SHIFT-INVARIANT ADAPTIVE LOCAL TRIGONOMETRIC DECOMPOSITION

Israel Cohen, Shalom Raz and David Malah

Department of Electrical Engineering
Technion — Israel Institute of Technology
Technion City, Haifa 32000, Israel
E-mail: Israel Cohen <cisrael@techunix.technion.ac.il>

ABSTRACT

A general formulation of shift-invariant "best-basis" expansions is presented. Specifically, we construct an extended library of smooth local trigonometric bases, and introduce a suitable "best-basis" search algorithm. We prove that the resultant decomposition is shift-invariant, orthonormal and characterized by a reduced information cost. The shift-invariance is derived from an adaptive relative shift of expansions in distinct resolution levels. We show that at any resolution level ℓ it suffices to examine and select one of two relative shift options — a zero shift or a $2^{-\ell-1}$ shift. A variable folding operator, whose polarity is locally adapted to the parity properties of the signal, extra enhances the representation.

1. INTRODUCTION

Bases whose elements are well localized in time and frequency are useful for signal analysis and compression. Coifman and Meyer [1] have introduced a library of orthonormal local trigonometric bases having a binary tree structure, where the "best basis" is efficiently searched for a prescribed signal, relative to a specified information cost function [2]. The "best basis" coefficients provide a compact signature of the original signal, implying signal compression and identification applications [3, 4]. A major drawback is the lack of shift-invariance. Both, the local trigonometric decomposition (LTD) of Coifman and Wickerhauser [2] as well as the time-varying wavelet packet decomposition proposed by Herley et al. [5], are sensitive to the initial phase of the signal. Shift-invariant multiresolution representations that exist are either non-orthogonal, non-unique [6] or entail high oversampling rates [7, 8].

Recently we have developed an orthonormal shift invariant wavelet packet decomposition [9]. In this work, similar principles are applied to smooth local trigonometric bases. We introduce a best-basis search algorithm, namely *shift-invariant adapted-polarity local trigonometric decomposition* (SIAP-LTD), that leads to an orthonormal shift-invariant representation.

The shift-invariance so acquired stems from a relative shift between expansions in distinct resolution levels. It is proved that at any resolution level ℓ it suffices to examine and select one of two relative shift options — a zero shift or a $2^{-\ell-1}$ shift. The resultant best-basis decomposition is not only shift-invariant, but also characterized by a lower information cost when compared to the LTD. Its quality is further enhanced by applying an adaptive-polarity folding operator which splits the prescribed signal and "folds" *adaptively* overlapping parts back into the segments. The polarity of the folding operation is locally adapted to the signal at the finest resolution level, and a recursive sequence is carried out towards the coarsest resolution level merging segments where beneficial. Each segment of the signal is then represented by a trigonometric basis which possesses the same parity properties at the end-points.

In the next section, the shift-invariance of SIAP-LTD is demonstrated, and its quality (information cost) is compared with and verified to be superior to that of LTD.

2. SHIFT-INVARIANT DECOMPOSITIONS

For simplicity, we shall restrict ourselves to $L_2[0,1]$, the set of square integrable functions on the circle $[0,1]$.

Definition 1 *$f, g \in L_2[0,1]$ are said to be identical to within a resolution J time-shift ($J > 0$) if there exists $q \in \mathbb{Z}$, $0 \le q < 2^J$, such that $g(t) = f(t - 2^{-J}q)$ for all $t \in [0,1]$.*

Let $\mathcal{B}$ denote a library of orthonormal bases in $L_2[0,1]$, $\mathcal{M}$ an additive information cost functional and $\mathcal{M}(Bg)$ the information cost of representing $g \in L_2[0,1]$ on a basis $B \in \mathcal{B}$. The best basis for g, relative to a library of bases $\mathcal{B}$ and an information cost functional $\mathcal{M}$, is that $B \in \mathcal{B}$ for which $\mathcal{M}(Bg)$ is minimal.

Definition 2 *Bases $B_1, B_2 \in \mathcal{B}$ are said to be identical to within a resolution J time-shift ($J > 0$) if there exists $q \in \mathbb{Z}$, $0 \le q < 2^J$, such that $\psi(t - 2^{-J}q) \in B_2$ for any $\psi(t) \in B_1$.*

Definition 3 *A best-basis decomposition is said to be shift-invariant up to a resolution level J ($J > 0$) if for any $f, g \in L_2[0,1]$ which are identical to within a resolution J time-shift their respective best bases, B_f and B_g, are identical to within the* same *time-shift.*

Notice that for uniformly sampled discrete functions of length $N = 2^J$, there is an equivalence between an invariance to discrete translation and shift-invariance up to a resolution level J. To demonstrate the shift-invariant properties of SIAP-LTD, compared to LTD which lacks this feature, we refer to the expansions of the signals $g(t)$ (Fig. 1) and $g(t-5\cdot 2^{-7})$. These signals contain $2^7 = 128$ samples. For definiteness, we choose entropy as the cost function. Figs. 2 and 3 depict the "best-basis" expansions under the LTD and the SIAP-LTD algorithms, respectively. A comparison of Fig. 2a and Fig. 2b readily reveals the sensitivity of LTD to temporal shifts while the "best-basis" SIAP-LTD representation is indeed shift-invariant and characterized by a lower entropy (Fig. 3).

3. SMOOTH LOCAL TRIGONOMETRIC BASES

In this section we construct a library of orthonormal bases of $L_2[0,1]$ which consist of sines or cosines multiplied by smooth compactly supported functions.

Let $r \in C^s(\mathbb{R})$ be a rising cutoff function [10], *i.e.,*

$$|r(t)|^2 + |r(-t)|^2 = 1 \quad \text{for all } t \in \mathbb{R}$$

$$r(t) = \begin{cases} 0, & \text{if } t \leq -1 \\ 1, & \text{if } t > 1\,. \end{cases} \tag{1}$$

Then $r(\frac{t-\alpha_I}{\epsilon})r(\frac{\beta_I - t}{\epsilon})$ is a window function supported on $[\alpha_I - \epsilon, \beta_I + \epsilon]$. A local trigonometric function subordinate to the interval $I = [\alpha_I, \beta_I) \stackrel{\triangle}{=} [2^{-\ell}n + 2^{-J}m, 2^{-\ell}(n+1)+2^{-J}m)$ can be defined by

$$\phi^{\rho_0,\rho_1}_{\ell,n,m,k}(t) = r(\frac{t-\alpha_I}{\epsilon})r(\frac{\beta_I - t}{\epsilon})C^{\rho_0,\rho_1}_{I,k}(t) \tag{2}$$

where

$$C^{\rho_0,\rho_1}_{I,k}(t) = 2^{\frac{\ell+1}{2}} h(k + \frac{1+\rho_0+\rho_1}{2})\cos[\pi 2^\ell \cdot (k + \tfrac{1+\rho_0+\rho_1}{2})(t-\alpha_I) - \rho_0 \tfrac{\pi}{2}]$$

is a trigonometric function whose parities at the end points α_I and β_I are specified by ρ_0 and ρ_1, respectively (even-even for $\rho_0 = 0$ and $\rho_1 = 1$, even-odd for 0-0, odd-even for 1-1 and odd-odd for 1-0), and

$$h(j) = \begin{cases} 1/\sqrt{2}, & j = 0\,, \\ 1, & j \neq 0\,. \end{cases}$$

are weight factors needed to insure orthonormality. We call ℓ the resolution-level index ($0 \leq \ell \leq L \leq J$), n position index ($0 \leq n < 2^\ell$), m shift index ($0 \leq m < 2^{J-\ell}$), k frequency index ($k \in \mathbb{Z}_+$) and $\rho_0, \rho_1 \in \{0,1\}$ polarity indices. Since we consider functions defined on the circle $[0,1]$, the basis functions are of the form

$$\psi^{\rho_0,\rho_1}_{\ell,n,m,k}(t) = \chi_{[0,1]} \sum_{q=-1}^{1} \phi^{\rho_0,\rho_1}_{\ell,n,m,k}(t+q) \tag{3}$$

where χ_I is an indicator function for the interval I, *i.e.,* the function that is 1 in I and 0 elsewhere. The role of $\epsilon > 0$ in (2) is to allow overlap of windows, and thus control the smoothness of the window function. ϵ must be small enough ($\epsilon < 2^{-L-1}$) so that every pair of adjacent intervals are compatible [11], *i.e.,* distinct overlapping intervals are disjoint. In order to implement a fast search for the best basis, we organize the library in a tree structure, where each node, indexed by the triplet (ℓ, n, m), represents a subspace with different time-frequency localization characteristics:

$$B^{\rho_0,\rho_1}_{\ell,n,m} = \left\{ \psi^{\rho_0,\rho_1}_{\ell,n,m,k} \;:\; k \in \mathbb{Z}_+ \right\} \tag{4}$$

$$V^{\rho_0,\rho_1}_{\ell,n,m} = \text{clos}_{L_2[0,1]} \langle B^{\rho_0,\rho_1}_{\ell,n,m} \rangle \tag{5}$$

To simplify notation in Lemma 1, the triplet (ℓ, n, m) is replaced by its corresponding interval I. We can expand a parent-node into children-nodes as follows:

Lemma 1 *If I' and I'' are adjacent compatible intervals, then $V^{\rho_0,\rho_1}_{I'} \oplus V^{\rho_1,\rho_2}_{I''} = V^{\rho_0,\rho_2}_{I' \cup I''}$.*

This implies that we can switch from a basis on the interval $I' \cup I''$ to bases on I' and I''.

The inner product of a function g with a basis function is efficiently computed by introducing a folding operator,

$$F(\alpha,\rho)g(t) = \begin{cases} r(\frac{t-\alpha}{\epsilon})g(t) + (-1)^\rho r(\frac{\alpha-t}{\epsilon})g(2\alpha - t), & \text{if } \alpha < t < \alpha + \epsilon \\ \bar{r}(\frac{\alpha-t}{\epsilon})g(t) - (-1)^\rho \bar{r}(\frac{t-\alpha}{\epsilon})g(2\alpha - t), & \text{if } \alpha - \epsilon < t < \alpha \\ g(t), & \text{otherwise} \end{cases} \tag{6}$$

and observing that

$$\phi^{\rho_0,\rho_1}_{\ell,n,m,k}(t) = F^*(\alpha_I,\rho_0)F^*(\beta_I,\rho_1)\chi_I C^{\rho_0,\rho_1}_{I,k}(t) \tag{7}$$

$$\left\langle \psi^{\rho_0,\rho_1}_{\ell,n,m,k}, g \right\rangle = \left\langle \chi_I C^{\rho_0,\rho_1}_{I,k}, F(\alpha_I,\rho_0)F(\beta_I,\rho_1)g \right\rangle \tag{8}$$

where F^* is the adjoint of F. F^*, the unfolding operator, is also the inverse of F owing to unitarity. $F(\alpha,\rho)$ has an odd-even ($\rho = 0$) or even-odd ($\rho = 1$) polarity around $t = \alpha$. That is, if g is smooth, then $\chi_{(-\infty,\alpha]}F(\alpha,0)g$ is a function that is smooth when extended odd to the right and $\chi_{[\alpha,\infty)}F(\alpha,0)g$ is a function that is smooth when extended even to the left. Expression (8) for the coefficients has great importance, since it implies that g can be preprocessed by folding and then represented by a trigonometric basis which reflects the parity properties at the end-points (DCT-II for even-even, DCT-IV for even-odd, DST-II for odd-odd and DST-IV for odd-even parity; all having fast implementation algorithms [12]).

Proposition 1 *Let $E = \{(\ell,n,m)\}$ denote a collection of indices $0 \leq \ell \leq L$, $0 \leq n < 2^\ell$ and $0 \leq m < 2^{J-\ell}$ satisfying*

(i) The segments $\{I_{\ell,n,m} : (\ell,n,m) \in E\}$ are a disjoint cover of $[a, a+1)$, for some $0 \leq a < 1$.

(ii) Nodes (ℓ,n_1,m_1), $(\ell,n_2,m_2) \in E$ at the same resolution level have identical shift index, $m_1 = m_2$.

Then for any $0 \leq P < 2^{2^L}$,

$$\{B^{\rho(\alpha_I),\rho(\beta_I)}_{\ell,n,m} \;:\; (\ell,n,m) \in E\}$$

forms an orthonormal basis of $L_2[0,1]$, where $\rho(\alpha_I)$ and $\rho(\beta_I)$ are the polarities at the end-points α_I and β_I, respectively, given by $\rho(\alpha) = p[2^L(\alpha - a)]$, and $\{p(i)\}$ are defined by $P = \sum_{i=0}^{2^L-1} p(i)2^i$; $\quad p(2^L) \stackrel{\triangle}{=} p(0)$.

The set of all (E, P) specified above generates a library of bases of $L_2[0,1]$. Condition (ii) is supplementary and further restricts the number of bases belonging to the library. It facilitates a reduction in the computational complexity of the best-basis algorithm while retaining shift-invariance.

4. THE BEST BASIS SELECTION

The shift-invariance and the adapted-polarity stem from independent degrees of freedom which are incorporated into the best-basis search algorithm. These are a relative shift between expansions in distinct resolution levels and a variable folding operator whose polarity is locally-adapted at the finest resolution level. The adaptive folding has nothing to do with shift-invariance, but with a reduction of the information cost.

Denote by $A^{\rho_0,\rho_1}_{\ell,n,m}$ the best basis for g restricted to the subspace $V^{\rho_0,\rho_1}_{\ell,n,m}$. Accordingly, $A^{p,p}_{0,0,m}$ for some $0 \le m < 2^J$ and $p \in \{0,1\}$ constitutes the best basis for g. These parameters, namely m and p, are determined recursively together with $A^{p,p}_{0,0,m}$. Let $m_0 = m$ and $P_0 = p_0(0) = p$. Suppose that at the resolution level ℓ we have found m_ℓ, P_ℓ and $A^{p_\ell(n),p_\ell(n+1)}_{\ell,n,m_\ell}$ for all $0 \le n < 2^\ell$, where $\{p_\ell(i)\}$ are defined by $P_\ell = \sum_{i=0}^{2^\ell-1} p_\ell(i)2^i$; $\quad p_\ell(2^\ell + i) \triangleq p_\ell(i)$. Then we will choose $m_{\ell-1}$, $P_{\ell-1}$ and $A^{p_{\ell-1}(n),p_{\ell-1}(n+1)}_{\ell-1,n,m_{\ell-1}}$ for $0 \le n < 2^{\ell-1}$ to minimize the information cost. First, we heed the disjoint union

$$I_{\ell-1,n,m} = I_{\ell,2n+\alpha,m_c} \cup I_{\ell,2n+1+\alpha,m_c} \tag{9}$$

where[1] $\alpha = m \text{ div } 2^{J-\ell}$, $m_c = m \bmod 2^{J-\ell}$ and recall $I_{\ell,n,m} \triangleq [2^{-\ell}n + 2^{-J}m, 2^{-\ell}(n+1) + 2^{-J}m)$. Therefore, it follows from Lemma 1 that

$$A^{\rho_0,\rho_1}_{\ell-1,n,m} = \begin{cases} B^{\rho_0,\rho_1}_{\ell-1,n,m}, \text{ if } \mathcal{M}_B \le \mathcal{M}_A\,, \\ A^{\rho_0,\rho_2}_{\ell,2n+\alpha,m_c} \oplus A^{\rho_2,\rho_1}_{\ell,2n+1+\alpha,m_c}, \text{else} \end{cases} \tag{10}$$

where $\mathcal{M}_A = \mathcal{M}(A^{\rho_0,\rho_2}_{\ell,2n+\alpha,m_c} g) + \mathcal{M}(A^{\rho_2,\rho_1}_{\ell,2n+1+\alpha,m_c} g)$ is the information cost of the children, $\mathcal{M}_B = \mathcal{M}(B^{\rho_0,\rho_1}_{\ell-1,n,m} g)$ the information cost of the parent, and $\rho_2 = p_\ell(2n+1+\alpha)$ is the right polarity of the left child and left polarity of the right child. Now, to acquire shift-invariance it is sufficient to consider two optional values of $m_{\ell-1}$: m_ℓ and $m_\ell + 2^{J-\ell}$. The respective information costs of g when expanded at the resolution level $\ell - 1$ are

$$\mathcal{M}'_{\ell-1} = \sum_{n=0}^{2^{\ell-1}-1} \mathcal{M}(A^{p_\ell(2n),p_\ell(2n+2)}_{\ell-1,n,m_\ell} g) \tag{11}$$

$$\mathcal{M}''_{\ell-1} = \sum_{n=0}^{2^{\ell-1}-1} \mathcal{M}(A^{p_\ell(2n+1),p_\ell(2n+3)}_{\ell-1,n,m_\ell+2^{J-\ell}} g) \tag{12}$$

So we decide on that value of $m_{\ell-1}$ which yields a cheaper representation, *i.e.*,

$$m_{\ell-1} = \begin{cases} m_\ell, \text{ if } \mathcal{M}'_{\ell-1} \le \mathcal{M}''_{\ell-1}\,, \\ m_\ell + 2^{J-\ell}, \quad \text{else.} \end{cases} \tag{13}$$

The polarity at the resolution level $\ell - 1$ is plainly obtained by keeping those bits which correspond to end-points of the same level intervals, namely, for $0 \le n < 2^{\ell-1}$

$$p_{\ell-1}(n) = \begin{cases} p_\ell(2n), \text{ if } \mathcal{M}'_{\ell-1} \le \mathcal{M}''_{\ell-1}\,, \\ p_\ell(2n+1), \quad \text{else} \end{cases} \tag{14}$$

[1] x *div* y denotes the integer part of the ratio x/y, and $x \bmod y$ represents its remainder.

and $p_{\ell-1}(2^{\ell-1} + n) = p_{\ell-1}(n)$.

The recursive procedure is carried out down to a specified level $\ell = L$ $(L \le J)$, where we impose

$$A^{\rho_0,\rho_1}_{L,n,m} = B^{\rho_0,\rho_1}_{L,n,m} \tag{15}$$

and pick a combination of shift and polarity by

$$(m_L, P_L) = Arg \min_{\substack{0 \le m < 2^{J-L} \\ 0 \le P < 2^{2^L}}} \{\sum_{n=0}^{2^L-1} \mathcal{M}(B^{p(n),p(n+1)}_{L,n,m} g)\}. \tag{16}$$

In practice, pursuing a global minimum of the information cost at the finest resolution level, as advised in (16), is worthless, because a sequential consideration of 2^{2^L} polarity values is inoperable. Instead, one should be satisfied with a locally adapted polarity. Fix the shift index m and denote the local information cost by

$$C_{m,n}(\rho) = \min_{\rho_0,\rho_1 \in \{0,1\}} \{\mathcal{M}(B^{\rho_0,\rho}_{L,n,m} g) + \mathcal{M}(B^{\rho,\rho_1}_{L,n+1,m} g)\},$$

$\rho \in \{0,1\}$, $0 \le n < 2^L$. Then the shift and polarity are given by

$$m_L = Arg \min_{0 \le m < 2^{J-L}} \{\sum_{n=0}^{2^L-1} \mathcal{M}(B^{\pi_m(n),\pi_m(n+1)}_{L,n,m} g)\} \tag{17}$$

$$p_L(n) = \pi_{m_L}(n), \quad 0 \le n < 2^L \tag{18}$$

where

$$\pi_m(n) = \begin{cases} 0, \text{ if } C_{m,n}(0) \le C_{m,n}(1) \\ 1, \quad \text{else.} \end{cases}$$

Notice that an ill-adapted polarity-bit is likely to be expunged at coarser resolution levels by merging intervals around it.

Proposition 2 *The best basis expansion stemming from the previously described recursive algorithm is shift-invariant up to a resolution level J.*

The computational complexity of executing SIAP-LTD is $O(N(L + 2^{J-L+1}) \log_2 N)$, where N denotes the length of the signal. This complexity is comparable to that of LTD [2] ($O(NL\log_2 N)$) with the benefits of shift-invariance and a higher quality (lower "information cost") "best-basis".

5. CONCLUSION

The attainment of shift-invariance in "best-basis" expansions necessitates an extended library of bases that includes all shifted versions of bases within the library. Such a library of smooth local trigonometric bases was formed, and an appropriate fast "best-basis" search algorithm was introduced. The gained properties of the generated best-basis representation, namely its shift-invariance, compactness and orthonormality, can be used advantageously in areas such as signal analysis, identification and compression applications.

6. REFERENCES

[1] Y. Meyer, *Wavelets: Algorithms and Applications*, SIAM, Philadelphia, 1993.

[2] R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection", IEEE Trans. Inform. Theory, Vol. 38, pp. 713–718, March 1992.

[3] E. Wesfreid and M.V. Wickerhauser, "Adapted local trigonometric transforms and speech processing", IEEE Trans. SP, Vol. 41, pp. 3596–3600, Dec. 1993.

[4] B. Jawerth, Y. Liu and W. Sweldens, "Signal compression with smooth local trigonometric bases", Optical Eng., Vol. 33, pp. 2125–2135, July 1994.

[5] C. Herley, J. Kovačević, K. Ramchandran and M. Vetterli, "Tilings of the time-frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms", IEEE Trans. Signal Proc., Vol. 41, pp. 3341–3359, Dec. 1993.

[6] S. Mallat and Z. Zhang, "Characterization of signals from multiscale edges", IEEE Trans. PAMI, Vol. 14, pp. 710–732, July 1992.

[7] R. Kronland-Martinet, J. Morlet and A. Grossman, "Analysis of sound patterns through wavelet transforms", Int. J. Patt. Rec. Art. Intell., Vol. 1, pp. 273–301, 1987.

[8] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries", IEEE Trans. Signal Proc., Vol. 41, Dec. 1993.

[9] I. Cohen, S. Raz and D. Malah, "Shift invariant wavelet packet bases", Proc. 20th ICASSP, May 1995.

[10] M.V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, AK Peters, Ltd, 1994.

[11] P. Auscher, G. Weiss and M.V. Wickerhauser, "Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets", in C.K Chui (ed.) *Wavelets — A Tutorial in Theory and Applications*, Academic Press, Inc., 1992, pp. 237–256.

[12] K.R. Rao and P. Yip, *Discrete Cosine Transform*, Academic Press Inc., 1990.

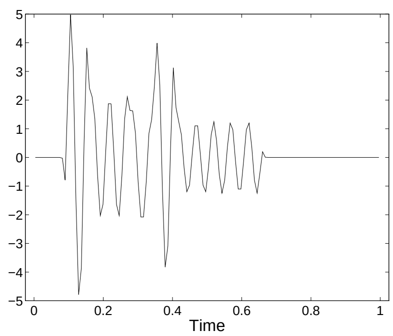

Figure 1: The signal $g(t)$ (2^7 samples).

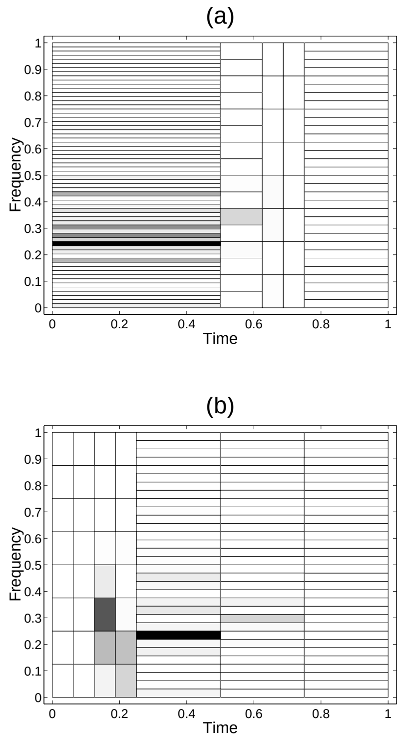

Figure 2: A LTD "best basis" expansions of: (a) $g(t)$, Entropy=2.57. (b) $g(t - 5 \cdot 2^{-7})$, Entropy=2.39.

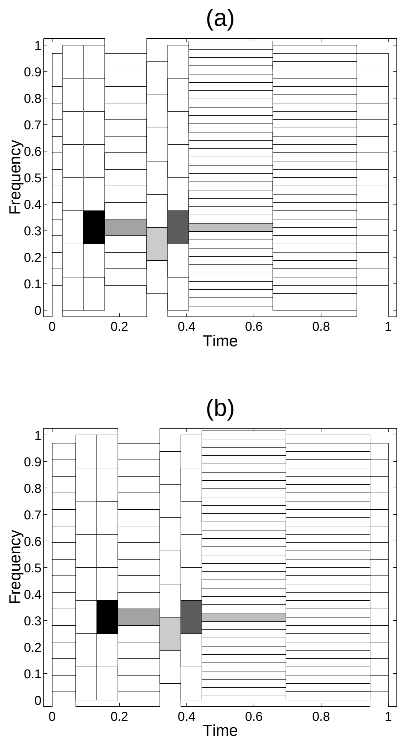

Figure 3: A SIAP-LTD "best basis" expansions of: (a) $g(t)$, Entropy=1.44. (b) $g(t - 5 \cdot 2^{-7})$, Entropy=1.44.