LOCAL COSINE BASES IN TWO DIMENSIONS

Jelena Kovačević
Signal Processing Research Department
AT&T Bell Laboratories
600 Mountain Avenue, Murray Hill, NJ 07974

Abstract

We construct two-dimensional local cosine bases in discrete and continuous time. Solutions are offered both for rectangular and nonrectangular lattices. In the case of nonrectangular lattices, the problem is solved by mapping it into a one-dimensional equivalent problem.

1 Introduction

Discrete-time cosine modulated filter banks, or, modulated lapped transforms (MLT's), have been in use for some time [1, 2, 3, 4]. Due to a few of their properties, they have become quite popular; For example, all filters (basis functions) of a filter bank are obtained by appropriate modulation of a single prototype filter. Then, fast algorithms exist, making them very attractive for implementation. Finally, they have been used recently to achieve time-varying splittings of the time-frequency plane [5].

Their continuous counterpart, termed "Malvar's wavelets", has found use in decomposing a signal into a linear combination of time-frequency atoms [6].

Modulated lapped transforms have been used extensively in audio coding [7]. They have also found use in image coding, due to the reduction of blocking effects [8] when compared to the DCT. Some video works contain modulated lapped transforms as well [9]. However, in all these applications, one-dimensional MLT's are used separately. We develop here true two-dimensional nonseparable MLT's, which offer more degrees of freedom. We consider both rectangular and nonrectangular sampling structures, and offer solutions for both. In continuous time, similar analysis is performed.

2 Review of Local Cosine Bases

Discrete-Time Case: By local cosine bases, or, modulated lapped transforms we will denote a class of perfect reconstruction filter banks which uses a single prototype filter, window, $w[n]$ of length $2N$ (where N is the number of channels and is even) to construct all of the filters $h_0, \ldots, h_{N-1}$ as follows:

$$h_k[n] \;=\; \frac{w[n]}{\sqrt{N}} \cdot \cos[\frac{2k+1}{4N}(2n-N+1)\pi], \qquad (1)$$

with $k = 0, \ldots, N-1,\ n = 0, \ldots, 2N-1$, and where the prototype lowpass filter $w[n]$ is symmetric, that is, $w[n] \;=\; w[2N-1-n], n = N, \ldots, 2N-1$, and satisfies the following [2]:

$$w^2[n] + w^2[N-1-n] \;=\; 2, \quad n = 0, \ldots, N-1. \qquad (2)$$

This last condition, imposed on the window, ensures that the resulting modulated lapped transform is orthogonal. The two symmetric halves of the window are called "tails".

A convenient way of analyzing filter banks in the time domain, uses infinite matrices, which describe the action of the filters on the input signal. For modulated lapped transforms, such an infinite matrix matrix $\mathbf{T}$ can be written as

$$\mathbf{T} \;=\; \begin{bmatrix} \ddots & & & & \\ & \mathbf{A}_0 & \mathbf{A}_1 & & \\ & & \mathbf{A}_0 & \mathbf{A}_1 & \\ & & & & \ddots \end{bmatrix}, \qquad (3)$$

where blocks $\mathbf{A}_0, \mathbf{A}_1$ are of sizes $N \times N$, and contain the impulse responses of the filters. Note that the filter length is twice the number of channels. For example, the jth row of $\mathbf{A}_i$ is $(h_j[2N-1-iN] \ldots h_j[N-iN])$ for $i = 0, 1$. For an orthogonal, perfect reconstruction solution, the matrix $\mathbf{T}$ has to be unitary, which is equivalent to the following [10]:

$$\mathbf{A}_0^T\mathbf{A}_0 \;+\; \mathbf{A}_1^T\mathbf{A}_1 \;=\; \mathbf{I}, \qquad (4)$$
$$\mathbf{A}_1^T\mathbf{A}_0 \;=\; \mathbf{A}_0^T\mathbf{A}_1 \;=\; \mathbf{0}. \qquad (5)$$

The conditions (5) are called the "orthogonality of tails" conditions [10]. One more fact will be of use

0-7803-2431-5/95 $4.00 © 1995 IEEE

later. Call $\mathbf{B}_i$ the blocks when no windowing is used, or $w[n] = 1$, $n = 0, \ldots, 2N-1$. That is, these blocks will just contin the cosines. Then

$$\mathbf{A}_0 = \mathbf{B}_0 \cdot \mathbf{W}, \quad \mathbf{A}_1 = \mathbf{B}_1 \cdot \mathbf{JWJ}, \qquad (6)$$

where $\mathbf{W}$ is a diagonal matrix with window coefficients on the diagonal $w[0], \ldots, w[N-1]$. The blocks $\mathbf{B}_i$ also satisfy

$$\mathbf{B}_0^T \mathbf{B}_0 = \frac{1}{2}(\mathbf{I} - \mathbf{J}), \quad \mathbf{B}_1^T \mathbf{B}_1 = \frac{1}{2}(\mathbf{I} + \mathbf{J}).$$

Continuous-Time Case: In continuous time, consider the following set of basis functions:

$$\varphi_{j,k}(t) = \sqrt{\frac{2}{L_j}}\, w_j(t)\, \cos[\frac{\pi}{L_j}(k + \frac{1}{2})(t - a_j)], \quad (7)$$

for $k = 0, 1, 2, \ldots$, and $j \in \mathcal{Z}$ and the window function $w_j(t)$ is centered around the interval $[a_j, a_{j+1}]$. As can be seen, (7) is the continuous-time counterpart of (1) seen in the discrete-time case. The a_j's are an increasing sequence of real numbers with $j \in \mathcal{Z}$, $\ldots a_{j-1} < a_j < a_{j+1} \ldots$ We will denote by L_j the distance between a_{j+1} and a_j, $L_j = a_{j+1} - a_j$. We will also assume that we are given a sequence of numbers $\eta_j > 0$ such that $0 < \eta_j, \eta_j + \eta_{j+1} < L_j, j \in \mathcal{Z}$. The windows $w_j(t)$ given will be derivable (possibly infinitely), and of compact support, with the following requirements:

1. $0 \leq w_j(t) \leq 1$, $w_j(t) = 1$ if $a_j + \eta_j \leq t \leq a_{j+1} - \eta_{j+1}$,
2. $w_j(t)$ is supported within $[a_j - \eta_j, a_{j+1} + \eta_{j+1}]$,
3. if $\mid t - a_j \mid \leq \eta_j$ then $w_{j-1}(t) = w_j(2a_j - t)$, and $w_{j-1}^2(t) + w_j^2(t) = 1$.

This last condition ensures that the "tails" of the adjacent windows are power complementary. With these conditions, the set of functions as in (7) forms an orthonormal basis for $L_2(\mathcal{R})$. Therefore, in this most general case, the window can go anywhere from length $2L$ to length L (being a constant window in this latter case of height 1), and is arbitrary as long as it satisfies the above three conditions.

The time-domain functions obtained here are local and smooth, and their Fourier transforms have arbitrary polynomial decay (depending on the smoothness or derivability of the window). Thus, the time-bandwidth product is now finite, and we have a local modulated basis with good time-frequency localization.

3 Discrete-Time Two-Dimensional Cosine Bases

Rectangular Sampling: We assume that we have rectangular sampling, N_1 in horizontal dimension, and N_2 in the vertical one. We will construct the filters as follows:

$$h_{ij}[n_1, n_2] = \frac{w[n_1, n_2]}{\sqrt{N_1 N_2}} \cdot m[n_1, n_2], \qquad (8)$$

where

$$m[n_1, n_2] = \cos[\frac{2i+1}{4N_1}(2n_1 - N_1 + 1)\pi] \cdot \cos[\frac{2j+1}{4N_2}(2n_2 - N_2 + 1)\pi],$$

with $i = 0, \ldots, N_1 - 1, j = 0, \ldots, N_2 - 1$, and $n_1 = 0, \ldots, 2N_1 - 1, n_2 = 0, \ldots, 2N_2 - 1$. The corresponding filters are of size $2N_1 \times 2N_2$.

The counterpart of one block-row of the matrix $\mathbf{T}$ from (3) is

$$\mathbf{T}_1 = (\mathbf{D}_0 \quad \mathbf{D}_1 \quad \mathbf{D}_2 \quad \mathbf{D}_3),$$

where each block $\mathbf{D}_i$ is of size $N_1 N_2 \times N_1 N_2$, and $\mathbf{D}_i$ are given by:

$$\mathbf{D}_0 = \mathbf{C}_0 \cdot \mathbf{W}_0, \quad \mathbf{D}_2 = \mathbf{C}_2 \cdot \mathbf{JW}_0\mathbf{J},$$
$$\mathbf{D}_1 = \mathbf{C}_1 \cdot \mathbf{W}_1, \quad \mathbf{D}_3 = \mathbf{C}_3 \cdot \mathbf{JW}_1\mathbf{J}$$

Here, diagonal matrices $\mathbf{W}_i$ contain appropriately placed coefficients of the two-dimensional, persymmetric window function $w[n_1, n_2]$ as:

$$\mathbf{W} = \begin{bmatrix} \mathbf{JW}_1\mathbf{J} & \mathbf{JW}_0\mathbf{J} \\ \mathbf{W}_0 & \mathbf{W}_1 \end{bmatrix}$$

Window matrices $\mathbf{W}_i$ are then obtained as:

$$\mathbf{W}_0 = \begin{bmatrix} w[0,0] & & \\ & \ddots & \\ & & w[N_1 - 1, N_2 - 1] \end{bmatrix},$$

$$\mathbf{W}_1 = \begin{bmatrix} w[N_1, 0] & & \\ & \ddots & \\ & & w[2N_1 - 1, N_2 - 1] \end{bmatrix},$$

where we grow the horizontal dimension first. Blocks $\mathbf{C}_i$ are given by

$$\mathbf{C}_0 = \mathbf{B}_{0_2} \otimes \mathbf{B}_{0_1}, \quad \mathbf{C}_1 = \mathbf{B}_{0_2} \otimes \mathbf{B}_{1_1},$$
$$\mathbf{C}_2 = \mathbf{B}_{1_2} \otimes \mathbf{B}_{0_1}, \quad \mathbf{C}_3 = \mathbf{B}_{1_2} \otimes \mathbf{B}_{1_1},$$

where block $\mathbf{B}_{i_j}$ is the ith block in dimension j as in (6). The conditions for perfect reconstruction are

$$\mathbf{D}_0^T\mathbf{D}_0 + \mathbf{D}_1^T\mathbf{D}_1 + \mathbf{D}_2^T\mathbf{D}_2 + \mathbf{D}_3^T\mathbf{D}_3 = \mathbf{I}, \quad (9)$$
$$\mathbf{D}_1^T\mathbf{D}_0 + \mathbf{D}_3^T\mathbf{D}_2 = \mathbf{0}, \quad (10)$$
$$\mathbf{D}_3^T\mathbf{D}_1 + \mathbf{D}_2^T\mathbf{D}_0 = \mathbf{0}, \quad (11)$$
$$\mathbf{D}_3^T\mathbf{D}_0 = \mathbf{0}, \quad (12)$$
$$\mathbf{D}_2^T\mathbf{D}_1 = \mathbf{0}. \quad (13)$$

Conditions (10)-(13) can be easily verified while (9) will lead to the conditions on the window, the first one being the two-dimensional counterpart of (2)

$$\mathbf{W}_0^2 + \mathbf{J}\mathbf{W}_0^2\mathbf{J} + \mathbf{W}_1^2 + \mathbf{J}\mathbf{W}_1^2\mathbf{J} = \mathbf{I}, \quad (14)$$

$$\begin{aligned} & -\mathbf{W}_0(\mathbf{I}\otimes\mathbf{J})\mathbf{W}_0 + \mathbf{J}\mathbf{W}_0\mathbf{J}(\mathbf{I}\otimes\mathbf{J})\mathbf{J}\mathbf{W}_0\mathbf{J} + \\ + \ & \mathbf{W}_1(\mathbf{I}\otimes\mathbf{J})\mathbf{W}_1 - \mathbf{J}\mathbf{W}_1\mathbf{J}(\mathbf{I}\otimes\mathbf{J})\mathbf{J}\mathbf{W}_1\mathbf{J} = \mathbf{0}, \\ & \mathbf{W}_0\mathbf{J}\mathbf{W}_0 - \mathbf{W}_1\mathbf{J}\mathbf{W}_1 = \mathbf{0}, \\ & -\mathbf{W}_0(\mathbf{J}\otimes\mathbf{I})\mathbf{W}_0 + \mathbf{J}\mathbf{W}_0\mathbf{J}(\mathbf{J}\otimes\mathbf{I})\mathbf{J}\mathbf{W}_0\mathbf{J} - \\ - \ & \mathbf{W}_1(\mathbf{J}\otimes\mathbf{I})\mathbf{W}_1 + \mathbf{J}\mathbf{W}_1\mathbf{J}(\mathbf{J}\otimes\mathbf{I})\mathbf{J}\mathbf{W}_1\mathbf{J} = \mathbf{0}. \end{aligned}$$

To summarize, in the two-dimensional case with rectangular sampling, we will have up to $(N_1N_2)/2$ free variables. Compare that to $(N_1+N_2)/2$ free variables in the two-dimensional case with separable sampling.

Nonrectangular Sampling: The nonrectangular case is a more difficult one. We offer a solution that will use a particular mapping from one dimension into two dimensions Note that this solution would mean that the filters are obtained by shifting the prototype filter along a line, and that it is very similar to what was done in [11]. It will hold for an even sampling density N. First, we find the upper-triangular form of the sampling matrix

$$\mathbf{D} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix},$$

with $N = \det(\mathbf{D})$ and we assume that b and c do not have common factors. For the support of our filters, we will take two unit cells, the ones located at points $[0,0]$ and $[a,0]$. Then, define the filters as follows:

$$h_k[n_1,n_2] = \frac{w[n_1,n_2]}{\sqrt{N}}\cos[\frac{2k+1}{4N}(2(cn_1-bn_2)-N+1)\pi], \quad (15)$$

with $k = 0,\ldots,N-1$, and n_1, n_2 belonging to the unit cell as explained above. By doing this, we have mapped the problem into the one-dimensional problem, that is

$$\begin{aligned} h_k[0,0] &= h_k[0], \\ h_k[\frac{b+1}{c},1] &= h_k[1], \\ &\vdots \\ h_k[2a-1+\frac{(c-1)(b+1)}{c},c-1] &= h_k[2N-1]. \end{aligned}$$

All the proofs are now equivalent. The only thing left is the condition on the window. The window has to be persymmetric and

$$w^2[n_1,n_2] + w^2[\frac{(c-1)(b+1)}{c}+a-1-n,c-1-n_2] = 2,$$

since

$$\mathbf{W}^2 + \mathbf{J}\mathbf{W}^2\mathbf{J} = 2\mathbf{I}.$$

In the quincunx case, for example, this scheme would lead to one-dimensional filters. However, if we replace $cn_1 - bn_2$ with $n_1 + n_2$ with $[n_1,n_2] \in \{[0,0],[1,0],[1,1],[2,1]\}$, the whole problem is again mapped into a one-dimensional problem and thus easily solved.

4 Continuous-Time Two-Dimensional Cosine Bases

The construction in continuous time is very similar to what we have just presented in discrete time. Thus, we will give a short overview, while for more details and proof of the basis property, refer to [12].

In continuous time, consider the following set of basis functions:

$$\varphi_{j_1,j_2,k_1,k_2}(t_1,t_2) = \sqrt{\frac{2}{L_{j_1}L_{j_2}}}\, w_{j_1,j_2}(t_1,t_2)\, m(t_1,t_2), \quad (16)$$

where

$$\begin{aligned} m(t_1,t_2) &= \cos[\frac{\pi}{L_{j_1}}(k_1+\frac{1}{2})(t_1-a_{j_1})]\cdot \\ &\quad \cos[\frac{\pi}{L_{j_2}}(k_2+\frac{1}{2})(t_2-a_{j_2})], \end{aligned}$$

for $k_i = 0,1,2,\ldots, i=1,2$, and $j \in \mathcal{Z}$. As can be seen, (16) is the continuous-time counterpart of (8) seen in the discrete-time case. The a_{j_i}'s are increasing sequences of real numbers with $j_i \in \mathcal{Z}$, $\ldots a_{j_i-1} < a_{j_i} < a_{j_i+1}\ldots$ We will denote by L_{j_i} the distance between a_{j_i+1} and a_{j_i}, $L_{j_i} = a_{j_i+1} - a_{j_i}$. We will also assume that we are given a sequence of numbers $\eta_{j_i} > 0$ such that $0 < \eta_{j_i} \le L_{j_i}/2$, $0 < \eta_{j_i} \le L_{j_i-1}, j_i \in \mathcal{Z}$. The windows $w_{j_1,j_2}(t_1,t_2)$ given will be derivable (possibly infinitely), and of compact support, with the following requirements:

1. $0 \le w_{j_1,j_2}(t_1,t_2) \le 1$, $w_{j_1,j_2}(t_1,t_2) = 1$ if $a_{j_i} + \eta_{j_i} \le t_i \le a_{j_i+1} - \eta_{j_i+1}$,

2. $w_j(t)$ is supported within $[a_{j_1} - \eta_{j_1}, a_{j_1+1} + \eta_{j_1+1}] \times [a_{j_2} - \eta_{j_2}, a_{j_2+1} + \eta_{j_2+1}]$,

3. tails of adjacent windows are power complementary (see [12] for more details).

With these conditions, the set of functions as in (16) forms an orthonormal basis for $L_2(\mathcal{R}^2)$ [12]. Note that as presented above this would be the counterpart of the rectangular case in discrete time. For a more general tiling of the space, that is, counterpart of the nonrectangular case, see [12]. Therefore, in this most general case, the window can go anywhere from length $2L_1 \times 2L_2$ to length $L_1 \times L_2$ (being a constant window in this latter case of height 1), and is arbitrary as long as it satisfies the above three conditions.

5 Conclusions

Two-dimensional local cosine bases were presented both in discrete and in continuous time. In discrete time, we examine both rectangular and nonrectangular lattices. Although solutions for the rectangular lattices are more important in practice, those for the nonrectangular ones are a difficult challenge. In that case, the problem is solved by mapping it inteo an equivalent one-dimensional problem. As a result, solutions easily follow, however, the resulting filters will be obtained by modulation along a single line. More general modulation structures are a topic of current research. In continuous time, similar analysis was performed, leading to the two-dimensional counterpart of the "Malvar's" wavelets.

Acknowledgement: I would like to thank Tsuhan Chen of AT&T Bell Laboratories for his useful comments.

References

[1] H. Malvar, *Signal Processing with Lapped Transforms.* Norwood, MA: Artech House, 1992.

[2] J. Princen, A. Johnson, and A. Bradley, "Subband transform coding using filter bank designs based on time domain aliasing cancellation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, (Dallas, TX), pp. 2161–2164, April 1987.

[3] T.A. Ramstad and J. Tanem, "Cosine-modulated analysis-synthesis filterbank with critical sampling and perfect reconstruction," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, (Toronto, Canada), pp. 1789–1792, May 1991.

[4] R.D. Koilpillai and P.P. Vaidyanathan, "New results on cosine-modulated FIR filter banks satisfying perfetc reconstruction," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, (Toronto, Canada), pp. 1793–1796, May 1991.

[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., Special Issue on Wavelets and Signal Processing*, vol. 41, pp. 3341–3359, December 1993.

[6] Y. Meyer, *Wavelets: Algorithms and Applications.* Philadelphia, PA: SIAM, 1993. Translated and revised by R.D. Ryan.

[7] C. Todd, G. Davidson, M. Davis, L. Fielder, B. Link, and S. Vernon, "AC-3: Flexible preceptual coding for audio transmission and storage," in *Convention of the AES*, (Amsterdam), February 1994.

[8] J. Kovačević, D. LeGall, and M. Vetterli, "Image coding with windowed modulated filter banks," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, (Glasgow, UK), pp. 1949–1952, May 1989.

[9] A. Johnson, J. Princen, and M. Chan, "Frequency scalable video coding using the MDCT," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, (Adelaide, Australia), 1994.

[10] M. Vetterli and D. LeGall, "Perfect reconstruction FIR filter banks: Some properties and factorizations," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. 37, pp. 1057–1071, July 1989.

[11] A. Tabatabai, "Some results on two-dimensional pseudoquadrature mirror filters," *IEEE Trans. Circ. and Syst.*, vol. 34, pp. 988–992, August 1987.

[12] R. Bernardini and J. Kovačević, "Local cosine bases in two dimensions," *Multidim. Syst. and Sign. Proc., special issue on Wavelets and Multiresolution Signal Processing*, December 1995. Invited paper.