

Region Adaptive Single-image Super-resolution Using Wavelet Transform

Oh-Jin Kwon¹ and Je-Ho Park²

¹*Department of Electronics Engineering, Sejong University, Seoul, Korea*

²*Department of Computer Science and Engineering, Dankook University, Cheonan-si, Chungnam, Korea*

¹*ojkwon@sejong.ac.kr, ²dk_jhpark@dankook.ac.kr*

Abstract

An efficient single-image super-resolution based on the discrete wavelet transform is proposed. The low-resolution image to be enhanced is assumed to be the low-frequency subband of the high-resolution image to be reconstructed. A support vector machine that synthesizes the high-frequency subband based on the inter-subband correlation of the directional edges is designed. Subband coefficients are classified into one of homogeneous region, edge region, and textured region. The support vector machine is adaptively trained for each region. Experimental results of sample images show that the proposed system offers improvements in terms of both measured distortion and subjective appearance.

Keywords: *super-resolution, image resolution enhancement, discrete wavelet transform, support vector machine*

1. Introduction

Super-resolution (SR) technology, enhancing the image resolution of a low-resolution (LR) observation to obtain a high-resolution (HR) image, has traditionally been used for surveillance, remote sensing, and medical purposes but is becoming more common in consumer products such as cellular phone, digital camera, high-definition television (HDTV), and closed-circuit television (CCTV). Whenever images taken from LR sensors are enlarged or moved to HR displays, SR technology is essential [1-17].

Based on the number of available LR images, SR algorithms can be classified into two types: single reference SR (SRSR) and multiple references SR (MRSR). Technical overviews, problem formulations, and recent advances in MRSR can be found in the literatures [1-7]. MRSR algorithms perform SR by assuming that a set of LR images obtained from different viewpoints is available, and most algorithms formulate the problem as two signal processing aspects: how the multiple LR images from the set are aligned (or affined) in a common coordinate system and how the HR image is reconstructed from the aligned LR images.

However, in SRSR, as only one image is available, the problem is reduced to image resolution enhancement of the given image. Traditionally, image interpolation techniques such as nearest-neighbor interpolation, bilinear interpolation, and bicubic interpolation have been used. These techniques are known to work well in smooth regions but tend to blur edges and sharp details in the images [8-17].

Recently, many researchers have investigated SRSR techniques based on the discrete wavelet transform (DWT) which have proven to be superior to interpolation-based techniques [9-14]. DWT-based SRSR algorithms normally assume that the LR image to

be enhanced is the low-frequency subband of the HR image to be reconstructed. These algorithms perform SR by estimating or synthesizing the DWT coefficients in the high-frequency subbands of the HR image from the LR image. Carey, *et al.*, [9] observed the regularity of edges by measuring the decay of the DWT coefficients across scales and proposed an algorithm for synthesizing an image twice the size of the original by extrapolating the fine-scale subbands preserving the underlying regularity. Kinebuchi, *et al.*, [10] modeled the statistical relationship between the DWT coefficients using a hidden Markov tree and used this model to interpolate the fine-scale coefficients. Temizel and Vlachos [11] exploited the DWT coefficient correlation in a local-neighborhood sense and employed a linear least-squares regression to estimate the unknown fine-scale coefficients. Piao, *et al.*, [12] utilized the correlation of the subbands with different sampling phases in the DWT and estimated the high-frequency subbands by using filters designed in the lower level under the assumption that filters connecting two bands are similar in different levels. Demirel and Anbarjafari [13] proposed an algorithm to correct the initially interpolated fine-scale coefficients by using the high-frequency subbands achieved by the stationary wavelet transform of the LR image. Gajjar and Joshi [14] proposed a learning-based approach; given an LR image to be enhanced and a database consisting of LR and HR images, the authors first obtained an initial HR estimate by learning the high-frequency details from the available database. They modeled the LR image as the aliased and noisy version of the corresponding HR image and obtained the final HR image by estimating the aliasing matrix entries using an inhomogeneous Gaussian Markov random field as the prior model

This paper proposes a support vector machine (SVM) synthesizing the high-frequency subband based on the inter-subband correlation of the directional edges. The LR image to be enhanced is assumed to be the low-frequency subband of the HR image to be reconstructed. Edge detection is performed on the LR image. Based on the edges, the LR image is classified into three regions: homogeneous region, edge region, and textured region. The introduction of the importance of edge directionality for exploiting the inter-subband correlation in each region is the contribution of our system. Most DWT-based SR algorithms [9-14] have exploited their own inter-subband correlation for the algorithm design. However, the edge direction for each region has not been related to the inter-subband correlation. A relative distribution of large coefficients in the high-frequency subbands is shown to be closely related to the edge strength and direction at a given position in each region. Motivated by the success of the learning-based approach [14-16] and the support vector regression method [17] for SRSR, our SVM is designed to perform SR for each region based on the inter-subband correlation of the directional edges.

The energy properties of the DWT coefficients and the proposed SRSR system are described in Section 2. The experimental results are presented in Section 3. Finally, Section 4 concludes the paper.

2. Proposed System

The DWT decomposes an input image into multiple subbands: LL, LH, HL, and HH. The input image is simultaneously passed through a low-pass filter (LPF) and a high-pass filter (HPF) and then down-sampled by 2:1 in the horizontal and vertical directions, respectively, to generate the DWT coefficients. If the LPF and HPF are suitably designed, the inverse DWT (IDWT), which is the process to reconstruct the input image perfectly from the DWT coefficients, is possible. Multi-level decomposition is achieved

by obtaining a higher-level decomposition from recursively applying the transform on the LL band of the lower level.

An example of one and two-level DWT is shown in Figure 1(a) and Figure 1(b), respectively. The DWT provides a scalable representation of images with the higher and lower decomposition levels representing the coarser and finer characteristics, respectively. In this paper, the k 'th row and the l 'th column of the DWT coefficient of the LL, LH, HL, and HH bands of decomposition level s are denoted by $x_{LL}^{(s)}(k,l)$, $x_{LH}^{(s)}(k,l)$, $x_{HL}^{(s)}(k,l)$, and $x_{HH}^{(s)}(k,l)$, respectively. The notation for Figure 1(a) and 1(b) are shown in Figure 1(c) and 1(d), respectively. Using this notation, the design problem for the SRSR system is to reconstruct the high-frequency subbands of HR image corresponding to $x_{LH}^{(1)}(k,l)$, $x_{HL}^{(1)}(k,l)$, and $x_{HH}^{(1)}(k,l)$ by assuming that the given LR image is $x_{LL}^{(1)}(k,l)$, which is equivalently a set of $x_{LL}^{(2)}(m,n)$, $x_{LH}^{(2)}(m,n)$, $x_{HL}^{(2)}(m,n)$, and $x_{HH}^{(2)}(m,n)$.

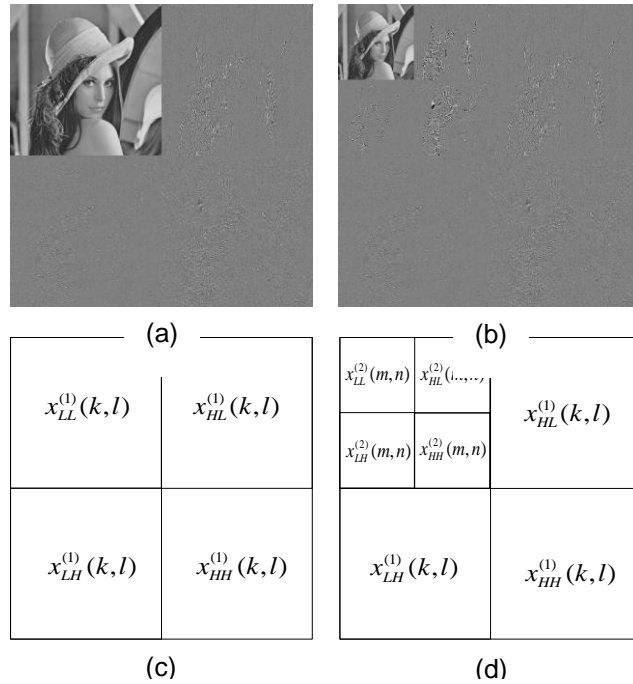


Figure 1. Examples of the DWT: (a) One-level DWT and (b) Two-level DWT. (c) and (d) Show our Notation for (a) and (b), Respectively

In this paper, the energy distribution of the DWT coefficients for the SRSR system design is the particular focus. The number of high-energy DWT coefficients is observed to increase in the edge regions. Our experiments for this observation have been performed by using the Sobel edge detector [18]. The edge region in $x_{LL}^{(1)}(k,l)$ is extracted by applying a Sobel mask and thresholding the large values. Figure 2(a) shows the extracted edge regions. It is observed that the large coefficients of $x_{LH}^{(1)}(k,l)$, $x_{HL}^{(1)}(k,l)$, and $x_{HH}^{(1)}(k,l)$ shown in Figure 1(a) are concentrated in the edge region shown in Figure 2(a). Then, based on the ratio of the horizontal to vertical Sobel masking [18], the edge region can be further separated into horizontal and vertical components as shown in Figure 2(b) and 2(c), respectively. The large-coefficients distribution is seen to be highly dependent on the directionality of edges. For example, $x_{LH}^{(1)}(k,l)$ and $x_{HL}^{(1)}(k,l)$ are

more closely related with the horizontal and vertical edge regions, respectively. Thus, the conclusions useful for the SRSR system design can be summarized as follows

1) The energy of the DWT coefficients in the high-frequency subband is concentrated in the edge region, and the DWT coefficients are large in magnitude in this region.

2) The relative distribution of the large coefficients in the LH, HL, and HH subbands is closely related to the edge strength and direction at a given position. Therefore, reconstructing the high-frequency DWT coefficients based on the edge strength and direction information at a given position is highly recommended.

3) The inter-subband correlation between the DWT coefficients in the mother and child bands is almost negligible when the distance is greater than 2 pixels in the mother band [10, 11, 14]. Therefore, it is sufficient for the SRSR to reconstruct the DWT coefficient in the child band by only using its corresponding coefficient and the 8 neighboring coefficients in the mother band.

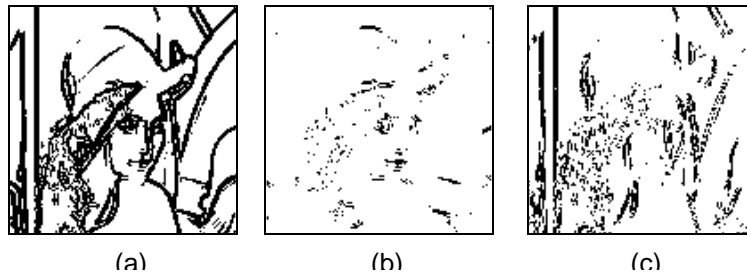


Figure 2. Examples of the Edge Regions: (a) Edge Region, (b) Horizontal Edge Region, and (c) Vertical Edge Region

This then allows an SRSR system for reconstructing an image that is twice the size of the original to be proposed. The basic structure of the system is shown in Figure 3 and described as follows.

1) Assume that the given LR image to be enhanced is the low-frequency subband of the HR image to be reconstructed and denote the LR image as $x_{LL}^{(1)}(k,l)$.

2) Perform one-level DWT on $x_{LL}^{(1)}(k,l)$ and obtain the DWT coefficients. Denote the resulting coefficients in the LH, HL, and HH bands by $x_{LH}^{(2)}(m,n)$, $x_{HL}^{(2)}(m,n)$, and $x_{HH}^{(2)}(m,n)$, respectively.

3) Perform edge detection on $x_{LL}^{(1)}(k,l)$ and denote the edge strength and direction by $|e(k,l)|$ and $\angle e(k,l)$, respectively.

4) Based on the edge strength and direction, classify $x_{LL}^{(1)}(k,l)$ into one of homogeneous region, edge region, and textured region.

5) Generate the high-frequency subbands of the HR image: $x_{LH}^{(1)}(k,l)$, $x_{HL}^{(1)}(k,l)$, and $x_{HH}^{(1)}(k,l)$, adaptively for each region by using the SVM whose input vector is $\{|e(i,j)|, \angle e(i,j), x_{LH}^{(2)}(m,n), x_{HL}^{(2)}(m,n), x_{HH}^{(2)}(m,n)\}; k-1 \leq i \leq k+1, l-1 \leq j \leq l+1, \lceil k/2 \rceil - 1 \leq m \leq \lceil k/2 \rceil + 1, \lceil l/2 \rceil - 1 \leq n \leq \lceil l/2 \rceil + 1\}$, where $\lceil z \rceil$ is the maximum integer value not greater than z .

6) Perform IDWT on $\{x_{LL}^{(1)}(k,l), x_{LH}^{(1)}(k,l), x_{HL}^{(1)}(k,l), x_{HH}^{(1)}(k,l)\}$ to obtain the HR image.

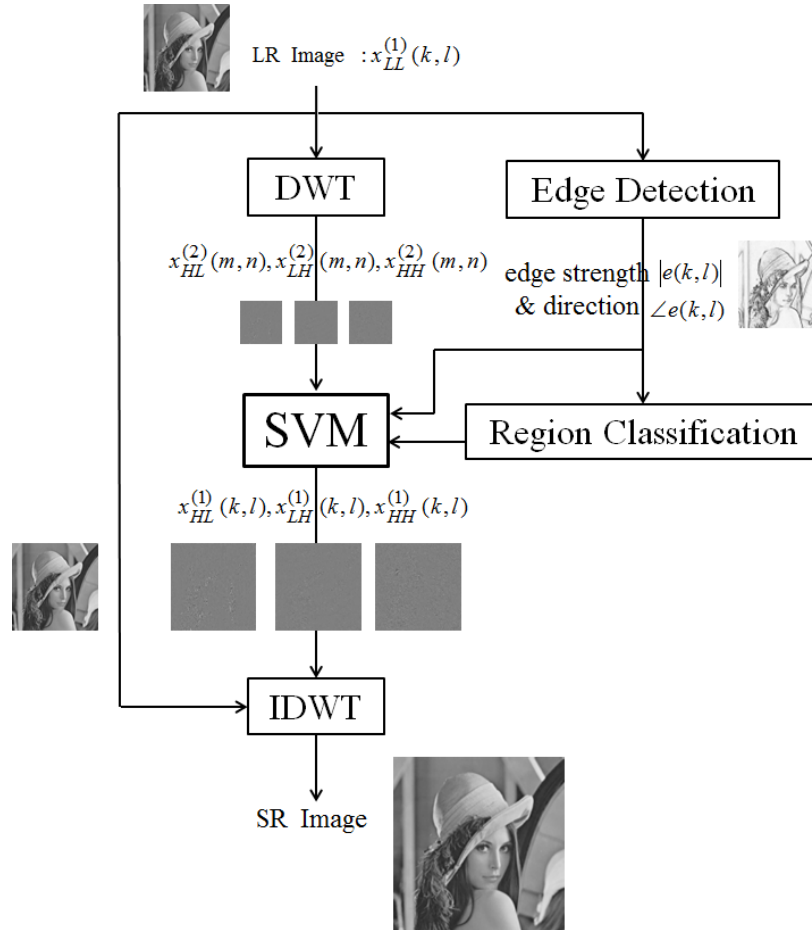


Figure 3. Basic Structure of the Proposed SRSR System

We have designed our system based on the inter-subband correlation between one-level mother and child DWT coefficients. It may be claimed that the usage of higher level inter-subband correlation is more helpful. However, it is noted that our experimental results for the usage of higher level inter-subband correlation on various images have shown that the performance improvement is almost negligible and it only increases the system complexity.

3. Experimental Results

In the experiments, the well-known Daubechies 9/7 analysis-synthesis filters are used for the implementation of the DWT and the Sobel edge detector is used for the implementation of the edge detection. A sophisticated edge detection method, such as the Canny edge detector [19], may be considered as a better edge detection choice. But it is at the expense of increased complexity and our experiments for various test images have shown that the improvement by using the better edge detector is almost negligible for our SRSR system performance.

We convolve the LR image $x_{LL}^{(1)}(k,l)$ with 3×3 horizontal and vertical Sobel masks [18]; denote the resulting values of the convolution at position (k, l) by $e_h(k,l)$ and

$e_v(k,l)$, respectively; and obtain the edge strength and direction by calculating $|e(k,l)| = \sqrt{e_h^2(k,l) + e_v^2(k,l)}$ and $\angle e(k,l) = \tan^{-1}[e_v(k,l)/e_h(k,l)]$, respectively.

To extract textured regions, the recursive thresholding method using several texture features [20] is employed. Edge strength randomness, edge direction randomness, edge contrast, edge density, and the number of horizontal and vertical peaks and valleys in a region are selected as texture features. Textured regions are extracted by obtaining the histograms of these features, selecting one of the histograms showing the best clusters, and thresholding it recursively until none of the histograms shows clustering. The more detailed algorithm for extracting textured regions is preferably referred to Reference [20]. After textured regions are extracted, edge regions are extracted by selecting the pixels whose edge strength is greater than a threshold value and clustering the selected pixels showing similar edge direction. All the pixels not included in any of textured regions and edge regions are declared as homogeneous regions.

In order to train the least dependent SVM on selected image samples, a random collect of 400 images with sizes larger than 512×512 pixels was assembled; the sample images include a variety of indoor and outdoor shots of buildings, streets, people, faces, animals, landscapes, *etc.*, taken from cellular phones, digital cameras or camcorders, HDTV, and CCTV. For simple implementation and fast processing, the input and output vectors of the SVM: $|e(i,j)|$, $x_{LH}^{(2)}(m,n)$, $x_{HL}^{(2)}(m,n)$, $x_{HH}^{(2)}(m,n)$, $x_{LH}^{(1)}(k,l)$, $x_{HL}^{(1)}(k,l)$, and $x_{HH}^{(1)}(k,l)$, are set to integer values. The edge direction $\angle e(i,j)$ is also quantized to 36 levels. No detectable degradation of the performance was observed as a result of this quantization.

For the performance tests, four well-known test images were used: Lena, Elaine, Baboon, and Peppers, which were not included in the sample set for the SVM training. All the sample images are 512×512 pixels. For objective tests, one-level DWT on the test images was performed and the LL band images were assumed to give LR images for a $2 \times$ resolution enhancement. The original images were taken as the ground truth and the performance was compared in terms of the peak signal-to-noise ratio (PSNR) metric. The PSNR results are listed in Table 1. The results are compared with two popular interpolation techniques: bilinear interpolation and bicubic interpolation. For comparison purposes only, the PSNR results obtained by Temizel and Vlachos¹¹ that are, to the authors' knowledge, regarded as one of the best set of results among the DWT-based SRSR systems are also listed. In Table 1, 'Wavelet-0' refers to the basic reconstruction method where unknown high-frequency subbands: $x_{LH}^{(1)}(k,l)$, $x_{HL}^{(1)}(k,l)$, and $x_{HH}^{(1)}(k,l)$, are estimated as zeroes. The results in Table 1 show that the proposed SRSR system gives an improved metric compared with the other techniques for all test images.

For subjective tests, the experimental results of the Lena and Baboon images are presented as examples. Results for the other images were similar. Figure 4 and 5 show the reconstructed HR images for Lena and Baboon, with details shown in Figure 6 and 7, respectively. The proposed system can be seen to improve the blurring and ringing distortions seen in the images reconstructed by the conventional bicubic interpolation method and the 'Wavelet-0' method.

Table 1. PSNR Comparison with other Techniques (dB)

Techniques	Lena	Elaine	Baboon	Peppers
Bilinear	30.84	31.10	22.44	29.35
Bicubic	31.15	31.30	22.64	29.41
Wavelet-0	35.92	33.45	24.08	32.27
Temizel & Vlachos	35.39	33.40	24.52	34.46
Proposed	37.89	34.12	26.61	34.98



Figure 4. Experimental Results for the Lena Image: (a) the Original Image and the Image Reconstructed by (b) Bicubic Interpolation, (c) 'Wavelet-0', and (d) the Proposed Technique

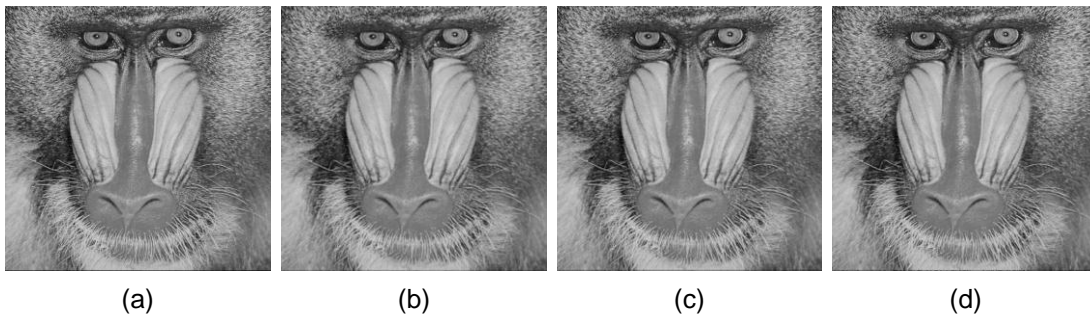


Figure 5. Experimental Results for the Baboon Image: (a) the Original Image and the Image Reconstructed by (b) Bicubic Interpolation, (c) 'Wavelet-0', and (d) the Proposed Technique

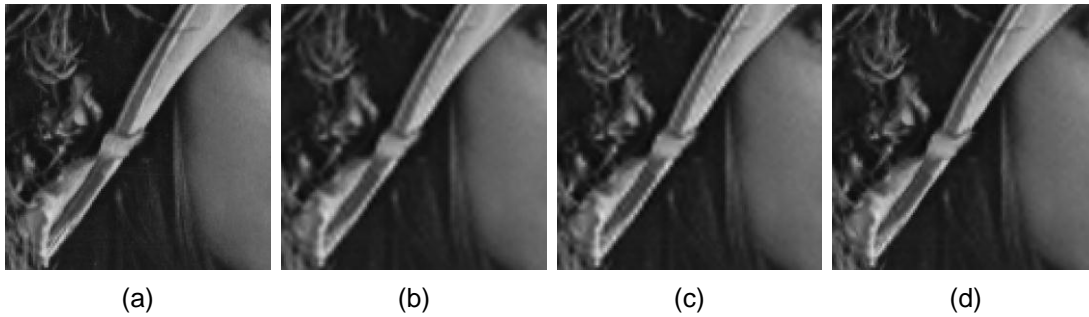


Figure 6. Enlarged Part of Figure 4: (a) the Original Image and the Image Reconstructed by (b) Bicubic Interpolation, (c) 'Wavelet-0', and (d) the Proposed Technique

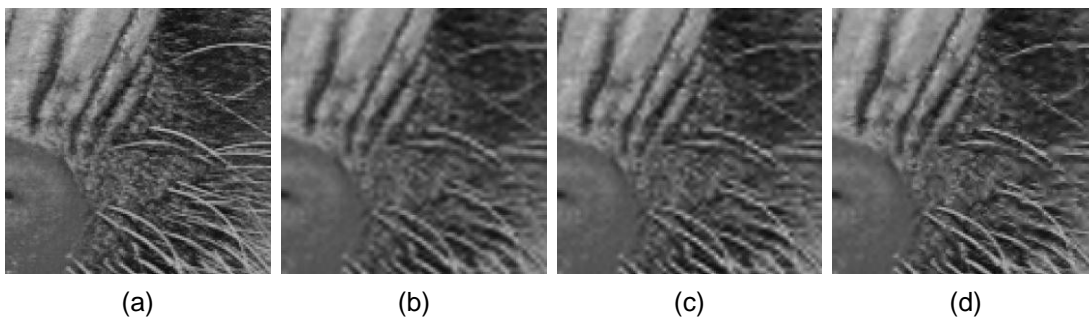


Figure 7. Enlarged Part of Figure 5: (a) the Original Image and the Image Reconstructed by (b) Bicubic Interpolation, (c) 'Wavelet-0', and (d) the Proposed Technique

Table 2. Computing Time for Processing Sample Images (Seconds)

Techniques	Lena	Elaine	Baboon	Peppers
Bilinear			0.1098	
Bicubic			0.2319	
Wavelet-0			0.5506	
Temizel & Vlachos			1.2589	
Proposed	1.0241	1.0184	1.0281	1.0149

Besides the quality of reconstructed HR images, computing efficiency is also important in real applications. All the experiments were conducted on a computer equipped with Duo 2.4 GHz CPU and 8.0 GB RAM memory. Table 2 shows the execution time of the proposed system compared to the other methods of Table 1. The execution time of all the other compared methods is sample image independent. The execution time of our system may vary based on the result of edge detection and region classification for each test image. But it has been known that the variations are almost negligible. It is seen that the proposed method is fast enough to be used for applications in which images are required to be scaled-up very quickly and shows better computing efficiency compared to the method proposed by Temizel and Vlachos [11].

4. Conclusion

A wavelet-based SRSR system using inter-subband correlation of the directional edge has been proposed. The system assumes that the LR image to be enhanced is the LL band of the HR image to be reconstructed. The LR image is divided by homogeneous region, edge region, and textured region. The high-frequency subbands of the HR image are adaptively generated for each region by the SVM whose input vectors are the mother band DWT coefficients and the edge strength and direction parameters. Experimental results showed that the system outperforms conventional image enhancement methods by improving the blurring and ringing distortions normally appearing in those methods.

The system only enhances the image to twice the size of the original, and future research will be focused on developing an optimal SRSR system that can control the size of the enhanced image based on the proposed system.

Acknowledgements

This research was supported by the ICT Standardization program of MISIP(The Ministry of Science, ICT & Future Planning).

References

- [1]. S. C. Park, M. K. Park and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, (2003).
- [2]. Z. Wei, F. Tao and W. Jun, "Kalman filter-based method for image superresolution using a sequence of low-resolution images," *J. Electron. Imaging*, vol. 23, no. 1, (2013).
- [3]. U. Mudenagudi, S. Banerjee and P. K. Kalra, "Space-time super-resolution using graph-cut optimization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, (2011).
- [4]. H. Su, L. Tang, Y. Wu, D. Treter and J. Zhou, "Spatially adaptive block-based super-resolution," *IEEE Trans. Image Processing*, vol. 21, no. 3, (2012).
- [5]. H. Ji and C. Fermuller, "Robust wavelet-based super-resolution reconstruction: theory and algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, (2009).
- [6]. B. Ning and X. Gao, "Multi-frame image super-resolution reconstruction using sparse co-occurrence prior and sub-pixel registration," *Neurocomputing*, vol. 117, (2013).
- [7]. W. Liu and S. Li, "Multi-morphology image super-resolution via sparse representation," *Neurocomputing*, vol. 120, (2013).
- [8]. D. f. Shen and C. W. Chiu, "Fundamental techniques for resolution enhancement of average subsampled images," *J. Electron. Imaging*, vol. 21, no. 3, (2012).
- [9]. W. K. Carey, D. B. Chuang and S. S. Hemami, "Regularity-preserving image interpolation," *IEEE Trans. Image Processing*, vol. 8, no. 9, (1999).
- [10]. K. Kinebuchi, D. D. Muresan and T. W. Parks, "Image interpolation using wavelet-based hidden Markov trees," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, (2001).
- [11]. A. Temizel and T. Vlachos, "Wavelet domain image resolution enhancement," *IEE Proc.-Vis. Image Signal Process*, vol. 153, no. 1, (2006).
- [12]. Y. Piao, I. Shin and H. W. Park, "Image resolution enhancement using inter-subband correlation in wavelet domain," *Proc. IEEE Int. Conf. Image Process.*, (2007), pp. 1-445-448.
- [13]. H. Demirel and G. Anbarjafari, "Image resolution enhancement by using discrete and stationary wavelet decomposition," *IEEE Trans. Image Processing*, vol. 20, no. 5, (2011).
- [14]. P. P. Gajjar and M. V. Joshi, "New learning based super-resolution: use of DWT and IGMRF prior," *IEEE Trans. Image Processing*, vol. 19, no. 5, (2010).
- [15]. S. Zhang and Y. Lu, "Image resolution enhancement via image restoration using neural network," *J. Electron. Imaging*, vol. 20, no. 2, (2011).
- [16]. W. Wu, Z. Liu, X. He, and W. Gueaieb, "Single-image super-resolution based on Markov random field and contourlet transform," *J. Electron. Imaging*, vol. 20, no. 2, (2011).
- [17]. K. S. Ni and T. Q. Nguyen, "Image superresolution using support vector regression," *IEEE Trans. Image Processing*, vol. 16, no. 6, (2007).

- [18]. R. Crane, *A Simplified Approach to Image Processing: Classical and Modern Techniques in C*, Prentice Hall, New Jersey, (1997).
- [19]. J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, (1986).
- [20]. O. Kwon, R. Chellappa, "Region adaptive subband image coding," *IEEE Trans. Image Processing*, vol. 7, no. 5, (1998).

Authors

Oh-Jin Kwon



He received a B.S. degree from Hanyang University, Seoul, Korea, in 1984, an M.S. degree from the University of Southern California, Los Angeles, 1991, and a Ph.D. degree from the University of Maryland, College Park, in 1994, all in electrical engineering. From 1984 to 1989, he was a research staff member at the Agency for Defense Development, Korea, and from 1995 to 1999, he was the head of the Media Laboratory at Samsung SDS, Seoul. Since 1999, he has been a faculty member at Sejong University, Seoul, Korea, where he is currently an Associate Professor. His research interests are image and video coding, watermarking, analyzing, and processing.



Je-Ho Park

He is an associate professor in the Computer Science and Engineering Department at Dankook University, South Korea. His research interests include multimedia database system, storage architecture, multimedia metadata, and indexing. Currently, he is a member of JPEG as co-editor of JP Search Part 2 and 6 as well as chair of JPEG Augmented Reality Ad hoc Group. Park has a PhD in multi-layered database architectures from the Polytechnic Institute at New York University.