

# Single-Image Super-Resolution Using Sparsity Constraints and Non-Local Similarities at Multiple Resolution Scales

Hiêp Q. Luong, Tijana Ruzic, Aleksandra Pižurica and Wilfried Philips

Ghent University, TELIN-IPI-IBBT, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

## ABSTRACT

Traditional super-resolution methods produce a clean high-resolution image from several observed degraded low-resolution images following an acquisition or degradation model. Such a model describes how each output pixel is related to one or more input pixels and it is called data fidelity term in the regularization framework. Additionally, prior knowledge such as piecewise smoothness can be incorporated to improve the image restoration result. The impact of an observed pixel on the restored pixels is thus local according to the degradation model and the prior knowledge. Therefore, the traditional methods only exploit the spatial redundancy in a local neighborhood and are therefore referred to as local methods.

Recently, non-local methods, which make use of similarities between image patches across the whole image, have gained popularity in image restoration in general. In super-resolution literature they are often referred to as exemplar-based methods. In this paper, we exploit the similarity of patches within the same scale (which is related to the class of non-local methods) and across different resolution scales of the same image (which is also related to the fractal-based methods). For patch fusion, we employ a kernel regression algorithm, which yields a blurry and noisy version of the desired high-resolution image. For the final reconstruction step, we develop a novel restoration algorithm. The joint deconvolution/denoising algorithm is based on the split Bregman iterations and, as prior knowledge, the algorithm exploits the sparsity of the image in the shearlet-transformed domain. Initial results indicate an improvement over both classical local and state-of-the-art non-local super-resolution methods.

**Keywords:** Super-resolution, non-local similarities, sparsity, Bregman iterations, kernel regression

## 1. INTRODUCTION

Since the early 1980s, the idea of self-similarity plays an important role in mathematics and physics. Fractal-based methods assume that many natural objects possess fractalness, i.e., parts of the image repeat themselves on an ever-diminishing scale, hence the term self-similarity. Thanks to the scale-invariance property of the fractal-based methods, self-similarity lends itself to image resolution enhancement applications. In practical applications, the self-similarity property is exploited via a contractive affine transformation of image blocks across different scales.<sup>1</sup>

Unlike fractal-based methods, non-local methods exploit the similarity of small patches at the same scale, without rotation or photometrical corrections. Repetitive structures could be regarded as multiple noisy observations of the same structure. The concept of repetitive structures was successfully introduced in the image denoising field as the non-local means algorithm by Buades et al.<sup>2,3</sup> Recently, a few single-image super-resolution techniques have been proposed based on this concept.<sup>4-6</sup>

Closely related to the non-local method, is the training-based (or learning-based) approach, which restores images based on image blocks or *examples* that are retrieved from ideal images. Therefore, these techniques are also referred to as *exemplar-based* methods. Codebooks with examples are built by applying the degradation model on these ideal (degradation-free) images. These image blocks and their ideal counterparts are then used to guide the restoration process.<sup>7-9</sup> If the learning process only takes the input image into account, we refer them to as *self-examples*. In the application of super-resolution, self-examples come from other scales of the same image, which makes the technique on its turn related to the fractal-based methods. This concept is also exploited in single-image super-resolution frameworks.<sup>6,10</sup>

---

Further author information: (Send correspondence to H. Q. Luong)

H. Q. Luong: E-mail: [hiiep.luong@telin.ugent.be](mailto:hiiep.luong@telin.ugent.be), Telephone: +32 (0) 9 264 79 66

T. Ruzic: E-mail: [tijana.ruzic@telin.ugent.be](mailto:tijana.ruzic@telin.ugent.be), Telephone: +32 (0) 9 264 34 16

In this paper, we will adopt the idea of self-examples in our framework and we exploit the similarity of patches within the same scale (which is related to the class of non-local methods) and across different resolution scales of the same image (which is also related to the fractal-based methods). For patch fusion, we employ a kernel regression algorithm, which yields a blurry and noisy version of the desired high-resolution image. For the final reconstruction step, we develop a novel restoration algorithm based on the split Bregman iterations and, as prior knowledge, the algorithm exploits the sparsity of the image in the shearlet-transformed domain. The main difference between our approach and the work of Ebrahimi and Vrscay,<sup>6</sup> is that we do not limit ourselves to a single scale (which is the same as the desired magnification factor). Often interesting scale-redundant properties do not occur at a single or particular scale. Glasner et al.<sup>10</sup> only takes the best match into account while we work with all similar patches. Another drawback is that they effectively build a Gaussian pyramid in their framework such that some possible matches can be lost, while we keep a low-resolution representation for each position of the input image. Also our proposed fusion and deconvolution components are quite different compared to their work.

The paper is organized as follows: in Section 2, a short introduction is given on the non-local means concept. In Section 3 our single-image super-resolution algorithm is explained. Finally, we present the experimental results in Section 4 and a conclusion in Section 5.

## 2. NON-LOCAL MEANS CONCEPT

The motivation to develop non-local methods is to exploit similar patterns and structures in an image. This relatively new class of denoising methods originates from the non-local means<sup>2,3</sup> which we now will describe very briefly.

Consider a noisy image  $\mathbf{v}$  on a discrete grid  $I$  (which is typically two or three dimensional):

$$\mathbf{v} = \{v_i | i \in I\} \quad (1)$$

where  $v_i$  denotes the intensity of the pixel at position  $i$  in the image. The denoised intensity  $\hat{v}(i)$  of a pixel  $i$  is computed as a weighted average of all the pixels in the image, which can be seen as a linear spatially adaptive filter:

$$\hat{v}_i = \frac{\sum_{j \in I} w_{i,j} v_j}{\sum_{j \in I} w_{i,j}}. \quad (2)$$

Let  $\mathbf{N}_i$  denote the vector of pixel intensities of a square window of size  $B \times B$  (e.g. obtained using the column-stack ordering, i.e. the columns of each image or window are stacked into one vector), centered at position  $i$ . The weights  $w_{i,j}$  in equation (2) depend on the similarity between  $\mathbf{N}_i$  and  $\mathbf{N}_j$ . The similarity between neighborhoods is computed by means of a *mean squared difference* (MSD)  $\|\mathbf{N}_i - \mathbf{N}_j\|^2$ , i.e. the Euclidean distance. A particular choice of weighting function is given by:

$$w_{i,j} = e^{-\frac{\|\mathbf{N}_i - \mathbf{N}_j\|^2}{h^2}} \quad (3)$$

where  $h$  is a constant, proportional to the noise variance. It turns out that this weighting function appears to be the Le Clerc robust loss function. Substantial improvements in quality can be obtained by replacing the weighting function with other robust loss functions.<sup>11</sup> The weights  $w_{i,j}$  decay at an exponential rate, which results in large weights for a small MSD (similar windows) and small weights for a large MSD (non-similar windows). We refer the interested reader to our previous work<sup>11,12</sup> on ways to improve the non-local means algorithm, both in computational complexity and image quality. In the next Section, we will generalize this concept to image zooming and use this as a component of our single-image super-resolution framework.

## 3. PROPOSED SINGLE-IMAGE SUPER-RESOLUTION FRAMEWORK

We propose a single-image super-resolution method which exploits repetitive structures within and across different scales. The proposed scheme is illustrated in Figure 1 and consists of three consecutive steps. The following sections will discuss each component carefully.

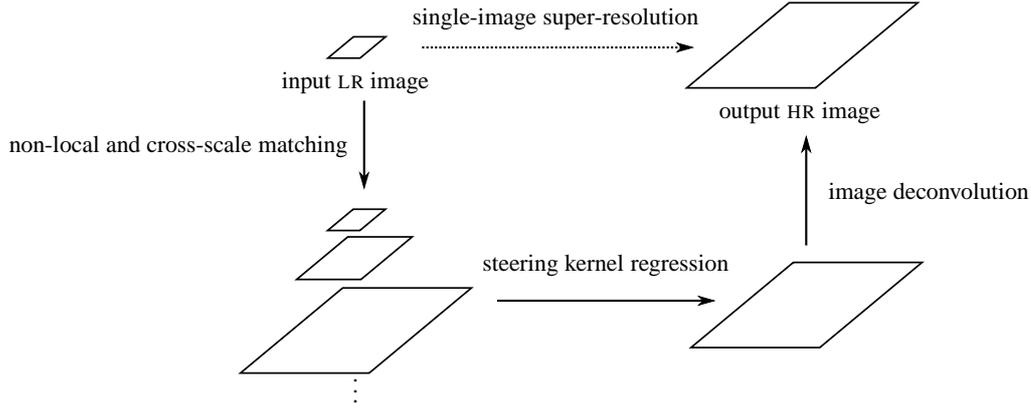


Figure 1. The proposed single-image super-resolution framework consisting of three consecutive steps.

### 3.1 Extension of non-local means to zooming

Similar to the work of Ebrahimi and Vrscay,<sup>6</sup> we extend the non-local means to image zooming. The high-resolution (HR) image is constructed as a weighted average of image patches, similar to the filter in equation (2) with the difference that  $k \times k$  image patches ( $k$  being the magnification factor) are used instead of pixels:

$$\hat{\mathbf{v}}_i = \frac{\sum_{j \in I} w_{i,j} \mathbf{v}_j}{\sum_{j \in I} w_{i,j}}, \quad (4)$$

where  $i$  are the positions of the non-overlapping  $k \times k$  patches in the HR image, which correspond to each pixel in the original input image. The grid  $I$  represents a search space consisting of Gaussian prefiltered and subsampled versions of the original input image (referred to as the low-resolution (LR) version), where all possible subsampling schemes are taken into account. In that way, we do not discard any possible matches in contrast to the work of Glasner et al.,<sup>10</sup> where they explicitly use a Gaussian pyramid.

The search unit is a  $B \times B$  neighborhood. Therefore, for each pixel at position  $i$  in the original input image, its surrounding  $B \times B$  neighborhood  $\mathbf{N}_i$  is compared with all possible neighborhoods  $\mathbf{N}_d$  in the search space  $I$ . The weighting function  $w_{i,j}$  is computed as:

$$w_{i,j} = e^{-\frac{\|\mathbf{N}_i - \mathbf{N}_{d,j}\|^2}{h^2}} \quad (5)$$

where  $j \in I$  denotes the central position of the matched neighborhood. Finally, the  $k \times k$  patch  $\mathbf{v}_j$  is extracted from the original input image as a *parent* patch corresponding to the location  $j$ . Therefore, each  $k \times k$  patch in the HR image is obtained as a weighted average of parent patches which correspond to the central positions of the matched neighborhoods upon which weights are computed. The cross-scale relationships are illustrated in Figure 2. The non-local means algorithm, described in the previous section, is actually a special case of this zooming process: in case that  $k = 1$  (i.e., the same scale as the input image), we perform non-local image denoising and the parent patch is simply the central pixel of the matched neighborhood itself.

An improvement of image quality towards the non-local algorithm is to ignore the contributions from dissimilar neighborhoods. Even though their weights are very small at first sight, the estimated pixel /patch values can be severely biased due to many small contributions. This bad influence of dissimilar neighborhoods can be eliminated by setting their corresponding weights to zero.

Our implementation involves non-local matching within the same scale and cross-scale matching simultaneously, i.e. across different scales ( $s = 1, \dots, S$ ). In this way we are more likely to find similar and relevant structures than just by searching at scale  $s = k$  that corresponds to the magnification factor (as in the work of Ebrahimi and Vrscay<sup>6</sup>).

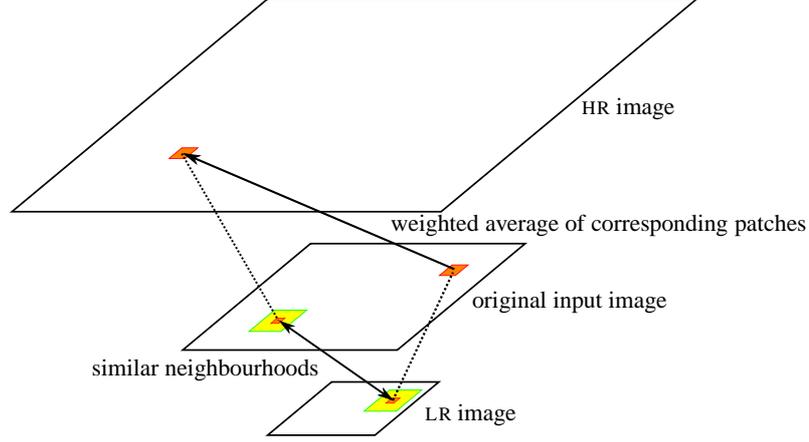


Figure 2. The process of cross-scale matching of similar neighborhoods (yellow) and the use of the corresponding image patch (orange) for the weighted average.

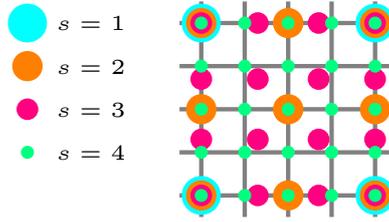


Figure 3. Illustration of the non-uniformly distributed samples for different scales  $s$ . The grid of the target image with magnification factor 4 of is also illustrated as the underlying grey lines.

A naive implementation of searching for similar patches will lead to an excessive computation time. An important improvement is the use of extension of the moving averaging filter for the MSD calculations which gives us a speed-up of a factor hundred. However, the issue of speed was not our primal concern in this paper. For more details and other improvements on computational complexity we refer the reader to our previous work.<sup>11,12</sup>

### 3.2 Patch fusion through steering kernel regression

The non-local and cross-scale matching is applied for different scales ( $s = 1, \dots, S$ ) resulting in multiple image data on different Cartesian grids. If we put all these grids on top of each other, the image data is distributed in a non-uniform way as illustrated in Figure 3 for magnification factor of 4 and  $S = 4$ . This problem also occurs in classical super-resolution frameworks and many solutions have been proposed in literature such as non-uniform resampling, Delaunay triangulation, POCS-based, shift-and-add algorithms, etc. In this paper, we will focus on the powerful kernel regression techniques.

We briefly describe the kernel regression method for solving the resampling problem in the ordinary least square sense as proposed by Takeda et al.<sup>13</sup> Suppose that we have to estimate the pixel value  $f(\mathbf{x})$  at position  $\mathbf{x}$  on the HR grid. In the surrounding neighbourhood, we have a set of  $p$  noisy measurements  $g_i$  at irregularly sampled positions  $\mathbf{x}_i$ . The data measurement model is then given by:

$$g_i = f(\mathbf{x}_i) + n_i, \quad i = 1, \dots, p, \quad (6)$$

where  $f(\cdot)$  is the unknown HR image, which is also referred to as the *regression function*, and  $n_i$  are independently and identically distributed zero-mean noise values. In a local neighbourhood, we can approximate the regression function  $f(\cdot)$  by its local second order Taylor's series expansion ( $N = 2$ ), which is denoted by:

$$\begin{aligned} f(\mathbf{x}_i) &\approx f(\mathbf{x}) + \{\nabla f(\mathbf{x})\}^T (\mathbf{x}_i - \mathbf{x}) + \frac{1}{2} (\mathbf{x}_i - \mathbf{x})^T \{\mathcal{H}f(\mathbf{x})\} (\mathbf{x}_i - \mathbf{x}) \\ &\approx \beta_0 + \beta_1^T (\mathbf{x}_i - \mathbf{x}) + (\mathbf{x}_i - \mathbf{x})^T \beta_2 (\mathbf{x}_i - \mathbf{x}), \end{aligned} \quad (7)$$

where  $\nabla$  and  $\mathcal{H}$  are the gradient and Hessian operators, respectively. The coefficients of this polynomial are estimated by the following weighted least-squares optimization problem ( $\hat{\beta} = [\beta_0 \ \beta_1 \ \beta_2]$ ):

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^p \left[ g_i - \beta_0 - \beta_1^T (\mathbf{x}_i - \mathbf{x}) - (\mathbf{x}_i - \mathbf{x})^T \beta_2 (\mathbf{x}_i - \mathbf{x}) \right]^2 k_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x}), \quad (8)$$

which can be solved easily using Gauss-Newton algorithms.  $\hat{f}(\mathbf{x}) = \hat{\beta}_0$  is the required estimated pixel value at the position  $\mathbf{x}$  on the HR grid. This estimator also models edges, ridges and blobs very well in opposite to most resampling techniques. The kernel function  $k_{\mathbf{H}}(\cdot)$  (which is typically of Gaussian or exponential form) penalizes distances further away from the grid position and its strength is controlled by the smoothing matrix  $\mathbf{H}$ :

$$k_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x}) = |\mathbf{H}|^{-1} k(\mathbf{H}^{-1} [\mathbf{x}_i - \mathbf{x}]). \quad (9)$$

In most applications, the  $2 \times 2$  smoothing matrix  $\mathbf{H}$  is equal to  $h\mathbf{I}$  with  $h$  being the bandwidth parameter such that the kernel's footprint is isotropic. Adapting the kernel's footprint locally according to the samples can prevent oversmoothing across edges in the same way as the existing geometry-driven tools. Therefore, the use of geometry-driven kernel functions (with anisotropic footprints) is referred to as *steering kernel regression*. Such a modification requires at least two iterations in the estimation algorithm: in the first step, the construction of such oriented kernels employs the first-order derivatives of the unknown HR image, which can be estimated from the non-uniformly distributed samples via equation (8) as  $\hat{\beta}_1$ . In the subsequent steps, the desired pixel value  $\hat{\beta}_0$  and the refined first-order derivatives  $\hat{\beta}_1$  are estimated with the newly modified smoothing matrices. For a more detailed discussion of steering kernel regression, we refer the interested reader to the paper of Takeda et al.<sup>13</sup> Another important benefit of kernel regression is that we can upscale the image with any magnification factor independent from the input of the non-local zooming.

### 3.3 Shearlet-based deconvolution

#### 3.3.1 Problem formulation

The most common model for the image deconvolution problem is given by the following matrix-vector formulation:

$$\mathbf{v} = \mathbf{A}\mathbf{u} + \mathbf{n}, \quad (10)$$

where  $\mathbf{v}$ ,  $\mathbf{u}$ ,  $\mathbf{n} \in \mathbb{R}^n$  are the observed blurry data, the ideal (or hypothetical desired) and additive noise images respectively in a column-stack ordering (i.e. the columns of each image are stacked into one vector).  $\mathbf{A}$  denotes the blur operator in the matrix-vector notation.

For the deblurring application, we consider a constrained optimization problem of the analysis-based form (also known as the *basis pursuit* problem<sup>14</sup>):

$$\arg \min_{\mathbf{u}} J(\mathbf{u}) \quad \text{s.t.} \quad H(\mathbf{u}) = 0, \quad (11)$$

or

$$\arg \min_{\mathbf{u}} J(\mathbf{u}) \quad \text{s.t.} \quad H(\mathbf{u}) < \sigma, \quad (12)$$

where the data fidelity is denoted by  $H(\mathbf{u}) = \frac{1}{2} \|\mathbf{A}\mathbf{u} - \mathbf{v}\|^2$  (which is convex, differentiable, and  $\min H(\mathbf{u}) = 0$ ).  $J$  represents the sparsity prior associated with the transform under which the image has a sparse representation or approximation. It is well-known that this problem is difficult to be solved numerically when  $J$  is non-differentiable. For the basis pursuit problem in this paper, we take the convex  $l_1$ -norm ( $|\cdot|$ ) regularization to promote sparsity in the transformed domain:

$$J(\mathbf{u}) = |\mathbf{S}\mathbf{u}|, \quad (13)$$

where  $\mathbf{S}$  denotes the sparsifying transform or the analysis operator. Popular choices for this transform are discrete gradient operators (e.g. total variation), wavelet or framelet transforms (e.g. the Besov norm with respect to the Haar wavelet transform), discrete Fourier transforms or local cosine transforms. In this paper, we apply the discrete shearlet transform because the shearlet transform provides both a multiresolution analysis (such as the wavelet transform), and at the same time an optimally sparse image-independent representation for images containing edges.<sup>15</sup> For implementation issues on the shearlet transform with a low redundancy factor, we refer the reader to our previous work.<sup>16</sup>

### 3.3.2 Split Bregman iterations

We wish to find a solution  $\hat{\mathbf{u}}$  for the constrained convex minimization problem stated in (11). The Lagrangian formulation of this constrained problem leads to the following unconstrained problem:

$$\arg \min_{\mathbf{u}} J(\mathbf{u}) + \lambda H(\mathbf{u}), \quad (14)$$

where  $\lambda$  is called the Lagrange multiplier or regularization parameter. A common way to reach the solution  $\hat{\mathbf{u}}$  for the problem (11) is to apply continuation where we solve sequentially the unconstrained problems (14) by performing gradient descent steps with varying regularization parameters  $\lambda_0 < \lambda_1 < \dots < \lambda_n \rightarrow +\infty$ . However because the regularization parameter tends to infinity or is very large in order to enforce that  $H(\mathbf{u}) \approx 0$ , continuation results in (14) having numerical instabilities for many applications.<sup>17</sup>

The idea of the Bregman iteration is also to transfer the constrained problem (11) into a sequence of unconstrained ones (14). But instead of varying the regularization parameter as in continuation methods, the Bregman iteration fixes this parameter and varies the observation data ( $\mathbf{v}_i$ ) in an iterative way. Recently, a variant on Bregman iterations was proposed by Goldstein and Osher.<sup>17</sup> By introducing an intermediate variable  $\mathbf{d}$  such that  $\mathbf{S}\mathbf{u} = \mathbf{d}$ , the terms in the prior  $J(\mathbf{u})$  and data fidelity  $H(\mathbf{u})$  are separable, and hence, the problem becomes more easy to minimize. We briefly summarize the iterative deconvolution algorithm (for more details on split Bregman iterations, we refer to the paper of Goldstein and Osher<sup>17</sup>), which is composed of four steps for each iteration:

$$\begin{cases} \mathbf{u}_{i+1} = \arg \min_{\mathbf{u}} \frac{\lambda}{2} \|\mathbf{A}\mathbf{u} - \mathbf{v}_i\|^2 + \frac{\mu}{2} \|\mathbf{d}_i - \mathbf{S}\mathbf{u} - \mathbf{b}_i\|^2 \\ \mathbf{d}_{i+1} = \arg \min_{\mathbf{d}} |\mathbf{d}| + \frac{\mu}{2} \|\mathbf{d} - \mathbf{S}\mathbf{u}_{i+1} - \mathbf{b}_i\|^2 \\ \mathbf{b}_{i+1} = \mathbf{b}_i + \mathbf{S}\mathbf{u}_{i+1} - \mathbf{d}_{i+1} \\ \mathbf{v}_{i+1} = \mathbf{v}_i + \mathbf{v} - \mathbf{A}\mathbf{u}_{i+1}. \end{cases} \quad (15)$$

The first step (which solely consists of  $l_2$ -terms) can simply be solved using Fourier techniques. The solution to the minimization with respect to  $\mathbf{d}$  in the second step is given by for example proximity operators. The proximity operator has well-known (closed-form) solutions for some models, we refer the interested reader to the work of Combettes<sup>18</sup> for a discussion on the proximal theory. In this paper, the regularization functional  $J$  is the  $l_1$ -norm, for which the solution is obtained by a soft thresholding operator:

$$\mathbf{d}_{i+1} = \Phi(\mathbf{S}\mathbf{u}_{i+1} - \mathbf{b}_i, \frac{1}{\mu}) \quad \text{where} \quad \Phi(\mathbf{u}, \gamma) = \frac{\mathbf{u}}{|\mathbf{u}|} \max(|\mathbf{u}| - \gamma, 0). \quad (16)$$

The outcomes of the third and fourth step of algorithm (15) can be computed in a straightforward way.

## 4. EXPERIMENTAL RESULTS

In the first experiment, we demonstrate the potential of single-image super-resolution. In Figure 4, we show the  $2 \times$  enlargement of a small window for the bicubic interpolation, the single-image super-resolution scheme of Glasner et al.<sup>10</sup> and our own super-resolution scheme. We can clearly notice that both super-resolution schemes are able to recover the fine rails in the intermediate arched windows, while that is not possible using interpolation schemes.

As a second experiment, we create a synthetic  $256 \times 256$  LR image from the  $1024 \times 1024$  *pentagon* standard test image. The HR image is first convolved with a Gaussian blur kernel before the subsampling. We enlarge the LR image with a magnification factor of 4. Figure 5 shows the enlarged results from the *pentagon* image for several techniques. We compare the images of the proposed single-image super-resolution method to those of the linear Blackman-Harris windowed sinc interpolation, the *iterated function systems* (based on fractals) upscaling<sup>19</sup> and the non-local interpolation.<sup>4</sup>

We notice that our super-resolution method produces less disturbing artefacts compared to the sinc interpolation and iterated function systems. The straight edges in the image are reconstructed much better with the proposed and non-local method: almost all annoying staircasing effects are removed. If we compare our proposed method to the non-local method, we can clearly notice more enriched texture at some places (e.g. at the left part of the image). Unfortunately, it does not always mean that the real high frequencies are restored (see groundtruth).

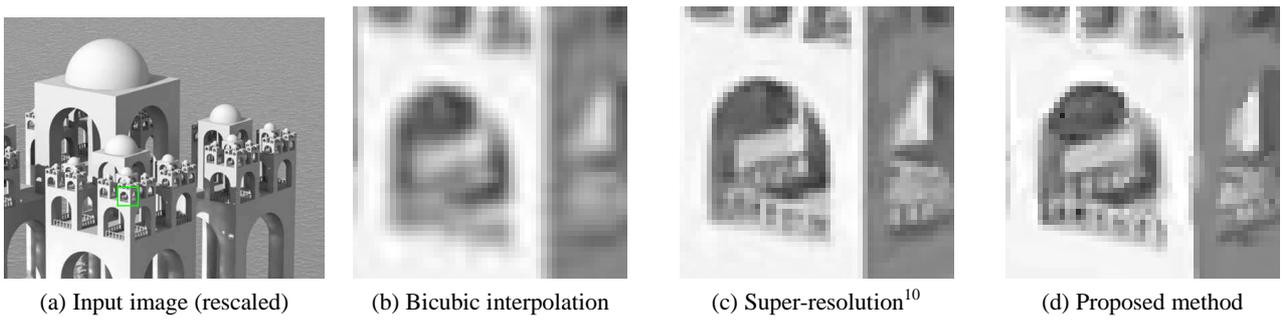


Figure 4. Towers image upscaling results (2× enlargement).

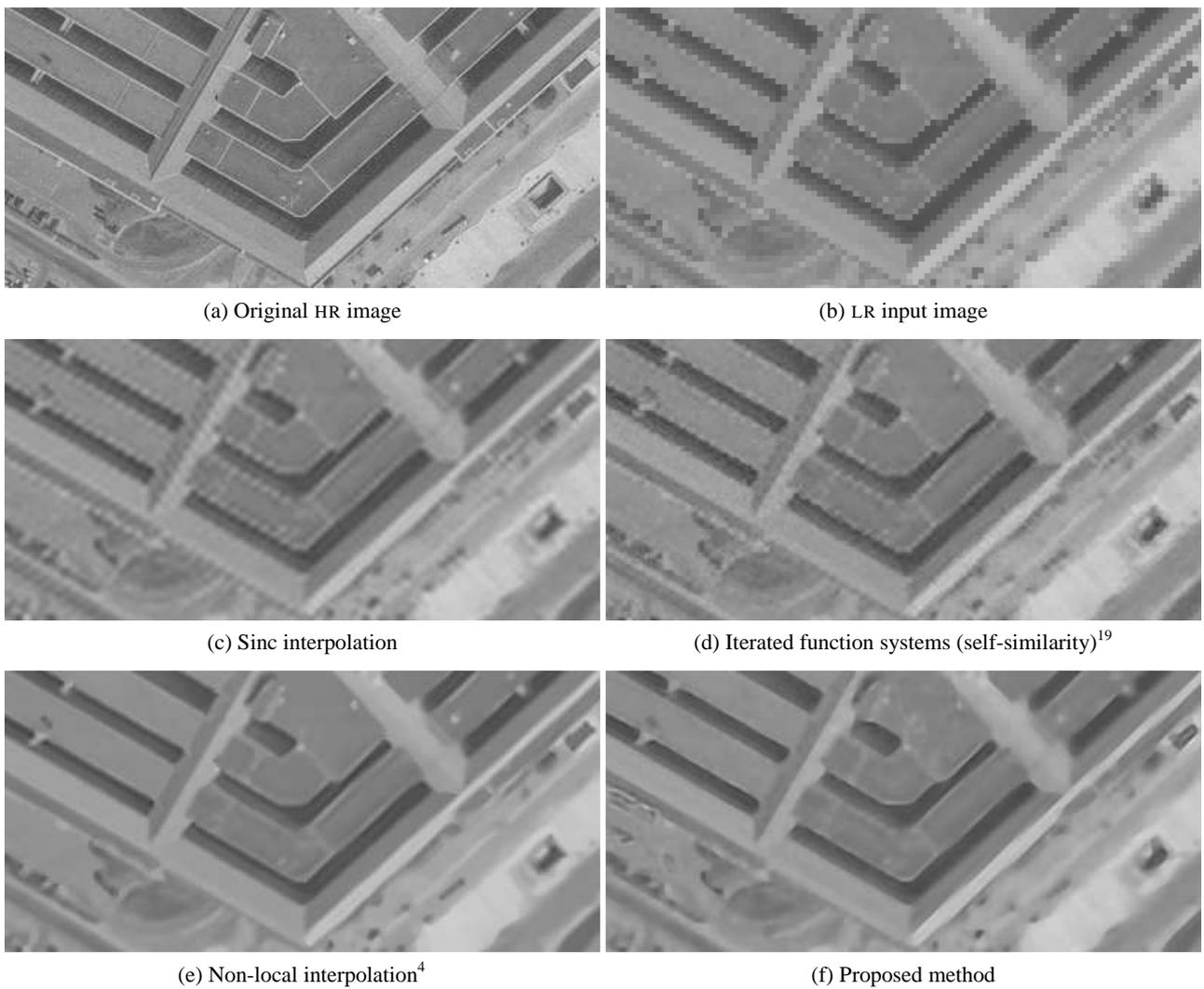


Figure 5. Pentagon image upscaling results (4× enlargement).

## 5. CONCLUSION

In this paper we have presented a novel single-image super-resolution scheme that exploits non-local and cross-scale redundancy in the image. The different image patches are fused via steering kernel regression and the output image is then obtained via a novel shearlet-based deblurring algorithm. Results show the effectiveness of our proposed image resolution enhancement technique. The parameter selection is an important issue which has not been addressed in sufficient detail in this paper, but is a part of our ongoing research.

## REFERENCES

- [1] Jacquin, A., “Image coding based on a fractal theory of iterated contractive image transformations,” *IEEE Transactions on Image Processing* **1**, 18–30 (Jan. 1992).
- [2] Buades, A., Coll, B., and Morel, J., “A review of image denoising algorithms, with a new one,” *SIAM Interdisciplinary Journal: Multiscale Modeling and Simulation* **4**(2), 290–530 (2005).
- [3] Buades, A., Coll, B., and Morel, J., “A non local algorithm for image denoising,” in [*Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*], **2**, 60–65 (2005).
- [4] Luong, H., Ledda, A., and Philips, W., “An image interpolation scheme for repetitive structures,” in [*Proceedings of International Conference On Image Analysis And Recognition (ICIAR)*], *Lecture Notes in Computer Science* **4141**, 104–115, Springer-Verlag (2006).
- [5] Luong, H., Ledda, A., and Philips, W., “Non-local image interpolation,” in [*Proceedings of IEEE International Conference on Image Processing (ICIP)*], 693–696 (2006).
- [6] Ebrahimi, M. and Vrscay, E., “Solving the inverse problem of image zooming using “self-examples”,” in [*Proceedings of International Conference On Image Analysis And Recognition (ICIAR)*], *Lecture Notes in Computer Science* **4633**, 117–130, Springer-Verlag (2007).
- [7] Baker, S. and Kanade, T., “Limits on super-resolution and how to break them,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(9), 1167–1183 (2002).
- [8] Freeman, W., Jones, T., and Pasztor, E., “Example-based super-resolution,” *IEEE Computer Graphics and Applications* **22**(2), 55–65 (2002).
- [9] Datsenko, D. and Elad, M., “Example-based single image super-resolution: A global MAP approach with outlier rejection,” *Journal of Multidimensional Systems and Signal Processing* **18**(2-3), 103–121 (2007).
- [10] Glasner, D., Bagon, S., and Irani, M., “Super-resolution from a single image,” in [*International Conference on Computer Vision*], (2009).
- [11] Goossens, B., Luong, H., Pižurica, A., and Philips, W., “An improved non-local denoising algorithm,” in [*Proceedings of International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*], (2008).
- [12] Dauwe, A., Goossens, B., Luong, H., and Philips, W., “A fast non-local image denoising algorithm,” in [*Proceedings of SPIE Electronic Imaging*], **6812** (2008).
- [13] Takeda, H., Farsiu, S., and Milanfar, P., “Kernel regression for image processing and reconstruction,” *IEEE Transactions on Image Processing* **16**, 349–366 (Feb. 2007).
- [14] Cai, J.-F., Osher, S., and Shen, Z., “Split Bregman methods and frame based image restoration,” Tech. Rep. 08-06, Computational and Applied Math, University of California (2008).
- [15] Guo, K. and Labate, D., “Optimally sparse multidimensional representation using shearlets,” *SIAM Journal Mathematical Analysis* **39**, 298–318 (2007).
- [16] Goossens, B., Aelterman, J., Luong, H., Pižurica, A., and Philips, W., “Efficient design of a low redundant discrete shearlet transform,” in [*Proceedings of International Workshop on Local and Non-Local Approximation in Image Processing (LNLA)*], (2009).
- [17] Goldstein, T. and Osher, S., “The split Bregman method for L1 regularized problems,” Tech. Rep. 08-29, Computational and Applied Math, University of California (2008).
- [18] Combettes, P. and Pesquet, J.-C., “A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery,” *IEEE Journal of Selected Topics in Signal Processing* **1**, 1–12 (Dec. 2007).
- [19] Gharavi-Alkhansari, M., Denardo, R., Tenda, Y., and Huang, T., “Resolution enhancement of images using fractal coding,” *Proceedings of SPIE, The International Society for Optical Engineering* **3024**(2), 1089–1100 (1997).