論文 / 著書情報
Article / Book Information

題目(和文)	重複変換および画像処理におけるその応用
Title(English)	Lapped transforms and their applications in image processing
著者(和文)	田中聡久
Author(English)	
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第5055号, 授与年月日:2002年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第5055号, Conferred date:2002/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Lapped Transforms and Their Applications in Image Processing

Toshihisa TANAKA

Department of International Development Engineering
Graduate School of Science and Engineering
Tokyo Institute of Technology

February 2002

Abstract

This dissertation deals with the theory and design of various classes of lapped transforms. We address problems of 1) the derivation of analytic solution of lapped transforms, 2) the design of adaptive lapped transforms and their image coding applications, and 3) the theory of a class of oversampled lapped transforms.

We introduce a biorthogonal lapped transform that consists of long and short basis functions (VLLBT). The design criterion is formulated as an approximation problem. In order to solve this, we provide the theory of the Karhunen-Loève transform in a subspace (SKLT), and we show that when the biorthogonal long basis functions of the VLLBT are given, the optimal short basis functions in the energy compaction sense are found. Therefore, the degree of freedom for the VLLBT is reduced to that for the long basis functions of the VLLBT.

By extending the VLLBT and the SKLT, we present the theory and design of two-dimensional adaptive lapped biorthogonal transforms for image coding. The proposed transform is an natural extension of the one-dimensional VLLBT. The adaptation is performed by making the short basis functions variable. Those short basis functions are derived by the SKLT. We show an orientation adaptive example, where each adaptive transform is characterized by the angle of edges in image blocks. Moreover, we illustrate image coding applications for several transforms proposed in this dissertation. Comparisons among the proposed transforms and the existing transforms are carried out. To be fair, we adopt the same techniques in the coding step. Specifically, transform coefficients are uniformly quantized and encoded with the run-length/Huffman tables used in the baseline JPEG. Through comparison, we illustrate benefit of the present transforms.

In this next step, we present a minimal lattice structure for a special class of N-channel oversampled linear-phase perfect reconstruction filter banks, which is called the generalized lapped pseudo-biorthogonal transform (GLPBT). The GLPBT is a generalization of a very large class of lapped transforms. Moreover, we provide an alternative lattice which includes a building block

suppressing noise added in the transform domain. We also show several design examples and simulations.

Finally, we state our contribution, and clarify open problems and directions for future research.

Acknowledgments

I wish to express my appreciation and gratitude to my advisor, Prof. Yukihiko YAMASHITA. His support and guidance have make the completion of this thesis possible. He has provided an environment conducive to learning and quality research. He has introduced me to the field of research on signal and image processing from a mathematical point of view.

I also thank Prof. Hidemitsu OGAWA. He has not directly advised me on my study. However, his attitude and philosophy toward research influenced me. His class on pseudo-biorthogonal bases inspired me to study oversampled filter banks. I am grateful to Prof. Itsuo KUMAZAWA. He gave me some pieces of useful advice for not only research also entering the doctoral course. I would like to express my gratitude to Prof. Akinori NISHIHARA. He has participated in the Wavelet seminar which I organized as an advisor since it started. He also gave me useful advice and nice support on IEEE student activities as the Branch Counselor during I am the Chair of the IEEE Student Branch at TITech. I wish to thank Prof. Isao YAMADA and Prof. Toshiyuki YOSHIDA. I enjoyed discussion on signal and image processing with them.

I would like to thank Dr. Takayuki NAGAI of the University of Electro-Communications for a helpful advice on the filter design. I also thank Dr. Ricardo De Queiroz of Xerox Corporation for giving me useful information about lapped transforms.

I am also grateful to Prof. Sung-Jea Ko of Korea University for allowing me to work on wavelet-based image coding during my stay in Korea University from September 1997 to August 1998 as a visiting graduate student. He and his students have greatly influenced my attitude to research.

I am grateful to my friend Dr. Masashi SUGIYAMA. His field of research is not exactly the same as mine, but we have sometimes exciting and useful discussions. I also would like to thank my colleagues supervised by Prof. Yamashita, Prof. Ogawa, and Prof. Kumazawa for having nice days together. Finally, I would like to acknowledge the support of my parents.

This work is partially supported in part by JSPS Grant-in-Aid for JSPS Fellows 1210283.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Image transformation is an well-used technique for image processing such as coding, restoration, recognition, and so on. When we transform a signal with large number of samples, we usually segment the signal into blocks with smaller number of samples. This block is dealt as a finite signal and processed independently. The block-based processing has advantages in the reductions of computational complexity and processing non-stational signals. However, the independent processing causes an artificial discontinuity between adjacent blocks. In image compression, for example, this discontinuity is called a *blocking artifact*, which appears in a decoded image at a low bit rate as illustrated in Fig. 1.1. Several studies on the reduction of blocking artifacts were carried out. For example, an overlapping and a filtering methods has been discussed in [1]. In the overlapping method, an extended block in which the boundary samples are overlapped is used. As a result, information which must be transmitted has redundancy. In the filtering method, a low-pass filter is applied only to the boundary pixels. Therefore, this method can be regarded as one of postprocessing techniques. A lot of excellent postprocessing methods for the reduction of blocking have been proposed [2, 3, 4, 5, 6, 7]. However, these postprocessing methods would lead to blurring at block boundaries.

In the context of reduction of the blocking effects, the lapped orthogonal transform (LOT) developed by Malvar *et al.* [8, 9, 10] made an great impact on the signal processing community. When the number of samples in a block is M, this new transform consists of M basis functions, which are of length $2M$. Hence, the transform generates M coefficients from $2M$ consecutive samples. The total number of transformed coefficients is the same as that of original samples.

Figure 1.1: Blocking artifacts

Moreover, the LOT can achieve perfect reconstruction. This is a desirable property for image coding, since there is no increase in the bit rate. The LOT was extended to more generalized forms, and these variations are referred to as *lapped transforms*. Their applications do not lie only in the field of image coding but also in that of image and signal processing, video processing, restoration, recognition, and adaptive filtering [11, 12, 13]. In this dissertation, several extensions and generalizations of the lapped transform are established for image coding and processing.

1.2 Brief History

The lapped transform was pioneered by Malvar and co-workers [8, 9]. The most elementary form of lapped transforms is the LOT, which has been generalized from various perspectives. We review in this section various types of generalization and improvements of lapped transforms from three viewpoints. We also review adaptive transforms for image coding and their relation to lapped transforms for further argument.

1.2.1 Generalization on Length of Basis Functions

The original LOT consists of M basis functions of length $2M$. The basis functions are orthogonal to each other, and the overlapping parts are also orthogonal. As pointed out by Vetterli and Le Gall [14], the LOT is equivalent to an M-channel maximally decimated linear phase (LP) paraunitary (PU) filter bank (FB), where all filters are of length $2M$. Queiroz *et al.* [15] developed a class of lapped orthogonal transforms where all basis functions have length KM, which is an integer multiple of M. This class of lapped transforms is called a *generalized lapped orthogonal transform* (GenLOT). In [15], a factorization and a lattice structure are also described. A similar lattice was proposed in [16]. These generalized versions of the LOT can be regarded as a subclass of M-channel maximally decimated LPPUFBs, and an investigation into complete and minimal M-channel LP PU lattice structures have been reported in [17].

All lapped transforms described above have basis functions of the same length which is a multiple of the number of channels. Although the overlapping basis functions can considerably reduce the blocking effect, they tend to spread the quantization error or the noise over adjacent blocks as shown in Fig. 1.2. To avoid this effect, variable-length (unequal-length) lapped transforms have been proposed [18, 19, 20]. In these lapped transforms (called *VLLOTs* [18] or *GULLOTs* [20]), longer basis functions can reduce the blocking, whereas the shorter basis functions can restrict the ringing within their supports. The VLLOT and the GULLOT are similar to wavelet transforms [12, 13] in that they have basis functions of variable lengths. However, since the VLLOT and the GULLOT are extensions of the GenLOT, they have efficient fast algorithms based on lattice structures. A history of generalization on the filter lengths is summarized in Table 1.1.

1.2.2 Biorthogonalization

Another generalization for the LOT is biorthogonalization. Chan replaced some orthogonal matrices in [8] with cascades of invertible block diagonal matrices [21]. Malvar suggested the lapped biorthogonal transform by introducing a $\sqrt{2}$ scaling of the first antisymmetric basis function of the DCT. However, these structures do not provide the general solution. Tran *et al*. have shown that a lapped biorthogonal transform can be implemented as an M-channel LP perfect reconstruction (PR) FB with minimal lattice structures [22]. This transform is referred to as the generalized lapped biorthogonal transform (GLBT).

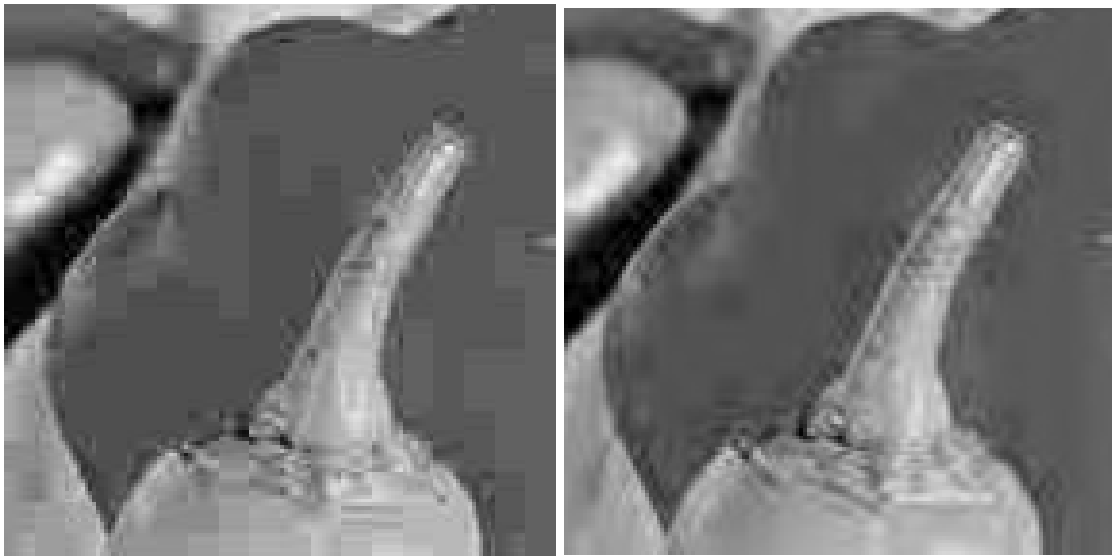


Figure 1.2: Difference of decoded images with the DCT (left) and the lapped transform (right): the decoded image with the lapped transform contains ringing artifacts around edges caused by long basis functions

Also, the VLLOT has been extended to a biorthogonal transform called the VLGLBT [23]. This can be an extension of the GLBT. Moreover, fast algorithms for the VLGLBT have been also developed [19].

1.2.3 Extension to Overcomplete Representation

Generalization for the number of basis functions or channels has been conducted [24]. In this generalization, the number of basis functions is greater than the number of samples in a block or a *decimation factor*. This lapped transform provides an *overcomplete* representation, and is related to redundant signal expansions [25, 26, 27, 28]. The oversampled FBs have some advantages such as their improved design freedom and noise immunity [29, 30]. However, these oversampled systems involve increased computational complexity. Therefore, oversampled DFT FBs [31, 12, 32, 33, 34] and oversampled cosine-modulated FBs [29, 35] have been developed for a fast and efficient implementation by a factorization. These FBs belong to a category of modulated FBs. For application in image processing, the linear-phase property is very significant. From this point of view, recently, a complete factorization of oversampled paraunitary (*pseudo-*

Table 1.1: Generalizations of the LOT in terms of the filter length and the biorthogonalization

Filter Length	Orthogonal	Biorthogonal
$2M$	Malvar *et al.* [8, 9] (LOT)	Chan [21] (GLT), Malvar [36] (LBT)
KM	Soman *et al.* [17], de Queiroz *et al.* [15] (GenLOT)	Tran *et al.* [22] (GLBT)
$KM+\beta$	Tran *et al.* [37]	

Table 1.2: Generalizations of the variable-length LOT in terms of the adaptivity and the biorthogonalization

Adaptivity	Orthogonal	Biorthogonal
Nonadaptive	Tran *et al.* [18] (VLLOT), Nagai and Ikehara [38] (ULLOT), Nagai *et al.* [20] (GULLOT)	Tran *et al.* [23] (VLGLBT), **Chapter 3**
Adaptive	**Chapter 4**	**Chapter 4**

orthogonal) FBs [1] yielding LP filters has been proposed [24].

1.2.4 Adaptation

Image coding using multiple transforms is known as one of efficient coding methods. This method is based on the notion that different blocks depend on different statistic models, respectively. For example, the JPEG [39], the MPEG-x, and the H.26x [40] use the discrete cosine transform (DCT) [41, 42] that is derived from one model, the first-order Markov model. How-

Table 1.3: A summary of previous works for a minimal complete lattice structure of an N-channel LPPRFB with decimation factor M: $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ denote the polyphase matrices of the analysis bank and the synthesis bank, respectively.

	$\boldsymbol{R}(z)=\boldsymbol{E}^T(z^{-1})$	$\boldsymbol{R}(z)\neq\boldsymbol{E}^T(z^{-1})$
$M=N$	Paraunitary or Orthogonal [17]	Biorthogonal [22]
$M<N$	Paraunitary or Pseudo-orthogonal [24]	Pseudo-biorthogonal (**Chapter 6**)

[1] In the paper [24], the authors termed the proposed FBs *paraunitary*. The polyphase matrix for the synthesis FB does not give such an inverse as defined for a full-rank square matrix but a left-inverse, which will be defined later. In order to distinguish those FBs from maximally decimated PU FBs, we will use the terminology *pseudo-orthogonal* for such oversampled FBs.

ever, the use of the DCT for fast varying blocks as well as plane blocks may be unreasonable. Signals in those plain and texture regions must be different stochastic processes. It is therefore quite natural that we use a transform derived from a suitable model for an input block. A coding method based on this strategy is called *adaptive transform* or *multiple transform*. The adaptive transform enables us to obtain a decoded image where edges and lines are well preserved at low bit rates. Therefore, it may be relevant to use a transform depending on each block. This has motivated several studies on transform coding. Bjøntegaard, who firstly suggested the use of multiple transforms for image coding, introduced *a priori* classes which are characterized by directionality such as edges [43]. Furthermore, Tanaka and Yamashita improved his adaptive transform using the so-called vector-embedded Karhunen-Loève transform (VEKLT) [44]. Instead of the use of *a priori* classes, some self-organizing methods with training of input signals has been proposed [45, 46, 47]. Helsingius *et al.* discussed benefit of the use of multiple transforms with image coding by empirical comparisons.

In the field of lapped transforms, de Queiroz *et al.* proposed the time-varying lapped transform [48], in which the adaptation is accomplished by replacing orthogonal matrices in the lattice structure to time-varying ones. Klausutis and Madisetti made length of basis functions adaptive [49].

1.3 Goal of This Dissertation

As seen in the previous section, lapped transforms have been generalized to various forms. However, there exist several unsolved problems. In this dissertation, the following problems will be addressed:

1.3.1 Theoretical Parameter Reduction in Design of Variable-Length Lapped Transforms

As discovered in [8, 17, 15, 22], all coefficients of lapped orthogonal transforms can be characterized by Givens rotation angles [50] in lattice structures. (In the biorthogonal case, additional positive values are required [22].) All existing design methods for design of those lapped transforms find their parameters by minimizing some cost function with numerical search. This iterative optimization requires much computation time to reach the solution. In most of the cases,

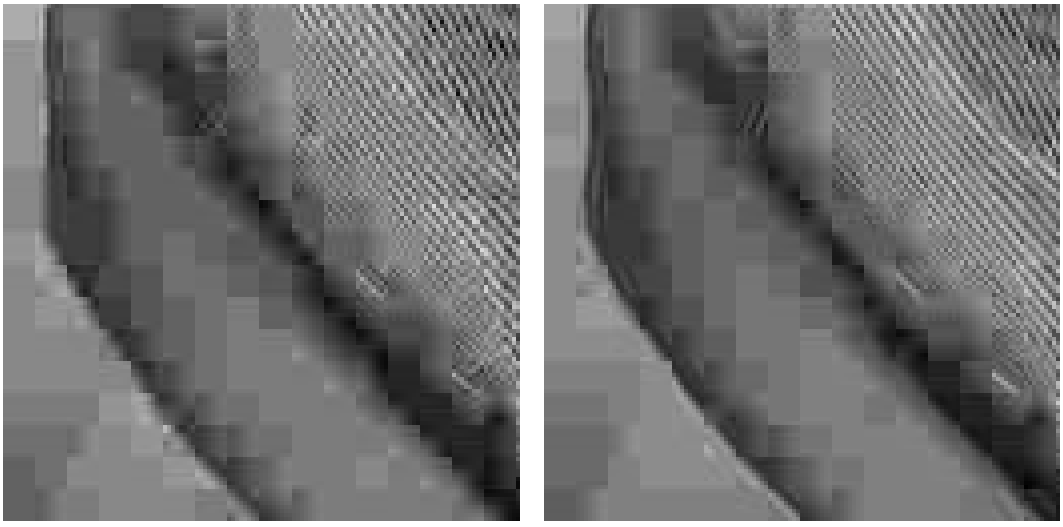
moreover, cost functions are not convex and hence there is no guarantee of obtaining a global minimum. Therefore, an explicit solution minimizing a cost function is desirable. Furthermore, the initial building block in the lattice structure is assumed to be the DCT to reduce free parameters. Indeed, this assumption is reported to be effective in design. However, there is no theoretical guarantee for the use of the DCT. The problem in this dissertation is to find an alternative method for parameter reduction in design of variable-length lapped transforms in terms of signal approximation. For the purpose, an extension of the Karhunen-Loève transform which is called the *subspace Karhunen-Loève transform* (SKLT) is proposed. The SKLT is formulated as the KLT in a given subspace. The formulation and the solution are presented. By using the SKLT we show that the number of free parameters can be reduced in the sense of the mean square error.

1.3.2 Formulation and Design of Adaptive Lapped Transforms

Figure 1.3 shows the effect of the orientation adaptation in block-based image coding [44]. However, as seen in Table 1.2, despite of the effectiveness of the adaptation, this technique has not been applied to the design of lapped transforms due to their strict constraints for perfect reconstruction. We solve this problem in this dissertation by introducing a class of lapped transforms consisting of overlapping and non-overlapping basis functions, which is indeed a subclass of variable-length lapped transforms. The adaptation is applied to the non-overlapping basis functions. The relationship between the conventional and the proposed transforms is listed in Table 1.2.

1.3.3 Parameterization of Overcomplete Lapped Transforms

Lapped transform produces the same number of transform coefficients as that of input samples in terms of polyphase matrices. If this constraint is softened, those lapped transforms are regarded as *overcomplete* systems. In the field of filter banks [11, 12, 13], they are interpreted as *oversampled* perfect reconstruction (PR) filter banks (FBs). The block size are called a *decimation factor*, and the number of output samples are called a *channel*. If we limit our discussion to the linear-phase (LP) case, which is indeed essential condition for image processing, factorizations with lattice structures are of particular interest. We address this problem for oversampled LP



(a) A decoded image with a block-based transform (the DCT)

(b) A decoded image with an orientation adapted transform (the VEKLT [44])

Figure 1.3: The effect of the orientation adaptation

PR FBs with filters [2] of length equal to an integer multiple of the block size or the decimation factor. Table 1.3 shows a position of this work in the field of oversampled PR FBs. We provide more generalized lattice structure which can cover a very wide range of oversampled LPPRFBs. Moreover, we address the problem to find a lattice structure for the case where noise is added to the transformed coefficients.

1.4 Organization of This Dissertation

This dissertation is organized as follows. An image of the relationship among chapters is depicted in Fig. 1.4.

In Chapter 2, we define notations and review fundamental theories which are necessary to understand this dissertation.

In Chapter 3, we present a new framework of design for a biorthogonal lapped transform that consists of overlapping (long) and non-overlapping (short) basis functions (VLLBT). We

[2] In the overcomplete case, we can no longer use the term "basis function," because those functions are linearly dependent. Those functions (called filters usually) yields a *frame* [26, 28].

formulate the VLLBT by extending conventional lapped transforms. Then, we provide a theory of the Karhunen-Loève transform in a subspace (SKLT). Using the theory of the SKLT, we show that when the biorthogonal long basis functions of the VLLBT are given, the optimal short basis functions in the energy compaction sense are derived by solving an eigenvalue problem without iterative searching techniques. This implies that the number of free parameter for the VLLBT to be determined is reduced to that for the long basis functions of the VLLBT, although in general, biorthogonalization leads to a large increase in the degree of freedom. We also provide design examples. The resulting VLLBT attains high coding gain comparing to other lapped transforms. These results suggest that the proposed VLLBT with the optimal short functions is a promising technique in the field of image coding.

In Chapter 4, we present the theory and design of an adaptive lapped biorthogonal transform for image coding. The proposed transform consists of basis functions overlapping across adjacent blocks and non-overlapping basis functions. The overlapping basis functions have samples whose number is an integer multiple of the traditional block size. We introduce two types of this transform: Type-E and Type-O. The former type has overlapping basis functions whose length is an even multiple of the traditional block size. In the latter type, overlapping basis functions' length is an odd multiple of the block size. Type-E requires special care for the image boundary to avoid the border distortion. In Type-O, on the other hand, basis functions' centers of symmetry are aligned. Therefore, we can use the symmetric extension method at image boundaries when we transform an input image. We next propose an adaptive lapped transform by extending the 1D VLLBT to a 2D transform. The adaptation is applied to non-overlapping basis functions, which are 2D and non-separable. It is shown that their derivation is also enabled by the SKLT. We further show an orientation adaptive example, where each adaptive transform is characterized by the angle of edges in image blocks.

In Chapter 5, we illustrate image coding applications for several transforms developed in Chapters 3 and 4. Comparisons among the proposed transforms and the existing transforms are carried out. To be fair, we adopt the same techniques in the coding step. Specifically, transform coefficients are uniformly quantized and encoded with the run-length/Huffman tables used in the baseline JPEG [39]. Through comparison, we illustrate benefit of the proposed transforms.

In Chapter 6, we investigate a lattice structure for a special class of N-channel oversampled linear-phase perfect reconstruction filter banks with a decimation factor M smaller than N. We

deal with systems in which all analysis and synthesis filters have the same FIR length and share the same center of symmetry. We provide the minimal lattice factorization of a polyphase matrix of a particular class of these oversampled filter banks. All filter coefficients are parameterized by rotation angles and positive values. The resulting lattice structure is able to provide fast implementation and allows us to determine the filter coefficients by solving an unconstrained optimization problem. We consider next the case where we give the GLPBT lattice structure with specific parameters and we *a priori* know the correlation matrix of noise which is added in the transform domain. In this case, we provide an alternative lattice structure which suppress the noise. We show that the proposed systems with the lattice structure cover a wide range of linear-phase perfect reconstruction filter banks. We also show several design examples and their properties.

In Chapter 7, we conclude this work. We summarize contributions of this dissertation, and clarify open problems and directions for future research.

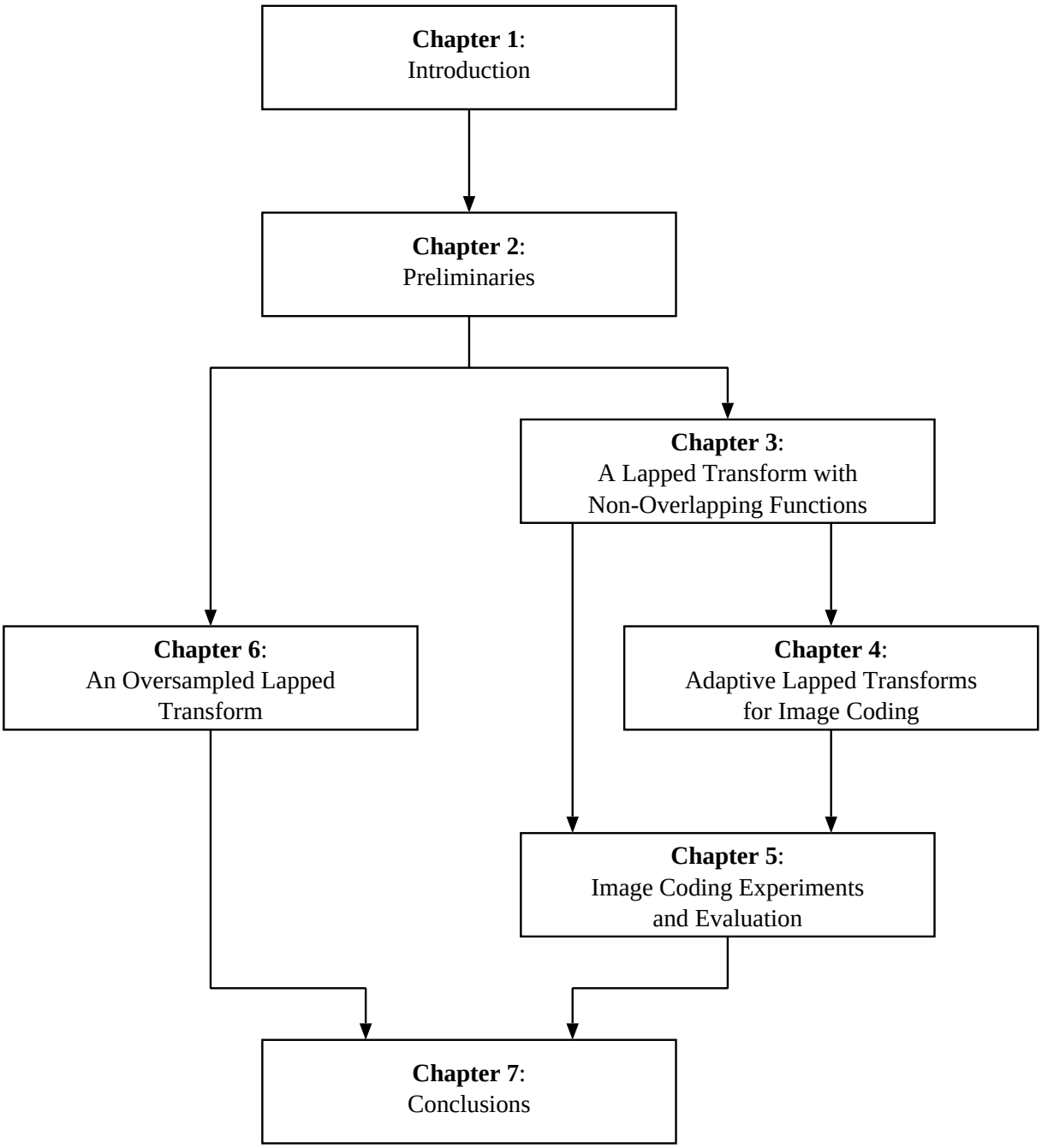


Figure 1.4: The relation among chapters in this dissertation

Chapter 2

Preliminaries

This chapter defines notation used throughout this dissertation, and reviews fundamental theories to analysis lapped transforms and to understand image coding principles.

2.1 Notation

The conventions listed in Table 2.1 are adopted in terms of notation. Bold-faced characters are used to denote vectors and matrices. We sometimes omit the subscript of these matrices if the size is obvious.

As with the Euclidean norm of a vector, we introduce the *Frobenius norm* or the *Schmidt norm* of a matrix as the following:

Definition 1 (The Frobenius norm [50]) *Let $\boldsymbol{A}$ be a matrix of size $M \times N$. The Frobenius norm $||\boldsymbol{A}||_F$ of $\boldsymbol{A}$ is defined as*

$$||\boldsymbol{A}||_F = \sqrt{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |a_{mn}|^2}, \tag{2.1}$$

where a_{mn} denotes the (m, n) entry of $\boldsymbol{A}$.

Then, the following important relation holds:

$$||\boldsymbol{A}||_F^2 = \mathrm{tr}[\boldsymbol{A}\boldsymbol{A}^T]. \tag{2.2}$$

We introduce the notion of a *left-inverse* related to overcomplete representation.

$\mathbb{R}^N$	N-dimensional Euclidean space
$\langle \boldsymbol{f}, \boldsymbol{g} \rangle$	inner product of two vectors $\boldsymbol{f}$ and $\boldsymbol{g}$
$\|\boldsymbol{f}\|$	Euclidean norm of $\boldsymbol{f}$
$f(i)$	i-th component of $\boldsymbol{f}$
$\boldsymbol{I}_n$	$n \times n$ identity matrix
$\boldsymbol{J}_n$	$n \times n$ reversal matrix
$\boldsymbol{0}_n$	$n \times n$ null matrix
$\boldsymbol{0}_{m \times n}$	$m \times n$ null matrix
$\boldsymbol{A}^T$	transposition of a matrix $\boldsymbol{A}$
$[\boldsymbol{A}]_{i,j}$	(i, j) element of $\boldsymbol{A}$
$\mathrm{tr}[\boldsymbol{A}]$	trace of $\boldsymbol{A}$
$R(\boldsymbol{A})$	range of $\boldsymbol{A}$
$N(\boldsymbol{A})$	null space of $\boldsymbol{A}$
$\mathrm{rank}(\boldsymbol{A})$	rank of $\boldsymbol{A}$
$\dim(\cdot)$	dimension of a linear space
$H(z)$	z-transform of a vector or a sequence $\boldsymbol{h}$
$\boldsymbol{H}(z)$	polyphase matrix
$\|\boldsymbol{A}(z)\|$	determinant of $\boldsymbol{A}(z)$
$\deg[\boldsymbol{A}(z)]$	degree of $\boldsymbol{A}(z)$
$(\downarrow M)$	downsampling operator with M
$(\uparrow M)$	upsampling operator with M

Table 2.1: Notations used throughout this dissertation

Definition 2 (Left-inverses) *A matrix $\boldsymbol{A}$ is called left-invertible if there exists a matrix $\boldsymbol{X}$ such that*

$$\boldsymbol{X}\boldsymbol{A} = \boldsymbol{I}. \tag{2.3}$$

The matrix $\boldsymbol{X}$ is denoted by $\boldsymbol{A}^-$ and called a left-inverse of $\boldsymbol{A}$.

Keep in mind that given a matrix $\boldsymbol{A}$, its left-inverse $\boldsymbol{A}^-$ is not uniquely determined. For any $\boldsymbol{A}^-$, moreover, it does not hold that $\boldsymbol{A}\boldsymbol{A}^- = \boldsymbol{I}$ in general. A left-inverse is included in a special class of pseudo (generalized) inverses [51]. Let $\boldsymbol{A}$ be a left-invertible matrix of size $n \times m$. Then, n must be greater than or equal to m, i.e. $n \geq m$, and $\mathrm{rank}(\boldsymbol{A}) = m$.

2.2 Orthogonal Transforms

2.2.1 Karhunen-Loève Transform (KLT)

We start with the Karhunen-Loève transform that packs most energy into the first k coefficients among all orthogonal transforms.

We can remove correlations between pixels using an orthogonal linear transform called the *Karhunen-Loève transform* (KLT), also known as the *Hotelling transform* [52, 53, 54]. Suppose we create N-dimensional vectors from a given image by taking blocks of N pixels. Let $\boldsymbol{f} = [f(0), \ldots, f(N-1)]^T$ be a vector of the original data samples in $\mathbb{R}^N$. The correlation matrix $\boldsymbol{R}_{\boldsymbol{ff}}$ with respect to $\boldsymbol{f}$ is given by

$$\boldsymbol{R}_{\boldsymbol{ff}} = E_{\boldsymbol{f}}[\boldsymbol{f}\boldsymbol{f}^T], \tag{2.4}$$

where we assume rank($\boldsymbol{R}$) = N. The matrix $\boldsymbol{R}_{\boldsymbol{ff}}$ is real and symmetric, hence there exist eigenvalues [55] $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{N-1} > 0$ and corresponding eigenvectors $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{N-1}$ such that $\{\boldsymbol{u}_i\}_{i=0}^{N-1}$ is an orthonormal basis of $\mathbb{R}^N$.

Definition 3 (The KLT [56]) *The Karhunen-Loève (KL) transform is defined as a matrix $\boldsymbol{X}$ which minimizes*

$$J_{\mathrm{KLT}} = E_{\boldsymbol{f}}\|\boldsymbol{f} - \boldsymbol{X}\boldsymbol{f}\|^2 \tag{2.5}$$

under the condition that the rank of $\boldsymbol{X}$ is equal to r, where $r \leq N$.

Theorem 1 *The KLT $\boldsymbol{X}$ of rank r is given by*

$$\boldsymbol{X} = \sum_{i=0}^{r-1} \boldsymbol{u}_i \boldsymbol{u}_i^T. \tag{2.6}$$

Proof: Clearly, rank($\boldsymbol{X}$) = dim($R(\boldsymbol{X})$) = r. Therefore, we seek the optimal approximation in the subspace $R(\boldsymbol{X})$. From the projection theorem [57], $\boldsymbol{X}$ should be an orthogonal projection matrix. Hence, the criterion J_{KLT} yields

$$\begin{aligned} J_{\mathrm{KLT}} &= E_{\boldsymbol{f}}\|\boldsymbol{f} - \boldsymbol{X}\boldsymbol{f}\|^2, \\ &= E_{\boldsymbol{f}}\|\boldsymbol{f}\|^2 - E_{\boldsymbol{f}}\|\boldsymbol{X}\boldsymbol{f}\|^2. \end{aligned} \tag{2.7}$$

Since $E_{\boldsymbol{f}}\|\boldsymbol{f}\|^2$ is constant, the matrix $\boldsymbol{X}$ maximizing $E_{\boldsymbol{f}}\|\boldsymbol{X}\boldsymbol{f}\|^2$ gives the solution. We have

$$\begin{aligned} E_{\boldsymbol{f}}\|\boldsymbol{X}\boldsymbol{f}\|^2 &= E_{\boldsymbol{f}}\left[\mathrm{tr}[(\boldsymbol{X}\boldsymbol{f})(\boldsymbol{X}\boldsymbol{f})^T]\right], \\ &= \mathrm{tr}[\boldsymbol{X}\boldsymbol{R}\boldsymbol{X}]. \end{aligned} \tag{2.8}$$

With the eigenvalue decomposition of $\boldsymbol{R}$, the above equation can be written

$$\begin{aligned} E_{\boldsymbol{f}}\|\boldsymbol{X}\boldsymbol{f}\|^2 &= \mathrm{tr}\left[\boldsymbol{X}\left(\sum_{i=0}^{N-1}\lambda_i\boldsymbol{u}_i\boldsymbol{u}_i^T\right)\boldsymbol{X}\right], \\ &= \sum_{i=0}^{N-1}\lambda_i\|\boldsymbol{X}\boldsymbol{u}_i\|^2. \end{aligned} \tag{2.9}$$

Because $\boldsymbol{X}$ is the orthogonal projection matrix such that $\mathrm{rank}(\boldsymbol{X}) = M$, the following is held:

$$0 \le \|\boldsymbol{X}\boldsymbol{u}_i\|^2 \le 1, \tag{2.10}$$

$$\sum_{i=0}^{N-1}\|\boldsymbol{X}\boldsymbol{u}_i\|^2 = M. \tag{2.11}$$

If we set

$$\|\boldsymbol{X}\boldsymbol{u}_i\| = \begin{cases} 1 & i \le r-1 \\ 0 & r \le i \le N-1 \end{cases} \tag{2.12}$$

$E_{\boldsymbol{f}}\|\boldsymbol{X}\boldsymbol{f}\|^2$ is maximized. □

Various proofs of the KLT providing the best approximation in the mean square sense have been shown. However, Ogawa has pointed out that those proofs are incomplete, and shown the exact proof [58]. Generally, the KLT is considered as an impractical transform because it depends on input signals. Therefore, it is usual to use an appropriate correlation matrix such as a first-order Markov model, that is,

$$[\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}]_{ij} = \rho^{|i-j|}, \tag{2.13}$$

where ρ is the correlation coefficient between adjacent pixels. The matrix $\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}$ leads to the fixed suboptimal KLT. For typical natural images, each pixel is strongly correlated ($0.9 < \rho < 1$). It has been shown that the KLT leads to the discrete cosine transform (DCT) as $\rho \to 1$ [41, 42, 59]. This fact will be reviewed in the next section. Moreover, an approach to the approximation of KLT, that does not restrict itself to a specific class of stationary process (such as the Markov-1 family), has been presented in [60].

2.2.2 Discrete Cosine Transform (DCT)

In 1974, Ahmed *et al.* has proposed the discrete cosine transform (DCT) [41].

Definition 4 (The DCT [41]) *The transform kernel of the discrete cosine transform (DCT) is defined as the form*

$$u_i(k) = \begin{cases} \dfrac{1}{\sqrt{N}} & i = 0 \qquad k = 0, \ldots, N-1 \\ \sqrt{\dfrac{2}{N}} \cos \dfrac{\pi(2k+1)i}{2N} & i = 1, \ldots, N-1 \quad k = 0, \ldots, N-1. \end{cases} \tag{2.14}$$

This definition of the DCT is known as type-II (DCT-II), which is the most widely used. Various types of DCTs have been developed [42, 40].

The DCT has several advantages over the DFT. First, unlike the DFT, the DCT is a real-valued transform that generates real coefficients from real-valued data. Second, the ability of the DCT and the DFT to pack signal energy into a small number of coefficients is a function of the global smoothness of these signals. In the class of transforms with a known fast computational algorithm, the DCT has a superior energy compaction property [41, 61, 42, 40].

The DCT has a close relationship with the KLT of which the correlation matrix is under the Markov model. We will show that the DCT may be derived from the KLT in the limiting case as the adjacent element correlation tends to unity [59].

The correlation matrix of the first-order Markov process $\boldsymbol{R}$ is given by

$$(\boldsymbol{R})_{i,k} = \rho^{|i-k|} \quad i, k = 0, 1, \ldots, N-1, \tag{2.15}$$

where ρ is the intersample correlation coefficient. Ray and Driver [62] have given the following solution

$$u_m(n) = \left[\frac{2}{N+\lambda_m}\right]^{1/2} \sin\left\{\omega_m\left[(n+1) - \frac{N+1}{2}\right] + (m+1)\frac{\pi}{2}\right\} \tag{2.16}$$
$$m, n = 0, 1, \ldots, N-1,$$

where λ_m's are N eigenvectors given as

$$\lambda_m = \frac{(1-\rho^2)}{1 - 2\cos\omega_m + \rho^2}. \tag{2.17}$$

ω_m is the real positive roots of the following

$$\tan(N\omega) = -\frac{(1-\rho^2)\sin\omega}{\cos\omega - 2\rho + \rho^2\cos\omega}. \tag{2.18}$$

In (2.16), (2.17), and (2.18), setting $\rho \to 1$, we have $\tan(N\omega) = 0$. Then,

$$\omega_k = \frac{k\pi}{N}, \quad \text{for } k = 0, 1, \ldots, N-1. \tag{2.19}$$

The eigenvalues are $\lambda_m = 0$ when $\omega_m \neq 0$. We have

$$\mathrm{tr}[\boldsymbol{A}] = \sum_{m=0}^{N-1} \lambda_m. \tag{2.20}$$

Since $[\boldsymbol{R}_{\boldsymbol{ff}}]_{m,m} = 1$, we have $\lambda_0 = N$. Substituting these relations in (2.16), we obtain the DCT-II as follows.

$$\begin{aligned} u_0(n) &= \frac{1}{\sqrt{N}}, \\ u_m(n) &= \sqrt{\left(\frac{2}{N}\right)} \sin\left[m\left(n+\frac{1}{2}\right)\frac{\pi}{N} + \frac{\pi}{2}\right] \\ &= \sqrt{\left(\frac{2}{N}\right)} \cos\left[m\left(n+\frac{1}{2}\right)\frac{\pi}{N}\right]. \end{aligned} \tag{2.21}$$

2.3 Filter Banks and Subband Transforms

2.3.1 Down/Up-Sampling

Let $h = \{h(i)\}$ be a sequence of real or complex values. The number of elements of h can be either finite or infinite. Consider sampling rate changes in the discrete-time domain. *Multirate signal processing* deals with discrete-time sequences taken at different rates, and has been investigated well [11]. We here review this topic briefly. Further details appear in [11].

Let $x(n)$ be an original sequence and let M_d be an integer. When a sequence $y(n)$ is given by

$$y(n) = x(nM_d), \tag{2.22}$$

we call this processing *downsampling* or *subsampling* a sequence $x(x)$ by M_d. In the z-transform domain, downsampling is written as

$$Y(z) = \frac{1}{M_d} \sum_{k=0}^{M_d-1} X(W_{M_d}^k z^{1/M_d}), \tag{2.23}$$

where $W_{M_d} = \exp(-j2\pi/M_d)$.

The converse of downsampling is *upsampling* by an integer M_u. An upsampled sequence is obtained by inserting $M_u - 1$ zeros between consecutive samples of the input sequence. Specifically, the upsampled sequence $y(n)$ is written as

$$y(n) = \begin{cases} x(n/M_u) & n = kM_u, k \in \mathcal{Z} \\ 0 & \text{otherwise.} \end{cases} \tag{2.24}$$

The above equation is written in z-transform domain as

$$Y(z) = X(z^{M_u}). \tag{2.25}$$

2.3.2 Perfect Reconstruction Filter Banks

A *filter bank* (FB) is a set of filters, usually associated with *downsamplers* or *decimators*. An *N-channel filter bank with the decimation factor M* is a system described in Fig. 2.1. FBs have their applications in areas of signal processing such as speech and image compression, digital watermarking, denoising, feature extraction, and so on.

If $M = N$ then it is called a *maximally decimated* or a *critically sampled* FB. We consider here only the case where the factor M is an integer. This system consists of two major parts: *analysis* and *synthesis* parts. Each channel in the analysis part has a analysis filter $H_n(z)$ followed by the downsampler ($\downarrow M$). In contrast, each channel in the synthesis part has the upsampler ($\uparrow M$) followed by a synthesis filter $F_i(z)$. To decide the filters, the condition for perfect reconstruction (PR) $x(n) = \hat{x}(n - l)$ are mainly imposed. There are several excellent references on maximally decimated PR FBs [11, 12, 13].

Transforms implemented by FBs as shown in Fig. 2.1 are called *subband transforms*. The subband transform can be interpreted as a generalization of block-based transforms. Figure 2.2 shows matrices for a block-based transform and for a typical filter bank.

2.3.3 Polyphase Representation

Polyphase representation is useful to analyze FBs since one can a system by a vector-matrix notation. The polyphase decomposition of the analysis filters $H_n(z)$ is described as

$$H_n(z) = \sum_{m=0}^{M-1} e_{n,m}(z^M) z^{-(M-1-m)}, \tag{2.26}$$

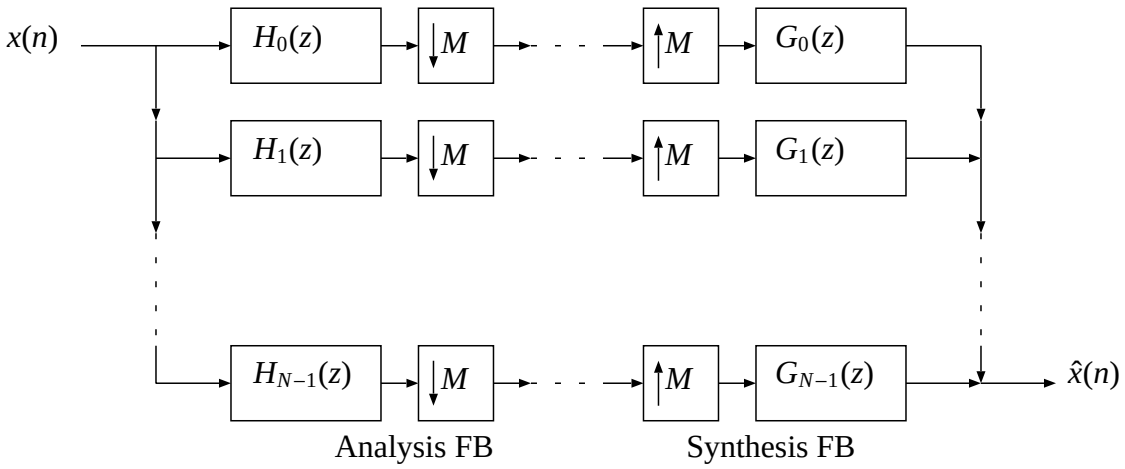


Figure 2.1: An N-channel filter bank with downsampling by the integer factor M

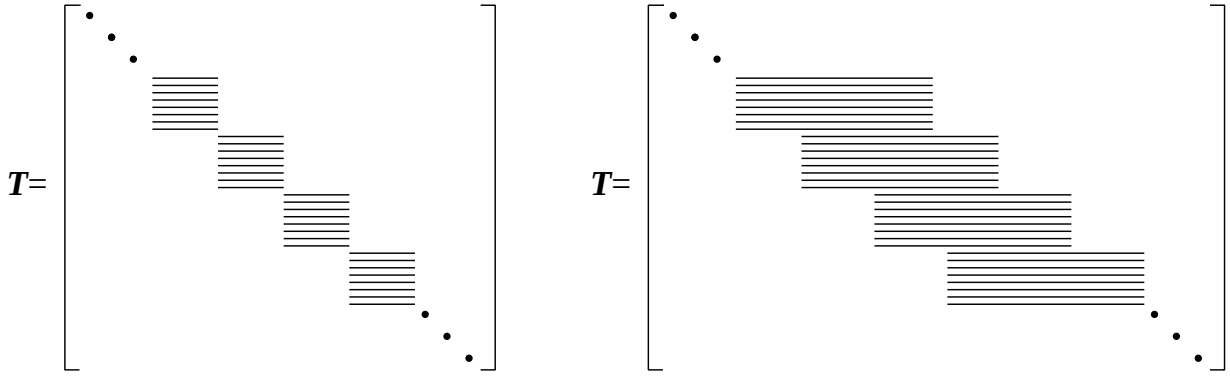


Figure 2.2: Matrices for blockwise transforms and filter banks

with

$$e_{n,m}(z) = \sum_{k=0}^{K-1} h_n(Mk + M - 1 - m)z^{-k}, \tag{2.27}$$

which is Type II polyphase [11]. The $N \times M$ analysis polyphase matrix $\boldsymbol{E}(z)$ is defined as $[\boldsymbol{E}(z)]_{n,m} = e_{n,m}(z)$. Similarly, the synthesis filters $G_n(z)$ can be decomposed as

$$G_n(z) = \sum_{m=0}^{M-1} r_{m,n}(z^M)z^{-m}, \tag{2.28}$$

with

$$r_{m,n}(z) = \sum_{k=0}^{K-1} g_n(Mk + m)z^{-k}, \tag{2.29}$$

which is Type I polyphase. The $M \times N$ synthesis polyphase matrix $\boldsymbol{R}(z)$ is defined as $[\boldsymbol{R}(z)]_{m,n} = r_{m,n}(z)$. The polyphase representation will be used for the analysis of oversampled FBs in Chapter 6.

2.3.4 Lapped Transforms

Lapped Orthogonal Transform

Lapped orthogonal transforms (LOTs), which are developed by Malvar *et al.* [9, 8, 10], are a subclass of M-channel LP PU FBs. The LOT has basis functions of the same length $2M$. The basis functions should be orthogonal, and the overlapping part of each basis function should also be orthogonal.

Let $\boldsymbol{E}_0$ and $\boldsymbol{E}_1$ be matrices of size $M \times M$. Then, the matrix defined as

$$\boldsymbol{E} = [\boldsymbol{E}_1 \ \boldsymbol{E}_0] \tag{2.30}$$

is of size $M \times 2M$. We introduce a transform matrix which is double block-diagonal as follows:

$$\boldsymbol{T}_a = \begin{bmatrix} & \vdots & \vdots & \vdots & \vdots & \\ \cdots & \boldsymbol{E}_1 & \boldsymbol{E}_0 & \boldsymbol{0} & \boldsymbol{0} & \cdots \\ \cdots & \boldsymbol{0} & \boldsymbol{E}_1 & \boldsymbol{E}_0 & \boldsymbol{0} & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}. \tag{2.31}$$

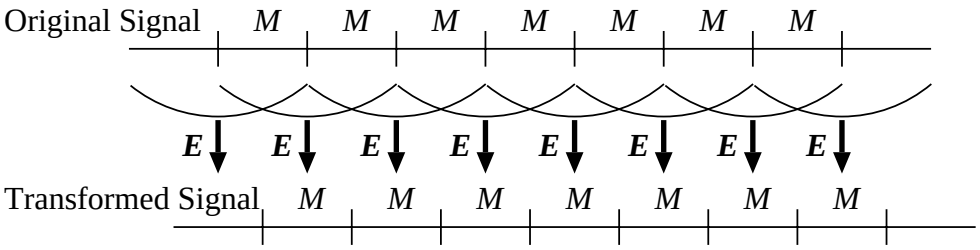


Figure 2.3: The lapped orthogonal transform

Consecutive two blocks are transformed and M transform coefficients are obtained as shown in Figs. 2.3 and 2.4. If the matrix $\boldsymbol{T}$ is orthogonal, this is called the *lapped orthogonal transform* (LOT). The orthogonal constraint $\boldsymbol{T}_a^T \boldsymbol{T}_a = \boldsymbol{I}$ yields that

$$\boldsymbol{E}_1^T \boldsymbol{E}_1 + \boldsymbol{E}_0^T \boldsymbol{E}_0 = \boldsymbol{I}_M, \tag{2.32}$$

$$\boldsymbol{E}_1 \boldsymbol{E}_0^T = \boldsymbol{0}_M, \ \boldsymbol{E}_1^T \boldsymbol{E}_0 = \boldsymbol{0}_M. \tag{2.33}$$

The second equation (2.33) is called *orthogonality of tails*. Let $\boldsymbol{f}_i$ be the i-th block with M samples of an input signal. Then, the transform vector $\tilde{\boldsymbol{f}}_i$ containing transform coefficients is given by

$$\tilde{\boldsymbol{f}}_i = \boldsymbol{E}_0 \boldsymbol{f}_{i-1} + \boldsymbol{E}_1 \boldsymbol{f}_i. \tag{2.34}$$

Contrarily, the reconstructed signal $\hat{\boldsymbol{f}}_i$ is obtained by

$$\hat{\boldsymbol{f}}_i = \boldsymbol{E}_0^T \tilde{\boldsymbol{f}}_i + \boldsymbol{E}_1^T \tilde{\boldsymbol{f}}_{i+1}. \tag{2.35}$$

It is clear that the LOT conditions (2.32) and (2.33) yields that $\hat{\boldsymbol{f}}_i = \boldsymbol{f}_i$. The LOT reduces the blocking effect because of its extended basis functions which overlap adjacent blocks [1] Malver imposed the linear phase property on the above conditions, and obtained lattice structures [8, 10].

The GenLOT

It is natural to establish the LOT which has length KM, where K is an positive integer and $K = 2$ for the LOT. Consider a matrix

$$\boldsymbol{E} = [\boldsymbol{E}_{K-1} \ \cdots \ \boldsymbol{E}_0]. \tag{2.36}$$

[1] Indeed, even though the LOT is used for image compression, the blocking artifacts appear slightly, since the LOTs basis functions do not smoothly decay to zero, generally. However, one can avoid this drawback by the use of the modulated lapped transform (MLT) for speech processing or the lapped biorthogonal transform (LBT) for image compression [36].

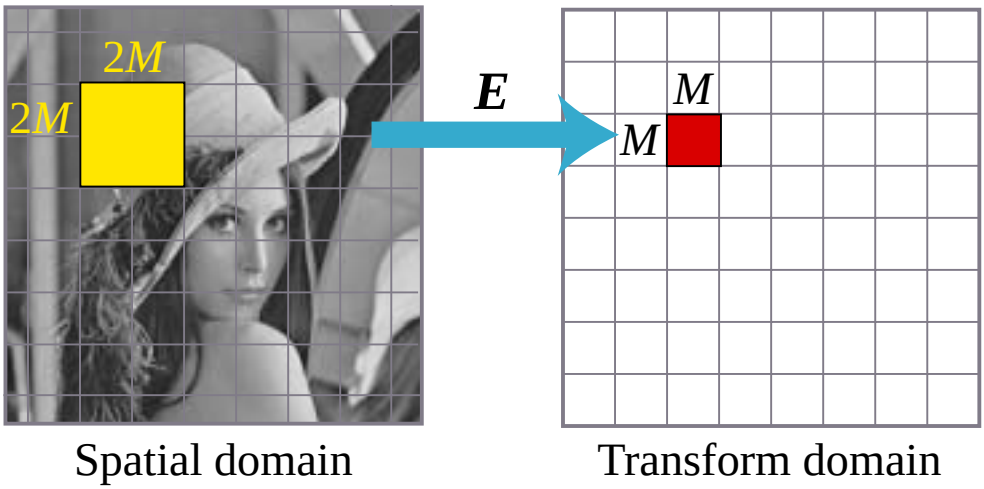


Figure 2.4: The two-dimensional lapped orthogonal transform

Then, this matrix is of size $M \times KM$. The generalized LOT $\boldsymbol{E}$ of size $LM \times M$ has to satisfy the following condition [13]:

$$\sum_{k=0}^{K-1} \boldsymbol{E}_k^T \boldsymbol{E}_k = \boldsymbol{I}_M, \tag{2.37}$$

$$\sum_{k=0}^{K-1-s} \boldsymbol{E}_k^T \boldsymbol{E}_{k+s} = \sum_{k=0}^{K-1-s} \boldsymbol{E}_{k+s}^T \boldsymbol{E}_k = \boldsymbol{0}_M, \quad \sum_{k=0}^{K-1-s} \boldsymbol{E}_k \boldsymbol{E}_{k+s}^T = \sum_{k=0}^{K-1-s} \boldsymbol{E}_{k+s} \boldsymbol{E}_k^T = \boldsymbol{0}_M, \tag{2.38}$$

where $s = 1, \ldots, K-1$. Equation (2.38) implies that the overlapping functions of neighboring blocks must also be orthogonal. In the field of linear-phase paraunitary filter banks, a design method with the lattice factorization has been presented by Queiroz *et al*, and the resulting LOT is called the GenLOT [15].

The linear-phase solution of the GenLOT is given as the polyphase matrix

$$\boldsymbol{E}(z) = \boldsymbol{K}_{K-1}(z)\boldsymbol{K}_{K-2}(z) \cdots \boldsymbol{K}_1(z)\boldsymbol{C}_M^{II}, \tag{2.39}$$

with

$$\boldsymbol{K}_i(z) = \frac{1}{2}\boldsymbol{\Phi}_i \boldsymbol{W} \boldsymbol{\Lambda}(z) \boldsymbol{W}, \tag{2.40}$$

where

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix}, \ \boldsymbol{\Lambda}(z) = \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & z^{-1}\boldsymbol{I}_{M/2} \end{bmatrix}, \ \boldsymbol{\Phi}_i = \begin{bmatrix} \boldsymbol{U}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_i \end{bmatrix}. \tag{2.41}$$

The matrix $\boldsymbol{C}_M^{II}$ denotes the M-point DCT-II matrix [42], the matrices $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ are $M/2 \times M/2$ orthogonal matrices.

2.4 Image Coding

One of the most major applications of image transforms is image coding. For transmission and storage of images, a enormous amount of memory is required. In the case of a square image with 512×512 pixels, for example, $512^2 \times 8 = 2,097,152$ bits are required! Video sequences are no longer stored and transmitted without compression.

Image coding consists of mapping digital images to sequences of binary digits. A desirable image coder is one that produces binary sequences whose lengths are on average much smaller than the original canonical representation of the image. In many imaging applications, it is not necessary to reproduce the image exactly. Therefore, one can perturb the image slightly to obtain a shorter representation. If this perturbation is much smaller than the blurring and noise introduced in the formation of the image in the first place, it is not necessary to use the more accurate representation. Such a coding strategy is called *lossy coding*. The goal of lossy coding is to reproduce a given image with minimum distortion, given some constraint on the total number of bits in the coded representation.

We can gain reductions in coded image size by discretizing images more coarsely, a process called *quantization*. By mapping visually indistinguishable images to the same code, we reduce the number of code words needed to encode images, at the price of a small amount of distortion. It is possible to quantize each pixel separately, a process known as *scalar quantization*. Quantizing a group of pixels together is known as *vector quantization* (VQ). VQ can theoretically capture the maximum compression. Although VQ is a very powerful theoretical paradigm, it can achieve optimality only asymptotically as its dimensions increase. But, the computational cost and delay grow exponentially with dimensionality, limiting the practicality of VQ. Due to these difficulties, most practical coding procedures have turned to *transform coding*. Transform coding is usually organized as three steps: linear transform, quantization, *entropy coding*. The success of transform coding depends on how well the basis functions of the transform represent the features of the signal.

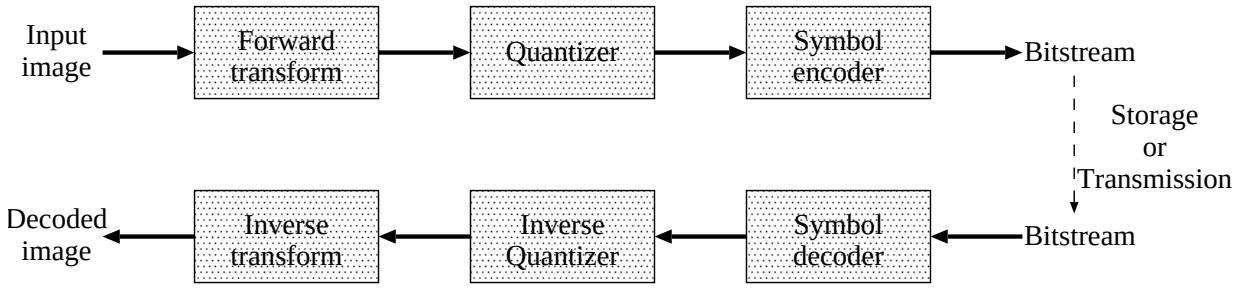


Figure 2.5: A transform coding system

2.4.1 Transform Coding System

Figure 2.5 shows a typical transform coding system. The decoder implements the inverse sequence of steps (with the exception of the quantization function) of the encoder, which performs three straightforward operations: transformation, quantization, and coding. Transform coding falls into two broad categories: *block transform coding* and *subband transform coding*. In the former case, an $N \times N$ input image is first subdivided into subimages of size $n \times n$, which are then transformed. In the latter case, the $N \times N$ image is transformed by a block Toeplitz matrix without being subdivided. Subband transforms are usually expressed as multi-rate systems consisting of three elements: filters, interpolators, and decimators. However, this dissertation sometimes deals with subband systems as linear transforms. Figure 2.2 illustrates matrices for blockwise transforms and filter banks. The goal of the transformation process is to decorrelate the pixels of each subimage, or to pack as much information as possible into the smallest number of transform coefficients.

Transform coding systems based on the Karhunen-Loève (KLT), discrete Fourier (DFT), Walsh-Hadamard (WHT), discrete cosine (DCT), subband (ST), wavelet (WT), and various other transforms have been constructed and/or studied extensively. The choice of a particular transform in a given application depends on the amount of reconstruction error that can be tolerated and the computational resources available. Compression is achieved during the quantization of the transformed coefficients (not during the transformation step).

2.5 Summary and References

We have reviewed several important notions and theories which are necessary to understand throughout this dissertation. Specifically, we have formulated the KLT and have reviewed the DCT as typical orthogonal transforms. The polyphase representation has been reviewed. This is a sophisticated way to understand filter banks. It allows us to treat filter banks as block-based transforms. Finally, a typical algorithm of transform coding has been described.

For more details, the following literatures are helpful to understand out review. The first description on the DCT appeared in [41]. Relation between the KLT and the DCT is firstly analyzed by Clarke [59]. A good reference for a DCT family is [42]. Other traditional transforms for image coding such as Fourier [63, 64], Hadamard [65], and Slant [66, 67] transforms are well studied in [68, 69].

Vaidyanathan [11] summarized the theory and design of filter banks. This book addresses a very wide range of filter banks. Wavelets and filter banks are well established in [12, 13].

The coding step in a transform image coder is well described and analyzed in [70]. Optimality of the KLT with respect to coding gain is also discussed in this book and in [11]. To understand the JPEG, [39] is widely referred.

Chapter 3

A Lapped Transform with Non-Overlapping Functions

3.1 Introduction

Although lapped transforms reduce the blocking effects, due to their long basis functions, the quantization error is spread out over adjacent blocks. To avoid the spread of high frequency noise, recently, the LOT with variable length functions (VLLOT) has been proposed [18, 38]. The VLLOT consists of overlapping (long) basis functions, which can reduce the blocking artifacts, and block-independent (short) basis functions, which can restrict the ringing artifacts in the block. Its biorthogonal version has also been studied [19], and these lapped transforms with variable length functions (VLLT) are based on lattice structures, where the initial building block is assumed to be the DCT to reduce the degree of freedom for design. Thus, the short basis functions are identical to higher DCT basis functions even in the biorthogonal case [19].

However, there is no theoretical reason that the DCT is adopted in the initial stage of the VLLT. It is used nothing but for fast implementation. In addition, it is very difficult to construct 2-D VLLTs within the conventional framework, since the degree of freedom greatly increases in the 2-D case. In this chapter, therefore, we provide a novel framework of lapped biorthogonal transforms with non-overlapping short basis functions (VLLBT), and its design method based on an eigenvalue problem. Using our proposed approach, the optimal short basis functions can be found without numerical searches when the overlapping basis functions are given. The chapter is organized as follows. In Section 3.2, we review the basic properties of lapped orthogonal transforms and give their factorized structure. In Section 3.2.1, we formulate the VLLBT by

biorthogonalizing the VLLOT. Then, we present theoretical preliminaries to derive the optimal solution for the short non-overlapping basis functions of the VLLBT. More specifically, we describe the transform (SKLT) that provides the optimal approximation of an original signal in the minimum mean square error (MSE) sense. In Section 3.4, using this transform, we show application of the SKLT to the derivation for the short basis functions when the biorthogonal basis functions are given. The optimal short functions can be derived as the solution to an eigenvalue problem. In Section 3.5, design methods for the VLLBT are presented. Coding gain of the resulting transform is superior to that of the orthogonal version. Finally, in Section 3.6, we summarize our proposed framework.

3.2 Variable-Length Lapped Transforms

We review here the linear-phase LOT and the variable-length LOT, which consists of overlapping and non-overlapping basis functions. Furthermore, we generalize it to the biorthogonal form. It is desirable in image processing to inflict transform basis functions on symmetric, that is, linear-phase property. Linear-phase lapped orthogonal transforms where the basis functions have length LM have been studied in [15, 16]. For simplicity, we will review the case $L = 2$ hereafter, but generalization to length LM can be easily attained.

Let $\boldsymbol{H}$ be an orthogonal matrix of size $M \times M$. When the rows of $\boldsymbol{H}$ are symmetric or antisymmetric, the orthogonal matrix $\boldsymbol{H}$ has the form

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_e \\ \boldsymbol{H}_o \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{U}_0 & \boldsymbol{U}_0 \boldsymbol{J}_{M/2} \\ -\boldsymbol{V}_0 & \boldsymbol{V}_0 \boldsymbol{J}_{M/2} \end{bmatrix}, \tag{3.1}$$

where $\boldsymbol{H}_e$ and $\boldsymbol{H}_o$ are $M \times M/2$ matrices containing the even and the odd rows of $\boldsymbol{H}$, respectively, and $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$ are $M/2 \times M/2$ orthogonal matrices. Let $\boldsymbol{P}$ be the Haar butterfly given as

$$\boldsymbol{P} = \frac{1}{2} \begin{bmatrix} \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \\ -\boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \end{bmatrix}, \tag{3.2}$$

which gives an orthogonal projection matrix. Let $\boldsymbol{Z}$ be the orthogonal matrix defined as

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{U}_1 & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{V}_1 \end{bmatrix}, \tag{3.3}$$

where both $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$ are $M/2 \times M/2$ orthogonal matrices. With the symmetric orthogonal matrix $\boldsymbol{H}$ and the orthogonal projection matrix $\boldsymbol{P}$, the general form of the linear-phase lapped

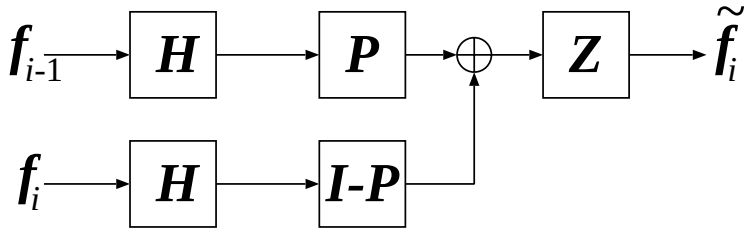


Figure 3.1: Factorized structure of lapped transforms

orthogonal transform is given by

$$\begin{aligned} \boldsymbol{E} &= \boldsymbol{Z}[\boldsymbol{P}\boldsymbol{H}\ (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{H}] && (3.4)\\ &= \frac{1}{2}\begin{bmatrix} \boldsymbol{U}_1(\boldsymbol{H}_e-\boldsymbol{H}_o) & \boldsymbol{U}_1(\boldsymbol{H}_e-\boldsymbol{H}_o)\boldsymbol{J}_M \\ -\boldsymbol{V}_1(\boldsymbol{H}_e-\boldsymbol{H}_o) & \boldsymbol{V}_1(\boldsymbol{H}_e-\boldsymbol{H}_o)\boldsymbol{J}_M \end{bmatrix}. && (3.5)\end{aligned}$$

The block diagram of the factorized LOT is illustrated in Fig. 3.1. It can be easily checked that the LOT conditions (2.32) and (2.33) are imposed on $\boldsymbol{E}$ given as in (3.4). Malvar proposed a fast LOT where $\boldsymbol{H}$ corresponds to the DCT matrix [8, 71]. In this fast LOT, the free parameters are the orthogonal matrices $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$. In order to decrease the number of free parameters, Malvar suggested that $\boldsymbol{U}_1 = \boldsymbol{I}$, and $\boldsymbol{V}_1$ is the product of plane rotations [8] or the product of DCT-II and DST-IV [71].

The lapped orthogonal transform with variable length (VLLOT) [18, 38] has developed in order to avoid the spread of high-frequency distortion into neighboring blocks. The most basic VLLOT consists of N_L long and $(M - N_L)$ short basis functions of length $2M$ and M, respectively. In the VLLOT, according to the necessary conditions for an existing linear phase perfect reconstruction (LPPR) filter bank [13, 18], N_L must be even. The VLLOT matrix can be also written by the factorized form as given as in (3.4). Then, the projection matrix for the VLLOT $\boldsymbol{P}$ is written by

$$\boldsymbol{P} = \frac{1}{2}\begin{bmatrix} \boldsymbol{I}_{N_L/2} & \boldsymbol{0} & -\boldsymbol{I}_{N_L/2} & \boldsymbol{0} \\ \boldsymbol{0} & 2\boldsymbol{I}_{(M-N_L)/2} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{I}_{N_L/2} & \boldsymbol{0} & \boldsymbol{I}_{N_L/2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \tag{3.6}$$

and the orthogonal matrices of the last stage $\boldsymbol{Z}$ are given by

$$\boldsymbol{U}_1 = \begin{bmatrix} \hat{\boldsymbol{U}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{(M-N_L)/2} \end{bmatrix}, \tag{3.7}$$

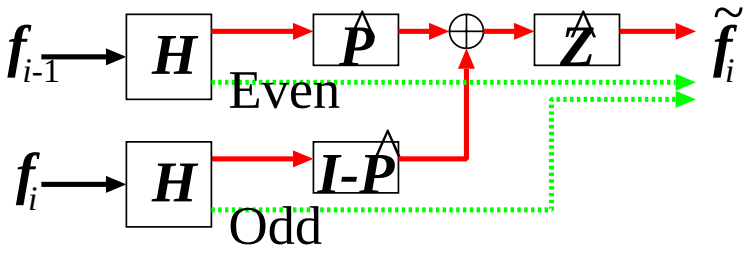


Figure 3.2: VLLOT factorized structure: a projection matrix $\boldsymbol{P}$ is given as in (3.6).

and

$$\boldsymbol{V}_1 = \begin{bmatrix} \hat{\boldsymbol{V}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{(M-N_L)/2} \end{bmatrix}, \tag{3.8}$$

where $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{V}}_1$ are $N_L/2 \times N_L/2$ orthogonal matrices. Substituting $\boldsymbol{P}$, $\boldsymbol{U}_1$, and $\boldsymbol{V}_1$ into (3.4), we have

$$\boldsymbol{E} = \frac{1}{2}\begin{bmatrix} \hat{\boldsymbol{U}}_1(\hat{\boldsymbol{H}}_e - \hat{\boldsymbol{H}}_o) & \hat{\boldsymbol{U}}_1(\hat{\boldsymbol{H}}_e - \hat{\boldsymbol{H}}_o)\boldsymbol{J}_M \\ 2\bar{\boldsymbol{H}}_e & \boldsymbol{0} \\ -\hat{\boldsymbol{V}}_1(\hat{\boldsymbol{H}}_e - \hat{\boldsymbol{H}}_o) & \hat{\boldsymbol{V}}_1(\hat{\boldsymbol{H}}_e - \hat{\boldsymbol{H}}_o)\boldsymbol{J}_M \\ \boldsymbol{0} & 2\bar{\boldsymbol{H}}_o \end{bmatrix}, \tag{3.9}$$

where $\hat{\boldsymbol{H}}_e$ and $\hat{\boldsymbol{H}}_o$ are $M \times N_L/2$ matrices and $\bar{\boldsymbol{H}}_e$ and $\bar{\boldsymbol{H}}_o$ are $M \times (M - N_L)/2$ matrices such that $\boldsymbol{H}_e = \begin{bmatrix} \hat{\boldsymbol{H}}_e \\ \bar{\boldsymbol{H}}_e \end{bmatrix}$ and $\boldsymbol{H}_o = \begin{bmatrix} \hat{\boldsymbol{H}}_o \\ \bar{\boldsymbol{H}}_o \end{bmatrix}$. The above form implies that the N_L rows of $\boldsymbol{H}$ generates the N_L long basis functions of size $2M$. The block diagram of the factorized VLLOT is depicted in Fig. 3.2.

Let us consider the degree of freedom for the VLLOT. By the theory of linear algebra, $\boldsymbol{U}_0$ and $\boldsymbol{V}_1$ each in the first stage has $\binom{M/2}{2}$ degrees of freedom. Moreover, $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{V}}_1$ each contains $\binom{N_L/2}{2}$ degrees of freedom. Therefore, the linear-phase VLLOT being of the generalized form has $2\binom{M/2}{2} + 2\binom{N_L/2}{2}$ degrees of freedom. In a fashion similar to the fast LOT, Nagai and Ikehara proposed the fast VLLOT [38] where $\boldsymbol{H}$ is set to the DCT, $\hat{\boldsymbol{U}}_1 = \boldsymbol{I}_{N_L/2}$, and $\hat{\boldsymbol{V}}_1$ is the product of plane rotations. Their proposed fast VLLOT reduces the degree of freedom to $N_L/2-1$.

3.2.1 Biorthogonalization

Biorthogonalization of orthogonal lapped transforms can be easily accomplished as follows. In this case, $\boldsymbol{E}$ given as in (3.4) is used for the forward transformation. The inverse transformation is defined as a matrix $\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{R}_0 \end{bmatrix}$ in the same fashion as the definition of $\boldsymbol{E}$. The condition for perfect reconstruction is obtained by rewriting the orthogonal constraints (2.32) and (2.33). Thus, the transform matrices $\boldsymbol{E}$ and $\boldsymbol{R}$ must satisfy the following:

$$\boldsymbol{R}_0\boldsymbol{E}_0 + \boldsymbol{R}_1\boldsymbol{E}_1 = \boldsymbol{I}_M, \tag{3.10}$$

$$\boldsymbol{R}_1\boldsymbol{E}_0 = \boldsymbol{R}_0\boldsymbol{E}_1 = \boldsymbol{0}_M, \boldsymbol{R}_1^T\boldsymbol{E}_0^T = \boldsymbol{R}_0^T\boldsymbol{E}_1^T = \boldsymbol{0}_M. \tag{3.11}$$

Equation (3.10) implies biorthogonality of long basis functions, and (3.11) describes birothogonality of tails. In this case, the transform vector $\tilde{\boldsymbol{f}}_i$ containing transform coefficients is given by

$$\tilde{\boldsymbol{f}}_i = \boldsymbol{E}_1^T\boldsymbol{f}_{i-1} + \boldsymbol{E}_1^T\boldsymbol{f}_i. \tag{3.12}$$

Contrarily, the reconstructed signal $\hat{\boldsymbol{f}}_i$ is obtained by

$$\hat{\boldsymbol{f}}_i = \boldsymbol{R}_0\tilde{\boldsymbol{f}}_i + \boldsymbol{R}_1\tilde{\boldsymbol{f}}_{i+1}. \tag{3.13}$$

The general form of the linear-phase lapped biorthogonal transform can be achieved by relaxing orthogonality of the matrix $\boldsymbol{H}$ in (3.4). Assume that $\boldsymbol{H}$ given as in (3.1) is non-singular. Moreover, we assume that $\boldsymbol{Z}$ given as in (3.3) is also non-singular, that is, both $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$ involved in $\boldsymbol{H}$ are non-singular. Then, forward and inverse transform matrices are given by

$$\boldsymbol{E} = \boldsymbol{Z}[\boldsymbol{P}\boldsymbol{H}\ (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{H}] \tag{3.14}$$

for the forward transform, and

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{G}\boldsymbol{P} \\ \boldsymbol{G}(\boldsymbol{I} - \boldsymbol{P}) \end{bmatrix}\boldsymbol{Z}^{-1}, \tag{3.15}$$

for the inverse transform, where

$$\boldsymbol{G} = \boldsymbol{H}^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{U}_0^{-1} & -\boldsymbol{V}_0^{-1} \\ \boldsymbol{J}\boldsymbol{U}_0^{-1} & \boldsymbol{J}\boldsymbol{V}_0^{-1} \end{bmatrix} \tag{3.16}$$

and,

$$\boldsymbol{Z}^{-1} = \begin{bmatrix} \boldsymbol{U}_1^{-1} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{V}_1^{-1} \end{bmatrix}, \tag{3.17}$$

When we choose $\boldsymbol{P}$ given as in (3.6), then $\boldsymbol{E}$ and $\boldsymbol{R}$ fulfill the requirements for perfect reconstruction described in (3.10) and (3.11). Let $\boldsymbol{e}_i$ and $\boldsymbol{r}_i$ be the i-th columns of the transform matrices $\boldsymbol{E}^T$ and $\boldsymbol{R}$, respectively. Then, a pair of vector sets $\{\boldsymbol{e}_i\}_{i=0}^{M}$ and $\{\boldsymbol{r}_i\}_{i=0}^{M}$ satisfies both biorthogonality and shift-biorthogonality. In the context of filter banks, $\boldsymbol{e}_i$ and $\boldsymbol{r}_i$ are called analysis and synthesis filters of biorthogonal filter banks, respectively.

A lapped biorthogonal transform with variable length basis functions (VLLBT) can be defined in the same manner as the VLLOT. The projection matrix $\boldsymbol{P}$ is given in (3.6). In the last stage, $\boldsymbol{U}_1^{-1}$ and $\boldsymbol{V}_1^{-1}$ included in $\boldsymbol{Z}^{-1}$ are expressed as

$$\boldsymbol{U}_1^{-1} = \begin{bmatrix} \hat{\boldsymbol{U}}_1^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{(M-N_L)/2} \end{bmatrix}, \tag{3.18}$$

and

$$\boldsymbol{V}_1^{-1} = \begin{bmatrix} \hat{\boldsymbol{V}}_1^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{(M-N_L)/2} \end{bmatrix}, \tag{3.19}$$

Evidently, the degree of freedom of the VLLBT increases comparing to that of the VLLOT. The degree of freedom of $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$ each is $(M/2)^2$. In the next stage, $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{V}}_1$ each has $(N_L/2)^2$ degrees of freedom. In total, the linear-phase VLLBT has $\dfrac{(M^2 + N_L^2)}{2}$ degrees of freedom.

3.3 Subspace Karhunen-Loève Transform

This section gives theoretical preliminaries to derive the optimal short basis functions of the VLLBT. Let $\boldsymbol{f}$ be a vector of M consecutive samples of a real wide-sense stationary random process. The well-known Karhunen-Loève transform (KLT) provides the optimal approximation of $\boldsymbol{f}$ [70, 12]. Moreover, among all block transforms, the KLT is indeed the best possible transform for minimizing the overall distortion for a given bit allocation, that is, the KLT provides the maximum coding gain [70]. Assume that the M-dimensional Euclidean space $\mathbb{R}^M$ is a direct sum of two spaces $\mathbb{S}_1$ and $\mathbb{S}_2$, that is, $\mathbb{R}^M = \mathbb{S}_1 \oplus \mathbb{S}_2$. Notice that it is unnecessary that $\mathbb{S}_1$ and $\mathbb{S}_2$ are orthogonal (see Fig. 3.3). Then, when the subspace $\mathbb{S}_1$ is given, the transform that provides the optimal approximation in $\mathbb{S}_2$ is called the subspace Karhunen-Loève transform (SKLT), which we propose in this chapter. This SKLT is defined as an extension of the vector-embedded Karhunen-Loève transform (VEKLT) proposed by Tanaka and Yamashita [72, 44]. In the SKLT, orthogonality for $\mathbb{S}_1$ and $\mathbb{S}_2$ of the VEKLT is softened.

Let $\boldsymbol{L}$ be the projection matrix onto $\mathbb{S}_1$. The projection matrix $\boldsymbol{L}$ is not necessarily orthogonal.

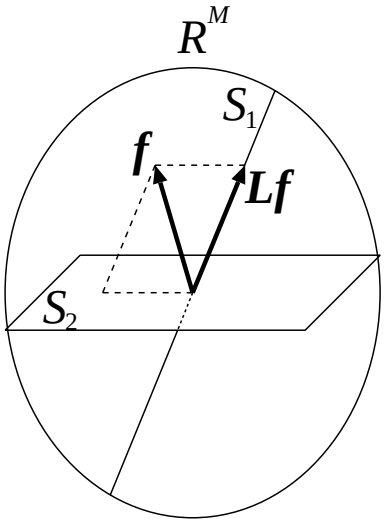


Figure 3.3: The case the subspaces $\mathbb{S}_1$ and $\mathbb{S}_2$ are not orthogonal

Definition 5 *When a projection matrix $\boldsymbol{L}$ whose rank is $N_L < M$ is given, the subspace Karhunen-Loève transform (SKLT) $\boldsymbol{X}$ minimizes the functional*

$$J[\boldsymbol{X}] = E_{\boldsymbol{f}}\|\boldsymbol{f} - (\boldsymbol{L} + \boldsymbol{X})\boldsymbol{f}\|^2 \tag{3.20}$$

under the condition $\mathrm{rank}(\boldsymbol{X}) = N$ *for any* $N \leq M - N_L$.

When the subspace $\mathbb{S}_1$ is given, the SKLT provides the optimal approximation in $\mathbb{S}_2$. Since letting $\boldsymbol{L} = \boldsymbol{0}$ leads to the criterion for the KLT [56], the KLT is a subclass of SKLTs.

Fortunately, the analytic solution of the above problem can be derived as will be shown. Let $\boldsymbol{R}_{\boldsymbol{ff}} = E_{\boldsymbol{f}}[\boldsymbol{f}\boldsymbol{f}^T]$ be the correlation matrix of the input vector $\boldsymbol{f}$.

Lemma 1 *Assume $\boldsymbol{R}_{\boldsymbol{ff}}$ has full rank. Then, there exist the $(M - N_L)$ non-zero eigenvalues of*

$$\boldsymbol{Q} = (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I} - \boldsymbol{L})^T \tag{3.21}$$

such that $\xi_0 \geq \cdots \geq \xi_{M-N_L-1} > 0$, and the corresponding eigenvectors $\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_{M-N_L-1}$.

Lemma 2 ([55, 56]) *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ matrices of rank M and N_L, respectively, where $N_L \leq M$, and the singular value decomposition (SVD) of $\boldsymbol{A}$ is given as*

$$\boldsymbol{A} = \sum_{i=0}^{M-1} \lambda_i \boldsymbol{\varphi}_i^* \boldsymbol{\varphi}_i^T, \tag{3.22}$$

where $\lambda_0 \geq \cdots \geq \lambda_{M-1}$*, and* $\{\boldsymbol{\varphi}_i\}_{i=0}^{M-1}$ *and* $\{\boldsymbol{\varphi}_i^*\}_{i=0}^{M-1}$ *are orthonormal bases in* $\mathbb{R}^M$*. Then, a matrix* $\boldsymbol{B}$ *minimizes a functional*

$$J_1[\boldsymbol{B}] = \|\boldsymbol{B} - \boldsymbol{A}\|_F \tag{3.23}$$

subject to $\text{rank}(\boldsymbol{B}) \leq N_L$ *if and only if*

$$\boldsymbol{B} = \sum_{i=0}^{N_L-1} \lambda_i \boldsymbol{\varphi}_i^* \boldsymbol{\varphi}_i^T. \tag{3.24}$$

Theorem 2 *Let* $\boldsymbol{\phi}_i$ *be eigenvectors of* $\boldsymbol{Q}$ *with respect to non-zero eigenvalues. Assume that we choose the eigenvectors* $\boldsymbol{\phi}_i$ *such that* $\{\boldsymbol{\phi}_i\}_{i=0}^{M-N_L-1}$ *forms an orthonormal system. Then, the functional* $J[\boldsymbol{X}]$ *in (3.20) is minimized by*

$$\boldsymbol{X} = \sum_{n=0}^{N-1} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^{*T}, \tag{3.25}$$

where

$$\boldsymbol{\phi}_n^* = (\boldsymbol{I} - \boldsymbol{L})^T \boldsymbol{\phi}_n. \tag{3.26}$$

Proof: Equation (3.20) can be rewritten as

$$\begin{aligned} J[\boldsymbol{X}] &= E_{\boldsymbol{f}} \|(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{f} - \boldsymbol{X}\boldsymbol{f}\|^2 \\ &= \text{tr}[(\boldsymbol{I} - \boldsymbol{L} - \boldsymbol{X}) E_{\boldsymbol{f}}(\boldsymbol{f}\boldsymbol{f}^T)(\boldsymbol{I} - \boldsymbol{L} - \boldsymbol{X})^T] \\ &= \text{tr}[(\boldsymbol{I} - \boldsymbol{L} - \boldsymbol{X}) \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}} (\boldsymbol{I} - \boldsymbol{L} - \boldsymbol{X})^T] \\ &= \|(\boldsymbol{I} - \boldsymbol{L} - \boldsymbol{X}) \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}\|_F^2 \\ &= \|(\boldsymbol{I} - \boldsymbol{L}) \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2} - \boldsymbol{X} \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}\|_F^2, \end{aligned} \tag{3.27}$$

where $\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}$ is a positive semidefinite matrix such that $\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2} \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2} = \boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}$. Since the correlation matrix $\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}$ is assumed to be non-singular, the rank of $\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}$ is also M. Therefore, we have $\text{rank}((\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}) = M - N_L$. The matrix $(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2}$ has singular value decomposition (SVD)

$$(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R}_{\boldsymbol{f}\boldsymbol{f}}^{1/2} = \sum_{i=0}^{M-N_L-1} \mu_i \boldsymbol{\phi}_i \boldsymbol{\psi}_i^T, \tag{3.28}$$

where $\boldsymbol{\phi}_i$ and $\boldsymbol{\psi}_i$ satisfy the following equations:

$$(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}\boldsymbol{\psi}_i = \mu_i\boldsymbol{\phi}_i, \tag{3.29}$$

$$[(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}]^T\boldsymbol{\phi}_i = \mu_i\boldsymbol{\psi}_i. \tag{3.30}$$

Since the rank of $\boldsymbol{X}$ is assumed to be $N \le M - N_L$, we have rank$(\boldsymbol{X}\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}) = N$. According to Lemma 2, the functional $J[\boldsymbol{X}]$ is minimized if and only if

$$\boldsymbol{X}\boldsymbol{R}_{\boldsymbol{ff}}^{1/2} = \sum_{i=0}^{N-1}\mu_i\boldsymbol{\phi}_i\boldsymbol{\psi}_i^T. \tag{3.31}$$

Since $\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}$ has full rank, there exists its inverse $\boldsymbol{R}_{\boldsymbol{ff}}^{-1/2}$. Therefore, (3.31) and (3.30) yield that

$$\begin{aligned}
\boldsymbol{X} &= \sum_{i=0}^{N-1}\mu_i\boldsymbol{\phi}_i\boldsymbol{\psi}_i^T\boldsymbol{R}_{\boldsymbol{ff}}^{-1/2} \\
&= \sum_{i=0}^{N-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^T[(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}]\boldsymbol{R}_{\boldsymbol{ff}}^{-1/2} \\
&= \sum_{i=0}^{N-1}\boldsymbol{\phi}_i[(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{\phi}_i]^T \\
&= \sum_{i=0}^{N-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^{*T}.
\end{aligned} \tag{3.32}$$

On the other hand, pre-multiplying both sides of (3.30) by $(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}$, we have

$$\begin{aligned}
(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}[(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}]^T\boldsymbol{\phi}_i &= \mu_i(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}\boldsymbol{\psi}_i \\
&= \mu_i^2\boldsymbol{\phi}_i,
\end{aligned} \tag{3.33}$$

with using (3.29). This yields

$$(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{\phi}_i = \mu_i^2\boldsymbol{\phi}_i, \tag{3.34}$$

which implies that $\boldsymbol{\phi}_i$ and μ_i^2 are an eigenvector and an eigenvalue of the matrix $(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I}-\boldsymbol{L})^T$, respectively. This completes the proof. □

We shall call this matrix $\boldsymbol{X}$ the *subspace Karhunen-Loève transform* (SKLT) of rank N with respect to $\boldsymbol{L}$.

Next, we have the following result on biorthogonality of the SKLT.

Lemma 3 *For $i = 0, \ldots, M - N_L - 1$, the eigenvector $\boldsymbol{\phi}_i$ of $\boldsymbol{Q}$ is in the null space of $\boldsymbol{L}$, that is,*

$$\boldsymbol{L}\boldsymbol{\phi}_i = \boldsymbol{0}. \tag{3.35}$$

The vector sets $\{\boldsymbol{\phi}_i\}_{i=0}^{M-N_L-1}$ and $\{\boldsymbol{\phi}_i^*\}_{i=0}^{M-N_L-1}$ have the following property.

Proposition 1 *Vector sets $\{\boldsymbol{\phi}_i\}_{i=0}^{M-N_L-1}$ and $\{\boldsymbol{\phi}_i^*\}_{i=0}^{M-N_L-1}$ construct a biorthonormal system, that is,*

$$\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j^* \rangle = \delta_{i,j}, \tag{3.36}$$

for $i, j = 0, \ldots, M - N_L - 1$.

Proof: Since the eigenvectors of $\boldsymbol{Q}$ are orthogonal, (3.35) yields

$$\langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j^* \rangle = \langle \boldsymbol{\phi}_i, (\boldsymbol{I} - \boldsymbol{L})^T \boldsymbol{\phi}_j \rangle = \langle (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\phi}_i, \boldsymbol{\phi}_j \rangle = \langle \boldsymbol{\phi}_i, \boldsymbol{\phi}_j \rangle = \delta_{i,j}. \tag{3.37}$$

□

In the following sections, it will be shown that we can apply this SKLT to design of lapped transform with variable length basis functions.

3.4 Application of the SKLT in VLLBT

Let $\boldsymbol{h}_i$ and $\boldsymbol{g}_i$ be i-th column functions of the matrices $\boldsymbol{H}^T$ and $\boldsymbol{G}$, respectively. In this subsection, when the long basis functions of the VLLBT determined by $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=0}^{N_L-1}$ are given, we discuss the design method for short basis functions.

Let $\boldsymbol{u}_i$ and $\boldsymbol{u}_i^*$ be columns of $\boldsymbol{U}_0^T$ and $\boldsymbol{U}_0^{-1}$, respectively, and let $\boldsymbol{v}_i$ and $\boldsymbol{v}_i^*$ be columns of $\boldsymbol{V}_0^T$ and $\boldsymbol{V}_0^{-1}$, respectively. Assume that biorthogonal systems $\{\boldsymbol{u}_i, \boldsymbol{u}_i^*\}_{i=0}^{N_L/2-1}$ and $\{\boldsymbol{v}_i, \boldsymbol{v}_i^*\}_{i=0}^{N_L/2-1}$ are given. Then, we obtain a biorthogonal system $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=0}^{N_L-1}$ such that for $i = 0, \ldots, N_L/2 - 1$,

$$\boldsymbol{h}_{2i} = \begin{bmatrix} \boldsymbol{u}_i \\ \boldsymbol{J}\boldsymbol{u}_i \end{bmatrix}, \boldsymbol{h}_{2i+1} = \begin{bmatrix} \boldsymbol{v}_i \\ -\boldsymbol{J}\boldsymbol{v}_i \end{bmatrix}, \tag{3.38}$$

and

$$\boldsymbol{g}_{2i} = \begin{bmatrix} \boldsymbol{u}_i^* \\ \boldsymbol{J}\boldsymbol{u}_i^* \end{bmatrix}, \boldsymbol{g}_{2i+1} = \begin{bmatrix} \boldsymbol{v}_i^* \\ -\boldsymbol{J}\boldsymbol{v}_i^* \end{bmatrix}, \tag{3.39}$$

From the theory of the VLLBT, $\{\boldsymbol{h}_i\}_{i=0}^{N_L-1}$ and $\{\boldsymbol{g}_i\}_{i=0}^{N_L-1}$ produce the long basis functions for the forward and the inverse transforms, respectively. With these functions $\boldsymbol{h}_i$ and $\boldsymbol{g}_i$, let

$$\boldsymbol{L} = \sum_{i=0}^{N_L-1} \boldsymbol{g}_i \boldsymbol{h}_i^T. \tag{3.40}$$

When we apply the projection matrix $\boldsymbol{L}$ to Theorem 2, we obtain a biorthonormal system $\{\boldsymbol{\phi}_i, \boldsymbol{\phi}_i^*\}_{i=0}^{M-N_L-1}$, which leads to the SKLT. If $\boldsymbol{h}_i = \epsilon_i \boldsymbol{h}_i'$ and $\boldsymbol{g}_i = \frac{1}{\epsilon_i}\boldsymbol{g}_i'$, where $\boldsymbol{h}_i' = \boldsymbol{g}_i'$, $\boldsymbol{L}$ gives an orthogonal projection; otherwise, it gives an oblique projection. In the former case, we have $\boldsymbol{\phi}_i = \boldsymbol{\phi}_i^*$ for $i = 0, \ldots, M - N_L - 1$. Most of biorthogonal lapped transforms correspond to this case [38, 22, 19]. The subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ are orthogonal.

To show that $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_i^*$ are symmetric/antisymmetric vectors, we introduce the following theorem. Below, $[x]$ denotes the integer part of x.

Lemma 4 ([73, 74]) *Let $\boldsymbol{A}$ be a $M \times M$ persymmetric matrix, that is, $\boldsymbol{JAJ} = \boldsymbol{A}$. Let $\boldsymbol{A}$ have distinct eigenvalues. Then, $\boldsymbol{A}$ has $[(M+1)/2]$ symmetric eigenvectors and $[M/2]$ antisymmetric eigenvectors that span the eigenspace of $\boldsymbol{Q}$.*

Theorem 3 *Let $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=0}^{N_L-1}$ be a biorthogonal system, where N_L is even. Let $\boldsymbol{L} = \sum_{i=0}^{N_L-1} \boldsymbol{g}_i \boldsymbol{h}_i^T$. Assume that $\boldsymbol{h}_i$ and $\boldsymbol{g}_i$ are symmetric for even i, and antisymmetric for odd i. Assume also that $\boldsymbol{R_{ff}}$ is $M \times M$ persymmetric matrix, that is, $\boldsymbol{JR_{ff}J} = \boldsymbol{R_{ff}}$, and has distinct eigenvalues. Then, the matrix*

$$\boldsymbol{Q} = (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R_{ff}}(\boldsymbol{I} - \boldsymbol{L})^T \tag{3.41}$$

has $(M - N_L)/2$ symmetric eigenvectors and $(M - N_L)/2$ antisymmetric eigenvectors.

Proof: For $i = 0, \ldots, N_L - 1$, $\boldsymbol{h}_i$ is in the null space of $\boldsymbol{Q}$, since

$$\begin{aligned} \boldsymbol{Q}\boldsymbol{h}_i &= (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R_{ff}}\left(\boldsymbol{h}_i - \sum_{j=0}^{N_L-1} \boldsymbol{h}_j \boldsymbol{g}_j^T \boldsymbol{h}_i\right) \\ &= (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R_{ff}}(\boldsymbol{h}_i - \boldsymbol{h}_i) \\ &= \boldsymbol{0}. \end{aligned} \tag{3.42}$$

From Lemma 1, $\boldsymbol{Q}$ has $(M - N_L)$ non-zero eigenvalues. Let

$$\boldsymbol{S} = \boldsymbol{Q} + \sum_{i=0}^{N_L-1} \boldsymbol{h}_i \boldsymbol{g}_i^T. \tag{3.43}$$

Then, $\boldsymbol{h}_i$ is an eigenvector of $\boldsymbol{S}$ and the corresponding eigenvalue is 1, since

$$\begin{aligned} \boldsymbol{S}\boldsymbol{h}_i &= \boldsymbol{Q}\boldsymbol{h}_i + \sum_{j=0}^{N_L-1} \boldsymbol{h}_j \boldsymbol{g}_j^T \boldsymbol{h}_i, \\ &= \boldsymbol{h}_i. \end{aligned} \tag{3.44}$$

Therefore, $\boldsymbol{S}$ has M eigenvalues. Moreover, we can easily examine that $\boldsymbol{JSJ} = \boldsymbol{S}$, so that $\boldsymbol{S}$ is a persymmetric matrix. From Lemma 4, $\boldsymbol{S}$ has $N_L/2$ symmetric eigenvectors containing $\boldsymbol{h}_i$ for even i, and also has $N_L/2$ antisymmetric eigenvectors containing $\boldsymbol{h}_i$ for odd i. It follows that $(M - N_L)/2$ eigenvectors of $\boldsymbol{Q}$ are symmetric and the others are antisymmetric. □

Theorem 3 guarantees that in $\boldsymbol{H}$ and $\boldsymbol{G}$, we can set that

$$\boldsymbol{h}_i = \boldsymbol{\phi}^*_{i-N_L}, \tag{3.45}$$

and

$$\boldsymbol{g}_i = \boldsymbol{\phi}_{i-N_L}, \tag{3.46}$$

for $i = N_L, \ldots, M-1$. Then, $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=0}^{M-1}$ gives a biorthonormal basis for $\mathbb{R}^M$. Keep in mind that $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=N_L}^{M-1}$ are uniquely determined by $\{\boldsymbol{h}_i, \boldsymbol{g}_i\}_{i=0}^{N_L-1}$ and $\boldsymbol{R}_{\boldsymbol{ff}}$. Substituting $\boldsymbol{H}$ and $\boldsymbol{G}$ into (3.14) and (3.15), respectively, we obtain the transform matrices $\boldsymbol{E}$ and $\boldsymbol{R}$.

By the use of the SKLT, the degree of freedom of the linear-phase VLLBT is significantly decreased. In the first stage, since it is only necessary to find the N_L lowest basis functions, the degree of freedom for $\{\boldsymbol{u}_i, \boldsymbol{u}_i^*\}_{i=0}^{N_L/2-1}$ is $(M/2)^2 - (M/2 - N_L/2)^2 = \dfrac{M^2 - (M-N_L)^2}{4}$ and so is that for $\{\boldsymbol{v}_i, \boldsymbol{v}_i^*\}_{i=0}^{N_L/2-1}$. Therefore, the VLLBT via the SKLT contains $\dfrac{M^2 - (M-N_L)^2}{2} + \dfrac{N_L^2}{2} = MN_L$ degree of freedom. Since the difference of the degree of freedom between the VLLBT and the VLLBT via the SKLT is $\dfrac{M^2 + N_L^2}{2} - MN_L = \dfrac{(M-N_L)^2}{2}$, the reduction in the degree of freedom is more significant for larger M or smaller N_L.

In summary, we have clarified in this section that from a functional approximation of view, which is the most fundamental concept in image compression, the degree of freedom of the VLLBT can be reduced; we need not find all columns of $\boldsymbol{H}$ and $\boldsymbol{G}$. As a result, the short basis functions are automatically derived from the long basis functions. Therefore, we should find only the long ones. We compare the degrees of freedom of the VLLOT, the VLLBT, and the VLLBT via the SKLT in Table 3.1. This table also contains examples of the case that $M = 8, N_L = 2$ and the case that $M = 8, N_L = 4$. To reduce free parameters, of course, a lot of methods have been proposed [18, 38, 19]. However, recall that all of those methods aim fast implementation. Furthermore, the reduction of free parameters have been carried out empirically, not theoretically. In contrast, the proposed design method with the SKLT ensures the reduction of the degree of freedom in the sense of mean square error.

Table 3.1: Comparison of the degrees of freedom

	VLLOT	VLLBT	VLLBT via SKLT
General case	$2\left[\binom{M/2}{2}+\binom{N_L/2}{2}\right]$	$\frac{M^2+N_L^2}{2}$	MN_L
$M=8, N_L=2$	12	34	16
$M=8, N_L=4$	14	40	32

3.5 Design Method

Since the optimal short basis functions with respect to a given set of long basis functions have been found, we only need to determine suitable long basis functions. For application in image coding, we use coding gain as a cost function. Higher coding gain correlates most consistently with higher PSNR. Coding gain for a biorthogonal transform is given by [75, 76]

$$J_{CG} = 10\log_{10}\frac{\sigma_{\boldsymbol{f}}^2}{\left(\prod_{i=0}^{M-1}\sigma_{\tilde{\boldsymbol{f}}_i}^2\|\boldsymbol{r}_i\|^2\right)^{1/M}}, \tag{3.47}$$

where $\sigma_{\boldsymbol{f}}^2$ is the variance of the input signal and $\sigma_{\tilde{\boldsymbol{f}}_i}^2$ is the variance of the i-th transform coefficient. Assume that the signal is the first-order Markov process with the correlation coefficient $\rho = 0.95$ [59], which is widely used in image processing. The correlation matrix is given by $(\boldsymbol{C})_{i,j} = \sigma_{\boldsymbol{f}}^2\rho^{|i-j|}$, for $i, j = 0, \ldots, 2M-1$. This yields the following expression for $\sigma_{\tilde{\boldsymbol{f}}_i}^2$:

$$\sigma_{\tilde{\boldsymbol{f}}_i}^2 = E_{\boldsymbol{f}}|\langle \boldsymbol{e}_i, \boldsymbol{f}\rangle|^2 = E_{\boldsymbol{f}}[\boldsymbol{e}_i^T\boldsymbol{f}\boldsymbol{f}^T\boldsymbol{e}_i] = \sigma_{\boldsymbol{f}}^2\langle \boldsymbol{e}_i, \boldsymbol{C}\boldsymbol{e}_i\rangle. \tag{3.48}$$

Therefore, the criterion J_{CG} in (3.47) is written as

$$J_{CG} = 10\log_{10}\left(\prod_{i=0}^{M-1}\langle \boldsymbol{e}_i, \boldsymbol{C}\boldsymbol{e}_i\rangle\|\boldsymbol{r}_i\|^2\right)^{-1/M}, \tag{3.49}$$

where recall that $\boldsymbol{e}_i$ and $\boldsymbol{r}_i$ are given in (3.14) and (3.15), respectively.

Low DC Leakage For image compression purpose, "low DC leakage" is an essential requirement [13]. Assume that the input signal is a constant function, that is, a DC signal, and then consider the case that only the lowest transform coefficient is kept and the rest is set to zero. If

each transform coefficient except the lowest one is not zero, then the reconstructed signal is no longer identical to the original one. Such basis functions cause the *checkerboard artifact* [13]. Therefore, the inner product of each basis function except the lowest one with the original signal must be zero. In other words, let $\mathbf{1}_N$ be a DC vector such that all components are one, that is, $\mathbf{1}_N = (\underbrace{1, \ldots, 1}_{N})^T$. Then, for $i = 1, \ldots, M-1$, $\boldsymbol{e}_i$ has to satisfy that

$$\langle \boldsymbol{e}_i, \mathbf{1}_{2M} \rangle = 0. \tag{3.50}$$

A cost function for the low DC leakage is defined as

$$J_{DC} = \sum_{i=1}^{M-1} |\langle \boldsymbol{e}_i, \mathbf{1}_{2M} \rangle|^2. \tag{3.51}$$

Consequently, finding VLLBT is formulated as a non-linear optimization problem:

$$\underset{\{\boldsymbol{u}_i, \boldsymbol{u}_i^*, \boldsymbol{v}_i, \boldsymbol{v}_i^*\}_{i=0}^{N_L/2-1}, \hat{\boldsymbol{U}}_1, \hat{\boldsymbol{V}}_1}{\text{Minimize}} \quad -J_{CG} + \mu J_{DC}, \tag{3.52}$$

subject to $\langle \boldsymbol{u}_i, \boldsymbol{u}_j^* \rangle = \delta_{i,j}$ and $\langle \boldsymbol{v}_i, \boldsymbol{v}_j^* \rangle = \delta_{i,j}$, for $i, j = 0, \ldots, N_L/2 - 1$, which are biorthogonal constraints. However, the objective function J_{CG} can involve the constraints if we regard those constraint equations as linear equations with respect to $u_i(n)$ and $v_i(n)$ (or $u_i^*(n)$ and $v_i^*(n)$). For example, a constraint $\langle \boldsymbol{u}_0, \boldsymbol{u}_1^* \rangle = 0$ means that a component $u_1(0)$ is determined by the other components, i.e, $u_0(0) = -[u_0(1)u_1^*(1) + \cdots + u_0(M/2-1)u_1^*(M/2-1)]/u_1^*(0)$. Therefore, this optimization problem is reduced to an unconstrained optimization problem including free parameters of which number is listed in Table 3.1.

In every step of the optimization, the short basis functions are found via the SKLT with a correlation matrix of the Markov model, as expressed in Section 4.3, and then the cost function (3.52) into which $\boldsymbol{E}$ and $\boldsymbol{R}$ are substituted is evaluated.

3.5.1 A Design Example and Evaluation

Using these properties, for example, let us consider the case $M = 8$ and $N_L = 2$. The case $N_L = 2$ gives the minimum number of the long basis functions because of the existing condition as mentioned previously. In this case, $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{V}}_1$ included in the last stage $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$ are expressed as scalars, i.e, $\hat{\boldsymbol{U}}_1 = \alpha$ and $\hat{\boldsymbol{V}}_1 = \beta$. Obviously, we have that $\hat{\boldsymbol{U}}_1^{-1} = 1/\alpha$ and $\hat{\boldsymbol{V}}_1^{-1} = 1/\beta$. As a result, the last stages $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$ only multiply the first and the second long basis functions

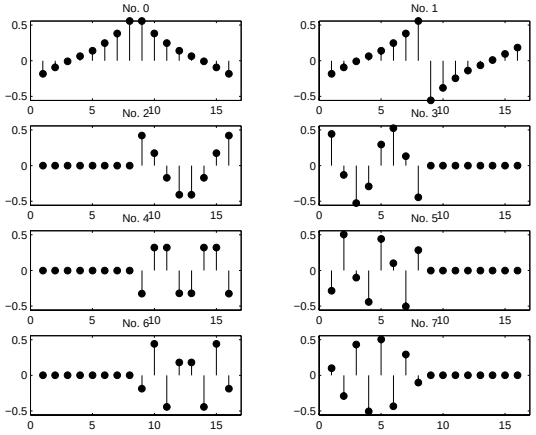
Table 3.2: Comparison of coding gain in dB for AR(1) with $\rho = 0.95$: We choose $M = 8$ for all cases.

Transforms	Coding Gain	Figure
DCT	8.826	
KLT	8.846	
LOT [10]	9.219	
VLLOT26 with DCT [38]	8.954	
VLLBT26 (all free parameters are optimized)	9.325	Fig. 3.4
VLLBT26 with DCT [23]	9.269	
VLLBT26 via SKLT	9.325	
VLLBT26 via SKLT (Low DC leakage)	9.320	Fig. 3.5

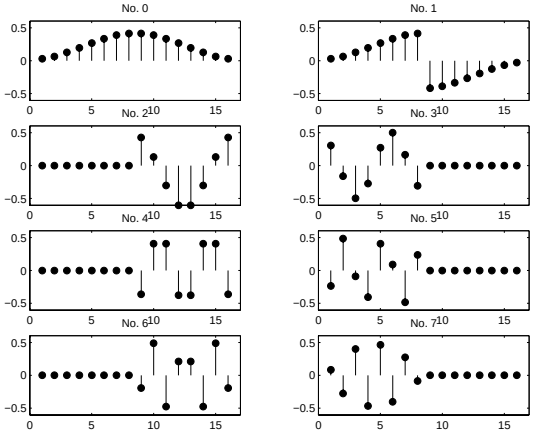
by a constant, respectively. Therefore, we fix the constants to 1, that is, $\alpha = \beta = 1$. We use the correlation matrix $[\boldsymbol{R}_{\boldsymbol{ff}}]_{i,j} = \rho^{|i-j|}$, for $i, j = 0, \ldots, M-1$ for Theorem 2 to find the short basis functions. Figures 3.4(a) and 3.4(b) illustrate the resulting basis functions $\boldsymbol{e}_i$ and $\boldsymbol{r}_i$ of the forward and the inverse transforms, respectively. Interestingly, the samples of the inverse long basis functions decay to zero at their ends, although we never impose any requirements of decay on transforms. This is greatly effective to reduce blocking. If we design the VLLBT ($N_L = 2$) without DC leakage, the unknown parameters are $\boldsymbol{u}_0, \boldsymbol{u}_0^*, \boldsymbol{v}_0$, and $\boldsymbol{v}_0^*$ where two of the 16 components are determined by the others. In this case, we can also rewrite (3.52) as the unconstrained optimization in the same manner. Table 3.2 shows comparison of coding gain of several transforms. In the table, "26" means that the transform matrix consistrs of two long and six short basis functions. Note that coding gain of the VLLBT where all free parameters are optimized is the same as that of the VLLBT via SKLT. Therefore, the use of the SKLT is effective in decreasing parameters to be determined. The basis functions and their frequency responses are depicted in Figs. 3.4 and 3.5. Moreover, the VLLBT via SKLT which is optimized only for coding gain gains 0.056 dB over the VLLBT26 with DCT. Indeed, the use of the DCT can reduce the number of parameters; however it leads to decline in coding gain.

3.6 Summary

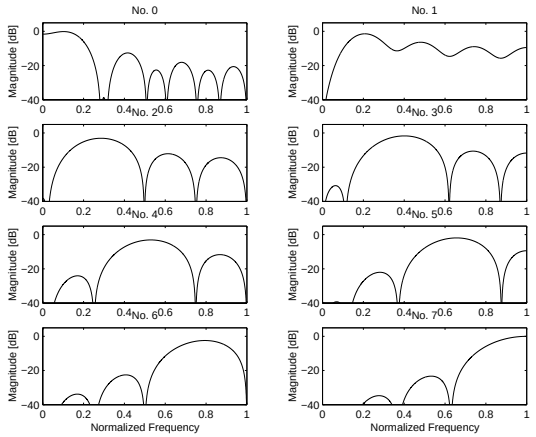
This chapter has presented a new framework for biorthogonal lapped transforms that consist of long and short basis functions called the VLLBT. Moreover, the transform that gives the optimal



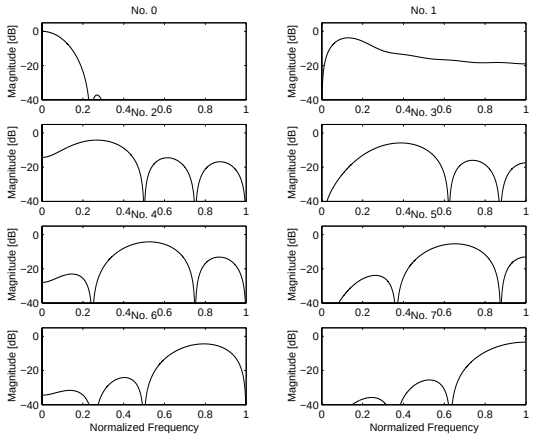
(a) The basis functions $\boldsymbol{e}_i$ of the forward transform



(b) The basis functions $\boldsymbol{r}_i$ of the inverse transform

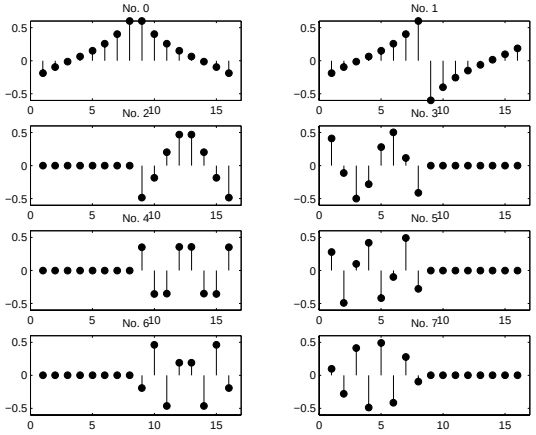


(c) Magnitude responses of the forward functions

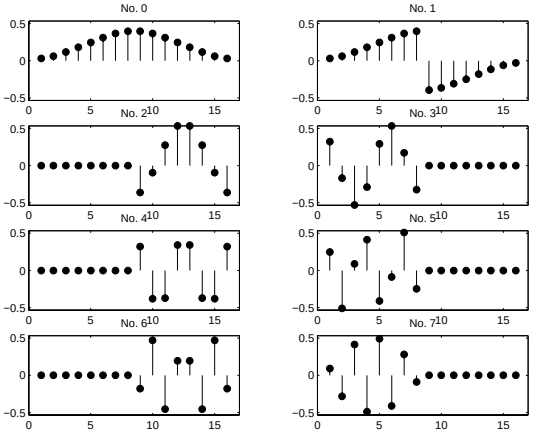


(d) Magnitude responses of the inverse functions

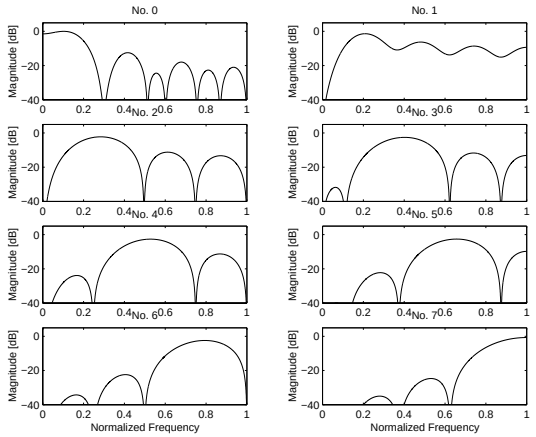
Figure 3.4: A design example VLLBT26 where all paremeters are optimized for coding gain: $M = 8, N_L = 2$ (Coding gain = 9.325 dB)



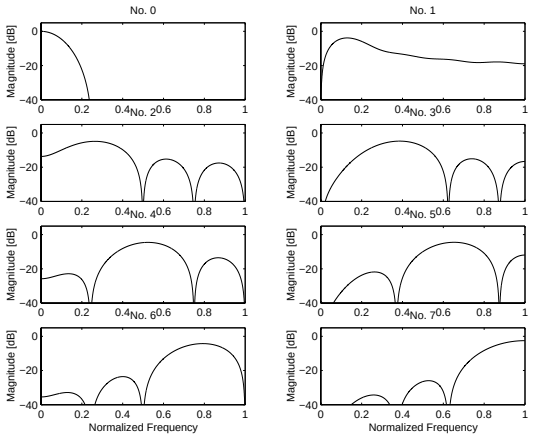
(a) The basis functions $\boldsymbol{e}_i$ of the forward transform



(b) The basis functions $\boldsymbol{r}_i$ of the inverse transform



(c) Magnitude responses of the forward functions



(d) Magnitude responses of the inverse functions

Figure 3.5: A design example VLLBT26 via SKLT with low DC leakage: $M = 8, N_L = 2$ (Coding gain = 9.320 dB)

approximation in a given subspace (SKLT) has been described. We have shown that when the biorthogonal long basis functions of the VLLBT are given, the optimal short basis functions in the energy compaction sense is derived by solving an eigenvalue problem which is resulted in the SKLT. Therefore, we can find the short basis functions uniquely without iterative searching, once the long basis functions are determined.

We have also provided design and image coding examples of the VLLBT. Coding gain of the VLLBT is higher than that of the VLLOT. Moreover, coding examples show that the proposed VLLBT is superior to other conventional transforms in terms of PSNR at high compression ratio. Furthermore, it significantly reduces the annoying blocking artifacts. These results may imply that the proposed VLLBT is a promising technique in the field of image coding.

The VLLBT designed by the method demonstrated in this chapter includes a potential for image coding using 1-D or 2-D adaptive transforms [47, 77, 78], which would be constructed by changing the correlation matrix $\boldsymbol{R}_{\boldsymbol{ff}}$ in (3.21), adaptively. This problem will be addressed in the next chapter.

3.7 Proofs

3.7.1 Proof of Lemma 1

For any $\boldsymbol{x} \in \mathbb{R}^M$, we have

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{Q}\boldsymbol{x} \rangle &= \langle \boldsymbol{x}, (\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{x} \rangle \\ &= \langle \boldsymbol{R}_{\boldsymbol{ff}}^{1/2}(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{x}, \boldsymbol{R}_{\boldsymbol{ff}}^{1/2}(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{x} \rangle \\ &= \|\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}(\boldsymbol{I}-\boldsymbol{L})^T\boldsymbol{x}\|^2 \geq 0. \end{aligned} \tag{3.53}$$

This implies that the matrix $\boldsymbol{Q}$ is positive semidefinite. On the other hand, since the rank of $\boldsymbol{R}_{\boldsymbol{ff}}$ is assumed to be M, that of $\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}$ is also M. Then, we have the rank of $\boldsymbol{I}-\boldsymbol{L}$ is $M-N_L$, we have

$$\begin{aligned} \mathrm{rank}(\boldsymbol{Q}) &= \mathrm{rank}((\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I}-\boldsymbol{L})^T) \\ &= \mathrm{rank}((\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}[(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}]^T) \\ &= \mathrm{rank}((\boldsymbol{I}-\boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}^{1/2}) \\ &= \mathrm{rank}(\boldsymbol{I}-\boldsymbol{L}) \\ &= M-N_L. \end{aligned} \tag{3.54}$$

Thus, the matrix $\boldsymbol{Q}$ has $(M - N_L)$ non-zero eigenvalues and the corresponding eigenvectors. □

3.7.2 Proof of Lemma 3

Pre-multiplying both sides of (3.34) by $(\boldsymbol{I} - \boldsymbol{L})^T$, we obtain

$$(\boldsymbol{I} - \boldsymbol{L})^2 \boldsymbol{R}_{\boldsymbol{ff}} (\boldsymbol{I} - \boldsymbol{L})^T \boldsymbol{\phi}_i = \mu_i^2 (\boldsymbol{I} - \boldsymbol{L}) \boldsymbol{\phi}_i. \tag{3.55}$$

Since $\boldsymbol{L}$ is a projection matrix, we have $(\boldsymbol{I} - \boldsymbol{L})^2 = \boldsymbol{I} - \boldsymbol{L}$. Therefore,

$$\begin{aligned} (\boldsymbol{I} - \boldsymbol{L}) \boldsymbol{R}_{\boldsymbol{ff}} (\boldsymbol{I} - \boldsymbol{L})^T \boldsymbol{\phi}_i &= \mu_i^2 (\boldsymbol{I} - \boldsymbol{L}) \boldsymbol{\phi}_i \\ &= \mu_i^2 \boldsymbol{\phi}_i - \mu_i^2 \boldsymbol{L} \boldsymbol{\phi}_i, \end{aligned} \tag{3.56}$$

which implies that $\boldsymbol{L}\boldsymbol{\phi}_i = \boldsymbol{0}$. □

Chapter 4

Adaptive Lapped Transforms for Image Coding

4.1 Introduction

The use of adaptation has resulted in significant improvements in both compression ratio and visual quality around edges. However, as long as we use block-transform coding, decoded images cannot be free from annoying *blocking effects* at low bit rates. By using wavelets or lapped transforms [13, 10], we can obtain reconstructed images without the blocking artifacts; however, they are usually blurred especially around edges. In the wavelet case, moreover, the construction of locally adaptive basis functions is very difficult since "double-shift" orthogonality/biorthogonality is required for perfect reconstruction. Also, lapped transforms (LT) are powerful tools for the reduction of the blocking artifacts in image compression [8, 15, 16, 36, 22]. The blocking artifacts are reduced by overlapping basis functions of which size is larger than the block size. It is however difficult to construct a space-varying LT [78] since the LT has a strong constraint such that the overlapping parts of the basis functions must be orthogonal or biorthogonal. Several studies on a space-varying LT have been conducted [48, 49]. However, they are one-dimensional transforms; therefore, difficulty in design of two-dimensional adaptive LTs such as orientation adaptive ones remains, since the degree of freedom extensively increases in the 2D case.

In this chapter, we propose an adaptive lapped biorthogonal transform for image coding. In Section 4.2, we formulate a general form of a one-dimensional lapped biorthogonal transform with variable length basis functions (VLLBT) for two cases: 1) Length of overlapping basis functions is an even-multiple of the block size; 2) the length is an odd-multiple. The former

is a generalization of the VLLBT dealt with in Chapter 3. In the field of filter banks, a linear-phase version of the VLLBT has been introduced [19], but we provide a vector-matrix form, which is easily extended to a 2-D non-separable form and is not limited to linear-phase. The SKLT introduced in Chapter 3 enables us to find the optimal non-overlapping basis functions without numerical searching. In Section 4.4, we show an orientation adaptation example. We construct a transform matrix of the VLLBT with respect to a class characterized by the angle of edges. The resulting VLLBT is two-dimensional and will be called an orientation adaptive lapped biorthogonal transform (OALBT). In Section 4.5, we show a method for designing the long basis functions and construct the OALBT.

4.2 Lapped Biorthogonal Transforms with Overlapping Basis Functions

Before derivation for the proposed adaptive lapped transform, we formulate in this section a generalized form of a lapped biorthogonal transform with overlapping and non-overlapping basis functions. A vector-matrix form will be utilized for the formulation.

4.2.1 Formulation

Consider two matrices $\boldsymbol{E} = \begin{bmatrix} \boldsymbol{E}_0^T & \boldsymbol{E}_1^T & \cdots & \boldsymbol{E}_{K-1}^T \end{bmatrix}^T$ and $\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_0^T & \boldsymbol{R}_1^T & \cdots & \boldsymbol{R}_{K-1}^T \end{bmatrix}^T$ of size $KM \times M$, which will be called a forward and an inverse transform matrices of a lapped biorthogonal transform (LBT), respectively. Keep in mind that for $i = 0, \ldots, K-1$, $\boldsymbol{E}_i$ and $\boldsymbol{R}_i$ are $M \times M$ matrices. Columns of $\boldsymbol{E}$ and $\boldsymbol{R}$ are basis functions. Let $\boldsymbol{f}_i$ be the i-th block with M samples of an input signal. Then, the transform vector $\boldsymbol{g}_i$ is obtained by

$$\tilde{\boldsymbol{f}}_i = \sum_{k=1}^{K} \boldsymbol{E}_{k-1}^T \boldsymbol{f}_{i-K+k}, \tag{4.1}$$

and the reconstructed block $\hat{\boldsymbol{f}}_i$ is obtained by

$$\hat{\boldsymbol{f}}_i = \sum_{k=1}^{K} \boldsymbol{R}_{k-1} \tilde{\boldsymbol{f}}_{i+K-k}. \tag{4.2}$$

For perfect reconstruction $\boldsymbol{f}_i = \hat{\boldsymbol{f}}_i$, the generalized LBT has to satisfy the following conditions [13]:

$$\sum_{k=0}^{K-1} \boldsymbol{R}_k^T \boldsymbol{E}_k = \sum_{k=0}^{K-1} \boldsymbol{R}_k \boldsymbol{E}_k^T = \boldsymbol{I}_M, \tag{4.3}$$

$$\sum_{k=0}^{K-1-s} \boldsymbol{R}_k^T \boldsymbol{E}_{k+s} = \sum_{k=0}^{K-1-s} \boldsymbol{R}_k \boldsymbol{E}_{k+s}^T = \boldsymbol{0}_M, \tag{4.4}$$

where $s = 1, \ldots, K-1$. Equation (4.3) describes biorthogonality of basis functions, and (4.4) means that the overlapping functions of neighboring blocks must also be biorthogonal.

Let us introduce two matrices $\hat{\boldsymbol{E}}_k$ and $\bar{\boldsymbol{E}}_k$ which contain the first N_L and the remaining $(M-N_L)$ column vectors of $\boldsymbol{E}_k$, respectively, that is,

$$\boldsymbol{E}_k = \left[\hat{\boldsymbol{E}}_k \;\; \bar{\boldsymbol{E}}_k\right], \quad k = 0, \ldots, K-1. \tag{4.5}$$

1. Type-E (K is even)

   Assume that K is even. Setting that $\bar{\boldsymbol{E}}_k = \boldsymbol{0}_M$ except for $k = (K-1)/2$, $\boldsymbol{E}$ becomes

   $$\boldsymbol{E} = \begin{bmatrix} \hat{\boldsymbol{E}}_0 & \hat{\boldsymbol{E}}_1 & \cdots & \hat{\boldsymbol{E}}_{K-1} \\ \bar{\boldsymbol{E}} & \boldsymbol{0}_M & \cdots & \boldsymbol{0}_M \end{bmatrix}, \tag{4.6}$$

   where $\bar{\boldsymbol{E}}_0$ is written as $\bar{\boldsymbol{E}}$ for simplicity. In this case, $\boldsymbol{E}$ is a forward transform such that N_L basis functions have length KM and the remaining $(M - N_L)$ functions have length M. The former will be called the *long basis functions*, and the latter will be called the *short basis functions*. In the same manner, we define an inverse transform as follows:

   $$\boldsymbol{R} = \begin{bmatrix} \hat{\boldsymbol{R}}_0 & \bar{\boldsymbol{R}} \\ \hat{\boldsymbol{R}}_1 & \boldsymbol{0}_M \\ \vdots & \vdots \\ \hat{\boldsymbol{R}}_{K-1} & \boldsymbol{0}_M \end{bmatrix}. \tag{4.7}$$

   The columns of $\hat{\boldsymbol{E}}$ and those of $\hat{\boldsymbol{R}}$ correspond to the long basis functions of the forward and the inverse transforms, respectively. Similarly, $\bar{\boldsymbol{E}}$ and $\bar{\boldsymbol{R}}$ contain the short basis

   Note that in this structure, if all basis functions are symmetric or antisymmetric, then the basis functions' center of symmetry are not aligned. This fact needs a special treatment at image boundaries when the transform is applied.

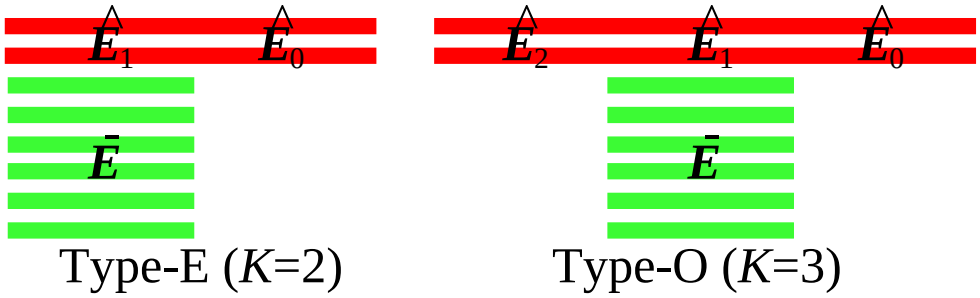


Figure 4.1: Images for transform matrices of Type-E and Type-O

2. Type-O (K is odd)

   Assume that K is odd. Setting that $\bar{\boldsymbol{E}}_k = \mathbf{0}_M$ except for $k = (K-1)/2$, $\boldsymbol{E}$ becomes

$$\boldsymbol{E} = \begin{bmatrix} \hat{\boldsymbol{E}}_0 & \cdots & \hat{\boldsymbol{E}}_{(K-1)/2} & \cdots & \hat{\boldsymbol{E}}_{K-1} \\ \mathbf{0}_M & \cdots & \bar{\boldsymbol{E}} & \cdots & \mathbf{0}_M \end{bmatrix}, \tag{4.8}$$

   where $\bar{\boldsymbol{E}}_{(K-1)/2}$ is written as $\bar{\boldsymbol{E}}$ for simplicity. In the same manner, we define an inverse transform as follows:

$$\boldsymbol{R} = \begin{bmatrix} \hat{\boldsymbol{R}}_0 & \mathbf{0}_M \\ \vdots & \vdots \\ \hat{\boldsymbol{R}}_{(K-1)/2} & \bar{\boldsymbol{R}} \\ \vdots & \vdots \\ \hat{\boldsymbol{R}}_{K-1} & \mathbf{0}_M \end{bmatrix}. \tag{4.9}$$

   Since K is odd, if all basis functions are symmetric or antisymmetric, then the basis functions' center of symmetry are aligned. Therefore, no special processing at image boundaries is needed.

Figure 4.1 illustrates images of Type-E and Type-O.

In the context of linear-phase perfect reconstruction filter banks, this transform can be regarded as particular cases of the generalized lapped biorthogonal transform with variable length functions [19]. We will use the term variable-length lapped transform (VLLT) to refer to the proposed transform. Especially, if the basis functions are orthogonal, the transform is called the VLLOT, and if the basis functions are biorthogonal, the transform is called the VLLBT. ("O" and "B" indicate orthogonal and biorthogonal, respectively.)

In the VLLT, the short basis functions never overlap across block boundaries, so that they enable us to easily construct adaptive lapped transforms. For both Type-E and Type-O VLLTs, inserting $\boldsymbol{E}$ and $\boldsymbol{R}$ into the LBT conditions (4.3) and (4.4), we obtain the following common conditions:

Condition 1 *The long basis functions of the VLLBT are required to be biorthogonal, and the overlapping parts of the basis functions of neighboring blocks must also be biorthogonal:*

$$\sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k^T \hat{\boldsymbol{E}}_k^T = \boldsymbol{I}_{N_L}, \tag{4.10}$$

$$\sum_{k=0}^{K-1-s} \hat{\boldsymbol{R}}_k^T \hat{\boldsymbol{E}}_{k+s}^T = \sum_{k=0}^{K-1-s} \hat{\boldsymbol{R}}_{k+s}^T \hat{\boldsymbol{E}}_k^T = \boldsymbol{0}_K, \quad s = 1, \ldots, K-1. \tag{4.11}$$

Furthermore, the following condition for overlapping parts is required:

$$\sum_{k=0}^{K-1-s} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_{k+s} = \sum_{k=0}^{K-1-s} \hat{\boldsymbol{R}}_{k+s} \hat{\boldsymbol{E}}_k = \boldsymbol{0}_M, \quad s = 1, \ldots, K-1. \tag{4.12}$$

Condition 2 *For the short basis functions, biorthogonality is required:*

$$\overline{\boldsymbol{R}}^T \overline{\boldsymbol{E}}^T = \boldsymbol{I}_{M-N_L}. \tag{4.13}$$

The long and short basis functions must have the following relation:

$$\sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_k + \overline{\boldsymbol{R}}\,\overline{\boldsymbol{E}} = \boldsymbol{I}_M. \tag{4.14}$$

If long basis functions meet Condition 1, we say that they are *feasible*.

4.3 Derivation of Short Basis Functions via the SKLT

Once $\hat{\boldsymbol{E}}$ and $\hat{\boldsymbol{R}}$ satisfy Condition 1 ((4.10), (4.11), and (4.12)), the short functions that fulfill Condition 2 ((4.13) and (4.14)) can be easily constructed by the SKLT introduced in Chapter 3. Let $\overline{\boldsymbol{E}}^T$ and $\overline{\boldsymbol{R}}$ be $M \times (M - N_L)$ matrices whose columns correspond to short basis functions. The following result leads to the optimal short basis functions in the sense of minimizing the criterion (3.20):

Proposition 2 $\sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_k$ *is a projection matrix of rank* K.

Proof appears in Section 4.7.1.

Proposition 2 guarantees data compression ability of short basis functions if suitable long basis functions are given. From Theorem 2, consequently, by setting that

$$\boldsymbol{L} = \sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_k \tag{4.15}$$

in (3.21), we obtain a biorthogonal system $\{\boldsymbol{\phi}_i, \boldsymbol{\phi}_i^*\}_{i=0}^{M-N_L-1}$, which leads to the optimal short basis functions. Let $\bar{\boldsymbol{e}}_i$ and $\bar{\boldsymbol{r}}_i$ be the i-th columns of $\bar{\boldsymbol{E}}^T$ and $\bar{\boldsymbol{R}}$.

Proposition 3 *If we set that*

$$\bar{\boldsymbol{e}}_i = \boldsymbol{\phi}_i^*, \tag{4.16}$$

and

$$\bar{\boldsymbol{r}}_i = \boldsymbol{\phi}_i, \tag{4.17}$$

the resulting VLLBT achieves perfect reconstruction, that is, $\bar{\boldsymbol{E}}$ and $\bar{\boldsymbol{R}}$ satisfies Condition 2.

Proof appears in Section 4.7.2.

4.3.1 Design Algorithm

The steps to design the short basis functions can be summarized as follows:

1. Choose a set of the feasible long basis functions

   $$\hat{\boldsymbol{E}} = \begin{bmatrix} \hat{\boldsymbol{E}}_0 & \cdots & \hat{\boldsymbol{E}}_{N_L-1} \end{bmatrix}$$

   and

   $$\hat{\boldsymbol{R}} = \begin{bmatrix} \hat{\boldsymbol{R}}_0^T & \cdots & \hat{\boldsymbol{R}}_{N_L-1}^T \end{bmatrix}^T.$$

2. Obtain the projection matrix $\boldsymbol{L}$.

3. Obtain a correlation matrix $\boldsymbol{R}_{\boldsymbol{ff}}$ of the signal.

4. From Lemma 1, obtain $(M - N_L)$ eigenvectors such that the corresponding eigenvalues are not zero.

5. Obtain the forward short basis functions from (4.16), and obtain the inverse short basis functions from (4.17).

4.4 Orientation Adaptation

4.4.1 Extension to 2-D transform

So far, it has been assumed that we are dealing with one-dimensional transforms. In the case of real-valued transforms, however, orientation adaptive transforms have the two-dimensional non-separable form. Fortunately, non-separable two-dimensional transforms can be easily constructed using the above algorithm when both a tensor product of the orthogonal projection matrix generated by the long basis functions and a two-dimensional correlation matrix are given. In the rest of this chapter, every $M \times M$ image block $\boldsymbol{f}$ is treated as an $M^2 \times 1$ vector in the space $\mathbb{R}^{M^2}$ by lexicographic ordering. In this case, as illustrated in Fig. 4.2, the "long" basis functions have "length" of $(KM)^2$. Similarly, the "short" ones have "length" of M^2. (We will use terms for the 1-D case such as "long," "short," and "length.") Let

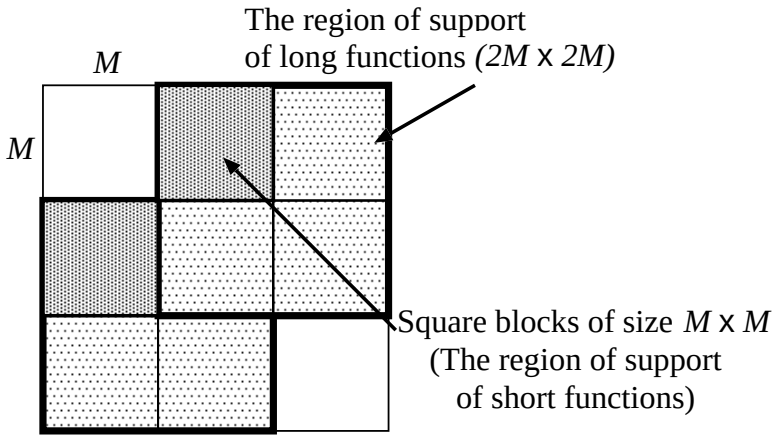
$$\boldsymbol{L} = \left(\sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_k \right) \otimes \left(\sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k \hat{\boldsymbol{E}}_k \right), \tag{4.18}$$

where $\otimes$ denotes the Kronecker's tensor product [50]. It can then be easily checked that $\boldsymbol{L}$ is still a projection matrix of rank N_L^2, and size $M^2 \times M^2$. By applying the result of Theorem 2, we obtain the $(M^2 - N_L^2)$ 2-D optimal short basis functions $\overline{\boldsymbol{e}}_i$ and $\overline{\boldsymbol{r}}_i$ of size M^2 for $i = 0, \ldots, M^2 - N_L^2 - 1$. The size of $\overline{\boldsymbol{E}}^T$ and $\overline{\boldsymbol{R}}$ is therefore $M^2 \times (M^2 - N_L^2)$.

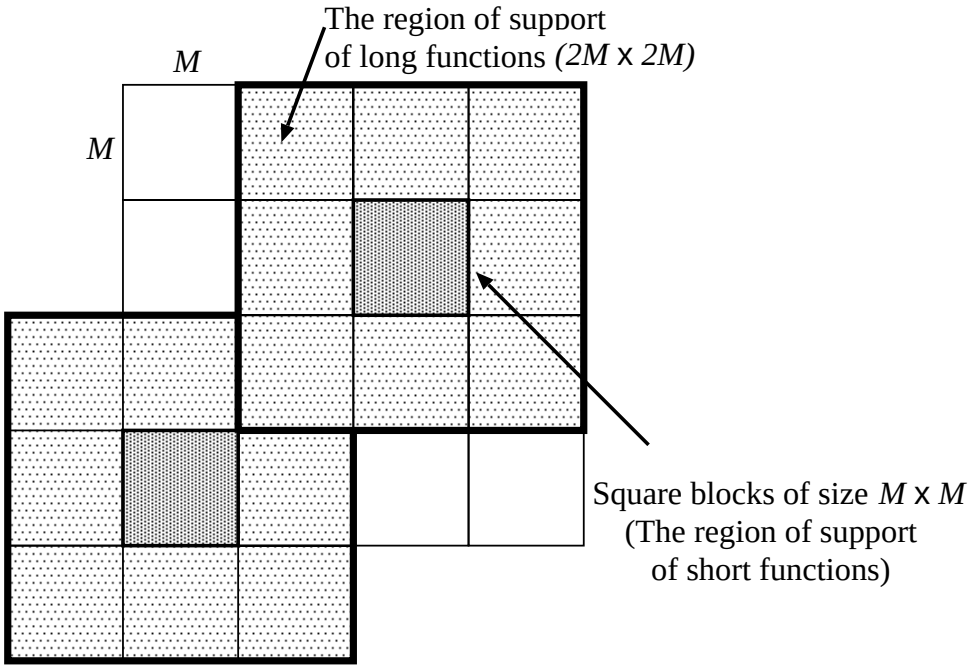
4.4.2 Orientation Adaptive Lapped Transforms

Since the short basis functions never overlap across block boundaries, they can easily generate space-varying lapped transforms via the SKLT. Let θ be a parameter with respect to the angle of an edge in each block. To simplify matters, the variation θ is discretized into J levels such that $\theta_j = -\frac{\pi}{2} + \frac{\pi}{J} j$ for $j = 0, \ldots, J-1$. For $j = 0, \ldots, J-1$, we define the class C_j which is characterized by the corresponding parameter θ_j. Image blocks are then classified into one of the classes $\{C_j\}_{j=0}^{J}$. The class C_J is introduced for image blocks which do not belong to C_j for $j = 0, \ldots, J-1$. We define the corresponding correlation matrix set $\{\boldsymbol{R}_{\boldsymbol{ff}}^{(j)}\}_{j=0}^{J-1}$, where $\boldsymbol{R}_{\boldsymbol{ff}}^{(j)}$ is an $M^2 \times M^2$ matrix of the "directional" Markov model [43] defined as

$$\begin{aligned} [\boldsymbol{R}_{\boldsymbol{ff}}^{(j)}]_{p+qM,p'+q'M} &= \alpha^{|d_x(\theta_j)|} \cdot \beta^{|d_y(\theta_j)|}, \\ & p, p', q, q' = 0, \ldots, M-1, \end{aligned} \tag{4.19}$$



(a) Type-E



(b) Type-O

Figure 4.2: Support of regions of long and short basis functions in the 2-D case

where α and β are the correlation coefficients and $d_x(\theta_j)$ and $d_y(\theta_j)$ are defined as

$$\begin{bmatrix} d_x(\theta_j) \\ d_y(\theta_j) \end{bmatrix} = \begin{bmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{bmatrix} \begin{bmatrix} p-p' \\ q-q' \end{bmatrix}. \tag{4.20}$$

The VLLT resulting from $\boldsymbol{L}$ and $\boldsymbol{R}_{\boldsymbol{ff}}^{(j)}$ will be called the *orientation adaptive lapped transform* (OALBT) for the direction θ_j. We write the short basis functions of the OALBT by a pair of $M^2 \times (M^2 - N_L^2)$ matrices $\bar{\boldsymbol{E}}^{(j)} = [\bar{\boldsymbol{e}}_0^{(j)}, \ldots, \bar{\boldsymbol{e}}_{M^2-N_L^2-1}^{(j)}]^T$ and $\bar{\boldsymbol{R}}^{(j)} = [\bar{\boldsymbol{r}}_0^{(j)}, \ldots, \bar{\boldsymbol{r}}_{M^2-N_L^2-1}^{(j)}]$. Similarly, we introduce a non-directional correlation matrix $\boldsymbol{R}_{\boldsymbol{ff}}^{(J)}$ with respect to the class C_J given by

$$\begin{aligned} (\boldsymbol{R}_{\boldsymbol{ff}}^{(J)})_{p+qM,p'+q'M} &= \rho^{\sqrt{|p-p'|^2+|q-q'|^2}}, \\ p, p', q, q' &= 0, \ldots, M-1, \end{aligned} \tag{4.21}$$

where ρ is the correlation coefficient. It generates the non-adaptive 2D VLLT with the optimal short basis functions (we will call this non-adaptive version just "VLLT" to distinguish from the OALBT in encoders). Its short functions are included in $\{\bar{\boldsymbol{E}}^{(J)}, \bar{\boldsymbol{R}}^{(J)}\}$ as columns.

4.5 Design Examples

4.5.1 Type-E OALBT

We consider the case that the block size is 8×8 ($M = 8$) and the length of long functions is 16 ($L = 2$). We use two long basis functions ($N_{L1} = N_{L2} = 2$) such that the transform matrices contain the most number of short functions. Therefore, the block signal $\boldsymbol{f}$ is in $\mathbb{R}^{64}$. This choice leads to four 2-D long functions and sixty short functions.

We choose two long functions from the Malvar's lapped biorthogonal transform (LBT) [36], where the particular choices for c generate synthesis basis functions whose asymptotic end values are exactly zero. The matrices containing the long functions are given as

$$\hat{\boldsymbol{E}} = \frac{1}{2}\begin{bmatrix} \boldsymbol{d}^{(0)} - c\boldsymbol{d}^{(1)} & \boldsymbol{d}^{(0)} - c\boldsymbol{d}^{(1)} \\ \boldsymbol{d}^{(0)} + c\boldsymbol{d}^{(1)} & -(\boldsymbol{d}^{(0)} + c\boldsymbol{d}^{(1)}) \end{bmatrix}, \tag{4.22}$$

for analysis, and

$$\hat{\boldsymbol{R}} = \frac{1}{2}\begin{bmatrix} \boldsymbol{d}^{(0)} - \frac{1}{c}\boldsymbol{d}^{(1)} & \boldsymbol{d}^{(0)} - \frac{1}{c}\boldsymbol{d}^{(1)} \\ \boldsymbol{d}^{(0)} + \frac{1}{c}\boldsymbol{d}^{(1)} & -(\boldsymbol{d}^{(0)} + \frac{1}{c}\boldsymbol{d}^{(1)}) \end{bmatrix}, \tag{4.23}$$

Figure 4.3: The first four short basis functions when $\theta = \pi/15$

for synthesis, where $\boldsymbol{d}^{(0)}$ and $\boldsymbol{d}^{(1)}$ are the first and the second basis functions of the M-point DCT, respectively, and c is a positive constant ($c = \sqrt{2}$ is chosen in [36]). This matrix satisfies Condition 1. It is moreover verified that $\hat{\boldsymbol{R}}_0\hat{\boldsymbol{E}}_0 + \hat{\boldsymbol{R}}_1\hat{\boldsymbol{E}}_1$ is a orthogonal projection matrix. The resulting 2-D orthogonal projection matrix $\boldsymbol{L}$ has rank four. Note that these functions are symmetric or antisymmetric, so that the resulting short functions are also symmetric or antisymmetric as shown in Theorem 3. This choice of the long basis functions leads to the fact that $\boldsymbol{R} = \boldsymbol{E}^T$, because the orthogonality of $\hat{\boldsymbol{R}}_0\hat{\boldsymbol{E}}_0 + \hat{\boldsymbol{R}}_1\hat{\boldsymbol{E}}_1$ yields the fact that $\boldsymbol{\phi}_i = \boldsymbol{\phi}_i^*$ when deriving the SKLT. In other words, although the long basis functions are biorthogonal, the short ones are orthogonal. This special choice is done to avoid complexity caused by special treatment at block boundaries as will be stated later. Some of the resulting short basis functions for $\theta = \pi/15$, are illustrated in Fig. 4.3. The correlation coefficients of $\boldsymbol{R}_{ff}^{(j)}$ as in (4.19) is chosen to $\alpha = 0.95$ and $\beta = 0.50$.

Special Processing at Image Boundaries Note that to process finite length images, special care is required for the image boundary to avoid the border distortion. Boundary functions have length $(M + M/2)$, that is, twelve in this case, and these are given as [8, 71]

$$\hat{\boldsymbol{E}}^{(\text{left})} = \frac{1}{2}\begin{bmatrix} 2\boldsymbol{d}_e^{(0)} & 2\boldsymbol{d}_e^{(0)} \\ \boldsymbol{d}^{(0)} + c\boldsymbol{d}^{(1)} & -(\boldsymbol{d}^{(0)} + c\boldsymbol{d}^{(1)}) \end{bmatrix}, \tag{4.24}$$

for left boundaries, and

$$\hat{\boldsymbol{E}}^{(\text{right})} = \frac{1}{2}\begin{bmatrix} \boldsymbol{d}^{(0)} - c\boldsymbol{d}^{(1)} & \boldsymbol{d}^{(0)} - c\boldsymbol{d}^{(1)} \\ 2\boldsymbol{J}_4\boldsymbol{d}_e^{(0)} & -2\boldsymbol{J}_4\boldsymbol{d}_e^{(0)} \end{bmatrix}, \tag{4.25}$$

for right boundaries. The function $\boldsymbol{d}_e^{(0)}$ contains the first DCT function as

$$\boldsymbol{d}^{(0)} = \begin{bmatrix} \boldsymbol{d}_e^{(0)} \\ \boldsymbol{J}_4\boldsymbol{d}_e^{(0)} \end{bmatrix}. \tag{4.26}$$

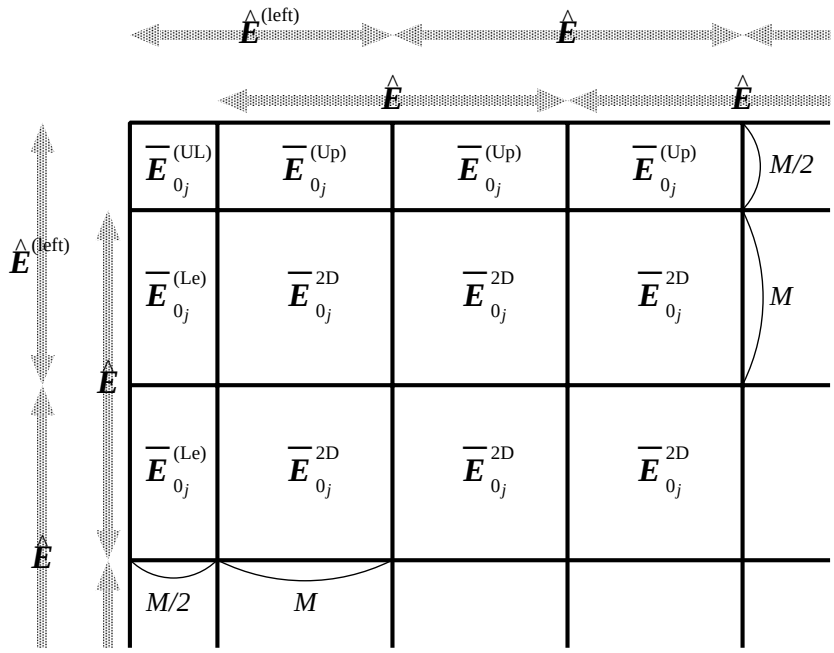


Figure 4.4: The block partition and short basis functions with respect to each block

We can also obtain $\hat{\boldsymbol{R}}^{(\text{left})}$ and $\hat{\boldsymbol{R}}^{(\text{right})}$ by replacing $c\boldsymbol{d}^{(1)}$ with $\frac{1}{c}\boldsymbol{d}^{(1)}$ in (4.24) and (4.25). Then, the orthogonal projection matrices $\boldsymbol{L}^{(\text{left})}$ and $\boldsymbol{L}^{(\text{right})}$ are given by

$$\boldsymbol{L}^{(\text{left})} = 2\boldsymbol{d}_e^{(0)}\boldsymbol{d}_e^{(0)^T} \tag{4.27}$$

$$\boldsymbol{L}^{(\text{right})} = 2\boldsymbol{J}_4\boldsymbol{d}_e^{(0)}\boldsymbol{d}_e^{(0)^T}\boldsymbol{J}_4. \tag{4.28}$$

This implies that the size of blocks at image boundaries is no longer 8×8 (see Fig. 4.4). For instance, $\boldsymbol{L}^{(\text{left})} \otimes \boldsymbol{L}^{(\text{left})}$ generates the 15 short basis functions for the upper left block of size 4×4. These are denoted by the matrix form $\overline{\boldsymbol{E}}^{(\text{UL})}_{0_j}$, where UL is initials of "upper left". Let us show another example in a similar manner. The orthogonal projection matrix $\boldsymbol{L}^{(\text{right})} \otimes \boldsymbol{L}_1$, whose rank is two, yields $\overline{\boldsymbol{E}}^{(\text{Lo})}_{0_j}$ for the lower blocks of size 8×4. The matrix $\overline{\boldsymbol{E}}^{(\text{Lo})}_{0_j}$ has size 32×30 and contains the 30 short basis functions. The relations between the 2-D orthogonal projection matrices and the corresponding short basis functions are summarized in Table 4.1 in our case, that is, $\boldsymbol{L}_1 = \boldsymbol{L}_2$ and $N_{L1} = 2$.

Table 4.1: The short basis functions for boundary blocks: OPM and # of SBF denote the 2-D orthogonal projection matrix and the number of the 2-D short basis functions, respectively.

Block Type	Size	OPM	Rank	# of SBF	Notation
Inside (normal)	8×8	$\boldsymbol{L}_1 \otimes \boldsymbol{L}_1$	4	60	$\overline{\boldsymbol{E}}_{0_j}^{\mathrm{2D}}$
Upper Left	4×4	$\boldsymbol{L}^{(\mathrm{left})} \otimes \boldsymbol{L}^{(\mathrm{left})}$	1	15	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{UL})}$
Upper Right	4×4	$\boldsymbol{L}^{(\mathrm{right})} \otimes \boldsymbol{L}^{(\mathrm{left})}$	1	15	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{UR})}$
Lower Left	4×4	$\boldsymbol{L}^{(\mathrm{left})} \otimes \boldsymbol{L}^{(\mathrm{right})}$	1	15	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{LL})}$
Lower Right	4×4	$\boldsymbol{L}^{(\mathrm{right})} \otimes \boldsymbol{L}^{(\mathrm{right})}$	1	15	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{LR})}$
Upper	4×8	$\boldsymbol{L}^{(\mathrm{left})} \otimes \boldsymbol{L}_1$	2	32	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{Up})}$
Lower	4×8	$\boldsymbol{L}^{(\mathrm{right})} \otimes \boldsymbol{L}_1$	2	32	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{Lo})}$
Left	8×4	$\boldsymbol{L}_1 \otimes \boldsymbol{L}^{(\mathrm{left})}$	2	32	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{Le})}$
Right	8×4	$\boldsymbol{L}_1 \otimes \boldsymbol{L}^{(\mathrm{right})}$	2	32	$\overline{\boldsymbol{E}}_{0_j}^{(\mathrm{Ri})}$

4.5.2 Type-O OALBT

Consider the case that the basis functions of the VLLBT are all symmetric/antisymmetric, that is, linear-phase. In the Type-E, since $K = 2$ was chosen, the symmetric extension method [79, 80] could not be used, and special care was required for the image boundary in order to avoid the border distortion.

The long basis functions are produced by the method like the GenLOT [15]: Let $\boldsymbol{h}_i$, $i = 0, \ldots, K-1$ be linear independent vectors of size M such that $\boldsymbol{h}_i$ is symmetric if i is even, and antisymmetric otherwise. Define

$$\hat{\boldsymbol{H}}_0 = \begin{bmatrix} \hat{\boldsymbol{H}}_e \\ \hat{\boldsymbol{H}}_o \end{bmatrix}, \boldsymbol{\Lambda}_i = \begin{bmatrix} \boldsymbol{U}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_i \end{bmatrix}, \boldsymbol{W} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_{K/2} & \boldsymbol{I}_{K/2} \\ \boldsymbol{I}_{K/2} & -\boldsymbol{I}_{K/2} \end{bmatrix}, \tag{4.29}$$

where $\hat{\boldsymbol{H}}_e$ and $\hat{\boldsymbol{H}}_o$ are $K/2 \times M$ matrices consisting of rows which are $\boldsymbol{h}_i^T$ for even i and $\boldsymbol{h}_i^T$ for odd i, respectively, and $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ are non-singular matrices of size $K/2 \times K/2$. Then, $\hat{\boldsymbol{E}}$ can be found from the following recursion:

$$\hat{\boldsymbol{H}}_i = \boldsymbol{\Lambda}_{i-1} \boldsymbol{W} \begin{bmatrix} \boldsymbol{I}_{K/2} & \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} \\ \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} & \boldsymbol{I}_{K/2} \end{bmatrix} \begin{bmatrix} \boldsymbol{W}\hat{\boldsymbol{H}}_{i-1} & \boldsymbol{0}_{K\times M} \\ \boldsymbol{0}_{K\times M} & \boldsymbol{W}\hat{\boldsymbol{H}}_{i-1} \end{bmatrix} \tag{4.30}$$

$$\hat{\boldsymbol{E}} = \hat{\boldsymbol{H}}_{L-1}. \tag{4.31}$$

The corresponding inverse long basis function matrix $\hat{\boldsymbol{R}}$ can be obtained by substituting $\boldsymbol{\Lambda}_i$ for

$\boldsymbol{\Lambda}_i^{-1}$ and $\hat{\boldsymbol{H}}_0$ for $\hat{\boldsymbol{G}}_0$ such that $\hat{\boldsymbol{G}}_0\hat{\boldsymbol{H}}_0 = \boldsymbol{I}_K$ in (4.30) and (4.31). Specifically, $\hat{\boldsymbol{R}}$ is written as

$$\hat{\boldsymbol{G}}_i = \begin{bmatrix} \hat{\boldsymbol{G}}_{i-1}\boldsymbol{W} & \boldsymbol{0}_{M\times K} \\ \boldsymbol{0}_{M\times K} & \hat{\boldsymbol{G}}_{i-1}\boldsymbol{W} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{K/2} & \boldsymbol{0}_{K/2} \\ \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} \\ \boldsymbol{0}_{K/2} & \boldsymbol{0}_{K/2} \\ \boldsymbol{0}_{K/2} & \boldsymbol{I}_{K/2} \end{bmatrix} \boldsymbol{W}\boldsymbol{\Lambda}_{i-1}^{-1}, \tag{4.32}$$

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{G}}_{L-1}. \tag{4.33}$$

These $\hat{\boldsymbol{E}}$ and $\hat{\boldsymbol{R}}$ meet the conditions (4.10), (4.11), and (4.12). As a result, the free parameters which we should find are $\hat{\boldsymbol{H}}_0$, $\hat{\boldsymbol{G}}_0$, and $\boldsymbol{\Lambda}_i (i = 0, \ldots, K-1)$.

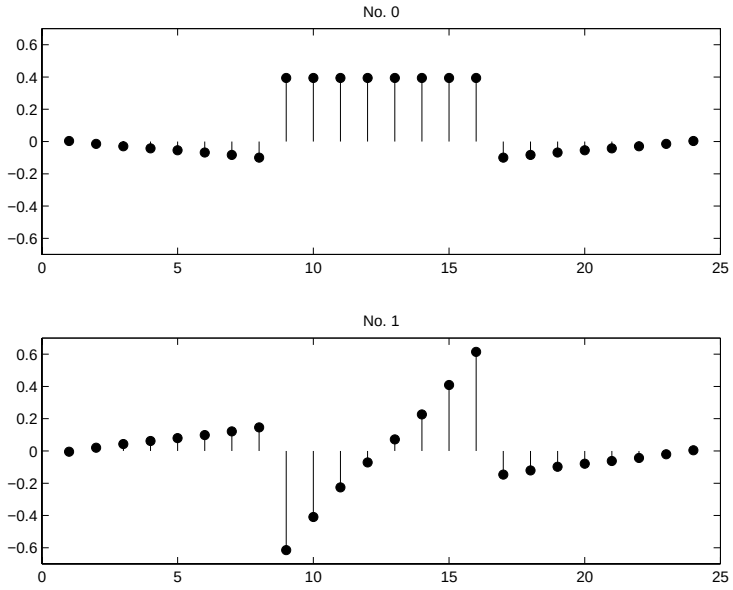
For application in image coding, we use coding gain and low DC leackage as cost functions. Finding the long basis functions is formulated as a non-linear optimization problem:

$$\underset{\hat{\boldsymbol{H}}_0, \hat{\boldsymbol{G}}_0, \boldsymbol{U}_i, \boldsymbol{V}_i}{\text{Minimize}} \quad -J_{CG} + \mu J_{DC}, \tag{4.34}$$

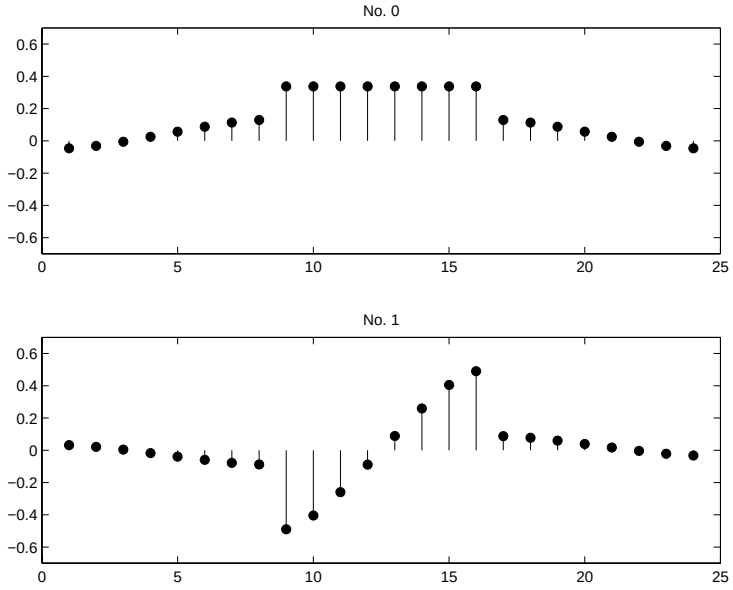
subject to $\hat{\boldsymbol{H}}_0\hat{\boldsymbol{G}}_0 = \boldsymbol{I}_{N_L}$, where μ is a weight for the combination of two cost functions. In every step of the optimization, the short basis functions are found via the SKLT with a correlation matrix of the Markov model, as expressed in Section 4.3, and then the cost function (4.34) into which $\boldsymbol{E}$ and $\boldsymbol{R}$ are substituted is evaluated.

In our test, we choose that $M = 8$, $L = 3$, and $N_L = 2$; there are two sets of long basis functions $\{\hat{\boldsymbol{e}}_0, \hat{\boldsymbol{e}}_1\}$ and $\{\hat{\boldsymbol{r}}_0, \hat{\boldsymbol{r}}_1\}$, where each function has length of 24. The case $N_L = 2$ gives the minimum number of the long basis functions because of the existing condition as described in [18]. In other words, this case gives the maximum number of adaptive basis functions. This may be a good choice for adaptive image coding purpose. In this case, $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ are scalars, and both $\hat{\boldsymbol{H}}_0^T$ and $\hat{\boldsymbol{G}}_0$ consist of two vectors of size M. This optimization problem can be reduced to an unconstrained optimization problem as described in Chapter 3. The resulting long basis functions are illustrated in Fig. 4.5. It is interesting that the inverse long basis functions decay to zero at their ends, even though we never impose any constraints for decay on the long basis functions. This property is effective to reduce the blocking effects.

Since the long basis functions have been already obtained, the short basis functions are automatically found as described in Section 4.4. By using $\boldsymbol{L}$ given as in (4.18), we can obtain the 2-D short basis functions with respect to the angle θ_j. For the correlation coefficients of $\boldsymbol{R}_{ff}^{(j)}$, we choose that $\alpha = 0.95$ and $\beta = 0.50$. As an example, the first eight short functions of the OALBT with respect to the angle $\theta_j = 2\pi/15$ are illustrated in Fig. 4.6.

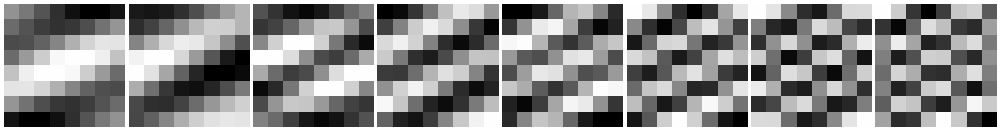


(a) The long basis functions of the forward transform

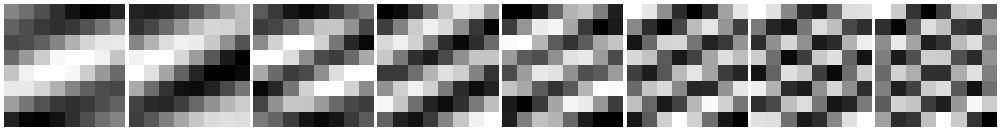


(b) The long basis functions of the inverse transform

Figure 4.5: The resulting long basis functions



(a) Short basis functions of the forward transform



(b) Short basis functions of the inverse transform

Figure 4.6: First eight short basis functions of the OALBT for $\theta = 2\pi/15$ (the correlation coefficients are set that $\alpha = 0.95$ and $\beta = 0.50$)

4.6 Summary

A novel adaptive lapped biorthogonal transform and its application in orientation adaptive coding have been proposed. The proposed transform consists of overlapping and non-overlapping basis functions, where the basis functions' centers of symmetry are aligned, so that we can treat the image boundaries without special processing. To construct non-overlapping basis functions, we have also introduced a transform that provides the optimal approximation of an original signal in a given subspace. In the encoder, an image block is selectively transformed by one of orientation adaptive transforms.

We have shown that the non-overlapping basis functions can be obtained from a correlation matrix by solving the eigenvalue problem. Any conventional adaptation procedures based on the Karhunen-Loève transform can be applied to the proposed design method of lapped transforms. Therefore, the proposed framework is powerful for the design of adaptive transforms without the blocking effect.

4.7 Proofs

4.7.1 Proof of Proposition 2

Let us introduce the transform matrices given by

$$\hat{\boldsymbol{T}}_f = \begin{bmatrix} \ddots & & & & & & \\ & \hat{\boldsymbol{E}}_{L-1} & \hat{\boldsymbol{E}}_{L-2} & \cdots & \hat{\boldsymbol{E}}_0 & & 0 \\ & & \hat{\boldsymbol{E}}_{L-1} & \cdots & \hat{\boldsymbol{E}}_1 & \hat{\boldsymbol{E}}_0 & \\ 0 & & & & & & \ddots \end{bmatrix}, \tag{4.35}$$

and

$$\hat{\boldsymbol{T}}_i = \begin{bmatrix} \ddots & & & 0 \\ & \hat{\boldsymbol{R}}_{L-1} & & \\ & \hat{\boldsymbol{R}}_{L-2} & \hat{\boldsymbol{R}}_{L-1} & \\ & \vdots & \vdots & \\ & \hat{\boldsymbol{R}}_1 & \hat{\boldsymbol{R}}_1 & \\ & & \hat{\boldsymbol{R}}_0 & \\ 0 & & & \ddots \end{bmatrix}, \tag{4.36}$$

Before we prove Proposition 2, we show the following lemma:

Lemma 5 *The matrix* $\boldsymbol{P}$ *defined by*

$$\boldsymbol{P} = \hat{\boldsymbol{T}}_i \hat{\boldsymbol{T}}_f \tag{4.37}$$

is a projection matrix.

Proof: We have

$$\boldsymbol{P}^2 = \hat{\boldsymbol{T}}_i \hat{\boldsymbol{T}}_f \hat{\boldsymbol{T}}_i \hat{\boldsymbol{T}}_f = \hat{\boldsymbol{T}}_i \hat{\boldsymbol{T}}_f = \boldsymbol{P}, \tag{4.38}$$

since the columns of $\hat{\boldsymbol{E}}$ and $\hat{\boldsymbol{R}}$ are biorthogonal from (4.10). Therefore, Lemma 5 holds. □

Now, we shall show the proof of Proposition 2.

Proof of Proposition 2: From (4.11), we can rewrite $\boldsymbol{P}$ as

$$\boldsymbol{P} = \begin{bmatrix} \ddots & & & 0 \\ & \hat{\boldsymbol{R}}_{L-1} & & \\ & \hat{\boldsymbol{R}}_{L-2} & \hat{\boldsymbol{R}}_{L-1} & \\ & \vdots & \vdots & \\ & \hat{\boldsymbol{R}}_0 & \hat{\boldsymbol{R}}_1 & \\ & & \hat{\boldsymbol{R}}_0 & \\ 0 & & & \ddots \end{bmatrix} \begin{bmatrix} \ddots & & & & & & \\ & \hat{\boldsymbol{E}}_{L-1} & \hat{\boldsymbol{E}}_{L-2} & \cdots & \hat{\boldsymbol{E}}_0 & & 0 \\ & & \hat{\boldsymbol{E}}_{L-1} & \cdots & \hat{\boldsymbol{E}}_1 & \hat{\boldsymbol{E}}_0 & \\ 0 & & & & & & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} \ddots & & & \\ & \sum_{l=0}^{L-1} \hat{\boldsymbol{R}}_l \hat{\boldsymbol{E}}_l & & 0 \\ & & \sum_{l=0}^{L-1} \hat{\boldsymbol{R}}_l \hat{\boldsymbol{E}}_l & \\ 0 & & & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} \ddots & & & \\ & \boldsymbol{L} & & 0 \\ & & \boldsymbol{L} & \\ 0 & & & \ddots \end{bmatrix}, \tag{4.39}$$

which implies that $\boldsymbol{P}$ is a block diagonal matrix. It follows from Lemma 5 that $\boldsymbol{L}^2 = \boldsymbol{L}$. Therefore, $\boldsymbol{L}$ is a projection matrix.

Next, because $\boldsymbol{L}$ is a projection matrix, we have

$$\mathrm{rank}(\boldsymbol{L}) = \mathrm{tr}(\boldsymbol{L}), \tag{4.40}$$

which can be rewritten as

$$\begin{aligned} \mathrm{rank}(\boldsymbol{L}) &= \mathrm{tr}(\boldsymbol{L}) \\ &= \mathrm{tr}\left(\sum_{l=0}^{L-1} \hat{\boldsymbol{R}}_l \hat{\boldsymbol{E}}_l\right) = \mathrm{tr}\left(\sum_{l=0}^{L-1} \hat{\boldsymbol{R}}_l^T \hat{\boldsymbol{E}}_l^T\right) = \mathrm{tr}[\boldsymbol{I}_{N_L}] = N_L, \end{aligned} \tag{4.41}$$

since (4.10) holds. This completes the proof. □

4.7.2 Proof of Proposition 3

It is evident from Proposition 2 that $\bar{\boldsymbol{E}}$ and $\bar{\boldsymbol{R}}$ satisfy the condition (4.13).

The SKLT $\boldsymbol{X}$ of rank $M - N_L$ is identical to $\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}$, since

$$\boldsymbol{X} = \sum_{n=0}^{M-N_L-1} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^{*T} = \bar{\boldsymbol{R}}\bar{\boldsymbol{E}}. \tag{4.42}$$

From the fact that $\boldsymbol{\phi}_i$ is an eigenvector of $\boldsymbol{Q}$, $\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}$ gives the projection into $R(\boldsymbol{Q})$. Hence, we have

$$\begin{aligned} R(\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}) &\subset R(\boldsymbol{Q}) \\ &= R((\boldsymbol{I} - \boldsymbol{L})\boldsymbol{R}_{\boldsymbol{ff}}(\boldsymbol{I} - \boldsymbol{L})^T) \\ &= R(\boldsymbol{I} - \boldsymbol{L}), \end{aligned} \tag{4.43}$$

since the rank of $\boldsymbol{R}_{\boldsymbol{ff}}$ is full. From Lemma 1, $\mathrm{rank}(\boldsymbol{Q}) = M - N_L$. Therefore, we have

$$R(\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}) = R(\boldsymbol{I} - \boldsymbol{L}). \tag{4.44}$$

On the other hand, because $\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}\boldsymbol{L} = \boldsymbol{0}$, we obtain $N(\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}) \supset R(\boldsymbol{L}) = N(\boldsymbol{I} - \boldsymbol{L})$. However, since the rank of $\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}$ is $M - N_L$, we have

$$N(\bar{\boldsymbol{R}}\bar{\boldsymbol{E}}) = N(\boldsymbol{I} - \boldsymbol{L}). \tag{4.45}$$

From (4.44) and (4.45),

$$\boldsymbol{I} - \boldsymbol{L} = \bar{\boldsymbol{R}}\bar{\boldsymbol{E}}, \tag{4.46}$$

which yields that (4.14) holds. □

Chapter 5

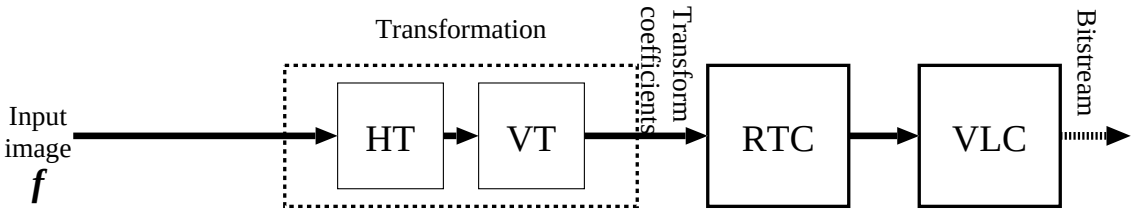
Image Coding Applications and Evaluation

5.1 Introduction

We have so far constructed the various classes of lapped transforms. Those lapped transforms have been designed for applications in image coding. In this chapter, we evaluate those performance by image coding experiments. We use a simple algorithm to encode transform coefficients. We describe details of coding methods, and show rate-distortion characteristics of the coders. Moreover, we also compare subjective qualities of decoded images.

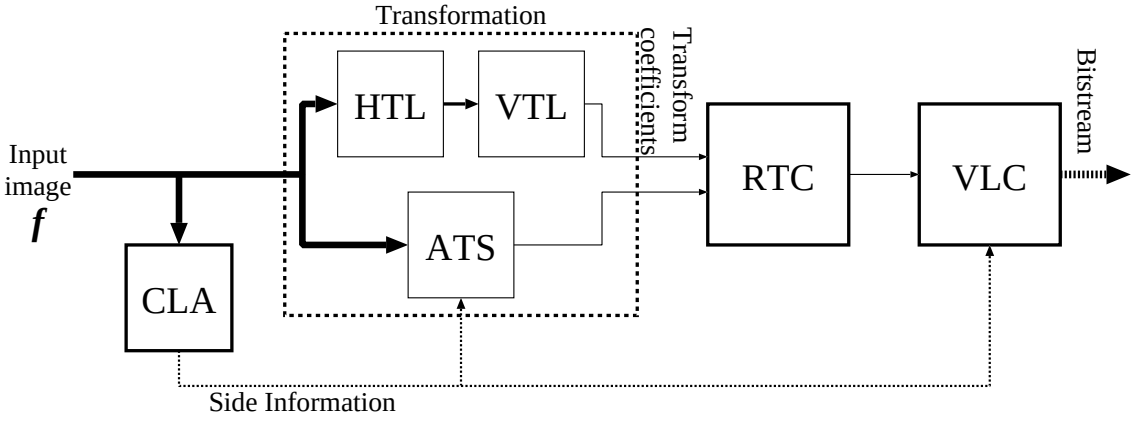
5.2 Image Coding Algorithms

The structures of the encoders used for comparison are illustrated in Fig. 5.1 and 5.2. The encoder shown in Fig. 5.1 has a typical structure as seen in Chapter 2. This is used for non-adaptive transform coding. Recall that this dissertation does not address the "coding step" but the "transform step." Therefore, we will use a simple and well-known algorithm to encode the transform coefficients. If the transform corresponds to the DCT, Fig. 5.1 represents a block diagram for the baseline-JPEG [39]. Figure 5.2 depicts the encoder structure when the adaptive transform is used. In other words, we use this algorithm for the OALBT. For fair comparison, the bit assignment parts (which is denoted by "VLC" in Figs. 5.1 and 5.2) in both the encoders are the same. However, we must choose an appropriate transform out of a set of adaptive transforms. Therefore, we need a classifier for input block signals. Moreover, short basis functions of adaptive transforms are of an 2D non-separable form. In the following, the details of the encoders for the VLLBT, the Type-E OALBT, and the Type-O OALBT are illustrated.



HT: Horizontal transformation
VT: Vertical transformation
RTC: Zig-zag reordering and quantization of the transform coefficients
VLC: Variable length coding (Run-length/Huffman encoding)

Figure 5.1: The structure of the non-adaptive encoder



CLA: Classification the blocks into (J +1) classes (J directional or one non-directional)
HTL: Horizontal transformation with the long basis functions
VTL: Vertical transformation with the long basis functions
ATS: Adaptive transformation with the short basis functions
RTC: Reordering and quantization of the transform coefficients
VLC: Variable length coding (Huffman encoding)

Figure 5.2: The structure of the adaptive encoder

5.2.1 VLLBT

In the VLLBT case, to process finite length signals, special care is required at the image boundary to avoid border distortion. Since the basis functions have different center of symmetry, the symmetric extension method is unable to be used for the proposed VLLBT. However, this problem has already discussed in [8, 38]. Thus, we apply the method introduced in [38] to boundary processing.

5.2.2 Type-E OALBT

The encoder consists of the following four steps

1. classification;
2. transformation;
3. reordering/quantization;
4. coding.

An input image is partitioned into 8×8-blocks. The block is classified into one of the classes and transformed by the corresponding adaptive transform. Side information on the transform selected in each block are coded by using the run-length and Huffman techniques. Due to boundary blocks whose size is not 8×8, scan method of transform coefficients is slightly tricky, The scanned coefficients are uniform quantized with the same step size for each coefficient. In order to code the transform coefficients, we use the same run-length/Huffman table as the baseline JPEG [39]. The details are described as follows.

Transformation

An input image is firstly transformed with the long basis functions. Transformation is performed in horizontal and vertical directions in the image as described in Fig. 4.4. When the input image is written as a matrix form $\boldsymbol{F}$, the matrix which contains the transform coefficients with respect to the long basis functions is obtained by

$$\hat{\boldsymbol{F}} = \hat{\boldsymbol{T}}^T \boldsymbol{F} \hat{\boldsymbol{T}}, \tag{5.1}$$

where

$$\hat{\boldsymbol{T}} = \begin{bmatrix} \hat{\boldsymbol{E}}^{(\mathrm{left})} & & & 0 \\ & \hat{\boldsymbol{E}} & & \\ & & \ddots & \\ & & & \hat{\boldsymbol{E}} \\ 0 & & & \hat{\boldsymbol{E}}^{(\mathrm{right})} \end{bmatrix}. \tag{5.2}$$

For example, when the image $\boldsymbol{F}$ is of size 512×512, the transform matrix of Malvar's long basis functions given in (4.22) produces the coefficient matrix $\hat{\boldsymbol{F}}$ of size 128×128.

On the other hand, the input image is partitioned into blocks as shown in Fig. 4.4. In this experiment, the use of adaptive transform is applied to normal 8×8 blocks. At boundary blocks, we assume that signals are the "non-directional" Markov process with the correlation matrix given by (4.21). At normal blocks, the signal is classified into one of the class set $\{C_j\}_{j=0}^{J}$ and transformed by the corresponding matrix $\overline{\boldsymbol{E}}_{0_j}^{\mathrm{2D}}$.

Classification

In order to transform blocks with the short basis functions, the encoder has to classify blocks into one of the classes C_j, $j = 0, \ldots, J-1$ for the OALBT and the class C_J for the VLLBT. Then, each block is transformed by the OALBT or the VLLBT derived from the corresponding correlation matrix $\boldsymbol{R}_{\boldsymbol{ff}}^{(j)}$.

Let $\boldsymbol{f}$ be an input block. If the block is a part of smooth image regions, that is, the variance of $\boldsymbol{f}$ is comparatively small, then it would not contain any strong edges. Therefore, it is natural that $\boldsymbol{f}$ belongs to C_J. Specifically, if the variance is below a certain threshold τ, then the input signal is assigned to the "non-directional" class C_J, i.e,

$$\sigma_{\boldsymbol{f}}^2 < \tau \Longrightarrow \boldsymbol{f} \in C_J, \tag{5.3}$$

where $\sigma_{\boldsymbol{f}}^2$ denotes the variance of $\boldsymbol{f}$. On the other hand, if the variance of $\boldsymbol{f}$ is comparatively large, that is, $\sigma_{\boldsymbol{f}}^2 \geq \tau$, the block would contain strong edges or textures. In this case, we classify the data $\boldsymbol{f}$ by the subspace method. The projection matrix with respect to the class C_j is defined as

$$\boldsymbol{P}_j = \sum_{r=0}^{R-1} \overline{\boldsymbol{r}}_r^{(j)} \overline{\boldsymbol{e}}_r^{(j)T}, \tag{5.4}$$

where $1 \le R \le M - N_L - 1$. In subspace methods, the projection matrix $\boldsymbol{P}_j$ characterizes each class C_j. Hence, an input block $\boldsymbol{f}$ is assigned to the class such that the norm of its projection is maximized, that is, for $i, j = 0, \ldots, J$,

$$\|\boldsymbol{P}_i \boldsymbol{f}\|^2 > \|\boldsymbol{P}_j \boldsymbol{f}\|^2 \quad \forall j \neq i \Longrightarrow \boldsymbol{f} \in C_i. \tag{5.5}$$

Assume that $\overline{\boldsymbol{r}}_i$, $i = 0, \ldots, M - N_L - 1$ is normalized so that $\|\overline{\boldsymbol{r}}_i\| = 1$. Then, the above expression can be reduced to

$$\sum_{r=0}^{R-1} |\langle \overline{\boldsymbol{e}}_r^{(i)}, \boldsymbol{f} \rangle|^2 > \sum_{r=0}^{R-1} |\langle \overline{\boldsymbol{e}}_r^{(j)}, \boldsymbol{f} \rangle|^2 \quad \forall j \neq i \Longrightarrow \boldsymbol{f} \in C_i, \tag{5.6}$$

since

$$\begin{aligned} \|\boldsymbol{P}_j \boldsymbol{f}\|^2 &= \left\| \sum_{r=0}^{R-1} \langle \overline{\boldsymbol{e}}_r^{(j)}, \boldsymbol{f} \rangle \overline{\boldsymbol{r}}_r^{(j)} \right\|^2 \\ &= \sum_{r=0}^{R-1} \sum_{s=0}^{R-1} \langle \overline{\boldsymbol{e}}_r^{(j)}, \boldsymbol{f} \rangle \langle \overline{\boldsymbol{e}}_s^{(j)}, \boldsymbol{f} \rangle \langle \overline{\boldsymbol{r}}_r^{(j)}, \overline{\boldsymbol{b}}_s^{(j)} \rangle \\ &= \sum_{r=0}^{R-1} |\langle \overline{\boldsymbol{e}}_r^{(j)}, \boldsymbol{f} \rangle|^2 \end{aligned} \tag{5.7}$$

by using the fact that inverse short basis functions $\overline{\boldsymbol{r}}_r$ are orthonormal because the matrix $\boldsymbol{Q}$ is symmetric.

Bit Allocation for Side Information

For decoding, side information which indicates the class of each block should be transmitted or stored. We build the Huffman codebook for run-length on the non-directional class. For a block which does not belong to the non-directional class C_J, five bits are allocated. When the rest of the blocks in the image are the non-directional class blocks, the end-of-header (EOH) symbol is used for efficient coding. The Huffman codes used for side information are listed in Table 5.1. Consequently, the larger threshold τ causes the shorter length of bitstream for side information, since the more blocks are classified into the non-directional class C_J. However, this leads to less coding performance. It is necessary to choose an appropriate threshold τ empirically.

Scanning Methods for Coefficients

The transform coefficients are coded by using the run-length and Huffman coder such as the baseline JPEG [39]. Because a few blocks, say, boundary blocks have different sizes from what

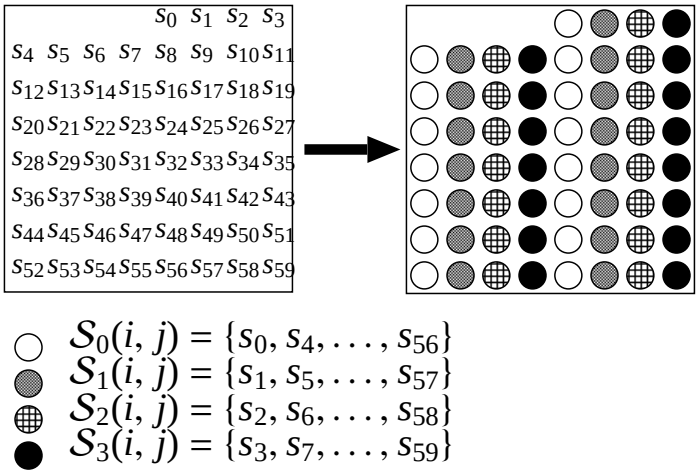


Figure 5.3: The four subsets $\mathcal{S}_m(i,j)$ that the set of coefficients $\mathcal{S}(i,j)$ are divided into

normal blocks have, it is necessary to use various run-length Huffman coders. However, by reorganizing the transform coefficients, we can apply a single Huffman coder. Figures 5.3 and 5.4 visualizes the block reorganizing method for the transform coefficients. The transform coefficients produced by $\hat{\boldsymbol{E}}_{0_j}^{2\mathrm{D}}$ are written as s_k, and they are illustrated as four kinds of marked circles, as shown in Fig. 5.3. Let $\mathcal{S}(i,j)$ be the set of the transform coefficients s_k of the normal block (i,j). In this example, $\mathcal{S}(i,j)$ has 60 elements, that is, $\mathcal{S}(i,j) = \{s_k\}_{k=0}^{59}$. For $m = 0,1,2,3$, let $\mathcal{S}_m(i,j)$ be the subset of $\mathcal{S}(i,j)$ defined as $\mathcal{S}_m(i,j) = \{s_k | k = 4n+m, 0 \le n \le 11\}$. To simplify the presentation, the elements of each subset $\mathcal{S}_m(i,j)$ are illustrated separately in Fig. 5.4 (b).

On the other hand, one can partition the transform coefficients matrix $\hat{\boldsymbol{F}}$ produced by $\hat{\boldsymbol{E}}$, $\hat{\boldsymbol{E}}^{(\mathrm{left})}$, and $\hat{\boldsymbol{E}}^{(\mathrm{right})}$ into $N_L \times N_L$-blocks (see Fig. 5.4 (c)). In this experiment, these blocks consist of one "DC" component and three "AC" components, and the coefficient set of the (i,j) block is written by $\mathcal{L}(i,j)$.

These sets organize the new coefficient set $\mathcal{B}(i,j)$ is given as

$$\mathcal{B}(i,j) = \{\mathcal{L}(i,j), \mathcal{S}_0(i+1,j+1), \mathcal{S}_1(i,j+1), \mathcal{S}_2(i+1,j), \mathcal{S}_3(i,j)\}, \tag{5.8}$$

as shown in Fig. 5.4 (d). After the quantization, in each block $\mathcal{B}(i,j)$, the run-length coder scans the coefficients of $\mathcal{L}$ first, and those of $\mathcal{B}$ in ascending order with respect to k.

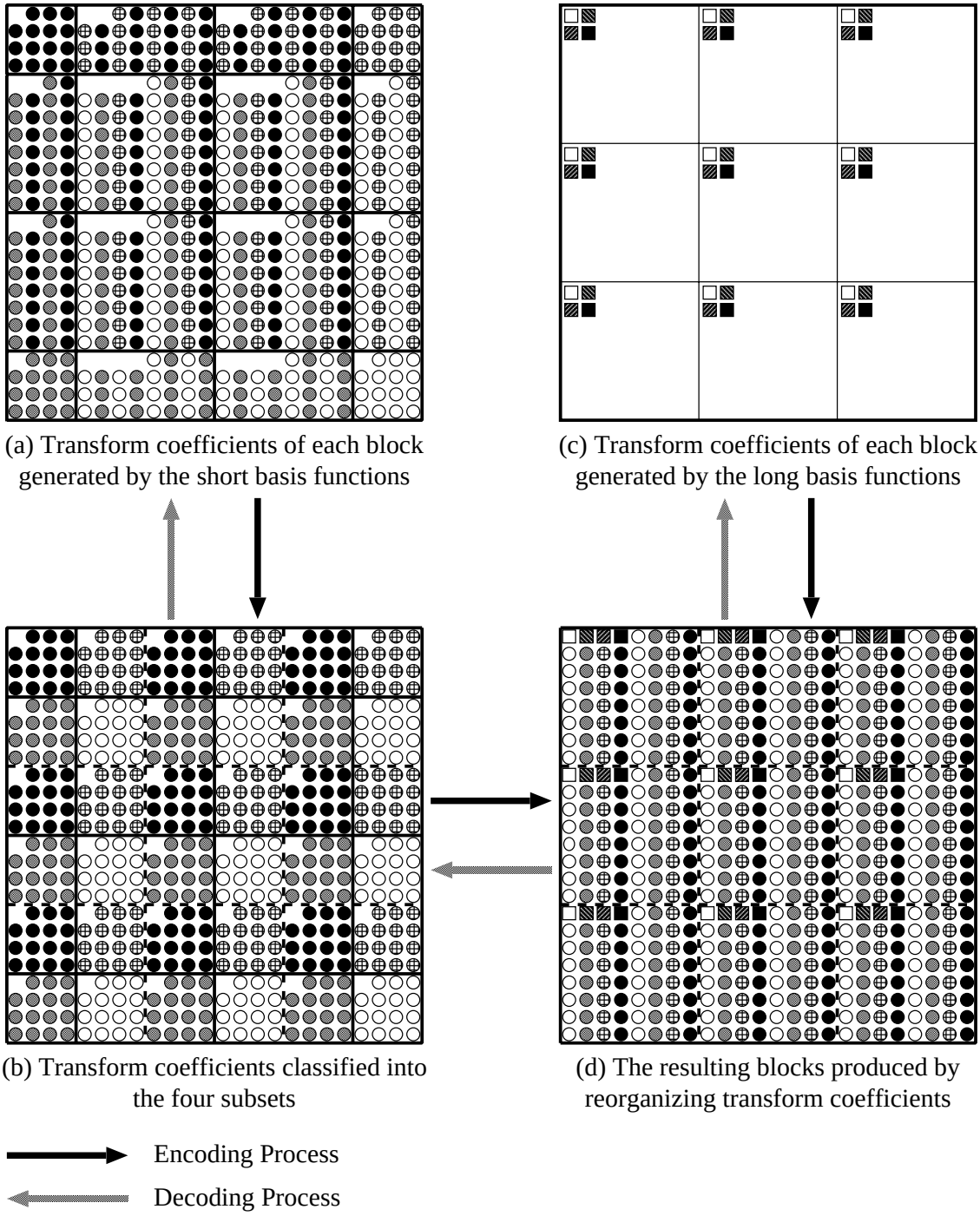


Figure 5.4: Block reorganizing

Table 5.1: Huffman codebook for run-length on the non-directional class to encode side information

Run	Code Length	Code Word	Hex
0	1	0	0000
1	2	10	0002
2	4	1100	000C
3	4	1101	000D
4	5	11100	001C
5	5	11110	001E
6	6	111110	003E
7	8	11111100	00FC
8	8	11111101	00FD
9	8	11111110	00FE
10	9	111111110	01FE
11	10	1111111110	03FE
12	12	111111111100	0FFC
13	12	111111111101	0FFD
14	12	111111111110	0FFE
15	13	1111111111110	1FFE
DRL	5	11101	001D
EOH	13	1111111111111	1FFF

5.2.3 Type-O OALBT

Because of the alignment property of basis functions of the Type-O OALBT, we need no special processing at image boundaries as needed for the VLLBT and the Type-E OALBT.

Transform, Classification, and Scanning

An input image is firstly transformed with the long basis functions. Transformation is performed in horizontal and vertical directions in the image. At image boundaries, the symmetric extension method [79, 80, 15] is employed. As a result, we obtain N_L^2 coefficients per block. Next, each block is classified into $J + 1$ classes by using the classification algorithm described in Section 5.2.2. Then, the block belonging to a class C_j is transformed with the short basis functions $\overline{\boldsymbol{E}}^{(j)}$. The short basis functions generate $(M^2 - N_L^2)$ transform coefficients at each block. Therefore, the total number of coefficients is the same as the size of the input image. Note that there is no redundancy.

For transform coefficients, in this test, we adopt a uniform scalar quantizer and the same run-length/Huffman codebook as the baseline JPEG [39]. However, zig-zag scanning is applied only to the coefficients produced by the long basis functions since the short basis functions have 2-D non-separable form.

5.3 Image Coding Results and Comparisons

The images used for the experiments are Barbara, Lena, and Pepper, which are standard, well-known 512×512 8-bit gray-scale test images. Table 5.2 shows comparison of PSNR (peak signal-to-noise ratio) in dB for the standard images at different bit rates in bit per pixel (bpp). Transforms listed in Table 5.2 are:

DCT the discrete cosine transform (DCT);

VLLOT26 the VLLOT ($N_L = 2$) designed in Chapter 3;

VLLBT26 the VLLBT ($N_L = 2$) designed in Chapter 3;

Type-E VLLBT the Type-E VLLBT ($N_L = 2$, $K = 2$) with 2D non-separable short basis functions designed in Chapter 4;

Type-E OALBT the Type-E OALBT ($N_L = 2$, $K = 2$) designed in Chapter 4 (Adaptive Coding for 32 orientations);

Type-O VLLBT the 2D separable Type-O VLLBT ($N_L = 2$, $K = 3$);

Type-O OALBT the Type-O OALBT ($K = 2$, $L = 3$) designed in Chapter 4 (Adaptive Coding for 32 orientations).

In the following, we describe arguments on each proposed transform.

5.3.1 VLLBT: Orthogonal vs Biorthogonal

The original and the decoded images for Barbara at 0.25 bpp are shown in Figs. 5.5 and 5.6. Table 5.2 shows that at 1.00 bpp and 0.50 bpp, comparatively higher bit rates, the coding performance of the DCT based method is superior to that of the other methods. At bit rates lower than 0.50 bpp, the VLLOT and the VLLBT work well, and the proposed VLLBT outperforms the VLLOT

Table 5.2: Comparison of PSNR (dB) results for 512×512 "Barbara," 512×512 "Lena," and 512×512 "Pepper" images at different bit rates (bpp)

Bit Rate	1.00	0.50	0.25	0.20
512×512 "Barbara"				
DCT	34.90	29.09	24.49	23.14
VLLOT26	33.75	28.67	24.50	23.28
VLLBT26	33.38	28.40	24.68	23.86
Type-E OALBT	34.02	29.44	25.80	24.01
Type-E VLLBT	33.43	28.02	24.45	23.48
Type-O OALBT	35.84	30.10	25.76	23.85
Type-O VLLBT	34.70	29.06	24.74	23.66
512×512 "Lena"				
DCT	38.52	34.99	30.73	29.18
VLLOT26	38.25	34.74	30.65	29.24
VLLBT26	38.01	34.85	31.13	29.81
Type-E OALBT	38.16	34.69	30.96	29.21
Type-E VLLBT	38.07	34.39	30.51	29.13
Type-O OALBT	38.35	35.28	31.32	29.39
Type-O VLLBT	38.47	35.17	31.17	29.72
512×512 "Pepper"				
DCT	36.00	33.88	30.49	28.95
VLLOT26	35.58	33.57	30.37	28.89
VLLBT26	35.76	33.86	31.15	29.91
Type-E OALBT	35.97	33.97	30.83	29.14
Type-E VLLBT	35.81	33.61	30.53	29.29
Type-O OALBT	35.84	34.08	31.21	29.48
Type-O VLLBT	35.83	34.00	31.17	29.84

in PSNR. It should be noted that the blocking artifacts of the VLLBT-decoded image are much less visible than those of the DCT- and the VLLOT-decoded images. This property is caused by the fact that end values of each long basis functions of the inverse transform $\boldsymbol{r}_i$ are almost zero as illustrated in Fig. 3.4(b).

5.3.2 OALBT: Non-Adaptive vs Adaptive

The correlation coefficient of $\boldsymbol{R}_{ff}^{(J)}$ for design of the VLLBT is set to 0.95 ($\rho = 0.95$). In the classifier, eight short basis functions are used ($R = 8$). The threshold is set to $\tau = 300$ for the Type-E and $\tau = 200$ for the Type-O. The total number of classes is 33 ($J = 32$), where the classes $C_j, j = 0, \ldots, 32$ denote 32 "directional" blocks and C_{32} denotes one non-directional block.

Type-E

Figure 5.7 shows the resulting class map of the test image "Barbara.". In Fig. 5.7, four kinds of white lines indicate directional blocks and its direction of the angle, where 32 directions are quantized into four directions to be seen easily.

The amount of side information and the percentage of directional blocks in the OALBT coder are shown in Table 5.3. It can be observed in Table 5.2 that the OALBT consistently outperforms the VLLBT although there exists side information for the OALBT. This fact indicates that more sophisticated scheme to encode side information than ours may lead to better performance. The OALBT has a higher coding gain than the DCT at rates below 0.5 bpp. For "Barbara" at 0.25 bpp, the OALBT gains about 1.39 dB over the DCT.

Figure 5.8 illustrates the decoded images 512×512 "Barbara" at 0.25 bpp, and Fig. 5.9 shows their magnified versions. Both the VLLBT and the OALBT have no blocking artifacts. However, it should be noted that the OALBT reduces ringing and blurring around strong edges and therefore provides clearer edges and lines than the other coders. Consequently, performance of adaptive transforms gains an advantage over that of non-adaptive transforms even though side information is needed.

Type-O

The amount of side information and the percentage of directional blocks (corresponding to classes C_j, $j = 0, \ldots, 31$) in the OALBT coder are shown in Table 5.4. Figure 5.10 shows

(a) The original “Barbara”

(b) The DCT (PSNR = 24.49 dB)

(c) The VLLOT (PSNR = 24.50 dB)

(d) The VLLBT (PSNR = 24.68 dB)

Figure 5.5: Comparison of the decoded images at rate 0.25 bpp

(a) The original “Barbara”

(b) The DCT (PSNR = 24.49 dB)

(c) The VLLOT (PSNR = 24.50 dB)

(d) The VLLBT (PSNR = 24.68 dB)

Figure 5.6: The magnified images of Fig. 5.5

Table 5.3: The amount of overhead and the percentage of directional regions for the Type-E OALBT coder

	Barbara	Lena	Pepper
Overhead in bpp	0.0439	0.0272	0.0251
Directional region	38.8 %	20.4 %	17.2 %

Figure 5.7: Classification map of “Barbara” in the Type-E coder: Each white line indicates the angle of the directional block.

(a) The VLLBT (PSNR = 24.45[dB]) (b) The Type-E OALBT (PSNR = 25.80dB)

Figure 5.8: Comparison of original and encode versions (0.25 bpp) of the 512×512 grey-scale "Barbara"

(a) The VLLBT (PSNR = 24.45[dB]) (b) The Type-E OALBT (PSNR = 25.80dB)

Figure 5.9: The magnified images of Fig. 5.8

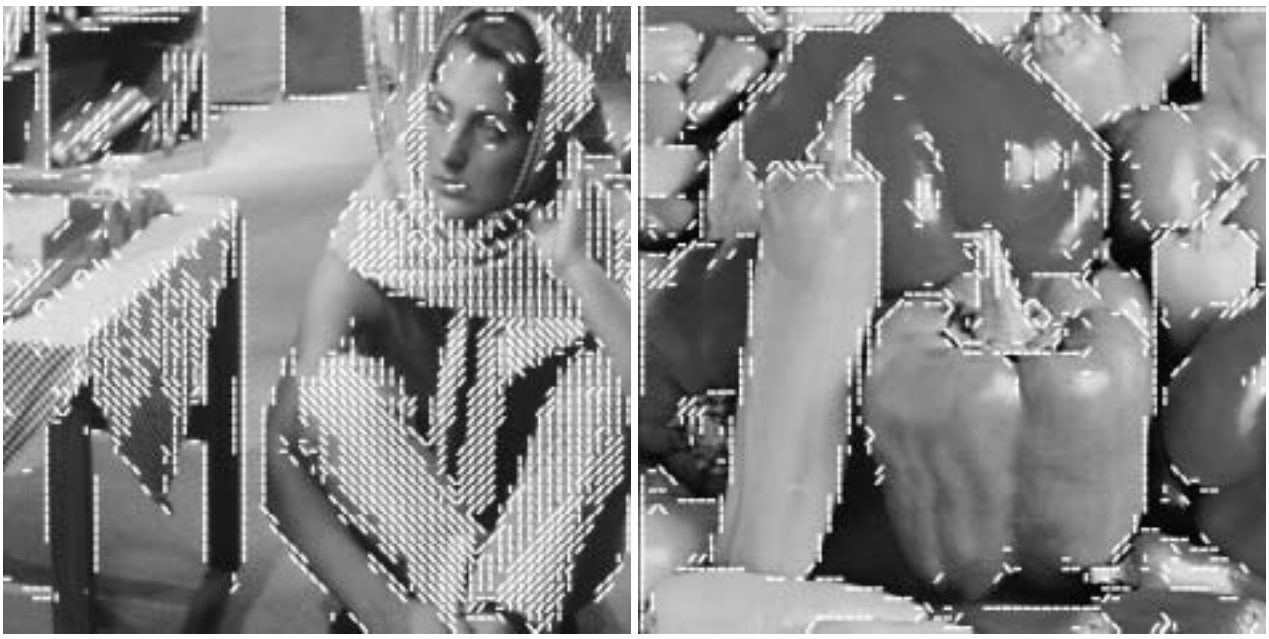
Table 5.4: The amount of overhead and the percentage of directional regions for the Type-O OALBT coder

	Barbara	Pepper	Lena
Overhead in bpp	0.0445	0.0276	0.0281
Directional region	39.1 %	19.3 %	20.2 %

the resulting map of the test images "Barbara" and "Pepper." In Fig. 5.10, four kinds of white line segments indicate directional blocks and its direction of the angle, where 32 directions are classified into four directions to be seen simply. In "Barbara," some blocks containing striped pattern are classified into directional classes, and others are not. Those blocks are judged to be plane regions by the encoder with (5.3). This decision depends on the threshold τ. For smaller threshold, they would be classified into directional regions. However, the smaller threshold results in the longer length of overhead and can lead to lower coding efficiency. The empirical discussion for the threshold in orientation adaptive coding is shown in [44].

Coding results and comparisons at different rates are illustrated in Table 5.2. In "Barbara," OALBT consistently outperforms both DCT and VLLBT in the PSNR sense, even though there exists side information. For example, at 0.25 bpp, OALBT gains 1.27 dB over DCT. In all images, at lower bit rates, both VLLBT and OALBT show higher PSNRs than DCT. However, PSNRs in "Pepper," "Lena," with OALBT are slightly lower than those with VLLBT. At around 0.20 bpp, OALBT results in lower performance than VLLBT. This can be explained by the existence of side information. For example, at the PSNR 30.00 dB, the bit rate (including side information) in OALBT is 0.2139 bpp, and that in VLLBT is 0.2049 bpp for "Pepper." Thus, VLLBT requires fewer bits than OALBT only by 0.0090 bits per pixel. In spite of this, 0.0276 bits are used for side information for "Pepper." This fact implies that a more sophisticated scheme to encode side information than ours may lead to better performance.

Figures 5.12 and 5.13 illustrate the original and the decoded images 512×512 "Barbara" and "Pepper" at 0.25 bpp, respectively. We can observe that OALBT provides better subjective quality of the decoded image compared to the other methods. It seems that OALBT reduces ringing and blurring around strong edges and therefore provides clearer edges and lines. These results may indicate the use of adaptive basis functions in the VLLBT can improve coding efficiency despite the fact that side information is required for decoding.



(a) “Barbara” (b) “Pepper”

Figure 5.10: Classification map in the Type-O coder: Each white line segment indicates the angle of the directional block.

(a) The VLLBT (PSNR = 24.74 dB) (b) The OALBT (PSNR = 25.76 dB)

Figure 5.11: Comparison of the decoded “Barbara” images at rate 0.25 bpp

5.3.3 Compariton of Type-O OALBT with Other Existing Lapped Transforms

As seen in the previous sections, the Type-O OALBT gives the most promising result among the proposed transforms. In this section, we compare it with the conventional LOT and LBT. The LBT used for comparison is Malvar’s LBT [36]. Rate-distortion characteristics for three coders and the DCT coder are plotted in Fig. 5.15. The Type-O OALBT consistently outperforms the DCT at rates lower than 0.70 bpp. Moreover, it provides higher PSNRs for the “Pepper” image at rates lower than approximately 0.60 bpp. However, the LBT consistently outperforms the OALBT. It provides better performance in the rate-distortion sense than the OALBT.

So, how is subjective quality? The answer is shown in Figs. 5.16 and 5.17. The decoded images with the LBT suffer from annoying ringing around edges compared to the ones with the OALBT as shown in Figs. 5.11, 5.12, 5.13, and 5.14. It is evident from those figures that the OALBT preserve edges and lines more clearly.

(a) The VLLBT (PSNR = 24.74 dB) (b) The OALBT (PSNR = 25.76 dB)

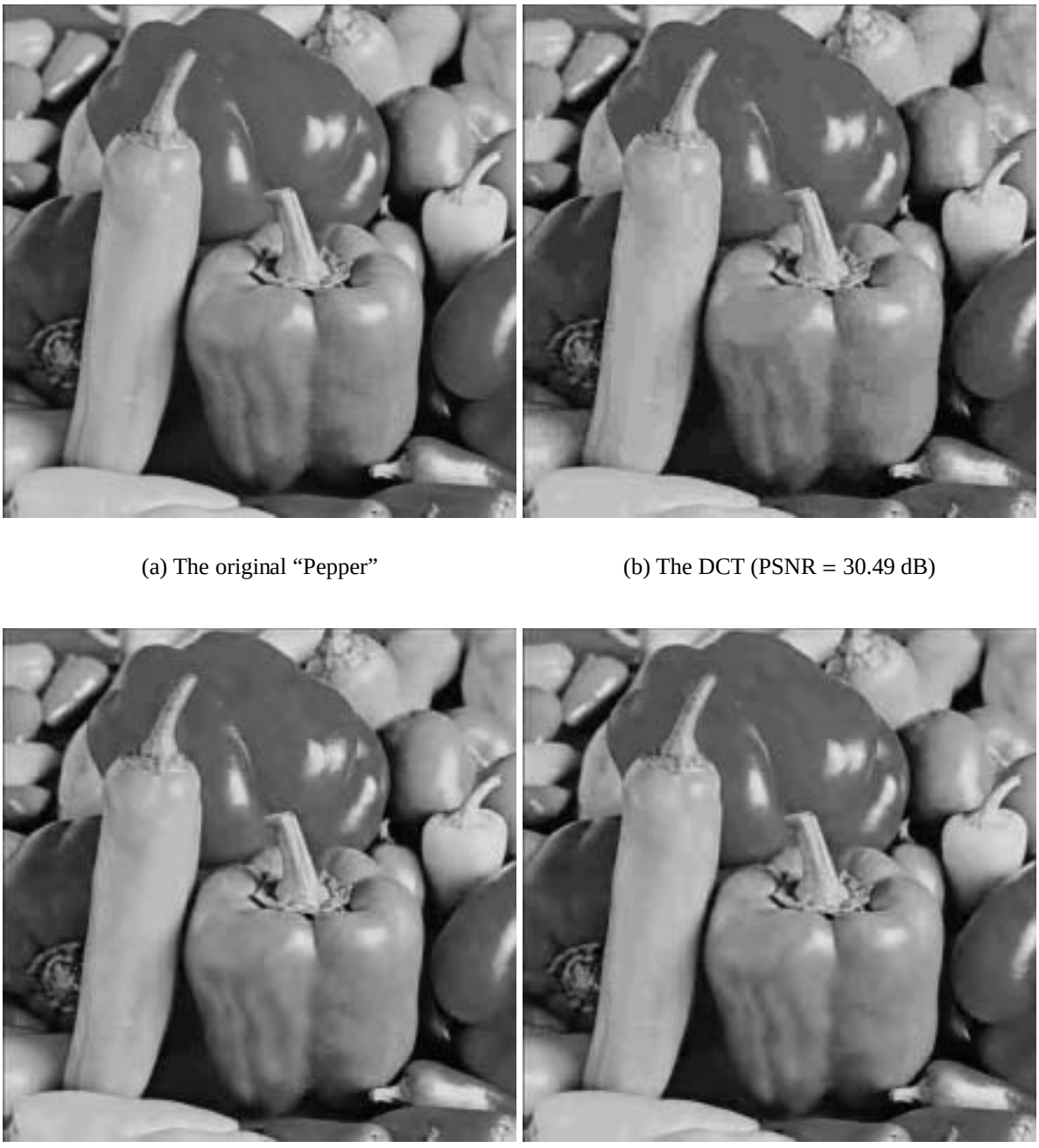
Figure 5.12: Magnified images in Fig. 5.11

5.4 Summary

We have presented image coding examples and their comparison. From the comparative study in this chapter, we can summarize those results as follows:

- Biorthogonal transforms provide better coding efficiency than orthogonal ones do.
- The orientation adaptation improves decoded images not only in preserving edges and lines, but also in PSNRs at low bit rates.
- Type-O OALBT outperforms Type-E OALBT in coding efficiency and coder complexity.

In addition, despite of the existence of side information accompanied by the adaptation, the OALBT provides high PSNRs and good visual quality. Therefore, the Type-O OALBT may be a promising transform in the field of image coding.



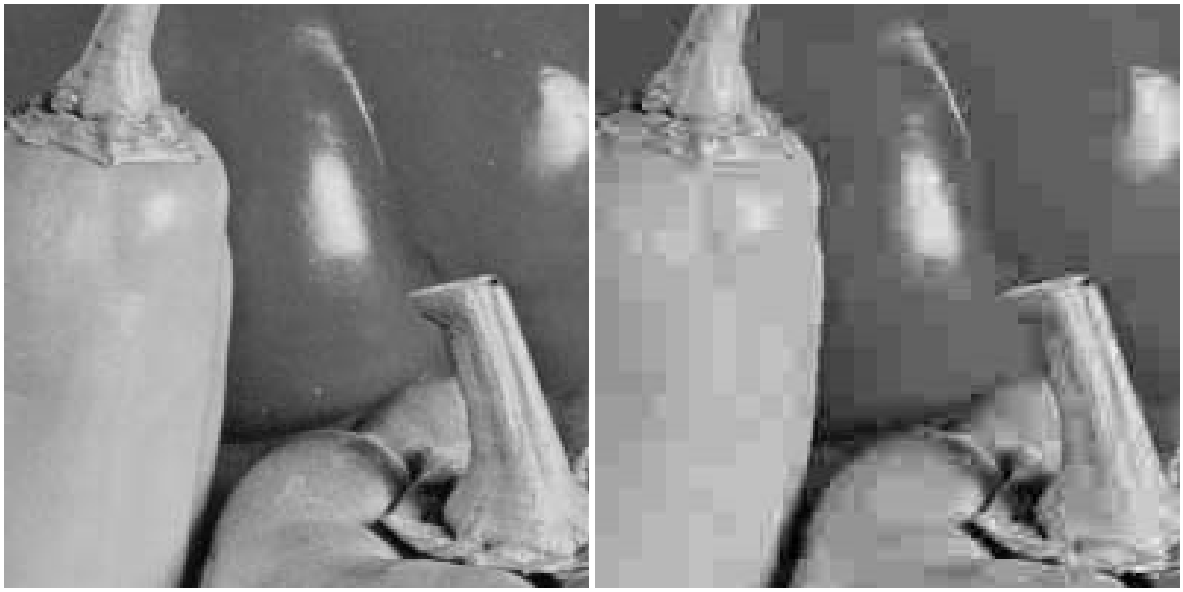
(a) The original “Pepper”

(b) The DCT (PSNR = 30.49 dB)

(c) The Type-O VLLBT (PSNR = 31.17 dB)

(d) The Type-O OALBT (PSNR = 31.21 dB)

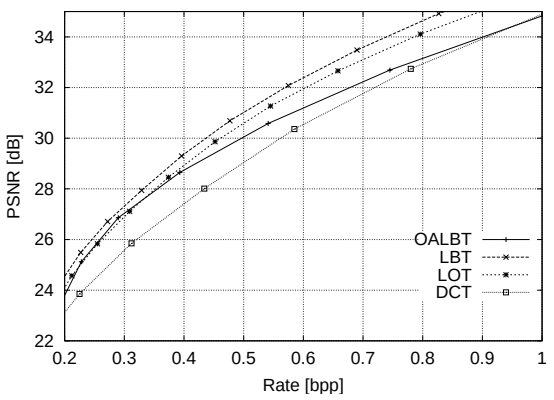
Figure 5.13: Comparison of the decoded “Pepper” images at rate 0.25 bpp



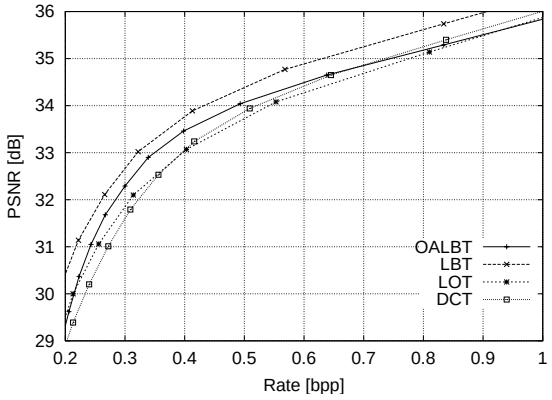
(a) The original “Pepper” (b) The DCT (PSNR = 30.49 dB)

(c) The Type-O VLLBT (PSNR = 31.17 dB) (d) The Type-O OALBT (PSNR = 31.21 dB)

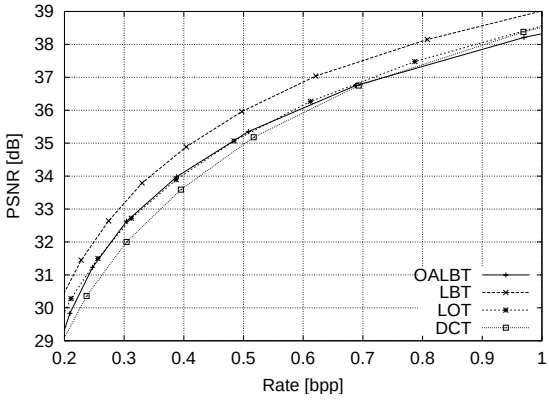
Figure 5.14: Magnified images in Fig. 5.13



(a) “Barbara”



(b) “Pepper”



(c) “Lena”

Figure 5.15: Comparison of PSNR (dB) results for 512×512 “Barbara,” “Lena,” and “Pepper” images at different bit rates (bpp)

(a) The LOT (PSNR = 25.70 dB)

(b) The LBT (PSNR = 26.11 dB)

(c) The LOT (PSNR = 30.94 dB)

(d) The LBT (PSNR = 31.76 dB)

Figure 5.16: The decoded “Barbara” and “Pepper” images with the LOT and the LBT at rate 0.25 bpp

(a) The LOT (PSNR = 25.70 dB)

(b) The LBT (PSNR = 26.11 dB)

(c) The LOT (PSNR = 30.94 dB)

(d) The LBT (PSNR = 31.76 dB)

Figure 5.17: Magnified images in Fig. 5.11

Chapter 6

An Oversampled Lapped Transform

6.1 Introduction

In all LPPRFBs with the lattice structure in the literature, the number of channels is equal to the decimation factor. On the other hand, FBs with a smaller decimation factor M than the number of channels N are called *oversampled* FBs. In this chapter, we make an attempt to construct N-channel ($N > M$) LPPRFBs based on the lattice structure. The oversampled FBs have some advantages such as their improved design freedom and noise immunity [29, 30]. In this chpter, our goal is to establish lattice structures which can even represent the existing GenLOT and GLBT. After preliminaries, a class of oversampled LPPRFBs is proposed in Section 6.2. From the lapped transform perspective, we call these FBs the generalized lapped pseudo-biorthogonal transform (GLPBT). This fundamental factorization is further parameterized by applying the singular value decomposition (SVD) in Section 6.2.2. The SVD enables us to characterize all filters by rotation angles and positive real numbers. For odd N, the factorization can be established in a similar fashion as shown in Section 6.2.6. The relation between the GLPBT and the conventional lapped transforms are discussed in Section 6.2.7. We also consider in Section 6.3 the noise robust GLPBT which has the final block in the lattice supressing noise added in the transform domain. We present some design examples in Section 6.4 and conclude this work in Section 6.5. This work can be regarded as a consequence of a generalization of [22] and [24], and can cover a wide range of possible LPPRFBs, as summarized in Table 1.3.

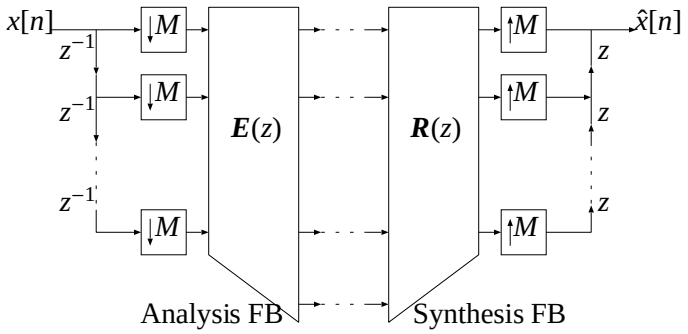


Figure 6.1: Polyphase representation.

6.1.1 LP and PR Conditions for Oversampled FBs

When the channel number N is greater than the decimation factor M, that is, $N > M$, such a FB is called an *oversampled FB*. Throughout this chapter, the polyphase matrices with respect to the analysis and the synthesis banks are written by $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$, respectively. The polyphase matrix $\boldsymbol{E}(z)$ is of size $N \times M$, and $\boldsymbol{R}(z)$ is of size $M \times N$. Figure 2.1 illustrates the oversampled N-channel FB that is dealt with throughout this chapter. This FB can be represented in terms of the corresponding polyphase matrices as shown in Fig. 6.1.

A FB system provides PR (with zero delay) if and only if

$$\boldsymbol{R}(z)\boldsymbol{E}(z) = \boldsymbol{I}_M. \tag{6.1}$$

If $\boldsymbol{R}(z)$ is the paraconjugate of $\boldsymbol{E}(z)$, i.e., $\boldsymbol{R}(z) = \boldsymbol{E}^T(z^{-1})$, we call this system *pseudo-orthogonal*; otherwise we call it *pseudo-biorthogonal*. The special case where $M = N$ gives paraunitary or orthogonal FBs if it holds that $\boldsymbol{R}(z) = \boldsymbol{E}^T(z^{-1})$, and gives biorthogonal FBs otherwise. It should be noted that the notion of pseudo-biorthogonality spans a very large space of PR FBs.

Let us consider an expression for the PR condition in the time domain, which is sometimes useful for understanding in a vector space. Let

$$\boldsymbol{E}(z) = \sum_{i=0}^{K-1} \boldsymbol{E}_i z^{-i}, \tag{6.2}$$

where $\boldsymbol{E}_i$ is a matrix with no delay whose size is the same as $\boldsymbol{E}(z)$. Similarly, let

$$\boldsymbol{R}(z) = \sum_{i=0}^{K-1} \boldsymbol{R}_i z^{i}, \tag{6.3}$$

where $\boldsymbol{R}_i$ is also a matrix with no delay. Substituting (6.2) and (6.3) into the PR condition (6.1), we obtain the equivalent condition in the time domain as follows [13]:

$$\sum_{i=0}^{K-1-s} \boldsymbol{R}_i \boldsymbol{E}_{i+s} = \sum_{i=0}^{K-1-s} \boldsymbol{R}_{i+s} \boldsymbol{E}_i = \delta_s \boldsymbol{I}_M, \quad s = 0, \ldots, K-1, \tag{6.4}$$

where $\delta_s = 1$ if $s = 0$; $\delta_s = 0$ otherwise.

In order that $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ has the LP property, it is required that

$$\begin{aligned} \boldsymbol{E}(z) &= z^{-(K-1)} \boldsymbol{D} \boldsymbol{E}(z^{-1}) \boldsymbol{J}, \\ \boldsymbol{R}(z) &= z^{K-1} \boldsymbol{J} \boldsymbol{R}(z^{-1}) \boldsymbol{D}, \end{aligned} \tag{6.5}$$

where $\boldsymbol{D}$ is the diagonal matrix whose entry is $+1$ when the corresponding filter is symmetric and -1 when the corresponding filter is antisymmetric.

6.2 Generalized Lapped Pseudo-Biorthogonal Transform

6.2.1 Even-Channel GLPBT

The generalized lapped pseudo-biorthogonal transform (GLPBT) is a class of oversampled LP-PRFBs and a natural extension of the existing lapped transforms with lattice structure as summarized in Table 1.3.

Define the following matrix:

$$\boldsymbol{G}(z) = \frac{1}{2} \boldsymbol{\Phi} \boldsymbol{W} \boldsymbol{\Lambda}(z) \boldsymbol{W}, \tag{6.6}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{U} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V} \end{bmatrix}, \ \boldsymbol{W} = \begin{bmatrix} \boldsymbol{I}_{N/2} & \boldsymbol{I}_{N/2} \\ \boldsymbol{I}_{N/2} & -\boldsymbol{I}_{N/2} \end{bmatrix}, \boldsymbol{\Lambda}(z) = \begin{bmatrix} \boldsymbol{I}_{N/2} & \boldsymbol{0}_{N/2} \\ \boldsymbol{0}_{N/2} & z^{-1} \boldsymbol{I}_{N/2} \end{bmatrix}, \tag{6.7}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $N \times N$ nonsingular matrices; therefore, $\boldsymbol{G}(z)$ is FIR invertible. Moreover, notice that

$$\boldsymbol{G}(z) = z^{-(K-1)} \boldsymbol{D} \boldsymbol{G}(z^{-1}) \boldsymbol{D}. \tag{6.8}$$

Definition 6 *The even-channel GLPBT is an oversampled LPPRFB defined by the $N \times M$ analysis and the $M \times N$ synthesis polyphase matrices with the factorization given as*

$$\begin{aligned} \boldsymbol{E}(z) &= \hat{\boldsymbol{E}}(z) \boldsymbol{\Phi}_0 \boldsymbol{S}, \\ \boldsymbol{R}(z) &= \boldsymbol{S}^T \boldsymbol{\Phi}_0^- \hat{\boldsymbol{R}}(z), \end{aligned} \tag{6.9}$$

respectively, where $\hat{\boldsymbol{E}}(z)$ *and* $\hat{\boldsymbol{R}}(z)$ *are given by*

$$\hat{\boldsymbol{E}}(z) = \prod_{i=K-1}^{1} \boldsymbol{G}_i(z),$$
$$\hat{\boldsymbol{R}}(z) = \prod_{i=1}^{K-1} \boldsymbol{G}_i^{-1}(z), \tag{6.10}$$

where $\boldsymbol{\Phi}_0$ *and* $\boldsymbol{S}$ *are defined as follows:*

- *if* M *is even,*

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \boldsymbol{U}_0 & \boldsymbol{0}_{\frac{N}{2}\times\frac{M}{2}} \\ \boldsymbol{0}_{\frac{N}{2}\times\frac{M}{2}} & \boldsymbol{V}_0 \end{bmatrix}, \ \boldsymbol{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_{\frac{M}{2}} & \boldsymbol{J}_{\frac{M}{2}} \\ \boldsymbol{I}_{\frac{M}{2}} & -\boldsymbol{J}_{\frac{M}{2}} \end{bmatrix}, \tag{6.11}$$

  where both $\boldsymbol{U}_0$ *and* $\boldsymbol{V}_0$ *are left-invertible matrices of size* $N/2 \times M/2$*;*

- *if* M *is odd,*

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \boldsymbol{U}_0 & \boldsymbol{0}_{\frac{N}{2}\times\frac{M-1}{2}} \\ \boldsymbol{0}_{\frac{N}{2}\times\frac{M+1}{2}} & \boldsymbol{V}_0 \end{bmatrix}, \ \boldsymbol{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_{\frac{M-1}{2}} & \boldsymbol{0}_{\frac{M-1}{2}\times 1} & \boldsymbol{J}_{\frac{M-1}{2}} \\ 0 & \sqrt{2} & 0 \\ \boldsymbol{I}_{\frac{M-1}{2}} & \boldsymbol{0}_{\frac{M-1}{2}\times 1} & -\boldsymbol{J}_{\frac{M-1}{2}} \end{bmatrix}, \tag{6.12}$$

  where both $\boldsymbol{U}_0$ *and* $\boldsymbol{V}_0$ *are left-invertible matrices of sizes of* $N/2 \times (M+1)/2$ *and* $N/2 \times (M-1)/2$*, respectively.*

The GLPBT has the following properties:

1. all analysis and synthesis filters are FIR with the same length $L = KM$ and have the same center of symmetry;

2. it consists of $N/2$ symmetric filters and $N/2$ antisymmetric filters if N is even.

Remarks

- The terminology *pseudo-biorthogonal* comes from a theory of pseudo-biorthogonal bases [25, 27], which is a particular class of frames [26, 28]. It can be shown that the filters of the analysis bank and the synthesis bank generate a frame or a pseudo-biorthogonal basis.

- In the maximally decimated case, i.e., $M = N$, the second restriction on symmetric filters is the unique solution for the number of filters in LPPRFBs [81, 37]. In the overcomplete case, however, the assumption does not cover all oversampled LPPRFBs [24]. Therefore, a factorization for another class of oversampled LPPRFBs is an open problem.

A structure is said to be minimal if the number of delays used is equal to the degree of the transfer function [17]. The following theorem guarantees that the factorization is minimal.

Theorem 4 *The factorization in (6.9) gives a minimal realization, i.e., the number of delays required for its implementation is minimal.*

Proof. From Definition 6, since $\boldsymbol{\Phi}_0\boldsymbol{S}$ has no delay, we have $\deg[\boldsymbol{E}(z)] = \deg[\hat{\boldsymbol{E}}(z)]$. It has been verified in [82] that for a causal square polyphase matrix having an anticausal inverse, the degree of the system is equal to the degree of its determinant. Therefore, we have

$$\begin{aligned} \deg[\boldsymbol{G}(z)] &= \deg[|\hat{\boldsymbol{E}}(z)|] \\ &= \deg[z^{-N(K-1)}|\boldsymbol{D}|\,|\hat{\boldsymbol{E}}(z^{-1})|\,|\boldsymbol{D}|] \\ &= N(K-1) - \deg[\hat{\boldsymbol{E}}(z)], \end{aligned} \tag{6.13}$$

where we used (6.8). It follows that $\deg[\hat{\boldsymbol{E}}(z)] = N(K-1)/2$. This degree is equal to the total number of delays employed in the structure. Therefore, the factorization is minimal. □

6.2.2 Parameterization of Each Block

In this section, we parameterize $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ in each building block $\boldsymbol{G}_i(z)$ with the lattice structure, leading fast implementation. The key technique to parameterization is the well-known singular value decomposition (SVD). By means of the SVD, any invertible matrix can be decomposed into two orthogonal matrices and one diagonal matrix consisting of positive parameters. Since an orthogonal matrix of size n is completely characterized by $\binom{n}{2}$ rotation angles, the invertible matrix is parameterized by rotation angles and positive multipliers [22]. The complete parameterization by the SVD enables us to obtain all filter coefficients by solving an unconstrained optimization problem.

However, the initial block $\boldsymbol{\Phi}$ is not invertible but left-invertible. As seen, a left-invertible matrix does not have the unique left-inverse. (This is a different point from what an invertible matrix has the unique inverse.) In the following, we provide a solution to this problem. The key technique is also the SVD parameterization.

6.2.3 Invertible Matrices

Parameterization of invertible matrices with the SVD has been used in [22]. For $i \geq 1$, the SVD decomposes every invertible matrix as $\boldsymbol{U}_i = \boldsymbol{U}_{i1}\boldsymbol{\Gamma}_i\boldsymbol{U}_{i0}$, where $\boldsymbol{U}_{i0}$ and $\boldsymbol{U}_{i1}$ are orthogonal matrices

and $\mathbf{\Gamma}_i$ is a diagonal matrix consisting of positive values [50]. Similarly, the invertible $\boldsymbol{V}_i$ can be written as the product of orthogonal matrices $\boldsymbol{V}_{i0}$ and $\boldsymbol{V}_{i1}$ and a diagonal matrix $\mathbf{\Delta}_i$: $\boldsymbol{V}_i = \boldsymbol{V}_{i1}\mathbf{\Delta}_i\boldsymbol{V}_{i0}$. Their inverses $\boldsymbol{U}_i^{-1}$ and $\boldsymbol{V}_i^{-1}$ are represented as the following factorized forms: $\boldsymbol{U}_i^{-1} = \boldsymbol{U}_{i0}^T\mathbf{\Gamma}_i^{-1}\boldsymbol{U}_{i1}^T$ and $\boldsymbol{V}_i^{-1} = \boldsymbol{V}_{i0}^T\mathbf{\Delta}_i^{-1}\boldsymbol{V}_{i1}^T$. Consequently, $\mathbf{\Phi}_i, i = 1, \ldots, K-1$ can be further factorized [22] as

$$\mathbf{\Phi}_i = \begin{bmatrix} \boldsymbol{U}_{i1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V}_{i1} \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{i0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V}_{i0} \end{bmatrix}, \tag{6.14}$$

where all of the orthogonal matrices $\boldsymbol{U}_{i0}$, $\boldsymbol{U}_{i1}$, $\boldsymbol{V}_{i0}$, and $\boldsymbol{V}_{i1}$ are of size $N/2 \times N/2$; and therefore each matrix can be characterized by $\binom{N/2}{2} = \frac{N(N-2)}{8}$ rotations. The diagonal matrices $\mathbf{\Gamma}_i$ and $\mathbf{\Delta}_i$ are parameterized by $N/2$ positive parameters each. Hence, the matrix $\mathbf{\Phi}_i$ is parameterized by $N^2/2$ free parameters, which indeed agrees with the sum of the degrees of freedom of two $N/2 \times N/2$ invertible (nonsingular) matrices.

6.2.4 Left-Invertible Matrices

As mentioned above, an invertible matrix can be easily parameterized. However, the problem here is whether or not every left-invertible matrix can be parameterized. The key to solve the problem is the SVD.

Let us consider the left-invertible matrix $\boldsymbol{U}$ of size $n \times m$, where $n > m$ in the oversampled case. Since the rank of the left-invertible matrix $\boldsymbol{U}$ is m, $\mathrm{rank}(\boldsymbol{U}^T\boldsymbol{U}) = \mathrm{rank}(\boldsymbol{U}\boldsymbol{U}^T) = m$. This implies that both $\boldsymbol{U}^T\boldsymbol{U}$ and $\boldsymbol{U}\boldsymbol{U}^T$ have the same m positive eigenvalues $\lambda_i > 0,\ i = 0, \ldots, m-1$. Let $\boldsymbol{U}_a$ be an $n \times m$ matrix such that ith column is an eigenvector of $\boldsymbol{U}\boldsymbol{U}^T$ with respect to λ_i. Similarly, let $\boldsymbol{U}_b$ be an $m \times m$ matrix such that ith row is an eigenvector of $\boldsymbol{U}^T\boldsymbol{U}$ with respect to λ_i. Since both $\boldsymbol{U}^T\boldsymbol{U}$ and $\boldsymbol{U}\boldsymbol{U}^T$ are real symmetric, we can choose columns of $\boldsymbol{U}_a$ and rows of $\boldsymbol{U}_b$ such that $\boldsymbol{U}_a^T\boldsymbol{U}_a = \boldsymbol{I}_m$ and $\boldsymbol{U}_b\boldsymbol{U}_b^T = \boldsymbol{I}_m$, respectively. If $\boldsymbol{U}^T\boldsymbol{U}$ and $\boldsymbol{U}\boldsymbol{U}^T$ have m distinct eigenvalues, the above orthogonal property holds automatically. Applying the SVD to $\boldsymbol{U}$, we have the following decomposed form:

$$\boldsymbol{U} = \boldsymbol{U}_a\mathbf{\Gamma}\boldsymbol{U}_b, \tag{6.15}$$

where $\mathbf{\Gamma}$ is a diagonal matrix with positive parameters $\sqrt{\lambda_i}$. Since the columns of $\boldsymbol{U}_a$ form an orthonormal system, $\boldsymbol{U}_a$ can be parameterized by $\binom{n}{2} - \binom{n-m}{2} = \frac{m(2n-m-1)}{2}$ plane rotations. The orthogonal matrix $\boldsymbol{U}_b$ is characterized by $\binom{m}{2} = \frac{m(m-1)}{2}$ rotations.

6.2.5 Straightforward Choice for the Left-Inverse

In order to parameterize a left-inverse, we adopt the product $\boldsymbol{U}_b^T \boldsymbol{\Gamma}^{-1} \boldsymbol{U}_a^T$ as a left-inverse matrix. It is easily confirmed that the product belongs to the collection of left-inverse matrices of $\boldsymbol{U}_0$. To specify that it is a special element of the set of left-inverse matrices, we write it by $\boldsymbol{U}_0^+$, that is,

$$\boldsymbol{U}_0^+ = \boldsymbol{U}_b^T \boldsymbol{\Gamma}^{-1} \boldsymbol{U}_a^T, \tag{6.16}$$

which is indeed referred to as the Moore-Penrose (MP) pseudoinverse [51], which is uniquely determined with respect to a given matrix. Using this left-inverse, we can further factorize the $N \times M$ matrix $\boldsymbol{\Phi}_0$ as in (6.11) and (6.12):

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \boldsymbol{U}_{01} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{01} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Gamma}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Delta}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{00} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{00} \end{bmatrix}, \tag{6.17}$$

and the corresponding MP pseudoinverse is given by

$$\boldsymbol{\Phi}_0^+ = \begin{bmatrix} \boldsymbol{U}_{00}^T & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{00}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\Gamma}_0^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Delta}_0^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{01}^T & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{01}^T \end{bmatrix}, \tag{6.18}$$

where if M is even,

- $\boldsymbol{U}_{00}$ and $\boldsymbol{V}_{00}$: $M/2 \times M/2$ orthogonal matrices;
- $\boldsymbol{\Gamma}_0$ and $\boldsymbol{\Delta}_0$: $M/2 \times M/2$ diagonal matrices;
- $\boldsymbol{U}_{01}$ and $\boldsymbol{V}_{01}$: $N/2 \times M/2$ matrices of which columns are orthonormal,

and if M is odd,

- $\boldsymbol{U}_{00}$ and $\boldsymbol{V}_{00}$: $(M+1)/2 \times (M+1)/2$ and $(M-1)/2 \times (M-1)/2$ orthogonal matrices, respectively;
- $\boldsymbol{\Gamma}_0$ and $\boldsymbol{\Delta}_0$: $(M+1)/2 \times (M+1)/2$ and $(M-1)/2 \times (M-1)/2$ diagonal matrices, respectively;
- $\boldsymbol{U}_{01}$ and $\boldsymbol{V}_{01}$: $N/2 \times (M+1)/2$ and $N/2 \times (M-1)/2$ matrices of which columns are orthonormal, respectively.

Both for even M and for odd M, the number of free parameters for $\boldsymbol{\Phi}_0$ is $MN/2$. Details of the initial building block are illustrated in Figs. 6.2 and 6.3.

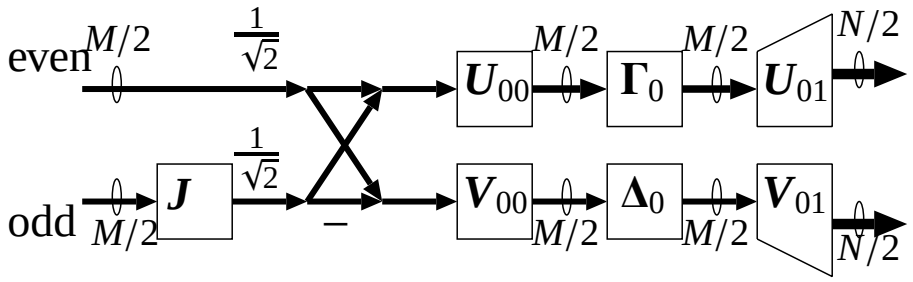


Figure 6.2: The initial block when both M and N are even

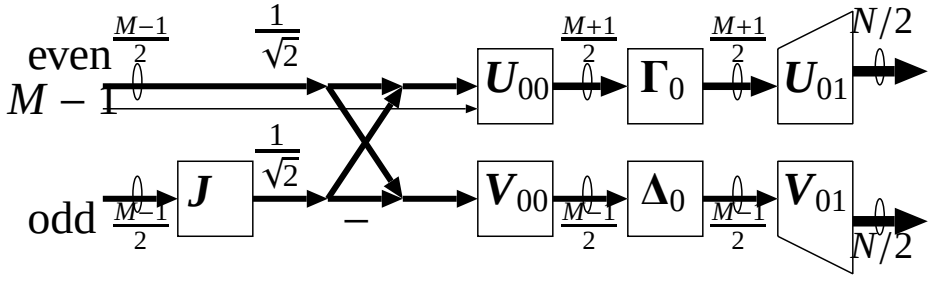


Figure 6.3: The initial block when M is odd and N is even

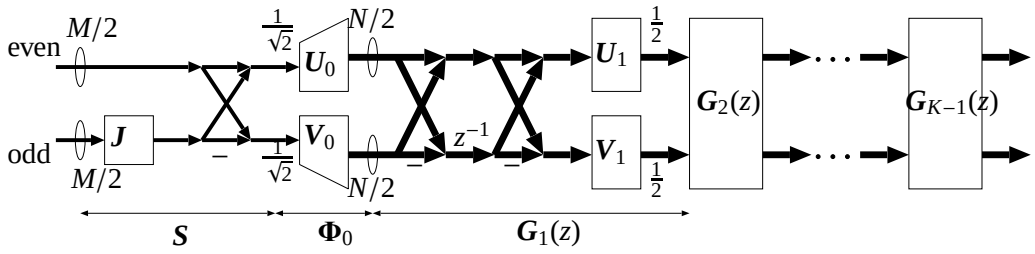


Figure 6.4: The lattice structure of the even-channel GLPBT for even M

Now, we have obtained a special synthesis polyphase matrix which gives an alternative form as follows:

$$\boldsymbol{R}(z) = \boldsymbol{S}^T \boldsymbol{\Phi}_0^+ \left(\prod_{i=1}^{K-1} \boldsymbol{G}_i^{-1}(z) \right). \tag{6.19}$$

This factorization for even M is depicted in Fig. 6.4. Actually, the MP pseudoinverse belongs to a subclass of the collection of left-inverses. For practical purpose, however, it may be effective and sufficient to adopt the MP pseudoinverse $\boldsymbol{\Phi}_0^+$ as a left-inverse, because it can be expressed by the lattice structure and can suppress the noise added in the transform domain.

6.2.6 Odd-Channel GLPBT

The decomposition can be easily established in the case that the number of filters N is odd. We provide only the results here.

$$\boldsymbol{G}_i(z) = \frac{1}{4} \boldsymbol{\Phi}_i \boldsymbol{W}_o \boldsymbol{\Lambda}_0(z) \boldsymbol{W}_o \boldsymbol{\Psi}_i \boldsymbol{W}_o \boldsymbol{\Lambda}_1(z) \boldsymbol{W}_o, \tag{6.20}$$

where

$$\begin{aligned} \boldsymbol{\Phi}_i = \begin{bmatrix} \boldsymbol{U}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_i \end{bmatrix}, \boldsymbol{\Psi}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & q_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_i \end{bmatrix}, \ \boldsymbol{W}_o = \begin{bmatrix} \boldsymbol{I}_{N/2} & \boldsymbol{0} & \boldsymbol{I}_{N/2} \\ \boldsymbol{0} & \sqrt{2} & \boldsymbol{0} \\ \boldsymbol{I}_{N/2} & \boldsymbol{0} & -\boldsymbol{I}_{N/2} \end{bmatrix}, \\ \boldsymbol{\Lambda}_0(z) = \begin{bmatrix} \boldsymbol{I}_{N/2} & \boldsymbol{0} & \boldsymbol{0}_{N/2} \\ \boldsymbol{0} & 1 & \boldsymbol{0} \\ \boldsymbol{0}_{N/2} & \boldsymbol{0} & z^{-1}\boldsymbol{I}_{N/2} \end{bmatrix}, \boldsymbol{\Lambda}_1(z) = \begin{bmatrix} \boldsymbol{I}_{N/2} & \boldsymbol{0} & \boldsymbol{0}_{N/2} \\ \boldsymbol{0} & z^{-1} & \boldsymbol{0} \\ \boldsymbol{0}_{N/2} & \boldsymbol{0} & z^{-1}\boldsymbol{I}_{N/2} \end{bmatrix}, \end{aligned} \tag{6.21}$$

- $\boldsymbol{V}_i$, $\boldsymbol{Q}_i$, and $\boldsymbol{R}_i$: $\frac{N-1}{2} \times \frac{N-1}{2}$ nonsingular matrices;
- $\boldsymbol{U}_i$: a $\frac{N+1}{2} \times \frac{N+1}{2}$ nonsingular matrix;

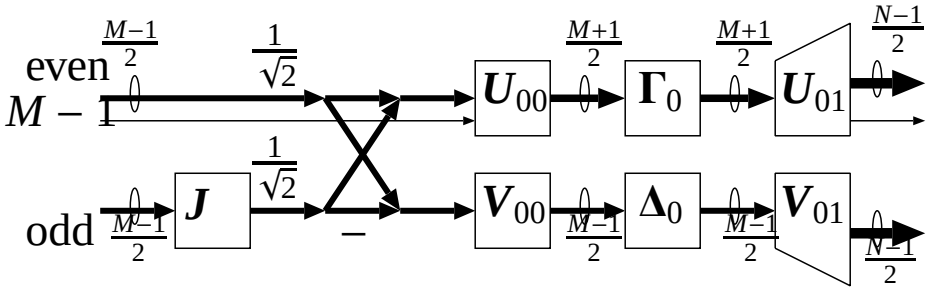


Figure 6.5: The initial block when both M and N are odd

- q_0: a scalar.

A polyphase matrix $\boldsymbol{E}(z)$ of an N-channel GLPBT with $(N+1)/2$ symmetric and $(N-1)/2$ antisymmetric filters can be factored as

$$\boldsymbol{E}(z) = \boldsymbol{G}_{K-2}(z)\boldsymbol{G}_{K-4}\cdots\boldsymbol{G}_3(z)\boldsymbol{G}_1(z)\boldsymbol{\Phi}_0\boldsymbol{S}, \tag{6.22}$$

where K must be odd [22].

For Even M

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \boldsymbol{U}_0 & \boldsymbol{0}_{\frac{N+1}{2}\times\frac{M}{2}} \\ \boldsymbol{0}_{\frac{N-1}{2}\times\frac{M}{2}} & \boldsymbol{V}_0 \end{bmatrix},\ \boldsymbol{S} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{I}_{\frac{M}{2}} & \boldsymbol{J}_{\frac{M}{2}} \\ \boldsymbol{I}_{\frac{M}{2}} & -\boldsymbol{J}_{\frac{M}{2}} \end{bmatrix}, \tag{6.23}$$

where $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$ are left-invertible matrices of size $(N+1)/2 \times M/2$ and $(N-1)/2 \times M/2$, respectively.

For Odd M

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \boldsymbol{U}_0 & \boldsymbol{0}_{\frac{N+1}{2}\times\frac{M-1}{2}} \\ \boldsymbol{0}_{\frac{N-1}{2}\times\frac{M+1}{2}} & \boldsymbol{V}_0 \end{bmatrix},\ \boldsymbol{S} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{I}_{\frac{M-1}{2}} & \boldsymbol{0}_{\frac{M-1}{2}\times 1} & \boldsymbol{J}_{\frac{M-1}{2}} \\ 0 & \sqrt{2} & 0 \\ \boldsymbol{I}_{\frac{M-1}{2}} & \boldsymbol{0}_{\frac{M-1}{2}\times 1} & -\boldsymbol{J}_{\frac{M-1}{2}} \end{bmatrix}, \tag{6.24}$$

where both $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$ are left-invertible matrices of sizes of $(N+1)/2 \times (M+1)/2$ and $(N-1)/2 \times (M-1)/2$, respectively. Details of the initial building block are depicted in Figs. 6.5 and 6.6.

It may be shown in a manner similar to that for even N, that non-singular matrices $\boldsymbol{\Phi}_i$ and $\boldsymbol{\Psi}_i$ can be parameterized with the SVD.

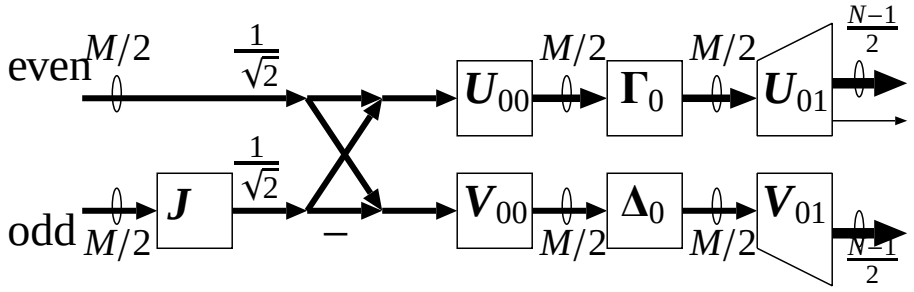


Figure 6.6: The initial block when M is even and N is odd

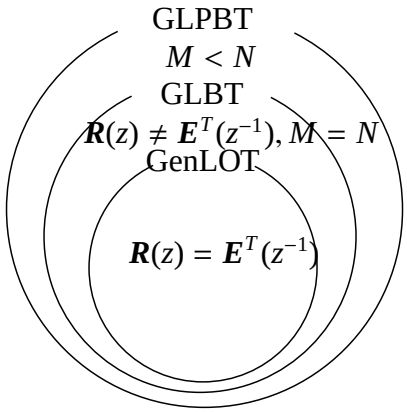


Figure 6.7: The relation between the GLPBT and other lapped transforms

6.2.7 Relation to the Conventional Lapped Transforms

As seen previously, the GLPBT can represent a very large class of LPPRFBs. If $M = N$, the GLPBT is identical to the GLBT [22]. If $M < N$ and $\boldsymbol{R}(z) = \boldsymbol{E}^T(z^{-1})$, the GLPBT represents the lattice structure of LP PU (pseudo-orthogonal) FBs proposed by Labeau *et al.* [24]. If $M = N$ and $\boldsymbol{R}(z) = \boldsymbol{E}^T(z^{-1})$, the GLPBT can represent the GenLOT [15, 16] and the lattice for LPPUFBs introduced by Soman *et al.* [17] Such a relation is illustrated in Fig. 6.7

6.3 Noise Robust GLPBT

Recall here that a benefit of the overcomplete representation is noise suppressing properties due to increased design freedom. Unfortunately, the lattice structures developed in [24] and in the above disregard the effect of noise added to the transformed signal. The use of the MP pseu-

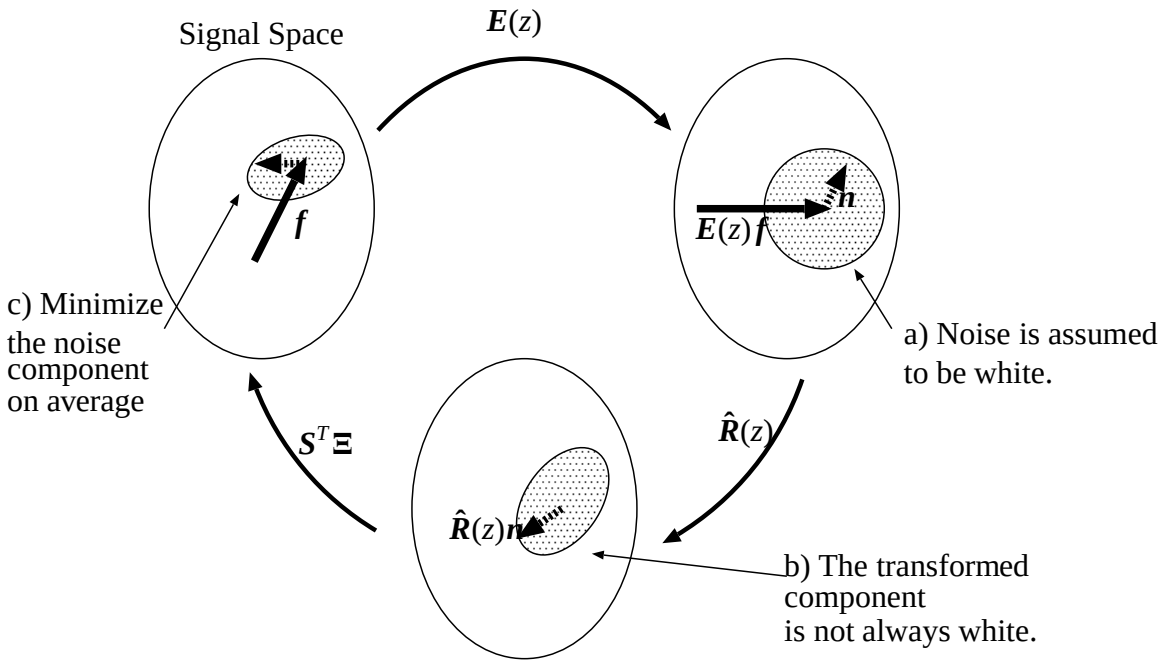


Figure 6.8: The noise robust GLPBT in subspaces: When noise added in the transform domain is white, the MP pseudoinverse may not optimally suppress the noise.

doinverse $\boldsymbol{\Phi}_0^+$, which corresponds to the use of a transposition $\boldsymbol{\Phi}_0^T$ in the pseudo-orthogonal (paraunitary) case [24], seems to reduce the effect of additive noise because of its minimal norm and least square properties. However, this use provides the optimal noise suppression only if the noise component transformed with $\hat{\boldsymbol{R}}(z)$ is white. Generally, this is not white, even although the noise added to a signal transformed with $\boldsymbol{E}(z)$ is white as depicted in Fig. 6.8. We consider in this section the case where we give the GLPBT lattice structure with specific parameters and we *a priori* know the correlation matrix of noise which is added in the transform domain.

The problem here is to find the appropriate synthesis polyphase matrix $\boldsymbol{X}(z)$ that reduces noise for a given analysis bank $\boldsymbol{E}(z)$. Let $\boldsymbol{f}_i \in \mathbb{R}^M$ and $\boldsymbol{n}_i \in \mathbb{R}^N$ be time series of random vectors, where $i \in \mathbb{Z}$. This $\boldsymbol{f}_i$ is transformed by $\boldsymbol{E}(z)$. Keep in mind that since z^{-k} is the delay operator defined by $z^{-n}\boldsymbol{f}_i = \boldsymbol{f}_{i-n}$, we have $\boldsymbol{E}(z)\boldsymbol{f}_i = \sum_{k=0}^{K-1} \boldsymbol{E}_k \boldsymbol{f}_{i-k}$. The signal $\boldsymbol{E}(z)\boldsymbol{f}$ will be degraded by additive noise $\boldsymbol{n}_i$ in the precess of transmission or storage. Therefore, the receiver or the observer obtains

$$\boldsymbol{g}_i = \boldsymbol{E}(z)\boldsymbol{f}_i + \boldsymbol{n}_i. \tag{6.25}$$

Note that $\boldsymbol{E}(z)$ provides overcomplete representation, and therefore there exist infinite number of synthesis polyphase matrices achieving PR. Let $\boldsymbol{X}(z)$ be determined by the submatrices $\boldsymbol{X}_i$, that is, $\boldsymbol{X}(z) = \sum_{k=0}^{K-1} \boldsymbol{X}_i z^k$. Moreover, let $\boldsymbol{X} = [\boldsymbol{X}_0 \ \cdots \ \boldsymbol{X}_{K-1}]$. Then, the reconstructed signal $\hat{\boldsymbol{f}}_i$ is obtained as

$$\begin{aligned} \hat{\boldsymbol{f}}_i &= \boldsymbol{X}(z)\boldsymbol{g}_i \\ &= \boldsymbol{f}_i + \boldsymbol{X}(z)\boldsymbol{n}_i. \end{aligned} \tag{6.26}$$

The approximation problem to be solved is to minimize

$$E_{\boldsymbol{n}} \| \hat{\boldsymbol{f}}_i - \boldsymbol{f}_i \|^2 \tag{6.27}$$

under the PR condition $\boldsymbol{X}(z)\boldsymbol{E}(z) = \boldsymbol{I}_M$, where $E_{\boldsymbol{n}}$ is the ensemble average on $\boldsymbol{n}_i, \ldots, \boldsymbol{n}_{i+K-1}$. Let $\boldsymbol{n} = [\boldsymbol{n}_i^T \ \cdots \ \boldsymbol{n}_{i+K-1}^T]^T$, and let $\boldsymbol{Q}$ be the correlation matrix of $\boldsymbol{n}$. Then, minimizing (6.27) is equivalent to minimizing the functional

$$\begin{aligned} \mathrm{MSE}[\boldsymbol{X}] &= E_{\boldsymbol{n}} \| \boldsymbol{X}(z)\boldsymbol{n}_i \|^2 \\ &= E_{\boldsymbol{n}} \left\| \boldsymbol{X} \begin{bmatrix} \boldsymbol{n}_i \\ \vdots \\ \boldsymbol{n}_{i+K-1} \end{bmatrix} \right\|^2 \\ &= \mathrm{tr}[\boldsymbol{X}\boldsymbol{Q}\boldsymbol{X}^T], \end{aligned} \tag{6.28}$$

under the PR condition

$$\sum_{i=0}^{K-1-s} \boldsymbol{X}_i \boldsymbol{E}_{i+s} = \sum_{i=0}^{K-1-s} \boldsymbol{X}_{i+s} \boldsymbol{E}_i = \delta_s \boldsymbol{I}_M, \quad s = 0, \ldots, K-1. \tag{6.29}$$

We should use an iterative optimization technique to obtain the solution. Moreover, it is not guaranteed that the synthesis polyphase matrix $\boldsymbol{X}^*(z)$ corresponding to the minimizer $\boldsymbol{X}^*$ can be factored into a lattice structure, although the analysis bank $\boldsymbol{E}(z)$ is organized as the lattice structure. This may lead to difficulty in implementation.

Therefore, let us consider here the following alternative problem. Assume that in a fashion similar to $\boldsymbol{R}(z)$, a synthesis polyphase matrix $\boldsymbol{X}(z)$ is given by the form

$$\boldsymbol{X}(z) = \boldsymbol{S}^T \boldsymbol{\Xi} \hat{\boldsymbol{R}}(z), \tag{6.30}$$

where $\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Xi}_U & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Xi}_V \end{bmatrix}$ and $\boldsymbol{\Xi}_U$ and $\boldsymbol{\Xi}_V$ are left-inverses of $\boldsymbol{U}_0$ and $\boldsymbol{V}_0$, respectively. Obviously, $\boldsymbol{X}(z)$ attains PR. Then, the cost functional becomes

$$\begin{aligned} \mathrm{MSE}[\boldsymbol{\Xi}_U, \boldsymbol{\Xi}_V] &= E_{\boldsymbol{n}} \|\boldsymbol{X}(z)\boldsymbol{n}_i\|^2 \\ &= E_{\boldsymbol{n}} \|\boldsymbol{S}^T \boldsymbol{\Xi} \hat{\boldsymbol{R}}(z)\boldsymbol{n}_i\|^2 \\ &= E_{\boldsymbol{n}} \|\boldsymbol{\Xi} \underbrace{[\hat{\boldsymbol{R}}_0 \ \cdots \ \hat{\boldsymbol{R}}_{K-1}]}_{\hat{\boldsymbol{R}}} \boldsymbol{n}\|^2 \\ &= E_{\boldsymbol{n}} \left\| \begin{bmatrix} \boldsymbol{\Xi}_U & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Xi}_V \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{R}}_u \\ \hat{\boldsymbol{R}}_l \end{bmatrix} \boldsymbol{n} \right\|^2 \\ &= \mathrm{tr}[\boldsymbol{\Xi}_U \hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T \boldsymbol{\Xi}_U^T] + \mathrm{tr}[\boldsymbol{\Xi}_V \hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T \boldsymbol{\Xi}_V^T], \end{aligned} \tag{6.31}$$

where $\hat{\boldsymbol{R}}(z) = \sum_{k=0}^{K-1} \hat{\boldsymbol{R}}_k z^k$ and $\hat{\boldsymbol{R}} = \begin{bmatrix} \hat{\boldsymbol{R}}_u \\ \hat{\boldsymbol{R}}_l \end{bmatrix}$. It is easily verified that the PR condition is reduced to

$$\boldsymbol{\Xi}_U \boldsymbol{U}_0 = \boldsymbol{\Xi}_V \boldsymbol{V}_0 = \boldsymbol{I}_{M/2}. \tag{6.32}$$

Since $\boldsymbol{\Xi}_U$ and $\boldsymbol{\Xi}_V$ are determined independently, we can divide the above problem into two independent problems. Therefore, the minimization problem for $\boldsymbol{\Xi}_U$ results in the following:

$$\begin{aligned} &\text{Minimize} && \mathrm{tr}[\boldsymbol{\Xi}_U \hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T \boldsymbol{\Xi}_U^T], \\ &\text{subject to} && \boldsymbol{\Xi}_U \boldsymbol{U}_0 = \boldsymbol{I}_{M/2}. \end{aligned}$$

Similarly, the minimization problem for $\boldsymbol{\Xi}_V$ is as follows:

$$\begin{aligned} &\text{Minimize} && \mathrm{tr}[\boldsymbol{\Xi}_V \hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T \boldsymbol{\Xi}_V^T], \\ &\text{subject to} && \boldsymbol{\Xi}_V \boldsymbol{V}_0 = \boldsymbol{I}_{M/2}. \end{aligned}$$

It is noted that the solution of the above problem is equivalent to the BLUE (best linear unbiased estimator) [83]. Therefore, we obtain the solutions

$$\begin{aligned} \boldsymbol{\Xi}_U^* &= (\boldsymbol{U}_0^T (\hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T)^+ \boldsymbol{U}_0)^+ \boldsymbol{U}_0^T (\hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T)^+, \\ \boldsymbol{\Xi}_V^* &= (\boldsymbol{V}_0^T (\hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T)^+ \boldsymbol{V}_0)^+ \boldsymbol{V}_0^T (\hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T)^+. \end{aligned} \tag{6.33}$$

In most cases, we can assume that $\boldsymbol{Q}$ is nonsingular, and therefore $(\hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T)^+$ and $(\hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T)^+$ can be replaced by $(\hat{\boldsymbol{R}}_u \boldsymbol{Q} \hat{\boldsymbol{R}}_u^T)^{-1}$ and $(\hat{\boldsymbol{R}}_l \boldsymbol{Q} \hat{\boldsymbol{R}}_l^T)^{-1}$, respectively.

In the general case, if white noise is added, the solution of the BLUE is provided by the MP pseudoinverse. We show however that in the GLPBT case, the choice of the MP pseudoinverse is generally incorrect. Assume that $\boldsymbol{Q} = \sigma^2 \boldsymbol{I}_N$ where σ^2 is the variance of the noise. The minimizer $\boldsymbol{\Xi}_U^*$ yields

$$\begin{aligned} \boldsymbol{\Xi}_U^* &= [\boldsymbol{U}_0^T(\sigma^2 \hat{\boldsymbol{R}}_u \hat{\boldsymbol{R}}_u^T)^+ \boldsymbol{U}_0]^+ \boldsymbol{U}_0^T(\sigma^2 \hat{\boldsymbol{R}}_u \hat{\boldsymbol{R}}_u^T)^+ \\ &= [(\hat{\boldsymbol{R}}_u^+ \boldsymbol{U}_0)^T \hat{\boldsymbol{R}}_u^+ \boldsymbol{U}_0]^+ (\hat{\boldsymbol{R}}_u^+ \boldsymbol{U}_0)^T \hat{\boldsymbol{R}}_u^+ \\ &= (\hat{\boldsymbol{R}}_u^+ \boldsymbol{U}_0)^+ \hat{\boldsymbol{R}}_u^+, \end{aligned} \tag{6.34}$$

where we used the fact that $\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{A}^T$. Therefore, the minimizer is not equal to $\boldsymbol{U}_0^+$ in general. In the pseudo-orthogonal (paraunitary) case, the PR condition implies that $\hat{\boldsymbol{R}} = [\hat{\boldsymbol{E}}_0^T \ \cdots \ \hat{\boldsymbol{E}}_{K-1}^T]$, which yields that

$$\hat{\boldsymbol{R}}\hat{\boldsymbol{R}}^T = \sum_{k=0}^{K-1} \hat{\boldsymbol{E}}_k^T \hat{\boldsymbol{E}}_k = \boldsymbol{I}_M. \tag{6.35}$$

Hence, we have

$$\boldsymbol{\Xi}_U^* = \boldsymbol{U}_0^+ = \boldsymbol{U}_0^T, \ \boldsymbol{\Xi}_V^* = \boldsymbol{V}_0^+ = \boldsymbol{V}_0^T. \tag{6.36}$$

However, keep in mind that this straightforward solution does not always hold when $\boldsymbol{Q} \neq \sigma^2 \boldsymbol{I}_N$ even in the pseudo-orthogonal case.

6.4 Design Examples

Some design examples are provided in this section. A cost function to design a filter bank will depend on its application. We use here the combination of normalized coding gain and stopband attenuation. The normalized coding gain is used to avoid a trivial solution such that all filter coefficients are zero, which leads to an infinite coding gain. In the maximally decimated case, this undesirable solution is never obtained because of the biorthogonal condition $\boldsymbol{E}\boldsymbol{R} = \boldsymbol{I}_N$. However, this constraint is not imposed on the overcomplete case. Therefore, normalization of coding gain is needed. Let $\boldsymbol{C}$ be a correlation matrix of an input signal $\boldsymbol{f}$. The normalized coding gain for oversampled filter banks is given as

$$J_{\mathrm{CG}} = 10 \log_{10} \left[\prod_{n=0}^{N-1} \left(\frac{M}{\langle \boldsymbol{h}_n, \boldsymbol{g}_n \rangle N} \right)^2 \langle \boldsymbol{h}_n, \boldsymbol{C}\boldsymbol{h}_n \rangle \|\boldsymbol{g}_n\|^2 \right]^{-\frac{1}{N}}, \tag{6.37}$$

where $\boldsymbol{h}_n$ and $\boldsymbol{g}_n$ are the impulse responses of the nth analysis and synthesis filters, respectively. If the filter bank is maximally decimated, that is, $M = N$ and $\langle \boldsymbol{h}_n, \boldsymbol{g}_n \rangle = 1$, then the normalized coding gain is equivalent to the conventional one. We assume that the input signal is the first-order Markov process (AR(1) process) with the correlation coefficient $\rho = 0.95$. The stopband attenuation cost is frequently used for filter design. Minimization of this cost makes each filter a bandpass filter. The stopband attenuation costs for the analysis and the synthesis filter banks are respectively given as

$$J_{\mathrm{A}} = \sum_{n=0}^{N-1} \int_{\omega \in \Omega_{\mathrm{stopband}}} W_a(\omega) |H_n(e^{j\omega})|^2 d\omega, \tag{6.38}$$

$$J_{\mathrm{S}} = \sum_{n=0}^{N-1} \int_{\omega \in \Omega_{\mathrm{stopband}}} W_s(\omega) |G_n(e^{j\omega})|^2 d\omega, \tag{6.39}$$

where $H_n(e^{j\omega})$ and $G_n(e^{j\omega})$ are the frequency responses of the nth analysis filter and the nth synthesis filter, respectively, and $W_a(\omega)$ and $W_s(\omega)$ are weighting functions. The cost function to be used for design is defined as a linear combination of these three costs:

$$J = \alpha_1 J_{\mathrm{CG}} + \alpha_2 J_{\mathrm{A}} + \alpha_3 J_{\mathrm{S}}. \tag{6.40}$$

All design examples here were obtained by unconstrained nonlinear optimization, where we used the routines provided by MATLAB [1] version 6.1. To obtain oversample PR FBs with bandpass filters, we should choose M and N such that N is an integer multiple of M due to the mechanism of alias cancellation. Figure 6.10 shows the filter coefficients and the corresponding frequency responses of an eight-channel GLPBT in which a decimation factor is four and all filters have length 16, i.e., $L = 16, M = 4$, and $N = 8$. These filters are optimized for stopband attenuation to design bandpass filters. The initial values for the optimization are given at random. The analysis and the synthesis filters seem almost the same, since the combination coefficients α_i are chosen so that $\alpha_1 = 0$ and $\alpha_2 = \alpha_3$.

The filters obtained by maximizing the normalized coding gain are illustrated in Fig. 6.11. The filters obtained in Fig. 6.10 are used as the initial values. All filters in Fig. 6.11 are normalized such that $\langle \boldsymbol{h}_n, \boldsymbol{g}_n \rangle = M/N$ and $\|\boldsymbol{g}_n\| = 1$. In this case, the GLPBT achieves an impressive coding gain of 18.32 dB. This implies that the overcomplete representation of signals is robust against quantization. We show comparison of coding gains for various transforms in Table 6.1.

[1] MATLAB is a trademark of The Math Works Inc.

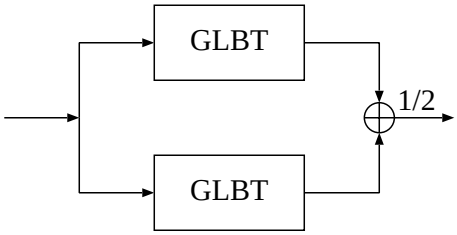


Figure 6.9: An oversampled LPPRFB organized by the parallel connection of two GLBTs

Table 6.1: Comparison of coding gains of various transforms: (M, N, L) indicates the decimation factor, the number of channels, and the length of filters, respectively. "p-GLBT" denotes the parallel connection of two GLBTs of coding gain 8.85 dB.

Transforms	DCT	LOT [8]	LBT [22]	GLBT	p-GLBT	GLPBT
(M, N, L)	$(8, 8, 8)$	$(8, 8, 16)$	$(8, 8, 16)$	$(4, 4, 16)$	$(4, 8, 16)$	$(4, 8, 16)$
Coding Gain (dB)	8.83	9.22	9.63	8.85	10.36	18.32

In order to evaluate the effect of the optimization, let us consider the eight-channel oversampled LPPRFB in which two four-channel GLBTs ($M = N = 4$) optimized for coding gain are connected in parallel as illustrated in Fig. 6.9. Each GLBT attains coding gain of 8.85 dB as listed in Table 6.1. Let J_{CG}[GLBT] be coding gain of the GLBT. Similarly, let J_{CG}[p-GLBT] be coding gain of the parallel connection of two GLBTs. Then, given a GLBT, we have the following:

$$J_{\mathrm{CG}}[\text{p-GLBT}] = J_{\mathrm{CG}}[\text{GLBT}] + \frac{20}{N} \log_{10} 2. \tag{6.41}$$

Since the GLBT attains coding gain of 8.85 dB, the above equation says that J_{CG}[p-GLBT] $\simeq$ 10.36dB, which is much lower than that of the GLPBT $(M, N, L) = (4, 8, 16)$ optimized for coding gain. It should be noted that although by the parallel connection, we can obtain higher coding gain compared to the maximally decimated case, it is much more effective to optimize free parameters in the GLPBT lattice structure.

Figures 6.12 and 6.13 illustrate a four-channel GLPBT $(M, N, L) = (3, 4, 12)$, and an eight-channel GLPBT $(M, N, L) = (7, 8, 21)$. All filters depicted in Figs. 6.12 and 6.13 are designed by optimizing the normalized coding gain with initial values which are optimized by minimizing stopband attenuation. As seen in these figures, the filters provide poor frequency selectivity compared to the FB shown in Fig. 6.10, since N is not an integer multiple of M. Coding gains

Table 6.2: Comparison of coding gains of the pseudo-orthogonal case (oversampled LPPUFBs) and the pseudo-biorthogonal case (oversampled LPPRFBs)

(M, N, L)	$(3, 4, 12)$	$(7, 8, 21)$	$(4, 8, 16)$
Pseudo-orthogonal	11.93	11.15	17.93
Pseudo-biorthogonal	12.47	11.45	18.32

attained in the pseudo-orthogonal (paraunitary) [24] and in the pseudo-biorthogonal cases are compared in Table 6.2. In all examples, the GLPBT achieves higher coding gain than the psudo-orthogonal transform. This is due to the increase of design freedom.

6.4.1 GLPBT with the Noise Robust Building Block

We confirm the effect of the noise robust GLPBT when the building block $\mathbf{\Phi}_0^+$ in the GLPBT is replaced by $\mathbf{\Xi}^*$. For comparison, we use two GLPBTs designed in the above subsection. All filter banks consist of filters of length 16 ($L = 16$), have eight channels ($N = 8$), and downsample the filtered signal by four ($M = 4$). One is a bandpass filter bank which is optimized for the stopband attenuation cost (GLPBT 1) as shown in Fig. 6.10. This is almost a pseudo-orthogonal (paraunitary) filter bank. The other is optimized for coding gain (GLPBT 2) as shown in Fig. 6.11. This is pseudo-biorthogonal.

We consider here two kinds of noise.

Case 1 The noise has the same correlation among channels. The noise is characterized by a block-diagonal matrix with the same entry, that is,

$$\boldsymbol{Q} = \mathrm{diag}[\underbrace{\mathbf{\Theta}, \ldots, \mathbf{\Theta}}_{K}], \tag{6.42}$$

where

$$[\mathbf{\Theta}]_{i,j} = \begin{cases} s_i & i = j \\ c & i \neq j \end{cases} \tag{6.43}$$

We used $\{s_i\} = \{10, 20, 30, 40, 40, 30, 20, 10\}$ and $c = 3$ in this test.

Case 2 All output coefficients are transmitted through one channel. The correlation function of the noise has exponential decay, that is,

$$[\boldsymbol{Q}]_{i,j} = e^{-\lambda|i-j|}. \tag{6.44}$$

Table 6.3: Difference in SNR (dB) between the noise robust GLPBT and the GLPBT

	White	Case 1	Case 2 ($\lambda = 0.1$)	($\lambda = 0.5$)
GLPBT 1	0	0.2177	4.5324	0.5161
GLPBT 2	0.2374	0.3868	2.6215	0.3193

We set that $\lambda = 0.1$ and 0.5 in this test.

For each case, we compare the minimal error produced by the noise robust GLPBT MSE[$\mathbf{\Xi}^*$] with the error produced by the GLPBT MSE[$\mathbf{\Phi}_0^+$]. Table 6.3 shows the difference in SNR (dB), that is,

$$-10(\log_{10} \text{MSE}[\mathbf{\Xi}^*] - \log_{10} \text{MSE}[\mathbf{\Phi}_0^+]). \tag{6.45}$$

The system can achieve PR, and therefore we can compute the difference of SNR in average on noise for any input. The table says that the differences in Case 2 ($\lambda = 0.1$) are greater than those in the other cases. This may be due to the fact the correlation matrix of Case 2 approaches to the identity matrix when λ increases.

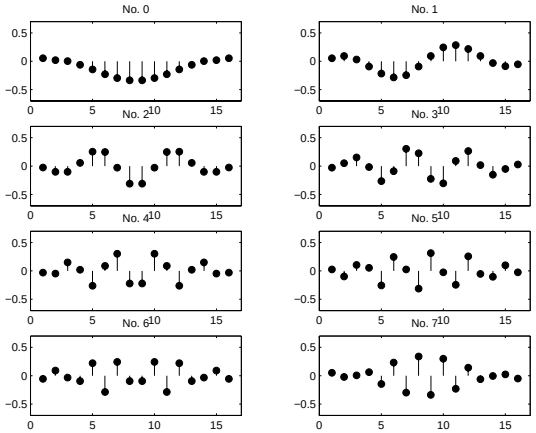
The table says that the differences in Case 2 ($\lambda = 0.1$) are greater than those in the other cases. This may be due to the fact the correlation matrix of Case 2 approaches to the identity matrix when λ increases.

6.5 Summary

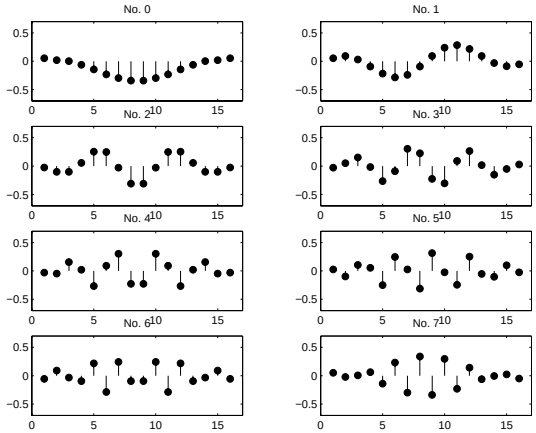
We have introduced the complete and minimal factorization of a special class of N-channel oversampled linear-phase perfect reconstruction filter banks (LPPRFBs). The analysis and synthesis filters yield a frame or a pseudo-biorthogonal basis; and therefore this filter bank is called the generalized lapped pseudo-biorthogonal transform (GLPBT) from the lapped transform perspective. The factorized filter banks are characterized by elementary rotation angles and scalar multiplications. Therefore, the lattice structure can provide fast implementation and enables us to determine the filter coefficients by solving an unconstrained optimization problem. Furthermore, we have considered the lattice structure of the GLPBT in presence of noise. We have also shown some design examples where coding gain and stopband attenuation are used as cost functions. To avoid a trivial solution, we have formulated the normalized coding gain. The characteristics

of the resulting filters are improved in the sense of coding gain because of increase in the degree of freedom. In the noisy case, given an analysis bank of the GLPBT and a correlation matrix of noise, the proposed synthesis bank suppresses the noise added in the transform domain. Experimental results show that this class of filter bank is more effective in the presence of colored noise than white or almost white noise.

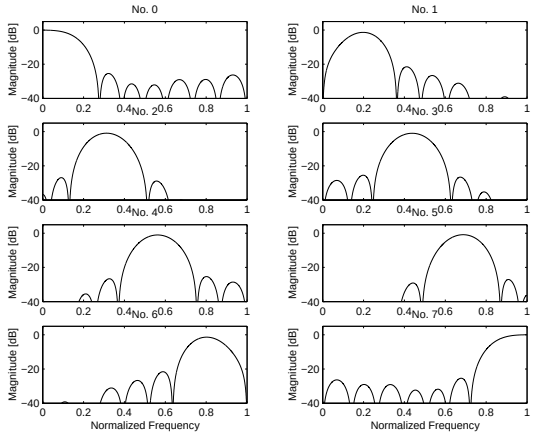
It should be noted that the theory in this chapter covers a very large class of LPPRFBs. A lot of conventional works can be regarded as a subclass of the proposed filter banks with the lattice structure. For future research, it is necessary to develop the factorization for all possible N-channel oversampled LPPRFBs.



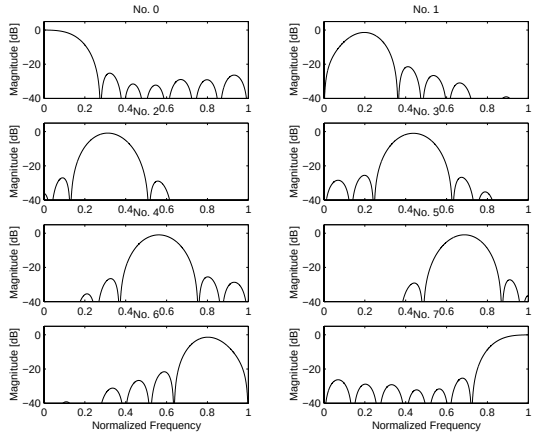
(a) Inpulse responses of the analysis filters



(b) Inpulse responses of the synthesis filters

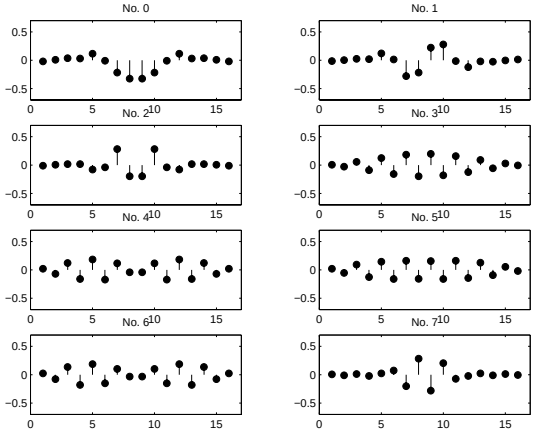


(c) Magnitude responses of the analysis filters

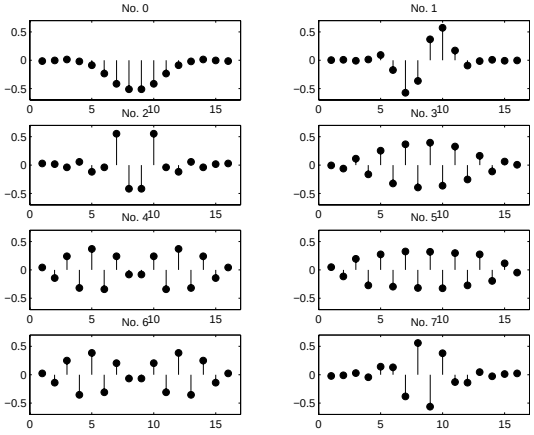


(d) Magnitude responses of the synthesis filters

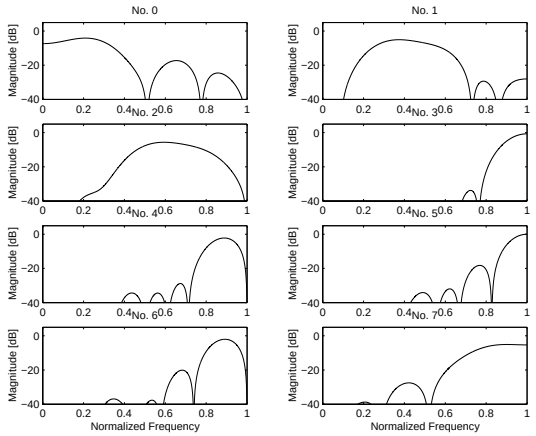
Figure 6.10: Design example for $M = 4, N = 8, L = 16$, which is optimized for stopband attenuation



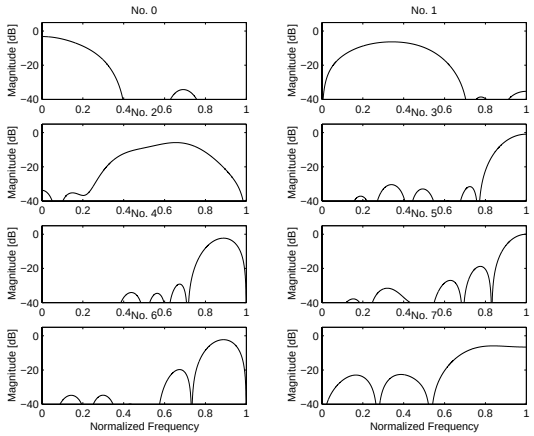
(a) Inpulse responses of the analysis filters



(b) Inpulse responses of the synthesis filters

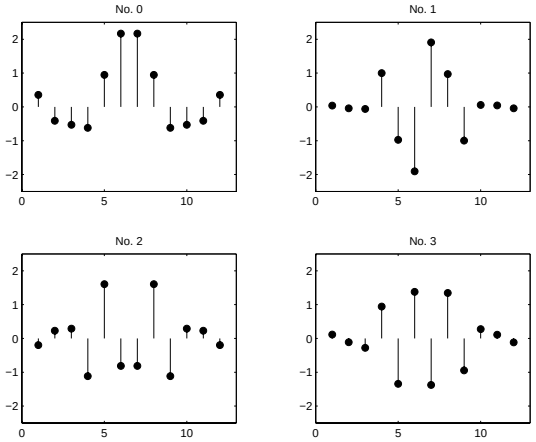


(c) Magnitude responses of the analysis filters

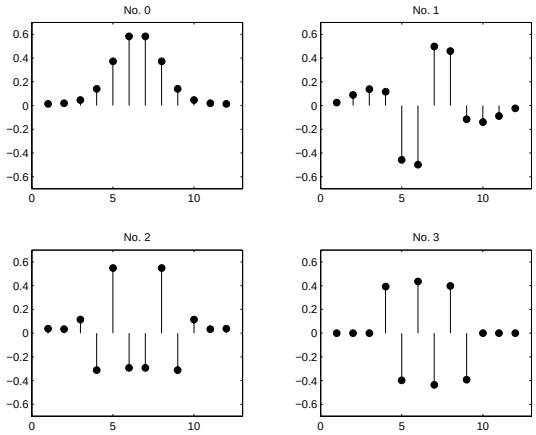


(d) Magnitude responses of the synthesis filters

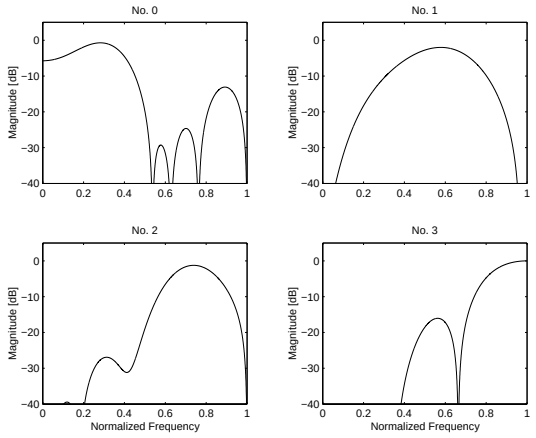
Figure 6.11: Design example for $M = 4, N = 8, L = 16$, which is optimized for coding gain



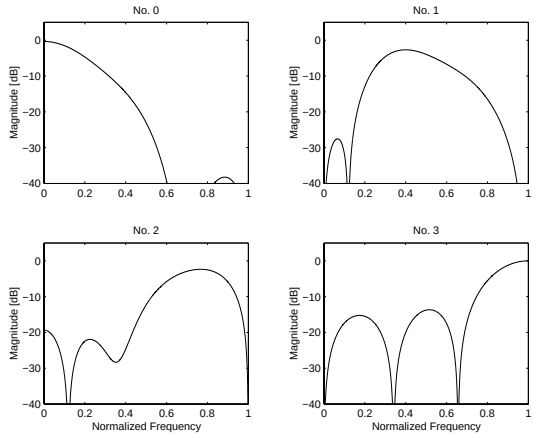
(a) Impulse responses of the analysis filters



(b) Impulse responses of the synthesis filters

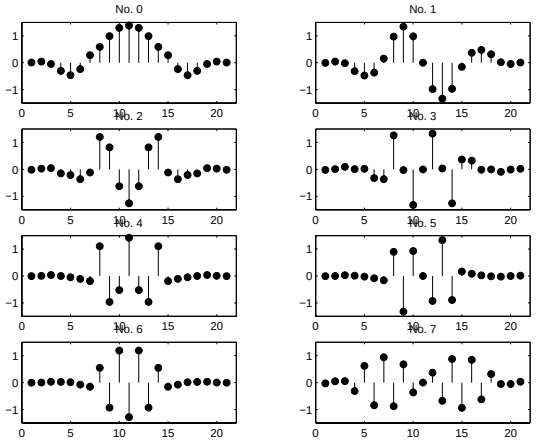


(c) Magnitude responses of the analysis filters

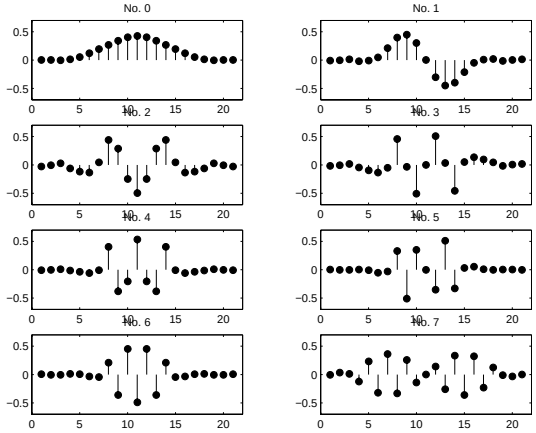


(d) Magnitude responses of the synthesis filters

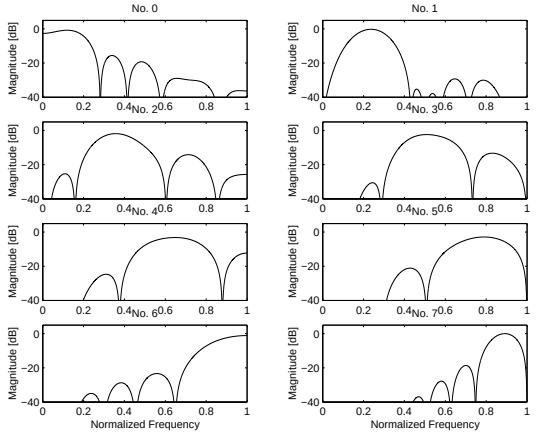
Figure 6.12: Design example for $M = 3, N = 4, L = 12$, which is optimized for coding gain



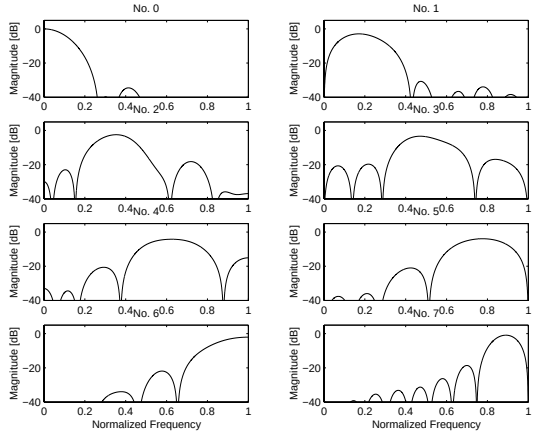
(a) Impulse responses of the analysis filters



(b) Impulse responses of the synthesis filters



(c) Magnitude responses of the analysis filters



(d) Magnitude responses of the synthesis filters

Figure 6.13: Design example for $M = 7, N = 8, L = 21$, which is optimized for coding gain

Chapter 7

Conclusions

7.1 Introduction

We have given solutions for the problems that are stated in Chapter 1. In this chapter, we conclude our works in this dissertation and describe open problems which have been clarified throughout this dissertation.

7.2 Summary

In Chapter 2, we define notations and review fundamental theories which are necessary to understand this dissertation.

In Chapter 3, we have introduced a biorthogonal lapped transform that consists of overlapping and non-overlapping basis functions (VLLBT). We have defined the VLLBT by extending conventional lapped transforms. Furthermore, we have provides a theory of the subspace Karhunen-Loève transform (SKLT). From the theory, when the biorthogonal long basis functions of the VLLBT are given, the optimal short basis functions in the energy compaction sense are derived by solving an eigenvalue problem without iterative searching techniques. This yields that the degree of freedom for the VLLBT is reduced to that for the long basis functions of the VLLBT, although in general, biorthogonalization leads to a large increase in the degree of freedom. It has been shown by comparison that the degree of freedom of the VLLBT with the SKLT is lower than that of the VLLBT without the SKLT. We also provide design examples. The resulting VLLBT attains high coding gain comparing to other lapped transforms.

In Chapter 4, we have presented the theory and design of an adaptive lapped biorthogonal

transform for image coding. The proposed transform is based on the extension of the VLLBT formulated in Chapter 3 on length of the basis functions. The overlapping basis functions have samples whose number is an integer multiple of the traditional block size. We have introduced two types of this transform: Type-E and Type-O. The Type-E requires special care for the image boundary to avoid the border distortion. In the Type-O, on the other hand, basis functions' centers of symmetry are aligned. Because of the alignment, we can use the symmetric extension method at image boundaries when we transform an input image. We show an orientation adaptive example, where each adaptive transform is characterized by the angle of edges in image blocks.

In Chapter 5, we have illustrated image coding applications for the transforms developed in Chapters 3 and 4. Comparisons among the proposed transforms and the existing transforms have been carried out. Through comparison, we have shown benefit of the proposed transforms.

In Chapter 6, we have developed a lattice structure for a special class of N-channel oversampled linear-phase perfect reconstruction filter banks with a desimation factor M smaller than N. We deal with systems in which all analysis and synthesis filters have the same FIR length and share the same center of symmetry. This class of LPPRFBs is called the generalized lapped pseudo-biorthogonal transform (GLPBT), since the present filter banks are generalizations of conventional lapped transforms. We provide the general lattice factorization of a polyphase matrix of the GLPBT. The lattice structure is based on the signular value decomposition for non-square matrices. The resulting lattice structure is able to provide fast implementation and allows us to determine the filter coefficients by solving an unconstrained optimization problem. Moreover, we have considered the case where we give the GLPBT lattice structure with specific parameters and we *a priori* know the correlation matrix of noise. We have shown that the present systems with the lattice structure cover a wide range of linear-phase perfect reconstruction filter banks. We have also shown design examples.

7.3 Open Problems

So far, we have addressed several problems, and have shown solutions of those problems throughout this dissertation. However, several new problems have arisen. We would summarize those open problems in this section, and suggest some possible directions for future work.

7.3.1 Analytic Solutions for Lapped Transforms

In Chapter 3, we have shown an analytic solution of the VLLBT when feasible long basis functions are given. However, we should find the long basis functions by numerical computation. Finding an analytic solution for all basis functions is still an open problem. Furthermore, it is more desirable to find analytic solutions for various classes of lapped transforms.

7.3.2 Adaptive Long Basis Functions

In this dissertation, the adaptation is applied only to short basis functions. However, constructing 2D non-separable long basis functions adapted for a 2D characteristic such as orientation would be very meaningful.

7.3.3 Self-Organized Classification for Adaptive Lapped Transform Coding

In this dissertation, we have shown only the orientation adaptation example. However, since the non-overlapping basis functions can be obtained from a correlation matrix by solving the eigenvalue problem, any conventional adaptation procedures based on the Karhunen-Loève transform can be applied to the proposed design method of lapped transforms. For instance, we can construct various classes of adaptive lapped transforms by training correlation matrices of a source of input signals. This self-organizing approach has been studied in the block transform cases [46, 47]. However, it may be more effective that this strategy is applied to our framework.

7.3.4 Classification Criteria for Adaptive Lapped Transforms

We choose the best transform out of a set of OALBTs by the subspace method in Chapter 5. However, this criterion never involve the amount of bits. In image coding applications, we should choose a transform which gives the best approximation for a fixed rate. Such a optimization is called the rate-distortion optimization [12]. This strategy would improve coding efficiency.

7.3.5 Complete Lattice for Oversampled LPPRFBs

The GLPBT defined in Chapter 6 is indeed a subclass of oversampled LPPRFBs. We have developed in this dissertation the complete lattice factorizaton for the GLPBT. However, The

development of a complete lattice structure is still an open problem.

7.3.6 Synthesis Polyphase Matrix for Oversampled LPPRFBs

In Chapter 6, we have introduced the final building block which can suppress noise. This choise seems the most practical and appropriate to fast implementation. However, this solution is not the global optimum of synthesis polyphase matrix. Therefore, we should find the optimal solution and construct its lattice structure for fast implementation if possible. Moreover, we can remove the PR condition to decide the synthesis FB. For example, using stastic knowledge on input signal, we can choose a synthesis FB which can provide the best approximation to the original signal such as the Wiener filter.

Bibliography

[1] H. C. Reeve III and J. S. Lim, "Reduction of blocking effects in image coding," *Oprical Eng.*, vol. 23, pp. 34–37, Jan./Feb. 1984.

[2] A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," *IEEE Trans. Circuits Syst. for Video Technol.*, vol. 2, pp. 91–95, Mar. 1992.

[3] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. Circuits Syst. for Video Technol.*, vol. 3, pp. 421–432, Dec. 1993.

[4] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images," *IEEE Trans. Image Processing*, vol. 4, pp. 896–908, July 1995.

[5] H. Paek, R.-C. Kim, and S.-U. Lee, "On the POCS-based postprocessing technique to reduce the blocking artifacts in transform coded images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, pp. 358–367, June 1998.

[6] J. Chou, M. Crouse, and K. Ramchandran, "A simple algorithm for removing blocking artifacts in block-transform coded images," *IEEE Signal Processing Letters*, vol. 4, pp. 33–35, Feb. 1998.

[7] Z. Wang and D. Zhang, "A novel approach for reduction of blocking effects in low-bit-rate image compression," *IEEE Trans. Commun.*, vol. 46, pp. 732–734, June 1998.

[8] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, pp. 553–559, Apr. 1989.

[9] P. M. Cassereau, D. H. Staelin, and G. de Jager, "Encoding of images based on a lapped orthogonal transform," *IEEE Trans. Commun.*, vol. 37, pp. 189–193, Feb. 1989.

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

[11] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Printice Hall, 1993.

[12] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. New Jersey: Prentice Hall PTR, 1995.

[13] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley MA: Wellesley-Cambridge Press, 1996.

[14] M. Vetterli and D. L. Gall, "Perfect reconstruction FIR filter banks: Some properties and factorizations," *IEEE Trans. Signal Processing*, vol. 37, pp. 1057–1071, July 1989.

[15] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The GenLOT: Generalized linear-phase lapped orthogonal transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 497–507, Mar. 1996.

[16] S. Muramatsu and H. Kiya, "A new factorization technique for the generalized linear-phase LOT and its fast implementation," *IEICE Trans. Fundamentals*, vol. E79-A, pp. 1173–1179, Aug. 1996.

[17] A. K. Soman, P. P. Vaidyanathan, and T. Q. Nguyen, "Linear phase paraunitary filter banks: Theory, factorizations and designs," *IEEE Trans. Signal Processing*, vol. 41, pp. 3480–3496, Dec. 1993.

[18] T. D. Tran, M. Ikehara, and T. Q. Nguyen, "Linear phase paraunitary filter bank with filters of different lengths and its application in image compression," *IEEE Trans. Signal Processing*, vol. 47, pp. 2730–2744, Oct. 1999.

[19] R. L. de Queiroz and T. D. Tran, "A fast lapped transform for image coding," in *Proc. IS&T/SPIE Symp. on Electronic Imaging, Image and Video Communications and Processing*, 2000.

[20] T. Nagai, M. Ikehara, M. Kaneko, and A. Kurematsu, "Generalized unequal length lapped orthogonal transform for subband image coding," *IEEE Trans. Signal Processing*, vol. 48, pp. 3365–3378, Dec. 2000.

[21] S. C. Chan, "The generalized lapped transform (GLT) for subband coding applications," in *Proc. 1995 IEEE Inter. Conf. Acoustics, Speech, and Signal Processing (ICASSP 95)*, pp. 1508–1511, May 1995.

[22] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear-phase perfect reconstruction filter bank: Lattice structure, design, and application in image coding," *IEEE Trans. Signal Processing*, vol. 48, pp. 133–147, Jan. 2000.

[23] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "Variable-length generalized lapped biorthogonal transform," in *Proc. 1998 IEEE Int. Conf. on Image Processing*, Oct. 1998.

[24] F. Labeau, L. Vandendorpe, and B. Macq, "Structures, factorizations, and design criteria for oversampled paraunitary filterbanks yielding linear-phase filters," *IEEE Trans. Signal Processing*, vol. 48, pp. 3062–3071, Nov. 2000.

[25] H. Ogawa, "A theory of pseudo biorthogonal bases," *Trans. IECE*, vol. J64-D, pp. 555–562, July 1981. (in Japanese).

[26] R. M. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic Press, 1980.

[27] H. Ogawa and N.-E. Berrached, "A theory of extended pseudo biorthogonal bases," *IEICE Trans. Inf. & Syst.*, vol. E76-D, pp. 890–897, Aug. 1993.

[28] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[29] H. Bölcdkei and F. Hlawatsch, "Oversampled cosine modulated filter banks with perfect reconstruction," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 1057–1071, Aug. 1998.

[30] H. Bölcdkei, F. Hlawatsch, and H. G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 3256–3268, Dec. 1998.

[31] M. J. T. Smith and I. T. P. Barnwell, "A new filter bank theory for time-frequency representation," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, pp. 314–327, Mar. 1987.

[32] Q.-G. Liu, B. Champagne, and D. K. C. Ho, "Simple design of oversampled uniform DFT filter banks with applications to subband acoustic echo cancellation," *Signal Processing*, vol. 80, pp. 831–847, May 2000.

[33] T. Strohmer, "Finite and infinite-dimensional models for oversampled filter banks," in *Modern Sampling Theory: Mathematics and Applications* (J. J. Benedetto and P. J. S. G. Ferreira, eds.), Boston: MA: Birkhäuser, 2001.

[34] K. Eneman and M. Moonen, "DFT modulated filter bank design for oversampled subband systems," *Signal Processing*, vol. 81, pp. 1947–1973, Sept. 2001.

[35] J. Kliewer and A. Mertins, "Oversampled cosine-modulated filter banks with arbitrary system delay," *IEEE Trans. Signal Processing*, vol. 46, pp. 941–955, Apr. 1998.

[36] H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Proc.*, vol. 46, pp. 1043–1053, Apr. 1998.

[37] T. D. Tran and T. Q. Nguyen, "On M-channel linear phase FIR filter banks and applicationin image compression," *IEEE Trans. Signal Processing*, vol. 45, pp. 2175–2187, Sept. 1997.

[38] T. Nagai and M. Ikehara, "Fast LOT with unequal length basis functions: Realization and application in subband image coding," *IEICE Trans. Fundamentals*, vol. E82-A, pp. 825–834, May 1999.

[39] W. B. Pennebaker and L. J. Mitchell, *JPEG Still Image Data Compression Standard*. New York, NY: Van Nostrand Reinhold, 1992.

[40] K. R. Rao and J. J. Hwang, *Techniques and standards for image, video, and audio coding*. New Jersey: Prentice Hall, 1996.

[41] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. COM-25, pp. 90–93, Jan. 1974.

[42] K. R. Rao and P. Yip, *Discrete cosine transform: algorithms, advantages, applications*. New York, NY: Academic Press, Inc, 1990.

[43] G. Bjøntegaard, "A novel method for compressing images using discrete directional transforms," in *Proc. SPIE Visual Commun. and Image Proc. '88*, vol. 1001, pp. 840–846, 1988.

[44] T. Tanaka and Y. Yamashita, "Vector-embedded Karhunen-Loève transform and its application in orientation adaptive coding of images," *IEICE Trans. Fundamentals*, vol. E73-A, June 2000.

[45] G. W. Wornell and D. H. Staelin, "Transform image coding with a new family of models," in *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Proc.*, pp. 777–780, 1988.

[46] K. Ohzeki, "Adaptive KL transform coding and its design method," *Trans. IEICE, Part B-I*, vol. J77-B-I, pp. 94–101, Feb. 1994. (in Japanese).

[47] R. D. Dony and S. Haykin, "Optimally adaptive transform coding," *IEEE Trans. Image Processing*, vol. 4, pp. 1358–1370, Oct. 1995.

[48] R. L. de Queiroz and K. R. Rao, "Time-varying lapped transforms and wavelet packets," *IEEE Trans. Signal Processing*, vol. 41, pp. 3293–3305, Dec. 1993.

[49] T. J. Klausutis and V. K. Madisetti, "Adaptive lapped transform-based image coding," *IEEE Signal Processing Letters*, vol. 4, pp. 245–247, Sept. 1997.

[50] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 1996.

[51] A. Ben-Israel and T. N. E. Greville, *Generalized Inverse: Theory and Applications*. New York: John Wiley & Sons, 1974.

[52] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educ. Psychology*, vol. 24, pp. 417–441, 498–520, 1933.

[53] H. Karhunen, "Über lineare Methoden in der Wahrscheinlich-Keisrechnung," *Ann. Acad. Sci. Fenn.*, Ser. A. I. 37, Helsinki, 1947.

[54] M. Loève, "Fonctions aléatoires de seconde ordre," in P. Lévy, *Processus Stochastiques et Mouvement Brownien,* Hermann, Paris, 1948.

[55] S. J. Leon, *Linear algebra with applications.* Englewood Cliffs, NJ: Prentice Hall, 1994.

[56] Y. Yamashita and H. Ogawa, "Relative Karhunen-Loève transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 371–378, Feb. 1996.

[57] D. G. Luengerger, *Optimization by vector space method.* New York, NY: Wiley, 1969.

[58] H. Ogawa, "Karhunen-Loève subspace," in *Proc. of the 12th Int. Conf. on Pattern Recognition*, vol. 2, (Hague, Netherlands), pp. 75–78, 1992.

[59] R. J. Clarke, "Relation between the Karhunen-Loève and cosine transforms," *IEE Proc. Pt. F*, vol. 128, pp. 359–360, Nov. 1981.

[60] M. Unser, "On the approximation of the discrete Karhunen-Loève transform for stationary processes," *Signal Processing*, vol. 7, pp. 231–249, Dec. 1984.

[61] W.-H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, vol. COM-25, pp. 1285–1295, Nov. 1977.

[62] W. D. Ray and R. M. Driver, "Futher decomposition of the Karhunen-Loève series representation of a stationary random process," *IEEE Trans. Information Theory*, vol. IT-16, pp. 663–668, Nov. 1970.

[63] H. C. Andrews and W. K. Pratt, "Fourier transform coding of images," in *Hawaii Internat'l Conf. on System Sciences*, pp. 677–679, Jan. 1968.

[64] G. B. Anderson and T. S. Huang, "Piecewise Fourier transformation for picture bandwith compression," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 133–140, Apr. 1971.

[65] W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard transform image coding," *Proc. IEEE*, vol. 57, pp. 58–68, Jan. 1969.

[66] W. K. Pratt, W. Chen, and L. R. Welch, "Slant transform image coding," *IEEE Trans. Commun.*, vol. COM-22, no. 8, pp. 1075–1093, 1974.

[67] H. Enomoto and K. Shibata, "Orthogonal transform coding system for television signals," *IEEE Trans. Electromagn. Compat.*, vol. EMC-13, pp. 11–17, Aug. 1971.

[68] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Addison-Wesley, 1992.

[69] W. K. Pratt, *Digital image processing*. New York, NY: J. Willy, 1991.

[70] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Boston, MA: Kluwer Academic Publishers, 1992.

[71] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoustics, Speech, and Signal Proc.*, vol. 38, pp. 969–978, June 1990.

[72] T. Tanaka and Y. Yamashita, "Image coding using vector-embedded Karhunen-Loève transform," in *Proc. 1999 IEEE Int. Conf. on Image Proc.*, Oct. 1999.

[73] A. Cantoni and P. Butler, "Properties of the eigenvectors of persymmetric matrices with applications to communication theory," *IEEE Trans. Commun.*, vol. COM-24, pp. 804–809, Aug. 1976.

[74] J. Makhoul, "On the eigenvectors of symmetric Toeplitz matrices," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. ASSP-29, pp. 804–809, Aug. 1981.

[75] J. Katto and Y. Yasuda, "Performance evaluation of subband coding and optimization of its filter coefficients," in *Proc. SPIE Conf. Visual Commun. and Image Processing*, vol. 1605, pp. 95–106, 1991.

[76] S. O. Aase and T. A. Ramstad, "On the optimality of nonunitary filter banks in subband coders," *IEEE Trans. Image Processing*, vol. 4, pp. 1585–1591, Dec. 1995.

[77] M. Helsingius, P. Kuosmanen, and J. Astola, "Image compression using multiple transforms," *Signal Process.: Image Commun.*, vol. 15, pp. 513–529, Mar. 2000.

[78] T. Tanaka and Y. Yamashita, "Adaptive transforms with overlapping basis functions for image coding," *J. Electronic Imaging*, vol. 10, pp. 706–719, July 2001.

[79] H. Kiya, K. Hishikawa, and M. Iwahashi, "A development of symmetric extension method for subband image coding," *IEEE Trans. Image Processing*, vol. 3, pp. 78–81, Jan. 1994.

[80] R. H. Bamberger, S. L. Eddins, and V. Nuri, "Generalized symmetric extension method for size-limited multirate filter banks," *IEEE Trans. Image Processing*, vol. 3, pp. 82–87, Jan. 1994.

[81] E. Kofidis, S. Theodoridis, and N. Kalouptsidis, "On the perfect reconstruction problem in *n*-band multirate maximally decimated FIR filter banks," *IEEE Trans. Signal Processing*, vol. 44, pp. 2439–2455, Oct. 1996.

[82] P. P. Vaidyanathan and T. Chen, "Role of anticausal inverses in multirate filter-banks—Part I: System-theoretic fundamentals," *IEEE Trans. Signal Processing*, vol. 43, pp. 1090–1102, May 1995.

[83] A. Albert, *Regression and the Moore-Penrose Pseudoinverse*. New York, NY/London: Academic Press, 1972.