LEARNING-BASED SUPER-RESOLUTION USING A MULTI-RESOLUTION WAVELET APPROACH

Changhyun Kim[*], *Kyuha Choi*[*], *Kyuyoung Hwang*[**], *and Jong Beom Ra*[*]

[*]School of Electrical Engineering and Computer Science, KAIST
Daejeon, Korea
[**]SAIT, Samsung Electronics Inc.
Suwon, Korea
E-mail: chkim@issserver.kaist.ac.kr

ABSTRACT

In this paper, we propose a learning-based super-resolution algorithm. In the proposed algorithm, a multi-resolution wavelet approach is adopted to perform the synthesis of local high-frequency features. To obtain a high-resolution image, wavelet coefficients of two dominant LH- and HL-bands are estimated based on wavelet frames. In order to prepare more efficient training sets, the proposed algorithm utilizes the LH-band and transposed HL-band. The training sets are then used for the estimation of wavelet coefficients for both LH- and HL-bands. Using the estimated high frequency bands, a high resolution image is reconstructed via the wavelet transform. Experimental results demonstrate that the proposed scheme can synthesize high-quality images.

Keywords: Super-resolution, learning-based, wavelet frames

1. INTRODUCTION

In many applications, an image of higher spatial resolution is preferred. Super-resolution (SR) refers to the process by which a higher-resolution image is synthesized from low-resolution (LR) image(s). Most SR algorithms require multiple LR images which are aligned within sub-pixel accuracy. In this paper, however, we focus on a SR algorithm using a single image, since it has more advantage in real-time applications.

Image enlargement can be done by using a simple interpolation technique. However, this approach cannot provide details in lines, edges, corners, and texture regions. Recently, there are several attempts for learning-based SR [1-5]. For example, Freeman *et al*. [1] prepared training sets based on the relationship between high-resolution (HR) images and their corresponding LR image, and produced SR images by using the sets. Sun *et al*. [2] developed a learning-based hallucination method by considering the similarity between image patches of low and high-resolution images. In determining the relationship of patches, they used image primitives such as edges, corners, and so on. Note here that learning-based methods generally share the common idea of using training sets for prior information [8]. Using the training sets, one can estimate the details of HR image from an input LR image.

In this study, we develop a learning-based SR algorithm based upon a wavelet synthesis approach. The relationship between an HR image and its LR image is closely correlated with the relationship between an image and its LL-band in the discrete wavelet transform (DWT) structure. In order to reconstruct an HR image, high bands of LH, HL, and HH are needed. However, since the HH-band is found not to be very critical for image quality, it is ignored in the proposed algorithm. For the prediction of LH- and HL-bands, we adopt the discrete wavelet frames (DWF) structure, in which coefficients in sub-bands are not sub-sampled unlike in the DWT [6].

This paper is organized as follows. Section 2 describes the proposed training process. In Section 3, the corresponding synthesis process is described. Section 4 shows experimental results and we conclude this paper in section 5.

2. TRAINING PROCESS

The SR system consists of two basic processes, namely a training process and a synthesis process. In the training process, we save an LR patch with its corresponding HR patch that contains fine details as a training set.

As was mentioned, an LR image is similar to the LL-band of an HR image in the DWT structure. According to the assumption that the high spatial-frequency components of an LR image are useful for predicting the details of HR image [7], LH- and HL-bands of the LL-band are used for estimating LH-and HL-bands of an HR image, instead of the LL-band itself. Note here that LH- and HL-bands are obtained by using the DWF technique without sub-sampling.

For efficient estimation, the LH- and HL-bands are divided into small patches. The patch size is typically 5×5 for sub-bands of the LL band and 10×10 for sub-bands of a training HR image. If the patch size is too large, noticeable artifacts may occur in the super-resolved image. Meanwhile, if the patch size is too small, it is hard to determine a proper patch. Note here that we adopt DWF instead DWT because the sub-sampling process in DWT makes difficult to find a proper patch. Furthermore, in DWT, it is hard to examine the relationship between adjacency patches.

HR patches of LH- and HL-bands of an HR image and their corresponding LR patches of LH- and HL-bands of the LL band are stored as training sets. Figure 1 presents a schematic illustration of the training process of the proposed SR system. In the figure, we can notice that LH- (or HL-) band has similar information to the HL- (or LH-) band besides its orientation. Hence, we can build training sets only for the LH-band by transposing (or rotating by 90°) the HL-band.

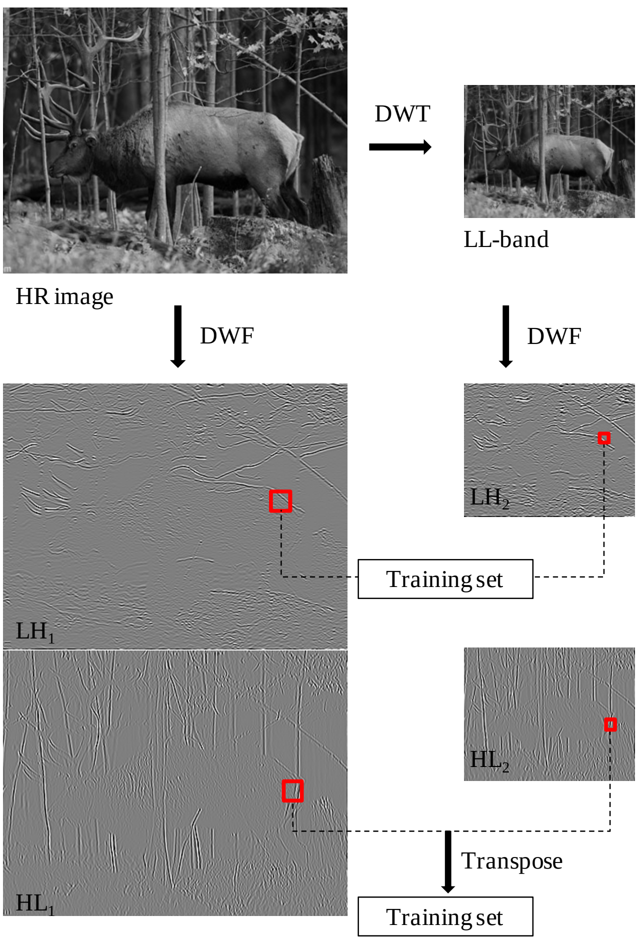

Figure 1. Training process

We store the HR patch corresponding to every possible LR patch in the LH- and transposed HL-bands. To reduce the size of database, by assuming that the characteristic of patches is independent of the brightness, we apply a local contrast normalization scheme. We also remove noisy or flat patches by examining the variance.

3. SYNTHESIS PROCESS

The overview of the synthesis process is shown in Figure 2. The first step of synthesis phase for SR is to decompose an input LR image into each band by using the DWF. We then estimate the LH- and HL- bands for an HR image by using the training database. Note that the HL-band is estimated after transposing it so that we may use the same training database as in the LH-band.

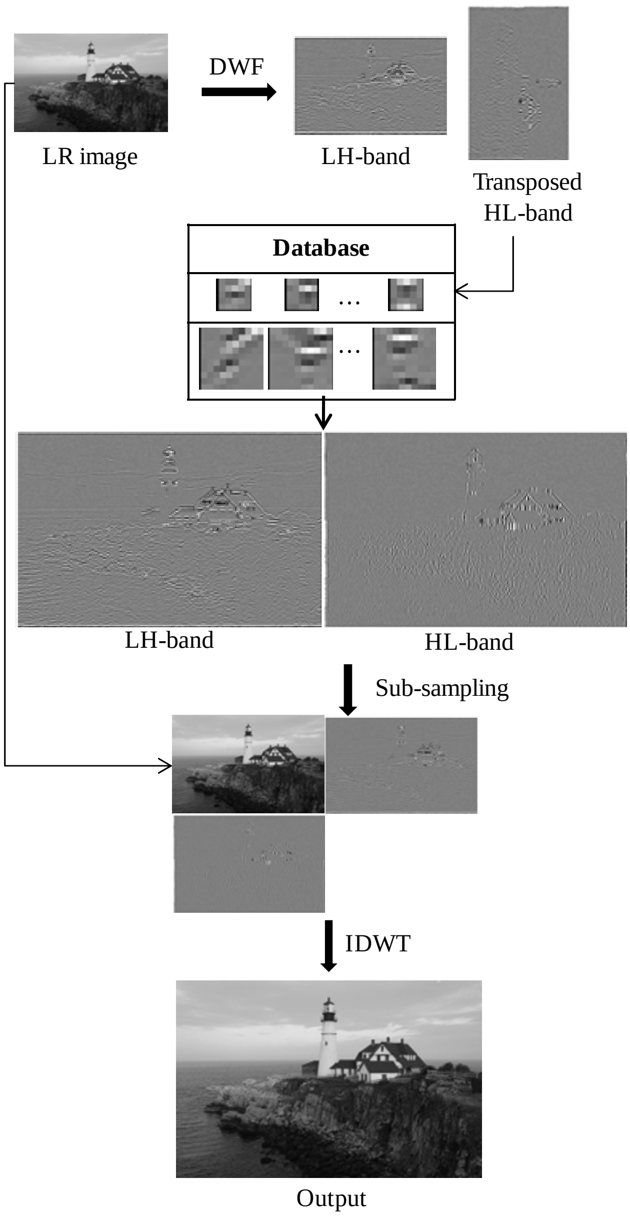

Figure 2. Synthesis process

For HR image reconstruction, we use an input LR image as the LL band. Since an LR image patch can be regarded as a patch in the LL-band of the DWT structure, we use the inverse DWT (IDWT) for HR image reconstruction. For the IDWT operation, we perform sub-sampling after estimating the whole LH- and HL-bands.

To determine HR patches, each band is divided into small patches of a size of 5×5. We then perform a local contrast normalization procedure and store normalization factor α for de-normalization. Examining the similarity of the input patch and the training LR patch based on the sum of difference (SAD) measure, we estimate an HR patch.

Since local image information itself is not sufficient to estimate HR details, we are enforced to use inter-patch relationship. For a given LR patch, we examine a training database in order to find the *K* candidates closest to the input patch. Among them, we select the best match by considering neighboring HR patches that are already estimated. This searching process starts from an input LR patch having the largest variance. This is because a patch having a large variance is more distinguishable than a patch having a small variance.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed algorithm, we select various twelve training images as shown in Figure 3. All the images are downloaded from a public web site of http://www.the-digital-picture.com /Gallery/. They were taken with a Canon EOS D60 digital camera with a resolution of 500×333 pixels. Using these images, about 100,000 training sets have been extracted.

Figure 3. Twelve training images

We use the 9/7 wavelet filters for DWF and DWT which are widely used in various image compression schemes. Note that two parameters are to be determined in the proposed method. In experiments, the number of nearest neighbors K is set to 8 and the overlap between adjacent patches is set to 2 pixels.

We compare our approach with bi-cubic interpolation. We perform SR only on the image intensity because it is known that the human visual system is more sensitive to the brightness. It is clearly seen in Figure 4 that bi-cubic interpolation provides less details or lower resolution images. On the other hand, our approach shows good SR results without introducing noticeable artifacts.

5. CONCLUSION

In this study, we have proposed a learning-based SR method for the reconstruction of an HR image from a single LR image. In the proposed method, a wavelet approach is used to synthesize local high-frequency details. The SR results obtained for various images clearly demonstrate perceptual improvements compared with the ones obtained from interpolation.

ACKNOWLEDGMENT

This research was supported by a grant from Samsung Electronics Inc., Korea.

REFERENCES

[1] W. Freeman, T. Jones, and E. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and applications*, 22(2), pp. 55-65, 2002.

[2] J. Sun, N. Zheng, H. Tao, and H. Shum, "Image hallucination with primal sketch priors," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pp. 729–736, 2003.

[3] W. Liu, D. Lin, and X. Tang, "Hallucinating faces: Tensor-patch super-resolution and coupled residue compensation," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pp. 478-484, 2005.

[4] S. Lui, J. Wu, H. Mao, and J. J. Lien, "Learning-based super-resolution system using single facial image and multi-resolution wavelet synthesis," *Asian conference on computer vision*, 4884, pp. 96-105, 2007.

[5] W. Fan and D. Yeung, "Image hallucination using neighbor embedding over visual primitive manifolds," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–7, 2007.

[6] M. Unser, "Texture classification and segmentation

(a) (b) (c)

Figure 4. Test images magnified by two times using (a) bi-cubic interpolation, (b) proposed scheme; (c) original HR image

using wavelet frames," *IEEE transactions on image processing*, volume 4, pp. 1549-1560, 1995.

[7] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, "Learning Low-Level Vision," *International Journal. Computer Vision*, 40(1), pp. 25-47, 2000.

[8] J. H. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proceeding of Royal Society*, 265(1394), pp. 359-366, 1998.