A CONFIDENCE GROWING MODEL FOR SUPER-RESOLUTION

Sina Lin[1,2] *Zengchang Qin*[1*] *Renjie Liao*[3] *and Tao Wan*[1,4]

[1]Intelligent Computing and Machine Leaning Lab
School of ASEE, Beihang University, Beijing, 100191, China
[2]Yamaha Corporation, Hamamatsu, 430-0904, Japan
[3]Department of CSE, Chinese University of Hong Kong, China
[4]School of Biological Science and Medical Engineering
Beihang University, Beijing, 100191, China

ABSTRACT

Single image super-resolution (SR) aims at generating a high-resolution (HR) image from one low-resolution (LR) input. In this paper, we focus on single image SR by using a *confidence growing* model based on an example-based super resolution approach. Compared to previous works that reconstruct high-resolution image in a raster scan order, the new proposed method reconstructs the patches using a new confidence measure. More confident reconstructions are propagated to neighboring areas by enforcing a smoothness constraint in selecting patches. We also adopt hierarchical clustering to construct a training set to speed up processing. Experimental results demonstrate that this simple method outperforms existing state-of-the-art algorithms on a the given benchmark SR test images.

Index Terms— super-resolution; confidence growing; example-based SR.

1. INTRODUCTION

The basic idea of Super-Resolution (SR) is to estimate a high resolution (HR) image from a single or several original low resolution (LR) images. This is an ill-posed problem since the mapping between HR image and LR image is many-to-one and information is lost in the HR-to-LR process. Three major paradigms for image super-resolution were well studied: (1) Interpolation based methods that generate HR image using single LR image [7, 8]. (2) Reconstruction-based methods using multiple LR images which describe the same scene to estimate the high-resolution image [1, 6] and (3) Example-based (also known as learning) methods that learn the LR/HR relation using a training set of LR/HR patch pairs [4, 13, 14].

The performance of classical reconstruction-based methods may degrade without adequate low-resolution patches. It is also influenced by the accuracy of registration procedure. In example-based super-resolution, LR and HR patch pairs learned from training dataset are used to estimate the high-resolution image. Higher SR factors have often been obtained by repeated applications of this process. Example-based SR has been shown to exceed the limits of classical SR. However, unlike classical SR, the high resolution details reconstructed (hallucinated) by example-based SR are not guaranteed to provide the true (unknown) HR details [5]. Selected HR patches are probably not compatible with its neighbors, which breaks the smoothness of final reconstructed image. Recently, some new approaches were proposed by using local learnable kernel regression [9] and first-order approximation of the nonlinear mapping based on self-similarity of patches [17].

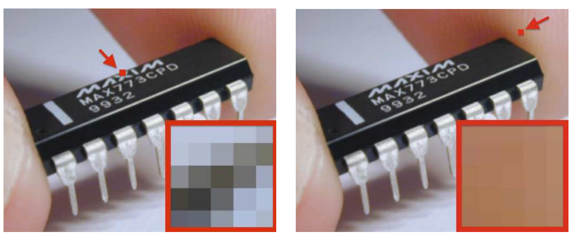
In previous work, Freeman *et al.* [4] employs the Markov Random Field (MRF) to describe both the probability of seeking correct HR patch and the compatibility between overlapped HR patches. In [3], Freeman *et al.* approximates the belief propagation (BP) inference in MRF with an example-based algorithm. It traverses all patches of the image in a raster-scan order and keeps edges of later-operated patch compatible with the ones of adjacent former-operated patches. However, it is the inherent weakness of this method that the error of one HR patch will propagate to the neighbor patches. To handle this problem, we propose a confidence growing model in this paper. Instead of the raster-scan order, we firstly choosing *confident* LR patches as starting points according to their uniqueness and then traverse across the image following the 4-neighbor seed growing scheme. As a result, with predefined overlap constraints, the propagated error can be eased.

The rest of this paper is organized as follows. In Section 2, we describe the confidence growing model in details. In Section 3, we explain how to learn the dictionary of LR/HR patch pairs. Experimental results are analyzed in Section 4. Conclusions and future work are given in Section 5.

* Corresponding author's email: zcqin@buaa.edu.cn. This work was partially funded by the National Science Foundation of China No. 61305047 and the SRF for ROCS, SEM, China.

2. CONFIDENCE GROWING MODEL

In example-based methods for super-resolution, local image information alone may not be sufficient to predict the missing high-resolution details, spatial neighborhood information is also crucial. For example, Freeman *et al.* [3, 4] used Markov network to model such spatial relation in HR domain by observing LR images. Based on the Markov model, a fast and simple one-pass, training-based super-resolution algorithm is proposed for creating plausible high frequency details in zoomed images. The iteration is discarded to speed up the process. However, the starting point will significantly influence the result of super-resolution. If there is a deviation of patch-matching resulted from noise at the very beginning, the later patch-matching will be disturbed. Better starting points are necessary therefore the patch pairs could be less affected by noise. We define these kinds of LR patches as *confidence patches*.



(a) Confidence patch (b) Normal patch

Fig. 1. Examples of a confidence patch and a normal patch. The patch in (a) is taken from the boundary of the chip and the patch in (b) is from the finger. Both of them are marked in red and details are shown in red squares at bottom.

2.1. Confidence of Patch

In HR-LR matching process, it should start from the patches with richness in information. For example, a corner, which always involves two edges, has the capability of including more information than a normal non-interest point. As it is a corner, information entropy will be larger than other patches. The information loss during the down-sampling will be less than others. Hence this leads to reduction of the error rate (e.g., see Fig. 1). Therefore, beginning from these carefully selected points, the probability of error spreading to other patches will be much smaller.

Our approach is based on the observation that for LR patches without salient texture, it is difficult to choose the corresponding high-resolution patches. Patches with less texture information may have more low-resolution patches from the dictionary than the ones with distinctive texture, although retrieved ones in fact have different high resolution patterns. It is thus difficult to reconstruct such positions with correct HR patches. On the contrary, patches with distinctive texture

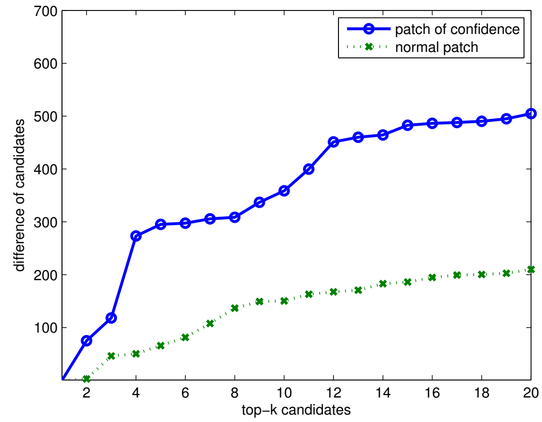


Fig. 2. Difference between the query patch and candidate patches. The x-axis denotes to the top 20 candidates, and the y-axis denotes to Euclidean distance of these candidates to the query low-resolution patch. For normal patches, the gaps of distance of candidates are smaller than confident patches.

can retrieve correct HR patch from the dictionary with a much higher probability.

Following this observation, we start the reconstruction from positions whose LR patches having distinctive texture. Then we grow them by reconstruct neighboring patches. This result of confidence is propagated to the whole image. In this paper, we use Kanade-Lucas-Tomasi (KLT) corner detector [12, 15] to find confident patches. We calculate the histogram difference between query patch and candidate patches from training set (whose details are available in Section 4) and the comparison results are demonstrated in Fig. 2.

2.2. Confidence Growing

The processing order will make a difference in high-resolution results. Raster-scan order is the most basic and easiest way. However, the noise will spread from the upper left of the image to the bottom right side of the image. If one LR patch gets its corresponding HR patch matching wrong, such information will lead to other wrong matching. To avoid such problems as far as possible, we employ a growing strategy. We begin with picking some key points and do the matching process in a different order. To be precisely, we process neighboring patches given the starting points. The patches far from the start points will be processed much later. Such a strategy for scanning patches is called confidence growing. For example, Fig. 3 shows an illustration of how the patches will be chosen based on some initial confident patches.

Following the method in [3], we define the overlap constraint as minimizing the difference of the overlap between two neighbored patches. Overlap constraint helps to determine the high-resolution patch from a set of candidates. In

Fig. 3. An illustration of the confidence growing model on the *Maddie* image. We show how the region expended from image in the bottom layer to the upper layer.

our method, after retrieve several low-resolution patches from the dictionary, we calculate the difference of these candidates from the query patch and choose the most compatible patch from them. As a result, former-processed patch can help choose right later-processed patch. The overlap constraint is important to guarantee smoothness of reconstructed HR images. In [4], overlap constraint is regarded as a compatibility matrix defined by Equation (1).

$$\phi_{ab}(i,j) = \exp\left(-\frac{d_{ij}(ab)^2}{2\sigma}\right) \quad (1)$$

where $d_{ij}(a,b)$ is the sum of squared differences between patch candidates a and b in their overlap regions at nodes i and j (the details of Markov model is introduced in [4]). σ is a predefined noise parameter.

3. HIERARCHICAL CLUSTERING

For example-based super-resolution, we need a database of LR-HR pairs for training. In this paper, we use the images from *Berkeley Segmentation Database*, by which we normalize every patch so that each pair in training represent the patches with same contrast. For each patch pair, we calculate the mean of the pixels of a patch and use $L2$ normalization (Equation 2).

$$\tilde{y} = \frac{y - \bar{y}}{\|y - \bar{y}\|_2} \quad (2)$$

Due to the redundancy among normalized patches, i.e., patch pairs in fact have similar texture in both low resolution and high resolution, we use the center patch pairs to represent a set of similar patches pairs by hierarchical clustering. Each patch pair vector C_i is defined as Equation 3.

$$C_i = L_i + H_i \quad (3)$$

where L_i denotes to low-resolution patch vector and H_i denotes to high-resolution patch vector.

By hierarchical clustering, each time two pairs (C_i , C_j) are gathered to one cluster. We search for the min Euclidean distance, which also used to retrieve the nearest neighbors in matching process, between two cluster centroid/elements, which means the corresponding two patch pairs are most similar with each other in Euclidean space. The nearest two patches will be merged into one. Then the distance from new merged cluster to other cluster/patches, will be calculated, and the merge process repeat. When the nearest distance between clusters/patches is further than the threshold we set, the merge process stops.

4. EXPERIMENTAL STUDIES

In this section, we describe the experiments in details and compare our method to several other state-of-the-art super-resolution algorithms.

4.1. Experimental Setting

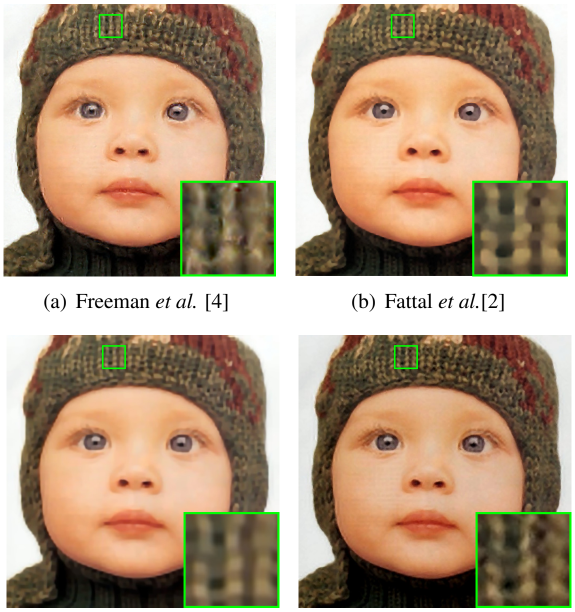
We extract 100,000 HR-LR patches from the *Berkeley Image Segmentation Database* [10] to build the training dataset. Specifically, we collect LR patches with size 3×3 and HR patches with size 4×4.

The corresponding lower spatial scale images are generated by blurring and downsampling the high-resolution images. Then we obtain some standard test images which are accompanied with ground truth from websites. As the conventional setting, we run most experiments with magnification factor 4. And one can run the algorithm recursively if results of greater magnification factor are required.

4.2. Experimental Results

We compare the new model to several other classical methods on the given standard training dataset. First, we test on the Infant image with magnification factor 4. The methods for comparisons are Freeman's Markov Random Field (MRF) super resolution method [4], Fattal's imposed statistics method [2] and Shan's fast method [11]. From Fig. 4, it is clear to see that, method in [4] produces some unrealistic textures. Though methods in [2] and [11] obtains more pleasing details, but the resolution is still lower than our method. The confidence growing model is obviously with clearer edges in detail.

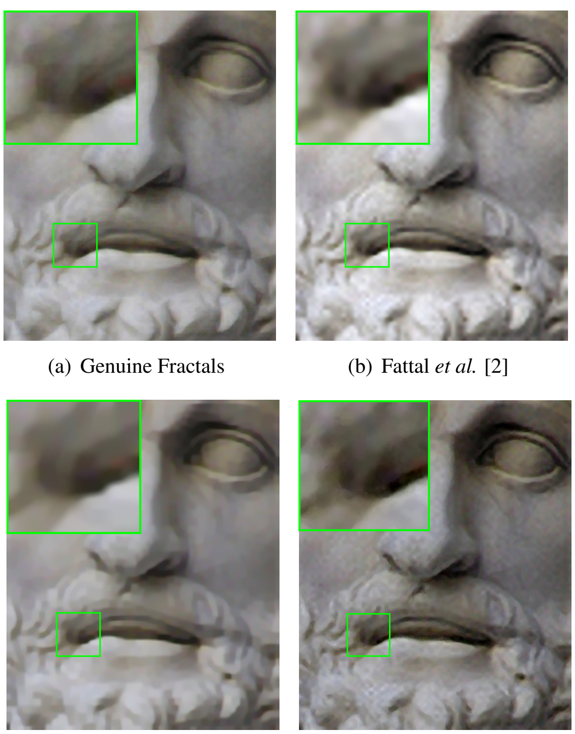
Then we increase the magnification factor to 8 and test it on the Statue image. We have done one more comparison experiment with the Genuine Fractals ^{TM}, which is a leading commercial plug-in for the Adobe Photoshop. Results are



(a) Freeman *et al.* [4] (b) Fattal *et al.*[2]

(c) Shan *et al.* [11] (d) Confidence Growing

Fig. 4. Experimental results of the *infant* image.



(a) Genuine Fractals (b) Fattal *et al.* [2]

(c) Shan *et al.*[11] (d) Confidence Growing

Fig. 5. Experimental results on the *statue* image.

shown in Fig. 5. We can find that, in our result, the contour of the statue's mouth is sharper than other methods.

To quantitatively evaluate the performance of our method and others, we adopt 3 conventional measures for super resolution: root mean square error (RMSE), peak signal to noise ratio (PSNR) and structure similarity (SSIM) [16]. Different measures of the infant image is listed in Table 1. From this table, we can easily see that our method achieve the smallest RMSE, the greatest PSNR and the greatest SSIM, thus outperforms other methods fore-mentioned.

We also compare the computational time of our model to other state-of-the art methods based on three benchmark test images. We run all these algorithms on the same machine. In order to make a fair comparison, we ignore training time and only consider the time of test procedure. The experimental results are shown in Table 2.

Method	PSNR	RMSE	SSIM
Freeman *et al.* [4]	22.0480	20.1437	0.5311
Fattal *et al.* [2]	22.0685	20.0962	0.5829
Shan *et al.*[11]	24.3879	15.3870	0.6494
Our Method	**24.7575**	**14.7457**	**0.6696**

Table 1. Experimental results in different measures on the *infant* image

Picture	Our method	MRF	Sparse
Infant	22.32s	92.45s	213.50s
Statue	22.34s	89.34s	543.22s
Maddie	22.23s	57.07s	1009.33s

Table 2. Computational complexity comparisons on given three images.

5. CONCLUSIONS AND FUTURE WORK

In this paper, based on the previous work of Freeman *et al.* [4], we propose a confidence growing model for single image super-resolution. We developed a new way of scanning LR-HR patch matchings to increase the model performance. Our model inherits the property of fast speed from example-based methods and improves the reconstructed ability of high resolution. Experimental results on a few benchmark test images show that the new proposed model outperforms a few classical models and with good efficiency.

Future work will focus on generalizing the algorithm to deal with less and compact patch dictionary without losing much accuracy. We will also improve the efficiency of the model to make it more applicable, even in real-time applications.

6. REFERENCES

[1] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, 2004.

[2] R. Fattal. Image upsampling via imposed edge statistics. *ACM Transactions on Graphics*, 26(3):95, 2007.

[3] W. Freeman, T. Jones, and E. Pasztor. Example-based super-resolution. *Computer Graphics and Applications*, 22(2):56–65, 2002.

[4] W. Freeman, E. Pasztor, and O. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, 2000.

[5] B. S. Glasner, D. and I. M. Super-resolution from a single image. In *ICCV*. IEEE, 2009.

[6] R. Hardie, K. Barnard, and E. Armstrong. Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633, 1997.

[7] H. Hou and H. Andrews. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26:508–517, 1978.

[8] X. Li and M. Orchard. New edge directed interpolation. In *ICIP*. Manchester, UK, 2000.

[9] R. Liao and Z. Qin. Image super-resolution using local learnable kernel regression. In *ACCV*, volume III, LNCS 7726, pages 349–360, 2013.

[10] 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*, volume 2, pages 416–423. IEEE, 2001.

[11] Q. Shan, Z. Li, J. Jia, and C. Tang. Fast image/video upsampling. In *ACM Transactions on Graphics*, volume 27, page 153. ACM, 2008.

[12] J. Shi and C. Tomasi. Good features to track. In *CVPR*, pages 593–600, 1994.

[13] J. Sun, Z. Xu, and H. Shum. Image super-resolution using gradient profile prior. In *CVPR*, pages 1–8. IEEE, 2008.

[14] S. Thurnhofer and S. Mitra. Edge-enhanced image zooming. *Optical Engineering*, 35:1862, 1996.

[15] C. Tomasi and T. Kanade. Detection and tracking of point features. In *Carnegie Mellon University, Tech. Report*, volume CMU-CS-91-132, 1991.

[16] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing,*, 13(4):600–612, 2004.

[17] J. Yang, Z. Lin, and C. S. Fast image super-resolution based on in-place example regression. In *CVPR*, pages 1059–1066. IEEE, 2013.