A NOVEL BOUNDARY HANDLING SCHEME FOR ARBITRARY OBJECT SHAPE IN VIDEO AND IMAGE COMPRESSION

Ka Shun Carson Pun and Truong Q. Nguyen

Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093-0407
carson@ieee.org, nguyent@ece.ucsd.edu

ABSTRACT

This paper presents a novel scheme for handling image boundary for a class of biorthogonal multi-dimensional (MD) perfect reconstruction (PR) filter banks (FB) using lifting scheme which finds important applications in image and video compression applications. An intuitive example is given to demonstrate our proposed for obtaining a nonexpansive transformation using nonseparable FB on arbitrary image shape.

1. INTRODUCTION

Handling the image boundaries for MD PR FB is a very important topic in non-expansive transformation. It is the central problem in non-expansive transformation which found important applications in image compression. In many image compression schemes, basically a 2-D image is first transformed by an analysis FB[1]. Then these subband samples are quantized and compressed by lossless entropy coding. To retrieve the image, the encoded bit-stream is decompressed to obtain the transformed coefficients. Then they are passed to a synthesis FB to reconstruct the original image. The above process requires two important properties: The first property is that the overall FB must be a perfect reconstruction (PR) FB[1,2]. The PR property guarantees that if there are no quantization and compression, the synthesis FB can reproduce the original input signal from the subband samples. The second property requires the transformation is non-expansive which is very important in compression applications. Non-expansive means that the number of input samples must equal to the number of subband samples. When the filtering is performed across the image boundary, the filtered image will spread over the edges. Therefore total number of subband samples are more than those of the original image.

[0]This work is supported by Skyworks Inc. and UC DiMi.

To obtain a non-expansive transformation, many researchers proposed various methods, like zero padding, circular convolution, symmetric extension, etc.[2]. In zero padding, the images are assumed to be zero outside the image borders. Therefore a large discontinuity appears at the image boundaries. The reconstructed pixels near the boundaries have very low pixel values which induce a large amount of errors. For the circular convolution, it is assumed that the image is periodic. Obviously, it is not the case and a large discontinuity will be generated at image boundaries. Using circular convolution, one can reconstruct the original image when there is no quantization. However, when quantization occurs, errors will generate at the image boundaries. In contrast, the symmetric extension mirrors the image across the image boundary. Therefore the extended image is continuous at the image border and combining symmetric filtering with 1 dimensional linear phase FB, one can easily obtain a non-expansive image transformation. In practical image compression engine, like in JPEG2000[3,4] and Adaptive Scanned Wavelets Difference Reduction (ASWDR)[5], the image is first mirror extended before filtering. Most of these boundary extension techniques only work with symmetric filters. For arbitrary filter impulse responses, like non-separable or even low-delay filters, to the authors best knowledge, obtaining an efficient non-expansive, invertible transform is non-trivial.

More recently, in video compression application, the object based video compression is of big interest. In object based video compression, an object is defined to be separately encoded without taking the background into the account. There are many methods for handling arbitrary shaped regions: two of the most popular methods are the shape adaptive DCT (SA-DCT) and the shape-adaptive DWT (SA-DWT). For SA-DCT [6], the object is first divided into many small blocks. For blocks which are totally within the object boundary, the conventional DCT is used for image transformation. When a block overlaps an edge, the pixels within that block will be flushed to an edge before the DCT transformation is performed. Hence the spatial correlation of the object is lost. Furthermore, this block based DCT suffers from blocking artifacts at low bit-rate. For the SA-DWT [7], the wavelet transform is performed row-wise and column-wise sequentially. In order to handle the boundary problem, for each image row or column, it is symmetric extended before performing the wavelet transform. Then the transformed coefficients outside the image boundary is removed. In these image compression schemes, the filters are all symmetric so as to handle the boundary very easily. On the other hand, for filters that are non-separable and without the symmetric properties, the object boundary handling is non-trivial.

In this paper, we propose a novel scheme to handle MD signals with arbitrary boundary for a class of bio-orthogonal FB [8] to obtain a non-expansive transformation. The FB can be designed as nonseparable FB through appropriate change of variables. Furthermore, this FB is structurally PR, because arbitrary operations at the lifting steps will not destroy the FB PR property. Due to this property, this FB is used as the basic building block for Directional FB (DFB) which finds many applications in directional image analysis [9–11] and potential image compression applications [11]. In [9–11], the DFB is constructed by a tree structure of the 2-channel biorthogonal 2D PR FB. A particular problem of the DFB is that its impulse responses of each channel are nonseparable. Therefore, one cannot easily obtain a non-expansive transformation using the DFB while having the PR property. To our best knowledge, there exists no previous results for solving the boundary problem for this class of DFB.

Furthermore, in the two-channel PR FB[8], filtering only occurs in lifting branches. For finite-length signals, after filtering, the signal length will be increased. Hence, in the analysis FB outputs, there are more samples. To keep the size of the input signal and the output signal to be the same, we utilize the FB structural PR property: since this FB is parameterized by arbitrary functions of $\alpha(z)$ and $\beta(z)$ (in Fig. 1), we can perform the same boundary extension and boundary removal in these subfilters without losing the PR property. To be more precise, once the subband signal goes into the lifting branches, its boundary is smoothly extended and filtered. Similar operations are performed at the synthesis FB in order to obtain a PR FB. This paper is organized as follows: In section 2, a brief review of bio-orthogonal PR FB will be given. Then in section 3, the proposed solutions will be given together with an intuitive example. Finally, in section 4, conclusions are drawn.

2. A REVIEW OF THE BIO-ORTHOGONAL PR FB

In this section, a review of the PR FB[8] is given. Detailed design information is given in [8, 12]. The 1-D 2-channel FB [8] is shown in Fig. 1.

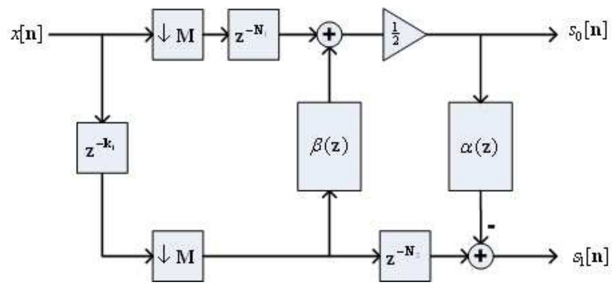

Fig. 1. FB proposed by Phoong *et. al.* [8].

where the lowpass filter $H_0(z)$ and the highpass filter $H_1(z)$ are given as follows,

$$\begin{aligned} H_0(z) &= \frac{1}{2}(z^{-2N_1} + z^{-1}\beta(z^2)), \\ H_1(z) &= -\alpha(z^2)H_0(z) + z^{-2N_2-1} \end{aligned} \tag{1}$$

From (1), the overall FB is defined by parameters $\beta(z)$, $\alpha(z)$, N_1 and N_2. By choosing the appropriate parameters, FB with very good selectivity, linear-phase or low-delay properties can be designed [12]. The ideal frequency responses of $\beta(e^{j\omega})$ and $\alpha(e^{j\omega})$ (denoted by $\beta_d(e^{j\omega})$ and $\alpha_d(e^{j\omega})$ respectively) are [12]:

$$\begin{aligned} \beta_d(e^{j\omega}) &= e^{-j(N_1-\frac{1}{2})\omega} \qquad \text{for } \omega \in [0, \pi] \\ \alpha_d(e^{j\omega}) &= e^{-j(N_2-N_1+\frac{1}{2})\omega} \qquad \text{for } \omega \in [0, \pi] \end{aligned} \tag{2}$$

In general, it is impossible to have ideal frequency responses using finite length FIR or IIR $\beta(z)$(and also $\alpha(z)$). However, very good frequency response can be obtained as the lengths of these subfilters are increased.

A simple method for obtaining a 2-D PR is by transforming the 1-D PR FB [8]. More precisely, $\alpha(z^2)$ and $\beta(z^2)$ are transformed to $\alpha(\mathbf{z}^{\mathbf{M_0}})\alpha(\mathbf{z}^{\mathbf{M_1}})$ and $\beta(\mathbf{z}^{\mathbf{M_0}})\beta(\mathbf{z}^{\mathbf{M_1}})$ respectively. The 1-D delay z^2 is replaced by $\mathbf{z}^{\mathbf{M}_0}\mathbf{z}^{\mathbf{M}_1}$. Careful examination shows that the same transformation can be carried out if $\alpha(z^2)$ and $\beta(z^2)$ are not the same function. With the sampling matrix $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, we have $H_0(z_0, z_1) = \frac{1}{2}(z_0^{-2N} + z_0^{-1}\beta(z_0 z_1)\beta(z_0 z_1^{-1}))$ and $H_1(z_0, z_1) = -\alpha(z_0 z_1)\alpha(z_0 z_1^{-1})H_0(z_0, z_1) + z_0^{-2M-1}$. For the diamond spectral support FB, $\beta(z)$ and $\alpha(z)$ will be chosen as given in (2). Other spectral support can be achieved through appropriate choice of transformations[11, 13]. Actually, the transformed FB can be further applied with a resampling matrix before the analysis and its inverse of the resampling matrix after the synthesis side without losing the PR property[11]. A resampling matrix $\mathbf{Q}$ is defined as $\det(\mathbf{Q}) = \pm 1$. i.e., for a decimator with a resampling matrix $\mathbf{Q}$, it will not drop any samples but only change their positions. At the synthesis side, the resampling matrix and the down sampling matrix can be combined together. Consider the resampling matrix $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. The effect is to have the original rectangular image sheared into a parallelogram shape before the sheared image is filtered. This type of structure occurs in DFB [11] which shears the image before passing it into a 2-channel FB. Since the original image becomes a sheared image, symmetric extension fails to provide a non-expansive transformation. In the next section, our proposed solution which can handle a wide variety of image shape will be presented.

3. PROPOSED SOLUTION AND EXAMPLE

In this section, the proposed method of handling boundary problem will be given. For this class of MD FB, the only filtering operations occur in the sub-filter $\alpha(z)$ and $\beta(z)$. When the image boundary passed through these pairs of sub-filter, the image boundaries will be spread out if no appropriate boundary handling is performed. Actually, we can remove the filtered samples which lay outside the image boundaries without losing the PR property. Doing so will not affect the PR condition because this operation still occurs within the lifting branches at both the synthesis and the analysis sides. The effect of the removals will be canceled at the synthesis side and the PR condition is held. In addition, any boundary extension can be performed within the lifting branches without losing the PR property.

If we denote the operation which removes the samples outside the image boundary as $\mathbf{B}_R(\cdot)$, and the operation which does boundary extension algorithm as $\mathbf{B}_E(\cdot)$ then the whole PR FB structure will be as shown in Fig. 3. The FB is structurally PR in the sense that, after operation occurs in the lifting branches at the analysis side can be undone with the corresponding branches at the synthesis side. Due to this special property, even though the coefficients within the lifting branches are quantized, or some nonlinear operations are performed, the whole FB is still PR.

3.1. Example

To better illustrate the idea, a detailed descriptive FB diagram is shown in Fig. 3. Consider a rectangular image first downsampled by a $\mathbf{M}$ with $|\det(\mathbf{M})| = 2$ which downsamples and rotates the image into a rotated parallelogram.

This FB in general can contain many lifting steps, but for ease of discussion, we focus on the first lifting step. The rotated parallelogram image is first boundary extended by the block $\mathbf{B}_E(\cdot)$ before filtered by $\beta(\mathbf{z})$. After filtering, the image size is increased. In order to achieve non-expansive transformation, the redundant samples are removed by the block $\mathbf{B}_R(\cdot)$. The role of the block $\mathbf{B}_R(\cdot)$ is to keep the size of the filtered image be exactly the same as the upper channel signal. Therefore these two signals can exactly overlap each other. When they are added or subtracted, the overall number of samples at each lifting stage outputs will be the same as the input signals. This process is repeated at the synthesis side with appropriate delay parameters. Finally the upsampling operator will rotate the parallelogram back to the original shape.

The choice of the boundary extension block $\mathbf{B}_E(\cdot)$ can be removed without losing PR condition. The role of this block is to smoothly extend the image signal so that in some applications, say, image compression, discontinuity across image boundary will not occur. This proposed structure can be applied to DFB proposed in [11], without major modification so as to achieve non-expansive directional image transformation. It is also noted that the proposed method can also handle arbitrary image shapes without major difficulties. Note that conventional SA-DWT[7] for object based processing cannot handle non-convex shape object efficiently. Consider the star image as shown in Fig. 2.

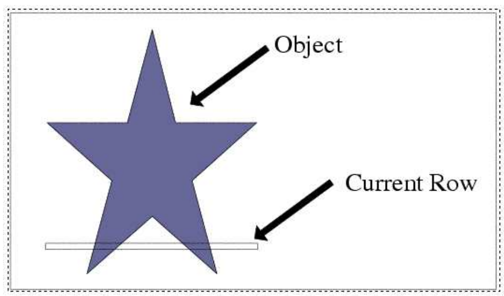

Fig. 2. A particular problematic image object

The star is the object to be transformed which is in gray color. The white areas are the background. The star object is going to be transformed without the background. If using conventional method, i.e. row-wise filtering and performing symmetric extension at the object boundaries, we will run into the problem when the current row (denoted by the lower black arrow) crosses 2 legs of the star. This is because, in the area between the 2 legs, the symmetric extended pixels from both sides may not be smoothly extended when they meet. Furthermore, imagine the case when the current row overlapped many boundaries within a finite length. Then the symmetric extension is rather complicated, because the filter has to cross over many intermittent object and background regions. However, using our proposed structure, by dropping out the transformed coefficients outside the image boundary (a very efficient operation), a non-expansive image transform is obtained without losing PR property.

4. CONCLUSION

In this paper, a novel scheme for handling arbitrary image shape for non-expansive image transformation using a class of biorthogonal PR MD FB is proposed. The proposed method is easy to implement and has a low complexity. One can remove the transformed coefficients which lies outside the boundary without losing the PR property. This work provides a solution for non-expansive DFB which have potential applications in image compression. The proposed structure can also be extended to other lifting structures (especially non-separable filters) and has potential applications in video object coding where arbitrary boundaries are needed to be handled. A practical codec is considered in future work.

5. REFERENCES

[1] P. P. Vaidyanathan, *Multirate systems and filter banks*. Englewood Cliffs, NJ, USA: Prentice Hall, 1993.

[2] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley Cambridge Press, 1996.

[3] *JPEG 2000 Part I: Final Draft International Standard*, (ISO/IEC FDIS 15444-1), ISO/IEC JTC1/SC29/WG1 N1855, Aug. 2000.

[4] *JPEG 2000 Requirements and Profiles*, ISO/IEC JTC1/SC29/WG1 N1271, Mar. 1999.

[5] J. S. Walker and T. Q. Nguyen, "Adaptive scanning methods for wavelet difference reduncion in lossy image compression," *Proc. IEEE ICIP*, vol. 3, pp. 182–185, Sept. 2000.

[6] T. Sikora and B. Makai, "Low complex shape-adaptive dct for generic and functional coding of segmented video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 59–62, Feb. 1995.

[7] S. Li and W. Li, "Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 725–743, Aug. 2000.

[8] S. M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," *IEEE Trans. Signal Processing*, vol. 43, pp. 649–665, Mar. 1995.

[9] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 882–893, Apr. 1992.

[10] S. Park, M. J. T. Smith, and J. J. Lee, "Fingerprint enhancement based on the directional filter bank," in *Proc. IEEE ICIP*, vol. 3, 2000, pp. 793–796.

[11] M. N. Do, "Directional multiresolution image representations," Ph.D Thesis, Swiss Federal Institute of Technology Lausanne, Nov. 2001.

[12] J. S. Mao, S. C. Chan, W. Liu, and K. L. Ho, "Design and multiplier-less implementation of a class of two-channel PR FIR filterbanks and wavelets with low system delay," *IEEE Trans. Signal Processing*, vol. 48, pp. 3379–3394, Dec. 2000.

[13] S. C. Chan, K. S. Pun, and K. L. Ho, "On the design and implementation of a class of multiplierless two-channel 1D and 2D nonseparable PR FIR filterbanks," in *Proc. IEEE ICIP*, vol. 2, Oct. 2001, pp. 241–224.

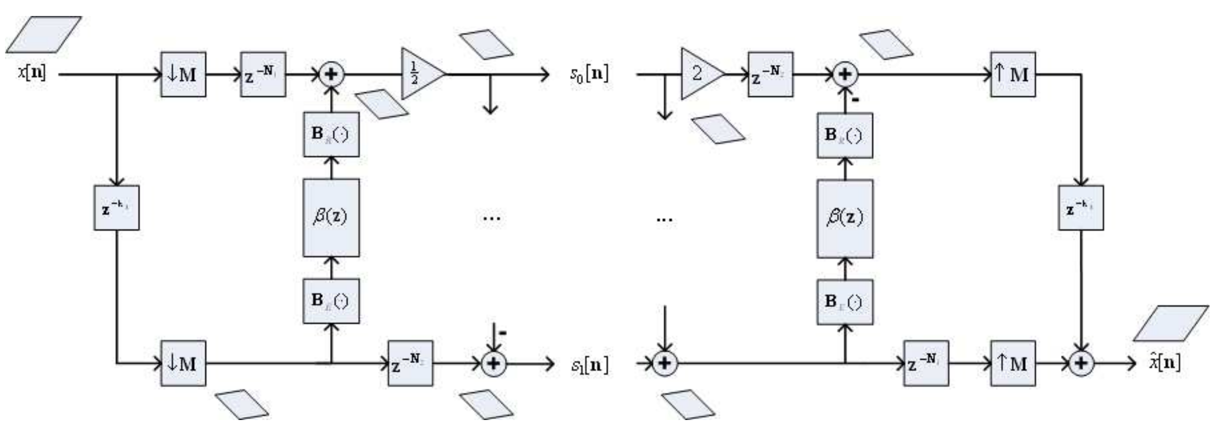

Fig. 3. The overall structure of the lifting operation and the boundary handling. (The blue images show the relative image size and orientation at various point of the FB.)