Local Consistency Constrained Adaptive Neighbor Embedding for Text Image Super-Resolution

Wei Fan, Jun Sun, Satoshi Naoi
Fujitsu Research and Development Center Co., Ltd.
Email: {fanwei, sunjun, naoi}@cn.fujitus.com

Akihiro Minagawa, Yoshinobu Hotta
Fujitsu Laboratories Ltd., Japan
Email: {minagawa.a, y.hotta}@jp.fujitsu.com

***Abstract*—This paper proposes a robust single-image super-resolution method for enlarging low quality camera captured text image. The contribution of this work is twofold. First, we point out the non-local reconstruction problem in neighbor embedding based super-resolution by statistical analysis on an empirical data set. Second, we introduce a local consistency constraint to explicitly regularize the linear reconstruction process, and adaptively generate the most possible candidates for the high-resolution image patch. For the non-consistent candidates, we rely on its adjacent overlapping patches for capability verification. Experimental results demonstrate that our solution produces visually pleasing enlargements for various text images.**

I. Introduction

With the Internet flourishing and the rapid progress in hand-held photographic devices, the scope of document imaging has increased. However, due to the low cost mobile cameras and server storage limitation, most document images exist in a poor quality degraded from the source, making the immediate recognition practically impossible. Super-resolution provides an algorithmic solution to the resolution enhancement problem. It refers to the process by which a higher-resolution enhanced image is synthesized from one or more low-resolution images.

This paper focuses on the issue of increasing the resolution of a single text image. Text images are a distinct class of images widely different from natural images. We study the distribution of small image patches in the text region and see what kinds of local primitive structures (e.g., stroke edges, blobs, or corners) are likely to occur in a document.

A. Related previous work

There has been a substantial amount of previous work in super-resolution for text images. Simple interpolation based methods such as cubic-spline interpolation suffer from blurring edges and image details, since the smoothing kernel is indiscriminate between text and non-text regions. The sharpened interpolation will introduce ringing or jaggy artifacts, especially along salient edges. Dalley *et al.* [3] employed a training-based method in a Bayesian framework. A database is built that indicates which high-res patch should be output given an input low-res patch. Park *et al.* [1] presented a prior model via Markov Random Field (MRF) framework for text image super-resolution, which can be benefited from strong prior knowledge of the image class. Banerjee *et al.* [12] presented an edge-directed super-resolution algorithm for document images without using any training set while explicitly encoding the text gradient information in a MRF framework.

Taking local information and spatial neighborhood effects into account, Freeman *et al.* [4] developed a one-pass example-based super-resolution algorithm which obtains sharper edges and richer textures. One disadvantage is that it introduces non-photo-realistic artifacts and amplifies noises from the input images. Assuming that image patches in the low- and high-res images share the similar local geometry, Chang *et al.* [5] proposed super-resolution through neighbor embedding in which the high-res test images can be estimated with a set of optimally weighted training patch pairs. However, some recent work in this field [7][8] pointed out that neighborhood preservation assumption for low- and high-res patches does not always hold, so that ambiguous reconstruction frequently occurs. An extension of [5] by Fan *et al.* [6] proposed an image hallucination method using neighbor embedding over visual primitive manifolds, showing that visual primitives [10] are more reliable for linear reconstruction.

B. The proposed method

Our method benefits from the above work of learning based super-resolution. To overcome the ambiguous linear reconstruction problem, we explicitly introduce the manifold consistency constraint to regularize the neighbor embedding process. For local consistent image patches such as straight edges, traditional linear reconstruction is used. For those non-consistent local patches corresponding to complex high-res text region, we rely on its adjacent patches for compatibility verification through a Markov network. From this point of view, the process of neighbor embedding is performed in an adaptive way. Figure 1 shows a flowchart of our approach.

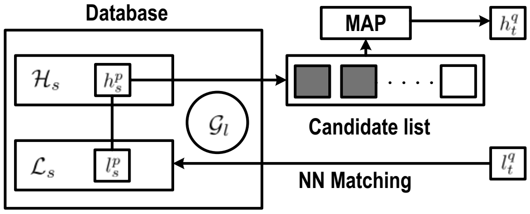


Fig. 1. A flowchart of the proposed approach

In the learning phase, large volumes of text primitive patches are extracted from both the low-res and high-res training

978-0-7695-4661-2/12 $26.00 © 2012 IEEE
DOI 10.1109/DAS.2012.52

images $\mathcal{L}_s$ and $\mathcal{H}_s$. The low-res patches are organized as a local consistency graph $\mathcal{G}_l$. During the synthesis phase, each patch l_t^q from a low-res image is presented to the system. A set of adaptively selected high-res candidates are retrieved from the training set by nearest neighbor (NN) matching. Each element h_t^q in the target high-res image comes from the MAP estimation of a Markov network based on the candidate list.

We first briefly introduce the neighbor embedding method [5] in section 2. We also point out the non-local reconstruction problem by statistical analysis. Details of the local consistency constrained solution are described in section 3. Experimental results are analyzed in section 4.

II. Neighbor embedding

A. The problem setting

The single-image super-resolution problem that we want to solve can be formulated as follows. Given a low-res image $\mathcal{L}_t$ as input, we estimate the target high-res image $\mathcal{H}_t$ with the help of a training set of one or more low-res images $\mathcal{L}_s$ and the corresponding high-res images $\mathcal{H}_s$.

In the neighbor embedding based super-resolution framework, each low-res or high-res image is represented as a set of small overlapping patches. We denote the sets of patches corresponding to $\mathcal{L}_t$, $\mathcal{H}_t$, $\mathcal{L}_s$ and $\mathcal{H}_s$ as $\{l_t^q\}_{q=1}^{N_t}$, $\{h_t^q\}_{q=1}^{N_t}$, $\{l_s^p\}_{p=1}^{N_s}$ and $\{h_s^p\}_{p=1}^{N_s}$, respectively. N_t and N_s are the number of patches in $\mathcal{L}_t$ (or $\mathcal{H}_t$) and $\mathcal{L}_s$ (or $\mathcal{H}_s$), which depend on the patch size and the degree of overlap between adjacent patches.

B. The neighbor embedding algorithm

Neighbor embedding based super-resolution reconstruction can be summarized in the following steps.

1. For each patch l_t^q in image $\mathcal{L}_t$
   (a) Find the set N_q of K nearest neighbors in $\mathcal{L}_s$.
   (b) Compute the reconstruction weights of the neighbors that minimize the error of reconstructing l_t^q,

$$\epsilon^q = \|l_t^q - \sum_{l_s^p \in N_q} \omega_{qp} l_s^p\|^2 \qquad (1)$$

   which is the squared distance between l_l^q and its reconstruction, subject to the constraints that $\sum_{l_s^p \in N_q} \omega_{qp} = 1$ and $\omega_{qp} = 0$ ($l_s^p \notin N_q$). Minimizing ϵ^q subject to the constraints is a constrained least squares problem which has closed-form solution.
   (c) Compute the initial high-res embedding h_t^q using the appropriate high-res features of the K nearest neighbors and the reconstruction weights.

$$h_t^q = \sum_{l_s^p \in N_q} \omega_{qp} h_s^p \qquad (2)$$

2. Construct the target high-res image $\mathcal{H}_t$ by enforcing the local compatibility and smoothness constraints between adjacent patches obtained in step 1(c).

C. The problem of non-local reconstruction

The underlying assumption of the neighbor embedding method is that small patches in the low-res and high-res images form manifolds with similar local geometry in the two corresponding feature spaces. However, image super-resolution is intrinsically an ill-posed problem since, theoretically, many high-res patches can give rise to the same low-res patch through the same degradation procedure. As Figure 2 shows, this one-to-multiple mapping from low-res to high-res feature space will violate the manifold assumption of neighbor embedding and cause non-local reconstruction problem.

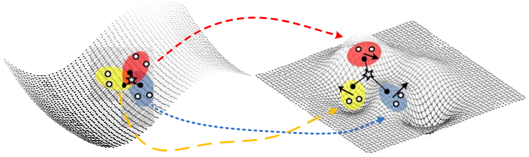


Fig. 2. One-to-multiple mapping from low-res (left) to high-res (right) feature space. The three nearest neighbors of the query patch (marked as a star) actually locate at different oriented linear regions in the high-res manifold.

We demonstrate this key observation by statistical analysis on an empirical data set. To measure the generalization capability of example-based pair matching, we define two terms. The first is *sufficiency*, which determines whether or not an input sample can find a good match in the training set. The second is *predictability*, which determines whether or not the high-res patch corresponding to the input sample's nearest neighbor in the training set is a good prediction of the target that we want to infer from the input sample.

To compare the pair matching accuracy, we use a Receiver Operating Characteristic (ROC) curve to demonstrate the tradeoff between match error and hit rate. For a given match error e, the hit rate h is the percentage of test data whose match error is less than e. Each test sample p's match error $e(p)$ is defined by a metric between p and the nearest sample p' in the training data. We define the match error as $e(p) = \frac{\|p-p'\|_2^2}{\|p\|_2^2}$. At a given match error, the higher hit rate represents the better sufficiency of the training dataset.

The prediction takes the following steps:

- Step 1: Find all those test samples whose nearest neighbor match error $e(p)$ is below a threshold.
- Step 2: For all the test samples in step 1, find its K nearest neighbors $p_1^{l'}, p_2^{l'}, \cdots, p_K^{l'}$ for each test sample p and the corresponding high-res patches $p_1^{h'}, p_2^{h'}, \cdots, p_K^{h'}$.
- Step 3: Find the nearest neighbor $q^{h'}$ among $\{p_i^{h'}, i = 1, \cdots, K\}$ for the test sample p's corresponding ground truth high-res patch q.

The prediction error is define as $e(q) = \frac{\|q-q^{h'}\|_2^2}{\|q\|_2^2}$

The above process is to measure both the sufficiency and predictability capability of using nearest neighbor matching in the low-res patch set to estimate the corresponding high-res patch. The following experiments also evaluate the performance for the inverse mapping, i.e. using high-res patch matching to estimate low-res patch.

An empirical data set, consisting of 10 document images captured by a digital camera, are divided equally into training images and test images. We collect two training patch sets $D_1 = \{D_1^h, D_1^l\}$ and $D_2 = \{D_2^h, D_2^l\}$ with different size. D_1 consists of 10^5 low- and high-res patch pairs and D_2 consists of 10^6 patch pairs. These patch pairs are uniformed sampled from the text region of the training images and their smoothed and down-sampled counterparts. About 10,000 test patches are randomly sampled from the testing images in a similar way.

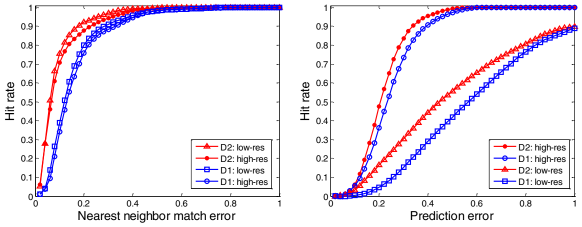


Fig. 3. ROC curves of nearest neighbor matching and prediction accuracy.

Two observations are found from the ROC curves in Figure 3. The nearest neighbor matching accuracy is similar for both the low-res patch set and high-res patch set. With an increasing number of training patches, higher matching accuracy, i.e. better sufficiency, can be achieved for the training data. However, the prediction accuracy is significantly lower when using low-res patch matching to find the high-res patch estimation than the inverse direction. In real scenario, it is infeasible to get the high-res patch as a query. However, for refining the training set, we can rely on the local relationship of high-res patches to find consistent regions on the two coupled manifolds.

III. The Proposed Method

A. Preprocessing

A high-res text image $\mathcal{H}$ (Figure 4(c)) is first blurred and sub-sampled to generate a corresponding low-res image $\mathcal{L}$ (Figure 4(a)). Applying an initial enhancement through bilinear interpolation to $\mathcal{L}$, we obtain an image $\mathcal{H}_l$ (Figure 4(b)) which has the same size as $\mathcal{H}$ but lacks the high-res details. In the training set, we only need to store the differences between $\mathcal{H}$ and $\mathcal{H}_l$ (Figure 4(d)), which correspond to the missing high-freq components caused by image degradation.

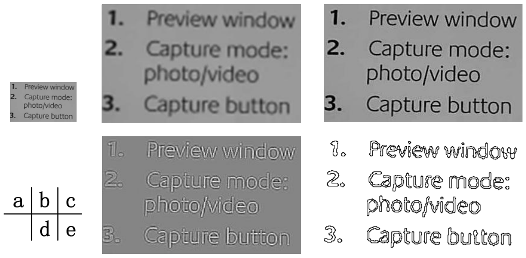


Fig. 4. Preprocessing steps for the training image

B. Training set preparation

An essential factor attributing to the success of example-based super-resolution approaches is how to construct a good training set, which is descriptive enough in giving useful information about the low and high-res relationships and is also compact enough for computational efficiency and good generalization.

When building our training set, we collect the patches centered on the text boundaries (Figure 4(e)) for three main reasons. First, text primitives are densely distributed over the text boundaries, while uniform background regions lack meaningful features to estimate the high-freq information. Focusing on the text regions can lead to significant speedup as fewer patches need to be transformed. Second, we believe the local neighborhood relationships between low-res and high-res text primitive patches in the two feature spaces are more consistent than those between general image patches. Third, the patch variation caused by translation can be reduced since each primitive we extract is centered on the text boundary.

To sum up, each example in the training set is in the form of a pair of text primitive patches. These pairs capture the statistical relationships that we are interested in.

C. Training set refinement by local consistency verification

The neighbor embedding based super-resolution assumes the reconstruction weights of a low-res patch should be similar with the weights of reconstructing the high-res counterpart by the corresponding high-res neighbors. However, this assumption can only hold in the locally consistent linear regions on the coupled manifolds. In the occasional case that multiple high-res patches with apparently different texture give rise to similar degraded low-res patches, the two respective linear regions are no longer consistent. If the reconstruction weights are forcibly estimated according to the low-res patches in such ambiguous region, the predicted high-res patch will severely deviate from the true target.

In the following subsection, we explicitly introduce a manifold consistency constraint to regularize the neighbor embedding process.

1) local cell construction: Different from [7], we refer the locally linear regions on the high-res manifold as local cells, since the mapping from high-res space to low-res space is more reliable in terms of local isometric preservation. Note nearby points in the low-res manifold may not lie on the same cell especially for those patches corresponding to complex texture. In the training set refinement step, each training patch pair $\{h_i, l_i\}$ is associated with a local cell $p_i = \{h_{1,i}, h_{2,i}, \cdots, h_{k,i}\}$ defined by the k nearest neighbors of h_i in $\mathcal{H}_s$.

2) consistent cell graph construction: We generate a graph representation $\mathcal{G}_l$ of the low-res patch set $\mathcal{L}_s$ by connecting any two patches l_i and l_j if their associated local cells p_i and p_j are consistent. Notice the consistency verification is only performed in the neighborhood of l_i and l_j, so the graph is relatively sparse.

We describe how to measure the consistency between two local cells p_i and p_j as follows.

3) local cell consistency estimation: Let us denote the data matrices for two cells p_i and p_j, respectively, as $X = [h_{1,i}, h_{2,i}, \cdots, h_{k,i}]$ and $Y = [h_{1,j}, h_{2,j}, \cdots, h_{k,j}]$, where each column of X or Y corresponds to one high-res patch in the corresponding local cells. The columns of X and Y define two linear subspaces $\mathcal{X} = span(X)$ and $\mathcal{Y} = span(Y)$ in the feature space. A distance measure for linear subspaces is the projection $\mathcal{L}_2$-norm:

$$dist_{\mathcal{L}_2}(\mathcal{X}, \mathcal{Y}) = \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2 \tag{3}$$

where $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ are the orthogonal projection matrices onto X and Y, respectively, and $\|.\|_2$ denotes the matrix $\mathcal{L}_2$-norm. The projection $\mathcal{L}_2$-norm is related to the largest canonical angle (or principal angle) between two subspaces. If the maximum canonical angle is small, the subspace are close to each other which means the two local cells are more consistent. To build the consistency graph $\mathcal{G}_l$, we simply compare the canonical angle of two nearby cells with an empirical threshold. A numerical stable algorithm to compute the canonical angles was proposed by Bjork and Golub [13] based on QR factorization of the data matrices X, Y and singular value decomposition (SVD).

D. MAP prediction of high-res image

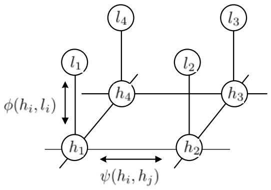
For a given low resolution image $\mathcal{L}$, we seek to obtain the *maximum a posteriori* (MAP) estimation of the posterior probability $P(\mathcal{H}|\mathcal{L}) = cP(\mathcal{L}|\mathcal{H})P(\mathcal{H})$ (the normalization, $c = \frac{1}{P(\mathcal{L})}$, is a constant over $\mathcal{H}$).

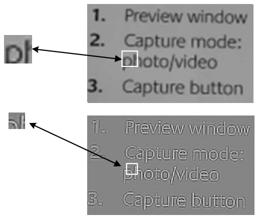
To make the MAP estimation tractable, we divide both the low- and high-res images into overlapping patches and model the spatial relationships between them using Markov network. In Figure 5, the circles represent network nodes, and the lines indicate statistical dependencies between nodes. We let the low-res image patches be observation nodes, l. For each input low-res patch, we select the K candidate high-res patches in a training dataset as the different states of the hidden nodes, h, that we seek to estimate. For this network, the probability of any given high-res patch choice for each node is proportional to the product of all sets of compatibility matrices ψ relating the possible states of each pair of neighboring hidden nodes, and vectors ϕ relating each observation to the underlying hidden states:

$$P(\mathcal{H}|\mathcal{L}) = c \prod_{(ij)} \psi(h_i, h_j) \prod_{(i)} \phi(h_i, l_i) \tag{4}$$

the first product is over all neighboring pairs of nodes, i and j. l_i and h_i are the observed low-res and estimated high-res patches at node i, respectively.

In the above MAP formulation, the prior probability $P(\mathcal{H})$ is encoded by the K candidate high-res patches retrieved from the training set. Traditional methods [4][10] typically set the candidate number K for each node as a constant. However, for text image super-resolution, a large proportional of patches to be estimated are located around text strokes which correspond to consistent local cells of the manifold. In such case, the neighbor embedding based reconstruction result can be reliably regarded as an optimal candidate. Otherwise, in ambiguous regions covering non-consistent local cells, all possible candidates should be kept and we rely on the Markov network to learn the most appropriate one which is compatible with its adjacent patches.




1. Preview window
2. Capture mode: photo/video
3. Capture button

1. Preview window
2. Capture mode: photo/video
3. Capture button

Fig. 5. Markov network model for super-resolution. The low-res patches at each node l_i are the observed input. The high-res patch at each node h_i is the quantity we want to estimate.

The procedure of adaptively selecting K candidates is illustrated as below.

For each query low-res patch l_t^q in image $\mathcal{L}_t$

- Find the set N_q of $K = 15$ nearest neighbors in $\mathcal{L}_s$.
- Referring to the consistent cell graph $\mathcal{G}_l$, we divide N_q into M subsets $\{N_q^1, N_q^2, \cdots, N_q^M\}$ ($M \leq K$) so that: a) The patches in any N_q^i are connected by certain edges as a subgraph; b) There is no edge connecting any subgraph N_q^i and N_q^j ($i \neq j$).
- For each subset N_q^i, if it contains multiple patches, we use neighbor embedding based method to reconstruct an optimal candidate. If the subset contains an isolate patch, its corresponding high-res patch in $\mathcal{H}_s$ is set as the candidate.
- After processing all M subsets, we get M candidates and update $K := M$.

In our implementation, The compatibility function $\psi(h_i, h_j)$ in (4) is defined by the compatibility of adjacent patches, $\psi(h_i, h_j) = \exp(-d(h_i, h_j)/\sigma_d^2)$, where $d(h_i, h_j)$ is the Sum Squared Difference (SSD) of the overlapping region between h_i and h_j and σ_d is a tuning variance. We use a similar quadratic penalty on differences between $h_i(k)$ and l_i to specify $\phi(h_i, l_i)$. The optimal MAP solution of (4) is obtained by running the Belief Propagation (BP) algorithm [14] with some biases.

IV. Experiments

We demonstrate the performance of our method on the document images captured by cell-phone cameras and web-cameras with relatively low imaging quality. Note that we only focus on the image intensity channel because the humans are more sensitive to the brightness information in a document image. The chrominance channels are simply interpolated by a bicubic function in the final stage.

To build a promising training set, we collect 20 high-res images by setting the capture device to its highest resolution. Some examples selected from the training set are illustrated in Figure 6. The corresponding low-res images are produced by blurring and downsampling the high-res images. The PSF of the blurring kernel is a Gaussian function with a standard variance of 1.4 corresponding to a magnification factor of 3. About 5,400,000 text primitive patches have been extracted from these training images.

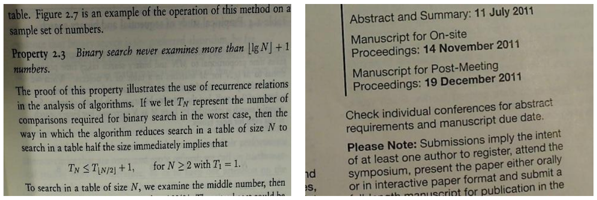


table. Figure 2.7 is an example of the operation of this method on a sample set of numbers.

Property 2.3 *Binary search never examines more than* $\lfloor \lg N \rfloor + 1$ *numbers.*

The proof of this property illustrates the use of recurrence relations in the analysis of algorithms. If we let T_N represent the number of comparisons required for binary search in the worst case, then the way in which the algorithm reduces search in a table of size N to search in a table half the size immediately implies that

$$T_N \leq T_{\lfloor N/2 \rfloor} + 1, \qquad \text{for } N \geq 2 \text{ with } T_1 = 1.$$

To search in a table of size N, we examine the middle number, then

Abstract and Summary: **11 July 2011**

Manuscript for On-site
Proceedings: **14 November 2011**

Manuscript for Post-Meeting
Proceedings: **19 December 2011**

Check individual conferences for abstract requirements and manuscript due date.

Please Note: Submissions imply the intent of at least one author to register, attend the symposium, present the paper either orally or in interactive paper format and submit a

Fig. 6. Example images selected from the training set.

We represent each low-res text primitive by a 7×7 patch P^l sampled from $\mathcal{H}_l$ (Figure 4(b)) and each high-res text primitive by a 5×5 patch P^h sampled from $\mathcal{H} - \mathcal{H}_l$ (Figure 4(d)). The corresponding low-res and high-res image patches are properly aligned by their geometrical centers in the image plane. We normalize P^l to get $\hat{P}^l$ according to the formula $\hat{P}^l = \frac{1}{c^l}(P^l - d^l)$, where DC bias d^l is estimated by the mean $E[P^l]$. The contrast c^l is estimated by $E[|P^l - E[P^l]|]$.

To show the effectiveness of our method, we compare it with several common methods, including bilinear, cubic-spline interpolation and Chang's method [5]. 10 test images are collected in a similar way as the training set. Figure 7 shows two cropped regions of the $3X$ enlargement results.

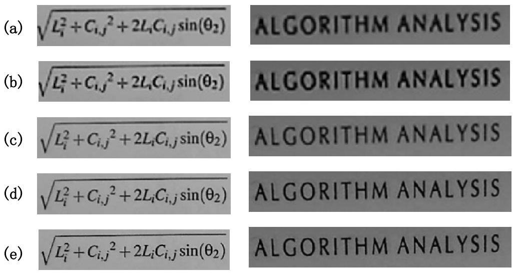
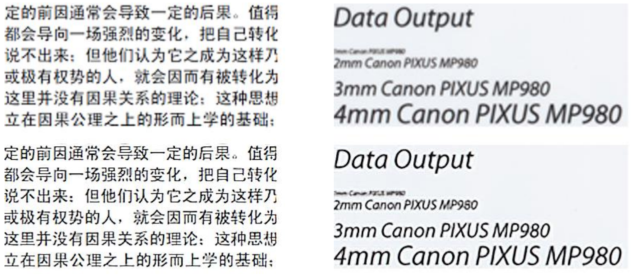


Fig. 7. Super-resolution results of four methods on two examples with $3X$ manification: (a) bilinear interpolation; (b) cubic-spline interpolation; (c) Chang's method; (d) our method; (e) ground-truth.

Figure 8 show another two examples of $3X$ enlargement results comparing Bicubic interpolation and our method.



定的前因通常会导致一定的后果。值得
都会导向一场强烈的变化，把自己转化
说不出来：但他们认为它之成为这样乃
或极有权势的人，就会因而有被转化为
这里并没有因果关系的理论：这种思想
立在因果公理之上的形而上学的基础：

Data Output

2mm Canon PIXUS MP980
3mm Canon PIXUS MP980
4mm Canon PIXUS MP980

定的前因通常会导致一定的后果。值得
都会导向一场强烈的变化，把自己转化
说不出来：但他们认为它之成为这样乃
或极有权势的人，就会因而有被转化为
这里并没有因果关系的理论：这种思想
立在因果公理之上的形而上学的基础：

Data Output

2mm Canon PIXUS MP980
3mm Canon PIXUS MP980
4mm Canon PIXUS MP980

Fig. 8. Super-resolution results of two methods on two examples with $3X$ manification: (Top row) bicubic interpolation; (Bottom row) our method.

TABLE I
RMSE AND RMSEB OF FOUR SUPER-RESOLUTION METHODS

	bilinear	cubic-spline	Chang's [5]	our method
RMSE	21.7	20.3	15.8	12.4
RMSES	35.4	32.2	24.7	23.1

We quantitatively demonstrate the superiority of the proposed model using two measures, RMSE (Root Mean Squared Error) between the true test image and the super-resolved result, and RMSEB (Root Mean Squared Error of Binarized images). Table 1 shows the average RMSE errors over 10 test images for different methods.

From both the quantitative measurements and qualitative comparison, our proposed method improves the performance of neighbor embedding based super-resolution methods.

V. Conclusion

This paper proposes a robust single-image super-resolution method for enlarging low quality camera captured text image. A local consistency constraint is introduced to explicitly regularize the linear reconstruction, and adaptively generate the most possible candidates for the high-res image patch. Experimental results demonstrate that our solution produces visually pleasing enlargements for various text images.

References

[1] J. Park, Y. Kwon and J. Kim, "An Example-based Prior Model for Text Image Super-resolution," in *International Conference on Document Analysis and Recognition*, 2005.

[2] H. Kim, "Binary operator design by k-nearest neighbor learning with application to image resolution increasing," in *International Journal Imaging Systems and Technology*, 2000, vol. 11, pp. 331-339.

[3] G. Dalley, B. Freeman, and J. Marks, "Single-frame text super-resolution: A bayesian approach," in *International Conference on Image Processing*, 2004.

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

[5] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2004, pp. 275-282.

[6] W. Fan and D. Yeung, "Image hallucination using neighbor embedding over visual primitive manifolds," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1-7.

[7] B. Li, H. Chang, S. Shan, and X. Chen, "Locality preserving constraints for super resolution with neighbor embedding," in *Proc. IEEE Int. Conf. Image Processing*, Nov. 2009, pp. 1189-1192.

[8] K. Su, Q. Tian, Q. Que, N. Sebe, and J. Ma,, "Neighborhood issue in single-frame image super-resolution," in *IEEE Internat. Conf. on Multimedia and Expo*, Amsterdam, 2005, pp. 1122-1125.

[9] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1-8.

[10] J. Sun, N. Zheng, H. Tao, and H. Shum. "Image hallucination with primal sketch priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, volume 2, pages 729-736, 18-20 June 2003.

[11] T. Chan, J. Zhang, J. Pu, and H. Huang, "Neighbor embedding based super-resolution algorithm through edge detection and feature selection," *Pattern Recognition Letters*, vol. 30, pp. 494-502, 2009.

[12] J. Banerjee and C. Jawahar, "Super-resolution of Text Images Using Edge-Directed Tangent Field," in *The Eighth IAPR Workshop on Document Analysis Systems*, 2008.

[13] A. Bjorck and G. Golub. "Numerical methods for computing angles between linear subspaces," *Mathematics of Computation*, 27(123):579-594, 1973.

[14] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference," 1988.