

# Image interpolation for high-resolution display based on the complex dual-tree wavelet transform and hidden Markov model

Ramesh Kumar Lama<sup>1</sup> · Moo-Rak Choi<sup>2,3</sup> · Goo-Rak Kwon<sup>1</sup>

Received: 2 March 2015 / Revised: 4 November 2015 / Accepted: 4 January 2016 /  
Published online: 13 February 2016  
© Springer Science+Business Media New York 2016

**Abstract** This paper presents a new method for the interpolation of a full high-definition (HD) image based on the dual-tree complex wavelet transform (DT-CWT) and hidden markov model (HMM). In the proposed method, the DT-CWT is used to decompose the low-resolution image into different subbands. In wavelet domain interpolation, given image is assumed as the low frequency LL subband of the wavelet coefficients of a high-resolution image. The proposed method estimates the higher band coefficients by learning the correlation between the coefficients across the scale. In this paper, the relationship between the wavelet coefficients across the scale is described by HMM, and each wavelet coefficient is modeled by a Gaussian mixture having multiple means and variances. Experimental results show that the proposed algorithm yields images that are sharper compared to several other methods that we have considered in this paper.

**Keywords** Image interpolation · Dual tree complex wavelet transform · Hidden markov model · Expectation maximization

## 1 Introduction

Image interpolation is a process that is often used to perform an image zoom or simply to increase the resolution of an original image from its down-sampled version [10]. Interpolation has been widely used in many image processing applications such as facial reconstruction,

---

✉ Goo-Rak Kwon  
grkwon@chosun.ac.kr

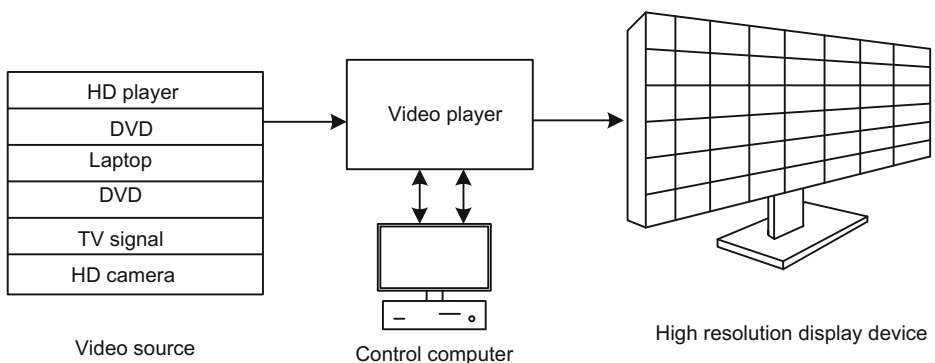
<sup>1</sup> Department Information and Communication Engineering, Chosun University, 375 Seosuk-Dong, Dong-Gu, Gwangju 501-759, Republic of Korea

<sup>2</sup> Division of IT Marketing, LG Electronics, 150-721 Seoul, Republic of Korea

<sup>3</sup> Division of Business Research, Sung Gwang, IT 555-809 Yeosu, Republic of Korea

multiple description coding (MDC), super resolution and digital high-definition television (HDTV) [3, 9, 18]. The rapid advancements in hardware and display devices have made it possible to process high-resolution digital color images as shown in Fig. 1. In order to efficiently utilize the display power of these state-of-the-art viewing devices, input signals from a lower-resolution source must be upscaled to higher resolutions through interpolation. However, the multichannel nature of color images demands efficient signal processing algorithms that take into account the existing inter-channel correlations when performing image size expansion. In this work, we propose a new technique that generates sharper super-resolved images. Various techniques have been proposed in order to scale the image from low resolution to high resolution and its reverse. Fundamentally, they can be classified into three categories: isotropic interpolation-based, edge-directed, and transform-based methods. The isotropic interpolation method [3, 9, 10, 18] considers the source image as a sub-sampled discrete version of the original “continuous” image. Nearest-neighbor, bilinear, and bi-cubic interpolation are some of the examples of this method. In the nearest-neighbor method, the source image is scaled by sampling the nearest pixels in the source image. This is also known as zero-order interpolation. It is a computationally efficient algorithm; however, the up-scaled image quality is degraded owing to the aliasing effect. In the bilinear and the bi-cubic interpolation methods, image quality is improved compared to the nearest neighbor method; however, the high-frequency content of the image still suffers from a blurring effect. Edge directed interpolation (EDI) [4, 11, 12, 19] uses statistical sampling to ensure quality while scaling the image instead of averaging the pixels, as in a bilinear interpolation average of the covariance of pixels. Hence, the sharpness and the continuity of the edges are preserved. Li [12] proposed an EDI using the statistical and geometrical properties of the pixel to interpolate the unknown pixel. A covariance of the pixel in a local block is required for the computation of the prediction coefficient. This algorithm works well for preserving the sharpness and continuity of edges. However, only four pixels are taken from the source image; hence, the prediction error is obvious in this algorithm. A modified training window structure has been proposed [16, 17] in order to overcome this shortcoming.

The training window in the second step of NEDI is modified to the form of sixth order linear interpolation with a  $5 \times 9$  training window. M. Li [11] suggested the Markov random field (MRF)-based model. The image is modeled with MRF, and edge estimation is extended in other directions by increasing the neighboring pixels in the kernel. Similarly, an improved edge-directed image interpolation algorithm with low time complexity which is the combination of Newton’s method and edge-directed method is proposed by Wang Xing-Yuan and Chen Zhi-Feng [20].



**Fig. 1** High resolution display video wall

This algorithm is able to preserve the sharpness of edges in the interpolated image. Transforms like DCT and wavelet transform are efficient for compressing the energy into a few transformed coefficients. As a result of multiplying these transformed coefficients, a scaled image can be obtained. The algorithm in [17] is based on the use of orthogonal wavelets as preconditioning transforms. Lama et al. in [8] suggested the adaptive directional to interpolate the image. In this method, directional lifting based wavelet is applied. The theory of projection onto convex sets (POCS) is the basis for the iterative technique proposed by K. Ratakonda in [14]. In this algorithm, scaling is accomplished in two steps. First the initial image is obtained by the conventional bilinear technique, then the iterative algorithm is applied on the initial image to obtain the final magnified image. Additionally, Chun-Ho Kim [5] proposed the winscale algorithm which is based on an area pixel model and domain filtering, and it uses a maximum of four pixels from the original image to calculate the pixel of a scaled image. The winscale algorithm uses the same hardware as the bilinear method; however, it is more complex in terms of the number of operations per pixel. Limitations of winscale in described in detail by E. Aho et al. [1].

Finally, we present a new algorithm that interpolates the image with consideration of preserving the original image edge and computational cost. The proposed technique uses (DT-CWT) to decompose the low-resolution image into different sub-band coefficients. We assume that a given image is the LL band of the wavelet coefficients of a high-resolution image as shown in Fig. 2. In order to predict the coefficients on the finest scale, we model the statistical relationship between the coefficients at coarser scales using a hidden markov tree (HMT). Finally, the magnified image is obtained combining the predicted high-frequency sub-band images and input image by using inverse discrete complex wavelet transform (ID-CWT).

## 2 Dual tree complex wavelet transform

The DT-CWT is an enhancement to the traditional critically sampled DWT to overcome the shortcomings with important additional properties: it is nearly shift invariant and offers higher directional selectivity [7]. There are six directional subbands capturing features along the lines at angles of  $\pm 15^\circ$ ,  $\pm 45^\circ$ , and  $\pm 75^\circ$ . The DT-CWT allows the characterization of images by providing explicit information about singularities in a broad range of orientations. The DT-CWT of a signal  $X(n)$  is implemented by using 2 critically sampled DWTs in parallel on the same data, as shown in Fig. 3. Since for  $N$  point signals, they generate  $2N$  DWT coefficients, the transform is doubly expansive. The filters are designed in a specific way such that the sub-band signals of the upper DWT can be interpreted as the real part of a complex wavelet transform, and the sub-band signals

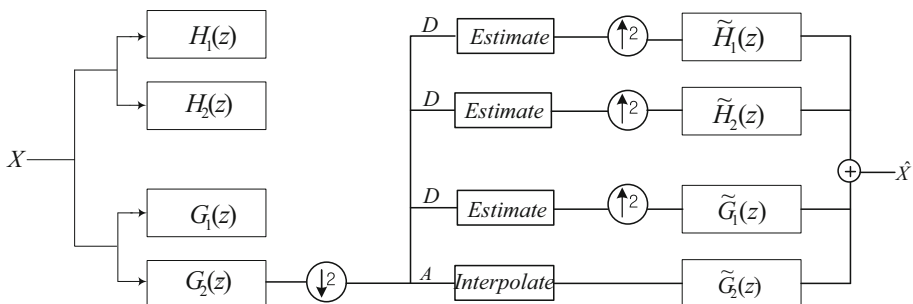
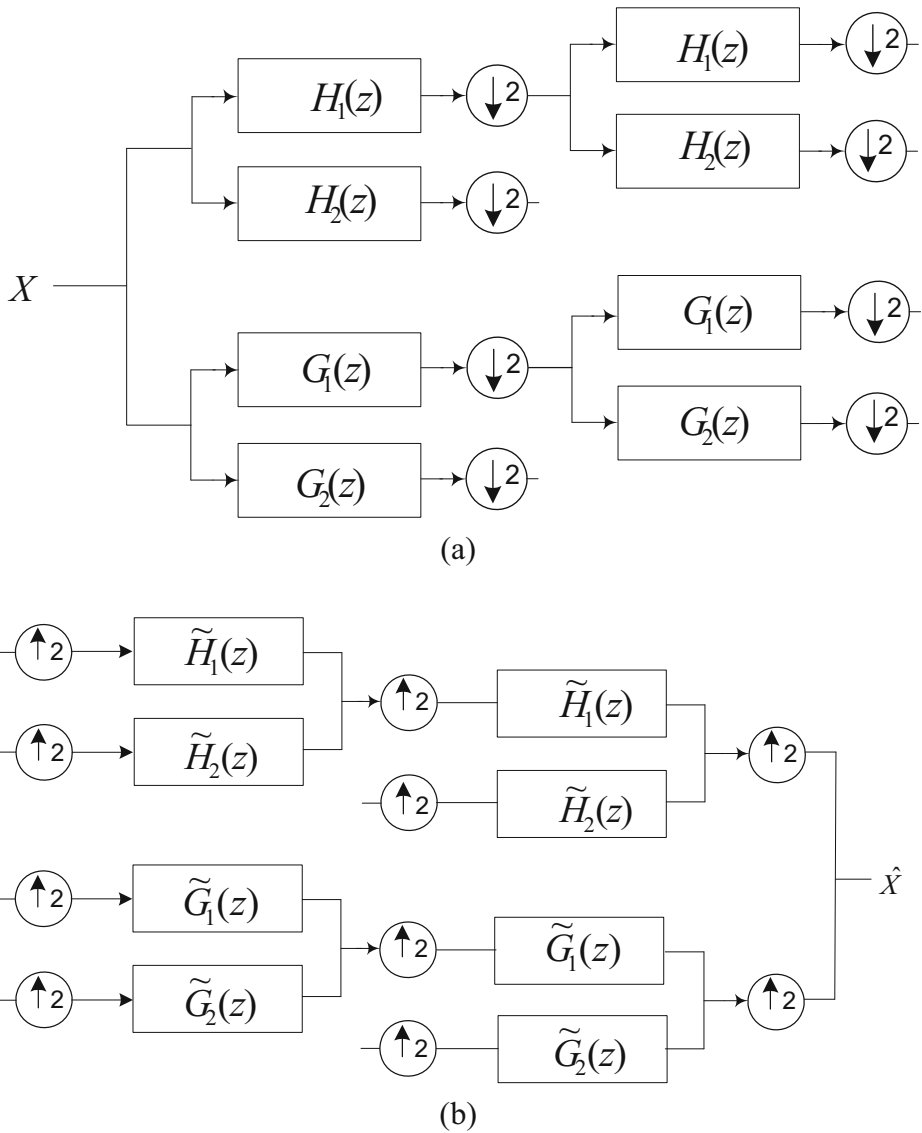


Fig. 2 Interpolation using DT-CWT and HMM



**Fig. 3** Dual-tree complex wavelet transform **a** Analysis **b** Synthesis

of the lower DWT can be interpreted as the imaginary part. When designed in this way, the DT-CWT is nearly shift invariant, in contrast to the classic DWT.

**2.1 Statistical modeling of complex wavelet transform**

The main idea of the wavelet domain interpolation is to exploit the properties of wavelet coefficients for estimating the extreme points in the higher frequency bands [13]. First, the intra-scale characteristic of coefficients is studied. Within the same scale, coefficients representing the smooth region in an image have small magnitudes while large magnitudes represent singular regions. Each wavelet coefficient is modeled by a gaussian mixture model

(GMM) that has multiple means and variances. The overall marginal probability distribution function (PDF) is given by a two-state model with hidden states ( $S_i = 1/2$ ), where one Gaussian is used to model the coefficients around zero for small region, while another is used for the higher-magnitude coefficients which constitute the singularities. These two states are illustrated in Fig. 4. Similarly, the PDF plot of wavelet coefficients of Lena image is shown in Fig. 5. Each coefficient is assumed to fall into one of these distributions, and HMT model is trained by the expectation maximization (EM). Then, the overall density function is given by:

$$f(w_i) = p(S_i = m)f(w_i/S_i = m), \tag{1}$$

where,  $p(S_i = m)$  is obtained using the EM. Given the state  $S_i$ , the conditional probability  $f(w_i | S_i = m)$  of the coefficient value  $w_i$  corresponding to the Gaussian distribution is obtained as:

$$f(w_i/S_i = m) = \frac{1}{\sqrt{2\pi\sigma_{i,m}^2}} \exp\left(-\frac{w_i^2}{2\sigma_{i,m}^2}\right) \tag{2}$$

Second, the unknown high-frequency coefficients are estimated by using the inter-scale relationship, which is defined by the Markov stochastic model [2]. This model uses both the persistency and non-Gaussianity properties of the wavelet coefficients. For any image, each parent wavelet coefficient in the HMT hierarchy has four children. Owing to persistence, the relative size of the coefficients propagates across the scale. To describe these dependencies, the two-state HMT model uses the state transition probabilities  $f(S_i = m | S_{p(i)} = n)$  between the hidden states  $S_i$  of the children, given that of the parent  $S_{p(i)}$ :

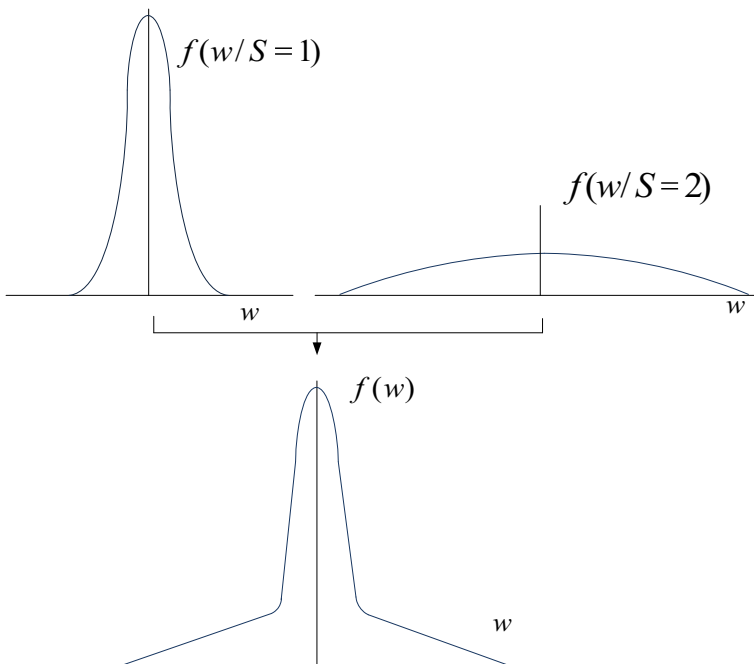


Fig. 4 Two state Gaussian model

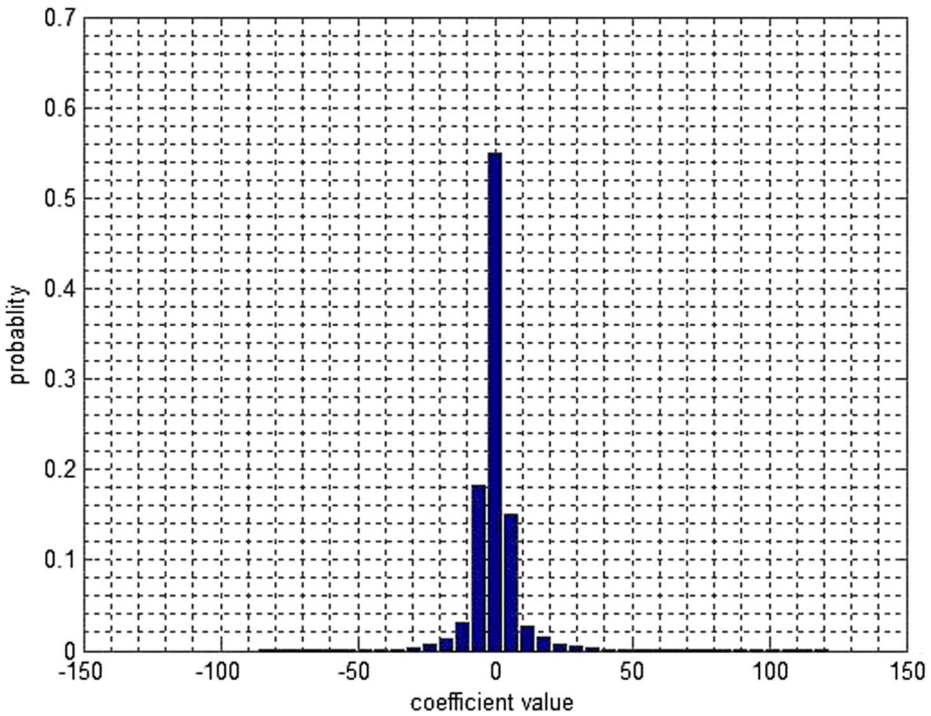


Fig. 5 PDF plot of wavelet coefficient in Lena image

$$f(S_i = m/S_{p(i)} = n) = \begin{pmatrix} f(S_i = 1|S_{p(i)} = 1)f(S_i = 1|S_{p(i)} = 2) \\ f(S_i = 2|S_{p(i)} = 1)f(S_i = 2|S_{p(i)} = 2) \end{pmatrix} \tag{3}$$

where according to the persistence assumption,  $f(S_i = 1|S_{p(i)} = 1) \gg f(S_i = 2|S_{p(i)} = 1)$  and

$$f(S_i = 2|S_{p(i)} = 2) \gg f(S_i = 2|S_{p(i)} = 1) \tag{4}$$

Therefore, the associated probabilities  $P$  are

$$\begin{aligned} P(S_i = 2|S_{p(i)} = 1) &= 0 \\ P(S_i = 1|S_{p(i)} = 1) &= 1 \end{aligned} \tag{5}$$

The probability that the large coefficient changes into the small coefficient converges to 0.5 across scales.

$$\begin{aligned} P(S_i = 1|S_{p(i)} = 2) &= 1/2 \\ P(S_i = 2|S_{p(i)} = 2) &= 1/2 \end{aligned} \tag{6}$$

Therefore, state transition probabilities  $\varepsilon$  can be approximated by following the model of universal HMT [15].

$$\begin{bmatrix} \varepsilon_{i,p(i)}^{11} & \varepsilon_{i,p(i)}^{12} \\ \varepsilon_{i,p(i)}^{21} & \varepsilon_{i,p(i)}^{22} \end{bmatrix} = \begin{bmatrix} 1 & 0.5 \\ 0 & 0.5 \end{bmatrix} \tag{7}$$

Once the state of the child coefficient is confirmed, we generate corresponding coefficient values based on different interpolation methods [6]. The state value of coefficients clearly indicates the pixel characteristics around the location i.e, the state value  $S_i = 1$ , indicates the coefficient belongs to low frequency region while the high  $S_i = 2$  corresponds to high frequency regions. In the proposed method unknown coefficients values belonging to the high frequency region are generated using a linear interpolation instead of random values based on the variance. Coefficients along the direction yielding the lowest high frequency energy among all the directions are selected for interpolation. While, in case of low frequency region, pixel intensity of neighboring pixels are almost identical. Thus we generate the child coefficients based from the variance same as its parents.

In this paper, the DWT wavelet coefficients are modeled using three independent HMT models. In this way, we tie together all trees belonging to each of the three detail subbands to decrease the computational complexity and prevent overfitting to the data.

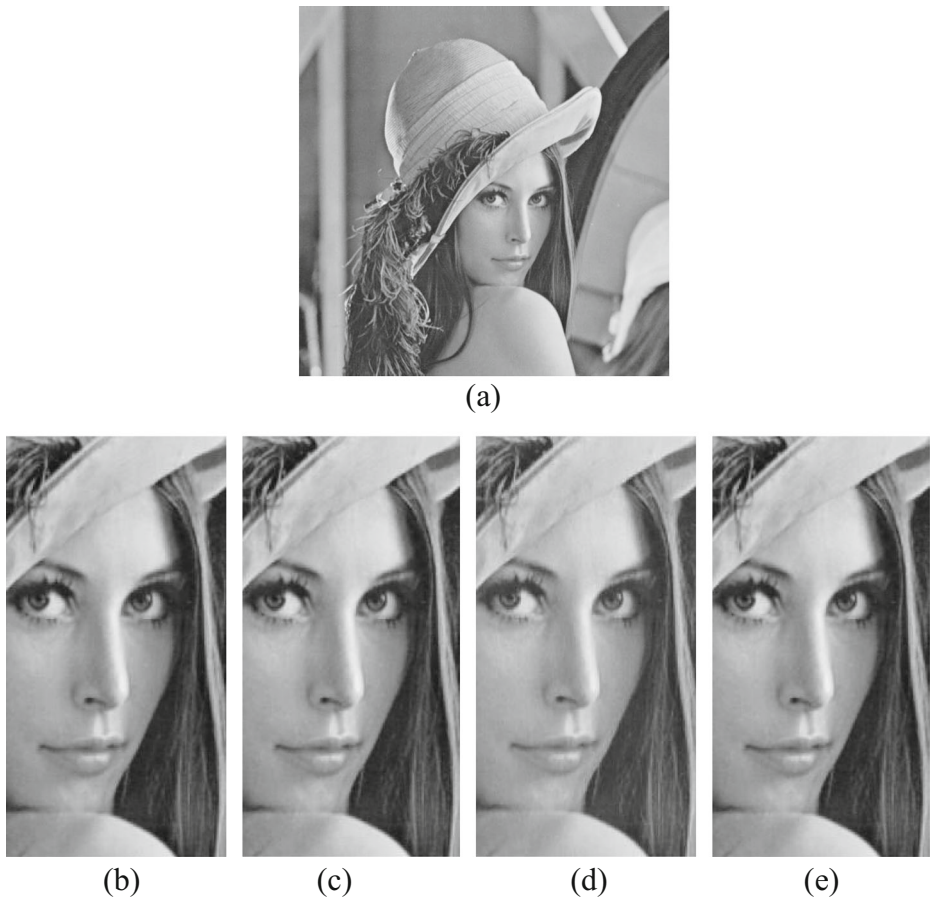
For the DT-CWT coefficients, the complex wavelet coefficient comprise real and complex components  $w = (w_r, w_c)$ . Thus, the conditional probability for the complex coefficient  $f(w_i | S_i = m)$  is expressed as:

$$f(w_i/S_i = m) = \frac{1}{\sqrt{2\pi\sigma_{i,m}^2}} \exp\left(-\frac{wr_i^2}{2\sigma_{i,m}^2}\right) \exp\left(-\frac{wc_i^2}{2\sigma_{i,m}^2}\right) \quad (8)$$

Scale-to-scale Markov transitions of the complex wavelet HMT and real DWT HMT have almost identical structures. The differences will be the use of six subband trees instead of three.

### 3 Experimental results

In this section, we present experimental results of evaluating the performance of the proposed algorithm using different test images, one of which is shown in Fig. 5. The performance is compared with several standard methods, such as bilinear interpolation, edge directed interpolation and in [12] and adaptive directional lifting based interpolation. We consider two types of performance metrics, which are visual quality and peak signal to noise ratio (PSNR). The visual comparison is performed using two images. We have selected gray scale Lena image of size  $512 \times 512$  and high resolution color image of size  $1920 \times 1080$ . Figure 6 shows the interpolation of Lena image using different interpolation method. Figure 6a shows the original image. Figure 6b shows an interpolated image using new edge directed interpolation (NEDI) [12]. Figure 6c shows a magnified image using the bilinear method. Figure 6d shows a magnified image using the adaptive directional lifting (ADL) based method. Finally, Fig. 6e shows the image upscaled using the proposed method. Visually, experiments from the three different filters yield very similar results and conclusions, but the PSNR tells quite a different story. In order to obtain PSNR measurement, we take the  $2N \times 2N$  and downsample it to obtain  $N \times N$  image. PSNR values of reconstructed images are shown in Table 1. The reconstructed images generally attain most of their quality in few iterations and practically do not change after 5–6 iterations of EM, both in visual quality and in PSNR measurements. The log likelihood of high frequency coefficients estimated during expectation maximization remains constant after 5–6 iterations as shown in Fig. 7. Since



**Fig. 6** Experimental results: **a** Original image Lena, **b** Interpolated image using the NEDI method, **c** Bilinear method, **d** ADL-based method, and **e** proposed method

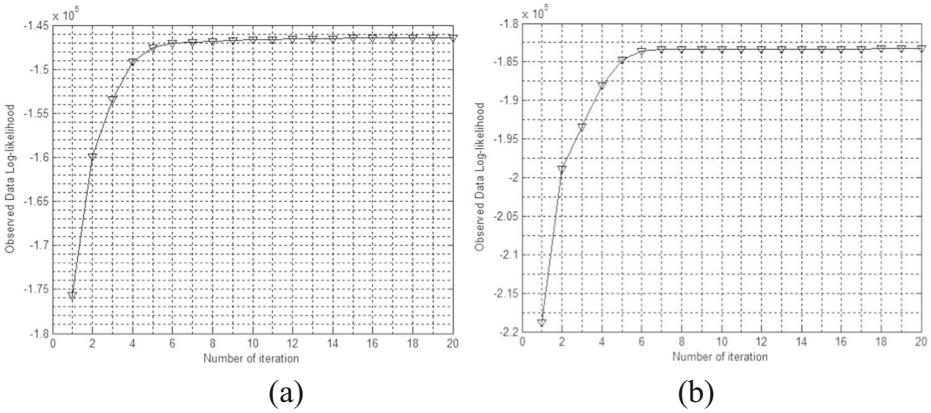
we obtain the variances and mean of coefficients using only few iterations, the use of EM does not make the overall system more complex. Although the PSNR is not a good indication of image quality, it is nevertheless frequently used, and the results are tabulated in Table 1 for the bilinear interpolation, NEDI methods and ADL method. The highest PSNR values among the different methods are highlighted in bold. From this

**Table 1** Competitive results

Image	Bilinear [dB]	NEDI [12] [dB]	ADL-based method [8] [dB]	Proposed method [dB]
Rocket	27.01	27.36	28.00	<b>28.19</b>
Flower	27.60	27.35	27.98	<b>28.10</b>
Lena	30.16	30.12	30.65	<b>30.81</b>
Barbara	24.33	23.85	24.53	<b>24.65</b>
Boat	27.50	27.18	27.73	<b>29.23</b>

The highest PSNR values among the different methods are highlighted in bold. From this table values, the proposed method outperforms the conventional algorithms in terms of PSNR



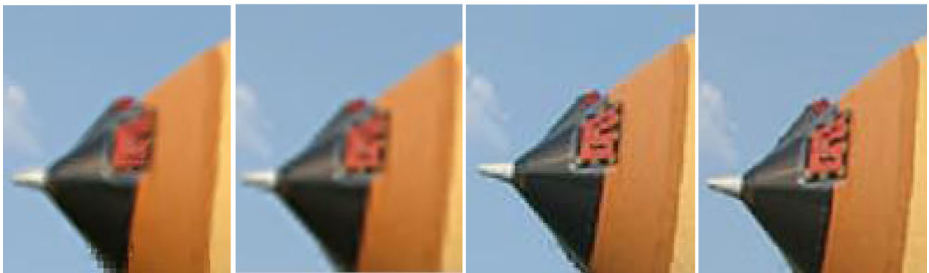


**Fig. 7** Log-likelihood vs number of iterations **a** Real HH coefficients **b** Imaginary HH coefficients

table values, the proposed method outperforms the conventional algorithms in terms of PSNR. Similarly, In Fig. 8, we demonstrate the interpolation of high resolution color image using different methods and their comparison with the proposed method. One can see that the proposed algorithm resizes the high-frequency component better without the ringing artifacts present in bilinear and other methods. We used a half-size downsampled image using nearest neighbor method to interpolate to the original size and compute the PSNR.



(a)



(b)

(c)

(d)

(e)

**Fig. 8** Experimental results: **a** Original image, **b** Interpolated image using the NEDI method, **c** Bilinear method, **d** ADL-based method, and **e** proposed method

## 4 Conclusions

In this paper, we proposed an efficient up-sampling algorithm based on the DT-CWT and HMM method. The proposed up-sampling method yields much better visual quality in high resolution color images and better PSNR for both the color and gray scale images compared to current state of the art in the literature of high-resolution display systems. The better performance comes at the expense of increase in complexity. However it is not a significantly high. We believe, many modifications can be done on the proposed method that can further improve its performance and reduce computational complexity. Finally, owing to the combined DT-CWT and HMM concept of the proposed method, it can be efficiently used for the interpolation of HD images.

**Acknowledgments** This study was supported (in part) by research funding from Chosun University, 2016.

## References

1. Aho E, Vanne J, Kuusolonna K, Hämäläinen TD (2005) Comments on Win scale, an image –scaling algorithm using an area Pixel model. *IEEE Trans Circuits Syst Video Technol* 15(3):454–455
2. Crouse MS, Nowak RD, Baraniuk RG (1998) Wavelet-based statistical signal processing using hidden Markov models. *IEEE Trans Signal Process* 46(4):886–902
3. El-Khamy SE, Hadhoud MM, Dessouky MI, Salam BM, Abd El-Samie FE (2005) Efficient implementation of image interpolation as an inverse problem. *Digit Signal Process* 15(2):137–152
4. Jeon G, Lee J, Kim W, Jeong J (2007) Edge direction-based simple re-sampling algorithm. In: *Proceedings of International Conference on Image Processing (ICIP)*, vol 7, pp 397–400
5. Kim CH, Seong SM, Lee JA, Kim LS (2003) Win scale, an image–scaling algorithm using an area Pixel model. *IEEE Trans Circuits Syst Video Technol* 13(6):549–553
6. Kinebuchi K, Muresan DD, Thomas W, Parks (2001) Image interpolation using wavelet based hidden Markov trees. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing, 2001. (ICASSP'01)*, vol 3, pp 1957–1960
7. Kingsbury NG (1999) Image processing with complex wavelets. *Philos Trans R Soc London A Math Phys Sci* 357(1760):2543–2560
8. Lama RK, Choi MR, Kim J-W, Pyun JY, Kwon GR (2014) Color image interpolation for high resolution display based on adaptive directional lifting based wavelet transform. In: *Proceedings of International Conference on Consumer Electronics (ICCE)*, pp 219–221
9. Lee S, Paik J (1993) Image interpolation using adaptive fast B-spline filtering. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, vol 5, pp 177–180
10. Lehmann TM, Gönner C, Spitzer K (1999) Survey: interpolation methods in medical image processing. *IEEE Trans Med Imaging* 18(11):1049–1075
11. Li M, Nguyen TQ (2008) Markov random field model-based edge-directed image interpolation. *IEEE Trans Image Process* 17(7):1121–1128
12. Li X, Orchard MT (2001) New edge-directed interpolation. *IEEE Trans Image Process* 10(10):1521–1527
13. Pentland AP (1994) Interpolation using wavelet bases. *IEEE Trans Pattern Anal Mach Intell* 16(4):673–688
14. Ratakonda K, Ahuja N (1998) POCS-based adaptive image magnification. In: *Proceedings of International Conference on Image Processing (ICIP)*, vol 3, 203–207
15. Romberg JK, Choi H, Baraniuk RG, Kingsbury NG (2000) Hidden Markov tree modeling for complex wavelet transforms. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, vol 1, pp 133–136
16. Tam WS (2009) Modified edge-directed interpolation. In: *Proceeding of 17th European Signal Processing Conference (EUSIPCO 2009)*, pp 283–287
17. Tam WS, Kok CW, Siu WC (2010) Modified edge-directed interpolation for images. *J Electron Imaging* 19(1):013011-1–013011-2
18. Thévenaz P, Blu T, Unser M (2000) Interpolation revisited. *IEEE Trans Med Imaging* 19(7):739–758
19. Wang Q, Ward R (2001) A new edge-directed image expansion scheme. In: *Proceedings of International Conference on Image Processing*, vol 3, pp 899–902
20. Yuan WX, Feng CZ (2009) An improved image interpolation algorithm. *Multidim Syst Signal Process* 20: 385–396



**Ramesh Kumar Lama** received the B.S. degree from Purwanchal University Nepal and Masters degree from Department of Information and Communication Engineering, Chosun university in 2004 and 2011. Currently, he is pursuing Ph. D. degree in Chosun university. His research interests are image and video coding, watermarking and multimedia security.



**Moo-Rak Choi** received the M.S. Degrees in Industrial Engineering from Hanyang University in 2000. He is in Ph. D Student in Mechatronics Engineering, Korea University and working as a research engineer in LG Electronics from 2000. He has an experience with development & design of: Multimedia Communication/ Embedded System Commercial Display System such as indoor & outdoor video wall Image/Video Quality Enhancement for Real time processing.



**Goo-Rak Kwon** received a Ph. D. degree from Korea University in Mechatronics in 2007. In 2008, he joined the Department of Information and Communication Engineering at Chosun University, where he is currently an associate professor. His interest research fields are A/V signal processing, multimedia communication, and applications. He is currently a Member in the IEEE, IEICE, and IS&T. He has been a Member in the Institute of Electronics Engineers of Korea (IEEK). And also he has served as an Editor for Journal of the Multimedia Society of Korea (KMMS).