Comments on "Phase-Shifting for Nonseparable 2-D Haar Wavelets"

Yiannis Andreopoulos

University College London
Dept. of Electronic and Electrical Engineering
Torrington Place, WC1E 7JE, London, UK
Tel: +44-20-76797303; fax: +44-20-73889325; email: iandreop@ee.ucl.ac.uk

Abstract

In their recent paper, Alnasser and Foroosh derive a wavelet-domain (in-band) method for phase-shifting of 2D "nonseparable" Haar transform coefficients. Their approach is parametrical to the (*a-priori* known) image translation. In this correspondence, we show that the utilized transform is in fact the separable Haar discrete wavelet transform (DWT). As such, wavelet-domain phase shifting can be performed using previously-proposed phase-shifting approaches that utilize the overcomplete DWT (ODWT), if the given image translation is mapped to the phase component and in-band position within the ODWT.

Keywords: *overcomplete discrete wavelet transforms*, *image translation*, *phase-shifting*

I. Introduction

IN a recent paper [1], a wavelet-domain (in-band) solution for the 2D phase-shifting of an input image was derived. The method is formulated for the multilevel 2D "nonseparable" Haar discrete wavelet transform (DWT) and it is parametrical to the (known) image translation (shift) in the spatial domain. Experiments are proposed that demonstrate that the proposed in-band solution provides significantly higher quality against conventional interpolation approaches when the input image undergoes a series of subpixel shifts [1].

In this correspondence, we show that the utilized transform of [1] is in-fact the conventional separable Haar transform (Section II). We also correct some minor mistakes made in the transform formulation of [1]. This means that the in-band phase shifting results obtained by [1] can be obtained with previously-known theoretical and practical results proposed in [2]-[4], as discussed in Section III.

II. COMMENTS ON TRANSFORM USED BY ALNASSER AND FOROOSH

We follow the notations of [1] and assume that the input is a 2D signal (image) with $2^N \times 2^N$ samples. Equation (2) of [1], which shows the single-level synthesis that produces the low-frequency (blur) coefficients, can be written as:

$$\begin{aligned} A^l_{2i,2j} &= A^{l-1}_{i,j} + X^l_{i,j} \\ A^l_{2i,2j+1} &= A^{l-1}_{i,j} + Y^l_{i,j} \\ A^l_{2i+1,2j} &= A^{l-1}_{i,j} + Z^l_{i,j} \\ A^l_{2i+1,2j+1} &= A^{l-1}_{i,j} + W^l_{i,j} \end{aligned} \tag{1}$$

where $l \in \{1,\dots,N\}$ indicates the synthesis level[1] and $i,j \in \{0,\dots,2^{l-1}-1\}$ are the image coordinates at level $l-1$ (coarse resolution). Coefficients $A^l_{2i+\{0,1\},2j+\{0,1\}}$ are the blur coefficients of level l. They are calculated in (1) by adding the blur coefficient of level $l-1$ at position (i,j) to the "composite" coefficients $X^l_{i,j}$, $Y^l_{i,j}$, $Z^l_{i,j}$, and $W^l_{i,j}$. These composite coefficients are actually calculated based on the detail coefficients of level $l-1$ [1]:

$$\begin{aligned} \widetilde{X}^l_{i,j} &= a^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} + b^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} + c^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} \\ \widetilde{Y}^l_{i,j} &= -a^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} + b^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} - c^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} \\ \widetilde{Z}^l_{i,j} &= a^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} - b^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} - c^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} \\ \widetilde{W}^l_{i,j} &= -a^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} - b^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} + c^{l-1}_{\lfloor i/2\rfloor,\lfloor j/2\rfloor} \end{aligned} \tag{2}$$

where a, b, and c are the horizontal, vertical and diagonal detail coefficients, $\lfloor w \rfloor$ is the largest integer that is less or equal to w, and notation $\widetilde{Q}$ is used to specify that quantity Q is derived by equation (1) of [1]. By replacing (2) in (1) we derive the following transform synthesis:

[1] We follow the convention used in [1], where the finest-resolution level that corresponds to the image pixels is level $l = N$, while the coarsest-resolution level is $l = 0$, which contains only one blur coefficient and three detail coefficients.

$$\begin{bmatrix} \tilde{A}^{l}_{2i,2j} \\ \tilde{A}^{l}_{2i,2j+1} \\ \tilde{A}^{l}_{2i+1,2j} \\ \tilde{A}^{l}_{2i+1,2j+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} A^{l-1}_{i,j} \\ a^{l-1}_{\lfloor i/2 \rfloor, \lfloor j/2 \rfloor} \\ b^{l-1}_{\lfloor i/2 \rfloor, \lfloor j/2 \rfloor} \\ c^{l-1}_{\lfloor i/2 \rfloor, \lfloor j/2 \rfloor} \end{bmatrix}. \tag{3}$$

Since we have $i, j \in \{0, \ldots, 2^{l-1} - 1\}$ for (1)-(3), the transform synthesis of (3) is erroneous as it uses only half of the detail coefficients of level $l-1$ to reconstruct the blur coefficients of level l, thereby not allowing for perfect reconstruction. For example, for $N = 3$ we have $l \in \{0, 1, 2\}$; for level $l = 2$ (finest decomposition level) we have $i, j \in \{0, 1\}$, which means that, based on (3), all high frequency coefficients $h^1_{0,1}$, $h^1_{1,0}$, $h^1_{1,1}$, with $h \in \{a, b, c\}$, are not used for the reconstruction of the blur coefficients of level two.

The problem is actually created by (2), i.e. equation (1) of [1], which should be written in [1] as:

$$\begin{aligned} X^l_{i,j} &= a^{l-1}_{i,j} + b^{l-1}_{i,j} + c^{l-1}_{i,j} \\ Y^l_{i,j} &= -a^{l-1}_{i,j} + b^{l-1}_{i,j} - c^{l-1}_{i,j} \\ Z^l_{i,j} &= a^{l-1}_{i,j} - b^{l-1}_{i,j} - c^{l-1}_{i,j} \\ W^l_{i,j} &= -a^{l-1}_{i,j} - b^{l-1}_{i,j} + c^{l-1}_{i,j} \end{aligned}. \tag{4}$$

with $i, j \in \{0, \ldots, 2^{l-1} - 1\}$.

Then, by replacing (4) in (1) we derive the correct transform synthesis as:

$$\begin{bmatrix} A^{l}_{2i,2j} \\ A^{l}_{2i,2j+1} \\ A^{l}_{2i+1,2j} \\ A^{l}_{2i+1,2j+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} A^{l-1}_{i,j} \\ a^{l-1}_{i,j} \\ b^{l-1}_{i,j} \\ c^{l-1}_{i,j} \end{bmatrix}. \tag{5}$$

The corresponding transform analysis is derived by inverting the system of (5), i.e.:

$$\begin{bmatrix} A^{l-1}_{i,j} \\ a^{l-1}_{i,j} \\ b^{l-1}_{i,j} \\ c^{l-1}_{i,j} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} A^{l}_{2i,2j} \\ A^{l}_{2i,2j+1} \\ A^{l}_{2i+1,2j} \\ A^{l}_{2i+1,2j+1} \end{bmatrix}. \tag{6}$$

Under this modification, the proposed phase shifting results of the paper follow, e.g. (8) (10), (11), (14) of [1]. Hence, the results of [1] are valid for the transform given by (6) and they are correct for this case.

However, it is straightforward to demonstrate that (6) corresponds to the scaled *separable* Haar DWT analysis. The row-column (separable) decomposition with the Haar DWT is given by:

$$\begin{bmatrix} A_{i,j,\mathrm{RC}}^{l-1} & a_{i,j,,\mathrm{RC}}^{l-1} \\ b_{i,j,\mathrm{RC}}^{l-1} & c_{i,j,\mathrm{RC}}^{l-1} \end{bmatrix} = \mathbf{E} \begin{bmatrix} A_{2i,2j}^{l} & A_{2i,2j+1}^{l} \\ A_{2i+1,2j}^{l} & A_{2i+1,2j+1}^{l} \end{bmatrix} \mathbf{E}^{\mathrm{T}} . \tag{7}$$

with $i,j = 0,\ldots,2^{l-1}-1$ and the analysis matrix:

$$\mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Equation (7) expands to:

$$\begin{bmatrix} A_{i,j,\mathrm{RC}}^{l-1} & a_{i,j,\mathrm{RC}}^{l-1} \\ b_{i,j,\mathrm{RC}}^{l-1} & c_{i,j,\mathrm{RC}}^{l-1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} A_{2i,2j}^{l} + A_{2i,2j+1}^{l} + A_{2i+1,2j}^{l} + A_{2i+1,2j+1}^{l} & A_{2i,2j}^{l} - A_{2i,2j+1}^{l} + A_{2i+1,2j}^{l} - A_{2i+1,2j+1}^{l} \\ A_{2i,2j}^{l} + A_{2i,2j+1}^{l} - A_{2i+1,2j}^{l} - A_{2i+1,2j+1}^{l} & A_{2i,2j}^{l} - A_{2i,2j+1}^{l} - A_{2i+1,2j}^{l} + A_{2i+1,2j+1}^{l} \end{bmatrix} \tag{8}$$

i.e. the decomposition of (6) scaled by 2.

Finally, in order to agree with Figure 1 of [1] and with equation (2) of [1], the three parts of Figure 2 of [1] should be altered to include only the pairs of 2×2 coefficients used in the actual computation of the algorithm shown by equations (6) and (9) of [1] and by the computations thereafter. In addition, the superscript of the summation that appears in equations (6) and (9) of [1] should be $2^k(i+1)-1$ instead of $2^k(i+1)$.

III. LINK TO PRIOR WORK ON IN-BAND PHASE SHIFTING OF THE DWT

The results of [1] are expressed via complicated formulae for the direct 2D phase shifting of the (separable) Haar transform due to the choice of the authors to utilize the direct 2D kernel of the Haar DWT given by (5) instead of the separable transform of (7). Hence, they provide rather limited insight on how to extend this process to a variety of wavelet transforms existing in the literature. In addition, the contribution of the high frequency coefficients of the multilevel critically-sampled decomposition is not evident from the proposed formulation since the terms $D_{i,j}^{l}$ [which depend on the composite coefficients of (4)] appear in the proposed phase shifting results.

Previous works [2]-[4] have already derived in-band phase shifting approaches for separable DWTs. By extending the notations of [1], having an input 2D signal (image) $x(m,n)$ and its N-level DWT consisting of $A_{0,0}^{0}$ and $a_{i,j}^{l}$, $b_{i,j}^{l}$, $c_{i,j}^{l}$ (with $l \in \{0,\ldots,N-1\}$ and $i,j \in \{0,\ldots,2^{l}-1\}$), based on [2]-[4] we can derive all the phase components $A_{\{r,c\},0,0}^{0}$ and $a_{\{r,c\},i,j}^{l}$, $b_{\{r,c\},i,j}^{l}$, $c_{\{r,c\},i,j}^{l}$ (with

$r, c \in \{0, \ldots, 2^{N-l} - 1\}$) of the ODWT by applying the 1D wavelet-domain phase-shifting formulation horizontally and vertically. The relationship between an image shift in the spatial domain and the equivalent positions within the subbands of a certain phase of level l was given by [2, Prediction Rules] and [5, eq. (3)]. For any level l, this relationship links the image shift to the phase components of the ODWT and also with the in-band translation within each subband. Hence, for any spatial translation of the image, the ODWT phase components and the in-band translation within each subband is established, and then the separable in-band phase-shifting of [2]-[4] can be applied directly. This has already been used for practical applications involving 2D DWT decompositions, e.g. for in-band motion estimation and compensation [2] [4] [5].

REFERENCES

[1] M. Alnasser and H. Foroosh, “Phase-shifting for nonseparable 2-D Haar wavelets,” IEEE Trans. on Image Process., vol. 17, no. 7, pp. 1061-1068, July 2008.

[2] G. Van der Auwera, A. Munteanu, P. Schelkens and J. Cornelis, “Bottom-up motion compensated prediction in the wavelet domain for spatially-scalable video”, *IEE Electron. Letters*, vol. 38, no. 21, pp. 1251-1253, Oct. 2002.

[3] X. Li, “New results of phase shifting in the wavelet space,” *IEEE Signal Processing Lett.*, vol. 10, no. 7, pp. 193-195, July 2003.

[4] Y. Andreopoulos, A. Munteanu, G. Van der Auwera, J. Cornelis and P. Schelkens, "Complete-to-overcomplete discrete wavelet transforms: theory and applications," *IEEE Trans. on Signal Process.*, vol. 53, no. 4, pp. 1398-1412, April 2005.

[5] H. W. Park and H. S. Kim, “Motion estimation using low-band-shift method for wavelet-based moving-picture coding,” *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 577-587, April 2000.