Fast Deconvolution-Based Image Super-Resolution Using Gradient Prior

Chun-Yu Lin, Chih-Chung Hsu, and Chia-Wen Lin^{*} Department of Electrical Engineering National Tsing Hua University Hsinchu, Taiwan cwlin@ee.nthu.edu.tw

Abstract—Single-image super-resolution (SR) is to reconstruct a high-resolution image from a low-resolution input image. Nevertheless, most SR algorithms are performed in an iterative manner and are therefore time-consuming. In this paper, we propose an iteration-free single-image SR algorithm based on fast deconvolution with gradient prior. Based on the prior calculated from the initially upsampled image via current approach (e.g., bicubic interpolation or example/learning-based approaches), we make the deconvolution process well-posed, which can be efficiently solved in FFT domain. Moreover, the proposed algorithm can be directly applied to video SR, where the temporal coherence can be automatically maintained. Experimental results demonstrate that the proposed method can simultaneously obtain significant acceleration and quality improvement over several existing SR methods.

I. INTRODUCTION

With the rapid development of multimedia and network technologies, sharing multimedia contents through heterogeneous devices of different capabilities has become more and more popular. In this paper, we focus on sharing or exchanging image contents among devices (e.g., handheld devices, TV, or computer monitor) with different display capabilities. The display size of a handheld device is typically much smaller than that of TV. To better display a low-resolution (LR) image acquired from a low-cost device (e.g., mobile phone) on a high-cost device (e.g., HDTV), image super-resolution (SR) technique is therefore required.

The goal of image SR is to recover a high-resolution (HR) image from one or multiple LR input images, which is essentially an inverse and ill-posed problem [1]. There are mainly two categories of approach for image SR: (i) classical approach and (ii) example/learning-based approach. In the classical approach, one sub-category is reconstruction-based approach, where a set of LR images of the same scene are aligned with sub-pixel accuracy to generate an HR image [2]. Such kind of approach is usually time-consuming and impractical due to multiple input LR images are required. Moreover, the allowed magnification factor is usually very limited [3]. The other sub-category of the classical approach is

Li-Wei Kang Institute of Information Science Academia Sinica Taipei, Taiwan lwkang@iis.sinica.edu.tw

the interpolation-based approach [4], which has been shown to generate overly smooth images with ringing and jagged artifacts.

The example/learning-based approach hallucinate the high frequency details of an LR image based on the co-occurrence prior between LR and HR image patches in a training set, which has been shown to outperform the classical approach [5]-[7]. Nevertheless, the HR details hallucinated by such kind of approach cannot be guaranteed to provide the true HR details. Moreover, the performance of this approach highly relies on the similarity between the training set and test set. It is still unclear how many training examples are sufficient for most images. Most example/learning-based methods typically require enormous databases of millions of HR and LR patch pairs, and are therefore computationally intensive. Although, the sparse coding-based approach [8] has been proposed to learn a compact representation for HR/LR patch pairs, substantially improving the speed, the computational complexity is still significant.

Different from the above-mentioned approaches, the statistical edge information was employed to help constraining the image SR problem [9]-[10], which has been shown to be much faster than the example/learning-based method [6] in terms of computation while outperform or be comparable to [6] in terms of image quality. In [9], a gradient profile prior describing the shape and sharpness of the image gradients was proposed for image SR. In [10], a statistical edge dependency relating certain edge features of two different resolutions was employed as the prior knowledge for image SR. Moreover, a very efficient method developing a feedback-control loop to automatically measure whether the output SR image satisfies the imaging model [3] through a deconvolution and reconvolution method was proposed in [11], which has been shown to outperform or be comparable to [10]. Nevertheless, the computational complexity of [11] is still significant due to its iterative manner.

In this paper, a fast deconvolution-based image SR method using gradient prior is proposed. The major contribution of

^{*}Corresponding author: Prof. Chia-Wen Lin, email: cwlin@ee.nthu.edu.tw This paper was supported in part by the National Science Council, Taiwan, under Grants NSC98-2221-E-007-080-MY3, NSC100-2218-E-001-007-MY3, NSC100-2811-E-001-005, NSC99-2218-E-001-010 and NSC99-2811-E-001-006.

this paper is three-fold: (i) our method is iteration-free and very fast; (ii) the used gradient prior information is generated from the input LR image itself without pre-collecting training images; and (iii) our method can be naturally extended to video SR, where the temporal coherence can be automatically maintained.

The rest of this paper is organized as follows. In Sec. 2, we briefly review the employed image formation process. Sec. 3 presents the proposed fast deconvolution-based image SR method. In Sec. 4, experimental results are demonstrated. Finally, Sec. 5 concludes this paper.

II. IMAGE FORMATION PROCESS

As shown in [3], [11], by letting H denote an HR image and f denote the camera's point spread function (PSF), the LR image L can be expressed as

$$L = (f \otimes H) \downarrow^d, \tag{1}$$

where \otimes is a convolution operator and \downarrow^d is a decimating (downsampling) operator with factor *d*. Eq. (1) can be viewed as two steps, describing a convolution step followed by a decimation process as

$$\widetilde{H} = f \otimes H, L = \widetilde{H} \downarrow^d, \tag{2}$$

where \tilde{H} is the linearly filtered HR image. If the two steps described in Eq. (2) can be properly inverted by first upsampling the image followed by deconvolving it using a linear filter, H can be adequately recovered. Nevertheless, the upsampling problem is severely ill-posed due to much information loss induced by only keeping one pixel for every $d \times d$ pixels in \tilde{H} in the decimation process. Moreover, for deconvolution process, it has been shown that the non-blind deconvolution problem can be well-posed by adding some prior term [12].

III. PROPOSED FAST DECONVOLUTION-BASED IMAGE SR METHOD USING GRADIENT PRIOR



Figure 1. The proposed fast image SR framework.

In this paper, we aim to develop a fast single-image SR algorithm. The proposed image SR framework shown in Figure 1 adopting an iteration-free method, consisting of the three steps including the initial upsampling, gradient prior calculation, and deconvolution steps. Given an input LR image L, we first apply a fast bicubic interpolation to initially upsample L to the desired resolution to obtain the linearly filtered HR version \tilde{H} of the image. Then, we calculate the gradient prior from \tilde{H} to make the deconvolution process well-posed for recovering the final output HR image H^* from \tilde{H} , which are described in detail in Sec. III-A and Sec. III-B.

A. Gradient Prior Calculation

Given the initially upsampled image \tilde{H} , the deconvolution process estimates the final output HR image H^* , subject to minimize the energy function defined as

$$E(H^*) = \left\| f \otimes H^* - \widetilde{H} \right\|_2^2.$$
(3)

Nevertheless, the direct minimization of Eq. (3) has been shown to be ill-posed [11]. In [11], a non-blind deconvolution method similar to [12] was adopted to make the deconvolution process well-posed by adding a prior term, where a gradient density distribution prior was used.

Instead, in this paper, we propose to simply calculate the gradient prior from the initially upsampled image \tilde{H} , generated by performing the bicubic interpolation to the input LR image L, as

$$G_x = \partial_x \widetilde{H}, G_y = \partial_y \widetilde{H},$$
 (4)

where G_x and G_y are the gradient values of the horizontal and vertical directions, respectively, of \tilde{H} , ∂_x and ∂_y are the gradient operators for the horizontal and vertical directions, respectively. In this paper, we use the Sobel operator to calculate the gradient information of \tilde{H} . Using the obtained gradient prior to make the deconvolution process well-posed is described in the next subsection.

B. Fast Deconvolution Using Gradient Prior in Frequency Domain

By incorporating the image gradient prior calculated by Eq. (4), the energy function defined in Eq. (3) can be converted to

$$E(H^*) = \|f \otimes H^* - \widetilde{H}\|_2^2 + \lambda \left(\|\partial_x H^* - G_x\|_2^2 + \|\partial_y H^* - G_y\|_2^2 \right),$$
(5)

where $\partial_x H^*$ and $\partial_y H^*$ are the gradient values of the horizontal and vertical directions, respectively, of H^* , and λ is a weight. Traditionally, Eq. (5) can be directly solved via gradient descent or matrix manipulation approaches, which is very time-consuming. Instead, similar to [12], we propose to employ Plancherel's theorem [13] to allow this computationally intensive optimization process to be efficiently performed in the frequency (Fourier transform) domain as follows.

Based on Plancherel's theorem [13] stating that the sum of the square of a function equals the sum of the square of its Fourier transform, Eq. (5) can be converted to

$$E(H^*) = \|\mathcal{F}(H) \circ \mathcal{F}(H^*) - \mathcal{F}(\tilde{H})\|_2^2 + \lambda \left(\|\mathcal{F}(\partial_x) \circ \mathcal{F}(H^*) - \mathcal{F}(G_x)\|_2^2 + \|\mathcal{F}(\partial_y) \circ \mathcal{F}(H^*) - \mathcal{F}(G_y)\|_2^2 \right),$$
(6)

where \mathcal{F} is the fast Fourier transform (FFT) operator and " \circ " is the element-wise multiplication operator. The major advantage for transforming Eq. (5) to FFT domain is to make the deconvolution process more efficient by converting the

convolution operation to the element-wise multiplication operation.

To solve the optimal $\mathcal{F}(H^*)$ for minimizing Eq. (6), we let $\partial E(H^*)/\partial H^* = 0$. Then, we can get

$$\frac{\partial E(H^*)}{\partial H^*} = \begin{bmatrix} \mathcal{F}(f) \circ \mathcal{F}(H^*) - \mathcal{F}(\widetilde{H}) \end{bmatrix} \circ \overline{\mathcal{F}(f)} + \\ \lambda \begin{bmatrix} \mathcal{F}(\partial_x) \circ \mathcal{F}(H^*) - \mathcal{F}(G_x) \end{bmatrix} \circ \overline{\mathcal{F}(\partial_x)} + \\ \lambda \begin{bmatrix} \mathcal{F}(\partial_y) \circ \mathcal{F}(H^*) - \mathcal{F}(G_y) \end{bmatrix} \circ \overline{\mathcal{F}(\partial_y)} = 0, \quad (7)$$

where $\overline{\mathcal{F}(f)}$, $\overline{\mathcal{F}(\partial_x)}$, and $\mathcal{F}(\partial_y)$ are the complex conjugates of $\mathcal{F}(f)$, $\mathcal{F}(\partial_x)$, and $\mathcal{F}(\partial_y)$, respectively. Based on Eq. (7), we can obtain

$$\begin{aligned} \mathcal{F}(H^*) &= \frac{\overline{\mathcal{F}(f)} \circ \mathcal{F}(\widetilde{H}) + \lambda \overline{\mathcal{F}(\partial_x)} \circ \mathcal{F}(G_x) + \lambda \overline{\mathcal{F}(\partial_y)} \circ \mathcal{F}(G_y)}{\overline{\mathcal{F}(f)} \circ \mathcal{F}(f) + \lambda \overline{\mathcal{F}(\partial_x)} \circ \mathcal{F}(\partial_x) + \lambda \overline{\mathcal{F}(\partial_y)} \circ \mathcal{F}(\partial_y)}. \end{aligned} (8)$$

Finally, the output HR image H^* can be obtained by performing the inverse FFT (\mathcal{F}^{-1}) to $\mathcal{F}(H^*)$ as

$$H^* = \mathcal{F}^{-1}\big(\mathcal{F}(H^*)\big). \tag{9}$$

C. Other Approaches of Gradient Prior Calculation

Besides the bicubic interpolation, any existing image SR method can be also used to initially upsample the input LR image and calculate the gradient prior. Nevertheless, if the employed existing SR method is time-consuming, the computational complexity of the proposed method will be increased accordingly.

In fact, our method can partially incorporate an existing SR method for calculating gradient prior and achieve similar performance with much less computations than this existing method. For example, when employing an example-based [6] or sparse coding-based [8] SR method, we select several patches with larger variances from the input LR image to be upsampled using the existing method while the rest patches are still upsampled using the bicubic interpolation.

IV. EXPERIMENTAL RESULTS

The proposed method was implemented in MATLAB[®] on a personal computer equipped with 2.4GHz CPU. To evaluate the performance of the proposed method, we first compare the proposed method with a fast image upsampling method also based on fast deconvolution [11] in terms of the computational complexity measured in run time (seconds). TABLE I shows the run time of the fast upsampling method [11], the GPUspeedup version based on the VGA 8600GT hardware of [11], and the proposed method using the bicubic interpolation (in default) for calculating the gradient prior for the four different target resolutions, where the run time for each resolution was obtained by averaging those of ten test images and the magnification factors are all set to 2. It can be observed from TABLE I that the proposed method can be significantly faster than the fast upsampling method [11] and its speedup version. The main reason is that the method proposed in [11] should iteratively perform deconvolution and re-convolution, which

is very time-consuming even though it can be speeded up by GPU. In contrast, the proposed method performs the deconvolution process only once by introducing the gradient prior calculated from the bicubic-upsampled image. Based on our experimental results, the gradient prior based on bicubic interpolation is usually good enough for the succeeding deconvolution process. We also subjectively compare the proposed method with the bicubic interpolation and the fast upsampling method [11], as shown in Figure 2. It can be also obtained from Figure 2 that the proposed method can outperform the method proposed in [11] in terms of both subjective image quality and computational speed.

 TABLE I. RUM-TIME (IN SECONDS) COMPARISONS AMONG DIFFERENT METHODS.

Target	Run Time (seconds)		
Resolution	Fast upsampling [11]	GPU-speedup for [11]	Proposed
800×600	35	8	0.54
1024×768	60	15	0.79
1280×1024	107	27	1.69
1920×1080	162	41	2.23



Figure 2. Image SR results: (a) the original image; and the selected region of SR images of (a) via the: (b) bicubic interpolation; (c) fast upsampling method (40 deconds) [11]; and (d) proposed method (0.57 seconds).

Moreover, we also evaluate the proposed method incorporating some example/learning-based SR method for gradient prior calculation. We employ the example-based SR method proposed in [6] by selecting 50% of patches with larger variances from the input LR image. Then, we initially upsample these patches via this SR method [6] and the rest patches via the bicubic interpolation to obtain the initially upsampled image, followed by calculating the gradient prior based on it. We also employ the sparse coding-based SR method proposed in [8] by performing the same gradient prior calculation process. TABLE II shows the objective performances in PSNR (peak signal-to-noise ratio in dB) and run time (in seconds) obtained by the example-based method [6], the sparse coding-based method [8], the proposed method with example-based gradient prior, the proposed method with

sparse coding-based gradient prior, and the proposed method (in default), respectively, for upsampling the *Butterfly* image from 129×120 to 387×360 . We also compare the subjective SR image quality obtained from these methods, as shown in Figure 3. It can be observed from TABLE II that the proposed method with example-based gradient prior can significantly outperform and be faster than the example-based gradient prior can be comparable and faster than the sparse coding-based method [8]. In addition, the proposed method in default (with bicubic gradient prior) can outperform or be comparable to, and significantly faster than the other methods for comparisons. In general, using the bicubic interpolation for calculating gradient prior usually provides acceptable results in terms of both performance and efficiency.

TABLE II. OBJECTIVE IMAGE QUALITY (IN DB) AND RUM-TIME (IN SECONDS) COMPARISONS AMONG DIFFERENT METHODS.

Methods	PSNR (in dB)	Run Time (seconds)
Example-based SR [6]	17.45	2160
Sparse coding-based SR [8]	24.53	26
Proposed with example- based gradient prior	24.13	1238
Proposed with sparse coding-based gradient prior	24.53	17
Proposed in default	24.37	0.15

Finally, we directly extend our method to video SR, where the temporal coherence can be automatically maintained. Figure 4 shows the SR results for a video frame obtained by our method and the bicubic interpolation.

V. CONCLUSIONS

In this paper, we have proposed a fast deconvolution-based image SR method, employing the bicubic interpolation for gradient prior calculation. By incorporating the prior, we make the deconvolution process well-posed, which can be efficiently solved in FFT domain. Experimental results demonstrate the efficiency of the proposed algorithm.

References

- S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21-36, 2003.
- [2] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327-1344, Oct. 2004.
- [3] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167-1183, 2002.
- [4] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Signal Process.*, vol. 26, no. 6, pp. 508–517, 1978.
- [5] W. T. Freeman, E. C. Pasztor, and O. Carmichael, "Learning low-level vision," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25-47, 2000.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based superresolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, Mar./Apr. 2002.
- [7] J. Sun, N. N. Zheng, H. Tao, and H. Y. Shum, "Image hallucination with primal sketch priors," in *Proc. IEEE CVPR*, June 2003.
- [8] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861-2873, 2010.

- [9] J. Sun, J. Sun, Z. B. Xu, and H. Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE CVPR*, June 2008.
- [10] R. Fattal, "Image upsampling via imposed edge statistics," ACM Trans. Graph., vol. 26, no. 3, July 2007.
- [11]Q. Shan, Z. Li, J. Jia, and C. K. Tang, "Fast image/video upsampling," ACM Trans. Graph., vol. 27, no. 5, Dec. 2008.
- [12]Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," ACM Trans. Graph., vol. 27, no. 3, Aug. 2008.
- [13] R. N. Bracewell, *The Fourier transform and its applications*, McGraw-Hill Science, 3 Edition, June 1999.



Figure 3. Image SR results: (a) the ground truth; and the SR images via the: (b) bicubic interpolation; (c) example-based SR [6]; (d) sparse coding-based SR [8]; (e) proposed with example-based gradient prior; and (f) proposed with sparse coding-based gradient prior.



Figure 4. Video frame SR results via left: the proposed in default and right: the bicubic interpolation.