Proceedings of ALGORITMY 2002
Conference on Scientific Computing, pp. 33–41

NEW VERSION OF MATCHING PURSUIT DECOMPOSITION WITH CORRECT REPRESENTATION OF LINEAR CHIRPS *

MONIKA KOVÁČOVÁ † AND MIRIAM KRISTEKOVÁ‡

Abstract. In this paper we describe a new version of matching pursuit algorithm, that decomposes any signal into a linear expansion of the functions selected from the redundant set called dictionary. A matching pursuit selects the signal structures that are coherent to the used dictionary.

Matching pursuit decomposition algorithm (MPD) with a dictionary based on the Gabor functions is an excellent tool for the time-frequency analysis of non-stationary signals. However, original Gabor dictionary does not represent signals with time-depending frequency modulations correctly. We present here a new generalization of the Gabor dictionary and corresponding extension of MPD algorithm that allows one to correctly and uniformly represent also the signals with linear chirps included in their structure.

Key words. Adaptive signal processing, approximation methods, time-frequency analysis, redundant systems, matching pursuit.

AMS subject classifications. 42C99, 42C15, 65T99

1. Introduction. The focus of this manuscript is the problem of obtaining efficient representations of functions. We seek to approximate functions with linear combination of a small number of unit functions from a family $\{g_\gamma\}_{\gamma\in\Gamma}$ in a Hilbert space H. For any $M > 0$, we want to minimize the error

$$\varepsilon(M) = \|f - \sum_{\gamma\in\Lambda} \alpha_\gamma \; g_\gamma\| \, ,$$

where $\Lambda \subset \Gamma$ is an index set of cardinality M. Representations of this form are central importance in numerous applications. If we can accurately approximate f with a linear combination of a small number of the functions g_γ, then we need store a small number of coefficients α_γ and indexes γ. For numerical methods, such representations can reduce the computational complexity on f to a small number of computations performed on each g_γ in the expansion of f, which means fast calculation.

When $\{g_\gamma\}_{\gamma\in\Gamma}$ is an orthogonal basis, we can minimize the approximation error $\varepsilon(M)$ by taking Λ to be functions corresponding to the largest M inner products $(|\langle f, g_\gamma\rangle|)_{\gamma\in\Gamma}$, since

$$\varepsilon(M) = \sum_{\gamma\in\Gamma-\Lambda} |\langle f, g_\gamma\rangle|^2 \, .$$

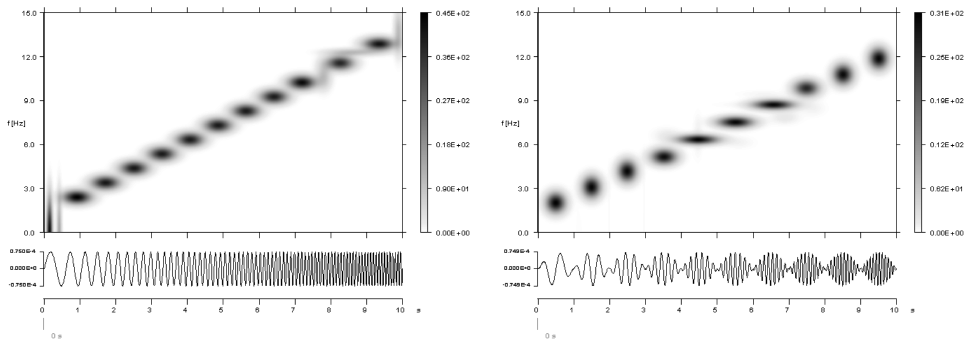
For the case that the Hilbert space H has a finite dimension N and the set Γ contains only the finite number of the orthogonal functions, the expansion which minimizes $\varepsilon(M)$ is not difficult to compute.

*This work was supported by Grants No.: 1-1090/2001, VEGA, Slovak Republic, European Commission Project No. EVG1-CT-2000-00026 SESAME, European Commission Project No. EVG1-CT-2001-00040 EUROSEIS-RISK.

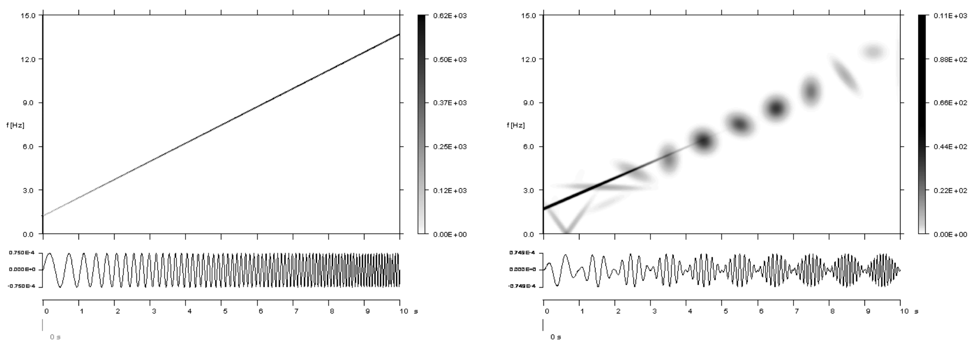
†Dept. of Mathematics, SjF STU, Nám. Slobody 17, 812 31 Bratislava, Slovak Republic (kovacova_v@sjf.stuba.sk), Dept. of Seismology, Geophysical Institute of Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovak Republic (geofmoko@savba.sk).

‡Dept. of Seismology, Geophysical Institute of Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovak Republic (geofmikr@savba.sk).

Decomposition of signals over family of functions that are well localized both in time and frequency have found many applications in signal processing and harmonic analysis. Such functions are called the *time-frequency atoms*. In this paper we will describe a new version of the algorithm, called *matching pursuit decomposition algorithm* (MPD), that decomposes any signal into a linear expansion of the functions, that are selected from the redundant set, called a *dictionary.* These functions are searched in order to best match the signal structure. Matching pursuits are general procedures to compute an adaptive signal representation. With a dictionary based on the Gabor functions the MPD algorithm is an excellent tool for the time-frequency analysis of non-stationary signals. However, the original Gabor dictionary does not represent signals with time-depending frequency modulations correctly. We observe for example, for two different signals very close time-frequency representations (see the next Figure). Consequently, we can not identify the presence of waveforms with



the frequency modulation. Because signals with nonlinear frequency modulation are of great importance in seismology, we have developed a new generalization of the Gabor dictionary and corresponding extension of the MPD algorithm that allow one to correctly represent also the signals with *linear chirps* and distinguish them from superpositions of waveforms with constant frequency modulation (see the next Figure).



2. Matching Pursuit in Hilbert Space. The general issue is to find algorithm to expand functions over a set of functions (or over the dictionary).The idea of the pursuit to represent a function has been known to statisticians for a long time. It can be found, for example, in a 1981 publication [4] by Friedman and Stuetzle, where the authors consider the problem of non-parametric regression. The general algorithm, called MPD, which performs such an adaptive decomposition for the time-frequency dictionary was described in [5], at first. We introduce here the short idea how it works.

Let H be a Hilbert space. The dictionary $D = (g_\gamma)_{\gamma\in\Gamma}$ was defined as a set of the functions in H, such that $\|g_\gamma\| = 1$. Let V be the closed linear span of the dictionary functions. We require that the finite linear expansions of functions in D are dense in space V. We say that the dictionary is complete if and only if $V = H$.

Let $f \in H$. We want to compute a linear expansion of f over a set of functions selected from D, in order to best match its structures. This done by successive approximations of f with orthogonal projections on elements of D. Let $g_{\gamma_0} \in D$. The function f can be decompose into

$$f = \langle f, g_{\gamma_0}\rangle g_{\gamma_0} + Rf\,, \tag{2.1}$$

where Rf is the residual function after approximating f into direction of g_{γ_0}. Clearly g_{γ_0} is orthogonal to Rf and we get energy conservation law for this step

$$\|f\|^2 = |\langle f, g_{\gamma_0}\rangle|^2 + \|Rf\|^2\,. \tag{2.2}$$

To minimize $\|Rf\|$, we must choose $g_{\gamma_0} \in D$ such that $|\langle f, g_{\gamma_0}\rangle|$ is maximum. In some cases, it is possible to find a function g_{γ_0} that is almost the best in the sense that

$$|\langle f, g_{\gamma_0}\rangle| \geq \alpha \sup_{\gamma\in\Gamma} |\langle f, g_\gamma\rangle|\,, \tag{2.3}$$

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

A matching pursuit is an iterative algorithm that sub-decomposes the residue Rf by projecting it into the family functions D that match Rf almost the best, as was done for f. This procedure is repeated each time on the following residue that is obtained. We mention that the choice of a function g_{γ_0} that satisfies (2.1) is not random. We will present in the next section, that the axiom of choice guaranties that there exists at least one choice function, but in practice there are many ways to define it, and it depends upon the numerical implementation.

Let us explain by induction, how the MPD-algorithm works. Let $R^0 f = f$. We suppose that we have computed the nth order residue $R^n f$, for $n \geq 0$. We choose, using the choice function C, an element $g_{\gamma_n} \in D$, which best match the residue $R^n f$ in the sense of the choice function

$$|\langle R^n f, g_{\gamma_n}\rangle| \geq \alpha \sup_{\gamma\in\Gamma} |\langle R^n f, g_\gamma\rangle|\,. \tag{2.4}$$

The residue $R^n f$ can be also decomposed into

$$R^n f = \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n} + R^{n+1} f\,, \tag{2.5}$$

which defines the residue at the order $n+1$. Since $R^{n+1} f$ is orthogonal to g_{γ_n} energy conservation law has the form

$$\|R^n f\|^2 = |\langle R^n f, g_{\gamma_n}\rangle|^2 + \|R^{n+1} f\|^2\,.$$

If we stop the decomposition process at the step m, the signal f is decomposed into the sum

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n} + R^m f\,. \tag{2.6}$$

Similarly, we maintain in the energy laws

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

The original function f is decomposed into a sum of dictionary elements, that are chosen to best match its residues. Although this decomposition is nonlinear, we maintain an energy conservation as if it was a linear orthogonal decomposition.

A major issue is to understand the behaviour of the residue $R^m f$ when m increases. The mathematical similarities between the Jones algorithm [2] and MPD algorithm allowed Mallat in [5] to prove the next theorem.

THEOREM 2.1. *Let V be a closed linear span of functions in D. Let us denote by W the orthogonal complement of V in H. The orthogonal projectors over V and W are written as P_V and P_W.*

Let $f \in H$. The residue $R^m f$ defined by the induction (2.6) satisfies

$$\lim_{m\to\infty} \|R^m f - P_W f\| = 0 .$$

Hence $P_V f = \sum_{n=0}^{\infty} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n}$. When the dictionary is complete, which means that $V = H$, then $P_V f = f$ and $P_W f = 0$.

After m iterations we decompose signal into the form (2.6). If we stop the algorithm at this stage and record the coefficients $(\langle R^n f, g_{\gamma_n}\rangle, g_{\gamma_n})_{0\le n<m}$, the decomposition error will be equal to $R^m f$. However, this sum is not linear expansion of the functions $(g_{\gamma_n})_{0\le n<m}$ that approximates f at best. We simply derive from (2.6) that

$$P_{V_m} f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n} + P_{V_m} R^m f .$$

If the family of functions $(g_{\gamma_n})_{0\le n<m}$ is not orthogonal, which is generally the case, then $P_{V_m} R^m f \neq 0$. So, we must compute $P_{V_m} R^m f$ once more. The computation of

$$P_{V_m} R^m f = \sum_{n=0}^{m-1} \alpha_n \, g_{\gamma_n} \tag{2.7}$$

is called a *back projection*. Instead of storing the inner products $\langle R^n f, g_{\gamma_n}\rangle$, we must to store $\langle R^n f, g_{\gamma_n}\rangle + \alpha_n$ in order to recover $P_{V_m} f$. Calculation of α_n requires to solve linear system of equation, which was described in detail in [1].

3. Matching Pursuit Algorithm with Time-Frequency atoms. The family $D = (g_\gamma)_{\gamma\in\Gamma}$ is extremely redundant, and its properties have been studied by [6] for the case of the Gabor dictionary [1, 5]. Elements were defined by set of parameters $\gamma = (s, u, \nu)$ by scaling (s), translating (u) and constant frequency modulating (ν) a single Gaussian window $g(t) = 2^{1/4} e^{-\pi t^2}$ and had the form $g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\nu t}$. $g_\gamma(.)$ are called Gabor functions. Our dictionary D, defined by functions (3.1), has the similar properties.

In this paper we put $H = L^2(R)$ such that $\|f\| = \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty$. Each dictionary element function g_γ is defined by

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i(\nu_0+\nu_1 t)t} , \tag{3.1}$$

where the index $\gamma = (s, u, \nu_0, \nu_1)$ is an element of the set $\Gamma = R^+ \times R^3$. The factor $\frac{1}{\sqrt{s}}$ normalizes to 1 the norm of $g_\gamma(t)$. The dictionary D is generated by scaling (s), translating (u) and modulating (ν_0, ν_1) a single Gaussian window function $g(t) \in L^2(R)\,, g(t) = 2^{1/4} e^{-\pi t^2}$. For $\nu_1 = 0\,,\ g_\gamma(.)$ in (3.1) defines original Gabor dictionary.

The aim of this paper is to include elements with linear frequency modulation within the dictionary elements. The larger collection of the dictionary elements may provide more freedom in the representation of the signals as the unique expansions to the linear combination of the basis elements.

The matching pursuit decomposition algorithm depends upon a choice function that selects at each iteration a function g_{γ_n} among all functions that satisfy (2.3). Unfortunately, the decomposition is in general not unique, the coefficients in the decomposition represent redundant information, and one has to compare several decompositions of the same signal and look for optimal representations.

We say that a subset $\Lambda \subset \Gamma$ is admissible and associated to $f \in L^2(R)$ if

$$\Lambda = \{\beta \in \Gamma, |\langle f, g_\beta\rangle| \geq \alpha \sup_{\gamma\in\Gamma} |\langle f, g_\gamma\rangle|\}\,. \tag{3.2}$$

Let Λ be an admissible set and $(a, c, b_0, b_1) \in R^+ \times R^3$. Let

$$\begin{aligned} \Lambda_{(a,c,b_0,b_1)} = \{\beta = (s, u, \nu_0, \nu_1) \in \Gamma :& \\ \left(\frac{s}{a}, \frac{u-c}{a}, a(\nu_0 - b_0) + 2ac(\nu_1 - b_1), a^2(\nu_1 - b_1)\right) \in \Lambda&\right\}\,. \end{aligned} \tag{3.3}$$

A choice function C is said to be covariant if and only if for any admissible set Λ, $C(\Lambda) = (s_0, u_0, \nu_{00}, \nu_{10})$, where C is a choice function, implies that

$$C(\Lambda_{(a,c,b_0,b_1)}) = \left(\frac{s_0}{a}, \frac{u_0 - c}{a}, a(\nu_{00} - b_0) + 2ac(\nu_{10} - b_1), a^2(\nu_{10} - b_1)\right)\,. \tag{3.4}$$

If we restrict our signal to the bounded and absolutely integrable functions , the matching pursuit residues will be also bounded and absolutely integrable. Covariant choice function $C(.)$ can be then defined as follows: For any admissible set Λ , $C(\Lambda) = (s_1, u_1, \nu_{01}, \nu_{11})$ such that

$$\begin{aligned} s_1 &= \sup\{s, \exists\ (u, \nu_0, \nu_1) \in R^3, (s, u, \nu_0, \nu_1) \in \Lambda\} \\ u_1 &= \sup\{u, \exists\ (\nu_0, \nu_1) \in R^2, (s_1, u, \nu_0, \nu_1) \in \Lambda\} \\ \nu_{01} &= \sup\{\nu_0, \exists\ \nu_1 \in R, (s_1, u_1, \nu_0, \nu_1) \in \Lambda\} \\ \nu_{11} &= \sup\{\nu_1, (s_1, u_1, \nu_{01}, \nu_1) \in \Lambda\} \end{aligned}$$

The following Theorem proves that the index $(s_1, u_1, \nu_{01}, \nu_{11})$ is well defined and belongs to Λ.

THEOREM 3.1. *For any admissible set Λ , associated to a bounded and absolutely integrable function, $(s_1, u_1, \nu_{01}, \nu_{11}) \in \Lambda$.*

Proof. We only give the main ideas of the proof.

The basic idea of the proof is to take Λ as an admissible index set associated to f. Since $g(t) = 2^{1/4} e^{-\pi t^2}$ is bounded and $f(t)$ is absolutely integrable, one can prove that

$$\lim_{s\to\infty} \sup_{u,\nu_0,\nu_1\in R^3} |\langle f, g_\gamma\rangle| = 0\,.$$

We can thus derive that there exists a finite s_1 that is the supremum of all s such that $(s,u,\nu_0,\nu_1)\in\Lambda$. Since $\lim\limits_{|t|\to\infty}|g(t)|=0$ and $f(t)$ is absolutely integrable, we can prove that for $\gamma=(s_1,u,\nu_0,\nu_1)$,

$$\lim_{u\to\infty}\sup_{\nu_0,\nu_1\in R^2}|\langle f,g_\gamma\rangle|=0\,.$$

We can then derive that there exists u_1 that is the supremum of all u such that $(s_1,u,\nu_0,\nu_1)\in\Lambda$. Since Λ is closed, there exists ν_0 such that $(s_1,u_1,\nu_0,\nu_1)\in\Lambda$. Similarly, we conclude from $\lim\limits_{|\omega|\to\infty}|\hat g(\omega)|=0$ and $\hat f(\omega)$ is absolutely integrable, that for $\gamma=(s_1,u_1,\nu_0,\nu_1)$

$$\lim_{\nu_0\to\infty}\sup_{\nu_1\in R}|\langle f,g_\gamma\rangle|=0\,,$$

hence that for $\gamma=(s_1,u_1,\nu_{01},\nu_1)$

$$\lim_{\nu_1\to\infty}\langle f,g_\gamma\rangle|=0\,,$$

We can finally derive that ν_{01} and ν_{11} are the supremas of all ν_0 resp. ν_1 fall into the closed Λ. $\square$

Let us prove the covariance of a matching pursuit based on the above defined covariant choice function. Let us define

$$f^1(t)=\frac{d}{\sqrt{a}}f^0\left(\frac{(t-c)}{a}\right)e^{i(b_0+b_1t)t}\,. \tag{3.5}$$

Let $\gamma^1=(s,u,\nu_0,\nu_1)$ and $\gamma^0=\left(\frac{s}{a},\frac{u-c}{a},a(\nu_0-b_0)+2ac(\nu_1-b_1),a^2(\nu_1-b_1)\right)$. We can show that

$$\langle f^1,g_{\gamma^1}\rangle=\langle f^0,de^{ic((\nu_0-b_0)+(\nu_1-b_1)t)}g_{\gamma^0}\rangle=de^{ic((b_0-\nu_0)+(b_1-\nu_1)t)}\langle f^0,g_{\gamma^0}\rangle\,. \tag{3.6}$$

Hence $\sup_{\gamma\in\Gamma}|\langle f^1,g_\gamma\rangle|=d\sup_{\gamma\in\Gamma}|\langle f^0,g_\gamma\rangle|$. According to the definition Λ (3.2) the equation (3.6) proves that the set $\Lambda_{(a,c,b_0,b_1)}$ also satisfies

$$\Lambda_{(a,c,b_0,b_1)}=\{\beta\in\Gamma,|\langle f^0,g_\beta\rangle|\ge\alpha\sup_{\gamma\in\Gamma}|\langle f^0,g_\gamma\rangle|\}\,.$$

The covariance of the choice function implies that if $C(\Lambda)=\gamma_0^1=(s_0,u_0,\nu_{00},\nu_{10})$ then $C(\Lambda_{(a,c,b_0,b_1)})=\gamma_0^0=\left(\frac{s_0}{a},\frac{u_0-c}{a},a(\nu_{00}-b_0)+2ac(\nu_{10}-b_1),a^2(\nu_{10}-b_1)\right)$. We can thus derive that

$$Rf^1(t)=\frac{d}{\sqrt{a}}Rf^0\left(\frac{(t-c)}{a}\right)e^{i(b_0+b_1t)t}\,.$$

Similarly we can prove by induction that

$$R^nf^1(t)=\frac{d}{\sqrt{a}}R^nf^0\left(\frac{(t-c)}{a}\right)e^{i(b_0+b_1t)t}\,,$$

and if $\gamma_n^1=(s_n,u_n,\nu_{0n},\nu_{1n})$ then

$$\gamma_n^0=\left(\frac{s_n}{a},\frac{u_n-c}{a},a(\nu_{0n}-b_0)+2ac(\nu_{1n}-b_1),a^2(\nu_{1n}-b_1)\right)$$

and

$$\langle R^nf^1,g_{\gamma_n^1}\rangle=de^{ic((b_0-\nu_{0n})+(b_1-\nu_{1n})t)}\langle R^nf^0,g_{\gamma_n^0}\rangle\,.$$

Hence a MPD based on covariance choice function (3.4) is covariant by dilation (s), translation (u) and modulation (ν_0,ν_1).

4. Numerical Experiments. For a single-component frequency modulated signal, our intuition suggests that the time-frequency representation should be concentrated near the curve of the instantaneous frequency $\nu_f(t)$. We can decompose our non stationary signal to a set of unimodular signals using the MPD algorithm. For simplicity can be supposed, that instantaneous amplitude is constant and equal to 1. The" ideal" situation would be that the time-frequency transformation has the form $C_{z_f}(t,\nu,\tilde{f}) = \delta(\nu - \nu_f(t))$, where $z_f(t)$ be the associates analytic signal to $f(t)$ and

$$C_f(t,\nu,\tilde{f}) = \iiint_{R^3} e^{i2\pi\xi(s-t)}\tilde{f}(\xi,\tau)f(s+\frac{\tau}{2})f^*(s-\frac{\tau}{2})e^{-i2\pi\nu\tau}\ d\xi\ ds\ d\tau$$

means the standard definition of Cohen's class transformations [3].

Our dictionary's element is a linear chirp characterized by $\nu_x(t) = \nu_0 + \nu_1 t$. Some direct computations show that

$$C_{z_f}(t,\nu,\tilde{f}) = \int_{-\infty}^{+\infty} \tilde{f}(\nu_1\tau,\tau)e^{-i2\pi(\nu-\nu_x(t))\tau}\ d\tau\ .$$

Hence for a nonzero slope ν_1 of the modulation, the only solution is given by $\tilde{f}(\xi,\tau) = 1$, and this characterizes the Wigner distribution. On the other hand, for a pure frequency $\nu_1 = 0$ every definition according to a parameter function with $\tilde{f}(0,\tau) = 1$ is suitable. This evidently agrees with the fact that this same condition guarantees a correct marginal distribution in frequency.

Since a time-frequency atom dictonary is complete, Theorem 2.1 and Theorem 3.1 prove that a matching pursuit decomposes any function $f \in L^2(R)$ into

$$f = \sum_{n=0}^{\infty} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n}\ , \tag{4.1}$$

where

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g(\frac{(t-u)}{s})e^{i(\nu_0+\nu_1 t)t}\ ,$$

and the index $\gamma = (s,u,\nu_0,\nu_1)$. These atoms are chosen to the best match the residues of f. In the previous section was proved that there exists a class of choice functions such that $f^1(t) = \frac{d}{\sqrt{a}} f^0(\frac{(t-c)}{a})e^{i(b_0+b_1t)t}$ if and only if for all $n \geq 0$

$$s_n^0 = \frac{s_n^1}{a},\ u_n^0 = \frac{u_n^1 - c}{a},\ \nu_{0n}^0 = a(\nu_{0n}^1 - b_0) + 2ac(\nu_{1n}^1 - b_1),\ \nu_{1n}^1 = a^2(\nu_{1n}^1 - b_1)\ ,$$

and

$$\langle R^n f^1, g_{\gamma_n^1}\rangle = de^{ic((b_0-\nu_{0n})+(b_1-\nu_{1n})t)}\langle R^n f^0, g_{\gamma_n^0}\rangle\ .$$

From the decomposition of any $f(t)$ we derive a new time-frequency energy distribution, by adding the Wigner distribution of each selected atom. According to the previous reasons this distribution match efficiently any signal structure and localizes it well in the time-frequency plane, regardless of whether this localization is in time or in frequency. So we can associate a time-frequency energy distribution with

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

where for our choice of the gaussian window $g(t) = 2^{1/4}e^{-\pi t^2}$ we get

$$Wg_{\gamma_n}(t,\omega) = 2exp\left[-2\pi\left(\frac{(t-u)^2}{s^2} + s^2\omega^2\right)\right].$$

Its energy is concentrated in the time and frequency domains where g_{γ_n} is localized.

As was mentioned before, the decay of $\|R^n f\|$ depends upon the correlation between the residues and the dictionary elements, i.e. how the dictionary is to appropriate for a given signal. In the Figure 4.1.(a) we can see the decay of $\|R^n f\|$ for original and our improved MPD for highly nonstationary signal with linear chirps (the same signal as in Figure 4.2.). The decay is faster for our improved "linear" MPD, hence the our new algorithm converges more quickly for such a type of complicated signal. Although average time for one iteration is greater in "linear" case ($\bar{t} = 7.37$) then for original MPD ($\bar{t} = 6.68$), due to its faster convergency the " linear" version is much more efficient and total computational time is shorter (Figure 4.1.(b)) for given level of energy represented.

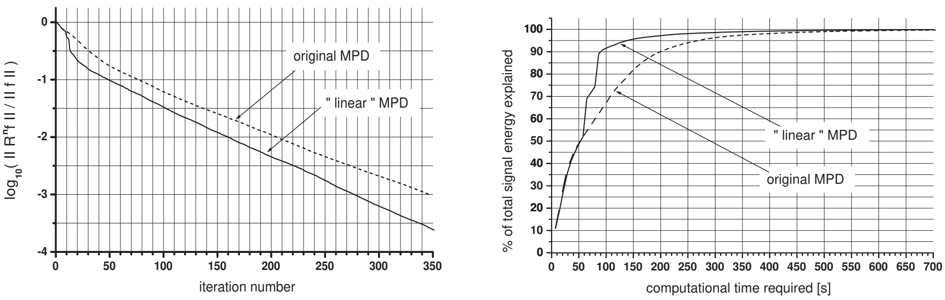


FIG. 4.1. *(a) Comparison of residual decay for "original" and our improved "linear" MPD, (b) Percentages of total signal energy represented* $\left(\frac{\sum_n (|\langle R^n f, g_\gamma\rangle|\cdot g_\gamma)^2}{\|f\|^2}\right)$ *versus computational time required.*

In Figure 4.2. the time-frequency planes are compared for the Mallat-Zhang highly non-stationary test signal, presented in [5]. Our time-frequency representation is more compact and the energy is better localized than with original Mallat's dictionary. The linear chirps are correctly represented.

REFERENCES

[1] G. DAVIS, S. MALLAT, AND M. AVELLANEDA, *Adaptive greedy approximations*, *Constr. Approx.*, vol. 13, no. 1, 1997, pp.57-98.
[2] L. K. JONES, *On a conjecture of Huber concerning the convergence of projection pursuit regression*, *Ann. Statistist.*, vol 15, no.2, 1987, pp. 880-882.
[3] P. FLANDRIN,*Time-Frequency/Time-Scale Analysis*, Acadmic Press, London, 1999.
[4] J. H. FRIEDMAN, AND W. STUETZLE, *Projection pursuit regression*, *J. Amer. Statist. Assoc.* 76, 1981, pp. 817-823.
[5] S. G. MALLAT, AND Z. ZHANG, *Matching Pursuit with Time-Frequency Dictionaries*,*IEEE Transactions on Signal Processing*, vol. 41, No. 12, 1993, pp. 3397-3415.
[6] B. TORRESANI, *Wavelets associated with representations of the affine Weyl-Heisenberg group*, *J. Math Physics*, vol. 32, 1991, pp.1273-1279.

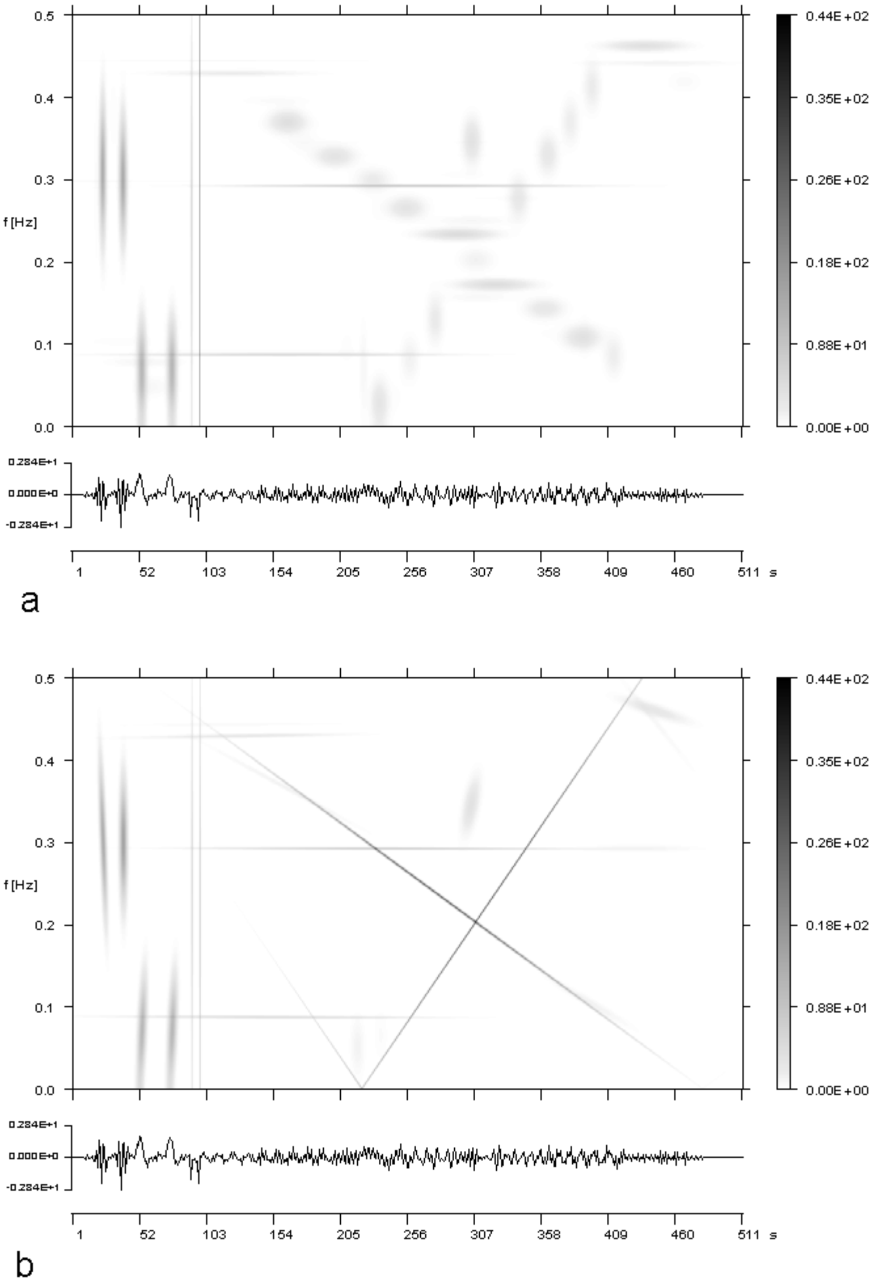


Fig. 4.2. *Time-frequency analysis of complicated nonstationary Mallat test signal (superposition of truncated sinusoids, Dirac impulses, waveforms centered at various time and frequency positions, waveforms with linear dispersion - chirps.) (a) Calculated with the original matching pursuit decomposition (result of 60 iterations) (b) Calculated with our improved matching pursuit decomposition (result of only 20 iterations).*