Super-Resolution Image Reconstruction Using Wavelet Based Patch and Discrete Wavelet Transform

Do Kyung Shin · Young Shik Moon

Abstract This paper proposes a novel method that combines the discrete wavelet transform (DWT) and example-based technique to reconstruct a high-resolution from a low-resolution image. Although previous interpolation- and example-based methods consider the reconstruction adaptive to edge directions, they still have a problem with aliasing and blurring effects around edges. In order to address these problems, in this paper, we utilize the frequency sub-bands of the DWT that has the feature of lossless compression. Our proposed method first extracts the frequency sub-bands (Low-Low, Low-High, High-Low, High-High) from an input low-resolution image by the DWT, and then the low-resolution image is inserted into the Low-Low sub-band. Since information in high-frequency sub-bands (Low-High, High-Low, and High-High) might be lost in the low-resolution image, they are reconstructed or estimated by using example-based method from image patch database. After that, we make a high-resolution image by performing the inverse DWT of reconstructed frequency sub-bands. In experimental results, we can show that the proposed method outperforms previous approaches in terms of edge enhancement, reduced aliasing effects, and reduced blurring effects.

Keywords Super-resolution · Patch-based · Discrete wavelet transform · Local binary pattern

D. K. Shin · Y. S. Moon (✉)
Department of Computer Science and Engineering,
Hanyang University, Ansan, Republic of Korea
e-mail: ysmoon@hanyang.ac.kr

D. K. Shin
e-mail: cmilre83@gmail.com

1 Introduction

Computer vision and pattern recognition techniques are utilized for recent digital image system. Researches on the pattern recognition, however, are in trouble due to the degradation of image quality and resolution. In order to address this difficulty, follow-up studies have been focusing on the approach to super-resolution (SR) image reconstruction. This technique is usually used in aviation, surveillance, medical images, and etc. SR image reconstruction techniques are categorized into a multi-frame SR [1, 2], which uses multiple low-resolution images taken from the same scene, and a single-frame SR, which uses a single low-resolution image.

There are two different methods for the single-frame SR. The first method uses pixel-interpolation techniques such as the nearest neighbor, bi-linear, B-Spline, cubic-convolution, cubic spline, and etc. [3, 4] Discrete Wavelet Transform (DWT)-based SR methods [5–9] are also proposed, which estimates coefficients in high-frequency sub-bands (Low-High: LH, High-Low: HL, and High-High: HH) by interpolating from the correlation between sub-bands. Mueller et al. [8] proposed the interpolation method based on the DWT, estimating coefficients in high-frequency sub-bands by iteratively removing noises while preserving important coefficients around edges from multi-scale geometric information generated by contourlet transform. Unfortunately, HR images generated by this method have shortfalls such as blurring edges and/or textures in high frequency sub-bands, and jagging artifacts at diagonal edges. In order to overcome these shortages, interpolation methods [10–17] adaptive to edge directions were proposed to preserve coefficients in high-frequency sub-bands. Even though these methods enable to preserve the distinctness of edge regions compared to iterative interpolation

methods, it is hard to preserve regions with the detailed texture.

Recently, example-based SR image reconstruction methods [18–25] have been receiving attention to address the above problem. This approach utilizes learning database which consists of pairs between a patch in a high-resolution image and a patch in a corresponding low-resolution image. By using the learning database, an image patch in an input image is compared with low-resolution patches, and then the found low-resolution patches are replaced to high-resolution image patches corresponding to found patches because they are linked each other. Freeman et al. [18] first proposed this technique which is a learning-based SR method selecting high-resolution patches corresponding to low-resolution patches by modeling the spatial relationship between two patches using Markov network. When high-frequency patches are selected, borders where overlaps occur are considered to select the best pair to increase the accuracy and connectivity between patches.

While the image is magnified, in general, the information loss brings blurring results. The loss of information in high-frequency areas with textures has a more decisive effect than in no texture areas. The reason is that the perception by the human eyes is more sensitive near edges where the difference of brightness exists when the distinctness of images is decided.

As described above, the distinctness of images is influenced by the reconstruction in the high-frequency areas. Freemans method does not consider the reconstruction in frequency domains because it just replaces low-resolution image patches to high-resolution image patches. In this regard, this paper proposes an SR image reconstruction method based on the DWT to estimate or reconstruct lost information in high-frequency domain. Our method is a novel example-based technique using wavelet patch-pairs to estimate or reconstruct coefficients in high-frequency sub-bands transformed by DWT.

2 Example-based Super-resolution Method using Discrete Wavelet Transform

Our SR image reconstruction method utilizes the Discrete Wavelet Transformation (DWT) to reconstruct information in high-frequency sub-bands (high-pass) as shown in Fig. 1. In this paper, we reconstruct a HR image magnified 2 times from a Low-Resolution (LR) image. An interpolated image is first magnified 2 times from an input LR image. Once one level DWT is applied to the interpolated image, the input image is decomposed into four sub-bands (Low-Low: LL, Low-High: LH, High-Low: HL, and High-High: HH) with 1/4 image size. The input LR image is inserted

into the LL sub-band, and other sub-bands are initialized as zero value (zero-padding) to estimate lost information in high frequency domain. This approach can reduce the loss of the input image because it is preserved in the LL band.

LH, HL, and HH as high-frequency sub-bands (high-pass) are obtained by the signed difference between horizontal, vertical, and diagonal directions when the input image is demagnified. This means that information in high-frequency sub-bands depends on pattern features in the LL sub-band because sub-bands are correlated each other. Accordingly, we can infer that pattern information in LH, HL, and HH sub-bands is similar to pattern features in LL sub-band if pattern features between a patch in an input image and a LL sub-band in database is similar. Using this relationship, we can estimate coefficients in LH, HL, and HH sub-bands by comparing between a patch in the LL sub-band and a wavelet patch-pair in learning database which consists of wavelet patch-pairs (a pair of LL, LH, HL, and HH wavelet patches) generated from HR images. In order to compare patches, we use Nearest Intensity - Local Binary Pattern (NI-LBP) which is used to find a class corresponding to a patch, and Subtraction of Center from Neighbors - MSE (SCN-MSE) which finds a wavelet patch-pair with the highest similarity from wavelet patch-pair database of the classified class. After searching wavelet patch-pairs corresponding to all patches in each high-frequency sub-band, each patch in the wavelet patch-pair is inserted into the same patch location corresponding to LH, HL, and HH sub-bands. Finally, we can make a super resolved HR image by applying the inverse DWT to four estimated sub-bands.

2.1 Generation of Training Dataset

In order to estimate coefficients in high-frequency sub-bands, we generated training database. The training database is generated by DWT from training image database as shown in Fig. 2. After an input image in training image database is decomposed to LL, LH, HL, and HH sub-bands by DWT, patches with the same location and size in sub-bands are divided. A wavelet patch-pair can be created and stored with a pair of (LL,LH,HL,HH). We repeat these processes until all the training images are completed. In this paper, we use Haar Wavelet and the patch size is 3 × 3. In addition, a pattern classification of wavelet patch-pair as shown in Fig. 2-④ is performed to reduce the retrieval cost and to increase the accuracy when we compare the similarity between a patch in LL sub-band and the LL patch of a wavelet patch-pair in training database.

This paper uses Neighbor Intensity of Local Binary Pattern (NI-LBP)[26] to classify wavelet patch-pairs. In
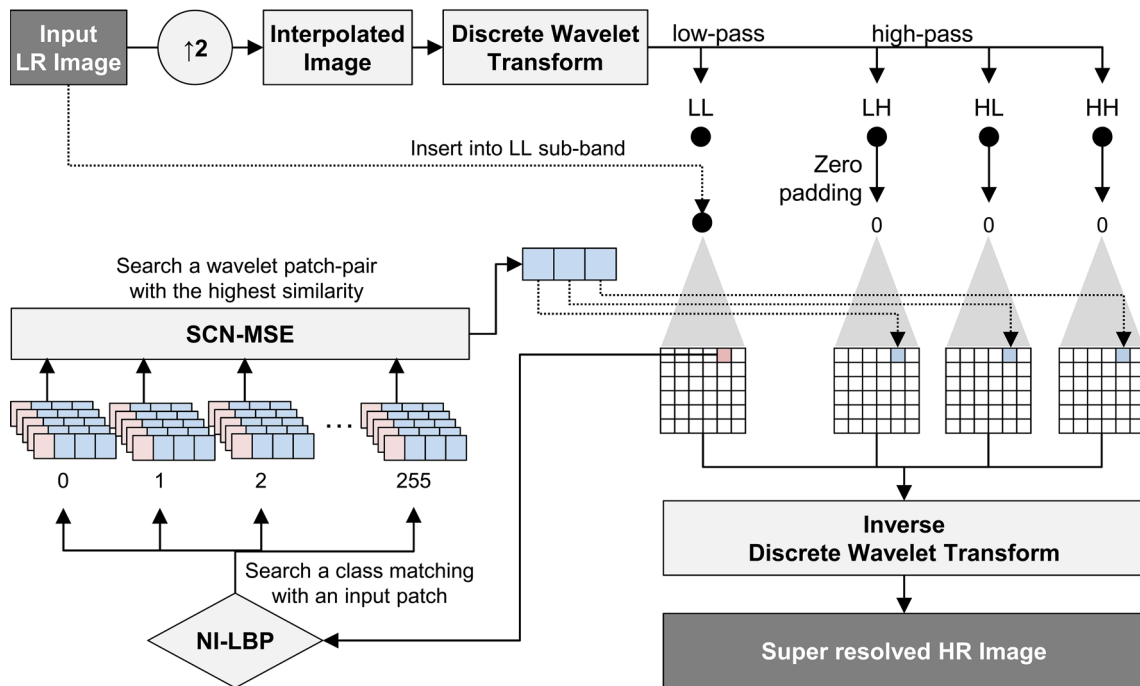
Figure 1 Overall flowchart of the proposed method.

general, Local Binary Pattern (LBP) descriptor [27] proposed by Ojala, which is based on the statistical features of texture, is used for the pattern recognition. This method has a good performance for the texture classification. LBP codes pixels values by using the relationship between the center pixel value of a patch and neighbor pixels values as shown in Eq. (1). If a neighbor pixel value is equal to or

greater than the center pixel value, then the code value is 1, otherwise, 0.

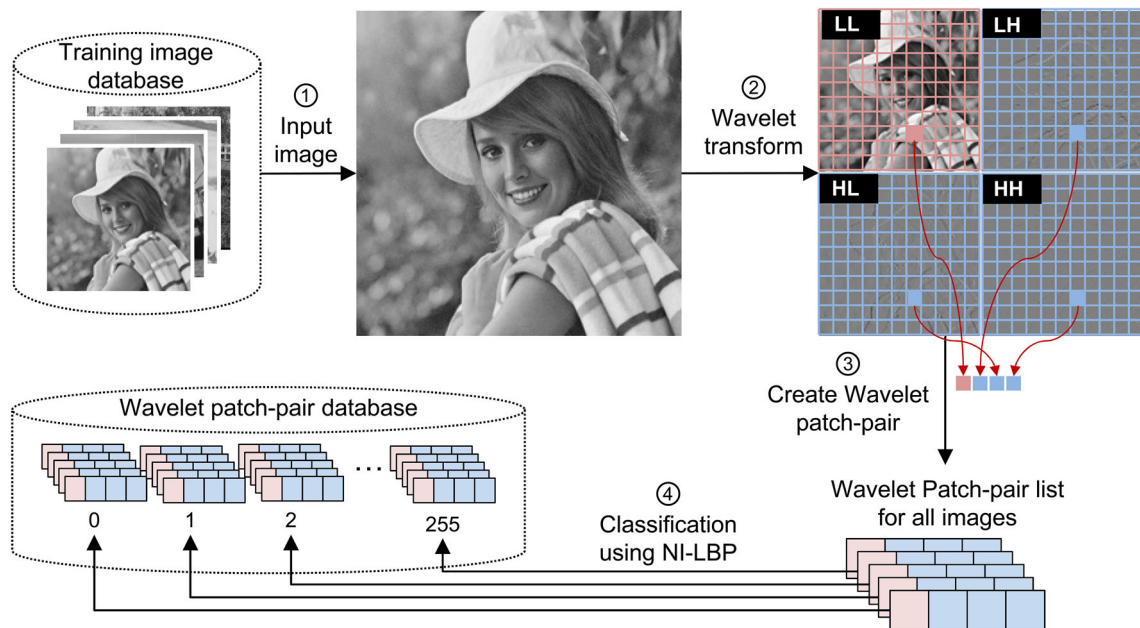$$LBP_{p,r} = \sum_{n=1}^{p-1} s(g_{r,n} - g_c)^{2^n} \tag{1}$$

Figure 2 Overall steps to generate the training database with wavelet patch-pair.

Since the LBP has a shortfall of recognizing edges with gradual intensity changes, this paper utilizes NI-LBP descriptor as shown in Eq. (2). This method uses the relationship between neighbor pixels value and the average intensity. If neighbor pixels value are equal to or greater than the average intensity, then the code value is 1, otherwise, 0.

$$NI - LBP_{p,r} = \sum_{n=1}^{p-1} s(g_{r,n} - \mu) 2^n \quad s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

where, $\mu = \frac{1}{2} \sum_{n=0}^{p-1} g_{r,n}$.

As shown in Fig. 3a, in case the gradual intensity change exists in the patch, human can divide it into three areas related with two edges. The LBP descriptor classifies it into the pattern with no consideration of edges, but the NI-LBP descriptor can separate it like the perception of the human eyes. In case of Fig. 3b, there is a diagonal edge from the left-bottom to the right-top. The LBP descriptor results in the different result although the intensity change is not significant. The reason is that the LBP descriptor distinguishes values from the intensity of the center pixel value. The result of using the NI-LBP descriptor, however, is similar to the pattern recognized by human.

Since the research on the SR image reconstruction is closely interrelated to the visual judgment of human, the pattern classification similar to the perception of the human eyes such as the NI-LBP descriptor can help increase the accuracy. In this paper, wavelet patch-pairs are classified into 256 classes by the NI-LBP descriptor as shown in Fig. 2.

2.2 Similarity Comparison between Patches in the LLSub-band

In this step, we propose a similarity function between a patch in the LL sub-band and a LL patch in wavelet patch-pairs of the learning database. The similarity comparison algorithm searches a class matching with a patch in the LL sub-band using the NI-LBP descriptor. In order to evaluate the similarity between the input patch and a patch in a matched class, we devise SCN-MSE which is based on the Mean Square Error (MSE) using Subtraction of Center from Neighbors (SCN) in pixels, as in Eq. (3).

$$SCN - MSE = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} ((I_{ij} - I_c) - (P_{ij} - P_c))^2 \quad (3)$$

where I and P are an input patch and an LL patch of wavelet patch-pairs in the matched class, I_c and P_c are the center pixel values of each patch, and N is the number of pixels in a patch. We repeat it until finding a patch with the highest similarity.

Notice that this paper estimates coefficients in LH, HL, and HH sub-bands which are the difference between gray pixels values in horizontal, vertical and diagonal directions when the input image is decomposed into four sub-bands by DWT. This means that the similarity comparison by gray pixel value may estimate wrong coefficients in sub-bands. Accordingly, the SCN-MSE considering the feature of DWT can resolve this problem. For example, Fig. 4a shows the gray-level of two input patches. Since two input patches have the same pattern by the NI-LBP descriptor, it is

Figure 3 The comparison between NI-LBP and LBP in terms of preserving weak edge patterns. (a) and (b) example patchs; **a1,b1** examples of coded value using NI-LBP; **a2,b2** examples of coded value using LBP.
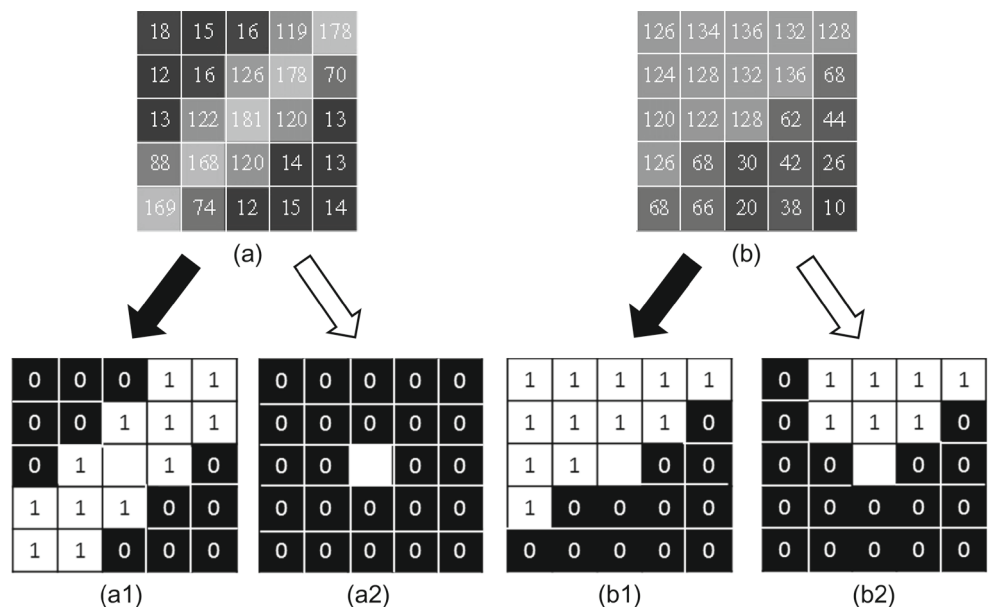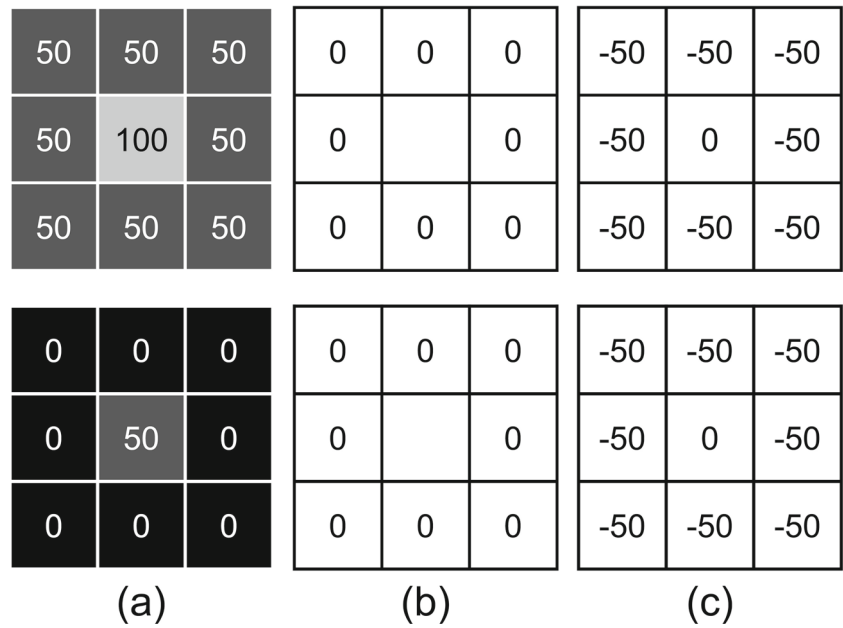
Figure 4 An example of the similarity comparison by using the proposed SCN-MSE. **a** Two image patches. **b** The result of applying the NI-LBP descriptor. **c** The result of applying SCN descriptor.

the same class and the similarity must be evaluated. Figure 4b depicts the result of the NI-LBP descriptor from those. When we apply the MSE to evaluate the similarity, the similarity between two patches is very low because it is 2500. In the MSE, the similarity is 0 if two patches are identical. However, Fig. 4c shows the result of the SCN that is the signed difference between neighbor pixels values and the center pixel value. This means the difference in horizontal, vertical, and diagonal directions. Accordingly, the SCN corresponds with the concept of DWT. The similarity calculated by MSE from the result of the SCN is 0, so that two patches are identical. Our SCN-MSE is utilized to find a

wavelet patch-pair with the highest similarity in the learning database. LH, HL, and HH patches of the wavelet patch-pair are used to estimate coefficients in LH, HL, and HH sub-bands of the input image.

2.3 Estimation of Coefficients in High-pass Sub-bands

We estimate coefficients in LH, HL, and HH sub-bands as shown in Fig. 5. The input LR image is inserted into the LL sub-band. The inserted LL sub-band is divided into 3×3 patches and each patch is compared to an LL patch of wavelet band-pairs in the learning database. If a wavelet
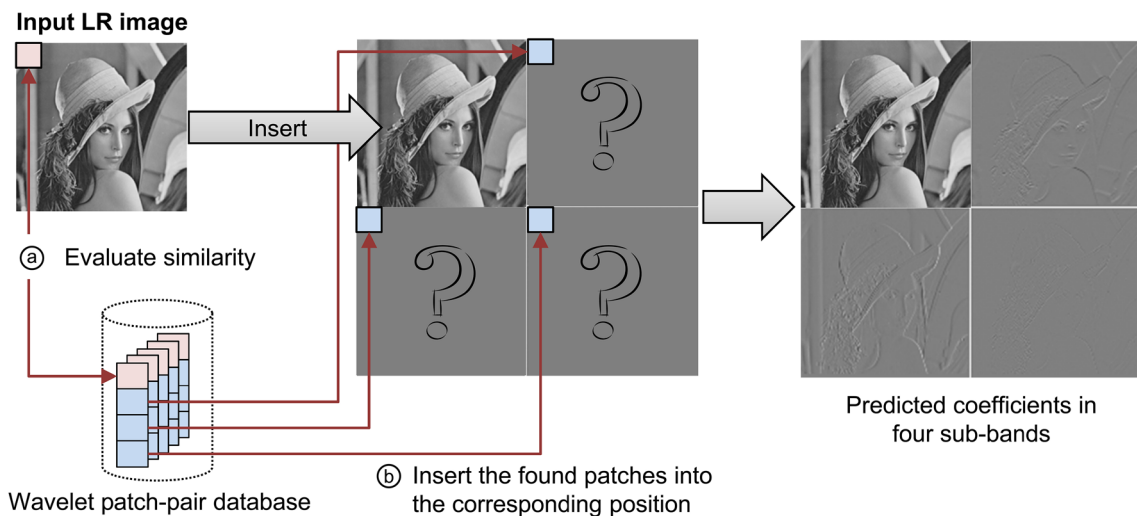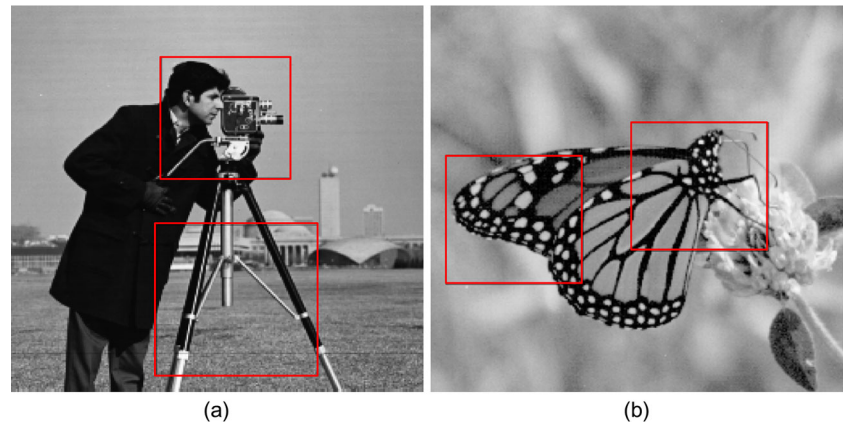

Figure 5 Overall steps to estimate coefficients in sub-bands through the similarity comparison from Wavelet band-pair database.

Figure 6 Cameramen and butterfly images used for experiments. **a** Cameramen image **b** Butterfly image; the red box indicates the interested regions to evaluate the proposed method and previous methods.

(a) (b)

patch-pair with the highest similarity by the SCN-MSE is found, LL, HL, and HH patches of the wavelet patch-pair are inserted into the corresponding patch location of LH, HL, and HH sub-bands, respectively. Finally, a high-resolution image is created by performing the inverse DWT from the estimated sub-bands.
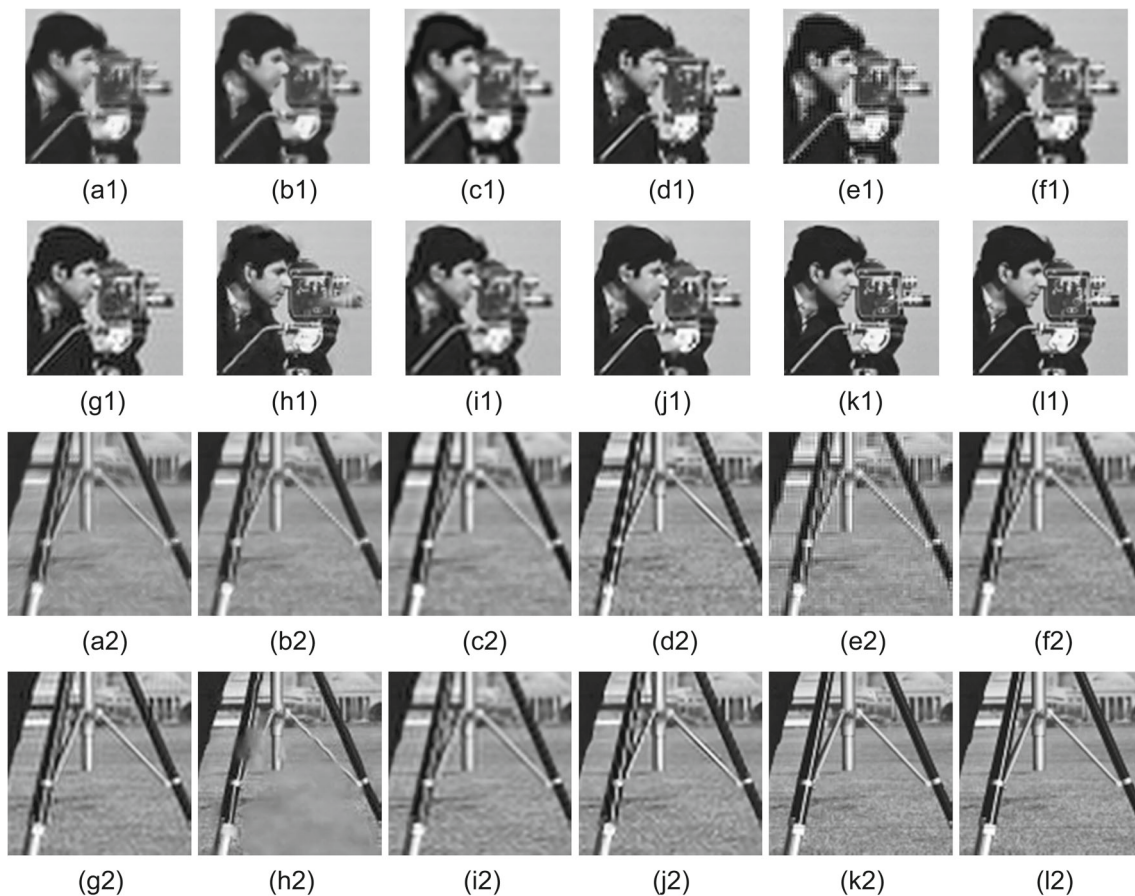

(a1) (b1) (c1) (d1) (e1) (f1)

(g1) (h1) (i1) (j1) (k1) (l1)

(a2) (b2) (c2) (d2) (e2) (f2)

(g2) (h2) (i2) (j2) (k2) (l2)

Figure 7 The comparison of previous methods and the proposed method in the face, camera region, and camera tripod of the camera-man image. **a1,a2** new edge-directed interpolation (NEDI) [11], **b1,b2** directional filtering and data fusion (DFDF) [13], **c1,c2** the single image detail synthesis with edge prior method [16], **d1,d2** sparse mixing estimators (SME) [17], **e1,e2** super-resolved image using DASR with haar wavelet function and bicubic interpolation [9], **f1,f2** super-resolved image using DASR with db.9/7 wavelet function and bicubic interpolation [9] **g1,g2** contourlet [8], **h1,h2** Freeman et al. [18], **i1,i2** Yang et al. [23], **j1,j2** Kim et al. [24]. **k1,k2** proposed method, **l1,l2** the truth high-resolution image.

3 Experimental Results

This paper proposed the SR reconstruction method based on Discrete Wavelet Transform (DWT) and the example-based technique using patches in order to keep the image discreteness and reduce blurring effects when the low-resolution (LR) image is magnified. We first built a learning database which is acquired from 50 images randomly selected from corel DB. In the experiments, the size of patches is set up to 3×3 and an LR image with the size of 128×128 is magnified to a Super Resolution (SR) image with the size of 256×256. Figure 6 shows the interested regions with the red box in the cameramen and butterfly images to compare the proposed method to previous interpolation methods. The image quality comparisons were evaluated using the regions cropped from HR images reconstructed.

In Figs. 7 and 8 (a)-(d) show the results of SR using the new edge directed interpolation (NEDI) method [11], the directional filtering and data fusion (DFDF) method [13], the single image detail synthesis with edge prior method16, and the sparse mixing estimators (SME) method [17], respectively. They are based on the interpolation method adaptive to edge directions. e and f in Figs. 7 and 8 show the results of SR using the haar wavelet and db.9/7 wavelet, respectively. g in Figs. 7 and 8 is the results of SR estimating coefficients in high-frequency sub-bands using the countourlet transform. As patch-based methods, h-j in Figs. 7 and 8 show the results of SR using Freeman method [18], Yang method [23], and Kim method [24], respectively. Finally, (l) indicates the truth HR image.

Figure 7(a1)-(l1) show the results of the HR image reconstructed from the LR cameramen image. When focusing on the region of the face and camera, the results of the
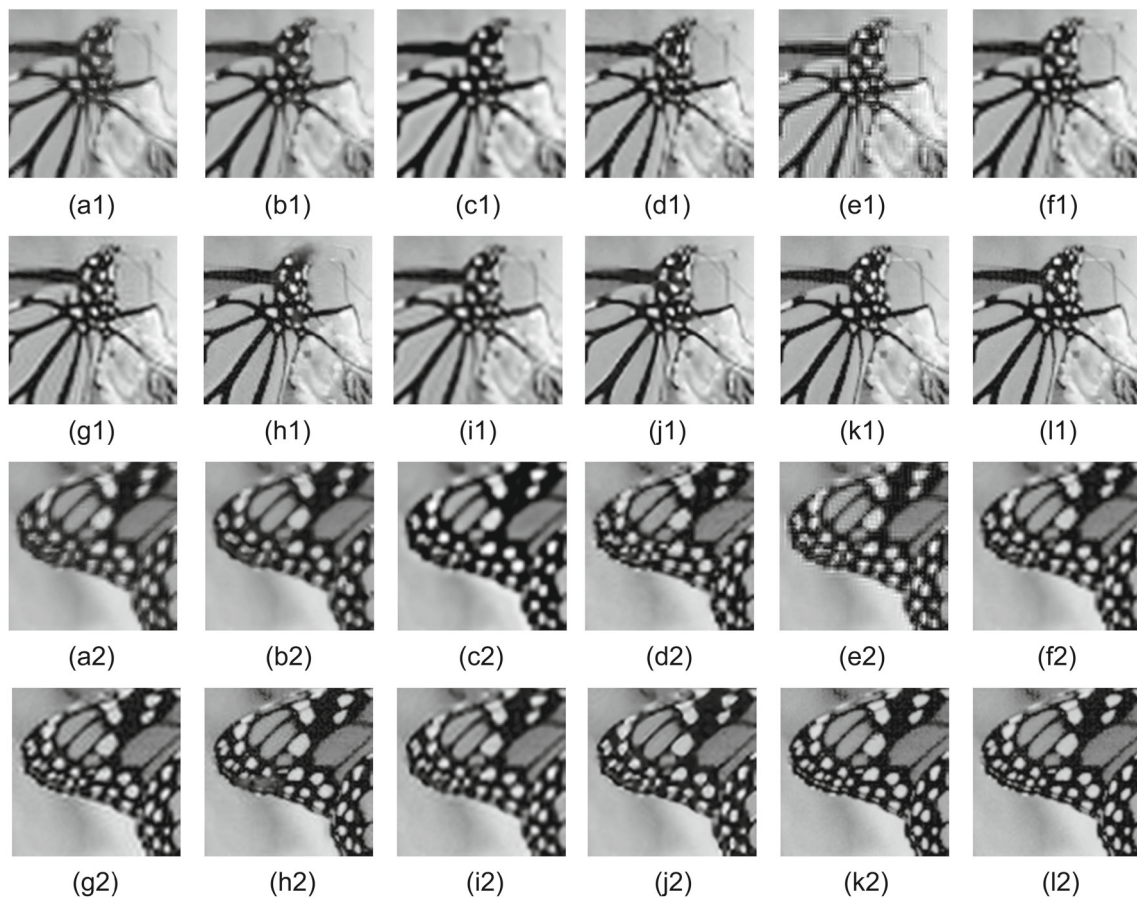
Figure 8 The comparison of previous methods and the proposed method in the face, camera region, and camera tripod of the camera-man image. **a1,a2** new edge-directed interpolation (NEDI) [11], **b1,b2** directional filtering and data fusion (DFDF) [13], **c1,c2** the single image detail synthesis with edge prior method [16], **d1,d2** sparse mixing estimators (SME) [17], **e1,e2** super-resolved image using DASR with haar wavelet function and bicubic interpolation [9], **f1,f2** super-resolved image using DASR with db.9/7 wavelet function and bicubic interpolation [9] **g1,g2** contourlet [8], **h1,h2** Freeman et al. [18], **i1,i2** Yang et al. [23], **j1,j2** Kim et al. [24]. **k1,k2** proposed method, **l1,l2** the truth high-resolution image.

previous methods might be blurred around edges. Although the result of Freeman method in Fig. 7(h1) is more discrete than previous methods in the face region, we can see the blurring effect around the hair region. In addition, the edges of the camera body are reconstructed well, but the lens region is blurred. On the other hand, we can see that the proposed method as shown in Fig. 7(k1) sharply reconstructed the eyes, nose, and mouth regions of the face and the detailed edges of the camera compared with the HR image of Fig. 7(l1).

Figure 7(a2)-(l2) show the results of HR image reconstruction for the region of the camera tripod in the cameraman image. In the edge-based methods, Fig. 7(a2)-(d2), the SME method of Fig. 7(d2) sharply reconstructed the edge regions than other edge-based methods compared to the high-resolution image, Fig. 7(l2). However, we can see the aliasing effect around the camera tripod. For Wavelet-based methods, Fig. 7(e2)-(g2), the contourlet method of Fig. 7(g2) outperforms others, but results in a little aliasing effect around the camera tripod and the blurring effect in horizontal edge direction on the building background. In the example-based methods, Fig. 7(h2)-(j2), the reconstruction of using Freeman method of Fig. 7(h2) has good

results around the horizontal and vertical edges in the building background, and less aliasing effect around the camera tripod. However, the reconstruction of the lawn with complex textures results in the severe blurring effect. As shown in the proposed methods result of Fig. 7(k2), our method outperforms previous methods in the result of the building background, the camera tripod, and the lawn. The reconstructed image quality is similar to the HR image of Fig. 7(l2).

Figure 8 shows the results of the cropped regions in the reconstructed HR butterfly image. Most of previous methods results in burring effects around the butterfly head as shown in Fig. 8(a1)-(l1). We can see less aliasing effects in case of the reconstruction of the butterfly wing patterns using the edge-based methods from Fig. 8(a1) to Fig. 8(d1). On the other hand, the Wavelet-based methods in Fig. 8(e1)-(g1) have more aliasing effects. Kim et al. method of Fig. 8(j1) has the most distinct result, but our method of Fig. 8(k1) is better than in the regions of the butterfly head and wing patterns. In addition, our method reconstructed two lines in the tip region of the butterfly wing as shown in Fig. 8(k2) compared to the truth HR image of Fig. 8(l2), but other methods from Fig. 8(a2)-(j2) did not.

Table 1 The results of PSNR for the reconstructed HR images using previous methods and the proposed method.

method	PSNR Index						
	cameraman	lena	butterfly	cat	filnstones	man	Avg.
NEDI	24.17	27.20	23.58	27.60	19.17	25.70	24.57
DFDF	24.25	27.13	23.63	27.72	19.13	25.66	24.59
EP	25.55	28.80	25.71	28.78	19.79	26.88	25.92
SME	26.23	32.19	27.13	32.91	21.53	28.72	28.12
Avg.	25.05	28.83	25.01	29.25	19.90	26.74	25.80
WZP(haar)	24.36	28.30	24.73	29.00	19.18	26.74	25.80
WZP(db.9/7)	24.36	27.31	24.07	28.01	19.23	25.80	24.81
Countourlet	27.35	32.28	28.85	33.78	22.41	29.34	29.02
Avg.	25.38	29.33	25.88	30.27	20.28	27.25	26.40
Freeman et al.	26.52	33.66	29.16	24.99	17.05	23.83	25.87
yang et al.	26.61	31.81	27.88	33.78	23.20	30.05	28.89
Kim et al.	27.52	32.48	29.22	33.48	23.55	30.24	29.42
Avg.	26.88	32.65	28.75	30.75	21.27	28.04	28.06
Proposed method	37.59	40.70	40.36	33.07	23.65	29.08	34.08

In order to quantitatively evaluate the image resolution quality, we utilize an evaluation measure based on the Mean Square Error (MSE) and the Peak Signal-to-Noise Ratio (PSNR). The MSE is used to measure the amount of data loss through the pixel value comparison. As the PSNR is derived from MSE, it is used to measure the image quality. The PSNR index indicates the ratio between the maximum possible pixel value of the original image and the pixel value by noise. The equation is as the following:

$$PSNR = 10 \log_{10}(\frac{R^2}{MSE}) \tag{4}$$

where R is the maximum possible pixel value of the input image (R is 255 represented by 8 bits) and MSE represents the MSE between the given input image I_{inp} and the original image I_{org} which can be obtained by the following :

$$MSE = \frac{\sum_{i,j}(I_{inp}(i,j) - I_{org}(i,j))^2}{M \times N} \tag{5}$$

where M and N are the size of the images.

The measurement by the ratio of signal noise simply transferred, however, is not enough because the last signal receiver is the human eyes in the image communication. In order to take it into account, Wang et al. [28] proposed a new theory that the human eyes extract structural information of the input signal when Human Visual System (HSV) recognizes images. They devised a Structural SIMilarity (SSIM) of images based on this theory. This method represents the original signal into $x = \{x_i | i = 1, 2, \ldots, m \times n\}$ and the distorted signal into $y = \{y_i | i = 1, 2, \ldots, m \times n\}$ in the size $m \times n$ of the window. The average $l(x, y)$, the contrast $c(x, y)$, and the correlation $s(x, y)$ are calculated as shown in Eq. (6). After that, the structural similarity is measured by multiplying each other.

$$
\begin{aligned}
l(x, y) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad C_1 = (K_1 L)^2 \\
c(x, y) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad C_2 = (K_2 L)^2 \\
s(x, y) &= \frac{\sigma_{xy} + C_3}{\sigma_x + \sigma_y + C_3}, \quad C_3 = c_2/2
\end{aligned}
\tag{6}
$$

where μ_x and μ_y are the mean of the signal x and y as the brightness, respectively, σ_x and σ_y are the variance of the

Table 2 The results of using MSSIM for the reconstructed HR images using previous methods and the proposed method.

method	PSNR Index						
	cameraman	lena	butterfly	cat	filnstones	man	Avg.
NEDI	0.82	0.87	0.86	0.89	0.72	0.79	0.83
DFDF	0.83	0.87	.0.87	0.89	0.73	0.80	0.83
EP	0.84	0.89	0.88	0.91	0.75	0.82	0.85
SME	0.87	0.94	0.90	0.95	0.84	0.87	0.90
Avg.	0.84	0.90	0.88	0.91	0.76	0.82	0.85
WZP(haar)	0.83	0.87	0.91	0.89	0.73	0.80	0.84
WZP(db.9/7)	0.83	0.89	0.88	0.90	0.74	0.81	0.84
Countourlet	0.88	0.94	0.92	0.96	0.83	0.89	0.90
Avg.	0.85	0.90	0.90	0.92	0.77	0.83	0.86
Freeman et al.	0.84	0.96	0.93	0.74	0.46	0.63	0.76
yang et al.	0.87	0.93	0.91	0.96	0.87	0.90	0.91
Kim et al.	0.89	0.95	0.93	0.96	0.82	0.90	0.91
Avg.	0.87	0.95	0.92	0.89	0.72	0.81	0.86
Proposed method	0.97	0.99	0.98	0.96	0.87	0.90	0.95

signal x and y as the contrast, respectively, σ_{xy} is the covariance between the signal x and y as the correlation between two signals, respectively, and L is the dynamic range of the pixel values. The SSIM measure uses the following parameter settings: $K_1 = 0.01$, $K_2 = 0.03$.

$$MSSIM(x, y) = \frac{1}{M} \sum_{j=1}^{M} SSIM(x_i, y_i) \qquad (7)$$

where x and y are the reference and distorted images, respectively, x_i and y_i are the image contents at the $j - th$ local window, and M is the number or local windows in the image.

Table 1 shows the results of measuring the image quality by the PSNR through six experimental images. The bold number indicates the highest PSNR index. The proposed method outperforms other methods in most results except the cat and man images. The PSNR index of our method is the average 8.28db higher than edge-based methods, the average 7.68db higher than wavelet-based methods, and the average 6.02db higher than example-based methods.

Table 2 shows the results of MSSIM to measure the structural similarity based on the perception of the human eyes. The higher similarity indicates that the MSSIM index is closer to 1. The bold number is the results with the highest similarity through six experimental images. The MSSIM index of our method is the heisted similarity as the average 0.95. It is the average 0.1 higher than edge-based methods, the average 0.09 higher than wavelet-based methods, and the average 0.09 higher than example-based methods.

4 Conclusions

We proposed a novel example-based Super Resolution (SR) image reconstruction method using the discrete wavelet transform. Our method estimates coefficients in the high-frequency sub-bands by searching the high-frequency patches matching with patches in the sub-bands of the input low-resolution image. As the experimental results, our method reduced blurring effects and aliasing effects. In addition, the high-resolution images were sharply reconstructed compared to the original one. For quantitative analysis, we used the PSNR to measure the amount of data loss and the MSSIM (Mean of Structural Similarity) to measure the structural similarity based on the

perception of the human eyes. The experimental results prove that the proposed method outperforms previous methods in most cases.

References

1. Anbarjafari, G., & Demirel, H. (2010). Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image. *ETRI Journal*, *32*(3), 390–394.
2. Carrato, S., Ramponi, G., Marsi, S. (1996). A simple edge-sensitive image interpolation filter. In *Proceeding IEEE conference on image processing* (Vol. 3, pp. 711–714).
3. Chang, H., Yeung, D., Xiong, Y. (2004). Super-resoultion through neighbor rm- bedding. In *Procedding IEEE conference on computer vision and pattern recognition(CVPR)* (pp. 275–282).
4. Chughtai, N.A., & Khattak, N. (2006). An edge preserving locally adaptive anti-aliasing zooming algorithm with diffused interpolation. In *Proceeding 3rd Canadian conference on computer robot vision* (pp. 49–55).
5. Datsenko, D., & Elad, M. (2007). Example-based single image super resolution: a global MAP approach with outlier rejection. *Journal of Multidimensional System and Signal Processing,* 18 (2-3), 103–121.
6. Farsin, S., Robinson, D., Elad, M., Milanfar, P. (2004). Fast and robust multi-frame super-resolution. In *IEEE transactions on image processing* (Vol. 13, no. 10, pp. 1327–1344).
7. Farsiu, S., Elad, M., Milanfar, P. (2006). Multiframe demosaicing and super-resolution of color images. In *IEEE transactions on image processing* (Vol. 15, no. 1, pp. 141–159).
8. Freeman, W.T., Jones, T.R., Pasztor, E.C. (2002). Example based super-resolution. In *IEEE computer graphics and applications* (Vol. 22, no. 2, pp. 56–65).
9. Keys, R. (1981). Cubic convolution interpolation for digital image processing. In *IEEE transactions on acoustics speech and signal processing* (Vol. 29, no. 6, pp. 1153–1160).
10. Kim, K.I., Kim, D.H., Kim, J.H. (2004). Example-based learning for image super-resolution. In *Procedding third Tsinghua-KAIST joint workshop pattern recognition* (pp. 140–148).
11. Kim, K.I., & Kwon, Y. (2008). Example-based learning for single image super-resolution. In: *Proceeding on DAGM symposium*, (pp. 456–465).
12. Kim, K.I., & Kwon, Y. (2010). Single-image super-resolution using sparse regression and natural image prior. In *IEEE transactions on pattern analysis and machine intelligence* (Vol. 32, no. 6, pp. 1127–1133).
13. Lee, S.W., & Paik, J.K. (1993). Image interpolation using adaptive fast B-spline filtering. In *IEEE international conference on acoustics, speech and signal processing* (Vol. 5, pp. 177–180).
14. Li, X., & Orchard, M.T. (2001). New edge-directed interpolation. In *IEEE transactions on image processing* (Vol. 10, no. 10, pp. 1521–1527).
15. Liu, L., Zhao, L., Long, Y., Kuang, G., Fieguth, P.W. (2012). Extended local binary patterns for texture classification. Journal of Image Vision and Computing, 30(2), 86–99.
16. Mallat, S., & Yu, G. (2010). Super-resolution with sparse mixing estimators. In *IEEE transactions on image processing* (Vol. 19, no. 11, pp. 2889–2900).

17. Malgouyres, F., & Guichard, F. (2002). Edge direction preserving image zooming: a mathematical and numerical analysis. Society for industrial and applied mathematics (SIAM). *Journal on Numerical Analysis*, 39(1), 1–37.
18. Mueller, N., Lu, Y., Do, M.N. (2007). Image interpolation using multi-scale geometric representations. In *Proceedings of SPIE computational imaging V* (vol. 6498, p. 64980A).
19. Ojala, T., Pietikainen, M., Maenpaa, T. (2002). Multiresolution grayscal and rotation invariant texture classification with local binary patterns. In *IEEE transactions on pattern analysis and machine intelligence* (Vol. 24, no. 7, pp. 971–987).
20. Tai, Y.W., Liu, S., Brown, S., Lin, S. (2010). Super resolution using edge prior and single image detail synthesis. In *Proceeding IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 2400–2407).
21. Temizel, A., & Vlachos, T. (2005). Wavelet domain image resolution enhancement using cycle-spinning. *Electronics Letters*, *41*(3), 119–121.
22. Temizel, A., & Vlachos, T. (2005). Image resolution upscaling in the wavelet domain using directional cycle spinning. *Journal of Electronic Imaging, 14*(4), 040501.
23. Wang, Q., Tang, X., Shum, H. (2005). Patch based blind image super resolution. In *Proceeding IEEE international conference on computer vision (ICCV)* (Vol. 1, pp. 709–716).
24. Wang, Q., & Ward, R.K. (2007). A new orientation-adaptive interpolation method. In *IEEE transactions on image processing* (Vol. 16, no. 4, pp. 889–900).
25. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P. (2004). Image quality assessment: from error measurement to structural similarity. In *IEEE transactions on image processing* (Vol. 13, no. 4, pp. 600–612).
26. Yang, J., Wright, J., Huang, T., Ma, Y. (2008). Image super-resolution as sparse representation of raw image patches. In *Proceeding IEEE conference on computer vision and pattern recognition(CVPR)* (pp. 1–8).
27. Zhang, L., & Wu, X. (2006). An edge-guided image interpolation algorithm via directional filtering and data fusion. In *IEEE transaction on image processing* (Vol. 15, no. 8, p. 2226).
28. Zhao, S., Han, H., Peng, S. (2003). Wavelet domain HMT-based image super resolution. In *IEEE international conference on image processing* (Vol. 2, pp. 933–936).

Do Kyung Shin Received her MS in computer science and engineering from Hanyang University, Republic of Korea, in 2007. She is currently working toward her PhD at the Computer Vision and Pattern Recognition Laboratory, Department of Computer Science and Engineering, Hanyang University, Republic of Korea. Her major interests include image super resolution, object segmentation, 3D reconstruction and multimedia processing.

Young Shik Moon Received the BS and MS in electronics engineering from Seoul National University and Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea, in 1980 and 1982, respectively, and Ph.D. degree in electrical and computer engineering from the University of California at Irvine, CA, in 1990. From 1982 to 1985, he had been a researcher at the Electronics and Telecommunications Research Institute, Daejon, Republic of Korea. In 1992 he joined the department of Computer Science and Engineering at Hanyang University, Republic of Korea, as an Assistant Professor, and is currently a Professor. His research interests include computer vision, image processing, pattern recognition, and computational photography.