

## MIT Open Access Articles

*Directional wavelet transforms for  
prediction residuals in video coding*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Kamisli, F., and J.S. Lim. "Directional wavelet transforms for prediction residuals in video coding." Image Processing (ICIP), 2009 16th IEEE International Conference on. 2009. 613-616. © 2009 IEEE

**As Published:** <http://dx.doi.org/10.1109/ICIP.2009.5413857>

**Publisher:** Institute of Electrical and Electronics Engineers

**Persistent URL:** <http://hdl.handle.net/1721.1/59407>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# DIRECTIONAL WAVELET TRANSFORMS FOR PREDICTION RESIDUALS IN VIDEO CODING

*Fatih Kamisli and Jae S. Lim*

Research Laboratory of Electronics  
Department of Electrical Engineering and Computer Science  
Massachusetts Institute of Technology

## ABSTRACT

Various directional transforms have been developed recently to improve image compression. In video compression, however, prediction residuals of image intensities, such as the motion compensation residual or the resolution enhancement residual, are transformed. The applicability of the directional transforms on prediction residuals have not been carefully investigated. In this paper, we briefly discuss differing characteristics of prediction residuals and images, and propose directional transforms specifically designed for prediction residuals. We compare these transforms with the directional transforms proposed for images using prediction residuals. The results of the comparison indicate that our proposed directional transforms can provide better compression of prediction residuals than the directional transforms proposed for images.

*Index Terms*— Video coding, Wavelet transforms

## 1. INTRODUCTION

Transforms are fundamental components of image and video compression systems. In image compression, the image intensities are transformed, whereas in video compression, prediction residuals of image intensities are often transformed. Examples of prediction residuals include the motion-compensation-residual (MC-residual), the resolution-enhancement-residual (RE-residual) in scalable video coding and the disparity-compensation-residual (DC-residual) in multiview coding. Typically, transforms used to compress images are also used to compress prediction residuals. For example, the Discrete Cosine Transform (DCT) is used to compress images in the JPEG standard and MC-residuals in the MPEG-2 standard. Another example of such transforms is the Discrete Wavelet Transform (DWT), which is used to compress images in the JPEG2000 standard and high-pass prediction residual frames in inter-frame wavelet coding [1]. However, prediction residuals may have different characteristics from images [2, 3] (See Figure 1 for a visual comparison). Therefore, it is of interest to study if transforms better than those used for images can be developed for prediction residuals.

Recently, significant research has been performed to develop transforms that can take advantage of locally anisotropic features in images [4, 5, 6, 7, 8]. Conventionally, the 2-D DCT or the 2-D DWT is carried out as a separable transform by cascading two 1-D transforms in the vertical and horizontal dimensions. This approach does not take advantage of the locally anisotropic features present in images because it favors horizontal or vertical features over others. These other approaches adapt to locally anisotropic features in images by performing the filtering along the direction where the image intensity variations are smaller. This is achieved, for example, by directional lifting implementations of the DWT [4]. Even though most

of the work is based on the DWT, similar ideas have been applied to DCT-based image compression [6]. It appears that these ideas have not been applied to modeling and compressing prediction residuals in video coding.

In [3], we proposed directional block transforms for the MC-residual and showed potential gains achievable with these transforms within the H.264/AVC codec. In this paper, we propose directional wavelet transforms for prediction residuals. We compare these transforms with the separable wavelet transform as well as some directional wavelet transforms proposed for images. The remainder of the paper is organized as follows. In the next section, we discuss differing characteristics of images and prediction residuals to motivate the proposed transforms. In Section 3, we first review some of the directional wavelet transforms proposed for images and then propose the directional wavelet transforms for prediction residuals. Some experimental results are then provided in Section 4, along with a conclusion in Section 5.

## 2. ANISOTROPIC FEATURES IN IMAGES AND PREDICTION RESIDUALS

Prediction residuals are the errors in predicting image intensities from previously encoded image intensities. It is well known that characteristics of images vary considerably from region to region in an image. Prediction residuals in video coding also have greatly varying characteristics. Figure 1 shows an image, its MC-residual and its RE-residual. In the MC-residual, prediction errors in smooth and slowly moving regions are smaller than in moving texture regions or edges. In the RE-residual, prediction errors in smooth regions are much smaller than in detailed texture regions or around edges.

In these prediction residuals, large prediction errors often concentrate in detailed texture regions or along edges. A simple inspection of the residual images indicates that one-dimensional structures along such edges occur in many regions of the residuals. Within a local neighborhood, many pixel intensities may be close to zero except those ones along edges. It appears that using two-dimensional transforms with basis functions that have square support is not the best choice for such regions. Therefore, we propose to use transforms with basis functions whose support follows the one-dimensional structures of the prediction residuals. Specifically, we propose to use one-dimensional directional DWT's on prediction residuals. These transforms are discussed in the next Section, along with a review of other directional wavelet transforms that were previously proposed for images.



**Fig. 1.** An image (mobile sequence at CIF resolution, frame 3), its MC-residual (predicted from frame 2) and its RE-residual (predicted from QCIF resolution).

### 3. DIRECTIONAL WAVELET TRANSFORMS

Many approaches have been developed to overcome the limitations of separable wavelet transforms in image processing. Some of these approaches are bandelets [5], directionlets [7], and lifting-based wavelet transforms with directional prediction [4, 8]. In this paper, we focus on the lifting-based approaches because our proposed directional transforms can be explained more easily within the lifting framework. We first briefly review the lifting implementation of the DWT and then discuss the lifting-based approaches with directional prediction proposed for images. Finally, we introduce the lifting-based approach with directional prediction for prediction residuals.

#### 3.1. Lifting-Based Wavelet Transform

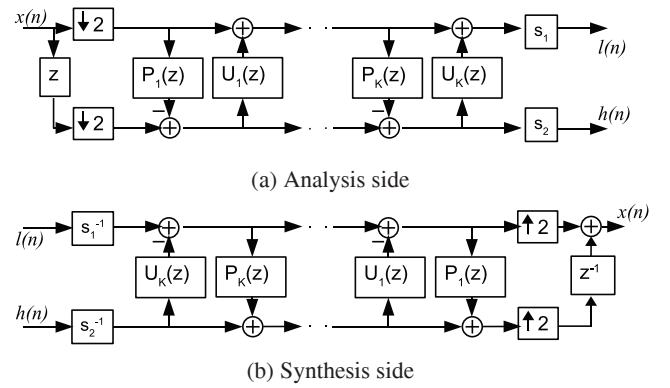
Lifting is a procedure to design wavelet transforms using a series of filtering steps called lifting steps [9]. As shown in Figure 2-a, the signal is first divided into even and odd samples and the odd samples are predicted from the even samples. The residual in that prediction is then used to update the even samples. Any number of prediction and update pairs can be cascaded until the final low-pass and high-pass signals of the transform are obtained. No matter how the prediction and update boxes in Figure 2 are chosen, this scheme is always invertible. The inverse transform is given in Figure 2-b.

The filters used for prediction and update determine the analysis and synthesis filters of the DWT. For example, the prediction and update filters shown below, result in the well-known 9/7 biorthogonal wavelet filters. A close inspection of these filters reveals that every odd sample is predicted by averaging and scaling the two neighboring even pixels, and every even pixel is updated by averaging and scaling the two neighboring odd pixels of the prediction residual.

$$\left. \begin{aligned} P_1(z) &= +1.58613(1 + z^{-1}) & U_1(z) &= -0.05298(1 + z^{+1}) \\ P_2(z) &= -0.88291(1 + z^{-1}) & U_2(z) &= +0.44350(1 + z^{+1}) \\ s_1 &= 1.23017 & s_2 &= 1/s_1 \end{aligned} \right\} (1)$$

#### 3.2. Lifting-Based 2-D Wavelet Transform with Directional Prediction (2D-dir-DWT)

To apply a separable 2-D DWT on a 2-D signal using the lifting implementation, 1-D DWT's with lifting implementation in the vertical and horizontal dimensions can be cascaded. Lifting-based wavelet



**Fig. 2.** Block diagram of the lifting implementation of the wavelet transform [9].

transform with directional prediction is performed by choosing the pixels from which a prediction (or update) is formed in an intelligent manner. These pixels are chosen along a direction which is not necessarily the horizontal or vertical direction as it is the case for the lifting implementation of the separable 2-D DWT. Figure 3-a shows several options that can be used along the vertical dimension. To predict the pixels in an odd row, fractional-pixels (interpolated from pixels in the same row) or full-pixels from even rows aligned along a particular direction can be used. In the update step, pixels in even rows are updated from the prediction residuals in odd rows aligned along the same direction. After subsampling in the vertical dimension to form the low-pass and high-pass signals, similar directional prediction and update operations can be performed along the horizontal dimension, separately on the low-pass (Figure 3-b) and high-pass signals. The low-low signal can be transformed again using directional lifting operations to obtain multilevel directional subband decompositions.

Figure 3 shows the directional prediction options proposed in [4]. Other prediction options have also been proposed [8]. In fact, to predict (update) a pixel in an odd (even) row, any pixel from any even (odd) row can be used. Typically, however, nearby pixels are likely to provide better prediction.

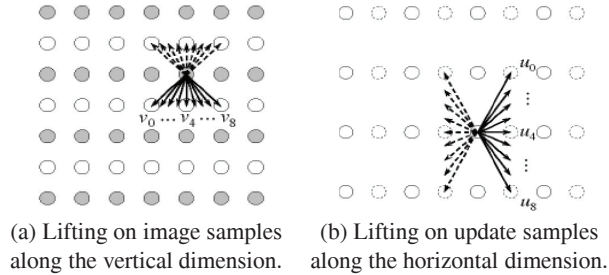


Fig. 3. Directional prediction options in [4].

### 3.3. Lifting-Based 1-D Wavelet Transform with Directional Prediction (1D-dir-DWT)

We propose lifting-based 1-D wavelet transform with directional prediction by applying the directional lifting steps only in the vertical or horizontal dimension. In other words, when performing the transform along the vertical dimension, either one of the prediction and update options in Figure 3-a is performed, or no prediction and update is performed. If no prediction and update is performed, then one of the prediction and update options in Figure 3-b is used along the horizontal dimension. If one of the options is performed, then no prediction and update is performed along the horizontal dimension.

For prediction residuals, directional 1-D DWTs can be superior to directional 2-D DWTs. As discussed in Section 2, the characteristics of prediction residuals are more coherent with the basis functions of directional 1-D DWTs. Much of the energy of prediction residuals concentrates along edges and object boundaries, forming one-dimensional structures. Transforms with basis functions whose support follow these one-dimensional structures can potentially perform better in approximating such regions of prediction residuals.

Even though 1-D directional transforms improve the compression of prediction residuals, 2-D transforms can also be used. There are regions of prediction residuals which can be better approximated with 2-D transforms. Therefore, in our experiments, we combine 1-D directional transforms with 2-D separable transforms. In other words, we perform locally (on a block basis) either a 1-D directional DWT, or a 2-D separable DWT, or no transform at all.

## 4. EXPERIMENTAL RESULTS

We present experimental results to compare the transforms discussed in Section 3 on MC-residuals and RE-residuals. We use a total of 10 CIF resolution video sequences for the experiments. From each sequence, we use a specific frame to compute its motion-compensated residual (P-frame with 8x8-pixel blocks and quarter-pixel motion estimation), and its resolution-enhancement residual (interpolated from QCIF resolution using the method in the reference software of H.264/AVC). Specifically, we compress a total of 20 (=10x2) prediction residual frames with each of the following transforms:

- 2-D separable DWT (2D-sep-DWT)
- 2-D directional DWTs (2D-dir-DWTs)
- 1-D directional DWTs and 2-D separable DWT (1D-dir-DWTs + 2D-sep-DWT).

We use only the prediction options that are shown in Figure 3. We select the best transform (i.e. prediction option) in a local region (8x8-pixel block) with rate-distortion optimization. Rate-distortion

optimization is performed by minimizing a Lagrangian cost function, formed from the mean-square-error and the number of nonzero transform coefficients, over each available prediction option. The lifting filters that were used in the experiments are the ones of the 9/7 biorthogonal wavelet filters.

We evaluate compression performance with PSNR and the number of nonzero transform coefficients (NONTC) after thresholding. Since we do not perform entropy coding, the bitrate is not used. The side information that would be needed to transmit the chosen transforms for each local region is also not taken into account. The effect of this side information is not likely to affect the main conclusion of this paper. Indeed, 2D-dir-DWTs require one out of 81 (= 9x9) transforms for each block, while 1D-dir-DWTs + 2D-sep-DWT require one out of 20 (= 2x9 + 1 + 1) transforms. The number of available choices is less, and the bitrate used for the side information is likely to be less. The 2D-sep-DWT, however, does not need transmission of side information. Typically, the increase in the bitrate due to the side information does not overshadow the gains achieved from using directional transforms [3, 4].

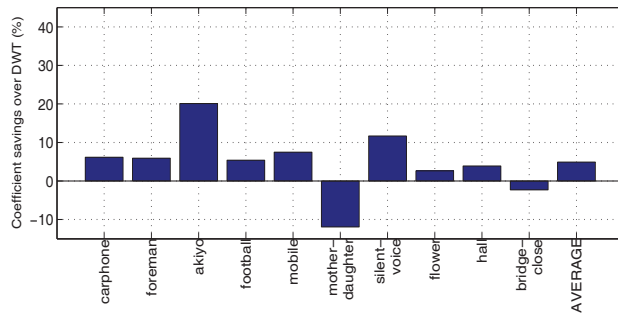
To summarize the comparison results at different compression ratios, we use a coefficient savings metric and the Bjontegaard-Delta PSNR (BD-PSNR) metric [10]. These metrics measure, respectively, the average percentage of coefficient savings and the average PSNR improvement between two PSNR-NONTC curves. Each curve is formed from four PSNR-NONTC data points with varying PSNR levels ranging from around 30dB to 45dB. For the comparisons, we take as one of the two PSNR-NONTC curves, the curve produced by the 2D-sep-DWT. The other curve is that produced by the transform-combination that is compared with the 2D-sep-DWT.

Figures 4 and 5 summarize the results of the experiments. Figure 4 shows the coefficient savings of 2D-dir-DWTs over the 2D-sep-DWT on MC-residuals and RE-residuals. Figure 5 shows the coefficient savings of our proposed 1D-dir-DWTs + 2D-sep-DWT over the 2D-sep-DWT on MC-residuals and RE-residuals.

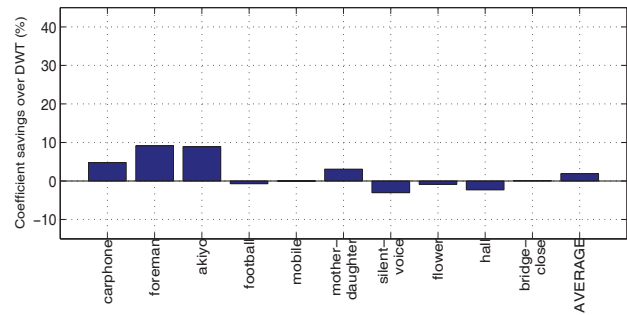
It can be observed from these Figures that 1D-dir-DWTs combined with the 2D-sep-DWT can perform much better at compressing prediction residuals than 2D-dir-DWTs. Specifically, in each of the 20 cases, our proposed method performed better than the 2D-dir-DWTs. On average, the 2D-dir-DWTs require 5% fewer coefficients than the 2D-sep-DWT for MC-residuals, and the 1D-dir-DWTs and 2D-sep-DWT combination requires 21% fewer coefficients than the 2D-sep-DWT. For RE-residuals, the savings are on average 2% if 2D-dir-DWTs are used, and 15% if the 1D-dir-DWTs and 2D-sep-DWT combination is used.

No plots are provided to show the comparison results with the BD-PSNR metric due to space limitations. The average improvements with this metric are as follows. While 2D-dir-DWTs provide on average 0.10dB PSNR improvement over the 2D-sep-DWT for MC-residuals, the 1D-dir-DWTs and 2D-sep-DWT combination provides on average 0.33dB PSNR improvement over the 2D-sep-DWT. For RE-residuals, the improvements are on average 0.05dB for the 2D-dir-DWTs, and 0.53dB for the 1D-dir-DWTs and 2D-sep-DWT combination.

These results clearly indicate that using one-dimensional directional wavelet transforms in addition to the two-dimensional separable wavelet transform can perform much better at compressing prediction residuals than using only two-dimensional directional wavelet transforms or the two-dimensional separable wavelet transform.

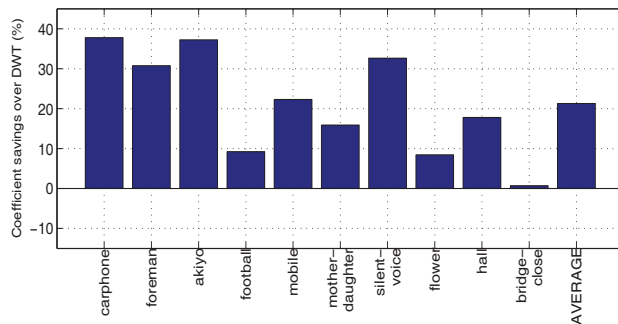


(a) MC-residual

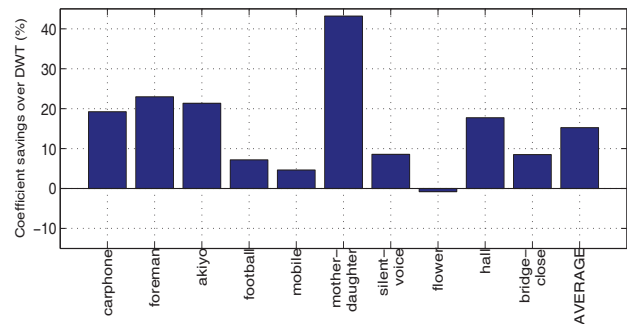


(b) RE-residual

**Fig. 4.** Coefficient savings of 2D-dir-DWTs over 2D-sep-DWT on MC- and RE-residuals. (1 residual-frame per sequence is used)



(a) MC-residual



(b) RE-residual

**Fig. 5.** Coefficient savings of 1D-dir-DWTs + 2D-sep-DWT over 2D-sep-DWT on MC- and RE-residuals. (1 residual-frame per sequence is used)

## 5. CONCLUSIONS

Typically, transforms used in image compression are also used to compress prediction residuals. However, prediction residuals may have different characteristics from images. Therefore it is useful to develop transforms that are adapted to prediction residuals. Various directional wavelet transforms have been developed recently to improve the compression of images. In this paper, we developed one-dimensional directional wavelet transforms specifically targeted for the compression of prediction residuals. To compare the compression performance of these transforms, we performed some preliminary experiments. The results indicate that using one-dimensional directional transforms in addition to two-dimensional separable transforms gives significantly better compression results for prediction residuals than using two-dimensional directional transforms or separable transforms.

## 6. REFERENCES

- [1] J. Ohm, M. v.d. Schaar, and J. W. Woods, "Interframe wavelet coding - motion picture representation for universal scalability," *EURASIP Signal Processing: Image Communication, Special Issue on Digital Cinema*, vol. 19, pp. 877–908, October 2004.
- [2] K.-C. Hui and W.-C. Siu, "Extended analysis of motion-compensated frame difference for block-based motion prediction error," *Image Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 1232–1245, May 2007.
- [3] F. Kamisli and J.S. Lim, "Transforms for the motion compensation residual," *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 789–792, April 2009.
- [4] F. Wu W. Ding and S. Li, "Lifting-based wavelet transform with directionally spatial prediction," *Picture Coding Symp.*, vol. 62, pp. 291–294, January 2004.
- [5] E. Le Pennec and S. Mallat, "Sparse geometric image representations with bandelets," *Image Processing, IEEE Transactions on*, vol. 14, no. 4, pp. 423–438, April 2005.
- [6] Bing Zeng and Jingjing Fu, "Directional discrete cosine transforms for image coding," *Multimedia and Expo, 2006 IEEE International Conference on*, pp. 721–724, 9-12 July 2006.
- [7] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P.L. Dragotti, "Directionlets: anisotropic multidirectional representation with separable filtering," *Image Processing, IEEE Transactions on*, vol. 15, no. 7, pp. 1916–1933, July 2006.
- [8] C.-L. Chang and B. Girod, "Direction-adaptive discrete wavelet transform for image compression," *Image Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 1289–1302, May 2007.
- [9] Wim Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM Journal on Mathematical Analysis*, vol. 29, no. 2, pp. 511–546, 1998.
- [10] G. Bjontegaard, "Calculation of average psnr differences between rd-curves," *VCEG Contribution VCEG-M33*, April 2001.