Self-Similarity Based Image Super-Resolution on Frequency Domain

Sae-Jin Park, Oh-Young Lee, Jong-Ok Kim School of Electrical Engineering, Korea University, Seoul, Korea E-mail: {bebequ, triple0, jokim}@korea.ac.kr

Abstract— Self-similarity has been popularly exploited for image super resolution in recent years. Image is decomposed into LF (low frequency) and HF (high frequency) components, and similar patches are searched in the LF domain across the pyramid scales of the original image. Once a similar LF patch is found, the LF is combined with the corresponding HR patch, and we reconstruct the HR (high resolution) version. In this paper, we separately search similar LR and HR patches in the LF and HF domains, respectively. In addition, self-similarity based SR is applied to the new structure-texture domain instead of the existing LF and HF. Experimental results show that the proposed method outperforms several conventional SR algorithms based on self-similarity.

I. INTRODUCTION

SR (super resolution) is a technique that reconstructs a HR (high resolution) image from one or more observed LR (low resolution) images. Owing to the rapid advance in the display technologies, the era of UHD (ultra high definition) TV has just begun recently, and SR is receiving even more interest.

In natural images, small patches often occur in a repetitive manner within the original scale as well as across different scales. This observation is referred to as self-similarity, and has been popularly exploited for image processing such as denoising, edge detection, and retargeting. It has been also applied to learning-based image super resolution in recent years [7][8].

Self-similarity based SR techniques can be classified into two categories, depending on the domain to which the technique is applied. One is on the spatial domain as investigated in [4], where similar patches are searched across image pyramid scales on the spatial domain. The other combines self-similarity with the LF (low frequency)-HF (high frequency) domain, and this is almost equal to the learning based method such as [1]. The only difference is to find the LF-HF patch pair within and across input image scales. In other words, self-similarity based SR on the LF-HF domain does not require any prior database (or examples) unlike the traditional example-based method [1]. Thus, it can reduce computational complexity and memory consumption, when compared to the conventional learning based approach. Suetake in [2] first applied self-similarity to the LF-HF domain, not spatial. An input LR image is decomposed into LF and HF components. Similar patches are searched in the in-scale LF domain only, and the best LF match is combined

with the corresponding HF patch for HR (high resolution) reconstruction. Chen in [3] extended the work in [2] by searching across pyramid scales as well as the original source image scale.

In this paper, we propose a new single SR method, based on self-similarity. In the existing approach, an input LR image is partitioned into LF and HF components, and the missing HF details of the target HR image are produced by finding the similar LF patches in the LF domain. In the proposed method, however, similar patches are separately searched in both LF and HF domains in order to restore the missing HF component beyond the Nyquist frequency more accurately. This is the key difference between the proposed and existing methods. In addition, we scale the HF intensity that is restored by self-similarity in order to boost sharpness effect.

The rest of this paper is organized as follows. In Section 2, the proposed method is presented. In Section 3, performance evaluations are shown, and then, Section 4 concludes the paper finally.

II. OVERALL SYSTEM

In this Section, we first present the proposed SR method based on the LF-HF domain and next, the structure-texture domain is applied to our algorithm, instead of the LF-HF domain. Fig. 1 illustrates the basic concept of the proposed method, comparing it with the existing approach. In the existing approach as shown in Fig. 1 (a), a similar LF patch is searched on the LF domain, and the corresponding HF patch is added to the LF patch. Meanwhile, in the proposed method (Fig. 1 (b)), a similar patch is separately searched on both LF and HF domains, respectively. In addition, LF and HF patches are hierarchically reconstructed on the pyramid scales. In other words, LR patches are gradually up-scaled by a factor of 1.25, while they are immediately enlarged at a time in the existing approach. These points become fundamental differences.

A. Overall Architecture

The proposed SR algorithm consists of three parts: generation of image pyramids for two LF and HF components, HR reconstruction based on self-similarity, and back projection after combing LF and HF components. First of all, we decompose an input LR image into LF and HF components using a *Gaussian* Filter (H), which is given by



Fig. 1 Comparison between (a) the existing approach and (b) the proposed method

$$LF_0 = I_0 * H \tag{1}$$

$$HF_0 = I_0 - LF_0 \tag{2}$$

After decomposition, pyramid image is constructed for each frequency component as shown in Fig. 1 (b) in order to exploit cross-scale patch redundancy. In the image pyramid, the LF-HF pair of k^{th} layer is denoted by LF_k-HF_k. An input image scale is represented by k=0, and the image scale is increased by a factor of fixed scale, in proportion to the value of k. In constructing the image pyramid for SR, it has been reported that an incremental coarse-grained method using a small scale factor produces a more elaborate result [4].

The negative value of k indicates the down-scale of the original input, and the down-scaled version is obtained by

$$LF_0 = D(LF_0 * \sigma_k) \tag{3}$$

where $D(\cdot)$ and σ indicate down-sampling operator and PSF(Point Spread Function), respectively.

For reconstruction of HR layers that are unknown in pyramid image, the input scale layer is divided into a number of small-sized patches, and they are searched in the lower layers of the image pyramid. If a similar patch is found, the corresponding patch in its upper layer is copied to the query patch's upper layer position that is to be reconstructed. When searching a similar patch or obtaining the corresponding patch, due to a small scaling factor, the patches' location is likely to be in factional coordinates. To obtain more accurate results, query LR patch (p_q) and the searched LR patches (p_b) are extracted in factional coordinates so that HR patches (p_r) in their upper layer can be located at integer coordinates.

We reconstruct up the image pyramid until we reach the desired magnification ratio using the same searching method.

The best-matched LF and HF HR patches are searched repeatedly, and final are obtained by

$$LF_{k} = \arg\min d(LF_{k-1}(p_{q}), LF_{k-2...}(p_{b}))$$
 (4)

$$HF_k = \arg\min d(HF_{k-1}(p_q), HF_{k-2\dots}(p_b))$$
(5)

After reconstructing LF and HF HR patches individually, we combine them by

$$I_T = LF_T + HF_T * \alpha \tag{6}$$

 α is a parameter that amplifies the HF component to obtain a more visually appealing image.

Finally, we carry out back projection as post processing in order to ensure the consistency with the input image.

$$I_{T} = I_{T} + U(I_{0} - D(I_{T} * \sigma_{T})) * B$$
(7)

where B is the back projection kernel which is typically assumed a Gaussian, and $U(\cdot)$ is an up-sampling operator.

B. Alternative : edge structure and texture domain

Whereas we decompose an image into separate frequency bands using a Gaussian filter in previous subsection, we consider another domain on which an image is divided into. It also consists of two components of geometrical edge structure and texture [9]. To extract two components from an image, we use popular Total Variation (TV) L1 model. In the TV-L1 model, edge and texture component are obtained by solving an optimization problem. For details, refer to [9].

Fig. 2 shows the decomposition of the image into the edgetexture components, and compares it with the LF-HF domain. Edge structure means smoothly varying intensity and texture means a small-scale oscillatory part. Unlike Fig. 2 (a) - (b), Fig. 2 (c)-(d) do not lose small texture component and achieve sharp edge preservation with no artifacts. Thus, reconstructing HR image on the edge-texture domain, based on selfsimilarity can make more reasonable and elaborate results.

Fig. 2 Comparison of decomposition methods for *Monarch*; (a) LF component, (b) HF component, (c) edge structure, (d) texture

III. EXPERIMENTAL RESULTS

In all experiments, we set the magnification factor to 2, scaling factor (between layers in the image pyramid) to 1.25, and the LR and HR patch sizes to 4 by 4 and 5 by 5, respectively. The test LR input image is obtained after applying *Gaussian* blur kernel to the original HR image. For color images, we convert them to YIQ space and apply our algorithm only to the Y channel, and color components are simply up-scaled, using conventional interpolation. Fig. 3 compares the ground truth patches with the best-match patches found from three different domains such as spatial, LF domain, and HF domain. Comparing (b) and (e) in Fig. 3, the best-matched patches in the 4th and 5th column specially show that we can obtain a result more similar to the ground truth than findings on the spatial domain by searching similar patches in LF and HF domain separately.

Fig. 4 shows the subjective quality comparisons of the reconstructed HR image to conventional algorithms. In Fig. 4, (b) - (g) shows the enlarged images of (a), *Lighthouse*, around the window. Fig. 4 (c) is the result from S.Chen [3], which is generally blurred in edge regions. Fig. 4 (d) has more HF components in the edge. However, artifacts exist at the same time. Fig. 4 (e) is the result of the SR technique on spatial some noise around the window frame. Fig. 4 (f) shows our result, which is sharp at the window frame and congruous to original images. From these experimental images, we can see that our proposed algorithm produces HR images with clear edges and fewer artifacts than other methods.

Fig. 3 Comparison of best-matched patches found on (b) spatial domain, (c) LF domain, and (d) HF domain. (e) is the combination of (c) and (d). (e) is the ground truth.

Fig. 4 Comparisons of the subjective quality for *Lighthouse*; (a) original image, and (b) enlargement of original image, (c) method [3], (d) method [5], (e) the proposed method on spatial domain (f) the proposed method 1 (LF-HF domain), and (g) the proposed method 2 (edge-texture domain).

Next, the proposed algorithm is objectively evaluated with PSNR and SSIM using a variety of well-known images. As confirmed in Table. 1, it is clear that our method outperforms other methods in most of the cases in terms of PSNR, and all cases in terms of SSIM. Throughout the above qualitative and quantitative comparison, we can confirm the effectiveness of our proposed method.

TABLE I PSNR AND SSIM COMPARISION

Image	Method [3]	Method [5]	Method [4]	Proposed Method 1	Proposed Method 2
Child	36.73	36.38	36.83	37.77	37.64
	0.9627	0.9576	0.9579	0.9657	0.9652
Koala	32.68	33.32	32.94	33.65	33.63
	0.9283	0.9262	0.9203	0.9306	0.9302
Lena	34.81	35.72	35.22	35.98	36.04
	0.9352	0.9353	0.9356	0.9371	0.9372
Light	37.87	38.51	39.00	39.40	39.78
	0.9587	0.9604	0.9614	0.96212	0.9632
Window	29.34	29.15	29.27	29.46	29.56
	0.8934	0.8913	0.8931	0.8971	0.8996
Zebra	26.90	26.89	27.83	28.36	28.45
	0.9099	0.9099	0.9112	0.9203	0.9205

Fig. 5 Comparison of the subjective quality with various alpha for *Child*; First row show HF components and second row shows resultant images with alpha value 1, 1.6, 2 along each column.

Finally, we show the influence of alpha (combination parameter of LF and HF in (5)) from our proposed method on the subjective quality. In our experiments, we found the best value of alpha to be 1.6, which is visually most appealing for several test images. In Fig. 7, the value of alpha used for (a), (b) and (c) are 1, 1.6 and 2 respectively. (For a fair comparison of the three HF components, the three HF components are scaled in the identical manner.

IV. CONCLUSION

This paper presents a novel super-resolution method based on both LF-HF and edge-texture domains. The searching mechanism of existing SR methods exploits a spatial domain only, which may lead to mismatching or blurring in the texture region, or non-singularity region. To solve this problem, we decompose the input image into two different components, and reconstruct HR image by using selfsimilarity. We confirm that our proposed SR method is more effective in reconstructing HR image than conventional methods in terms of PSNR and SSIM as well as the subjective quality.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0009957)

References

- W.T. Freedman, T.R Jones, and E.C. Pasztor "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, pp. 56-65, Mar. 2002.
- [2] N. Suetake, M. Sakano and E. Uchino "Image super-resolution based on local self-similarity," *Optical Review*, vol. 15, pp. 26-30, Jan. 2008.
- [3] S. Chen, H. Gong, and C. Li, "Super resolution from a single image based on self-similarity," *IEEE Int. Conf. on Computational and Information Sciences*, pp. 91-94, Oct. 2011
- [4] D. Glasner, S. Bagon, and M.Irani, "Super-resolution from a single image," 12th IEEE Int. Conf. on Computer Vision, pp. 349 – 356, Sept. 2009.
- [5] G. Freeman, and R. Fattal, "Image and video upscaling from local self-examples," *ACM Trans. on Graphics*, vol. 30, Article No. 12, April. 2011.
- [6] C.Y Yand, J.B Huang, and M.H Yang, "Exploiting selfsimilarities for single frame super-resolution," *Proc. of Asian Conf. on Computer Vision*, pp. 497-510, Nov, 2010.
- [7] A. Buades, B. Coll and J.M Morel, "A non-local algorithm for image denoising", *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 60 – 65, June. 2005.
- [8] L. Wolf, X. Huang, I. Martin, and D. Metaxas. "Patch-based texture edge and segmentation," *Proc. of the 9th European Conf.* on Computer Vision, vol. 2, pp. 481-493, May. 2006.
- [9] S. Takahiro, and K. Takashi, "Image-processing approach based on nonlinear image decomposition," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E92A No. 3, pp. 696-707, March. 2009.