

NEW ALGORITHMS FOR BEST LOCAL COSINE BASIS SEARCH

Y. Huang[†], I. Pollak[†], C.A. Bouman[†], and M.N. Do[‡]

[†] Purdue University
School of ECE
West Lafayette, IN 47907

[‡] University of Illinois at Urbana-Champaign
Department of ECE
Urbana, IL 61801

ABSTRACT

We propose a best basis search algorithm for local cosine dictionaries. We improve upon the classical best local cosine basis selection based on a dyadic tree [2], by considering a larger dictionary of bases. This results in more compact representations, lower costs, and approximate shift-invariance. We also provide a version of our algorithm which is strictly shift-invariant.

1. INTRODUCTION

Adaptive signal representation and approximation in overcomplete dictionaries have received much attention in recent years. The contributions of our paper are in the area of best basis search algorithms where the aim is to adaptively select, from a library of orthonormal bases, the basis which minimizes a cost for a given signal. Such methods have been demonstrated to be effective for compression [11], noise removal [5,6,9], and time-frequency analysis [3, 12].

The original paper on best basis search [2] exploited the fact that 1-D wavelet packet bases [1] and local cosine bases [8] can be organized on a single dyadic tree, making it possible to find the best basis (for an additive cost) via an efficient dynamic programming algorithm. In the case of local cosines, however, it required the restriction of local cosine bases to dyadic intervals. In this paper, we remove this restriction and propose an algorithm to find the best basis in an arbitrary collection of local cosine bases. As we show through several examples, this results in reduced costs, sparser time-frequency pictures, and approximate shift-invariance. We show that this algorithm can moreover be made strictly shift-invariant by using a procedure similar to the one developed in [3].

In a companion paper [10], we show that our dynamic programming algorithm is a 1-D version of a wide class of 2-D optimal basis search algorithms. In [10], we use these 2-D algorithms to find the best rectangular tiling of an image, and to find the best anisotropic wavelet packet basis [13].

Note that the original best basis paper [2] proposed using an entropy cost. Since then, a number of papers have proposed different criteria, e.g., based on the MDL principle [5, 9], Bayesian estimation [6], rate-distortion framework [11]. We do not address the issue of cost selection in this paper, and use the entropy cost for simplicity. Our algorithms, however, can be used in conjunction with any additive or multiplicative cost.

This work was supported in part by a National Science Foundation (NSF) CAREER award CCR-0093105, a Purdue Research Foundation grant, and an NSF CAREER award CCR-0237633. All plots were generated with the help of Wavelab 802 [4].

2. BACKGROUND

The general best basis search problem is formulated, for example, in [2, 7]. We consider a dictionary \mathcal{D} that is the union of orthonormal bases for \mathbb{C}^N , $\mathcal{D} = \bigcup_{\lambda \in \Lambda} B^\lambda$, where each basis B^λ is a family of N vectors, $B^\lambda = \{g_m^\lambda\}_{1 \leq m \leq N}$. The cost of representing a signal f in a basis B^λ is defined as

$$C(f, B^\lambda) = \sum_{m=1}^N \Phi \left(\frac{|\langle f, g_m^\lambda \rangle|^2}{\|f\|^2} \right), \quad (1)$$

where Φ is application dependent. Any basis which achieves the minimum of the cost $C(f, B^\lambda)$ over all the bases in the dictionary, is called the best basis. In this paper, we develop fast algorithms for finding the best basis in local cosine dictionaries.

3. MULTITREE LOCAL COSINE DECOMPOSITIONS.

3.1. A Local Cosine Dictionary

A local cosine family [8] is defined using cosine functions multiplied by overlapping smooth windows. For each discrete interval $[u, v-1] \subset \mathbb{Z}$ of length $a = v-u$, we define a window function $\beta_{u,v}$ which gradually ramps up from zero to one around u and goes down from one to zero around $v-1$:

$$\beta_{u,v}(t) = \begin{cases} r \left(\frac{t-(u-1/2)}{\eta} \right) & \text{if } u - \frac{1}{2} - \eta \leq t < u - \frac{1}{2} + \eta \\ 1 & \text{if } u - \frac{1}{2} + \eta \leq t < v - \frac{1}{2} - \eta \\ r \left(\frac{(v-1/2)-t}{\eta} \right) & \text{if } v - \frac{1}{2} - \eta \leq t \leq v - \frac{1}{2} + \eta \\ 0 & \text{otherwise,} \end{cases}$$

where the parameter $\eta \in \mathbb{R}$ controls how fast the window tapers off, and r is a monotonically increasing profile function,

$$r^2(t) + r^2(-t) = 1 \quad \forall t \in \mathbb{R}; \quad r(t) = \begin{cases} 0 & \text{if } t < -1 \\ 1 & \text{if } t > 1 \end{cases}$$

We always set $a \geq 2\eta$. Following [8], we define the discrete local cosine family $\mathcal{B}_{u,v}$ as follows:

$$\mathcal{B}_{u,v} = \left\{ \frac{\beta_{u,v}(n)\sqrt{2}}{\sqrt{v-u}} \cos \frac{\pi(k + \frac{1}{2})(n - (u - \frac{1}{2}))}{v-u} \right\}_{k=0}^{v-u-1},$$

where $n \in \mathbb{Z}$ is a discrete parameter. It can be shown [8] that this set of signals is orthonormal.

For a signal f of length N , we search for the best basis in the local cosine dictionary

$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} B^\lambda \quad (2)$$

which consists of the following local cosine bases:

$$B^\lambda = \bigcup_{k=0}^{K_\lambda-1} \mathcal{B}_{n_k, n_{k+1}}, \quad (3)$$

where λ is a set of partition points $\{n_k\}_{0 \leq k \leq K_\lambda}$ of the domain of f . We impose that the finest cell size be some fixed integer $M \geq 2\eta$, i.e., we require the partition points to be integer multiples of M . This restriction ensures that only adjacent windows overlap and improves the speed of calculation:

$$\begin{aligned} n_0 = 0 < n_1 < \dots < n_{K_\lambda-1} < n_{K_\lambda} = N \\ n_k \text{ is divisible by } M \text{ where } M \geq 2\eta \text{ is a fixed integer.} \end{aligned}$$

We note that the resulting dictionary is larger than the local cosine tree dictionary of [2]. In fact, if we choose M such that $N/M = 2^J$ where J is the maximum depth of the local cosine tree of [2], it can be easily shown that the local cosine tree dictionary of [2] will be a subset of our dictionary.

3.2. A Best Basis Algorithm.

In [2], the fact that the bases of the dictionary can be organized as a dyadic tree was exploited to yield a fast ($O(NJ \log N)$) dynamic programming algorithm for the best basis search. Although the bases of our dictionary cannot be organized as a single tree, we can still use dynamic programming for the best basis search. Let $f_{u,v}$ denote the following segment of a discrete signal f : $f(u), f(u+1), \dots, f(v-1)$, where $0 \leq u < v \leq N$, and let the best basis for $f_{u,v}$ be $\mathcal{O}_{u,v}$. For $v-u > M$,

$$\mathcal{O}_{u,v} = \begin{cases} \mathcal{B}_{u,d^*} \cup \mathcal{O}_{d^*,v}, & \text{if } C(f, \mathcal{B}_{u,d^*}) + C(f, \mathcal{O}_{d^*,v}) < C(f, \mathcal{B}_{u,v}) \\ \mathcal{B}_{u,v}, & \text{otherwise} \end{cases} \quad (4)$$

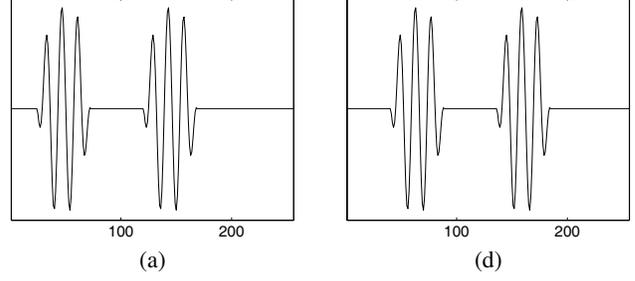
where

$$d^* = \arg \min_{d: u < d < v, d \text{ is a multiple of } M} C(f, \mathcal{B}_{u,d}) + C(f, \mathcal{O}_{d,v}).$$

(Note that, since the cost function is additive, the cost of $\mathcal{B}_{u,d} \cup \mathcal{O}_{d,v}$ is $C(f, \mathcal{B}_{u,d}) + C(f, \mathcal{O}_{d,v})$.) The initial condition is that for $v-u = M$, $\mathcal{O}_{u,v} = \mathcal{B}_{u,v}$.

Then the best basis $\mathcal{O}_{0,N}$ for signal f can be calculated recursively by applying (4). If $L = N/M$, then the recursion formula (4) takes $O(L^2)$ operations. The major computational burden is associated with computing the costs. The calculation of $C(f, \mathcal{B}_{u,v})$ via the definition (1) involves $O(a)$ additions where $a = v-u$, as well as the computation of the inner product of f with each basis function in $\mathcal{B}_{u,v}$ which requires $O(a \log a)$ operations using a fast local cosine transform algorithm [8]. In the process of calculating $\mathcal{O}_{0,N}$, we need the values for $C(f, \mathcal{B}_{u,v})$ with $u = pM, v = qM$ where $p = 0, 1, \dots, L-1$ and $q = p+1, p+2, \dots, L$. It is easy to show that this results in the overall time complexity of $O(L^2 N \log N) = O(N^3/M^2 \log N)$.

Since the bases in our dictionary can be organized using multiple trees (as opposed to a single tree in [2]), we call our method a multitree local cosine decomposition.



Time-frequency planes for best local cosine bases:

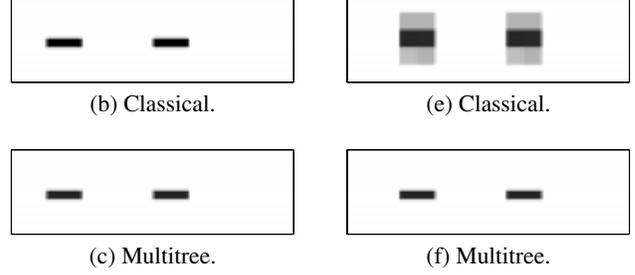
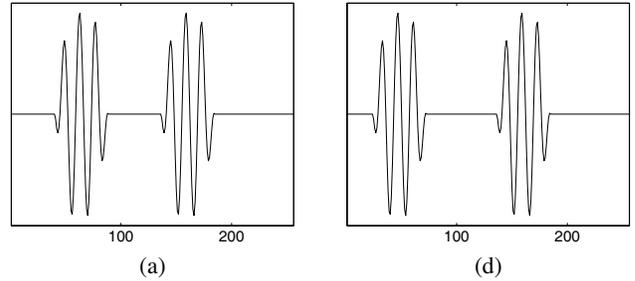


Fig. 1. The original signals and time-frequency representations of the best local cosine basis with smallest cell size $M = 16$: (a) a signal consisting of two local cosine basis functions; (b) the time-frequency tiling for the best local cosine basis of [2]; (c) the time-frequency tiling for the best multitree local cosine basis; (d-f) a similar experiment for the signal in (a) shifted by 16 samples. The darker the rectangle in (b,c,e,f) the larger the amplitude of the corresponding local cosine coefficient.



Time-frequency planes for best local cosine bases:

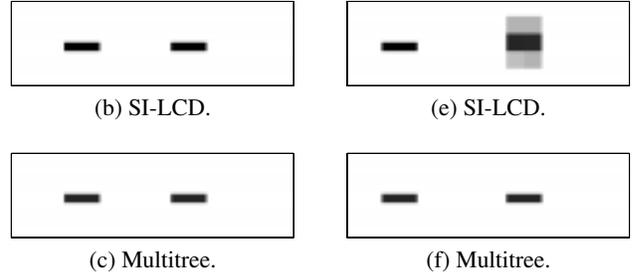


Fig. 2. The original signals and time-frequency representations of the best local cosine basis with smallest cell size $M = 16$: (a) The signal from Fig. 1(d); (b) the shift-invariant local cosine decomposition [3]; (c) a copy of Fig. 1(f), i.e., the best multitree local cosine basis; (d-f) a similar experiment for a signal where the two local cosine bumps are shifted by different amounts.

3.3. Shift-Invariance: A Qualitative Discussion.

We call a best basis search algorithm n_0 -*shift-invariant* if circularly shifting any signal by an arbitrary integer multiple of n_0 leads to shifting its best basis by the same multiple of n_0 . When $n_0 = 1$ —i.e., when the algorithm is invariant to any integer shift, we simply call it shift-invariant.

Our method described in the previous subsection is, strictly speaking, not M -shift-invariant, since we always require the leftmost basis function to start at the leftmost point of the signal. It is, however, M -shift-invariant, modulo these boundary effects: i.e., it is invariant to shifts by integer multiples of M for signals whose support is well within the interval $[0, N - 1]$.

The classical best local cosine basis algorithm of [2] is fundamentally not shift-invariant since it uses a dyadic tree. Its variant introduced in [3] is formally shift-invariant; however we now show that our method offers certain advantages.

To illustrate the shift-invariance properties of the algorithms, we use a 256-point signal depicted in Fig. 1(a) which consists of two local cosine basis functions, one with $u = 32$ and $v = 64$, and another one with $u = 128$ and $v = 160$. For each algorithm, the smallest cell size M is chosen to be 16. (For the single-tree methods, this means that the maximal tree depth is set to $J = \log_2 N - \log_2 M = 4$.) Figs. 1(b) and (c) show the time-frequency tilings for the best basis extracted by the classical method of [2] and our multitree method, respectively. These are identical. However, when the signal is shifted by 16 samples to the right (Fig. 1(d)), the result for our method stays the same (Fig. 1(f)) whereas the classical best basis changes drastically (Fig. 1(e)).

This can be fixed by the shift-invariant local cosine decomposition (SI-LCD) proposed in [3] which essentially considers N shifted versions of the dictionary, and is therefore shift-invariant. The signal of Fig. 2(a) is identical to that of Fig. 1(d). The best basis extracted by the SI-LCD algorithm Fig. 2(b) is identical to the best multitree basis. Let us now consider another signal, obtained by taking the signal of Fig. 1(a), retaining its first component as is, and shifting its second component to the right by 16, as shown in Fig. 2(d). Our algorithm is still invariant to this change, see Fig. 2(f). SI-LCD, however, produces a different basis.

3.4. A Strictly Shift-Invariant Algorithm.

The qualitative discussion of the previous subsection shows that our algorithm possesses the desired shift-invariant properties, even though it is not, strictly speaking, shift-invariant. We now show, in addition, that we can make it strictly invariant to any integer shift, using a method similar to [3].

For a discrete signal f of length N , we extend both the signal and the basis functions periodically with period N . We expand the dictionary of Subsection 3.1 by adding in the shifts of the basis signals. We define \mathcal{D}_0 to be the same as the dictionary of Eqs. (2,3), and let \mathcal{D}_s be \mathcal{D}_0 shifted by s to the left. The new dictionary \mathcal{D}_{SI} is defined to be the union of the N shifted subdictionaries:

$$\mathcal{D}_{SI} = \bigcup_{s=0}^{N-1} \mathcal{D}_s.$$

Now the best basis search involves finding the subdictionary \mathcal{D}_{s^*} that contains the best basis and searching for the best basis in \mathcal{D}_{s^*} . Using an argument similar to the one in Subsection 3.2, it can be shown that the optimal solution is achieved in $O(N^3 \log N)$. We

adopt a suboptimal solution based on the method in [3], to result in the time complexity similar to that of Subsection 3.2.

Recall that M is the size of the finest cell we are considering, and $L = N/M$ is the total number of such cells. We let $s = lM + m$ where l is the cell where s appears, $0 \leq l < L$, and m is the position of s within the cell, $0 \leq m < M$. Instead of finding s^* , i.e., optimizing over m and l jointly, we first optimize over m and then optimize over l . We optimize over m using the method described in [3]. For each shift $m = 0, 1, \dots, M - 1$, define the following basis B_m :

$$B_m = \bigcup_{l=0}^{L-1} \mathcal{B}_{lM+m, (l+1)M+m}.$$

We calculate the cost C_m of approximating f with the basis B_m :

$$C_m = \sum_{l=0}^{L-1} C(f, \mathcal{B}_{lM+m, (l+1)M+m}).$$

Then m^* is found by minimizing the cost C_m over m :

$$m^* = \arg \min_{0 \leq m \leq M-1} C_m.$$

Now we optimize over l . Let $\mathcal{O}_{lM+m^*, N+lM+m^*}$ be the best basis for the signal $f_{lM+m^*, N+lM+m^*}$ in the subdictionary \mathcal{D}_{lM+m^*} , for $l = 0, 1, \dots, L - 1$. The best basis for each l is calculated in the same way as in Subsection 3.2. Then we choose the best l :

$$l^* = \arg \min_{0 \leq l \leq L-1} C(f, \mathcal{O}_{lM+m^*, N+lM+m^*}).$$

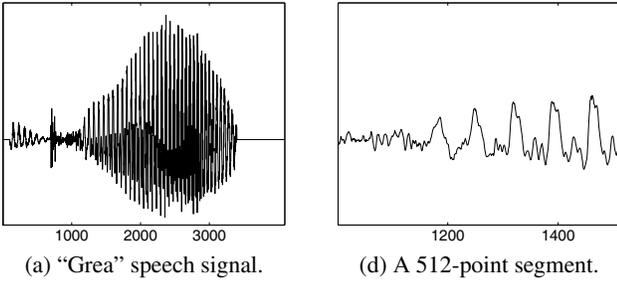
The corresponding (suboptimal) shift-invariant best basis \mathcal{O} for the signal f is $\mathcal{O}_{l^*M+m^*, N+l^*M+m^*}$. It can be shown that this results in the overall time complexity of $O((L^2 + M)N \log N) = O((N/M)^2 + M)N \log N$, which is the same as in Subsection 3.2 if M is not too large. For comparison, the algorithm of [3] is $O((\log(N/M) + M)N \log N)$.

4. EXAMPLES

To further illustrate our methods, we use two examples which compare our proposed multitree local cosine decomposition with the classical best local cosine basis selection based on a single dyadic tree [2]. We follow [2] and use the entropy cost function for all the experiments, i.e., $\Phi(x) = -x \log_e x$ in Eq. (1). We set $\eta = 8$.

Fig. 3(a-c) shows a speech signal of length $N = 4096$ and the time-frequency pictures for the best bases selected by the two methods. The minimal cell size for these experiments was set at $M = 16$ for both methods. The resulting costs are: 4.11 for the classical method and 3.51 for our method. In addition, note the more sparse time-frequency picture resulting from our method.

In Fig. 3(d), we zoom into the samples 1001 through 1512 of the signal in Fig. 3(a). For this 512-point segment, we compute the best basis with the two methods, again setting $M = 16$. This results in the following costs: 2.44 for the classical method and 2.02 for our method. Again, the representation resulting from our method corresponds to a more sparse time-frequency tiling. Moreover, the transition between two phonemes (in the neighborhood of the sample 1150) is missed by the best single-tree basis but is accurately captured by the best multitree basis.



Time-frequency planes for best local cosine bases:

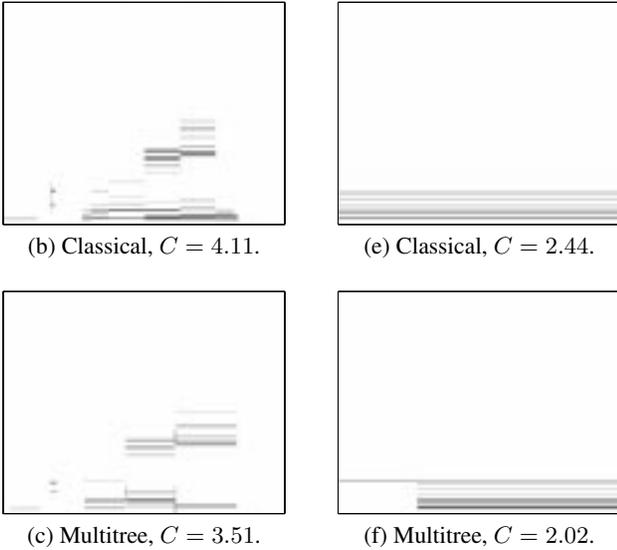


Fig. 3. Two signals and the time-frequency pictures of their best bases: (a) segment “grea” of the speech signal “greasy”; (b,c) the time-frequency tilings for the best local cosine basis of [2] and for the best multitree local cosine basis, respectively; (d-f) a similar experiment for a shorter segment of the speech signal.

Fig. 4 summarizes a larger experiment where the best basis was found for four different values of the minimal cell size M : 128, 64, 32, and 16. In addition to the classical single-tree algorithm and our multitree algorithm, we compared the results with the shift-invariant versions of the two algorithms: the shift-invariant local cosine decomposition (SI-LCD) [3] and our shift-invariant multitree algorithm described above. The resulting costs for the four algorithms are plotted as a function of M . Note that in both cases, the whole curve for the multitree algorithm is below each of the outcomes for the algorithms in [2, 3]. This is to be expected since we perform the search over a much larger dictionary. The price to pay is the time complexity of the algorithm, which, as indicated above, is higher than the time complexity for the classical best-basis search algorithm. Note, however, that for small signal lengths the running time of the two algorithms is similar. For example, in our experiment with the 512-point signal, the running times for $M = 128, 64, 32, 16$ are 0.01, 0.02, 0.02, and 0.03 seconds, respectively, for the classical algorithm and 0.01, 0.01, 0.05, and 0.19, respectively, for our algorithm. This suggests that the most practical way of using our algorithms is on small blocks. We are currently developing efficient methods for doing this.

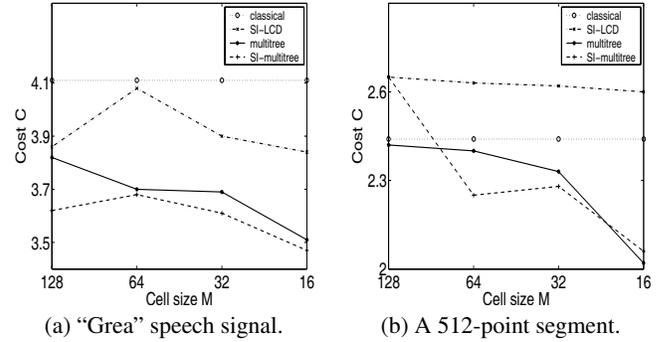


Fig. 4. The performance of four algorithms for extracting the best local cosine basis: classical single-tree algorithm [2] (dotted), shift-invariant LCD [3] (dash-dot), the proposed multitree algorithm (solid), and the proposed shift-invariant version of the multitree algorithm (dashed). The optimal cost is depicted as a function of the minimal allowed cell size: (a) 4096-point “grea” speech signal, (b) 512-point segment of the signal.

5. CONCLUSIONS

We have developed a new best basis search algorithm to adaptively compute the optimal multitree local cosine decomposition. Simple examples show that the algorithm yields lower costs, sparser representations, and better shift-invariance properties than the classical best basis search. In addition, it can better represent important time-frequency features.

6. REFERENCES

- [1] R.R. Coifman, Y. Meyer, and M.V. Wickerhauser. Wavelet analysis and signal processing. In *Wavelets and Their Applications*, M.B. Ruskai *et al.*, Eds., pp. 153-178. Jones and Bartlett, Boston, 1992.
- [2] R.R. Coifman and M.V. Wickerhauser. Entropy based algorithms for best basis selection. *IEEE Trans. Inf. Th.*, 38(2):713-718, March 1992.
- [3] I. Cohen, S. Raz, and D. Malah. Orthonormal shift-invariant adaptive local trigonometric decomposition. *Sig. Proc.*, 57(1):43-64, Feb. 1997.
- [4] D. Donoho, M.R. Duncan, X. Huo, O. Levi, J. Buckheit, M. Clerc, J. Kalifa, S.G. Mallat, T. Yu. Wavelab 802. www-stat.stanford.edu/~wavelab.
- [5] H. Krim and J.-C. Pesquet. On the statistics of best bases criteria. In *Wavelets and Statistics*, Lecture Notes in Statistics, A. Antoniadis, Ed., pp. 193-207. Springer-Verlag, 1995.
- [6] H. Krim, D. Tucker, S. Mallat, D. Donoho. On denoising and best signal representation. *IEEE Trans. Inf. Th.*, 45(7):2225-2238, Nov. 1999.
- [7] S.G. Mallat. *A Wavelet Tour of Signal Processing*, Second Edition. Academic Press, 1999.
- [8] H. Malvar. *Signal Processing with Lapped Transforms*. Artech House, 1992.
- [9] P. Moulin. Signal estimation using adapted tree-structured bases and the MDL principle. In *Proc. IEEE-SP Int. Symp. TFTS*, pp. 141-143, Paris, June 1996.
- [10] I. Pollak, M.N. Do, and C.A. Bouman. Optimal tilings and best basis search in large dictionaries. In *Proc. 37-th Asilomar Conf. Sig., Syst., and Comp.*, Nov. 9-12, 2003.
- [11] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. Im. Proc.*, 2(2):160-175, Apr. 1993.
- [12] L.F. Villemoes. Adapted bases of time-frequency local cosines. Preprint, Feb. 2000, www.math.kth.se/~larsv/publ.html.
- [13] D. Xu and M.N. Do. Anisotropic 2-D wavelet packets and rectangular tiling: theory and algorithms. In *Proc. SPIE Conf. on Wavelet Appl. in Sig. and Im. Proc. X*, San Diego, Aug. 2003.