WAVELET DOMAIN IMAGE RESOLUTION ENHANCEMENT USING CYCLE SPINNING AND EDGE MODELLING

Alptekin Temizel[1], *Theo Vlachos*[2]

[1]Honeywell Video Systems-Visioprime, 30 St. Johns Road, Woking, Surrey, GU21 7SA, UK, atemizel@visioprime.com
[2] Centre for Vision, Speech, and Signal Processing (CVSSP), University of Surrey, Guildford, GU2 7XH, UK, t.vlachos@surrey.ac.uk

ABSTRACT

In this paper we present a wavelet domain image resolution enhancement algorithm. An initial high-resolution approximation to the original image is obtained by means of zero-padding in the wavelet domain. This is further processed using the cycle-spinning methodology which reduces ringing. A critical element of the algorithm is the adoption of a simplified edge profile suitable for the description of edge degradations such as blurring due to loss of resolution. Linear regression using a minimal training set of high-resolution originals, is finally employed to rectify the degraded edges. Our results show that the proposed method outperforms conventional image interpolation approaches, both in objective and subjective terms, while it also compares favourably with state-of-the-art methods operating in the wavelet domain.

1. INTRODUCTION

Resolution enhancement of pictorial data is desirable in many applications such as monitoring, surveillance, medical imaging and remote sensing. It is a classic signal interpolation problem and conventional approaches such as zero-order interpolation (sample-and-hold) cause severe pixelation impairments while bilinear and spline interpolation invariably result in undesirable levels of smoothing across salient edges. Recently several efforts in the field have utilised wavelet-domain methodologies with the intention of overcoming some of the problems associated with conventional treatment. A common feature of these algorithms is the assumption that the low-resolution (LR) image to be enhanced is the low-pass filtered subband of a high-resolution (HR) image which has been subjected to a decimated wavelet transform. A trivial approach would be to reconstruct an approximation to the HR image by filling the unknown, so-called 'detail' subbands (normally containing high-pass spatial frequency information) with zeros followed by the application of the inverse wavelet transform (IWT). It is interesting to note that while this approach is capable of outperforming bilinear interpolation it has never appeared in the literature probably due to its simplicity. More sophisticated methods have attempted to estimate the unknown detail wavelet coefficients in an effort to improve the sharpness of the reconstructed images.

In [1] and [2] estimation was carried out by examining the evolution of wavelet transform extrema from finer to coarser subbands. Edges identified by an edge detection algorithm in lower frequency subbands were used to formulate a template for estimating edges in higher-frequency subbands. Only coefficients with significant magnitudes were estimated as the evolution of the wavelet coefficients among the scales was found to be difficult to model for other coefficients. Significant magnitude coefficients correspond to salient image discontinuities and consequently only the portrayal of those can be targeted with this approach while moderate activity detail escapes treatment. Furthermore, due to the fact that wavelet filters have support which spans a number of neighbouring coefficients, edge reconstruction is inevitably based on contributions from such neighbourhoods. As methods based on extrema evolution only target locations of coefficients with significant magnitudes, such neighbourhoods will inevitably provide incomplete information ultimately affecting the quality of edge reconstruction. Performance is also affected by the fact that the signs of estimated coefficients are replicated directly from 'parent' coefficients (in a quad-tree hierarchical decomposition sense) without any attempt being made to estimate the actual signs. This is contradictory to the commonly accepted fact that there is very low correlation between the signs of parent coefficients and their descendants. In a coding context for example, the signs of descendants were generally assumed to be random [3], [4]. As a result, the signs of the coefficients estimated using extrema evolution techniques cannot be relied upon.

In [5] a technique was proposed which takes into account the Hidden Markov Tree (HMT) approach of [6]. The latter was successfully applied to a different class of problems including image denoising and related applications. An extended version of this approach utilising super-resolution type of methodologies is presented in [7]. These methods model the unknown wavelet coefficients as belonging to mixed Gaussian distributions (states) which are symmetrical around the zero mean. HMT models are used to find out the most probable state for the coefficient to be estimated (i.e. to which distribution it belongs to). The posterior state is found using state-transition information from lower-resolution scales and the coefficient estimates are randomly generated using this distribution. Being symmetrical around zero, the probability of estimation of a coefficient with a negative sign is equal to that with a positive sign. Consequently sign changes between the scales are not taken into account and randomly generated signs are assigned to the estimated coefficients. Finally the HMT based method has been further developed so that it does not require any training data set [8].

In [9] and [10] a wavelet-based superresolution method was presented based on the Multiresolutional Basis Fitting Reconstruction (MBFR) technique in [11]. The algorithm exploits the interlaced sampling structure in the LR data in the existence of multiple LR images. Finally, a similar approach was proposed in [12] on the basis of the availability of a single LR image. The basis of this approach, MBFR technique, was designed to take advantage of the non-uniform sampling of a signal using sections with higher sampling rates to interpolate higher frequencies locally. However availability of only a single LR image, with implication that the sampling is uniform, prohibits taking full-advantage of this scheme.

A different approach to improve the perceptual quality of wavelet compressed images has been proposed in [13]. In this work, edges are represented using the edge model devised in [14]. The edges are then reconstructed by changing the edge parameters and making them more similar to the ideal edge model. As the inherited edge model is an artificial one, the algorithm results in non-natural looking edges. It should also be noted that this scheme targets wavelet compressed images and hasn't been applied to image resolution enhancement problems.

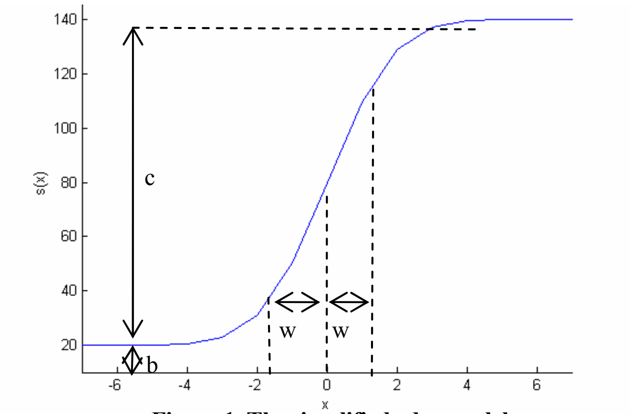

Figure 1. The simplified edge model

Recently it has been shown that the cycle-spinning methodology produces notable results when adapted to wavelet domain resolution enhancement problems [15]. In this paper, we adopt the standard degradation model where an LR image is obtained from an HR image as the low spatial frequency subband of a decimated discrete wavelet transform and show that the results obtained in [15] could be further improved upon by least-squares estimation of the parameters of a generic edge profile.

2. LOW-RESOLUTION IMAGE GENERATION AND EDGE MODELLING

As already stated it is assumed that the LR image to be enlarged is the LL (low-pass) subband of a quad-tree wavelet decomposition of the unknown HR image. This LL subband is the result of variable-separable (first horizontal, then vertical) low-pass wavelet filtering. The unavailability of high-frequency spatial information normally residing in the unknown subbands results invariably in blurring and ringing of salient image features such as sharp edges. Nevertheless, low-pass filtered versions of edges actually survive within the LL subband. While these are typically smoothed and widened versions of the original edges it is worth noting that, for simple edge profiles, their notional centre of symmetry (i.e. the midpoint between the minimum and the maximum intensity value) has not moved from its original location.

To describe edge evolution as it undergoes low-pass filtering we adopt the model proposed in [14]. According to this model image edges can be approximated by Gaussian-smoothed step functions. In one dimension (i.e. along an image scan line), if $h(x;b,c)$ is a step function and $g(x,w)$ is a Gaussian, then an edge $s(x)$ is approximated by:

$$s(x) \equiv s(x;b,c,w) = h(x;b,c)*g(x,w) \tag{1}$$

where c is the edge contrast the, b is the edge minimum and * denotes convolution. Parameter w describes the width of the edge (i.e. distance from the midpoint) and is implicitly related to the variance of the Gaussian smoothing function. The above profile is illustrated in Figure 1 for edge parameters b=20, c=120 and w=1.5.

Using the above model, smoothing due to low-pass filtering manifests itself as an increase of parameter w, while the other parameters remain relatively unaffected (assuming for simplicity wavelet filters of unity gain). If $s_0(x) = s(x;b_0,c_0,w_0)$ is an edge in the unavailable HR image, the corresponding surviving edge in the available image would be represented by $s_1(x) = s(x;b_0,c_0,w_1) + q_n(x)$ where $w_1 = \lambda w_0$ with an edge widening factor of λ and $q_n(x)$ is a term accounting for other residual degradations such as ringing. As explained in the next section, a critical component of our method is the estimation of edge width w. We account for edge distortion (primarily smoothing) by establishing a correspondence between available LR image data and a training set of HR image data using linear regression.

Image/Method	*Lena*	*Elaine*	*Baboon*	*Peppers*
Original	1.3086 ±0.3863	1.5197 ±0.2843	1.0206 ±0.2924	1.4627 ±0.3110
Bilinear	1.4672 ±0.4085	1.7326 ±0.3477	1.2373 ±0.3627	1.6970 ±0.3422
WZP (Db.9/7)	1.3635 ±0.402	1.6284 ±0.3217	1.1157 ±0.3355	1.5743 ±0.3302
WZP and CS	1.3867 ±0.3979	1.6683 ±0.3158	1.1373 ±0.3301	1.6174 ±0.3316
WZP,CS and ER	1.3325 ±0.3774	1.6058 ±0.3037	1.1564 ±0.3250	1.5514 ±0.3170

Table 1. Calculation of horizontal edge width parameter *w* for a variety of images and reconstruction techniques.

3. RESOLUTION ENHANCEMENT

The proposed algorithm consists of three steps. In the first step, an initial HR approximation is generated using wavelet domain zero padding. This approximation commonly exhibits artefacts such as smoothing and ringing. To reduce the ringing artefacts, a variant of the cycle spinning methodology is applied as a second step. Finally, the edges are rectified by re-adjusting their widths according to estimates obtained by processing LR image data. These steps are explained in more detail below.

3.1 Wavelet Domain Zero Padding (WZP)

An initial approximation to the unknown HR image is generated using wavelet-domain zero padding (WZP). Using a given LR image $\mathbf{x}$ of size m x n, the unknown HR image $\mathbf{y}$ is reconstructed by using zero padding of high-frequency subbands (i.e. setting all elements of these subbands to zeros) followed by inverse wavelet transform: $\hat{\mathbf{y}}_0 = W^{-1}\begin{bmatrix} \mathbf{x} & \mathbf{0}_{\mathbf{m,n}} \\ \mathbf{0}_{\mathbf{m,n}} & \mathbf{0}_{\mathbf{m,n}} \end{bmatrix}$ where $\mathbf{0}_{\mathbf{m,n}}$ is an all-zero submatrix of dimensions m x n and W^{-1} is the inverse discrete wavelet transform.

3.2 Cycle Spinning (CS)

The decimated wavelet transform is not shift-invariant and as a result, distortion of wavelet coefficients, due to quantisation of coefficients in compression applications or non-exact estimation of high-frequency coefficients in resolution enhancement applications (including zero padding of coefficients as in WZP), introduces cyclostationarity into the image which manifests itself as ringing in the neighbourhood of discontinuities. Cycle-spinning (CS) has been shown to be an effective method against ringing when used for denoising purposes in the wavelet domain [16] and also for reducing ringing and increasing the perceptual quality of compressed images. CS method aims to approximate shift-invariant statistics by averaging out the cyclostationarity. In [17] and [18], it was shown that CS applied as a post-processing operation after decompression results in significant improvements in the framework of JPEG and JPEG2000 image compression.

First a number of LR images $\hat{\mathbf{x}}_{i,j}$ are generated from $\hat{\mathbf{y}}_0$ by (i) spatial shifting, (ii) wavelet transforming and (iii) discarding the high frequency (HF) coefficients: $\hat{\mathbf{x}}_{i,j} = DWS_{i,j}\hat{\mathbf{y}}_0$ where D represents discarding of HF coefficients, W denotes wavelet transform and $S_{i,j}$ is a shift operator applying horizontal and vertical shifts of (i,j) for $i, j \in \{-k, -k+1, ..., k-1, k\}$. Second, WZP is applied to all $\hat{\mathbf{x}}_{i,j}$ yielding N $\hat{\mathbf{y}}_{i,j}$ images, where N=($2k$+1)($2k$+1). Finally, these intermediate HR images are re-aligned and averaged to give the

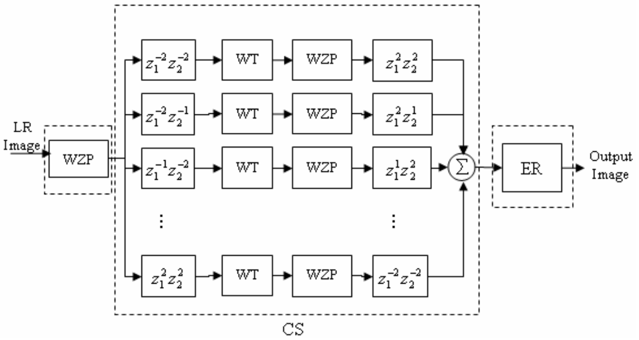

Figure 2. Simplified block diagram of the proposed method

final HR reconstructed image: $\hat{\mathbf{y}} = \frac{1}{N}\sum_{i=-k}^{k}\sum_{j=-k}^{k} S_{i,j}^{-1}\hat{\mathbf{y}}_{i,j}$ where $S_{i,j}^{-1}$ is the inverse of the shifting operator $S_{i,j}$.

3.3 Edge Rectification (ER)

As already stated the loss of high-frequency information results in larger values for parameter w –implying widening of edges- in reconstructed HR images compared with the original images, while parameters b and c remain relatively unaffected. Furthermore, application of cycle-spinning widens the edges further. Table 1 shows the mean value and standard deviation of w for a range of images obtained using various techniques. In this table Db.9/7 denotes the well-known Daubechies 9/7 discrete wavelet filterbank which was used throughout our experiments.

It would be reasonable to assume that a method readjusting (i.e. reducing) edge width while maintaining the edge profile otherwise would improve the quality of the reconstructed image. The method proposed in [13] attempts to achieve this by making edges more similar to the ideal edge model. However, since an artificial edge model is used, the algorithm results in non-natural looking edges. We propose an alternative method which estimates edge parameters directly from a training set of HR image data and further refines them using local neighbourhood information.

First we detect the edges in $\mathbf{y}$, the HR image used as a ground truth training set, using a Canny Edge Detector. Then we calculate the edge parameters using (1) and cluster the edges. Edges are clustered into 9 groups according to their w and c parameters. These groups are equally separated in between maximum and minimum values of w and c as illustrated in Figure 3. The following algorithm is applied to each cluster independently.

The edges in $\hat{\mathbf{y}}$, the output of the previous step, are degraded versions of the corresponding edges in $\mathbf{y}$. Linear regression is employed to estimate pixel values in the neighbourhood of the centre of symmetry of an edge. In particular we express edge pixel values in $\mathbf{y}$ as a weighted linear mix of neighbouring pixels in $\hat{\mathbf{y}}$. By denoting the coefficient at position (m,n) in $\mathbf{y}$ as $y_{m,n}$ and the coefficient at position (m,n) in $\hat{\mathbf{y}}$ as $\hat{y}_{m,n}$, the estimates of $y_{m-1,n}$, $y_{m,n}$ and $y_{m+1,n}$ are obtained according to the following:

$$y_{m-1,n} = a_0 + a_1\hat{y}_{m-2,n} + a_2\hat{y}_{m-1,n} + a_3\hat{y}_{m,n} \quad (2)$$

$$y_{m,n} = a_4 + a_5\hat{y}_{m-1,n} + a_6\hat{y}_{m,n} + a_7\hat{y}_{m+1,n} \quad (3)$$

$$y_{m+1,n} = a_8 + a_9\hat{y}_{m,n} + a_{10}\hat{y}_{m+1,n} + a_{11}\hat{y}_{m+2,n} \quad (4)$$

Using (2), (3) and (4) for all detected edge points falling into the current cluster, three over-determined equations could be constructed. These equations are solved by linear least squares regression to find the estimator weights sets $\{a_0, a_1, a_2, a_3\}$, $\{a_4, a_5, a_6, a_7\}$ and $\{a_8, a_9, a_{10}, a_{11}\}$ minimising the error. These estimated parameters are then used to rectify the edges in $\hat{\mathbf{y}}$. By repeating this process for each cluster, separate estimator sets for each cluster are obtained.

It is interesting to note that estimated parameters obtained from a very small training set, even a single image, work well for a wide

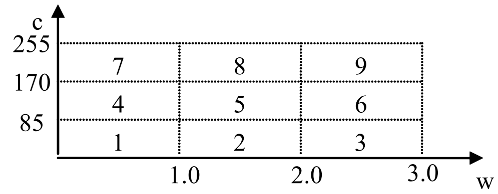

Figure 3. Edge cluster map

variety of natural images. This edge rectification process is subsequently referred to as ER. The algorithm works in a variable-separable way, first horizontally then vertically. As horizontal and vertical edges in natural images have the potential of different degrees of sharpness, for example due to non-isotropic sensors, it was found beneficial to calculate different estimator weights for each direction. A simplified block diagram for the overall algorithm is shown in Figure 2 where the notation for spatial shifting is in the z-domain.

Further levels of enlargement, such as a factor of 4, can be achieved by first generating the 4x enlarged image directly with WZP method and then applying the cycle-spinning method followed by enhancement of edges on this resulting image. Iterative application of the algorithm is also possible but it was observed that this method generates inferior quality images as well as being computationally more expensive.

4. EXPERIMENTAL RESULTS

We have experimented with a number of well-known test images including *Lena, Elaine, Baboon* and *Peppers*. As already stated a HR version of these (512x512) was used as ground truth for performance evaluation purposes. The training phase for linear regression is repeated using all the images in the data set with one image left out for evaluation.

The error between the ground truth high-resolution originals on the one hand and reconstructions on the other is expressed in terms of PSNR values and is tabulated in Tables 2 and 3 for 2x and 4x enlargement respectively. Objective comparisons are carried out with conventional bilinear interpolation and established wavelet based methods [2,5,7]. A non-wavelet scheme based on edge-directed interpolation [19] was also considered to provide a comparison with a state-of-the-art method not operating in the wavelet domain. Our results show that the proposed algorithm yields modest but consistent improvements over all competing methods something which was confirmed by visual inspection of the resolution enhanced images.

Figure 4 shows subjective comparisons with bilinear interpolation for *Lena*. Overall our results confirm that the portrayal of salient image features such as edges and contours is consistently improved while no perceivable artefacts are introduced.

Our experiments have also revealed that the choice of parameter k (the maximum shift to be applied in the cycle-spinning) has a significant effect on the algorithm efficiency and choosing $k = 4$ results in the best performance. Using a smaller neighbourhood suffers from not having sufficient data while a larger neighbourhood is liable to crosstalk from spatially uncorrelated image features.

5. CONCLUSIONS

An image resolution enhancement algorithm operating in the wavelet domain was presented. The main elements of this algorithm were zero-padding of high-frequency wavelet subbands, cycle spinning to reduce ringing arising from zero-padding and finally edge rectification to alleviate blurring due to the unavailability of high spatial

Image/Method	*Lena*	*Elaine*	*Baboon*	*Peppers*
Bilinear	30.13	30.60	22.85	30.01
Bicubic	31.34	31.17	22.98	30.28
NEDI[19]	34.10	32.89	23.87	33.54
WZP (Haar)	31.46	31.71	23.61	31.45
WZP (Db.9/7)	34.45	33.26	24.22	33.94
Carey et al.[2]	34.48	33.29	24.24	34.03
HMM [5]	34.52	33.31	24.24	34.04
HMM SR [7]	34.61	33.40	24.31	34.10
WZP and CS [15]	34.93	33.56	24.28	34.32
WZP, CS and ER	35.25	33.71	24.44	34.65

Table 2. PSNR(dB) results for 2x enlarged images (From 256x256 to 512x512)

Image/Method	*Lena*	*Elaine*	*Baboon*	*Peppers*
Bilinear	24.06	25.38	20.43	24.37
Bicubic	26.76	28.93	21.02	26.86
NEDI[19]	28.81	29.97	21.18	28.52
WZP (Haar)	26.67	28.06	21.11	26.89
WZP (Db. 9/7)	28.84	30.44	21.47	29.57
Carey et al.[2]	28.81	30.42	21.47	29.57
HMM [5]	28.86	30.46	21.47	29.58
HMM SR [7]	28.88	30.51	21.49	29.60
WZP and CS [15]	29.27	30.78	21.54	29.87
WZP, CS and ER	29.36	30.89	21.56	30.05

Table 3 – PSNR(dB) results for 4x enlarged images (From 128x128 to 512x512)

Figure 4. Extract from original *Lena* image (top), 4x reconstructions using bilinear interpolation (mid. left) and the proposed method (mid. right) and corresponding amplified error images (bottom).

frequency information. The proposed algorithm was based on the adoption of a simplified profile model and the parameterisation and rectification of blurring using linear regression. Our results have shown that the proposed method outperforms conventional image interpolation approaches, both in objective and subjective terms, while it also compares favourably with state-of-the-art methods operating in the wavelet domain.

REFERENCES

[1] S.G Chang, Z. Cvetkovic and M. Vetterli, “Resolution enhancement of images using wavelet transform extrema extrapolation”, *Proc.ICASSP‘95*, vol.4,pp.2379-2382, May 1995.

[2] W.K. Carey, D.B. Chuang and S.S. Hemami, “Regularity-Preserving Image Interpolation”, *IEEE Trans. Image Proc.*, vol.8, no.9, pp.1295- 1297, Sep. 1999.

[3] J.M. Shapiro, Embedded Image Coding Using Zerotrees of Wavelet Coefficients, *IEEE Trans. Signal Proc.,* vol.41, no.12, pp. 3445-3462, Dec. 1993.

[4] A. Said, W.A. Pearlman, A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees, *IEEE Trans. Circ. & Syst*., vol.6, pp.243-250, June 1996.

[5] K. Kinebuchi, D.D. Muresan and T.W. Parks, “Image Interpolation Using Wavelet-Based Hidden Markov Trees”, *Proc. ICASSP ‘01*, vol. 3, pp. 7-11, May 2001.

[6] M.S. Crouse, R.D. Nowak and R.G. Baraniuk, “Wavelet-Based Statistical Signal Processing Using Hidden Markov Models”, *IEEE Trans. Signal Proc*., vol.46,no.4, pp.886–902, Apr. 1998.

[7] S. Zhao, H. Han and S. Peng, “Wavelet Domain HMT-Based Image Superresolution”, *IEEE International Conference on Image Proc.*, vol. 2, pp. 933-936, Sep. 2003.

[8] D.H. Woo, I.K. Eom and Y.S. Kim, “Image Interpolation based on inter-scale dependency in wavelet domain”, *Proc. ICIP ‘04.,* Oct. 2004.

[9] N. Nguyen, “Numerical Techniques for Image Superresolution”, *Ph.D. dissert.*, Stanford Uni., Stanford, CA, Apr. 2000.

[10] N. Nguyen, P. Milanfar, “An efficient wavelet-based algorithm for image superresolution”, *Proc. ICIP ‘00*, vol.2, pp. 351-354, Sep. 2000.

[11] C. Ford and D.M. Etter, “Wavelet Basis Reconstruction of Nonuniformly Sampled Data”, *IEEE Trans. Circ. & Syst*., vol.45, no.8, pp.1165–1168, Aug. 1998.

[12] S. Mitevski and M. Bogdanov, “Application of Multiresolutional Basis Fitting Reconstruction in Image Magnifying”, *Proc. 9th Telecomnications Forum*, pp. 565-568, Nov. 2001.

[13] G. Fan and W. Cham, “Model-Based Edge Reconstruction for Low Bit-Rate Wavelet-Compressed Images”, *IEEE Trans. Circ.&Syst. for Vid.Tech.*, vol. 10, no.1, pp.120-132,Feb. 2000.

[14] P.J. Beek, P. J. L. Van, “Edge-Based Image Representation and Coding”, *Ph.D. Thesis*, Delft Uni. of Technology, 1995

[15] A. Temizel and T. Vlachos, “Wavelet Domain Image Resolution Enhancement Using Cycle-Spinning”, *IEE Electronics Letters*, vol. 41, no. 3, Feb. 2005

[16] R.R. Coifman and D.L. Donoho, “Translation-Invariant Denoising”, *Lecture Notes in Statistics*, pp. 125–150, Springer Verlag, 1995.

[17] A. Nosratinia, “Denoising of JPEG Images by Re-application of JPEG”, *Journal of VLSI Signal Processing*, vol. 27, no. 1, pp. 69–79, 2001.

[18] A. Nosratinia, “Post-Processing of JPEG-2000 Images to Remove Compression Artifacts”, *IEEE Signal Proc. Letters*, Vol. 10, No. 10, pp. 296-299, Oct. 2003.

[19] X. Li and M.T. Orchard, “New edge-directed interpolation”, *IEEE Trans. Image Proc.*, Vol. 10, No. 10, pp. 1521-1527, Oct. 2001.