# NEW IMAGE INTERPOLATION ALGORITHMS BASED ON DUAL-TREE COMPLEX WAVELET TRANSFORM AND MULTILAYER FEEDFORWARD NEURAL NETWORKS

SHIRIN SALEHI[1] AND HOMAYOUN MAHDAVI-NASAB[2]

[1]Young Researchers Club
[2]Department of Electrical Engineering
Najafabad Branch, Islamic Azad University
Isfahan, Iran
sh_salehi@sel.iaun.ac.ir; mahdavinasab@iaun.ac.ir

ABSTRACT. *In this paper, new image resolution enhancement algorithms based on complex wavelet transform and feedforward neural networks are proposed. The wavelet subbands corresponding to high-resolution images are estimated by neural networks using low-resolution counterparts. High-resolution images are then reconstructed employing the inverse transform. We take advantage of dual-tree complex wavelet transform, such as approximate shift invariance, substantial reduced aliasing and directional selectivity, to obtain a richer representation of local structures in interpolated images. These properties make the subband estimation process more effective and lead to more accurate reconstruction of texture and edge regions. We also present a simplified version of the proposed algorithm to reduce computational cost without significant performance reduction. Subjective comparisons and objective quality assessments indicate notable improvement over the conventional bicubic and bilinear interpolation techniques and some typical recently proposed methods.*
**Keywords:** Image interpolation, Image resolution, Dual-tree complex wavelet transform, Multilayer feedforward neural networks

1. **Introduction.** Resolution of a digital image is defined as the number of pixels per unit distance [1]. Super-resolution is the process of producing a high-resolution (HR) image using one or more low-resolution (LR) images or frames [2]. In this paper the concentration is on single-frame super-resolution which employs just one LR image as the source. This kind of super-resolution is known as image interpolation, resolution enhancement, scaling, zooming and enlargement. Image interpolation is necessary to magnify images so that output images contain more pixels. Due to the physical limitations of imaging hardware, image interpolation techniques are commonly used to construct an HR image from its LR version by estimating the unknown pixel intensities. The challenge is to process the image in such a way that keeps the interpolated image sharp and clear to human observer. Image interpolation has applications in medical imaging, remote sensing, digital photography, printing industry and video transmission [3-11]. It is an essential process in HDTV or medical image display technologies [9]. When the resolution of an image generated by a host PC is different from the screen resolution of a Liquid Crystal Display (LCD), image scaling process is necessary. Nowadays, satellite images are used in many applications such as geosciences, astronomy and geographical information systems. Image interpolation is a well-known method to increase the resolution of satellite digital images [4].

The conventional methods of interpolation, such as nearest neighbor, bilinear and bicubic, use the weighted sum of neighboring pixels to estimate the interpolated pixel. Although these methods are simply implemented, they suffer from blurring and blocking artifacts. They are also unable to generate the high-frequency or fine details [2,10-14]. These shortcomings have been the motivation of proposing new and more sophisticated methods to improve the visual quality of interpolated images.

Edge-directed methods use the LR edge information to estimate the edge mapping of the HR image [13-16]. The estimation helps to guide the process so that the edges are not smoothed. In [13], a nonlinear edge-directed interpolation algorithm using directional filtering and data fusion is proposed. In this algorithm the directional estimates of the missing samples are computed and fused for interpolation. The basic idea in [14] is to estimate local covariance coefficients from an LR image to adapt interpolation based on the geometric duality between LR and HR covariances.

Other important interpolation algorithms are those which employ neural networks. Multilayer feedforward perceptrons (MLPs) have been applied successfully to solve some difficult and diverse problems using a popular supervised training algorithm known as back propagation (BP) [17-19]. Different variations of BP are presented including gradient descent with momentum, variable learning rate, resilient, conjugate gradient, Quasi-Newton and Levenberg-Marquardt algorithms in order to optimize network parameters and avoid local minima [18,20]. In [21], an interpolation algorithm based on MLP is proposed. Classification-based neural networks are used for interpolation in [22]. A human visual system directed neural network-based scheme for natural images is proposed in [23]. Also, researchers of [24,25] have presented methods of improving the spatial resolution of images by means of MLP neural networks. In [26], a wavelet-based resolution enhancement algorithm is proposed which uses neural networks for estimating wavelet subbands. The HR image is finally produced by the wavelet synthesis procedure.

Recently, some complex wavelet domain resolution enhancement algorithms have been proposed for satellite images [4,5]. Complex wavelet transform (CWT) is recently emerged in image processing [27,28]. CWT of an image produces two low-frequency and six high-frequency complex-valued subbands. The high-frequency subbands are oriented at angles $\pm 15°$, $\pm 45°$ and $\pm 75°$. Each complex-valued subband is composed of a real and an imaginary part. In high-frequency subbands the real and imaginary parts are oriented at the same angle. In [4], CWT is used to decompose an LR image into different subbands. The high-frequency subbands are interpolated using bicubic interpolation. The input image is also interpolated with the half of the interpolation factor used for the high-frequency subbands. Finally, the interpolated subbands and input image are combined by inverse CWT. In [5], the initial estimate of the HR image is constructed by applying a cycle-spinning methodology introduced in [29]. CWT highpass coefficients together with the LR image are used to reconstruct the HR image by inverse CWT.

In this paper an image resolution algorithm based on dual-tree complex wavelet transform (DT-CWT) and MLP neural networks is proposed. Each of wavelet subbands corresponding to an HR image is estimated using a subband estimator from its LR version. The HR image is then obtained using inverse DT-CWT (IDT-CWT). Directional selectivity of DT-CWT makes the subband estimation process very successful and leads to more accurate reconstruction of texture and edge areas. Subjective comparisons and objective quality assessments indicate additional improvements over conventional and more recent interpolation algorithms, described in [13,14], as edge directed methods, and [26], as a neural network and wavelet-based method.

This paper is organized as follows. Section 2 reviews the main features of DT-CWT. Section 3 describes the proposed interpolation algorithm. Section 4 presents a simplified

version of the algorithm which is suitable for fast and practical implementations. Section 5 presents the experimental results. Section 6 concludes this paper.

2. **The Dual-Tree Complex Wavelet Transform.** DT-CWT is a relatively new enhancement to the DWT. It was proposed to overcome the shortcoming of DWT [28]. In DWT, wavelet coefficients tend to oscillate positive and negative around singularities which makes the singularity extraction complicated. A small shift of the signal also results in greatly different coefficients around singularities. It makes wavelet-based processing intricate, because the algorithms must be capable of coping with different wavelet patterns caused by shifted singularities. Any wavelet coefficient processing upsets the delicate balance between the forward and inverse transforms, leading to artifacts in the reconstructed signal. Furthermore, lack of directional selectivity greatly complicates processing of image features like ridges and edges [28]. DT-CWT nearly overcomes all these shortcomings at the cost of $2^d$ for a $d$-dimensional signal. DT-CWT uses two real DWTs. One produces the real and the other the imaginary part of the transform. The analysis and synthesis filter banks are illustrated in Figures 1 and 2. $h_0(n)$ and $h_1(n)$ are the lowpass and highpass filter pair for the upper filter bank and $g_0(n)$ and $g_1(n)$ are for the lower one. It is 2-times expansive in 1-D because the total output data rate is exactly twice the input data rate. The 1-D DWT decomposes the input signal $x(t)$ in terms of wavelets and scaling functions as:

$$x(t) = \sum_{n=-\infty}^{\infty} c(n)\phi(t - n) + \sum_{j=0}^{\infty} \sum_{n=-\infty}^{\infty} d(j,n)2^{j/2}\psi(2^j t - n) \tag{1}$$

The scaling coefficients $c(n)$ and wavelet coefficients $d(j,n)$ are computed via the inner products:

$$c(n) = \int_{-\infty}^{\infty} x(t)\phi(t - n)dt \tag{2}$$

$$d(j,n) = 2^{j/2} \int_{-\infty}^{\infty} x(t)\psi(2^j t - n)dt \tag{3}$$

CWT is considered as the above mentioned equations with complex-valued wavelet and scaling functions:

$$\psi_c(t) = \psi_r(t) + j\psi_i(t) \tag{4}$$

$$\phi_c(t) = \phi_r(t) + j\phi_i(t). \tag{5}$$

The 2-D DT-CWT gives rise to wavelets in six different directions of $\pm 15°$, $\pm 45°$ and $\pm 75°$. There are two wavelets in each direction. One can be considered as the real part of a complex-valued wavelet and the other as the imaginary part. In fact, a DT-CWT of an image produces two real-valued (imaginary-valued) low-frequency subbands and six real-valued (imaginary-valued) high-frequency subbands. Because the complex version has twice as many wavelets as the real transform, it is 4-times expansive. The complex 2-D dual-tree is implemented as four critically-sampled separable 2-D DWTs operating in parallel. However, different filter sets are used along the rows and columns. As in the real case, the sum and difference of subbands are used to obtain the oriented wavelets.

3. **Proposed Interpolation Algorithm.** In recent years, research interest in image interpolation has been more focused on visual quality of the interpolated images. According to a study carried out by van Quwerkerk [2], the algorithms which are based on neural networks are capable of producing images without the artifacts described before. It seems that neural networks make a promising field in image interpolation research. In one of
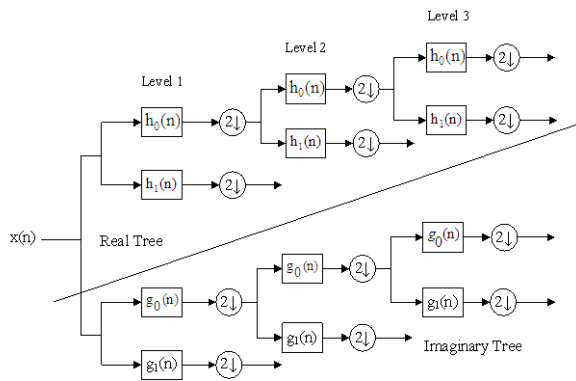
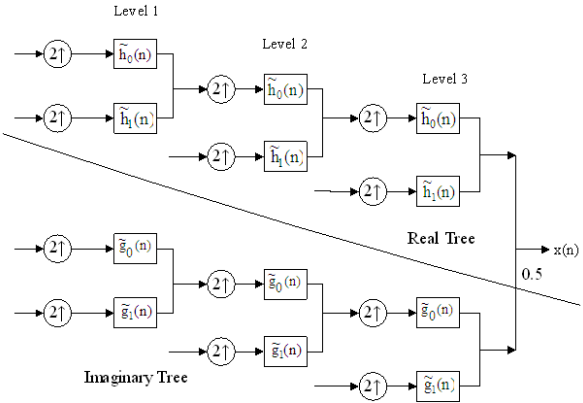FIGURE 1. Analysis filter bank for 1-D dual-tree complex wavelet transform

FIGURE 2. Synthesis filter bank for 1-D dual-tree complex wavelet transform

the earliest studies by Plaziac in 1999 [21], MLP neural networks are used for this purpose. There is no classification in this method. In almost all other algorithms, researchers employ some kinds of classification methods, which consider local structures of images. The purpose of data clustering is to partition the LR image neighborhoods into a finite number of clusters which are similar in some sense. This is especially of high importance in texture and edge areas, because of the human visual system sensitivity to the high contrast regions rather than smooth areas. In [32], the division of signal space to local regions is implemented via a vector quantization (VQ) algorithm. The VQ is accomplished using Kohonen's self-organizing map. In [22], in a pre-processing step the image blocks are classified into a number of classes by Adaptive Dynamic Range Coding (ADRC) [33]. The optimal coefficients of the neural interpolating filters are then obtained for each class by a supervised learning process which uses original HR and downscaled images as the training set. In [26], the neural networks do not accomplish the interpolation process directly. In fact, the task of MLP is to approximate the DWT subbands corresponding to the HR image. In this algorithm, the HR image is decomposed into vertical, horizontal and diagonal subbands and a separate neural network is used for estimating each subband. In the algorithm presented in [23], a fuzzy decision system inspired by human visual system is proposed to classify the input image into human perception nonsensitive regions and sensitive regions to select either the bilinear or the proposed neural-network interpolation module to operate in each region. It is observed that the accuracy of the interpolation process is greatly dependent on the result of the classification process. Neural networks generally present better performance when the solution space is reduced.

In this paper, DT-CWT with its high directional selectivity is used to decompose the HR image into directional subbands. The estimation process of edges in the corresponding directions is carried out by a specialized MLP estimator. Since high contrast regions such as edge structures are sensitive to human eye, the proposed algorithms can significantly enhance the edge structures and improve the quality of interpolated images. Unlike DWT, CWT coefficients do not have oscillatory behavior. This property simplifies the estimation of directional subbands. Shift invariance property of CWT coefficients also improves the estimation, since MLP estimators do not have to cope with different wavelet coefficients produced by shifted singularities. Furthermore, reduced aliasing in CWT reduces artifacts in the reconstructed signal.

The neural networks are initially trained to obtain optimal coefficients for each of the subband estimators. These coefficients are then used to estimate the wavelet subbands

corresponding to the HR image using its LR version. The HR image is then reconstructed by IDT-CWT.

3.1. **Training procedure.** The block diagram of the training procedure is shown in Figure 3. The down-scaled version of the training set is considered as the input to the neural networks. 16 subbands resulting from DT-CWT of the original HR training set are considered as targets. We employ a separate MLP estimator for each subband. The input vectors are the overlapping blocks of size $5 \times 5$ in the LR input and the central pixel of the corresponding block in each wavelet subband is considered as the target. As the size of the blocks is increased more information is provided around the approximation location. On the other hand, the size of the blocks determines the number of input elements and therefore the larger block size leads to higher network complexity. Therefore, we consider blocks of size $5 \times 5$ as a compromise between complexity and performance. Note that each wavelet subband of the HR image has the same size of the LR input. Here we consider one separate subband estimator for estimating each of the real or imaginary parts of the complex-valued subbands. The final weights for each estimator are obtained using resilient BP (RP) and Levenberg-Marquardt (LM) training algorithms beginning by random weight matrices. The neural networks characteristics are shown in Table 1. The hidden layers' number of neurons was selected heuristically. Different numbers of neurons in hidden layers have been tested and the numbers with the best performance are selected. The memory requirements for RP training algorithm are relatively small in comparison with the other algorithm. Although memory requirements for LM are larger, the mean square error is much lower than that achieved by RP algorithm.

3.2. **Interpolation process.** The block diagram of the interpolation process is shown in Figure 4. MLP subband estimators estimate 16 wavelet subbands corresponding to the HR image from the LR input which are all of the same size equal to the LR image. The HR image is then reconstructed by IDT-CWT. The structure of a typical subband
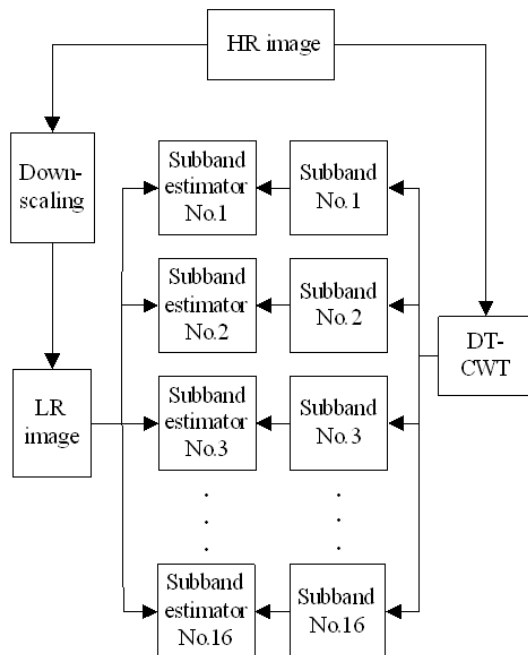


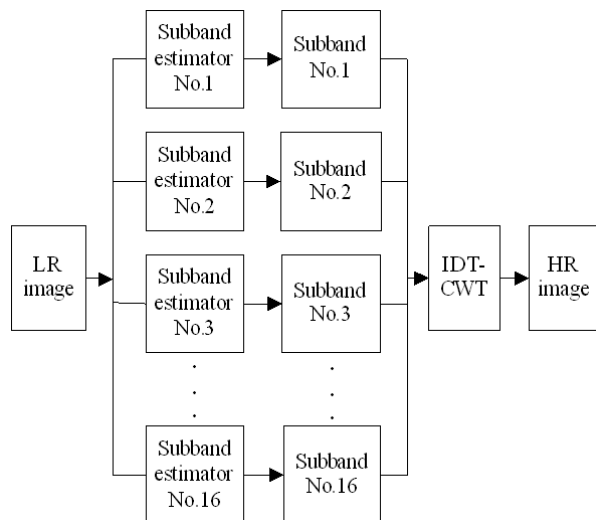FIGURE 3. Block diagram of the training procedure

FIGURE 4. Block diagram of the proposed interpolation process

TABLE 1. Neural network characteristics

| | |
|---|---|
| Number of input elements | 25 |
| Number of output neurons | 1 |
| Number of hidden layers | 2 |
| Number of neurons in first hidden layer | 22 |
| Number of neurons in second hidden layer | 12 |
| Transfer function of the first hidden layer | Hyperbolic tangent |
| Transfer function of the second hidden layer | Hyperbolic tangent |
| Transfer function of the output layer | Linear |
| Type of network | Feedforward |
| Network training algorithm | - Resilient BP<br>- Levenberg-Marquardt |
| Number of epochs | 100 (RP)<br>50 (LM) |
| Performance function | Mean squared error |
| Max error | $7 \times 10^{-4}$ (RP)<br>$4 \times 10^{-4}$ (LM) |
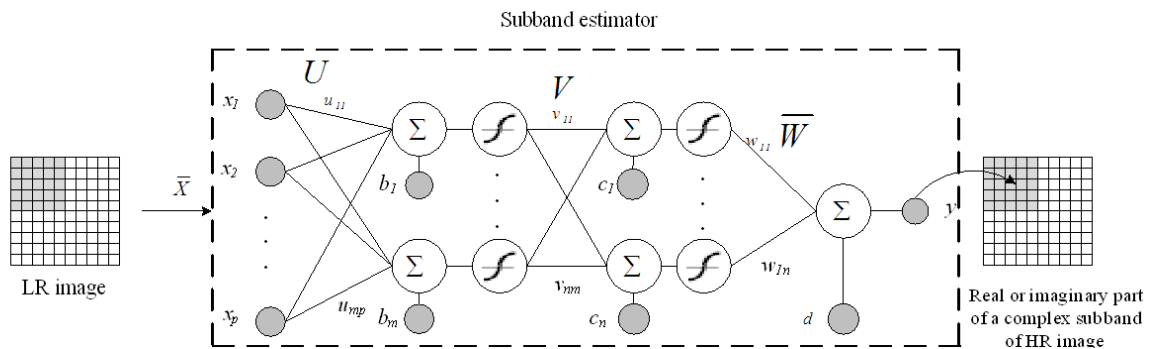


FIGURE 5. The structure of a typical subband neural estimator

estimator is shown in Figure 5. The output of the first hidden layer $\bar{O}^{-2}$ is:

$$\bar{O}^{-2} = \tanh(U.\bar{X} + \bar{B}). \tag{6}$$

where $\bar{X}$ is the input vector, $U$ is the weight matrix of the first hidden layer and $\bar{B}$ is the bias vector. The output of the second hidden layer $\bar{O}^{-1}$ is:

$$\bar{O}^{-1} = \tanh(V.\bar{O}^{-2} + \bar{C}). \tag{7}$$

where $V$ is the weight matrix of the second hidden layer and $\bar{C}$ is the bias vector of the second hidden layer. The estimated pixel value of the subband image $y$ is:

$$y = \overline{W}.\bar{O}^{-1} + d \tag{8}$$

where $\overline{W}$ the weight matrix of the output layer and $d$ is the bias of the output neuron.

Boundary pixels that cannot be estimated by MLP subband estimators are set to zero in high-frequency subbands, and are replaced from the LR image in low-frequency subbands so that the size of output image is exactly twice of the input.

Clearly this wavelet-based scheme supports a progressive interpolation. We can reconstruct the interpolated HR image by estimating the low-frequency subbands. Adding one or more high-frequency subbands can enhance the quality of the interpolated image.
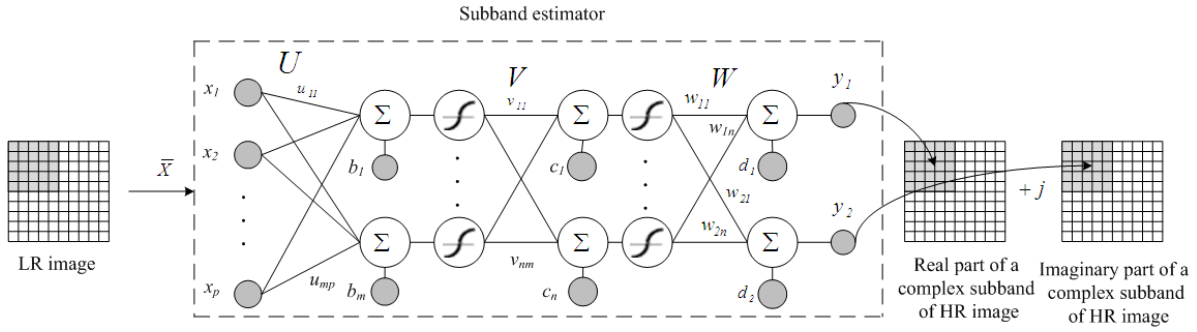
FIGURE 6. The structure of a typical subband neural estimator in the simplified algorithm

4. **Simplified Algorithm.** To reduce the structural and computational complexity of the proposed scheme, especially in the final interpolation step, we can consider 8 subband estimators instead of 16. The block diagram of the simplified training procedure is shown in Figure 7. The input is the down-scaled version of the training set as before and 8 complex-valued (with a real and an imaginary part) subbands resulting from DT-CWT of the original HR training set are considered as targets. The input vectors are $5 \times 5$ blocks in the LR input and the central pixels of the corresponding blocks in each complex-valued subband are the targets. To avoid any complex arithmetic, we consider each complex-valued subband consisting of two real and imaginary parts. Therefore, we have two central pixels with real values and we have 2-element target vectors. The optimal weights for each estimator are approximated by LM training algorithm. The other characteristics of the neural networks are as before. Block diagram of the simplified interpolation process and the structure of a typical subband estimator are shown in Figure 6 and Figure 8, respectively. The estimated pixel value of a complex-valued subband image $\bar{Y}$ is obtained by the following equations.

$$\bar{O}^{-2} = \tanh(U.\bar{X} + \bar{B}) \tag{9}$$

$$\bar{O}^{-1} = \tanh(V.\bar{O}^{-2} + \bar{C}) \tag{10}$$

$$\bar{Y} = W.\bar{O}^{-1} + \bar{D} \tag{11}$$

5. **Experimental Results.** The training set of 5 different gray-scale images including *Houses, Elaine, Couple, Boat* and *Bridge* of size $512 \times 512$ pixels is used for training. These images include a variety of natural images such as people, building and landscape. *Houses* and *Boat* have edges in different directions, *Bridge* has mainly horizontal edges, *Couple* has the various types of image components and *Elaine* is a portrait image. The number of training images is selected by experimentation to achieve better general performance. 321,552 samples are used for training of each subband estimator.

As shown in [30], for an acceptable and fair comparison of image interpolation algorithms, the down-scaling techniques, image contents and the quality metrics all need to be considered. So, in the following experiments we use three different down-scaled versions of HR images, named LR1, LR2 and LR3, two objective quality metrics along with subjective quality assessment. Also test images with different internal structures are employed; namely *Lena, Peppers, Baboon, Goldhill, Lighthouse* and *F*-16.

The first down-scaled version of the training set, LR1, is generated by lowpass filtering with the averaging filter of size $2 \times 2$ and then down-sampling of the HR image. The second version, LR2, is made by directly down-sampling of the HR training set, and the third, LR3 is generated through MATLAB *imresize* command. In fact LR3 is generated
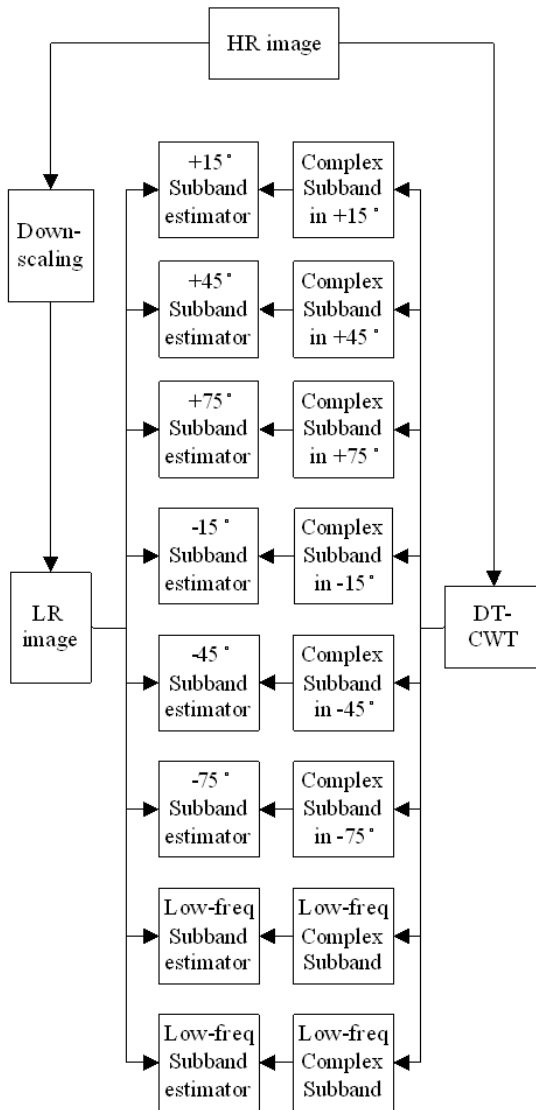
FIGURE 7. Block diagram of the simplified training procedure
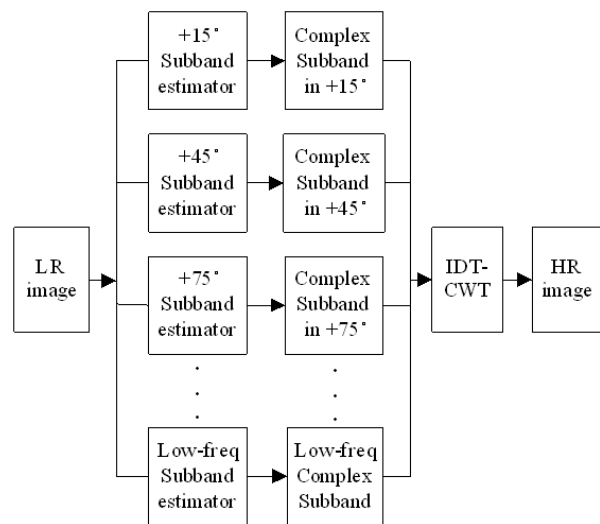


FIGURE 8. Block diagram of the simplified interpolation process

from decimated and anti-aliasing filtered HR image. For objective quality assessments we use Peak Signal to Noise Ratio (PSNR) and Mean Structural SIMilarity index (MSSIM). Structural Similarity (SSIM) index [31] is more consistent to perceived visual quality. The test images are with the original size of $512 \times 512$ and are not included in the training set to show the generalization ability.

We compare the performance of the proposed algorithms with that of bilinear and bicubic as two conventional interpolation algorithms and recently proposed methods named WMI [26], NEDI [13] and DFDF [14]. CWRPI, CWLMI and SCWI algorithms represent the proposed methods with 16 Complex Wavelet subband estimators and RP training Interpolation algorithm, 16 subband estimators and LM training and the simplified one with 8 subband estimators, respectively.

5.1. **Peak signal to noise ratio.** PSNR is defined in Equations (12) and (13), where $M$ and $N$ are numbers of rows and columns in an image and $y_{orig}$, $y_{inpl}$ are pixel intensities

of the original and interpolated image respectively. PSNR is widely used in image interpolation as a quantitative measurement of image quality. Greater PSNR corresponds to better quality of an interpolated image. The PSNR results for the test images are shown in Tables 2-4. A comparison of different interpolation algorithms on LR1 down-sampled training images is also shown in Table 5. The greatest value for each test image is shown in bold.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}, \, (dB) \tag{12}$$

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (y_{orig}(i,j) - y_{inpl}(i,j))^2 \tag{13}$$

As it is seen in Tables 2-4, the proposed methods are superior to other interpolation algorithms, both conventional and the more recently proposed, by an average of 0.78-3.81, 0.11-2.7 and 1.43-4.1dB for LR1, LR2 and LR3 down-sampling methods, respectively. Most image interpolation algorithms achieve higher PSNR when the down-scaling method is LR1 or LR3. This is because in LR1 and LR3 we apply a lowpass/anti-aliasing filter as a part of down-sampling process to avoid aliasing, while in directly down-sampling

TABLE 2. PSNR for LR1 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 32.69 | 34.26 | 31.12 | 30.32 | 30.43 | 34.59 | **35.61** | 35.59 |
| Peppers | 32.72 | 33.84 | 30.53 | 30.68 | 30.68 | 33.89 | **35.13** | 35.08 |
| Baboon | 23.37 | 23.98 | 23.24 | 22.70 | 22.68 | 24.06 | 24.29 | **24.30** |
| Goldhill | 28.96 | 29.84 | 27.86 | 27.47 | 27.65 | 29.96 | 30.29 | **30.30** |
| Lighthouse | 25.22 | 26.08 | 24.97 | 23.73 | 24.19 | 26.08 | 26.67 | **26.76** |
| F-16 | 31.66 | 33.45 | 29.32 | 29.23 | 29.55 | 33.73 | 34.91 | **34.93** |
| Average | 29.10 | 30.24 | 27.84 | 27.35 | 27.53 | 30.38 | 31.15 | **31.16** |

TABLE 3. PSNR for LR2 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 30.15 | 30.07 | 30.96 | 33.77 | 33.81 | 33.27 | **34.15** | 34.02 |
| Peppers | 30.21 | 30.03 | 30.17 | 32.94 | 32.99 | 32.40 | **33.44** | 33.36 |
| Baboon | 22.04 | 21.68 | 22.66 | **23.08** | **23.08** | 22.89 | **23.08** | 23.06 |
| Goldhill | 27.26 | 27.06 | 27.43 | 28.33 | 28.71 | 28.60 | **28.79** | 28.78 |
| Lighthouse | 23.09 | 23.49 | 24.41 | 24.74 | **25.38** | 25.08 | 25.25 | 25.24 |
| F-16 | 29.27 | 29.31 | 29.26 | 32.17 | **33.22** | 32.33 | 33.17 | 33.17 |
| Average | 27.00 | 26.94 | 27.48 | 29.17 | 29.53 | 20.09 | **29.64** | 29.60 |

TABLE 4. PSNR for LR3 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 32.64 | 34.08 | 31.58 | 30.34 | 30.43 | 34.85 | **35.89** | 35.82 |
| Peppers | 32.66 | 33.68 | 30.82 | 30.66 | 30.67 | 33.83 | **35.26** | 35.21 |
| Baboon | 23.24 | 23.84 | 23.64 | 22.71 | 22.69 | 24.23 | 24.53 | **24.54** |
| Goldhill | 28.79 | 29.64 | 28.25 | 27.51 | 27.66 | 30.13 | **30.56** | 30.55 |
| Lighthouse | 25.01 | 25.87 | 25.59 | 23.88 | 24.18 | 26.43 | **27.37** | 27.13 |
| F-16 | 31.59 | 33.21 | 29.44 | 29.23 | 29.53 | 34.01 | **35.30** | 35.26 |
| Average | 28.98 | 30.05 | 28.22 | 27.38 | 27.52 | 30.58 | **31.48** | 31.41 |

TABLE 5. PSNR for LR1 down-sampled *training* images

|        | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|--------|----------|---------|-------|-------|-------|-------|---------|-------|
| Houses | 23.01 | 23.88 | 22.45 | 21.50 | 21.67 | 24.46 | **26.07** | 25.81 |
| Elaine | 32.49 | 33.09 | 29.88 | 31.09 | 31.16 | 33.13 | **33.48** | 33.47 |
| Couple | 28.70 | 29.68 | 28.08 | 27.29 | 27.44 | 29.94 | **30.71** | 30.59 |
| Boat   | 29.02 | 30.07 | 28.21 | 27.50 | 27.60 | 30.46 | **31.54** | 31.43 |
| Bridge | 25.93 | 26.75 | 25.31 | 24.70 | 24.87 | 26.93 | **27.50** | 27.45 |
| Average | 27.83 | 28.69 | 26.78 | 26.41 | 26.54 | 28.98 | **29.86** | 29.75 |

the Nyquist condition for proper sampling may not be satisfied. However, it can be seen that NEDI and DFDF algorithms perform better when the images are directly down-sampled (LR2). Predictably, the LM training performs better than RP. However, even the simpler CWRPI is superior to existing methods in most situations; especially when more applicable LR1 and LR3 down-sampling are used. Finally, the proposed CWLMI and SCWI methods perform almost equally which means that we can employ SCWI with a significant lower complexity without noticeable performance reduction.

5.2. **Structural similarity index (SSIM).** The most widely used interpolation quality metric is the mean squared error (MSE), computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of peak signal to noise ratio (PSNR). These are appealing because they are simple and mathematically convenient in the context of optimization and have clear physical meanings. However, they are not well matched to perceived visual quality [31]. In [31], Wang et al. proposed a quality assessment metric, named SSIM, which is based on degradation of structural information. Here we use this metric as a supporting quality criterion to evaluate the interpolation techniques of this study.

Suppose that $X$ and $Y$ are two image signals. The default local weighting for computing statistics is a Gaussian weighting function of size $11 \times 11$ with standard deviation of 1.5. So the local statistics $\mu_x$, $\sigma_x$ and $\sigma_{xy}$ are defined as:

$$\mu_x = \sum_{i=1}^{N} w_i x_i \tag{14}$$

$$\sigma_x = \left( \sum_{i=1}^{N} w_i (x_i - \mu_x)^2 \right)^{1/2} \tag{15}$$

$$\sigma_{xy} = \sum_{i=1}^{N} w_i (x_i - \mu_x)(y_i - \mu_y) \tag{16}$$

The SSIM is defined as:

$$SSIM(X,Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{17}$$

where $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$. $L$ is the dynamic range of pixel values and $K_1 = 0.01$ and $K_2 = 0.03$. The overall image quality is evaluated by mean *SSIM (MSSIM)* as:

$$MSSIM(X,Y) = \frac{1}{M} \sum_{j=1}^{M} SSIM(x_j, y_j) \tag{18}$$

where $X$ and $Y$ are the reference and distorted images, respectively. $x_j$ and $y_j$ are the image contents at the $j$th local window and $M$ is the number of local windows in the image.

The value of MSSIM is between 0 and 1 and the higher value indicates the higher quality. The MSSIM values on different test images are shown in Tables 6-8. The greatest value for each test image is shown in bold. A comparison of different interpolation algorithms on LR1 down-sampled training images is also shown in Table 9.

TABLE 6. MSSIM for LR1 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 0.89 | 0.90 | 0.83 | 0.85 | 0.86 | 0.90 | **0.92** | 0.91 |
| Peppers | 0.86 | 0.87 | 0.81 | 0.84 | 0.85 | 0.87 | **0.88** | **0.88** |
| Baboon | 0.68 | 0.73 | 0.70 | 0.64 | 0.64 | 0.75 | **0.76** | 0.75 |
| Goldhill | 0.81 | 0.84 | 0.82 | 0.77 | 0.78 | 0.85 | **0.86** | **0.86** |
| Lighthouse | 0.76 | 0.79 | 0.73 | 0.72 | 0.72 | 0.79 | **0.81** | **0.81** |
| F-16 | 0.92 | **0.94** | 0.90 | 0.89 | 0.90 | 0.92 | **0.94** | **0.94** |
| Average | 0.82 | 0.845 | 0.798 | 0.785 | 0.791 | 0.846 | **0.861** | 0.858 |

TABLE 7. MSSIM for LR2 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 0.85 | 0.84 | 0.86 | 0.87 | 0.89 | 0.89 | **0.90** | 0.89 |
| Peppers | 0.82 | 0.81 | 0.83 | 0.86 | **0.86** | 0.85 | **0.86** | **0.86** |
| Baboon | 0.61 | 0.62 | 0.68 | **0.71** | **0.71** | 0.69 | **0.71** | 0.70 |
| Goldhill | 0.76 | 0.76 | 0.80 | 0.80 | 0.81 | 0.81 | **0.82** | **0.82** |
| Lighthouse | 0.71 | 0.71 | 0.75 | **0.78** | **0.78** | 0.77 | **0.78** | 0.77 |
| F-16 | 0.89 | 0.89 | 0.91 | 0.92 | **0.93** | 0.92 | **0.93** | 0.92 |
| Average | 0.773 | 0.771 | 0.821 | 0.805 | 0.830 | 0.821 | **0.833** | 0.826 |

TABLE 8. MSSIM for LR3 down-sampled test images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Lena | 0.89 | 0.90 | 0.87 | 0.85 | 0.86 | 0.89 | **0.92** | 0.91 |
| Peppers | 0.86 | 0.87 | 0.84 | 0.84 | 0.84 | 0.87 | **0.88** | **0.88** |
| Baboon | 0.67 | 0.72 | 0.72 | 0.63 | 0.63 | 0.74 | **0.76** | **0.76** |
| Goldhill | 0.80 | 0.83 | 0.83 | 0.76 | 0.77 | 0.85 | **0.86** | **0.86** |
| Lighthouse | 0.74 | 0.77 | 0.78 | 0.71 | 0.71 | 0.77 | **0.83** | 0.82 |
| F-16 | 0.91 | 0.93 | 0.90 | 0.88 | 0.89 | 0.93 | **0.94** | **0.94** |
| Average | 0.811 | 0.836 | 0.823 | 0.778 | 0.783 | 0.841 | **0.865** | 0.861 |

TABLE 9. MSSIM for LR1 down-sampled training images

|  | Bilinear | Bicubic | WMI | NEDI | DFDF | CWRPI | CWLMI | SCWI |
|---|---|---|---|---|---|---|---|---|
| Houses | 0.75 | 0.77 | 0.71 | 0.70 | 0.70 | 0.77 | **0.83** | 0.82 |
| Elaine | 0.77 | 0.79 | 0.76 | 0.75 | 0.76 | 0.79 | **0.80** | **0.80** |
| Couple | 0.83 | 0.85 | 0.82 | 0.79 | 0.80 | 0.86 | **0.87** | **0.87** |
| Boat | 0.83 | 0.85 | 0.80 | 0.79 | 0.80 | 0.86 | **0.87** | 0.86 |
| Bridge | 0.75 | 0.79 | 0.74 | 0.70 | 0.71 | 0.81 | **0.83** | **0.83** |
| Average | 0.786 | 0.810 | 0.766 | 0.746 | 0.754 | 0.818 | **0.840** | 0.836 |

FIGURE 9. (a) Original *Lena* image. Resolution enhanced *Lena* image by (b) Bilinear, (c) Bicubic, (d) WMI, (e) NEDI, (f) DFDF, (g) CWLMI and (h) SCWI.
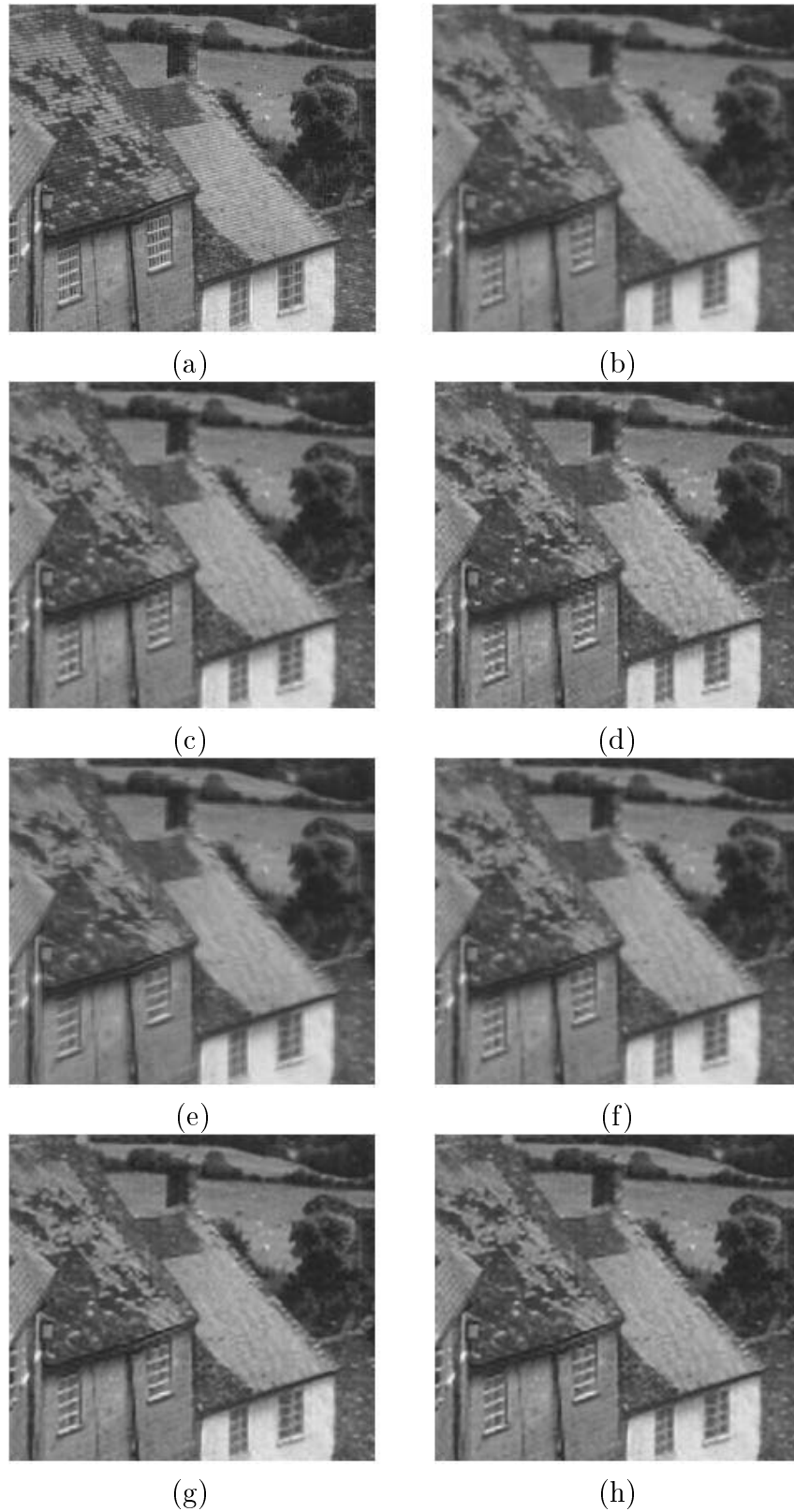
FIGURE 10. (a) Original *Goldhill* image. Resolution enhanced *Goldhill* image by (b) Bilinear, (c) Bicubic, (d) WMI, (e) NEDI, (f) DFDF, (g) CWLMI and (h) SCWI.
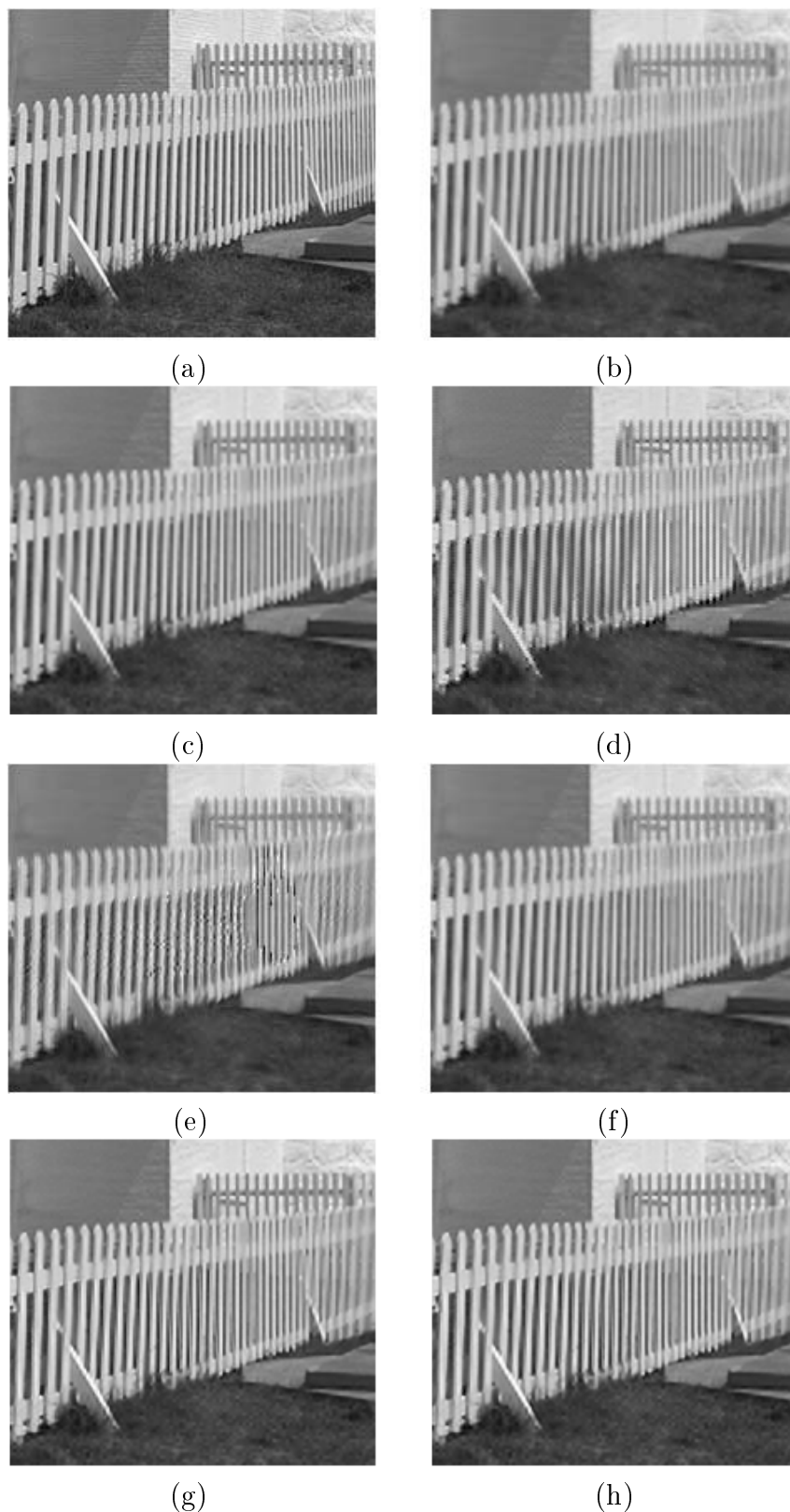
FIGURE 11. (a) Original *Lighthouse* image. Resolution enhanced *Lighthouse* image by (b) Bilinear, (c) Bicubic, (d) WMI, (e) NEDI, (f) DFDF, (g) CWLMI and (h) SCWI.
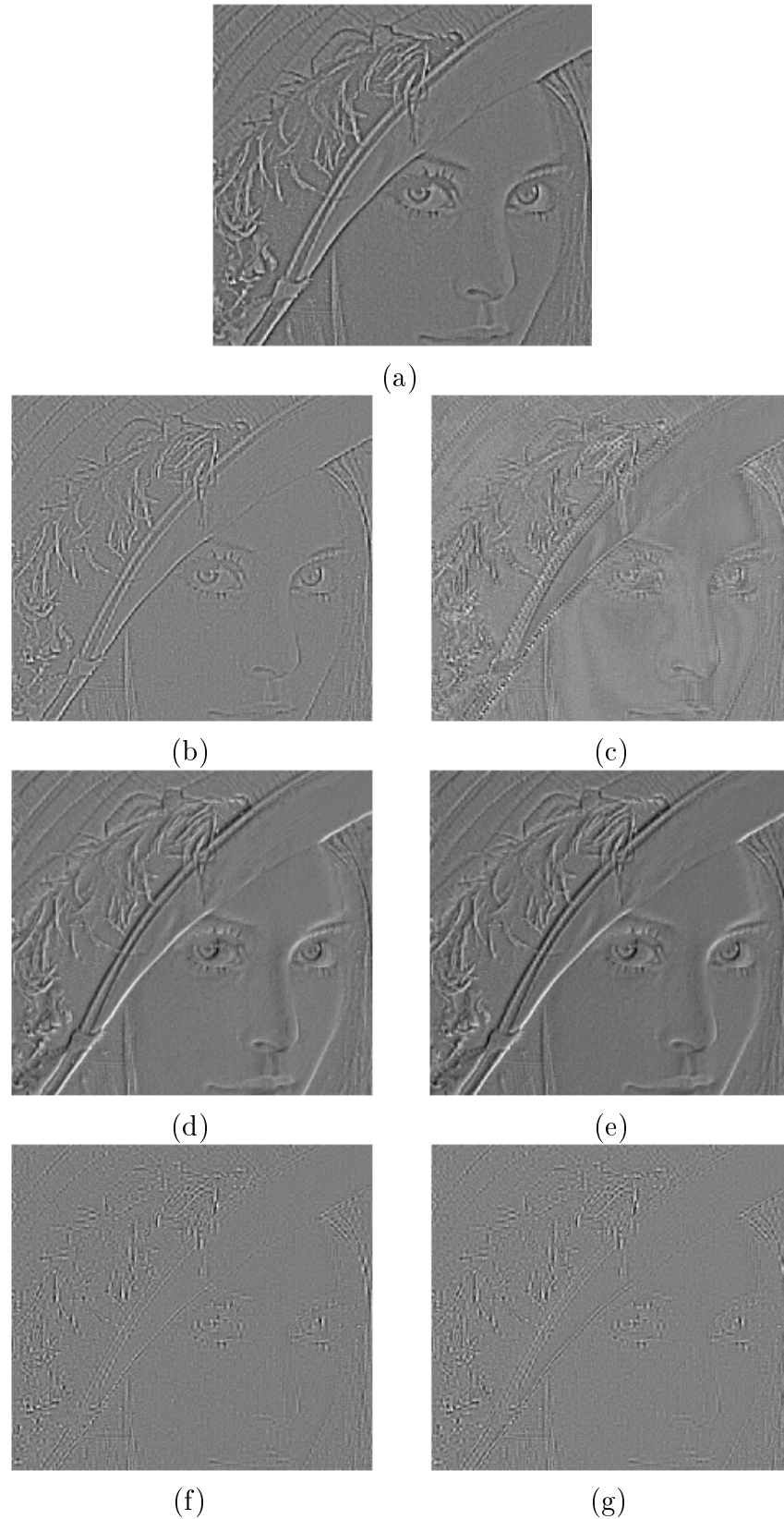
FIGURE 12. Difference images between the original *Lena* image and the resolution enhanced *Lena* image by (a) Bilinear, (b) Bicubic, (c) WMI, (d) NEDI, (e) DFDF, (f) CWLMI and (g) SCWI.

MSSIM results also show the superiority of the proposed algorithms by an average of 0.016-0.076, 0.003-0.062 and 0.029-0.087 for LR1, LR2 and LR3 down-sampled test images, respectively. Again, we see that CWLMI and SCWI perform almost equally.

5.3. **Subjective comparison.** In order to compare the performance of the proposed methods subjectively, portions of *Lena, Goldhill* and *Lighthouse* magnified by a factor of 2 are shown in Figures 9-11. It is obvious that the proposed methods are able to reconstruct the edges most sharply and produce less artifacts, leading to a better image quality. The images produced by bilinear and bicubic interpolations are blurred and WMI images suffer from blocking artifacts. In some cases, NEDI is not able to produce the interpolated image properly, as in *Lighthouse* image. DFDF algorithm suffers from blocking artifacts, as in the edge of the hat in *Lena* image. Figure 12 displays the difference images between the original *Lena* and the corresponding resolution enhanced images. The gray pixels indicate the error free area, while the darker or lighter areas indicate the interpolation errors. The proposed algorithms reduce the errors significantly and are able to reconstruct the original image nearly completely. The performance of the proposed interpolation algorithm is especially of considerable value in high contrast edge areas where the other interpolation algorithms fail to reconstruct the original image.

A comparison of computation time for generating a $512\times512$ HR image from a $256\times256$ LR input for different image interpolation algorithms is shown in Table 10. The programs are implemented in MATLAB. The simulations are done on a system with 2.5-GHz Intel Core 2 Duo CPU and 4-GB RAM and Windows Vista operating system.

TABLE 10. Comparison of computation time for different image interpolation algorithms

|           | WMI | NEDI | DFDF | CWRP(LM)I | SCWI |
|-----------|-----|------|------|-----------|------|
| Time (sec) | 6   | 24   | 16   | 28        | 14   |

6. **Conclusions.** We proposed hybrid image interpolation algorithms using DT-CWT and MLP neural networks. MLP subband estimators estimate the expected wavelet subbands corresponding to the HR image. The HR image is then reconstructed using the IDT-CWT. To reduce the computational complexity, a simplified version of the interpolation algorithm was presented. The simplified version has half of the complexity of the original algorithm but no significant performance reduction is observed. Taking advantage of CWT such as shift invariance, substantial reduced aliasing and directional selectivity the subband estimation process is most effective and leads to nearly accurate reconstruction of texture and edge areas. The experimental results on various images and down-scaling methods using different image quality assessments verify the superiority of the proposed algorithms.

**REFERENCES**

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd Edition, Prentice Hall, NJ, 2008.
[2] J. D. van Ouwerkerk, Image super-resolution survey, *Image Vision Comput.*, vol.24, no.10, pp.1039-1052, 2006.
[3] M. Unser, Splines: A perfect fit for signal and image processing, *IEEE Signal Process. Mag.*, no.11, pp.22-38, 1999.
[4] H. Demirel and G. Anbarjafari, Satellite image resolution enhancement using complex wavelet transform, *IEEE Geosci. Remote Sens. Lett.*, vol.7, no.1, pp.123-125, 2010.
[5] T. Celik and T. Tjahjadi, Image resolution enhancement using dual-tree complex wavelet transform, *IEEE Geosci. Remote Sens. Lett.*, vol.7, no.3, pp.554-557, 2010.

[6] E. Meijering, A chronology of interpolation: From ancient astronomy to modern signal and image processing, *Proc. of IEEE*, vol.90, no.3, pp.319-342, 2002.

[7] T. M. Lehmann, C. Gonner and K. Spitzer, Survey: Interpolation methods in medical image processing, *IEEE Trans. Med. Imaging*, vol.18, no.11, pp.1049-1075, 1999.

[8] W. L. Lee, C. C. Yang and H. T. Wu, Wavelet-based interpolation scheme for resolution enhancement of medical images, *J. Signal Process. Sys.*, vol.55, pp.251-265, 2009.

[9] J. Xiao, X. Zou, Z. Liu and X. Guo, A novel adaptive interpolation algorithm for image resizing, *International Journal of Innovative Computing, Information and Control*, vol.3, no.6(A), pp.1335-1345, 2007.

[10] Q. Wang and R. K. Ward, A new orientation-adaptive interpolation method, *IEEE Trans. Image Process.*, vol.16, no.4, pp.889-900, 2007.

[11] Y. Zhang, S. Gao, C. Zhang and J. Chi, Application of a bivariate rational interpolation in image zooming, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(B), pp.4299-4307, 2009.

[12] S. G. Chang, Z. Cvetkovic and M. Vetterli, Locally adaptive wavelet-based image interpolation, *IEEE Trans. Image Process.*, vol.15, no.6, pp.1471-1485, 2006.

[13] X. Li and M. T. Orchard, New edge-directed interpolation, *IEEE Trans. Image Process.*, vol.10, no.10, pp.1521-1527, 2001.

[14] L. Zhang and X. Wu, An edge-guided image interpolation algorithm via directional filtering and data fusion, *IEEE Trans. Image Process.*, vol.15, no.8, pp.2226-2238, 2006.

[15] F. Malgouyres and F. Guichard, Edge direction preserving image zooming: A mathematical and numerical analysis, *SIAM J. Numer. Anal.*, vol.39, pp.1-37, 2001.

[16] D. D. Muresan, Fast edge directed polynomial interpolation, *Proc. of Int. Conf. on Image Process.*, vol.2, pp.990-993, 2005.

[17] S. Haykin, *Neural Networks*, Macmillan, New York, 1994.

[18] M. T. Hagan, H. B. Demuth and M. H. Beale, *Neural Network Design*, PWS Publishing Company, Boston, MA, 1996.

[19] S. Gao, J. Zhang, X. Wang and Z. Tang, Multi-layer neural network learning algorithm based on random pattern search method, *International Journal of Innovative Computing, Information and Control*, vol.5. no.2, pp.489-502, 2009.

[20] T. Su, J. Jhang and C. Hou, A hybrid artificial neural networks and particle swarm optimization for function approximation, *International Journal of Innovative Computing, Information and Control*, vol.4, no.9, pp.2363-2374, 2008.

[21] N. Plaziac, Image interpolation using neural networks, *IEEE Trans. Image Process.*, vol.8, no.11, pp.381-390, 1999.

[22] H. Hu, P. M. Hofman and G. de Haan, Image interpolation using classification-based neural networks, *Proc. of ISCE*, Reading, UK, pp.133-137, 2004.

[23] C. T. Lin, K. W. Fan, H. C. Pu, S. M. Lu and S. F. Liang, An HVS-directed neural-network-based image resolution enhancement scheme for image resizing, *IEEE Trans. Fuzzy Syst.*, vol.15, no.4, pp.605-615, 2007.

[24] A. Panagiotopoulou and V. Anastassopoulos, Scanned images resolution improvement using neural networks, *Neural Comput. Appl.*, vol.17, pp.39-47, 2008.

[25] S. S. Kim, Y. S. Kim and I. K. Eom, Image interpolation using MLP neural network with phase compensation of wavelet coefficients, *Neural Comput. Appl.*, vol.18, pp.967-977, 2009.

[26] Y. L. Huang, Wavelet-based image interpolation using multilayer perceptron, *Neural Comput. Appl.*, vol.14, no.1, pp.1-10, 2005.

[27] N. G. Kingsbury, Image processing with complex wavelets, *Philos. Trans. R. Soc. London A, Math. Phys. Sci.*, vol.357, no.1760, pp.2543-2560, 1999.

[28] I. W. Selesnick, R. G. Baraniuk and N. G. Kingsbury, The dual tree complex wavelet transform, *IEEE Signal Process. Mag.*, vol.22, no.6, pp.123-151, 2005.

[29] A. Temizel and T. Vlachos, Wavelet domain image resolution enhancement using cycle-spinning, *Electron. Lett.*, vol.41, no.3, pp.119-121, 2005.

[30] E. Dumic, S. Grgic and M. Grgic, The use of wavelets in image interpolation: Possibilities and limitations, *Radioengineering*, vol.16, no.4, pp.101-109, 2007.

[31] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Process.*, vol.13, no.4, pp.600-612, 2004.

[32] F. M. Candocia and J. C. Principe, Super-resolution of images based on local correlations, *IEEE Trans. Neural Networks*, vol.2, no.2, pp.372-380, 1999.

[33] T. Kondo and K. Kawaguchi, Adaptive dynamic range encoding by apparatus, *US-Patent 5,444,487*, 1995.