SELF-SAMPLED IMAGE RESOLUTION ENHANCEMENT USING DUAL-TREE COMPLEX WAVELET TRANSFORM

Turgay Celik[a] *and Huseyin Kusetogullari*[b]

[a]Faculty of Science, Department of Chemistry, National University of Singapore, Singapore
email: chmcelik@nus.edu.sg
[b]School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom
email: h.kusetogullari@warwick.ac.uk

ABSTRACT

In this paper, a dual-tree complex wavelet transform domain image resolution enhancement method is proposed. The method estimates detail wavelet coefficients for the input low-resolution (LR) image using different types of deformations on the initial estimate of high-resolution (HR) image. Edge preserving smoothing filtering with different parameters is used in deformations. Decomposition of each deformed HR image results in a different set of detail wavelet coefficients for LR image, and the resultant HR image is computed by averaging the different reconstructions from LR image using different detail wavelet coefficient sets. The perceptual and objective quality of resolution enhanced images compare favorably with recently emerged methods in the field.

1. INTRODUCTION

Discrete wavelet transform (DWT) based techniques have been widely used for performing image interpolation. A common assumption of these techniques 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 order to estimate those detail wavelet coefficients more sophisticated methods have been applied.

DWT extreme evolution was employed in [1]. Only coefficients with significant magnitudes are estimated as the evolution of the wavelet coefficients among the scales. The performance is mainly affected from the fact that the signs of estimated coefficients are copied directly from parent coefficients without any attempt being made to estimate the actual signs. This is contradictory to the fact that there is very little correlation between the signs of the parent coefficients and their descendants. As a result, the signs of the coefficients estimated using extreme evolution techniques cannot be relied upon. Hidden Markov Tree (HMT) based method in [2] models the unknown wavelet coefficients as belonging to mixed Gaussian distributions which are symmetrical around the zero mean. HMT models are used to find out the most probable state for the coefficients to be estimated. The performance is mainly suffered from the sign changes between the scales.

The decimated wavelet transform is not shift-invariant and, as a result, suppression of wavelet coefficients, such as quantization of coefficients during the compression process or non-exact estimation of high-frequency subband coefficients, introduces cyclostationarity into the image which manifests itself as ringing in the neighbourhood of discontinuities [3]. In order to combat this drawback in image interpolation, cycle-spinning methodology was adopted in [3]. The perceptual and objective quality of the resolution enhanced images by their method compare favorably with recently emerged methods in the field [1, 2].

Standard DWTs suffer from several disadvantages that undermine its usage in many other applications. The most important drawbacks of the standard DWTs are: 1) lack of shift invariance; 2) limited directionality when extended to higher dimensions; and 3) lack of phase information. One of the most promising decomposition that removes the above drawbacks satisfactorily is the dual-tree complex wavelet transform (CWT) [4]. Two classical wavelet trees (with real filters) are developed in parallel, with the wavelets forming (approximate) Hilbert pairs. One can then interpret the wavelets in the two trees of the CWT as the real and imaginary parts of some complex wavelet.

In this paper, we employ the CWT for image resolution enhancement. We propose a sample based wavelet coefficient estimation technique in the CWT domain. The wavelet coefficients of the LR image is estimated from the CWT decomposition of deformed initial estimate of the HR image. The initial estimate of the HR image is reconstructed from the LR image and zero padding of wavelet coefficients using the inverse CWT (ICWT). We also adopt the technique given in [3] to the CWT domain.

The paper is organized as follows. Section 2 gives a brief review of the CWT. Section 3 describes the proposed CWT domain image resolution enhancement algorithm. Section 4 provides some experimental results of the proposed approach and comparisons with the other approaches in the literature. Section 5 concludes the paper.

2. COMPLEX WAVELETS

The ordinary DWT is not shift invariant because of the decimation operation exploited in the transform. As a result, a small shift in the input signal can cause a very different set of output wavelet coefficients. For that, Kingsbury [4] introduced a new kind of wavelet transform, CWT, that exhibits approximate shift invariant property and improved directional resolution when compared that of the DWT.

The CWT also yields perfect reconstruction by using two parallel decimated trees with real coefficients. The one-dimensional (1-D) CWT decomposes a signal $f(x)$ in terms of a complex shifted and dilated mother wavelet $\psi(x)$ and

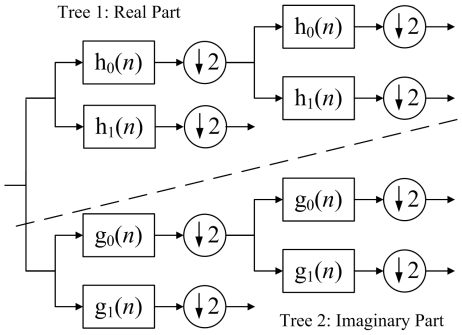


Figure 1: Two-level 1-D dual-tree complex wavelet transform (CWT).

scaling function $\phi(x)$, i.e.,

$$f(x) = \sum_{l \in Z} s_{j_0,l} \phi_{j_0,l}(x) + \sum_{j \geq j_0} \sum_{l \in Z} c_{j,l} \psi_{j,l}(x), \qquad (1)$$

where Z is the set of all integers, j and l refer to the shifts and dilations respectively, $s_{j_0,l}$ is the scaling coefficient and $c_{j,l}$ is the complex wavelet coefficient with $\phi_{j_0,l}(x) = \phi^r_{j_0,l}(x) + \sqrt{-1}\phi^i_{j_0,l}(x)$, and $\psi_{j,l}(x) = \psi^r_{j,l}(x) + \sqrt{-1}\psi^i_{j,l}(x)$. The complex wavelet transform is a combination of two real-valued wavelet transforms; in 1-D, the set $\{\phi^r_{j_0,l}, \phi^i_{j_0,l}, \psi^r_{j_0,l}, \psi^i_{j_0,l}\}$ forms a tight wavelet frame with two times of redundancy. The real and imaginary parts of the 1-D CWT are computed using separate filter banks with wavelet filters h_0 and h_1 for the real part, and g_0 and g_1 for the imaginary part, as illustrated in Figure 1 [4]. The outputs from the two trees in Figure 1 are interpreted as the real and imaginary parts of the complex coefficients.

Similar to the 1-D CWT, the two-dimensional (2-D) CWT decomposes an 2-D image $f(x,y)$ through a series of dilations and translations of a complex scaling function and six complex wavelet functions $\psi^\theta_{j,l}$, i.e.,

$$f(x,y) = \sum_{l \in Z^2} s_{j_0,l} \phi_{j_0,l}(x,y) + \sum_{\theta \in \Theta} \sum_{j \geq j_0} \sum_{l \in Z^2} c^\theta_{j,l} \psi^\theta_{j,l}(x,y). \quad (2)$$

where $\theta \in \Theta = \{\pm 15^\circ, \pm 45^\circ, \pm 75^\circ\}$ refers to the *directionality* of the complex wavelet function. The impulse response of six complex wavelets associated with the 2-D complex wavelet transform is illustrated in Figure 2. The complex wavelet transform can discriminate between features at positive and negative frequencies. Hence, there are six subbands capturing features along the lines with angles of $\theta \in \Theta$, respectively.

3. THE PROPOSED METHOD

Let us assume that the one-level CWT decomposition of an $W \times H$ image $\mathbf{I}$ results in a matrix of $\text{CWT}(\mathbf{I}) = [\mathbf{LP_I}\ \mathbf{HP_I}]$, and inverse CWT (ICWT) of $[\mathbf{LP_I}\ \mathbf{HP_I}]$ reconstructs the signal $\mathbf{I}$, i.e., $\text{ICWT}([\mathbf{LP_I}\ \mathbf{HP_I}]) = \mathbf{I}$. $\mathbf{LP_I}$ is a matrix of size $\frac{W}{2} \times \frac{H}{2}$ which is the complex-valued low-pass subband resulting from the one-level CWT decomposition of image $\mathbf{I}$, and $\mathbf{HP_I}$ is a matrix of size $\frac{W}{2} \times \frac{H}{2} \times 6$ which is the collection of all complex-valued high-pass subbands resulting from the one-level CWT decomposition of image $\mathbf{I}$. The simplified

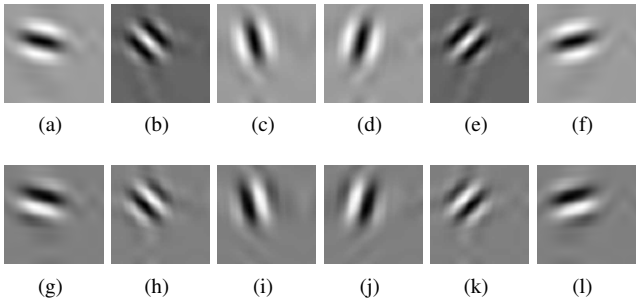


Figure 2: The real (R) and imaginary (I) parts of the impulse responses of the 2-D CWT filters for the 6 directional subbands: (a) R_{-15°; (b) R_{-45°; (c) R_{-75°; (d) R_{+75°; (e) R_{+45°; (f) R_{+15°; (g) I_{-15°; (h) I_{-45°; (i) I_{-75°; (j) I_{+75°; (k) I_{+45°; (l) I_{+15°.

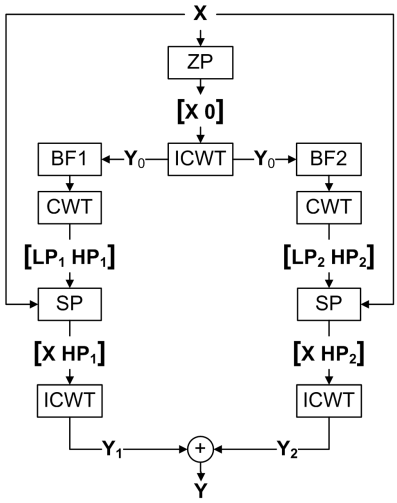


Figure 3: Block diagram of proposed method where **BF** is Bilateral Filtering, **ZP** is Zero Padding, and **SP** is Subband Padding.

block diagram of the proposed method is shown in Figure 3 and explained as follows.

Using a given LR image $\mathbf{X}$ of size $W \times H$, initial approximation $\mathbf{Y}_0$ to the unknown HR image $\mathbf{Y}$ is reconstructed using zero padding of high-frequency subbands (i.e. setting all elements of these subbands to zeros) followed by ICWT, i.e.,

$$\mathbf{Y}_0 = \text{ICWT}([\mathbf{X}\ \mathbf{0}]). \qquad (3)$$

where $\mathbf{0}$ is an all-zero matrix of size $W \times H \times 6$. $\mathbf{Y}_0$ can be used to make a better estimations to high-pass coefficients in (3). Deformations to initial estimation $\mathbf{Y}_0$ are applied using edge preserving smoothing filtering (EPSF). Tomasi and Manduchi's bilateral filter [5] is the state of the art for edge-preserving smoothing. The filter combines grey levels or colours of pixels based on both their geometric closeness and their photometric similarity. The filter needs parameters for domain filtering (σ_d) and range filtering (σ_r). It also necessary to specify the bilateral filter half-width (w). Different parameters generate different filtering results. In this paper, we apply two realizations of EPSF, i.e., ($w = 5$, $\sigma_d = 3$, $\sigma_r = 0.1$) (BF1 in Figure 3), ($w = 5$, $\sigma_d = 5$, $\sigma_r = 0.1$) (BF2

(a) (b) (c) $\varepsilon_R = 4.38$

(d) (e) $\varepsilon_R = 2.69$

Figure 4: Extracts from original and reconstructed Lena images: (a) Original image; (b)-(c) Reconstructed and residual images using bilinear interpolation; and (d-e) Reconstructed and residual images using the proposed method.

in Figure 3) to generate two filtered versions of $\mathbf{Y}_0$. Different set of realizations can be applied. Then the CWT decomposition of the two filtered HR images results in $[\mathbf{LP}_1\ \mathbf{HP}_1]$ and $[\mathbf{LP}_2\ \mathbf{HP}_2]$. This brings a better approximation to high-pass coefficients than all-zero matrix in (3). Two HR images are generated, i.e.,

$$\begin{aligned} \mathbf{Y}_1 &= \text{ICWT}\left([\mathbf{X}\ \mathbf{HP}_1]\right), \\ \mathbf{Y}_2 &= \text{ICWT}\left([\mathbf{X}\ \mathbf{HP}_2]\right). \end{aligned} \tag{4}$$

Finally, the resultant HR image $\mathbf{Y}$ of size $2W \times 2H$ is the average of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, i.e.,

$$\mathbf{Y} = \frac{\mathbf{Y}_1 + \mathbf{Y}_2}{2}. \tag{5}$$

The residual image ($\mathbf{I}_E$) between the original HR image ($\mathbf{I}$) and reconstructed HR image ($\hat{\mathbf{I}}$) is defined as $\mathbf{I}_E = |\mathbf{I} - \hat{\mathbf{I}}|$. The quantitative measure of the reconstruction error (ε_R) is computed using the average error, i.e.,

$$\varepsilon_R = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \mathbf{I}_E\,(i,j) \tag{6}$$

Figure 3 shows HR image reconstruction together with the residual images when the bilinear interpolation (see Figures 3 (b) and (c)) and the proposed method (see Figures 3 (d) and (e)) are used. The residual image are inverted in intensity for display purposes. It is clear from the residual images and corresponding reconstruction errors that the performance of the proposed method is much better than the performance of the widely used bilinear interpolation type HR image reconstruction.

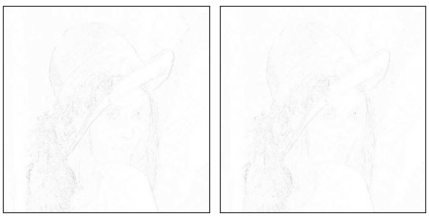
4. EXPERIMENTAL RESULTS

Experiments are conducted using six test images, as shown in Figure 5. In our simulations, the observed LR images are independently generated by convolving each ground-truth image with a point spread function (PSF), followed by a downsampling operation with a ratio 2 in both horizontal and vertical directions, respectively. The HR image, which is estimated version of ground-truth image, is reconstructed from the LR image.

(a) Bridge (b) Lena (c) Peppers (d) Barbara (e) Boat (f) Elaine

Figure 5: Test images used in experiments.

(a) $\varepsilon_R = 6.79$ (b) $\varepsilon_R = 6.63$ (c) $\varepsilon_R = 4.40$



(d) $\varepsilon_R = 2.77$ (e) $\varepsilon_R = 2.46$

Figure 6: Comparisons of the resultant residual images between ground-truth Lena image and reconstructed HR image from LR image which is constructed from the ground-truth Lena image using the nearest-neighbor interpolation as PSF: (a) The method of [7]; (b) The method of [8]; (c) The method of [3]; (d) The method of [3] using CWT; and (e) The proposed method.

To provide the objective performance comparisons between the ground-truth image $\mathbf{I}$ of size $H \times W$ and and the reconstructed image $\hat{\mathbf{I}}$, two image quality measurements are used in this paper: PSNR, i.e.,

$$\text{PSNR} = 10\log_{10} \frac{H \times W}{\sum_{i=1}^{H} \sum_{j=1}^{W} (\mathbf{I}(i,j) - \hat{\mathbf{I}}(i,j))^2}, \tag{7}$$

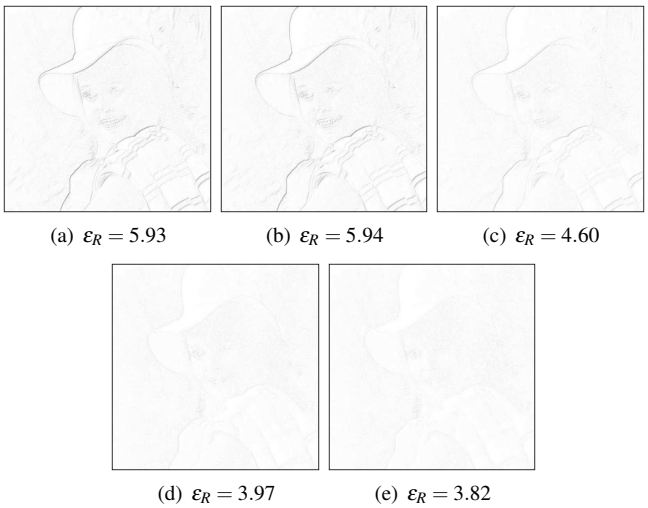
and universal image quality index (Q) [6]. The dynamic range of Q is $[-1, 1]$, the best value 1 is achieved if and only if $\mathbf{I} = \hat{\mathbf{I}}$, and the lowest value is -1. The universal image quality index models any distortion as a combination of three different factors: loss of correlation, luminance distortion, and contrast distortion [6].

Experiments are conducted to compare the performance of the proposed approach with that of recently proposed methods in [7], [8], and [3]. The implementations of the methods in [7] and [8] are received from the authors own implementations, on the other hand we implement the method in [3], and also we modify it to perform with the CWT instead of the DWT.

The performance of the image resolution enhancement

Table 1: The PSNR and Q performance comparisons when LR image is created from the ground-truth image using the nearest-neighbor interpolation as PSF.

	Method in [7]		Method in [8]		Method in [3]		Method [3] with CWT		Proposed Method	
Test Image	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q
Bridge	21.62	0.50	21.73	0.52	24.04	0.71	25.64	0.77	25.82	0.80
Lena	25.76	0.55	26.12	0.57	29.80	0.72	33.53	0.77	33.98	0.79
Peppers	26.39	0.55	26.55	0.55	29.79	0.65	32.65	0.68	32.90	0.70
Barbara	20.20	0.50	21.52	0.51	22.97	0.66	24.15	0.69	24.52	0.74
Boat	23.59	0.47	23.86	0.49	26.73	0.65	28.87	0.70	29.18	0.72
Elaine	27.47	0.51	27.55	0.52	29.66	0.60	31.56	0.64	31.57	0.65



(a) $\varepsilon_R = 5.93$ (b) $\varepsilon_R = 5.94$ (c) $\varepsilon_R = 4.60$

(d) $\varepsilon_R = 3.97$ (e) $\varepsilon_R = 3.82$

Figure 7: Comparisons of the resultant residual images between ground-truth Elaine image and reconstructed HR image from LR image which is constructed from the ground-truth Elaine image using average low-pass filter with a size of 3×3 as PSF: (a) The method of [7]; (b) The method of [8]; (c) The method of [3]; (d) The method of [3] using CWT; and (e) The proposed method.

method depends on the point spread function of the camera that maps HR image to LR image. Table 1 shows the performance of the different techniques when LR image is created from the ground-truth Lena image using the nearest-neighbor interpolation as the PSF. The reconstructed HR images from different techniques are compared with the ground-truth image. It is clear that the proposed method achieves the best performance. The performances of the methods are visualized in Figure 4 by showing the residual images between the ground-truth image and reconstructed HR image. It is clear that the minimum difference from the ground-truth image is achieved using the HR image resulted from the proposed method.

Table 2 shows the performance comparisons when the LR image is created by the PSF of an average low-pass filter with a size of 3×3. Similar to the Table 1, the performance of the proposed method is better than the other methods when PSF is a low-pass average filter. The resultant residual images between the ground-truth Elaine test image, and its reconstructed HR image for different methods are shown in Figure 4. It is clear from Figure 4 that, the numerical values shown in Table 1 are supported by the visual quality.

Table 2: The PSNR and Q performance comparisons when LR image is created from the ground-truth image using an average low-pass filter with a size of 3×3 as PSF.

	Method in [7]		Method in [8]		Method in [3]		Method [3] with CWT		Proposed Method	
Test Image	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q	PSNR (dB)	Q
Bridge	22.50	0.54	22.58	0.56	25.01	0.73	26.09	0.76	26.60	0.79
Lena	26.43	0.59	26.63	0.61	30.33	0.74	33.23	0.77	33.93	0.79
Peppers	26.99	0.59	27.09	0.60	30.56	0.70	33.11	0.72	33.75	0.73
Barbara	22.33	0.54	22.67	0.54	24.17	0.66	24.72	0.68	25.13	0.71
Boat	24.42	0.51	24.58	0.53	27.59	0.67	29.23	0.71	29.80	0.73
Elaine	28.32	0.57	28.39	0.57	31.06	0.66	32.71	0.67	33.00	0.69

Figure 8 presents various image results of the above five methods. As seen from the above Tables 1 and 2 and Figure 8, the proposed method always yields the best performance among the above-mentioned four methods.

5. CONCLUSION

A method for image resolution enhancement in the dual-tree complex wavelet domain is presented. Initial estimate of HR image is deformed using EPSF with different parameters. Deformed HR images are used to estimate complex wavelet coefficients for the input LR image. Estimated wavelet coefficients are used in parallel with the input LR image to estimate the resultant HR image. Intensive tests and comparisons with the techniques from the literature has shown the superiority of the proposed method.

REFERENCES

[1] W. K. Carey, D. B. Chuang, and S. S. Hemami, "Regularity-preserving image interpolation," *IEEE Trans. Image Proc.*, vol. 8, no. 9, pp. 1293–1297, 1999.

[2] Kentaro Kinebuchi, D. Darian Muresan, and Thomas W. Parks, "Image interpolation using wavelet-based hidden markov trees," in *Proc. IEEE Int. Conf. Acoust. Speech and Signal Process.*, 2001, pp. 1957–1960.

[3] A. Temizel and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning," *Electronics Letters*, vol. 41, no. 3, pp. 119–121, Feb 2005.

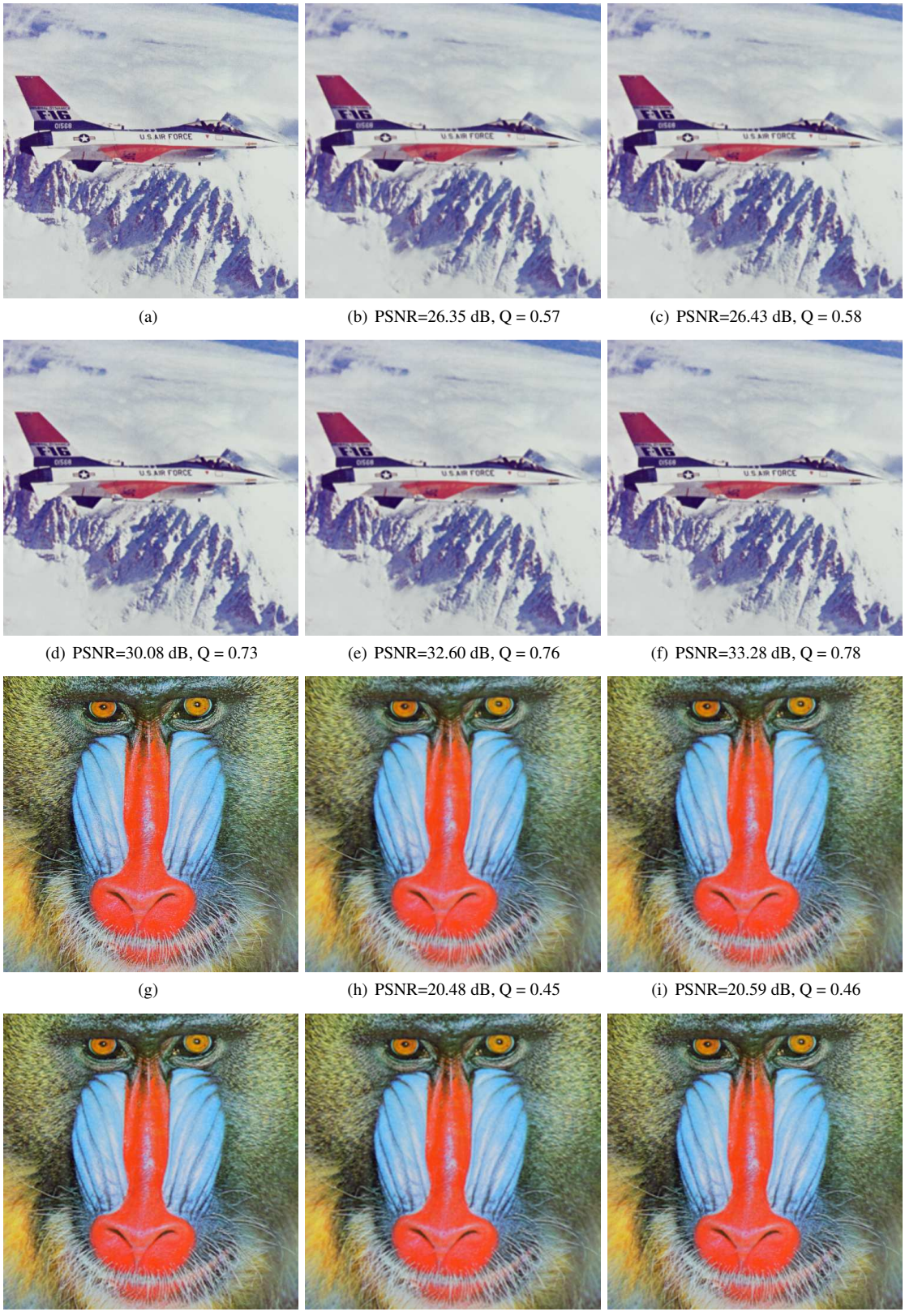
[4] Nick Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *J. Appl. Comput. Harmon. Anal.*, vol. 10, no. 3, pp. 234–253, 2001.

[5] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comp. Vis.*, 1998, pp. 839–846.

[6] Zhou Wang and A.C. Bovik, "A universal image quality index," *IEEE Sig. Proc. Letts.*, vol. 9, no. 3, pp. 81–84, Mar 2002.

[7] Xin Li and M.T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Proc.*, vol. 10, no. 10, pp. 1521–1527, Oct 2001.

[8] Lei Zhang and Xiaolin Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," *IEEE Trans. Image Proc.*, vol. 15, no. 8, pp. 2226–2238, Aug 2006.



(a) (b) PSNR=26.35 dB, Q = 0.57 (c) PSNR=26.43 dB, Q = 0.58

(d) PSNR=30.08 dB, Q = 0.73 (e) PSNR=32.60 dB, Q = 0.76 (f) PSNR=33.28 dB, Q = 0.78

(g) (h) PSNR=20.48 dB, Q = 0.45 (i) PSNR=20.59 dB, Q = 0.46

(j) PSNR=22.31 dB, Q = 0.63 (k) PSNR=22.86 dB, Q = 0.65 (l) PSNR=23.27 dB, Q = 0.68

Figure 8: The subjective and quantitative performance comparisons: (a,g) The original images; (b,h) The results of [7]; (c,i) The results of [8]; (d,j) The results of [3]; (e,k) The results of [3] using CWT; and (f,l) The results of the proposed method.