\gamma \in \mathbf{\Gamma}} |< f, g_\gamma >|. \tag{18}$$

Depending upon $\alpha$ and the dictionary redundancy, the set $\mathbf{\Gamma}_\alpha$ can be much smaller than $\mathbf{\Gamma}$. The matching pursuit is initialized by computing the inner products $(< f, g_\gamma >)_{\gamma \in \mathbf{\Gamma}_\alpha}$, and continues by induction as follow. Suppose that we have already computed $(< R^n f, g_\gamma >)_{\gamma \in \mathbf{\Gamma}_\alpha}$, for $n \geq 0$. We search in $\mathcal{D}_\alpha$ for an element $g_{\tilde{\gamma}_n}$ such that

$$|< R^n f, g_{\tilde{\gamma}_n} >| = \sup_{\gamma \in \mathbf{\Gamma}_\alpha} |< R^n f, g_\gamma >|. \tag{19}$$

To find a dictionary element that matches $f$ even better than $g_{\tilde{\gamma}_n}$, we then search with a Newton method for an index $\gamma_n$ in a neighborhood of $\tilde{\gamma}_n$ in $\mathbf{\Gamma}$ where $| < f, g_\gamma > |$ reaches a local maxima. Clearly

$$|< R^n f, g_{\gamma_n} >| \geq |< R^n f, g_{\tilde{\gamma}_n} >| \geq \alpha \sup_{\gamma \in \mathbf{\Gamma}} |< R^n f, g_\gamma >|. \tag{20}$$

Let us observe that the choice function mentioned in section 3 is defined indirectly by this double search strategy. Once the vector $g_{\gamma_n}$ is selected, we compute the inner product of the new residue $R^{n+1} f$ with any $g_\gamma \in \mathcal{D}_\alpha$, with an updating formula derived from equation (9)

$$< R^{n+1} f, g_\gamma > = < R^n f, g_\gamma > - < R^n f, g_{\gamma_n} > < g_{\gamma_n}, g_\gamma > . \tag{21}$$

Since we previously stored $< R^n f, g_\gamma >$ and $< R^n f, g_{\gamma_n} >$, this update requires only to compute $< g_{\gamma_n}, g_\gamma >$. Dictionaries are generally built so that this inner product is recovered with a small number of operations. We describe in[6] how to compute efficiently the inner product of two discrete Gabor atoms, in $O(1)$ operations.

The number of times we sub-decompose the residues of a given signal $f$ depends upon the desired precision $\epsilon$. The number of iterations is the minimum $p$ such that

$$\|R^p f\| = \|f - \sum_{n=0}^{p-1} < R^n f, g_{\gamma_n} > g_{\gamma_n}\| \leq \epsilon \|f\|. \tag{22}$$

The energy conservation (14) proves that this equation is equivalent to

$$\|f\|^2 - \sum_{n=0}^{p-1} |< R^n f, g_{\gamma_n} >|^2 \leq \epsilon^2 \|f\|^2. \tag{23}$$

Since we do not compute the residue $R^n f$, at each iteration we test the validity of (23) to stop the decomposition.

# 5 Matching Pursuit With Time-Frequency Dictionaries

For dictionaries of time-frequency atoms, a matching pursuit yields an adaptive time-frequency transform. It decomposes any function $f(t) \in \mathbf{L}^2(\mathbf{R})$ into a sum of complex time-frequency atoms that best match its residues. This section studies the properties of this particular matching pursuit decomposition. We derive a new type of time-frequency energy distribution by summing the Wigner distribution of each time-frequency atom.

Since a time-frequency atom dictionary is complete, Theorem 1 proves that a matching pursuit decomposes any function $f \in \mathbf{L}^2(\mathbf{R})$ into

$$f = \sum_{n=0}^{+\infty} < R^n f, g_{\gamma_n} > g_{\gamma_n}, \tag{24}$$

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

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{s_n}} g(\frac{t - u_n}{s_n}) e^{i\xi_n t}. \tag{25}$$

These atoms are chosen to best match the residues of $f$.

From the decomposition of any $f(t)$ within a time-frequency dictionary, we derive a new time-frequency energy distribution, by adding the Wigner distribution of each selected atom. Let us recall that the cross Wigner distribution of two functions $f(t)$ and $h(t)$ is defined by

$$W[f,h](t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t + \frac{\tau}{2}) \overline{h}(t - \frac{\tau}{2}) e^{-i\omega\tau} d\tau. \tag{26}$$

The Wigner distribution of $f(t)$ is $Wf(t,\omega) = W[f,f](t,\omega)$. Since the Wigner distribution is quadratic, we derive from the atomic decomposition (24) of $f(t)$ that

$$\begin{aligned} Wf(t,\omega) &= \sum_{n=0}^{+\infty} |< R^n f, g_{\gamma_n} >|^2 W g_{\gamma_n}(t,\omega) \\ &+ \sum_{n=0}^{+\infty} \sum_{m=0, m\neq n}^{+\infty} < R^n f, g_{\gamma_n} > \overline{< R^m f, g_{\gamma_m} >} W[g_{\gamma_n}, g_{\gamma_m}](t,\omega). \end{aligned} \tag{27}$$

The double sum corresponds to the cross terms of the Wigner distribution. It regroups the terms that one usually tries to remove in order to obtain a clear picture of the energy distribution of $f(t)$ in the time-frequency plane. We thus only keep the first sum and define

$$Ef(t,\omega) = \sum_{n=0}^{+\infty} |< R^n f, g_{\gamma_n} >|^2 W g_{\gamma_n}(t,\omega). \tag{28}$$

A similar decomposition algorithm over time-frequency atoms was derived independently by Qian and Chen,[7] in order to define this energy distribution in the time-frequency plane. From the well known dilation and translation properties of the Wigner distribution and the expression (25) of a time-frequency atom, we derive that for $\gamma = (s, \xi, u)$

$$W g_\gamma(t,\omega) = W g(\frac{t-u}{s}, s(\omega - \xi)), \tag{29}$$

and hence

$$Ef(t,\omega) = \sum_{n=0}^{+\infty} |< R^n f, g_{\gamma_n} >|^2 W g(\frac{t-u_n}{s_n}, s_n(\omega - \xi_n)). \tag{30}$$

The Wigner distribution also satisfies

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W g(t,\omega) dt \; d\omega = \|g\|^2 = 1, \tag{31}$$

so the energy conservation equation (17) implies

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Ef(t,\omega) dt \; d\omega = \|f\|^2. \tag{32}$$

We can thus interpret $Ef(t,\omega)$ as an energy density of $f$ in the time-frequency plane $(t,\omega)$. Unlike the Wigner and the Cohen class distributions, it does not include cross terms. It also remains positive if $Wg(t,\omega)$ is positive, which the case when $g(t)$ is Gaussian. On the other hand, the energy density $Ef(t,\omega)$ does not satisfy marginal properties, as opposed to certain Cohen class distributions.[1] The importance of these marginal properties for signal processing is however not clear. If $g(t)$ is the Gaussian window

$$g(t) = 2^{1/4} e^{-\pi t^2}, \tag{33}$$

then

$$Wg(t,\omega) = 2e^{-2\pi(t^2 + (\frac{\omega}{2\pi})^2)}. \tag{34}$$

The time-frequency atoms $g_\gamma(t)$ are then called Gabor functions. The time-frequency energy distribution $Ef(t,\omega)$ is a sum of Gaussian blobs whose locations and variances along the time and frequency axes depend upon the parameters $(s_n, u_n, \xi_n)$.

Fig. 1(a) is a signal $f$ of 512 samples that is built by adding chirps, truncated sinusoidal waves and waveforms of different time-frequency localizations. No Gabor function have been used to construct this signal. Fig. 1(b) shows the time-frequency energy distribution $Ef(t,\omega)$. Since $Ef(t,\omega) = Ef(t,-\omega)$, we only display its values for $\omega \geq 0$. Each Gabor time-frequency atom selected by the matching pursuit is an elongated Gaussian blob in the time-frequency plane. We clearly see appearing two chirps that cross each others, with a localized time-frequency waveform on the top of their crossing point. We can also detect closely spaced Diracs, and truncated sinusoidal waves having close frequencies. Several isolated localized time-frequency components also appear in this energy distribution.

Fig. 2(a) is the graph of a speech recording corresponding to the word "greasy", sampled at 16 kHz. From the time-frequency energy displayed in Fig. 2(b), we can see the low-frequency component of the "g" and the quick burst transition to the "ea". The "ea" has many harmonics that are lined up but we can also see localized

high-frequency impulses that correspond to the pitch. The "s" component has a time-frequency energy spread over a high-frequency interval. Most of the signal energy is characterized by few time-frequency atoms. For $n = 250$ atoms, $\frac{\|R^n f\|}{\|f\|} = .169$, although the signal has 5782 samples, and the sound recovered from these atoms is of excellent quality.

Fig. 3(a) shows a signal obtained by adding a Gaussian white noise to the speech recording given in Fig. 2(a), with a signal to noise ratio of 1.5 db. Fig. 3(b) is the time-frequency energy distribution of this noisy signal. The white noise generates time-frequency atoms spread across the whole time-frequency plane, but we can still distinguish the time-frequency structures of the original signal because their energy is better concentrated in this plane.

# 6 Chaos in Matching Pursuit

In finite dimension, the energy of the residue converges exponentially to zero. We renormalize this residue in order to study its properties when the number of iterations increases

$$\tilde{R}^n f = \frac{R^n f}{||R^n f||}.$$

The matching pursuit defines the non-linear map

$$\tilde{R}^{n+1} f = M(\tilde{R}^n f). \tag{35}$$

At each iteration, the renormalized matching pursuit map removes the largest dictionary component of the residue and then renormalizes the new residue. This action is much like that of a left-shift map acting on a string of digits. Let $\Sigma_m$ be the set of all strings of digits $\{0, 1, ...m-1\}$. The left-shift map $L_m$ transforms any string in $\Sigma_m$ by

$$L_m(d_1 d_2 d_3 ...) = d_2 d_3 d_4 ...$$

The left shift operator removes the most significant digit of the string and the string is shift to the left, which is equivalent to a renormalization. The left shift map is known to be a chaotic map. A matching pursuit shares the same properties and we proved[3] that for a particular dictionary in $\mathbf{R^3}$, a renormalized matching pursuit map is topologically equivalent to a shift map.

Experimental data suggest that the residues of a normalized matching pursuit converge to a chaotic equilibrium state, even in high dimensional spaces. This is proved[3] for a simple dictionary composed of Diracs and complex exponentials but the same result was observed for more complicated dictionaries such as the one composed of Gabor functions. The residues converge to realizations of a specific process that we call dictionary noise. If the dictionary is invariant when we translate its elements or multiply them by a complex exponential, one can then prove[3] that this process is white and stationary. This is the case for a Gabor dictionary. Realizations of a dictionary noise are signals whose inner products with elements of the dictionary are uniformly small. In other words, such signals have no structure that is particularly coherent with respect to the dictionary. To measure whether a given residue $R^m f$ is close to the attractor, we compute

$$\lambda(\tilde{R}^m f) = \sup_{\gamma \in \mathbf{\Gamma}} | < \tilde{R}^m f, g_\gamma > |$$

and compare it to $\lambda_e$, which is the expected value of $\lambda$ for realizations of the dictionary noise. If $\lambda(\tilde{R}^m f) \leq \lambda_e$, then $R^m f$ does not include any more coherent component with respect to the dictionary. The coherent components of $f$ are thus reduced to the first $m$ selected dictionary vectors $(g_{\gamma_n})_{0 \leq n < m}$.

The expected value $\lambda_e$ has been measured numerically for a Gabor dictionary.[6] The noisy greasy signal in Fig. 3(a) has $m = 76$ coherent structures, whose time-frequency distributions are shown in Fig. 4(a). Fig 4(b) is the

signal reconstructed from these time-frequency atoms. The SNR of the reconstructed signal is $6.8db$. The white noise has been removed and this signal has a good auditory quality because the main time-frequency structures of the original speech signal have been retained.

# 7 REFERENCES


[1] L. Cohen "Time-frequency distributions: a review" Proceedings of the IEEE, vol. 77, No. 7, p. 941-979, July 1989.

[2] I. Daubechies, "Ten lectures on wavelets", CBMS-NSF Series in Appl. Math., SIAM, 1991.

[3] G. Davis and S. Mallat, "Choatic Adaptive Time-Frequency Decompositions", Technical Report, Computer Science, NYU, April 1994.

[4] P. J. Huber, "Projection Pursuit", *The Annals of Statistics*, vol. 13, No. 2, p. 435-475, 1985.

[5] L. K. Jones, "On a conjecture of Huber concerning the convergence of projection pursuit regression", *The Annals of Statistics*, vol. 15, No. 2, p. 880-882, 1987.

[6] S. Mallat and Z. Zhang "Matching Pursuit with Time-Frequency Dictionaries", *IEEE Trans. on Signal Processing*, Dec. 1993.

[7] S. Qian and D. Chen, "Signal Representation via Adaptive Normalized Gaussian Functions," *IEEE Trans. on Signal Processing*, vol. 36, no. 1, Jan. 1994.

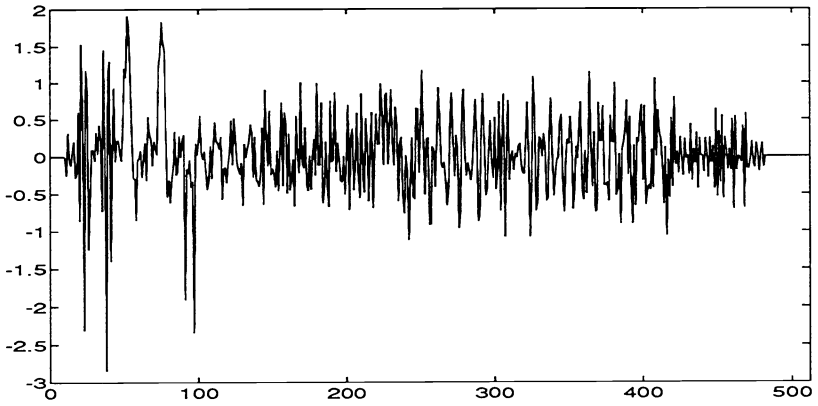


Figure 1(a): Signal of 512 samples built by adding chirps, truncated sinusoidal waves and waveforms of different time-frequency localizations.

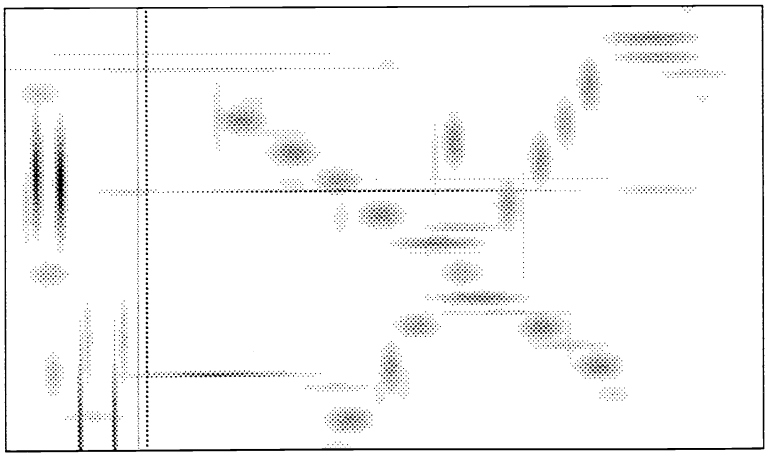

Figure 1(b): Time-frequency energy distribution $Ef(t,\omega)$ of the signal shown in (a). The horizontal axis is time. The vertical axis is frequency. The highest frequencies are on the top. The darkness of this time-frequency image increases with the value $Ef(t,\omega)$.

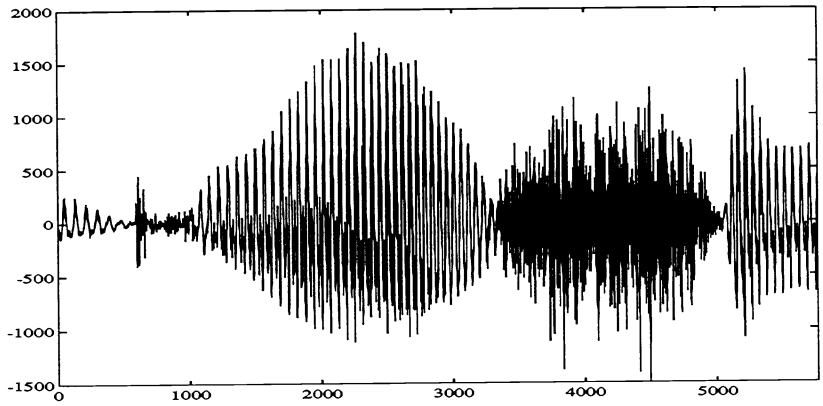


Figure 2(a): speech recording of the word "greasy", sampled at 16 kHz.

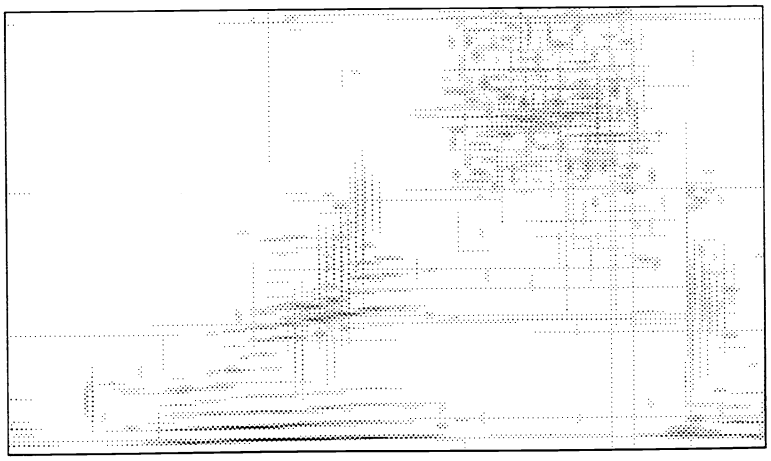

Figure 2(b): time-frequency energy distribution of the speech recording shown in (a). We see the low-frequency component of the "g", the quick burst transition to the "ea" and the harmonics of the "ea". The "s" has energy spread over high frequencies.

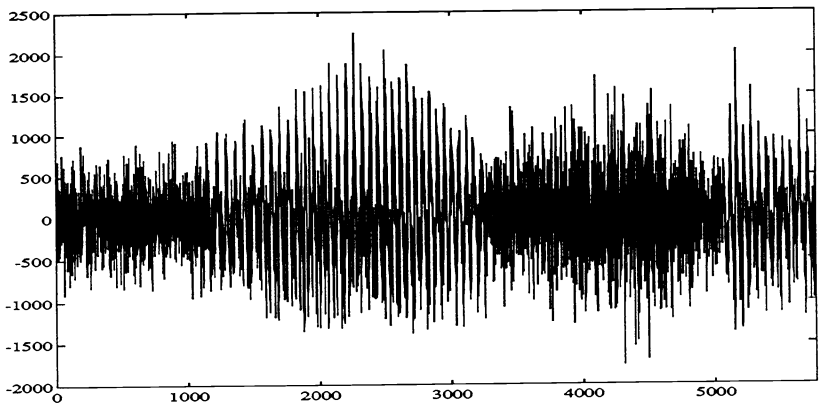


Figure 3(a): signal obtained by adding a Gaussian white noise to the speech recording shown in Fig.2(a). The signal to noise ratio is 1.5db.

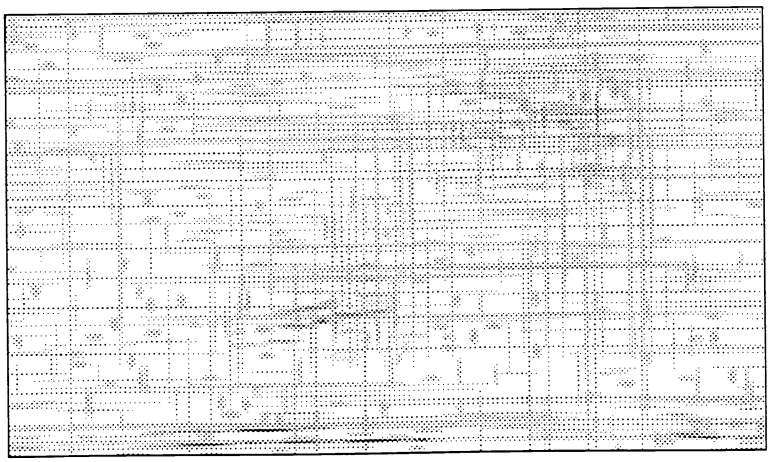

Figure 3(b): time-frequency energy distribution of the noisy speech signal. The energy distribution of the white noise is spread across the whole time-frequency plane.

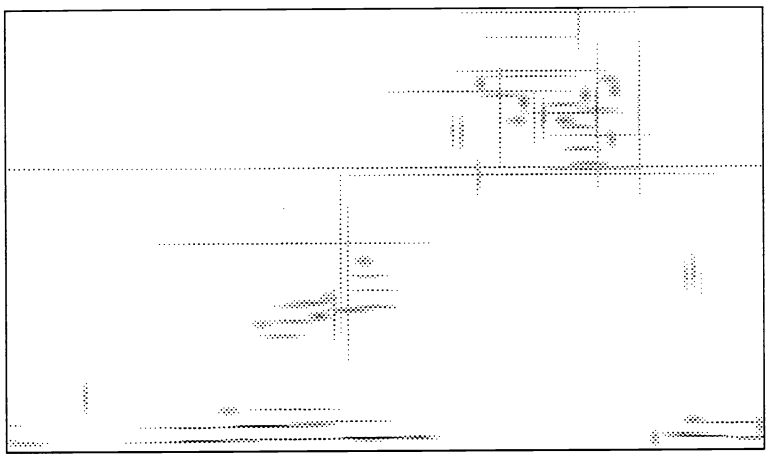

Figure 4(a): time-frequency energy distribution of the $m = 76$ coherent structures of the noisy speech signal shown in Fig.3(a).

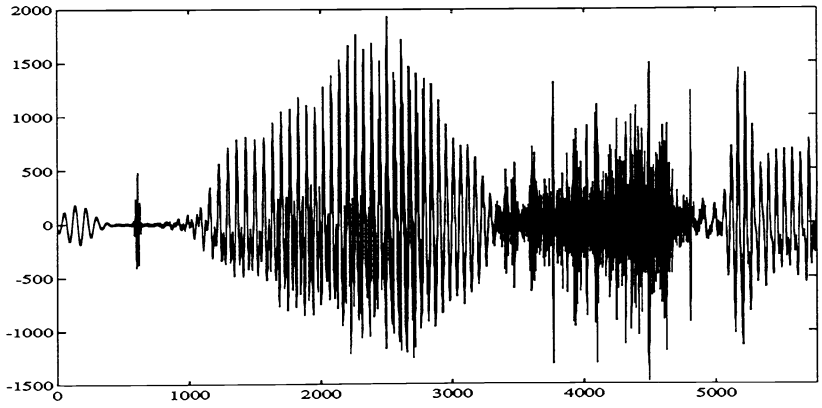


Figure 4(b): time-frequency energy distribution of the $m = 76$ coherent structures of the noisy speech signal shown in Fig. 3. (b): signal reconstructed from the 76 coherent structures shown in (a). The white noise has been removed.