IMAGE RESOLUTION ENHANCEMENT USING WAVELET DOMAIN HIDDEN MARKOV TREE AND COEFFICIENT SIGN ESTIMATION

Alptekin Temizel

Graduate School of Informatics, METU, Ankara, Turkey

ABSTRACT

Image resolution enhancement using wavelets is a relatively new subject and many new algorithms have been proposed recently. These algorithms assume that the low resolution image is the approximation subband of a higher resolution image and attempts to estimate the unknown detail coefficients to reconstruct a high resolution image. A subset of these recent approaches utilized probabilistic models to estimate these unknown coefficients. Particularly, Hidden Markov Tree (HMT) based methods using Gaussian mixture models have been shown to produce promising results. However, one drawback of these methods is that, as the Gaussian is symmetrical around zero, signs of the coefficients generated using this distribution function are inherently random, adversely affecting the resulting image quality. In this paper, we demonstrate that, sign information is an important element affecting the results and propose a method to estimate signs of these coefficients more accurately.

Index Terms— image resolution, wavelet transforms, image processing, image enhancement

1. INTRODUCTION

A common feature of the wavelet domain image resolution enhancement 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 is 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). 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 among the same type of sub-bands. Edges identified by an edge detection algorithm in lower frequency subbands were used to formulate a template for estimating edges in higher-frequency sub-bands. Only the 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. Algorithm performance is also affected by the fact that the signs of the 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 de-noising. An extended version of this approach 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 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, in effect, randomly generated signs are assigned to the estimated coefficients. The HMT based method has been further developed so that it does not require any training data set [8]. and modeling with higher number of Gaussian distribution functions have been investigated [9]. The authors also suggested using a maximum a posteriori (MAP) approach to refine the resulting images.

In this paper we show that improved results could be achieved by HMT based methods with more accurate coefficient sign estimation.

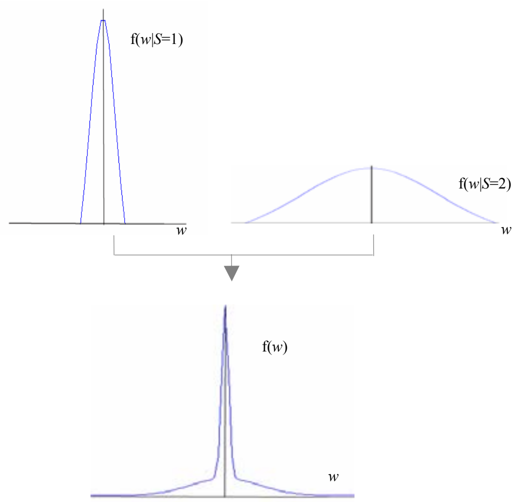

Figure 1: Two state Gaussian Model

In Section 2, we present the Hidden Markov Tree based approaches to image resolution enhancement and demonstrate their shortcomings. We propose a coefficient sign estimation scheme in Section 3 and then demonstrate the experimental results in Section 4. Concluding remarks are presented in Section 5.

2. IMAGE RESOLUTION ENHANCEMENT USING HIDDEN MARKOV TREE

Hidden Markov Tree (HMT) based methods make use of both the persistency and non-Gaussianity properties of wavelet coefficients. These methods model the unknown coefficients as belonging to mixed Gaussian distributions (states). The motivation being that the coefficient distributions which are heavy-tailed and have high density around zero could be modeled by a mixture of Gaussian distributions. Although in [7] higher number of states are used, generally a two state model where one Gaussian is used to model the coefficients around zero and one for the higher-magnitude coefficients, which constitutes the singularities is generally adopted. This two state model is illustrated in Figure 1. Each coefficient is assumed to fall into one of these distributions and the HMT model is trained using Expectation Maximization (EM) which finds the state transition parameters which are most likely to result in the coefficients in the observation set by iteration until a specified convergence error is achieved. In training, the coefficients in the same type of sub-band are tied together so that a single parameter is calculated for that type of sub-band to prevent over-fitting to the training image.
The 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 then the coefficient estimates are randomly generated using this distribution. Letting $p(S^{s-1}{}_l=m)$ denote the

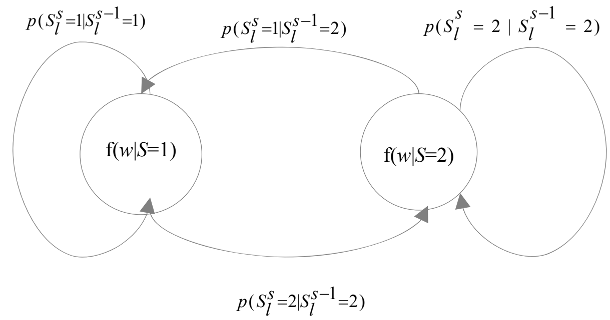

Figure 2: Coefficient magnitude state transitions.

probability of the state of coefficient l at scale s-1 being m, $p(S^s{}_l=m)$ which is the probability of the state of the coefficient l at scale s, that is to be estimated, can be written as (for two state model m= 1 or 2);

$$p(S_l^s=1)=p(S_l^s=1\,|\,S_l^{s-1}=1)\,p(S_l^{s-1}=1) + p(S_l^s=1\,|\,S_l^{s-1}=2)\,p(S_l^{s-1}=2) \tag{1}$$

$$p(S_l^s=2)=p(S_l^s=2\,|\,S_l^{s-1}=1)\,p(S_l^{s-1}=1) + p(S_l^s=2\,|\,S_l^{s-1}=2)\,p(S_l^{s-1}=2) \tag{2}$$

For each coefficient to be estimated, these state probabilities are calculated and the coefficient is assumed to have a state of highest probability. Figure 2 shows an illustration of the state transitions. The seminal work for this type of approach is the work by Crouse et al. [6] who applied the technique to signal denoising. Subsequently, this technique was adopted for image enlargement in [5].
An extended version of this approach is presented in [7]. The HMT based method has also been further developed so that it does not require any training data set [8]. In this method, state transition parameters are obtained from the low-resolution image in hand. This is achieved by using the coarser sub-bands of the image which are obtained by further applications of wavelet transform. Despite the need to calculate the state transition parameters for each image to be enlarged, the algorithm generates improved results as the parameters fit better to the particular image.
In the HMT methodology, once the states of the coefficients are estimated, coefficient magnitudes are assigned randomly using the Gaussian distribution which is associated with the state. As Gaussian distributions are symmetrical around zero, coefficients generated using these distributions have an equal chance of having assigned a negative or a positive sign. This does not pose a problem when the algorithm is used for denoising purposes such as in [6], as in these applications the coefficient signs in the finest scale already exist and does not need to be estimated. However, this is not ideal in resolution enhancement problems where correct sign estimation has an important effect on resulting image quality. This is demonstrated in Figure 3 where it is shown that incorrect sign information results in higher error than opting for not estimating the coefficients and using zero values instead.

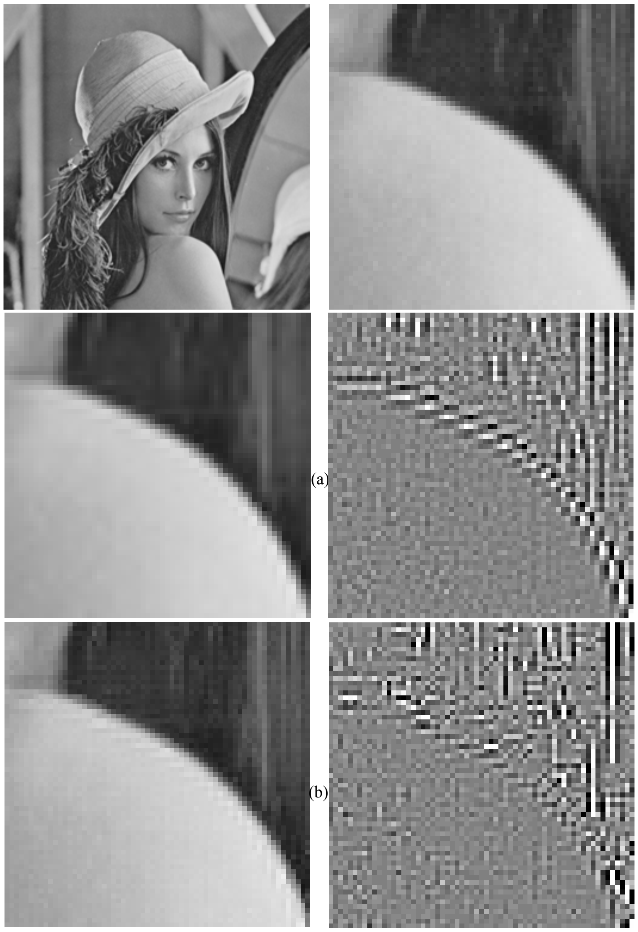

Figure 3: Original Lena image (top left) and an extract from this image (top right). (a) Extract from reconstructed image using zero for unknown coefficient values and amplified error. (b) Extract from reconstructed image using original wavelet coefficient magnitudes while coefficients signs are copied from their parents and amplified error images.

3. ESTIMATION OF SIGN INFORMATION

In this work, we make use of the fact that the coefficient sign and magnitude information are statistically independent, hence, we propose to estimate coefficient magnitude and sign separately. As the magnitude parameters are unknown, they are modeled using HMT which allows determining the hidden parameters from the observation as explained in the previous section. Once the state of the unknown coefficient is found, the magnitude is generated using a random number generator with this distribution function. For the coefficient sign estimation, we make use of the observation presented in [10] which states that there is higher correlation among the corresponding coefficients between high-pass wavelet filtered LR image and the unknown high-frequency subbands to be estimated. 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 Y 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 directions. Using this notation we can write that:

$$LL_0(z) = (\downarrow 2) L_{col}(z) L_{row}(z) Y(z) \quad (3)$$

$$HL_0(z) = (\downarrow 2) L_{col}(z) H_{row}(z) Y(z) \quad (4)$$

where LL_0 and HL_0 are respectively low-pass approximation subband and vertical detail subband of wavelet coefficients. We assume that LL_0 is the available LR image X whose resolution we seek to enhance and consequently try to estimate the elements of detail coefficients contained in detail subbands such as those in HL_0.

In the literature, the estimation is generally done using the next level detail subbands. Our approach is based on the assumption that a high-pass filtered undecimated version of the available LR image (LL_0):

$$HL_0'(z) = H_{row}(z) LL_0(z) \quad (5)$$

is sufficiently correlated with HL_0 to provide a basis for the estimation of the latter. HL_0' is obtained by row-wise high-pass filtering of LL_0, while the next level HL subband (HL_1) would be obtained by column-wise low-pass filtering and decimation in both directions following this. As the purpose of the estimation is to reconstruct the high-frequency details, we don't apply low-pass filtering and decimation to prevent loss of data and keep as much high-frequency information as possible. This is also intuitive as the salient image features in HL_0 are effectively in the same spatial locations as LL_0 and the HR image. The coefficient signs at the same spatial location in HL_0' are then used together with the HMT magnitude estimates. A similar treatment is possible for the LH_0 subband. In this case, high-pass filtering is applied in the vertical direction to LL_0 to obtain LH_0' which is then used as the estimation base for the coefficient signs.

Table 1 provides a comparison in terms of coefficient sign agreement achieved by the estimation process using Daubechies 9/7 filters. 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 the 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 in [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 sign agreement percentage drops significantly. To be able to carry out a fair comparison with the other methods, two different percentage values are shown for the Regularity Preserving Image Interpolation [2]; the first is relative to the total number of coefficients while the second (in brackets) is relative to the subset of estimated coefficients. The results confirm the observation that there is little correlation among signs of co-sited coefficients across scales as commonly expected in the literature [3,4]. However, the correlation is higher as the coefficient magnitude increases. As expected, the HMM based estimation [5], due to the inherent randomness of estimated signs, achieves just below 50% agreement, increasing slightly with coefficient magnitude.

4. EXPERIMENTAL RESULTS

For the experiments we have used the well known images of *Lena*, *Elaine*, *Baboon* and *Peppers*. The fact that the HH subband has a very small variance makes it difficult to estimate the coefficient states accurately. Also considering its relatively small impact on the resulting image quality, estimation of this subband is ignored and the HR images are reconstructed using the estimated LH and HL subbands followed by inverse wavelet transform.

Table 2 and 3 shows the Peak-Signal-to-Noise Ratio (PSNR). As well as wavelet based methods, the results have been compared against, bilinear and bicubic interpolation and New Edge Directed Interpolation (NEDI) [11]. Wavelet domain methods include WZP, where the unknown coefficients in high frequency subbands are filled with zeros, Regularity-Preserving Image Interpolation [2] and HMT method [5]. The visual quality improvement over HMT method is mostly perceptible in the neighborhood of the edges where correct sign information has the most visible effect. The results show that the proposed method improves on the HMT based methods and compares favorably with established methods in the literature.

5. CONCLUSIONS

In this paper, an image resolution enhancement algorithm operating in the wavelet domain was presented. The proposed algorithm aims to alleviate the main drawback of HMT based wavelet coefficient estimation methods resulting in inaccurate coefficient sign estimation. The results show that the results of this subset of algorithms could be improved by separating the magnitude and sign estimation and incorporating a more accurate sign estimation scheme.

6. REFERENCES

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

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

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

[4] Said A., Pearlman W.A., "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] Kinebuchi K., Muresan D.D. and Parks T.W., "Image Interpolation Using Wavelet-Based Hidden Markov Trees", Proc. ICASSP '01, vol. 3, pp. 7-11, May 2001.

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

[7] Zhao S., Han H. and Peng S., "Wavelet Domain HMT-Based Image Superres.", *Proc. ICIP '03*, vol. 2, pp. 933-936, Sep. 2003.

[8] Woo D.H., Eom I.K. and Kim Y.S., "Image Interpolation based on inter-scale dependency in wavelet domain", *Proc. ICIP '04*, vol. 3, 1687-1690, Oct. 2004.

[9] Kim J.H., Lee S.H. and Cho N.I., "Bayesian Image Interpolation Based on the Learning and Estimation of Higher Band Wavelet Coefficients", *Proc. ICIP '06,* Oct. 2006.

[10] Temizel A. and Vlachos T., "Wavelet Domain Image Resolution Enhancement", *IEE Proceedings Vision, Image and Signal Proc.*, vol. 153, no. 1, pp 25-30, Feb. 2006.

[11] Li X. and Orchard M.T., "New edge-directed interpolation", *IEEE Tr. Image Proc.*, Vol. 10, No. 10, pp. 1521-1527, Oct. 2001.

Image/ Method	%>θ	RP [2]	HMT[5]	Proposed
Lena	100	3.86 (52.80)	46.55	59.10
	20	5.01 (56.34)	50.97	69.05
	10	5.10 (59.23)	51.41	73.56
	2	5.30 (66.67)	50.80	84.51
Elaine	100	3.83 (51.10)	49.05	58.77
	20	5.25 (58.76)	49.79	68.24
	10	6.74 (64.15)	50.85	71.56
	2	15.13 (83.64)	51.69	77.86
Baboon	100	3.47 (48.34)	49.64	60.39
	20	3.56 (50.47)	49.34	69.63
	10	3.77 (55.88)	49.34	72.48
	2	1.65 (38.46)	47.23	77.37
Peppers	100	3.52 (49.17)	45.91	58.18
	20	5.35 (60.00)	49.29	67.10
	10	6.71 (69.39)	50.56	71.17
	2	11.59 (81.40)	54.34	84.34

Table 1: Correct coefficient sign percentages for various techniques. The calculations are done for the coefficients whose magnitude exceeds the threshold θ. The threshold θ is found by constraining the percentage of coefficients exceeding the θ to the values in the second column.

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[11]	34.10	32.89	23.87	33.54
WZP (Db.9/7)	34.45	33.26	24.22	33.94
Carey et al.[2]	34.48	33.29	24.24	34.03
HMT [5]	34.52	33.31	24.24	34.04
Proposed	34.68	33.40	24.24	34.18

Table 2: PSNR(dB) results for enlargement 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[11]	28.81	29.97	21.18	28.52
WZP (Db. 9/7)	28.84	30.44	21.47	29.57
Carey et al.[2]	28.81	30.42	21.47	29.57
HMT [5]	28.86	30.46	21.47	29.58
Proposed	28.96	30.58	21.47	29.74

Table 3: PSNR(dB) results for enlargement from 128x128 to 512x512.