

LETTER

Middle-Frequency Based Refinement for Image Super-Resolution

Jae-Hee JUN[†], Ji-Hoon CHOI[†], Nonmembers, and Jong-Ok KIM^{†a)}, Member

SUMMARY This letter proposes a novel post-processing method for self-similarity based super-resolution (SR). Existing back-projection (BP) methods enhance SR images by refining the reconstructed coarse high-frequency (HF) information. However, it causes artifacts due to interpolation and excessively smoothes small HF signals, particularly in texture regions. Motivated by these observations, we propose a novel post-processing method referred to as middle-frequency (MF) based refinement. The proposed method refines the reconstructed HF information in the MF domain rather than in the spatial domain, as in BP. In addition, it does not require an internal interpolation process, so it is free from the side-effects of interpolation. Experimental results show that the proposed algorithm provides superior performance in terms of both the quantity of reproduced HF information and the visual quality.

key words: back-projection, middle-frequency, post processing, super-resolution

1. Introduction

Super-resolution (SR) is a technique used to reconstruct a high-resolution (HR) image from one or more low-resolution (LR) images. As such, it attempts to recover the lost high-frequency (HF) information from an LR image.

Recently, the learning-based SR approach has been a popular focus of research [1]–[5]. It demonstrates better performance as it accesses a larger volume of available prior information. It establishes an image database that includes LR and HR image pairs to infer HR information from corresponding LR input. In particular, an SR approach based on self-similarity has exhibited superior reconstruction performance with low complexity. It finds image prior examples within an image itself, without the need for an external database.

In the self-similarity SR approach, an input image is decomposed into low-frequency (LF) and high-frequency (HF) components. The input image is also up-scaled to HR using a simple interpolation method, and the resulting up-scaled version is regarded as the LF component of an HR image. For each patch in the up-scaled image, a similar one is looked for in the low-resolution LF component. Once the best LF match has been found, its associated HF patch is combined with the query LF patch and an HR image is produced.

The existing self-similarity SR approach excels at re-

covering lost HF information from an input LR image. However, the HF information of an HR image is recovered by simply copying the HF signals corresponding to a similar patch found in the LR image. Therefore, some discrepancies still exist between original and super-resolved images due to inaccurate HF estimation. The inconsistency is particularly severe in the weak HF region (which is defined by the lower half of the HF component, and will be referred to as middle-frequency (MF) throughout the letter), which commonly corresponds to texture in an image. In order to refine the coarse HF signals reconstructed by self-similarity SR, post-processing is required. Back-projection has been widely used as a post-processing technique [6].

Back-projection projects an HR image as an LR one, and works iteratively to minimize the errors between the input LR and the projection of the HR image [7], [8]. It further enhances the SR image by refining the reconstructed coarse HF. However, it excessively smoothes HF signals and causes blurring, especially in the texture regions in the HR image. Small signal variations such as texture are easily affected by back-projection because it is carried out on a large-scale spatial domain.

Motivated by these observations, we propose a novel post-processing method for self-similarity SR. The key goal of SR is to recover the lost HF information; thus, post-processing should specifically concentrate on HF signals. The proposed method aims to refine the reconstructed middle-frequency (MF) of the HF domain, unlike the spatial domain as in back-projection. In addition, it does not include the process of up-scaling, which is the reason for additional artifacts in back-projection. These two features lead to fewer blurring artifacts than back-projection does, while also improving the sharpness in the texture region of the image.

The rest of this letter is organized as follows. Section 2 presents the self-similarity SR method with iterative back-projection. The proposed MF-based refinement method is presented in Sect. 3. Section 4 presents the experimental results. Section 5 concludes the letter.

2. Self-Similarity Based SR with Back-Projection

In natural images, small patches tend to recur redundantly across scale as well as in-scale. This phenomenon is referred to as self-similarity, and has been exploited for learning-based image SR in recent years [9], [10]. Self-similarity based SR in the LF-HF domain is similar to learning-based

Manuscript received August 13, 2015.

Manuscript revised September 30, 2015.

Manuscript publicized October 14, 2015.

[†]The authors are with School of Electrical Engineering, Korea University, Seoul, Korea.

a) E-mail: jokim@korea.ac.kr

DOI: 10.1587/transinf.2015EDL8180

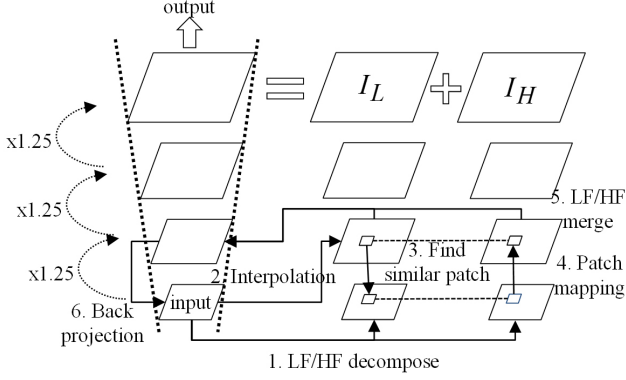


Fig. 1 The process of the self-similarity based SR algorithm.

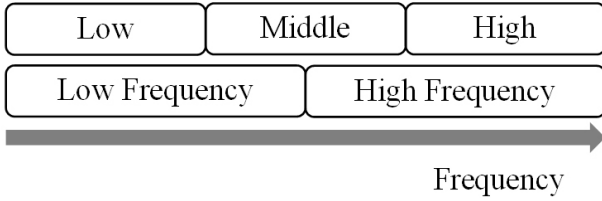


Fig. 2 The middle-frequency domain.

methods such as [11], [12]. However, unlike the traditional example-based approach, SR based on self-similarity does not require a prior database. Thus, it can reduce computational complexity and memory consumption in comparison to the conventional learning-based approach. The overall architecture of the self-similarity based SR method is shown in Fig. 1. In the first step, the input image is decomposed into both LF and HF components, which are given by

$$I_{LR,LF} = H(I_{LR}) \quad (1)$$

$$I_{LR,HF} = I_{LR} - I_{LR,LF} \quad (2)$$

where H is a blurring operator and $I_{LR,LF}$ is the LF component of the input LR image. After frequency decomposition, an initial estimation of the HR image is obtained by simply interpolating the input image. This HR image can be considered the LF component of an HR image, denoted as $I_{HR,LF}$. SR to a target scale is conducted repeatedly by a small scale factor (e.g., 1.25) until arriving at a target resolution. It has been previously reported that an incremental coarse-grained SR with a small scale factor produces a finer result [1].

Following this, for each image patch within the HR-LF image, we find the most similar patch within the LR-LF image, denoted as $I_{LR,LF}$. Once the most similar patch has been found, its HF pair is regarded as a potential estimate of the HR-HF image, denoted as $I_{HR,HF}$. By merging $I_{HR,LF}$ and $I_{HR,HF}$, we can reproduce a full HR image.

Finally, back-projection is carried out on the output HR image to compensate for reconstruction errors caused by the SR process. The super-resolved HR image is projected on the LR grid, and the difference between the LR projection of the HR and the original LR image is obtained by

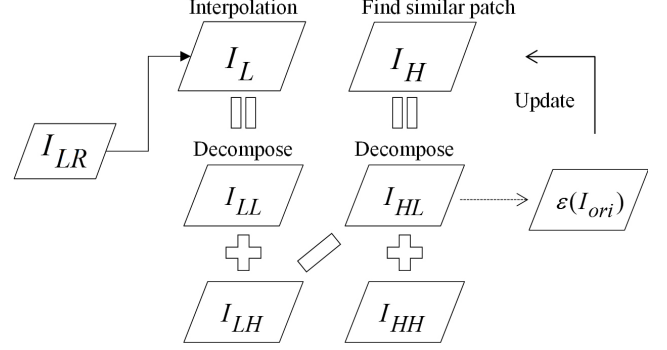


Fig. 3 The proposed MF-based refinement algorithm.

$$e_r(I) = I_{LR} - (I_{HR} * g) \downarrow s \quad (3)$$

where $e_r(I)$ is reconstruction error, g is a blurring filter, $*$ is the convolution operator, $\downarrow s$ is the down-sampling operator with scaling factor s , and I_{LR} and I_{HR} are the low-resolution and high-resolution images, respectively. Back-projection iteratively updates the HR image to minimize the errors in (3).

3. The Proposed Method

Back-projection is an efficient algorithm for strengthening the consistency of the reconstructed HR image in comparison to the original LR image. Despite this efficiency, however, it suffers from some fundamental side-effects such as smoothing, chessboard, and ringing artifacts, the latter two of which are mainly caused by interpolation in back-projection.

Back-projection hierarchically runs over the image pyramid for specific resolutions, so it requires up-scaling. It is the up-scaling process that induces the artifacts mentioned above, artifacts that are particularly severe in the MF texture regions of images. This provides the motivation to develop the proposed MF-based refinement method, whose key feature doesn't require the interpolation process. The proposed MF-based refinement method seeks to specifically improve the MF region. This is why more SR artifacts are observed in the texture region.

An image is decomposed into LF and HF components using typical loss pass filtering as in (1) and (2), and each component is decomposed into LF and HF components again. In other words, the original image is divided into four components as follows.

$$I_{LL} = I_L * g \quad (4)$$

$$I_{LH} = I_L - I_{LL} \quad (5)$$

$$I_{HL} = I_H * g \quad (6)$$

$$I_{HH} = I_H - I_{HL} \quad (7)$$

Of the four components in (4)-(7), both I_{LH} and I_{HL} overlap exactly between I_L and I_H , as illustrated in Fig. 2,

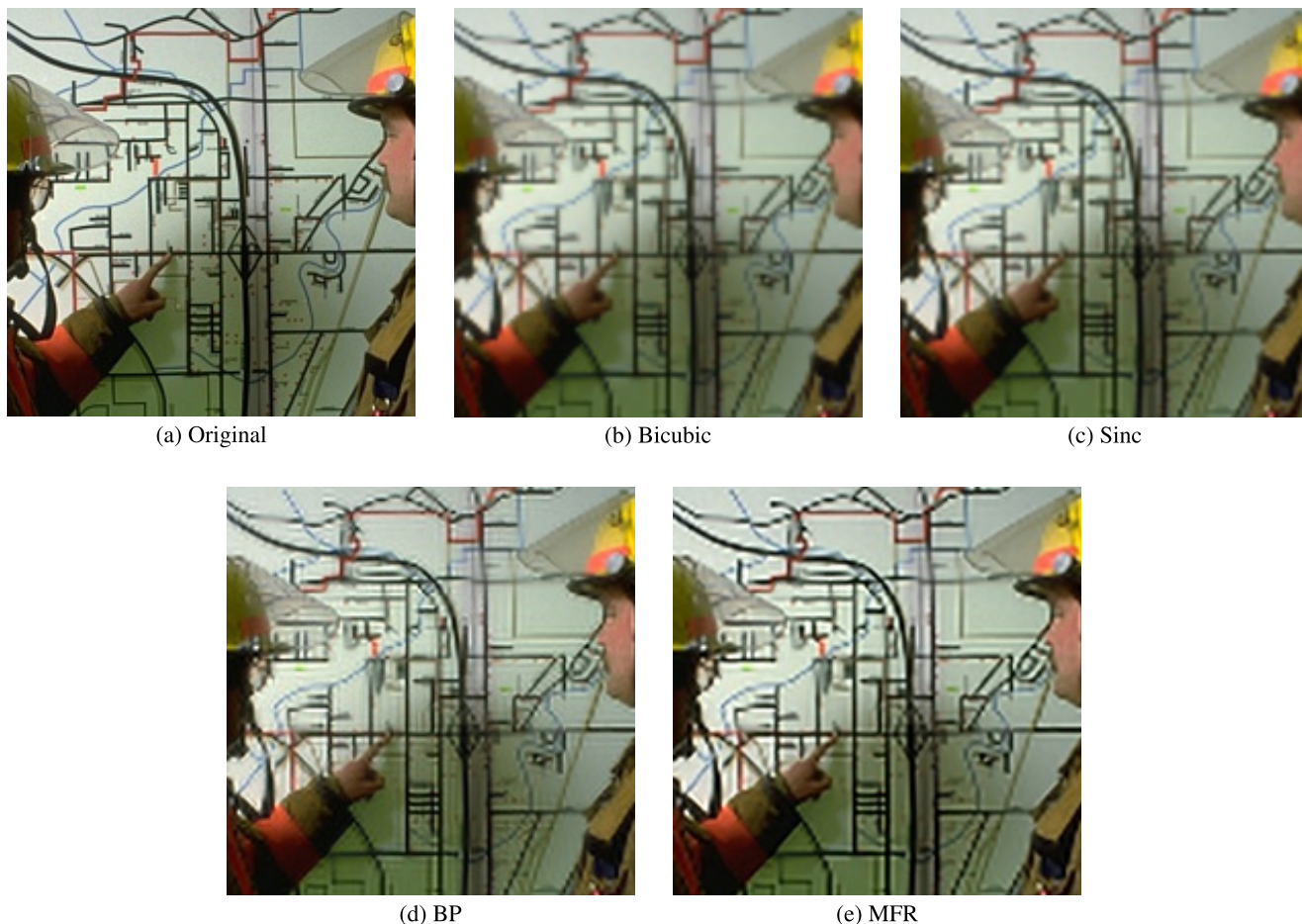


Fig. 4 Comparisons of subjective image quality for the test image, *workers*.

and they are equal to each other at a signal level. This overlapping region between I_L and I_H is defined as MF in this letter. This relationship can be proven as follows.

$$I_{LH} = I_L - I_{LL} = I_{ori} * g - I_{ori} * g * g \quad (8)$$

$$I_{HL} = I_H * g = I_{ori} * g - I_{ori} * g * g \quad (9)$$

It is assumed that the interpolated version of an LR image is equal to the I_L of an HR image. SR targets the recovery of I_H . The lower frequency signals (corresponding to I_{HL}) in the newly reconstructed I_H are refined using I_{LH} , which is regarded as the ground truth of an HR image.

The overall MF-based refinement process is illustrated in Fig. 3. The reconstruction errors produced by SR are defined by the difference between I_{LH} and I_{HL} . The newly estimated I_H is recursively compensated for in such a way that it should have a strong consistency with I_L in the overlapping frequency regions. This process can be expressed as

$$\varepsilon(I_{ori}) = (I_{LH} - I_{HL}) * p \quad (10)$$

$$I_H^{(t+1)} = I_H^{(t)} + \varepsilon(I_{ori}^{(t)}) \quad (11)$$

where p is a Gaussian kernel, the same as that used in back-projection.

4. Experimental Results

We evaluate the performance of the proposed method for various test images and present representative results in this section. The proposed MFR and BP methods are applied to the self-similarity based SR method shown in Fig. 1. Other common interpolation methods, such as Bicubic and Sinc interpolation, are also tested to allow for performance comparisons. The magnification factor is configured at 2 for all experiments. The test LR input image is obtained by applying a Gaussian blur kernel to the original HR image. Performance is assessed in terms of two quality factors: (1) the subjective and objective evaluation of the visual quality of the SR image, and (2) the quantitative measurement of the amount of reproduced HF information, because the ultimate goal of SR is to reconstruct lost HF information.

Firstly, the visual quality of the SR images is compared. Figures 4 and 5 show the resulting SR images of the proposed MFR and conventional Bicubic, Sinc, and BP methods. In Figs. 4 and 5, MFR and BP can reconstruct more realistic image details. It is because MFR and BP based on self-similarity can restore more accurate HF details than Bicubic and Sinc which are simple interpolation

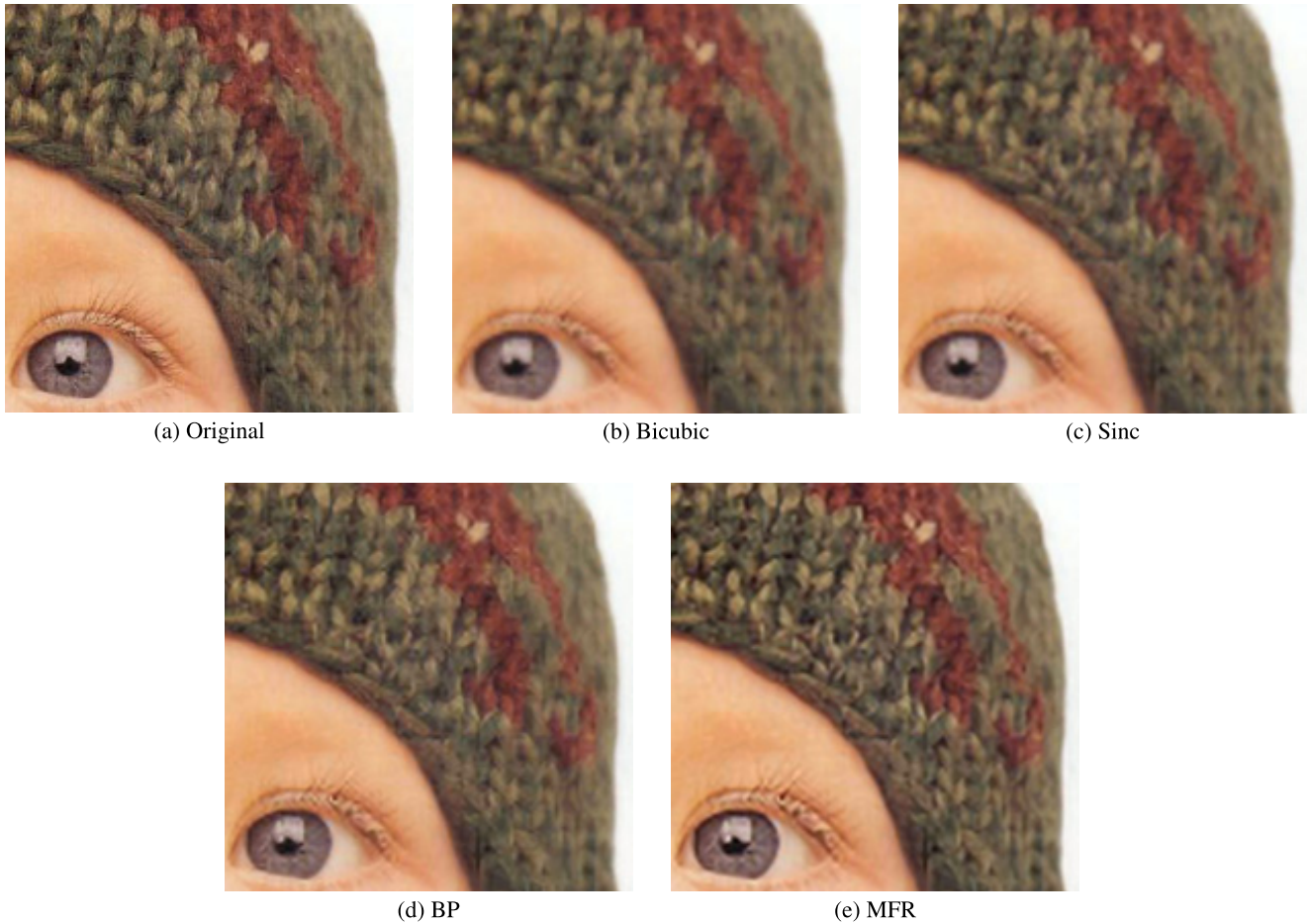


Fig. 5 Comparisons of subjective image quality for the test image, *child*.

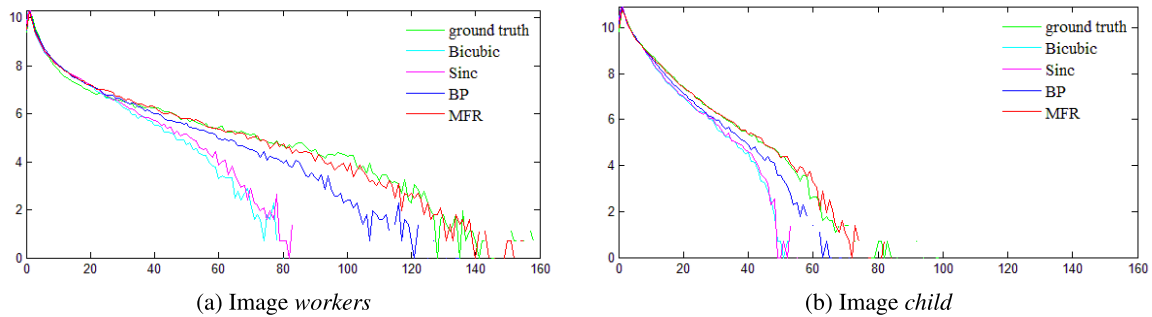


Fig. 6 Comparisons of gradient magnitude distributions.

methods. Between MFR and BP, MFR achieves a higher degree of sharpness (acutance) compared to BP. In particular, it improves the visual quality of texture region. For instance, Fig. 5 (e) is sharper than (d). Note that texture regions in the image usually lie in the MF domain. It is thus to be expected that the MF-based method is effective for texture regions. In addition, ringing artifacts are much more prevalent in conventional BP than in MFR (Figs. 4 and 5). MFR does not include interpolation process. The experimental results in Figs. 4 and 5 clearly confirm that the proposed MFR is effective in both improving texture quality and reducing the

side-effects of interpolation.

The visual quality is then measured numerically using PSNR and SSIM; the values are listed in Table 1. In general, the texture regions exhibit small, complex variations in terms of signal. It should be noted that it is difficult to quantitatively measure the sharpness of texture regions. However, although the benefits of the proposed method are not definitive in terms of PSNR, the SSIM values are superior for all of the test images.

Finally, we quantitatively evaluate the ability of the proposed middle-frequency refinement (MFR) post-

Table 1 The comparison of PSNR and SSIM

		<i>workers</i>	<i>church</i>	<i>flower</i>	<i>horse</i>	<i>child</i>
PSNR	Bicubic	24.47	29.09	30.51	28.31	34.29
	Sinc	24.60	29.31	30.77	28.66	34.54
	BP	24.92	30.31	31.25	29.37	34.94
	MFR	24.59	30.46	31.19	29.40	34.67
SSIM	Bicubic	0.8359	0.9268	0.9088	0.8623	0.9315
	Sinc	0.8418	0.9284	0.9148	0.8705	0.9351
	BP	0.8586	0.9330	0.9271	0.8834	0.9416
	MFR	0.8651	0.9345	0.9326	0.8883	0.9434

Table 2 Kullback-Leibler divergence values of gradient magnitude distributions

		<i>workers</i>	<i>church</i>	<i>flower</i>	<i>horse</i>	<i>child</i>
Value	Bicubic	0.8118	0.6974	0.3335	0.4673	0.1067
	Sinc	0.6805	0.5267	0.2707	0.3787	0.0947
	BP	0.0603	0.1297	0.0441	0.0673	0.0277
	MFR	0.0209	0.0610	0.0160	0.0048	0.0037

processing method to reproduce HF information. Gradient magnitude distributions are used to quantify the number of reconstructed HF signals. Figure 6 compares the plots of the gradient magnitude distributions for the test images of *workers* and *child*. When compared with conventional methods, the proposed method recovers more HF information, and is thus closer to the original HR-HF image. This can lead to improvement in SR image sharpness, as confirmed in the subjective quality comparisons in Figs. 4 and 5. The similarity between the gradient magnitude distributions is also calculated using Kullback-Leibler divergence (Table 2). For all test images, the K-L divergence of MFR is over 2 times smaller than that of other methods, indicating that the gradient distribution of MFR is more similar to the ground truth image.

5. Conclusion

This letter proposes a novel post-processing method for SR. The goal of SR is to reconstruct lost HF information accurately. In particular, it has been challenging to recover complex MF texture information in the SR process. This motivated the development of a refinement method for MF, which is based on the fact that the LF and HF regions in image decomposition partially overlap. The proposed method does not contain an interpolation step, unlike conventional BP, and thus does not suffer the side-effects of interpolation. Even though the proposed method is applied to self-similarity SR in this letter, it can be used for any SR method on the frequency domain.

Acknowledgements

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2015-H8501-15-1017) supervised by the IITP (Institute for Information & communications Technology Promotion).

This work was supported by the Brain Korea 21 Plus Project in 2015.

References

- [1] 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.
- [2] Q. Ning, K. Chen, L. Yi, C. Fan, Y. Lu, and J. Wen, "Image super-resolution via analysis sparse prior," Signal Processing Letters, IEEE, vol.20, no.4, pp.399–402, April 2013.
- [3] C. Kim, K. Choi, and J.B. Ra, "Example-based super-resolution via structure analysis of patches," Signal Processing Letters, IEEE, vol.20, no.4, pp.407–410, April 2013.
- [4] L. Guo, Y. Li, J. Yang, and L. Lu, "Exploration into Single Image Super-Resolution via Self Similarity by Sparse Representation," IEICE Trans. Inf. & Syst., vol.E93-D, no.11, pp.3144–3148, Nov. 2010.
- [5] J.-P. He, G.-D. Su, and J.-S. Chen, "Super-Resolution for Facial Images Based on Local Similarity Preserving," IEICE Trans. Inf. & Syst., vol.E95-D, no.3, pp.892–896, March 2012.
- [6] M. Irani and S. Peleg, "Motion analysis for image enhancement: resolution, occlusion and transparency," J. Visual Comm. Image Represent., vol.4, no.4, pp.324–335, Dec. 1993.
- [7] S. Dai, M. Han, Y. Wu, and Y. Gong, "Bilateral back-projection for single image super resolution," in Proc. Int. Conf. on Multimedia and Expo 2007, pp.1039–1042, July 2007.
- [8] W. Dong, L. Zhang, G. Shi, and X. Wu, "Nonlocal back-projection for adaptive image enlargement," in Proc. IEEE int. Conf. Image Process., pp.349–352, Cairo, Egypt, Nov. 2009.
- [9] G. Freeman and R. Fattal, "Image and video upscaling from local self-examples," ACM Trans. on Graphics, vol.30, Article No.12, April 2011.
- [10] N. Suetake, M. Sakano, and E. Uchino, "Image super-resolution based on local self-similarity," Optical Review, vol.15, no.1, pp.26–30, Jan. 2008.
- [11] W.T. Freedman, T.R. Jones, and E.C. Pasztor, "Example-based super-resolution," IEEE Computer Graphics and Applications, vol.22, no.2, pp.56–65, March 2002.
- [12] S.-J. Park, O.-Y. Lee, and J.-O. Kim, "Self-similarity based image super-resolution on frequency domain," Proc. of APSIPA ASC 2014, pp.1–4, Nov. 2013.