Wavelet domain image resolution enhancement

A. Temizel and T. Vlachos

Abstract: A wavelet-domain image resolution enhancement algorithm which is based on the estimation of detail wavelet coefficients at high resolution scales is proposed. The method exploits wavelet coefficient correlation in a local neighbourhood sense and employs linear least-squares regression to estimate the unknown detail coefficients. Results show that the proposed method is considerably superior to conventional image interpolation techniques, both in objective and subjective terms, while it also compares favourably with competing methods operating in the wavelet domain.

1 Introduction

Wavelet-based algorithms have proved successful for image resolution enhancement tasks and several methods have appeared recently in the literature. A common feature of these is the assumption that the image to be enhanced is the low-pass filtered subband of a wavelet-transformed high-resolution image. This type of approach requires the estimation of detail wavelet coefficients in subbands containing high-pass spatial frequency information.

In [1], detail coefficients are estimated using the evolution of wavelet transform extrema in coarser subbands. A similar but less computationally expensive approach is advocated in [2]. In [3] a technique is proposed which takes into account the hidden Markov tree (HMT) approach of [4]. The latter was successfully applied to a different class of problems namely image denoising and related applications. An extended version of this approach utilising superresolution-like methods is presented in [5]. In [6] and [7] a wavelet-based superresolution method is presented based on the multiresolutional basis fitting reconstruction (MBFR) technique in [8]. Finally, a similar approach is proposed in [9] using a single low-resolution image.

The methods described in [1] and [2] employ wavelet transform extrema evolution. Only coefficients with significant magnitudes are estimated as the evolution of the wavelet coefficients among the scales is found to be difficult to model for other coefficients. Significant magnitude coefficients correspond to image discontinuities and consequently only the portrayal of edges in the enhanced resolution image can be targeted while smoother regions cannot be accommodated with such an approach. Furthermore, owing 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 from the fact that the signs of estimated coefficients are copied 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 fact that, in the literature, it is commonly accepted that there is very little correlation between the signs of the parent coefficients and their descendants. In coding applications for example, the signs of descendants are generally assumed to be random [10, 11]. As a result, the signs of the coefficients estimated using extrema evolution techniques cannot be relied upon.

HMT-based methods in [3] and [5] 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. It should also be noted that none of the aforementioned methods in addition to the techniques in [6, 7, 9] utilise existing correlations between neighbouring coefficients which is an important element of the method we are proposing.

In this paper, we present a method for estimating higher resolution wavelet coefficient values by exploiting the coefficient correlation in a local neighbourhood. In Section 2 we present an overview of the problem and we also discuss the importance of correct estimation of coefficients signs. In Section 3 we present our resolution enhancement method including a discussion of its underlying signal processing elements. Section 4 provides experimental results for various test images. Finally concluding remarks are made in Section 5.

2 Background

As already discussed wavelet-domain resolution enhancement is commonly achieved by regarding the available image (i.e. the image whose resolution we seek to

IEE Proceedings online no. 20045056

doi:10.1049/ip-vis:20045056

Paper received 21st June 2004

A. Temizel is with Visioprime, 30 St. Johns Road, Woking, GU21 7SA, UK

T. Vlachos is with the Centre for Vision, Speech, and Signal Processing (CVSSP), University of Surrey, Guildford, GU2 7XH, UK

E-mail: atemizel@visioprime.com

enhance) as the low-resolution (LR) version (i.e. a low-pass wavelet-filtered image both horizontally and vertically) of the unknown high-resolution (HR) image. Then by zero-filling the unknown detail subbands (i.e. using zeros in place of the unknown high-order wavelet coefficients) and combining these with the LR image, an enhanced resolution image can be obtained by taking the inverse wavelet transform. This method is illustrated in Fig. 1 where X represents the HR image and LL_0 is the available LR image. It should be noted that, in this method, the interpolative reconstruction is achieved by the synthesis wavelet filter pairs and, as a consequence, the selection of a mother wavelet which better models the regularity of natural images yields better results. For example, the well-established in the wavelet literature Daubechies 9/7 filters are expected to generate better results than Haar wavelets. Alternatively it should also be noted that increasing the wavelet support too much results in over-smoothness. Our experiments suggest that Daubechies 9/7 filters achieve a good balance for natural images. Despite its simplicity, the zero-filling method works reasonably well and produces good quality images. However, if reliable estimation of the unknown coefficients could be carried out, high spatial frequency image features could be further enhanced.

The wavelet coefficients of natural images have two important properties: persistency and non-Gaussianity. Persistency refers to the observation that the magnitudes of wavelet coefficients corresponding to the same spatial location tend to propagate from lower resolution scales through to higher resolution scales (Fig. 2). This property is exploited in state-of-the-art wavelet-based image compression algorithms such as Shapiro's embedded zerotree wavelet (EZW) [10] and Said and Pearlman's set partitioning in hierarchical trees (SPIHT) [11]. Additionally, works reported in the wavelet-domain image resolution enhancement literature also exploit the persistency property. However, the persistency property does not apply to coefficient signs. While these are generally assumed random, their contribution to reconstructed image quality is significant as demonstrated in Fig. 3. The image shown was generated by (i) taking a 4-band (LL, LH, HL and HH) wavelet transform of the Lena' image, (ii) inverting the signs of randomly selected 50% of coefficients in LH, HL and HH subbands and (iii) taking the inverse wavelet transform.

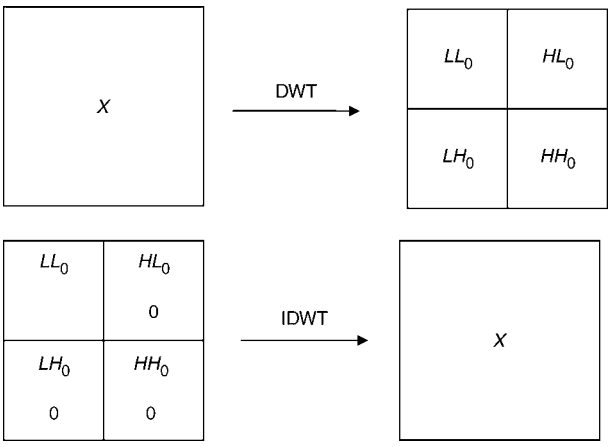

Fig. 1 *Conventional wavelet domain interpolation method*

The available image is assumed to be the low-resolution subband of the target high resolution image. The unknown subband elements (detail coefficients) are filled with 0s and resolution enhancement is achieved by inverse wavelet transform

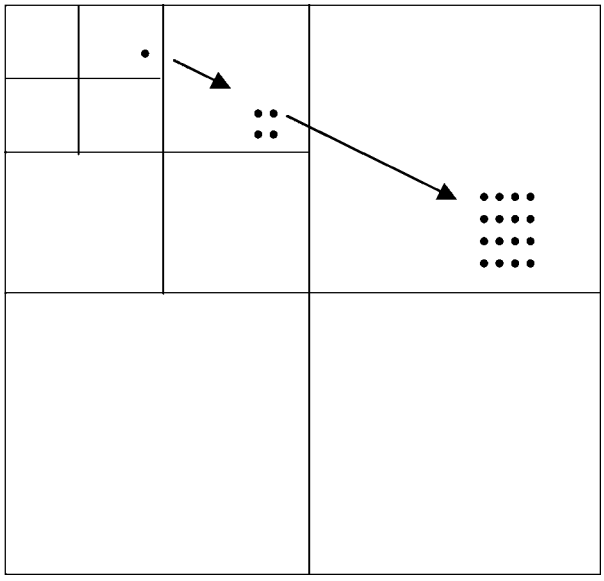

Fig. 2 *Persistency of wavelet coefficient magnitudes through different scales of wavelet decomposition*

The coefficient magnitudes at corresponding spatial locations decay at a similar rate from lower resolution scales to higher resolution scales

In the following Section, we show that the estimation results can be improved by exploiting correlation among neighbouring coefficients in addition to persistency.

3 Wavelet-domain resolution enhancement

Let $L(z)$ and $H(z)$ represent respectively the low-pass (LP) and high-pass (HP) filters constituting an analysis/synthesis filter pair for the discrete wavelet transform in the z-domain and let $X(z)$ be the unknown high-resolution image we seek to reconstruct. We use a notation in which the direction of filtering (i.e. row/column) is shown explicitly as a subscript while $(\downarrow 2)$ denotes decimation by a factor of 2 in both

Fig. 3 *'Lena' image reconstructed using original wavelet coefficient magnitudes while 50% of randomly selected coefficients with inverted signs (PSNR 27.71 dB)*

directions. Using this notation we can write that

$$LL_0(z) = (\downarrow 2)L_{col}(z)L_{row}(z)X(z) \quad (1)$$

and

$$HL_0(z) = (\downarrow 2)L_{col}(z)H_{row}(z)X(z) \quad (2)$$

where LL_0 and HL_0 are respectively low-pass and vertical detail subbands of wavelet coefficients. We assume that LL_0 is the available LR image whose resolution we seek to enhance and consequently our task is to estimate the elements of detail coefficients such as those contained in subband HL_0. Our approach is based on the assumption that a high-pass filtered undecimated version of the available LR image

$$HL_0'(z) = H_{row}(z)LL_0(z) \quad (3)$$

is sufficiently correlated with HL_0 horizontally to provide a basis for the estimation of the latter.

This assumption is tested by applying undecimated wavelet filtering on the available image (LL_0) to obtain low- and high-pass filtered versions LL_1 and HL_1

$$LL_1(z) = L_{col}(z)L_{row}(z)LL_0(z) \quad (4)$$

and

$$HL_1(z) = L_{col}(z)H_{row}(z)LL_0(z) \quad (5)$$

Subsequently, in accordance with (3), we obtain

$$HL_1'(z) = H_{row}(z)LL_1(z) \quad (6)$$

and demonstrate that substantial levels of horizontal correlation exist between HL_1 and HL_1'. Measured correlation values between HL_1 and HL_1' are given in Table 1. This Table shows the correlation between a coefficient located at (m, n) in HL_1 and coefficients in the neighbourhood of (m, n) in HL_1'. The highest correlation values are observed at locations $(m-1, n)$, (m, n), $(m+1, n)$ and $(m+2, n)$. These locations define the horizontal neighbourhood that will be used in the remainder of this work. Calculations on different test images confirm that the highest correlation values occur at the same locations.

The above observations have motivated an estimation approach which is based on linear least-squares regression. By denoting the coefficient at position (m, n) in HL_1 as $y_{m,n}$ and the coefficient at position (m, n) in HL_1' as $x_{m,n}$ an estimate of $y_{m,n}$ is obtained according to the following

$$y_{m,n} = a_0 + a_1x_{m-1,n} + a_2x_{m,n} + a_3x_{m+1,n} + a_4x_{m+2,n}. \quad (7)$$

Applying (7) to all coefficients of HL_1 in the specified neighbourhood, we can solve the over-determined problem by linear least-squares regression to obtain the weights a_0, a_1, a_2, a_3 and a_4 that minimise the approximation error. These weights are subsequently used to obtain an estimate of the unknown HL_0 coefficients as a linear combination of neighbouring HL_0' coefficients obtained by (3).

Table 1: Correlation values between HL_1 and HL_1' subbands for 'Lena'

	$x-2$	$x-1$	x	$x+1$	$x+2$	$x+3$
$y-2$	0.002	−0.045	0.045	0.007	−0.011	0.015
$y-1$	0.053	−0.148	0.034	0.181	−0.161	0.008
y	0.092	−0.327	0.206	0.208	−0.324	0.122
$y+1$	0.015	−0.165	0.185	0.028	−0.137	0.064
$y+2$	0.034	−0.005	0.018	0.034	−0.041	0.006

A similar treatment is possible for the LH_0 subband. In this case, as expected, the highest correlation values are found in the vertical neighbourhood and the same weights as for HL_0 could be used. Estimation of HH_0 coefficients was omitted as its contribution in subjective terms was found to be negligible. This is not inconsistent with human visual perception since the HH_0 subband is generated by high-pass filtering the image horizontally and vertically. Such filtering promotes the retention of the spatially diagonal image activity and it is widely accepted that the human visual system is relatively less sensitive to diagonal detail. This is an established consideration in image coding systems which invariably tolerate higher levels of quality loss in diagonal detail.

After coefficient estimation is carried out as above, the high-resolution image is obtained by applying the inverse wavelet transform. Our experiments show that using a training set consisting of a single image to drive the estimation process yields weights which are reliable and stable over a wide range of natural images for both 2× and 4× factors of magnification.

The baseline algorithm described above generates HR images which are 2× the original LR image size. Further factors of enlargement can be achieved by applying the same algorithm iteratively exploiting the hierarchy of the quad-tree wavelet decomposition.

3.1 Discussion

It could be argued that the operation described by (3) is self-contradictory i.e. it amounts to high-pass filtering image data which are essentially low-pass frequency in nature. Nevertheless it should be taken into account that all filters involved are non-ideal (i.e. brickwall) and consequently there will be frequency components that will survive this operation. Such components correspond to the aliased part of the image spectrum. In that sense our method seeks to exploit correlation between aliased frequency components in low-frequency subband images on the one hand and frequency components which are unknown and reside in high-frequency subband images i.e. the wavelet coefficients which we endeavour to estimate. Our measurements have established that such a degree of correlation exists. A similar case can be argued for the operation described by (6).

4 Experimental results

In our experiments we have used four well-known test images: 'Lena', 'Elaine', 'Baboon' and 'Peppers'. We have filtered and downsampled the original images by factors of 2 and 4 respectively for 2× and 4× resolution enhancement and used those as the available LR images. We have used the original images as ground truth and we have expressed the quality of the resolution enhancement process in terms of the well-established 'in image processing' peak signal-to-noise ratio (PSNR) metric.

Experimental results obtained by simulating the above algorithm are shown in Tables 2 and 3 where PSNR performance is tabulated for 2× and 4× enlargement factors respectively. In both Tables we compare the standard bilinear interpolation approach with a number of alternative approaches in the wavelet domain. In those Tables 'Wavelet' refers to a basic reconstruction method (using two well-known classes of analysis-synthesis filters namely Haar and Daubechies 9/7) where the unknown

Table 2: PSNR results for 2× enlarged images (from 256 × 256 to 512 × 512)

Image/Method	Bilinear	Wavelet (Haar)	Wavelet (Daub. 9/7)	Carey *et al.* [2]	HMM [3]	HMM Superres. [4]	Proposed technique
Lena	30.13	31.46	34.45	34.48	34.52	34.61	35.39
Elaine	30.60	31.71	33.26	33.29	33.31	33.40	33.40
Baboon	22.85	23.61	24.22	24.24	24.24	24.31	24.52
Peppers	30.01	31.45	33.94	34.03	34.04	34.10	34.46

Table 3: PSNR results for 4× enlarged images (from 128 × 128 to 512 × 512)

Image/Method	Bilinear	Wavelet (Haar)	Wavelet (Daub. 9/7)	Carey *et al.* [2]	HMM [3]	HMM Superres. [4]	Proposed technique
Lena	24.06	26.67	28.84	28.81	28.86	28.88	29.44
Elaine	25.38	28.06	30.44	30.42	30.46	30.51	30.92
Baboon	20.43	21.11	21.47	21.47	21.47	21.49	21.59
Peppers	24.37	26.89	29.57	29.57	29.58	29.60	30.16

detail coefficients are estimated as zeros. Results obtained by using three state-of-the-art methods which appeared in the literature relatively recently and reported in [2–4], are also shown. Our results demonstrate that the proposed method outperforms competing methods by as much as 0.7 dB in terms of measured distortion.

A subjective evaluation can be found in Figs. 4, 5 and 6. Figure 4 shows results obtained using a quadrant of a 'Zone Plate' test image. Zone plates are obtained by appropriate modulation of sinusoids so that a continuous variation of spatial detail in cycles per picture width/height is achieved. They are often used to measure directly the frequency response of processing algorithms and, as such, they are particularly useful for assessing resolution enhancement methods. We show differences between the original HR zone plate (used as ground truth) and reconstructed images obtained by applying three competing algorithms on a LR version of the original. In those difference images mid-grey tones represent zero error while darker

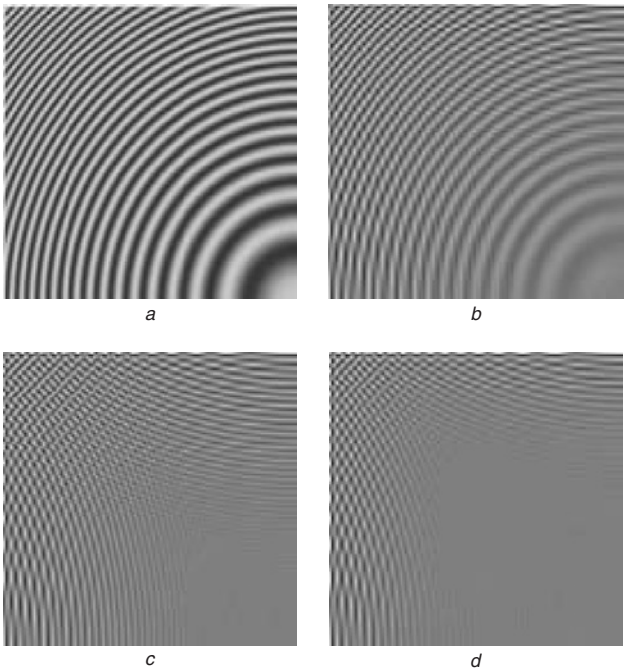

Fig. 4 *Result obtained using a 'Zone Plate' test image*

a Top-left quadrant of original 'Zone Plate' image
b Difference image using bilinear interpolation for resolution enhancement
c Difference image using 0s in the 'unknown' subbands
d Difference image using predicted coefficient values in the 'unknown' subbands

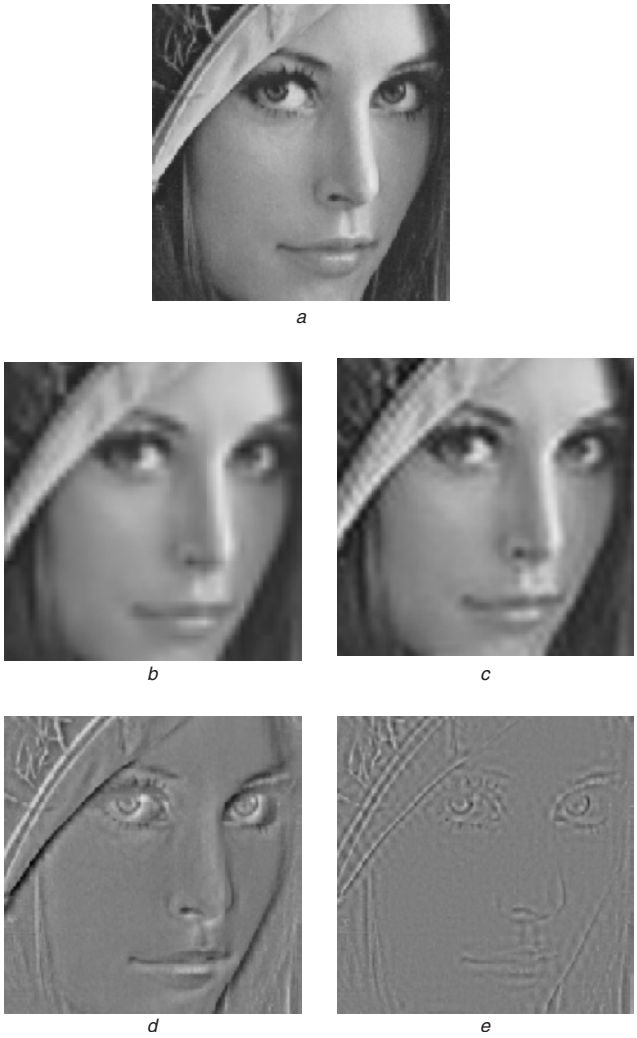

Fig. 5 *Comparison with bilinear interpolation*

a Extract from original 'Lena' image
b 4× reconstructions using bilinear interpolations
c The proposed method
d,e Error images corresponding to *b* and *c*, respectively

and lighter tones negative and positive errors respectively. Our results clearly indicate that the proposed technique safeguards the integrity of a wider range of spatial frequencies. Figure 5 shows a comparison with bilinear interpolation for a 4× reconstruction of an extract of 'Lena'. It can be seen that the proposed method offers considerable improvements both in terms of measured distortion and subjective appearance. An appreciation of the added detail contributed by our method is shown in Fig. 6. This is a comparison between the actual detail contained in HL_0 and the estimated detail using our method. Such added detail images were obtained by setting to zero all the elements of the LL_0 subband, keeping either the original or the estimated information in all remaining subbands (other than LL_0) and finally applying the inverse wavelet transform. In this way, the added high frequency information can be visualised as grey level variations on a mid-grey canvas.

It should be noted that the proposed algorithm works better with images having stronger edges. As can be seen from Tables 2 and 3, highly textured images such as 'Baboon' are reconstructed with a relatively lower PSNR gain of 1.42 dB. Alternatively, for natural images with more pronounced transitions such as 'Elaine', a PSNR improvement of 6.3 dB over bilinear interpolation was possible.

Table 4 provides a comparison in terms of coefficient sign agreement achieved by the estimation process. This is expressed as a percentage of correctly estimated signs relative to the total number of coefficients. Results are additionally classified according to coefficient magnitude. Only coefficients exceeding a pre-specified magnitude threshold θ are contributing to the results. This threshold is adjusted to achieve the percentage of contributing coefficients shown in the second column. It should be noted that the method proposed by Carey *et al.* [2] only estimates coefficients at extrema points while the other coefficients are assigned zero values. As a consequence, when all coefficients are taken into account, the coefficient percentage drops significantly. To be able to carry out a fair comparison with the other methods, two different percentage values are shown for [2]; the first is relative to the total number of coefficients while the second (in brackets) is relative to the subset of estimated coefficients. These results also confirm the observation that there is little correlation among signs of co-sited coefficients across scales. However, it is interesting to note that there is increasing correlation as the coefficient magnitude increases. As expected, the HMM based estimation, owing to the inherent randomness of estimated signs, achieves just under 50% agreement, increasing slightly with coefficient magnitude. Our results demonstrate that the exploitation of correlation in the proposed neighbourhood improves significantly coefficient sign agreement and for images such as 'Elaine', up to 93.36% agreement was achieved.

Table 4: Correct coefficient sign percentage for various techniques

Image/Method	% > θ	Carey *et al.* [2]	HMM [3]	Proposed technique
Lena	100	3.86 (52.80)	46.55	63.51
	20	5.01 (56.34)	50.97	72.58
	10	5.10 (59.23)	51.41	75.88
	2	5.30 (66.67)	50.80	85.52
Elaine	100	3.83 (51.10)	49.05	62.32
	20	5.25 (58.76)	49.79	75.38
	10	6.74 (64.15)	50.85	80.15
	2	15.13 (83.64)	51.69	93.36
Baboon	100	3.47 (48.34)	49.64	60.83
	20	3.56 (50.47)	49.34	69.54
	10	3.77 (55.88)	49.34	72.20
	2	1.65 (38.46)	47.23	70.30
Peppers	100	3.52 (49.17)	45.91	61.29
	20	5.35 (60.00)	49.29	75.94
	10	6.71 (69.39)	50.56	84.71
	2	11.59 (81.40)	54.34	92.43

The calculations are done for the coefficients whose magnitude exceeds the threshold θ. The threshold θ is found by constraining the percentage of coefficients extending the θ to the values in the second column.

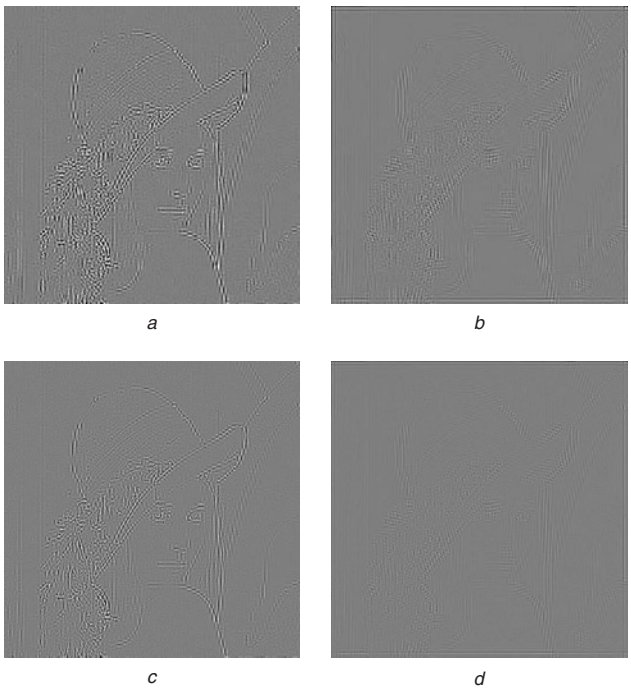

Fig. 6 *Added detail to 'Lena'*

a Original coefficients of the detail subbands for 2× enlargement
b Predicted coefficients for 2× enlargement
c Original coefficients of the detail subbands for 4× enlargement
d Predicted coefficients for 4× enlargement

5 Conclusions

We have presented a wavelet-domain image resolution enhancement algorithm which operates in a quad-tree decomposition framework and exploits wavelet coefficient correlation in a local neighbourhood sense. The proposed method employs linear least-squares regression to estimate the unknown detail coefficients. Our results show that our method outperforms conventional image resolution enhancement methods such as bilinear interpolation, for a wide range of standard test images. More importantly our method compares favourably with state-of-the-art competing methods operating in the wavelet domain both in objective and subjective terms. The availability of fast wavelet transform and filtering implementations in addition to the fact that key estimation parameters, such as linear regression weights, are relatively insensitive to image content makes our scheme attractive for real-time and on-line applications.

References

1 Chang, S.G., Cvetkovic, Z., and Vetterli, M.: 'Resolution enhancement of images using wavelet transform extrema extrapolation'. Proc. ICASSP95, May 1995, **4**, pp. 2379–2382

2 Carey, W.K., Chuang, D.B., and Hemami, S.S.: 'Regularity-preserving image interpolation', *IEEE Trans. Image Process.*, 1999, **8**, (9), pp. 1295–1297

3 Kinebuchi, K., Muresan, D.D., and Parks, T.W.: 'Image interpolation using wavelet-based hidden Markov trees'. Proc. ICASSP01, May 2001, **3**, pp. 7–11

4 Crouse, M.S., Nowak, R.D., and Baraniuk, R.G.: 'Wavelet-based statistical signal processing using hidden Markov models', *IEEE Trans. Signal Process.*, 1998, **46**, (4), pp. 886–902

5 Zhao, S., Han, H., and Peng, S.: 'Wavelet domain HMT-based image superresolution'. Proc. ICIP03, Sept. 2003, **2**, pp. 933–936

6 Nguyen, N.: 'Numerical techniques for image superresolution'. PhD Thesis, Stanford University, 2000

7 Nguyen, N., and Milanfar, P.: 'An efficient wavelet-based algorithm for image superresolution'. Proc. ICIP00, Sept. 2000 **2**, pp. 351–354

8 Ford, C., and Etter, D.M.: 'Wavelet basis reconstruction of nonuniformly sampled data', *IEEE Trans. Circuits Syst.*, 1998, **45**, (8), pp. 1165–1168

9 Mitevski, S., and Bogdanov, M.: 'Application of multiresolutional basis fitting reconstruction in image magnifying'. Proc. 9th Telecommunications Forum TELFOR 2001, Nov. 2001, pp. 565–568

10 Shapiro, J.M.: 'Embedded image coding using zerotrees of wavelet coefficients', *IEEE Trans. Signal Process.*, 1993, **41**, (12), pp. 3445–3462

11 Said, A., and Pearlman, W.A.: 'A new fast and efficient image codec based on set partitioning in hierarchical trees', *IEEE Trans. Circuits Syst.*, 1996, **6**, pp. 243–250