SELF-EXAMPLE BASED SUPER-RESOLUTION WITH FRACTAL-BASED GRADIENT ENHANCEMENT

Licheng Yu, Yi Xu, Hongteng Xu, Xiaokang Yang

Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China
Shanghai Key Lab of Digital Media Processing and Communication
xuyi@sjtu.edu.cn

ABSTRACT

Recently, the example-based super-resolution method has been extensively studied due to its vivid perception. However, this kind of method directly transfers the high-frequency details of the examples to the low-resolution image, incurring false structures and over-sharpness around the texture regions. In this paper, the problem in the example-based method is investigated from an analytic discussion. Then we propose a super-resolution method that reconstructs sharp edges using the redundancy properties. The super-resolution problem is formulated as a unified regularization scheme which adaptively emphasizes the importance of high-frequency residuals in structural examples and scale invariant fractal property in textural regions. The experimental results show that the highlights of our method exist in the enhanced visual quality with sharp edges, natural textures and few artifacts.

Index Terms— Super-resolution, fractal analysis, self-example, gradient enhancement

1. INTRODUCTION

Super-resolution (SR) techniques overcome the limitations of the imaging system without the need for additional hardware and find increasing applications in digital TV, movie restoration and video surveillance. Generally, super-resolution refers to the process of obtaining higher-resolution images from several lower-resolution ones with fractional-pixel displacements between images. Accordingly, the research of single image super-resolution is developed to estimate the high-resolution (HR) image from only one low-resolution (LR) image. The current single image based SR methods can be divided into three categories: the interpolation-based SR method, the reconstruction-based SR method and the example-based SR method.

The interpolation-based approaches like bi-linear and bi-cubic are commonly used in real-time super-resolution tasks. However, they are inclined to produce blurry and jaggy effects. To well preserve the sharpness of edge structures, directional interpolation methods are recently proposed, such as edge adaptive NEDI[1], iterative curvature based interpolation[2] and auto-regression based interpolation[3].

Reconstruction-based method is proposed to estimate the super-resolution result as an inverse procedure of a SR-to-LR degradation process. Since many HR images may produce the same LR images using the degradation model, the prior of the original HR image is needed. The local smoothness prior of gradient is commonly used to keep the piece-wise smooth gradient surface[4]. As human eyes are more sensitive to edges, many prior models have been developed to achieve image gradient enhancement. Usually, the statistics of local gradient field is used for image upsampling[5]. To enhance the sharpness of local gradient, the concept of gradient profile was proposed in [6] and the reconstruction based on the statistical profile transformation between LR and SR achieved better SR results with faster speed. To achieve comparable sharpness enhancement of local gradient in even shorter time, the work in [7] exploited a patch-by-patch transformation to enhance the gradient map.

Compared with the interpolation-based methods and reconstruction-based methods, which depend on priors to estimate the missing details of LR images, the example-based methods are intensively studied in the last decade, which take efforts to fill in the missing high-frequency band from the external database of the example patches. Usually, the example patches are decomposed into their smoothed version and the residual high-frequency band[8, 9, 10], where the former ones are used for sample matching and the latter ones for filling the missing information. However, the filled-in examples would produce fairly noisy SR results[8] if they are of low relevance to the LR image content. To take advantage of self-similarity, some recent works proposed to construct the example database from the input LR image itself[11, 12], which is promising to produce sharp and fine edges, but the wrongly hallucinated textures are meanwhile introduced.

In this paper, we propose a unified super-resolution framework that merges the gradient regularization into self-

This work is supported in part by NSFC (60932006, 60902073), the 111 project (B07022), STCSM (12DZ2272600) and the High Technology Research and Development Program of China (2011AA01A107, 2012AA011702).

example synthesis. This framework emphasizes the importance of both high-frequency residuals in structural examples and scale invariant fractal property in textural regions. As compared with current reconstruction-based SR methods, instead of recovering the gradient map statistically in texture regions, we employ the fractal invariance property[13] to adaptively obtain a good estimation of the gradient map from a geometrical viewpoint.

The reminder of our paper is organized as follows: Section 2 discusses the problems of the example-based methods from an analytic viewpoint. Section 3 proposes a unified S-R framework that adaptively emphasizes the importance of both high-frequency residuals in structural examples and scale invariant property in textural regions. And in Section 4, the proposed super-resolution approach is compared with the state-of-the-art super-resolution methods. Finally, the conclusion of this work is given in Section 5.

2. EXAMPLE-BASED SUPER-RESOLUTION IN TEXTURE AND STRUCTURE REGIONS

The example-based methods aim to fill in the high-frequency band of the SR image. The choice of the example database is of great importance, as the examples that are irrelevant to the image content would make the SR result noisy[8]. To overcome this problem, there are currently two ways to make example-based method more practical. The first way is to segment the input image into different parts, and perform example synthesis on them using the most related databases[10]. However, the requirement of accurate segmentation and selection of dataset makes the method complicate, usually involving the users' interaction[10, 15]. In addition, the strong reliance on numerous external databases makes the method hard to obtain general application.

The second way is to exploit self-similarity across different scales, performing the hallucination using the database extracted from the input itself[11, 12]. As most singularity

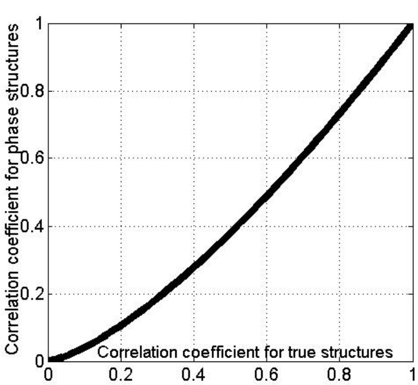

Fig. 1: Correlation coefficients for true structures and phase reconstructed structures[14].

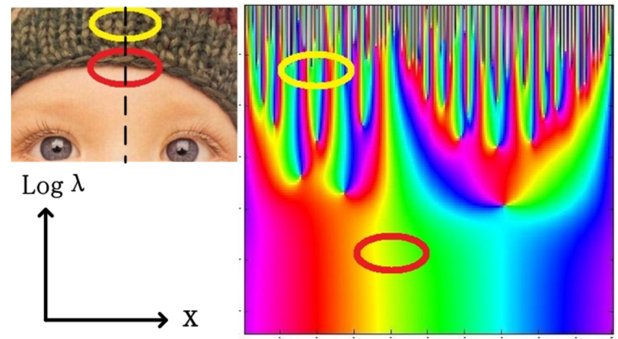

Fig. 2: Phase invariance in Gabor scale space.

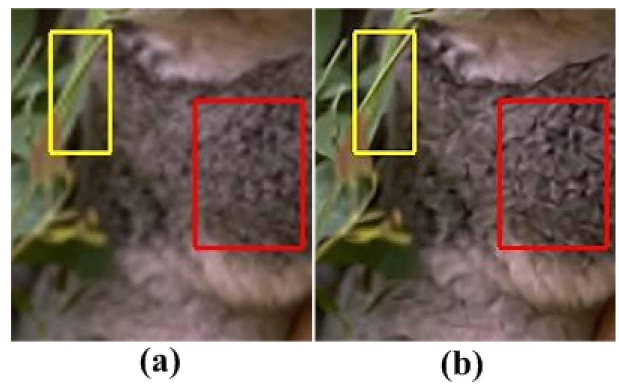

Fig. 3: 4X results of "Koala". (a) Bi-cubic. (b) Fattal 11[12]

structures such as corners and lines are very similar to themselves upon successive magnification with small factors, the missing high-frequency information of the required size can be predicted from the input image itself through iterative scaling. During each up-sampling stage, the local small window of the smoothed version of current input LR image constructs the database for the low-frequency patch matching. Then the residual high-frequency contents of the matched patch at the finer scales are used to fill in the missing high-frequency band of the up-sampled image at the coarser scales. As the localized matching is more accurate and considerably faster than searching matches in a large external database, it opens up a new direction in the example-based methods.

However, self-similarity is still a problematic universal assumption. The scale invariance only holds accurately for the singularities like edges, corners and contour structures, but not for the textures. As we know, phases encode the essential structures in an image. As shown in Figure 1, the correlation of phase-only reconstructions follows very closely the true correlation between the original structures[14]. Therefore, we investigate the invariance of phase structures in scale space to analyze the performance of example-based SR methods in structural regions and textural regions. In Figure 2, we use a log Gabor scale space to observe the self-similarity of phase structure across scales, where λ denotes the scale and x locates the pixel position. In the structural regions marked with a red circle, the phases remain constant within a large scale space range. In contrast, the invariance of phases can be observed only in rather a small scale space range in the textural region, which is located with yellow circle. As a result, a strong redundancy of repetitive image contents can be found in structural regions across different scales, which is consistent with the assumption of self-example SR methods. But it is not a stable case in the textural regions since the phases will change a lot beyond a small scale range. Figure 3 gives such effects. As can be seen, the self-example based method[12] is able to present sharp structural edges compared with bicubic, however, the textural fur region is added with false line artifacts.

3. SELF-EXAMPLE BASED SUPER-RESOLUTION WITH FRACTAL-BASED GRADIENT ENHANCEMENT

In the reconstruction-based SR methods, as above-mentioned, gradient plays an important to obtain an enhanced visual perception. Generally, a statistical gradient transformation[5, 6, 4, 7] is used for estimating the gradient map of the SR image. Although the SR results are fairly encouraging, they still suffer the loss of details in the texture parts. The reason is that the gradient statistics focuses on preserving edges while leaving relatively "smooth" texture regions untouched. In [13], the scale invariance of fractal features in the local gradient field is exploited to obtain reliable reconstruction of the gradient map, rather than directly modeling a statistical transformation of the gradient itself. As compared with example-based SR method, this method is highlighted in achieving more vivid textures, as shown in the red rectangles in Figure 4. However, the proposed parametric reconstruction model [13] cannot accurately recover the high-frequency components of varied edge structures due to unknown anti-aliasing operation on LR image and the potential biased estimates when compared with example-based method[12], as shown in the yellow rectangles

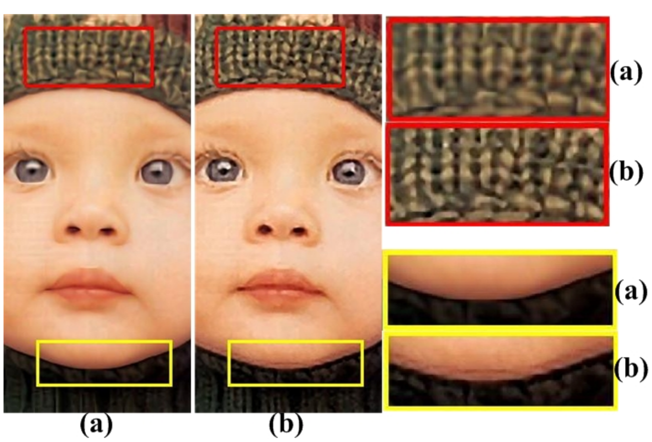

Fig. 4: 4X results of "Child face". (a) Fattal 11[12]. (b) Xu 12[13].

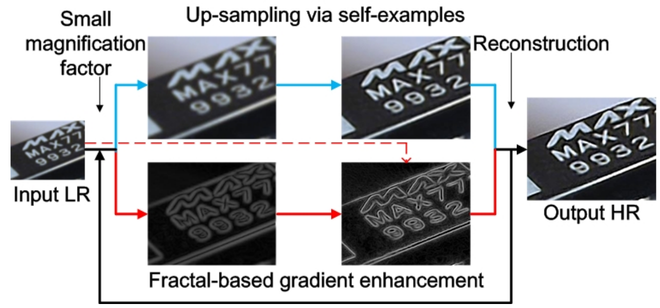

Fig. 5: Our SR scheme.

in Figure 4.

Motivated by the complementary advantages of self-example based SR method and reconstruction-based SR method using invariant fractal features, in this paper, we set up a unified framework which aims to preserve sharp edges and meanwhile reproduce vivid textures. The whole scheme is drawn in Figure 5. Given the input LR image I_0, we exploit the local self-similarities to upsample I_0 iteratively,

$$Y_0(I_0) \overset{s_1}{\to} Y_1 \overset{s_2}{\to} \cdots \overset{s_i}{\to} Y_i \overset{s_{i+1}}{\to} Y_{i+1} \overset{s_{i+2}}{\to} \cdots \to Y_{out}, \quad (1)$$

where Y_i is the i-th upsampled image, s_i is the small magnification factor for the i-th stage. Different from [11, 12], Y_{i+1} is no longer directly gotten by the redundancy of multiple self-examples in our scheme. The SR result is the optimal solution to the following function,

$$Y_{i+1} = \underset{Y}{\text{argmin}} \left\| (Y * f) \downarrow - Y_i \right\| + \lambda_1 \left\| Y - \mathbf{E}(Y_i) \right\| + \lambda_2 \left\| grad(Y) - \mathbf{F}(grad(Y_i)) \right\|, \quad (2)$$

where

$$\mathbf{F}(grad(Y_i(y)) = \beta \frac{\|grad(Y_i(y))\|}{\|grad(Y_i(y))^\alpha\| + \varepsilon} grad(Y_i(y))^\alpha. \quad (3)$$

In (2) and (3), f is the empirical blurring kernel that is usually described as a Gaussian kernel, '$grad$' denotes the gradient map, $\mathbf{E}(\cdot)$ is the self-example based up-scaler for each small magnification factor s_{i+1}[10, 12], $\mathbf{F}(\cdot)$ is the gradient transformation operator, in which two parameters α and β keep the scale invariance of fractal dimension and fractal length of local gradients[13]. Coefficients $\lambda_{1,2}$ control the weights of the intensity and gradient regularization.

As formulated in (2), the first term enforces the reconstruction constraint describing the SR→LR degradation process, i.e., the smoothed and down-sampled version of the estimated HR image should be consistent with its LR image. The second term imposes the regularization constraint on intensity domain, so that Y_{i+1} is able to reconstruct the fine edges from self-examples. And the third term introduces similarity restraints between Y_{i+1}'s gradient field and the enhanced Y_i's gradient field.

(a) Fattal 11 (b) Ours

(c) Xu 12 (d) Ours

Fig. 6: Comparisons of 4X results among Fattal 11[12], Xu 12[13] and ours.

Different from the traditional gradient enhancement models [4, 5, 6, 7], which dominantly focuses on edge sharpening, the fractal-based gradient recovery in (3) can achieve both detail enhancement and edge sharpening. Especially for the texture parts, the gradient regularization in (2) is effective in restraining the over-smoothness and false artifacts. As can be seen in Figure 6(a) and 6(b), more texture details are presented using our method. In addition, the regularization term using self-examples in (2) promotes the proposed scheme in overcoming the drawback shared in the reconstruction-based methods that the sharpened edges would appear coarse, as shown in the lip part and the chin part of Figure 6(c). This is because of the blurry effects during interpolation before gradient enhancement. Instead, such ill-effects would disappear in our framework. The main reason is that the redundancy of self-examples across scales provides sufficient high-frequency band for the reconstruction of fine-scale structures, which can be further enhanced by the fractal-based gradient recovery, as shown in Figure 6(d).

4. EXPERIMENTS AND DISCUSSIONS

To demonstrate the performance of the proposed method, we compare our approach with the state-of-art methods. In the experiments, the Gaussian kernel f in (1) is a 7×7 mask with deviation 0.8, the image patch is set to be 5×5, the searching window of examples has the size of size 10×10, and λ_1 and λ_2 are empirically selected to be 3 and $0.18i$ in the i-th stage, which are acceptable in general case.

In Figure 7, our method is compared on "Einstein" with the recent reconstruction-based methods. More fine line structures are presented in our scheme, as those lines are recovered with high-frequency information from self-examples, thus avoiding being blurred by the initial interpolation. Figure 8 gives another comparison on "chip" with the self-example based SR methods. Due to the gradient regularization, our edges look sharper and meanwhile kick out jaggy effect, as compared with [11].

Table 1: RMS/SSIM values of different methods on the child face

Method	RMS/SSIM	Method	RMS/SSIM
Bi-cubic	15.847/0.705	Freeman02[8]	23.728/0.576
ICBI[**?**]	21.654/0.624	Fattal11[12]	24.065/0.594
Fattal07[5]	23.713/0.624	Xu12[13]	22.141/0.620
Shan08[4]	18.265/0.681	Ours	21.119/0.602

Figure 9 presents the visual results of "Child". To our knowledge, "Child" is the only public available image having both ground truth and SR results of all the testing methods, thus the RMS (Root Mean Square) and SSIM (Structural SIMilarity) values are computed only on this image. As listed in Table I, the interpolation-based methods like Bi-cubic, ICBI and Shan08[1] have better performance in both RMS and SSIM as they emphasize on the consistency with the LR input, while the example-based methods have larger RMS values due to their heavy dependence on the database which has no constraint on consistency. Our approach, yields a lower RMS and higher SSIM value than the other example-based methods, as we take the gradient-based reconstruction model into consideration. Due to the gradient enhancement in our scheme, the proposed method is able to present clearer and sharper edges around the face, as compared with [13]. Meanwhile, the hat details appear move vivid and visible compared with the others.

For validating the robustness of our method, more comparisons of "Can" and "Wheel" are presented in Figure 10.

5. CONCLUSION

In this paper, we propose a single image SR framework that enforces fractal-based gradient regularization into self-example synthesis. While local self-examples provide the S-R result with the fine line structures in the intensity field, the fractal-based gradient regularization contributes to the artifact revision and detail enhancement. Imposing the constraints on both intensity and gradient field, we are able to get a promising SR result with both sharp edges and enhanced textures. In

[1]The SR technique of Shan08 combines the interpolation and reconstruction-based methods.

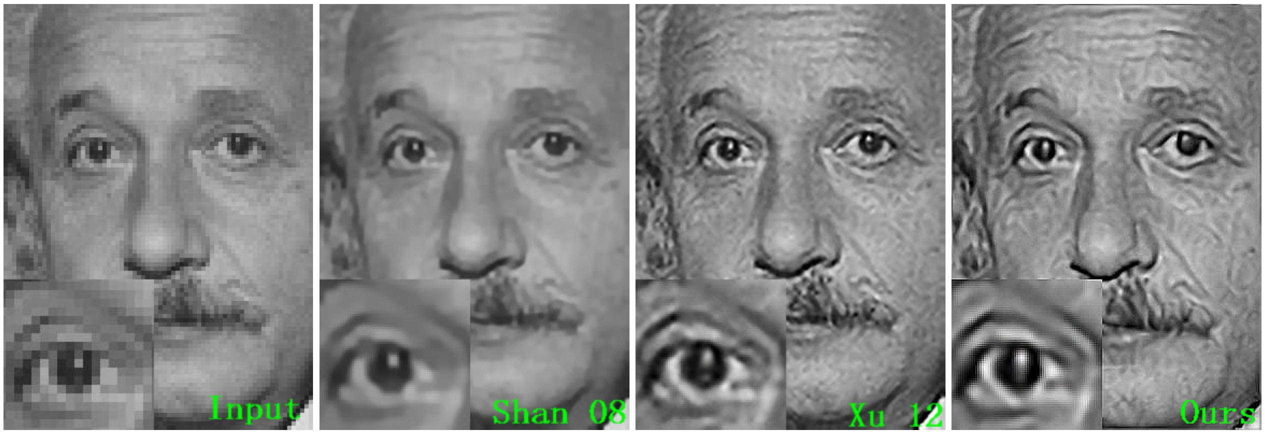

Fig. 7: Comparisons of 4X results on ”Einstein” with the reconstruction-based methods[4, 13].

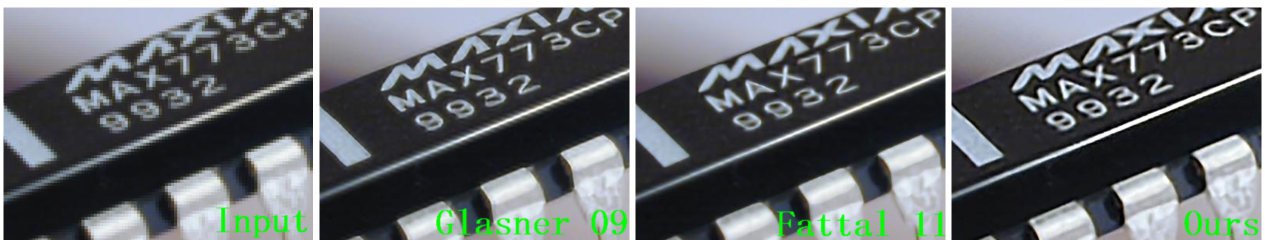

Fig. 8: Comparisons of 4X results on ”Chip” with the example-based methods[12, 11].

Fig. 9: Comparisons of 4X results on ”Child” using different methods.

future works, we will compute the regularization coefficients according to the kind of image patch, including structural image patch and textural image patch, making the SR method more adaptive to the extensive image contents.

6. REFERENCES

[1] X. Li and M.T. Orchard, “New edge-directed interpolation,” *Image Processing, IEEE Transactions on*, vol. 10, no. 10, pp. 1521–1527, 2001.

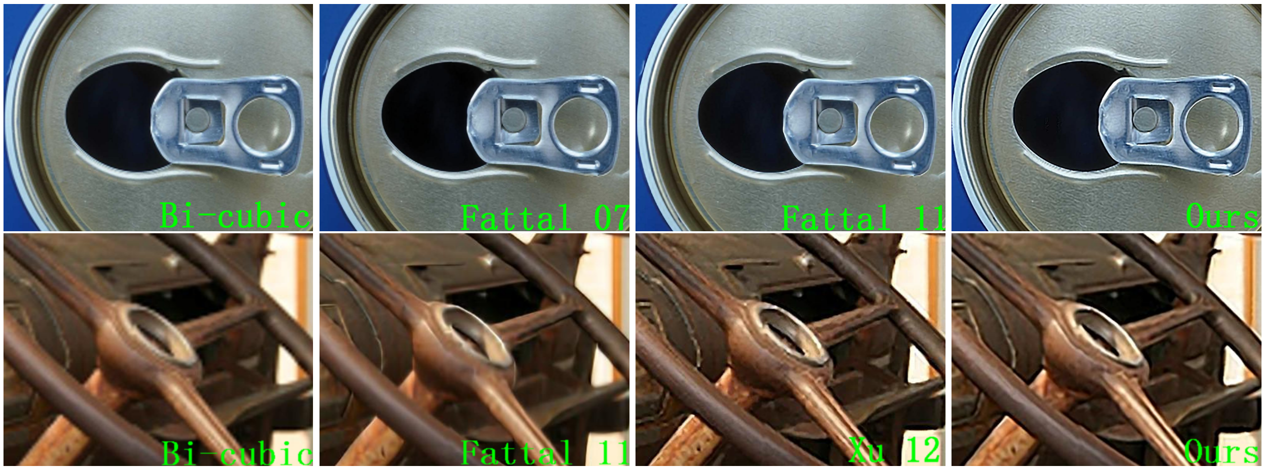

Fig. 10: More comparisons of 4X results on "Can" and "Wheel" using different methods.

[2] Andrea Giachetti and Nicola Asuni, "Real-time artifact-free image upscaling," *Image Processing, IEEE Transactions on*, vol. 20, no. 10, pp. 2760–2768, 2011.

[3] X. Zhang and X. Wu, "Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation," *Image Processing, IEEE Transactions on*, vol. 17, no. 6, pp. 887–896, 2008.

[4] Q. Shan, Z. Li, J. Jia, and C.K. Tang, "Fast image/video upsampling," in *ACM Transactions on Graphics (TOG)*. ACM, 2008, vol. 27, p. 153.

[5] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 95, 2007.

[6] J. Sun, Z. Xu, and H.Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *Image Processing, IEEE Transactions on*, vol. 20, no. 6, pp. 1529–1542, 2011.

[7] Licheng Yu, Hongteng Xu, Yi Xu, and Xiaokang Yang, "Robust single image super-resolution based on gradient enhancement," in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*. IEEE, 2012, pp. 1–6.

[8] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-based super-resolution," *Computer Graphics and Applications, IEEE*, vol. 22, no. 2, pp. 56–65, 2002.

[9] J. Yang, J. Wright, T.S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.

[10] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *Computational Photography (ICCP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–8.

[11] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 349–356.

[12] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 2, pp. 12, 2011.

[13] H.T. Xu, G.T. Zhai, and X.K. Yang, "Single image super-resolution with detail enhancement based on local fractal analysis of gradient," in *Circuits and Systems for Video Technology, IEEE Transactions on*. IEEE, 2012.

[14] Alan V Oppenheim and Jae S Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[15] Y.W. Tai, S. Liu, M.S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2400–2407.