Adaptive Self-Similarity Based Image Super-Resolution Using Non Local Means

Ji-Hoon Choi *, Sae-Jin Park *, Dae-Yeol Lee†, Sung-Chang Lim†, Jin-Soo Choi† and Jong-Ok Kim *
* School of Electrical Engineering, Korea University, Seoul, Korea
† Realistic Broadcasting Media Research Department, Electronics and Telecommunications Research Institute (ETRI), Seoul, Korea
E-mail: {risty, bebequ, jokim}@korea.ac.kr Tel: +82-2-3290-3520
E-mail: {daelee711, sclim, jschoi}@etri.re.kr Tel: +82-42-860-6541

Abstract—Self-similarity based SR method on the LF-HF domain basically relies on the assumption that LF patch is highly correlated with HF patch. However, this assumption is content-specific significantly, and we can occasionally observe little correlation between LF and HF patch especially for texture region of natural images. In this paper we propose a new self-similarity based SR method to reflect this observation. There are two differences between the proposed and existing methods. First, HF details of target HR image are recovered by finding the similar patches on the MF domain in case of texture region. Second, the proposed method performs pixel-based reconstruction by adopting the concept of non-local means. Experimental results show that the proposed method can reconstruct more realistic image details.

I. INTRODUCTION
The goal of image super-resolution (SR) is to reconstruct a HR (high resolution) image from one or more observed LR (low resolution) images. As UHD (ultra high definition) TV appears recently, SR is receiving even more interest.

A couple of researches on self-similarity based image super-resolution have been done in recent years [1-6]. Self-similarity assumption is that small patches with similar structure in natural images tend to redundantly recur across different scales as well as in-scale [1]. This does not require any prior database (or examples) unlike the traditional example-based method [7]. Thus, it can reduce computational complexity and memory consumption, when compared to the conventional learning based approach.

SR methods based on self-similarity can be classified into two categories, depending on the domain to which the self-similarity assumption is applied; spatial domain and LF (low frequency)–HF (high frequency) domain. Some researches use MF (middle frequency) instead of LF. In [1], they search similar image patches in and across different scales on the image pyramid, on the spatial domain. Suetake [2] was the first to adopt the self-similarity concept to the LF-HF domain decomposition approach, not spatial. They decompose the input LR image into LF and HF components. Similar image patches are searched only in the LF domain, in the in-scale LF domain only. Once the best LF match is found, the HF patch corresponding to the best LF match is combined with the query LF patch for HR (high resolution) reconstruction. In [3], Chen extended the previous work [2] via searching across pyramid scales in addition to the original source image scale.

Self-similarity based SR method on the LF-HF domain basically relies on the assumption that HF patch is highly correlated with LF patch. However, this assumption is content-specific significantly, and we can occasionally observe little correlation between LF and HF patch in natural images [4]. This phenomenon occurs especially in texture region of image. This motivates us to investigate the correlation between MF and HF patches. The result of our initial experiments shows that the correlation between MF and HF patch is higher than the correlation between LF and HF patch in texture region.

In this paper we propose a new single SR method based on self-similarity. There are two differences between the proposed and existing methods. First, 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 either only finding the similar LF patches in the LF domain or only finding the similar MF patches in the MF domain. In the proposed method, however, the local image structure of the query patch determines the domain of the similar patches to be found. HF details of target HR image are produced by finding the similar MF patches in the MF domain in the case of texture region. Second, the proposed method performs pixel-based reconstruction by adopting the concept of non-local means [8]. In the existing approach HF of target HR image is produced every patch basis, but in the proposed method, the only center pixel of a patch is reconstructed by patch based processing. Above-mentioned two points are novel contributions in this paper.

The paper is organized as follows. Section II describes the proposed SR method. In Section III, performance evaluations are shown, and then, Section IV concludes the paper finally.

II. PROPOSED METHOD
In this Section, we first introduce the overall architecture of the proposed self-similarity SR method and next, the concept of non-local means is adopted for recovering HR pixels in the framework of self-similarity SR in Section II-A. In Section II-B, we describe the adaptive switching of domain on which self-similar patches are searched.
The overall architecture of the proposed self-similarity SR method is depicted in Fig. 1. An input LR image $I_0$ is upsampled by bi-cubic interpolator $U$. This initial up-scaled image is lacking in HF component (denoted by $H_1$), whose reconstruction is an ultimate goal of SR techniques. Thus, it is denoted by $L_0=U(I_0)$ which corresponds to the LF component of the image $I_1$. An input $I_0$ can be decomposed into $L_0$ and $H_0$ as follows.

$$L_0 = U(D(I_0))$$  \hspace{1cm} (1) \\
$$H_0 = I_0 - L_0$$  \hspace{1cm} (2)

where $D$ is a downsampling operator, $L_0$ is a smoothed version of $I_0$ which corresponds to the LF component of $I_0$ and $H_0$ indicates the HF component of $I_0$. The LF component in (1) is further decomposed into LF and HF, and the second LF component extracted from the first LF is defined by MF as follows.

$$M_0 = L_0 - U(D(L_0))$$  \hspace{1cm} (6)

For the reconstruction of $H_1$ in the existing method, for every patch $p_i$ in the LF image $L_1$, the most similar patch $p_j$ is searched on the LF domain, $L_0$ or for every patch $p_i$ in the MF image $M_1$, the most similar patch $p_j$ is searched on the MF domain, $M_0$. The corresponding $H_0$ patch is then copied to query patch $p_i$’s position as follows.

$$I_i(p_i) = L_i(p_i) + H_i(p_i)$$
$$H_i(p_i) = H_0(p_j)$$  \hspace{1cm} (3)

But in the proposed method, as shown in Fig. 1, the only center pixel of a patch is used for HF reconstruction and the domain on which self-similar patches are searched is adaptively switched.

As shown in Fig. 1, an input LR image is gradually up-scaled by a factor of 1.25 because self-similarity assumption holds better for small scaling factors [5]. The HR images, $I_2$ and $I_3$ are hierarchically reconstructed in the same manner as $I_1$. In this way, hierarchical HF reconstructions in the image pyramid are repeated until we arrive at the target HR. Then, we obtain the HR image by combining HF image into LF.

### A. NLM (Non-local means)

In the existing approach, the HF component of the target HR image is produced on a patch basis, but the proposed method performs pixel-based reconstruction by adopting the concept of non-local means. Namely, the only center pixel of a patch is used for actual HF reconstruction. HF pixels are restored by the weighted sum of the center pixels of similar patches and the weight is proportional to the similarity.

The NLM (non-local means) filter is simple yet effective for image denoising. This approach assumes that local image structures tend to repeat themselves within some similar neighbors across image pyramid [9]. This assumption is basically equal to self-similarity in SR methods. From this aspect, we borrow the NLM idea and it is directly applied to the self-similarity SR method as follows.

$$H_i(i) = \frac{\sum_{j \in \Omega} w_i^N H_0(j)}{\sum_{j \in \Omega} w_i^N}$$  \hspace{1cm} (4)

where

$$w_i^N = \exp(-\left\|L_i(p_i) - L_0(p_j)\right\|^2 / h^2)$$  \hspace{1cm} (5)

where $H_i(i)$ is the $i$-th reconstructed HF pixel, $H_0(j)$ is the $j$-th HF pixel of LR and $\Omega$ is an index set of similar patches.

The proposed method performs pixel-based reconstruction, and one may worry about algorithm complexity. But the complexity of the proposed method is almost equal to that of the existing approach because the HF patch in the existing method is reconstructed while shifted by a single pixel in a raster order.

### B. Mid Low Frequency

The self-similarity based SR method assumes that LF patch is highly correlated with HF patch. However, this assumption does not hold for all the regions in image, and the correlation between LF and HF patches is hardly observed in some regions of natural images [4]. This phenomenon typically
occurs especially for texture region of image. Motivated by this observation, we consider the correlation between MF and HF patches.

In Fig. 2, we compare HF patches searched on the LF domain with HF on the MF domain. As shown in Fig. 2, the HF patch searched on the MF domain is surely more similar to ground truth HF (Fig. 2 (c)) than that on the LF domain. Therefore, it is expected that the proposed method can restore more accurate HF details of target HR image by finding more similar MF patches because MF is more correlated with HF than LF. However, except for texture region (i.e., for edge and flat regions) LF has stronger correlation than MF oppositely. Thus, searching similar patches in edge and flat regions is performed on the LF domain. In other words, the search domain should be adaptively determined according to local image structure. For this adaptive operation, all pixels in an image should be classified into either texture or not.

To determine whether a pixel is on texture region or not, the texture segmentation method is needed. Chen in [10] proposed an edge indicator, which is so called difference curvature to use the second derivatives in the direction of both gradient and perpendicular to gradient. This indicator is adopted for our proposed method.

III. EXPERIMENTAL RESULTS

The proposed method is evaluated with various test images and some of representative results are shown in this Section.

In all experiments, the scaling factor at each hierarchical step is fixed by 1.25 to achieve a magnification factor of 2 in total as illustrated in Fig. 1. Thus, the process of finding a similar patch is repeated 3 times. The LF component of an image is obtained by bi-cubic up-scaling with a factor of 1.25. Patch size of LR and HR images is equally set to ‘5 by 5’. The SR algorithm is applied to Y channel only, and the other color channels are simply up-scaled by bi-cubic interpolation.

Figure 3 shows the texture segmentation of test image using difference curvature which is calculated for each pixel [10]. The segmentation results are shown in Fig. 3 (a) where white pixels are classified by texture and black pixels are by flat or edge. Due to pixel-based segmentation, there happen some holes as illustrated in Fig. 3 (a). For removing those holes, morphology filtering is done additionally, and the filtered version is shown in Fig. 3 (b).

Table I shows subjective quality comparisons of the reconstructed HR image with conventional methods. For performance comparisons, we consider two methods to reconstruct an HR pixel in the existing SR approach. One is to use the best similar patch only (denoted by ‘BP’), and the other method (denoted by ‘MP’) reconstructs HR by weighted sum of multiple similar patches.

Compared with the results generated by existing SR approach, the proposed method can reconstruct more realistic image details. The fur and the whiskers of the baboon in Fig. 4 (c) have more details than Fig. 4 (a) or Fig. 4 (b). It is because our proposed method can restore more accurate HF details than existing method, e.g., HF image in Fig. 4 (f) can produce HR image has more sharper edge and finer details than other HF images especially in texture region.

Next, we evaluated quantitatively the proposed method with PSNR and SSIM using a variety of well-known images. As confirmed in Table I, PSNR and SSIM of our method are higher a little than the existing methods. Although the objective quality gain is marginal, subjective quality is quite better as shown in Fig. 4. Recall that the proposed method primarily targets at improving texture visual quality during SR process. The objective quality measure such as PSNR and SSIM are not appropriate for texture region in particular, and we may rely on subjective quality measurement. Nevertheless, the proposed method achieves higher quality index for both PSNR and SSIM.

Finally, we evaluated the amount of the reconstructed high frequencies for test SR images in order to confirm that the proposed method can really recover more image details in

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>36.17</td>
<td>36.00</td>
<td>36.27</td>
</tr>
<tr>
<td></td>
<td>0.9510</td>
<td>0.9488</td>
<td>0.9531</td>
</tr>
<tr>
<td>Lenna</td>
<td>33.96</td>
<td>33.75</td>
<td>34.09</td>
</tr>
<tr>
<td></td>
<td>0.9016</td>
<td>0.8988</td>
<td>0.9023</td>
</tr>
<tr>
<td>Pepper</td>
<td>31.28</td>
<td>31.18</td>
<td>32.28</td>
</tr>
<tr>
<td></td>
<td>0.8950</td>
<td>0.8928</td>
<td>0.8939</td>
</tr>
<tr>
<td>Fruits</td>
<td>36.66</td>
<td>36.35</td>
<td>37.14</td>
</tr>
<tr>
<td></td>
<td>0.9546</td>
<td>0.9514</td>
<td>0.9589</td>
</tr>
<tr>
<td>Baboon</td>
<td>23.50</td>
<td>23.43</td>
<td>23.62</td>
</tr>
<tr>
<td></td>
<td>0.6924</td>
<td>0.6847</td>
<td>0.7115</td>
</tr>
</tbody>
</table>
texture region. The SR image is converted into frequency domain with 2D Fourier transform. We calculate the sum of all vertical frequency amplitudes for each horizontal frequency. Figure 5 shows the recovered high frequency characteristic for each method (black line – original, green line – BP, red line – MP, blue line – the proposed method). On the whole, the proposed method recovers much larger amount of high frequencies (image details) than conventional methods. Note that the frequency gain of the test image pepper is relatively marginal because it originally contains less image details.

IV. CONCLUSIONS

In this paper, we proposed a novel self-similarity based SR method. In the SR approach to exploit the property of self-similarity, the most important thing is to search self-similar signals for HR reconstruction. We apply the concept of non-local means which performs pixel-based reconstruction, unlike the existing patch based one. In addition, we use the adaptive switching between LF and MF domain for searching similar patches instead of using only LF or MF domain in order to find more accurate HF details. This is particularly effective for texture region in an image. Experimental results confirm that the proposed method can recover more realistic image details.

ACKNOWLEDGMENT

This work was supported by the ICT R&D program of MSIP/IITP. [13-912-02-002, Development of Cloud Computing Based Realistic Media Production Technology]
REFERENCES


