ROBUST INTERNAL EXEMPLAR-BASED IMAGE ENHANCEMENT

Yang Xian[1] *and Yingli Tian*[1,2]

[1]The Graduate Center, [2]The City College,
The City University of New York, New York,
Email: yxian@gc.cuny.edu; ytian@ccny.cuny.edu

ABSTRACT

Image enhancement aims to modify images to achieve a better perception for human visual system or a more suitable representation for further analysis. Based on different attributes of given input images, tasks vary, e.g., noise removal, deblurring, resolution enhancement, prediction of missing pixels, etc. The latter two are usually referred to as image super-resolution and image inpainting. There exist complicated circumstances where low-quality input images suffer from insufficient resolution with missing regions. In this paper, we propose a novel uniform framework to accomplish both image super-resolution and inpainting simultaneously. The proposed approach adopts internal exemplar similarities in image level and gradient level where later enhancement results from both levels are fed into a pre-defined cost function to restore the final output. Experimental results demonstrate that our method is capable of generating visually plausible, natural-looking results with clear edges and realistic textures.

***Index Terms*—** image enhancement, super-resolution, image inpainting, exemplar-based, gradient across-scale similarity

1. INTRODUCTION

Image enhancement has drawn increasing attention in improving image quality or interpretability. After enhancement, images are more visually pleasing for human perceptual system or more suitable for further analysis in many applications such as medical imaging, remote sensing, and video surveillance. Based on different characteristics of the low-quality input images, enhancement tasks vary, for example, noise reduction, inpainting, deblur, and super-resolution.

Among the numerous situations of image enhancement, there are complicated circumstances where part of an input low-resolution (LR) image is undesired or unavailable, e.g., an overexposed spot, an unwanted shadow or scratch within a digital photo, a satellite image partly covered by cloud. For these cases, in addition to the enhancement of image resolution, image inpainting techniques are also needed to reconstruct more visually plausible result. The underlying goal for both image super-resolution and inpainting is to predict the unknown pixel values within an image. They are numerically ill-posed problems and rely on additional assumptions or priors to finalize the output among all the possible solutions.

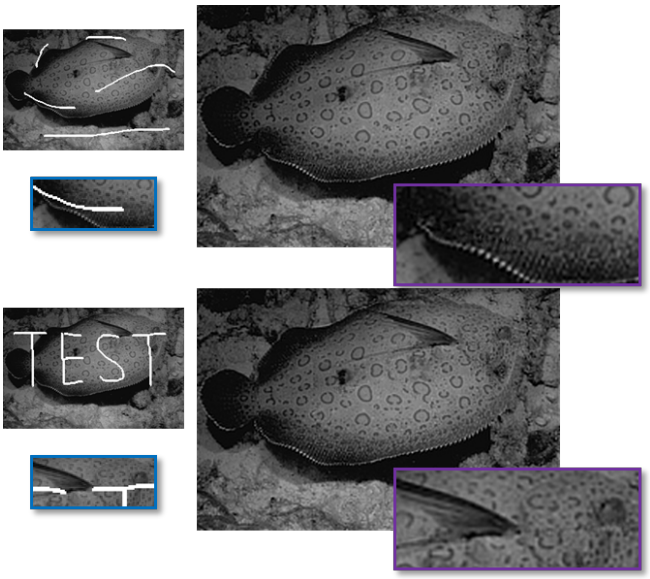

Fig. 1. Enhancement results of the proposed framework on image 'fish' ($\times 2$) with different masks. Left: input LR images with missing regions (marked in white); Right: generated HR results with clear contours and natural textures. The figure is better viewed on screen with HR display.

Single image super-resolution aims at estimating a high quality fine-resolution image from one coarse-resolution image. Recently, exemplar learning-based techniques have been widely used in image super-resolution approaches by exploring the relationship between high-resolution (HR) exemplars and their corresponding LR exemplars. The learning can be performed either through an external dataset [1, 2, 3, 4, 5] or within the input image [6, 7]. Different from external exemplar-based learning which depends on a large dataset of natural images, internal exemplar-based super-resolution is based on the observation that small image patches in a natu-

This work was supported in part by ONR grant N000141310450 and NSF grants EFRI-1137172, IIP-1343402.

ral image tend to appear redundantly within the input image itself and across different scales. Glasner *et al.* proposed an image super-resolution framework based on the nearest neighbor search within a patch pool formed with internal patches collected through a pyramid structure utilizing only the input image at different resolutions. In [8], Zontak and Irani have demonstrated the effectiveness and powerfulness of internal exemplars in image enhancement tasks including image super-resolution. Compared with internal exemplar-based approaches, external based methods usually require a dataset with hundreds of training images to achieve comparable enhancement performance.

Different from internal exemplar-based image super-resolution which incorporates self-similarities across different scales, exemplar-based image inpainting approaches [9, 10, 11, 12, 13] deal with patch recurrence within the same resolution. They replace incomplete patches with similar ones in the known region and have been successful in replicating visually plausible background textures. Moreover, super-resolution techniques have already been utilized to assistant image inpainting. Le Meur and Guillemot [13] proposed a super-resolution-aided inpainting method which firstly performed inpainting in a coarse-version of the input image followed by a super-resolution process as to retain the original resolution.

In this paper, we propose a novel and straightforward algorithm for image enhancement which performs super-resolution and inpainting simultaneously. Given an input LR image and a mask representing the missing region(s), we perform enhancement in both gradient level and image level. The input LR gradients in x and y directions are upscaled with inpainting embedded. The resulting HR gradients, along with the inpainted LR image, are fed into an easily-optimized cost function to reconstruct the final HR enhanced image.

The contributions of the proposed image enhancement framework are fourfold: 1) A uniform image enhancement framework is proposed to accomplish both super-resolution and inpainting given a LR input image with unavailable region(s). 2) Both gradient-level and image-level enhancement are adopted to ensure the robust performance. 3) A straightforward energy function is utilized to incorporate the enhanced gradients while maintaining the consistency to the input image. 4) Experimental results demonstrate that our algorithm is capable of generating natural and visually pleasing results.

2. EXEMPLAR-BASED IMAGE ENHANCEMENT

In this section, we present our proposed image enhancement framework based on internal exemplar similarity within the same scale and across different resolutions. Both gradient-level and image-level enhancement are employed to better preserve intensity changes and ensure robust performance. Fig. 2 illustrates the schematic pipeline of our approach. Given an input LR image and a mask indicating the region(s) with missing pixels, the system consists of three components to fulfill both super-resolution and inpainting: gradient-level upscaling with inpainting, image-level inpainting, and final HR image reconstruction.

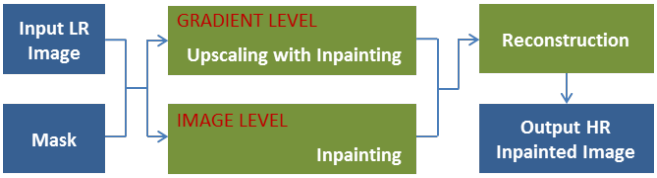

Fig. 2. Flowchart of the proposed internal exemplar-based image enhancement method. Given an input LR image with a mask indicating the missing region(s), the LR gradients are upsampled with inpainting embedded. The HR gradients, along with the inpainted LR image, are fed into an energy function to reconstruct the final HR output image.

2.1. Gradient-Level Upscaling with Inpainting

Internal across-scale gradient similarity is utilized to accomplish super-resolution and inpainting for input LR gradients simultaneously. It is based on the observation that since for small image patches in a natural image, self-similarities exist within the image itself and across different scales, we should expect the similar redundancy for gradient patches.

Given a greyscale input LR image L, the mask M, and the scaling factor s, we denote the gradients of L in x and y directions as L_x and L_y. The calculated HR gradients are represented as H_x and H_y. L_x is decomposed into a set of overlapping patches with size $a \times a$. Patches with unknown pixels are upscaled first. Among the patches with missing regions, the upsampling priority for patch P centered at pixel q is determined as follows:

$$Pri(q) = C(q) \cdot D(q) = \frac{\sum_{i \subseteq \{P(q) \cap \bar{M}\}} C(i)}{a^2} \cdot \sqrt{\frac{\nabla L_{x_q}^{\perp} \cdot u_q}{N}}, \quad (1)$$

where $P(q)$ represents the instant patch centered at pixel q, $\bar{M}$ indicates the unmasked region, N is a normalization factor (255 for grey images), u_q stands for a unit vector orthogonal to the front at pixel q. The initialization for $C(i)$ is set to $C(i) = 0$ if pixel value i is unknown and $C(i) = 1$ otherwise. L_{x_q} represents the value at pixel q in L_x. As indicated in Eq. (1), priority at a given pixel is measured as the product of two terms: the confidence term $C(\cdot)$ and the data term $D(\cdot)$. Both terms are normalized to range between 0 and 1.

Different from the priority computation in [9], we assign more credit to the data term in the calculation of final priority by modifying it to a squared form. The confidence term remains unchanged. In general, the confidence term measures the amount of reliable information surrounding a given pixel.

The data term detects how strongly an isophote at that pixel collides and the contour at the same pixel.

After calculating the priority for each pixel along the boundary of the masked region, the patch with the highest priority at its center pixel is selected as the query patch p to be upsampled. We then downsample L_x by the scaling factor s to obtain LL_x. A gradient patch pool $\wp_x$ is formed with all the patches in LL_x (size $a \times a$) whose pixel values are all known. To ensure a more expressive representation, all the patches in $\wp_x$ as well as the query patch p are normalized to have zero mean and unit variance.

Given a query patch p, its k most similar patches are searched within the patch pool $\wp_x$. The similarity between two patches is measured in the mean square error (MSE) with only the available pixels. After obtaining the k similar patches in LL_x, their corresponding 'parent' patches in L_x are extracted and combined weightedly. The combined patch is then readjusted according to the original mean and variance of p and 'pasted' to the corresponding position in H_x. After updating the confidence term and data term, the above process is repeated until all patches which have overlaps with the mask M are upsampled. Then the rest patches in L_x are upscaled in a similar manner without the necessity of computing the patch priority. H_y is computed in the same structure utilizing L_y.

2.2. Image Level Inpainting

To ensure a robust enhancement performance in constructing the final HR enhanced image, we also execute the image level inpainting on the input LR image L before the final reconstruction step.

Priorities for every pixel along the boundary of the mask region in L are calculated according to Eq. (1). We then form a patch pool with all the patches (size $a \times a$) in the unmasked region of L. The k most similar patches of patch p whose center pixel has the highest priority are searched within the patch pool. Then the k found patches are combined weightedly based on their similarity with the query patch. The unknown pixel values in patch p are filled with the corresponding values in the combined patch.

After updating the confidence term and data term for the filling pixels, the above process is repeated until all the pixel values within the mask region are predicted. Finally, the inpainted image L_I along with the HR gradients H_x, H_y are utilized for the final reconstruction of the target HR image H_E.

2.3. Final Image Reconstruction

After obtaining the inpainted LR image L_I and the HR gradients H_x, H_y, the output HR image H_E is reconstructed through minimizing the following energy function:

Fig. 3. Enhancement result on image 'snow' ($\times 2$). Left: input LR image; Right: result after enhancement. The figure is better viewed on screen with HR display.

$$H_E = \underset{H}{\operatorname{argmin}}\{\lambda|\nabla H - \nabla H_D|^2 + |(H * G)\downarrow_s - L_I|^2\}, \quad (2)$$

where ∇H_D indicates the computed H_x and H_y. λ is the weighting factor between two terms of the cost function. G represents a Gaussian kernel whose standard variance σ is set according to the scaling factor s: $\sigma = \{0.8, 1.2, 1.6\}$ if $s = \{2, 3, 4\}$.

Two terms are included in the energy function shown in Eq. (2): the first term poses a constraint on the gradients of the target HR image with the gradients calculated after 'gradient-level upscaling with inpainting' step. The second term ensures the consistency between the output HR image and the completed input LR image. The cost function can be easily optimized through the gradient descent algorithm iteratively with:

$$H^{t+1} = H^t - \delta \cdot (((H^t * G)\downarrow - L_I)\uparrow * G - \lambda(\nabla^2 H^t - \nabla^2 H_D)), \quad (3)$$

where t represents the counter of iteration and δ stands for the step width. As illustrated in Fig. 1, under different masks, the proposed framework well enhances the resolution of the input LR image and predicts the missing pixel values in a way that is visually plausible with natural and realistic textures.

3. EXPERIMENTAL RESULTS

In this section, the proposed internal exemplar-based image enhancement method is evaluated on multiple natural images

Fig. 4. Enhancement results on images in BSDS[14] ($\times 2$) with different masks. Left: input LR images with missing region(s). Right: output HR images after enhancement. The figure is better viewed on screen with HR display.

with different masks at a magnification factor of 2. Generally, patch size used in internal exemplar-based super-resolution approaches is not large since patch recurrence occur among small patches within the same image or across scales. However, for patch-based image inpainting algorithms, we need a relatively large patch to retain local structure information and ensure a robust inpainting performance. Therefore, in the proposed framework, to balance between these two tasks, we set the patch size a to 7. In calculation of the pixel priority, normalization factor N is 255 for grey images. During the nearest neighbor search, number k of similar patches extracted is set to 10. The weighting factor λ is 0.1 in final image reconstruction step.

Fig. 3 presents the image enhancement result on image 'snow'. As illustrated by the zoom-in region, the missing pixels are restored and upsampled with textures consistent with the overall structure. After image enhancement, sharper edges and more realistic textures are restored. Fig. 4 provides more results on images in the Berkeley Segmentation Dataset (BSDS) [14] under different masks. Enhancement for certain input low-quality images is challenging due to the existence of complicated textures, e.g., fur, trees, body patterns. Zoom-in areas clearly indicate that our method correctly predicts the missing pixel values and reconstructs more natural contours and clearer details with minimal visual artifacts.

4. CONCLUSION

In this paper, we have proposed a novel robust image enhancement framework which accomplishes both super-resolution and inpainting given a LR input image with missing pixels. The input LR gradients are upscaled and inpainted utilizing internal across-scale patch similarities. Along with the inpainted LR image, the HR enhanced gradients are incorporated into a straightforward cost function to reconstruct the final output image. As demonstrated by the extensive experimental results, the proposed approach is robust and capable of generating visually pleasing results with sharp edges and natural textures.

5. REFERENCES

[1] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25–47, 2000.

[2] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *CVPR*, 2008.

[3] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *ICCP*, 2010.

[4] R. Timofte, V. D. Smet, and L. V. Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *ICCV*, 2013.

[5] Y. Zhu, Y. Zhang, and A. L. Yuille, "Single image super-resolution using deformable patches," in *CVPR*, 2014.

[6] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *ICCV*, 2009.

[7] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Transactions on Graphics*, vol. 28, no. 3, pp. 1–10, 2010.

[8] M. Zontak and M. Irani, "Internal statistics of a single natural image," in *CVPR*, 2011.

[9] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," in *CVPR*, 2003.

[10] Nikos Komodakis and Georgios Tziritas, "Image completion using global optimization," in *CVPR*, 2006.

[11] Alexander Wong and Jeff Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *ICIP*, 2008.

[12] Zongben Xu and Jian Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1153–1165, 2010.

[13] Olivier Le Meur and Christine Guillemot, "Super-resolution-based inpainting," in *ECCV*, 2012.

[14] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV*, 2001.