

SoftCuts: A Soft Edge Smoothness Prior for Color Image Super-Resolution

Shengyang Dai*, *Student Member, IEEE*, Mei Han, Wei Xu, Ying Wu, *Senior Member, IEEE*,
Yihong Gong, Aggelos K. Katsaggelos, *Fellow, IEEE*.

Abstract—Designing effective image priors is of great interest to image super-resolution, which is a severely under-determined problem. An edge smoothness prior is favored since it is able to suppress the jagged edge artifact effectively. However, for soft image edges with gradual intensity transitions, it is generally difficult to obtain analytical forms for evaluating their smoothness. This paper characterizes soft edge smoothness based on a novel *SoftCuts* metric by generalizing the *Geocuts* method [1]. The proposed soft edge smoothness measure can approximate the average length of all level lines in an intensity image. Thus the total length of all level lines can be minimized effectively by integrating this new form of prior. In addition, this paper presents a novel combination of this soft edge smoothness prior and the alpha matting technique for color image super-resolution, by adaptively normalizing image edges according to their α -channel description. This leads to the adaptive *SoftCuts* algorithm, which represents a unified treatment of edges with different contrasts and scales. Experimental results are presented which demonstrate the effectiveness of the proposed method.

Index Terms—SoftCuts, super-resolution, edge smoothness, α -channel description.

I. INTRODUCTION

IMAGE super-resolution (SR) [2] is the problem of recovering high resolution (HR) images from low resolution (LR) inputs. This problem is of great practical interest, especially to video communication, object recognition, HDTV, image compression, etc. This is a very challenging task, especially when only one single low resolution input image is available (see for example [3]), as is very often the case. This paper mainly focuses on this case.

In theory, the generation process of low resolution images can be characterized by smoothing and down-sampling the high resolution scenes with low-quality image sensors. Recovering the original high resolution image from low resolution inputs is an inverse process. One intuitive approach to address this inverse problem is to minimize the reconstruction error, which is defined as the difference between the observation and the result obtained by putting the recovered HR image through the same generation process. In other words, the result which can produce the closed low resolution image to the

Shengyang Dai, Ying Wu, and Aggelos K. Katsaggelos are with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208, United States, Tel: (847) 491-4466, Fax: (847) 491-4455, Email: {sda690, yingwu, agk}@eecs.northwestern.edu. Mei Han is with Google Inc., Mountain View, CA 94043, United States, Tel: (650) 214-6606, Fax: (650)253-8447, Email: meihan@google.com. Wei Xu and Yihong Gong are with NEC Laboratories America, Inc., Cupertino, CA 95014, United States, Tel: (408) 996-0561, Fax: (408) 863-6099, Email: {xw, ygong}@sv.nec-labs.com.

observed one is preferred. Such a reconstruction error can be efficiently optimized by the back-projection method [4] in an iterative way. However, researchers have found that the super-resolution problem is essentially under-determined [5], [6]. Given the LR input or inputs, there might be multiple solutions that can minimize the reconstruction error. Thus, by simply minimizing the reconstruction error, the result may converge to an unsatisfying solution. To overcome this difficulty, it is necessary to regularize the under-determined inverse problem by incorporating effective image priors.

Simple interpolation-based methods, such as bilinear or bicubic interpolation tend to produce high resolution images with jagged edges, which is also a common artifacts for many super-resolution algorithms. This observation motivates us to design an *edge smoothness prior* that favors an HR image with smooth edges. In addition, this prior is also consistent with human perception. However, designing and incorporating an effective edge smoothness prior is a very challenging task due to the following two main difficulties in practice:

1. It is very difficult to quantitatively characterize the smoothness of an edge, especially for edges in natural images. In most real situations, an image edge exhibits a gradual intensity transition, which is in general much more complex than a simple geometric curve. In this paper, we refer to such an edge as a *soft edge* (see Fig. 1(b) for an example), as opposed to a *hard edge*, as shown in Fig. 1(a). Working with soft edges poses an extra difficulty in measuring and quantifying the edge smoothness in an analytical form.

2. Color image edges are determined by the information from all three color channels simultaneously. Besides, edges in natural images exhibit a large variation with different contrasts and scales. How to explore the 3D color information and treat those various edges in a unified way is of great importance to color image SR.

This paper is mainly focused on addressing the above two issues. The main contributions are as follows:

1. To measure and quantify the edge smoothness, we introduce the *Geocuts* method [1], which can approximate the Euclidean length of a *hard edge* with a *cut metric* on the image grid. More importantly, in order to handle soft edges, we propose an extension of the *SoftCuts* method based on a *soft edge cut metric*, which can measure the smoothness of soft edges in an intensity image. This extension is significant because it leads to a new analytical form for the soft edge smoothness prior. The new smoothness measure is used to regularize the objective function of the SR task, and produce impressive results. This metric also has a nice geometric

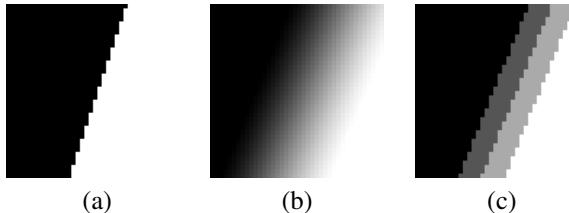


Fig. 1. Examples of (a) hard edge, (b) soft edge, and (c) level lines (boundaries between different intensities) for (b) with quantization step size equal to 64 on image intensities.

property, since it approximates the average length of all level lines in the intensity image, where a level line represents the boundary between pixels with intensities smaller and larger than a given value (an example is shown in Fig. 1(c)).

2. To handle various edges in color images, a novel image representation scheme is proposed based on the alpha matting technique from the computer graphics literature. We transform the problem of color edge super-resolution to a combination of alpha matting decomposition and α -channel super-resolution. This makes possible the application of the soft edge smoothness prior to natural color images.

3. To process the entire image in a unified way, an adaptive strategy is further proposed to utilize the *SoftCuts* for all image edges simultaneously by using an α -channel image description obtained with a multiple layer image matting decomposition method [7]. Due to this decomposition, the *SoftCuts* regularization term for different edges is normalized, such that the smoothness measure of all level lines in the α channel has the same weight.

The proposed SR algorithm has the following benefits: (1) Due to the geometry property of the proposed *SoftCuts* measure, the length of all image level lines can be minimized simultaneously for the super-resolution task. Thus, results with smooth edges can be obtained. In the mean time, the edge-preserving property of the proposed prior term can also make the resulting edges have sharp transitions. (2) All three color channels are utilized simultaneously with the α -channel SR scheme, and the adaptive strategy provides a unified treatment of edges with different contrasts and scales.

In this paper, the related work is summarized in Sec. II. The *SoftCuts* method is proposed in Sec. III, and the geometric explanation is presented. This *SoftCuts* metric can be used as a regularization term in the super-resolution problem to favor soft smooth edges. To make use of this term adaptively, we first introduce the α -channel edge/image description method in Sec. IV. Based on that, an adaptation scheme is derived in Sec. V, which applies the *SoftCuts* regularization term with the same weight on different edge transitions. Experiments are shown in Sec. VI and Sec. VII concludes this paper.

II. RELATED WORK

Due to the under-determined nature of the super-resolution problem, an image prior needs to be utilized to regularize the restoration process. Extensive work has been done on super-resolution given multiple LR images [2], [4], [8], [9]. The prior term (or image regularization term) in these works could also be applied to the single image SR problem directly. So in this section, we will also review the methods used when

multiple LR images are available. In general, having multiple LR input images imposes more constraints on the likelihood term (or image fidelity term). Thus, it is argued that the use of an effective image prior is even more critical for single image SR.

Existing works on the use of image priors for super-resolution tasks can be placed into two categories: priors from image modeling and priors from learning.

In the first case, based on the observation that neighboring pixels tend to have the same color, the image smoothness prior is widely explored, such as in various filtering or interpolation algorithms (e.g., bilinear and bicubic interpolation). Such methods usually produce blurry images, since the image edges are also smoothed. Edge-preserving regularization terms [9], [10], [11] are designed to address the over-smoothness problem at image boundaries. In [9], the bilateral prior is used for single channel images. The high order neighborhood is also investigated in this work, but from a different perspective from ours. This work is extended to color images [12] by optimizing a multi-term cost function, which is a combination of regularization terms on illuminance and chrominance. In [10], the sparse derivative prior is applied, which can preserve sharp edges by modeling the image derivative with a heavy-tailed distribution, instead of traditional Gaussian distribution. In [11], a TV prior is used along with a variational approximation for obtaining the solution. Another popular approach is using edge directed interpolation [13], [14], [15], which is trying to infer sub-pixel edges positions, to further prevent cross-edge interpolation. Locating the edge position with high precision is of great importance for those methods. Another common observable artifact of super-resolution algorithms is the jagged edge effect. It is reasonable to assume an edge smoothness prior without any other prior knowledge on the image. This prior is also consistent with human perception, which seems to also favor smooth curves in natural images. Based on this assumption, a lot of algorithms have been proposed to obtain smooth edges. Various techniques are investigated in the literature to obtain smooth image boundaries, such as Level-set [16], multi-scale tensor voting [17], and Snake-based vectorization [18] techniques.

Instead of modeling the image prior explicitly, many researchers try to learn the prior from natural images, especially by learning from image exemplars directly. One common way is to infer the missing high frequency information for the HR images for each position based on the middle frequency information. This approach is adopted in [19] for the image restoration problem. The Markov Random Fields model [20], [21], [22] usually is utilized to enforce spatial consistency. Extension of this research direction includes video SR [23] and learning from domain-specific exemplars [24]. In such methods, two key problems usually need to be addressed. The first is how to search HR candidate patches efficiently. Locality Sensitive Hashing [25] and KD-tree [24] are applied to speed up the searching. The second is how to solve the optimization problem efficiently. Belief propagation can be used, and the image primal sketch [26] method can simplify the problem to a chain structure. Other learning algorithms can also be used

to infer the high frequency information, such as locally linear embedding [27].

This work is an extension of the authors' previous work on α -channel edge super-resolution [28]. This paper represents a more detailed and comprehensive treatment of the topic, and introduces the local adaptive scheme under the framework of α -channel super-resolution, in order to enable a unified treatment of the entire image.

III. SOFT EDGE SMOOTHNESS PRIOR

In this section, we first review the Geocuts method which is designed for image segmentation tasks, and then present our work of the *SoftCuts* method for measuring the soft edge smoothness.

A. Geocuts

Our work is partially motivated by the Geocuts method [1], which is briefly summarized in this section.

The basic idea of Geocuts is to use a cut metric to approximate the Euclidean length of a curve. The cut metric is defined on the image grid graph, with edges connecting pixels in a large neighborhood system, and proper edge weights. Since such a cut metric has an analytic form, the smooth hard edges can be obtained by minimizing this metric.

More formally, the weighted grid-graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ is designed in the following way. \mathcal{V} is the set of all discretized image pixel positions. The edge set \mathcal{E} is defined according to the neighborhood system represented by a set of vectors $\{e_k \mid 1 \leq k \leq n_{\mathcal{G}}\}$, where $n_{\mathcal{G}}$ is the neighborhood order, and the e_k 's are chosen as the relative position (taking integer values as its components, and the unit is the grid interval) of the nearest $n_{\mathcal{G}}$ neighbors with different directions. They are ordered by their corresponding angle ϕ_k w.r.t. the $+x$ axis, such that $0 \leq \phi_1 < \phi_2 < \dots < \phi_{n_{\mathcal{G}}} < \pi$. For example, when $n_{\mathcal{G}} = 4$, we have $e_1 = (1, 0)$, $e_2 = (1, 1)$, $e_3 = (0, 1)$, and $e_4 = (-1, 1)$. The set of pixel pairs with relative position e_k is denoted by \mathcal{N}_k , assuming that $\mathcal{N} = \cup_{k=1}^{n_{\mathcal{G}}} \mathcal{N}_k$. All pixel pairs with relative position inside the set \mathcal{N} are connected by edges. Figure 2 shows some example neighborhood systems. After defining the edges, the edge weights corresponding to \mathcal{N}_k are defined as

$$w_k = \frac{\delta^2 \Delta \phi_k}{2|e_k|}, \quad (1)$$

where $\Delta \phi_k = \phi_{k+1} - \phi_k$ (set $\phi_{n_{\mathcal{G}}+1} = \pi$), and δ is the size of the grid interval.

Given the weighted grid-graph \mathcal{G} and a curve C in \mathbb{R}^2 overlayed on \mathcal{G} , we denote by \mathcal{E}_C the set of edges intersecting this curve. The *cut metric* of C is defined as

$$|C|_{\mathcal{G}} = \sum_{e \in \mathcal{E}_C} w_e, \quad (2)$$

where w_e is the weight for edge e . In another words, $|C|_{\mathcal{G}}$ is the summation of the weights of all edges intersecting C . Denoting by $|C|_{\mathcal{E}}$ the Euclidean length of curve C , the following theorem is derived in [1]:

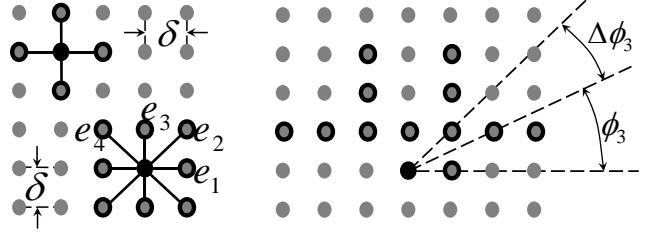


Fig. 2. Neighborhood system for $n_{\mathcal{G}} = 2, 4$ (left) and $n_{\mathcal{G}} = 12$ (right, only the neighbors on the upper plane are shown).

Theorem 1: If C is a continuously differentiable regular curve in \mathbb{R}^2 intersecting each straight line a finite number of times then

$$|C|_{\mathcal{G}} \rightarrow |C|_{\mathcal{E}}$$

as $\delta, \sup_k |\Delta \phi_k|$, and $\sup_k |e_k|$ go to zero [1].

Theorem 1 means that the length of a curve can be approximated by its cut metric on a weighted image grid graph. It is derived from the Cauchy-Crofton formula in Integral Geometry, which relates the Euclidean length of a curve with the expectation of the number of intersections with a random line. This expectation is further approximated by the cut metric defined in Eqn. 2 with the image grid graph and appropriately chosen edge weights. Roughly speaking, the approximation error converges to zero as the grid size goes to zero and a larger neighborhood is considered. This method can be generalized to 3D and arbitrary Riemannian metric. The global minimum of the cut metric can be obtained in a close-to-linear time by the Graph Cuts method [29], [30]. As its name suggests, Geocuts reveals the underlining relationship between two well-known segmentation algorithms, *i.e.*, Geodesic active contours and Graph Cuts. Geocuts also provides a principled way to choose the edge weights for using higher order neighborhoods.

By integrating the cut metric into an objective function, the hard edge smoothness prior can be added. Curves with smaller Euclidean length are preferred by minimizing such an objective function, thus smooth curves are obtained.

B. SoftCuts for smoothness measure of soft edges

Now, we present the proposed *SoftCuts* method as a generalization of the Geocuts method.

A cut metric can be defined on any set of disjoint closed curves C , or equivalently, a binary valued characteristic function χ^C on \mathbb{R}^2 which equals 1 inside C , and 0 otherwise. Geocuts is only applicable to a binary valued function χ^C on the image plane. To handle soft edges, which are gradual transitions in an intensity image, we first rewrite the definition of the cut metric in Eqn. 2 w.r.t. curve C (or equivalently, function χ^C) as follows

$$|C|_{\mathcal{G}} = |\chi^C|_{\mathcal{G}} \quad (3)$$

$$= \sum_{1 \leq k \leq n_{\mathcal{G}}} (w_k \sum_{(p,q) \in \mathcal{N}_k} |\chi_p^C - \chi_q^C|), \quad (4)$$

where p and q are pixel positions on the grid \mathcal{G} ; they are used as subscripts to indicate the function value at that pixel

for simplicity throughout this paper. We simplify the above equation as

$$|C|_{\mathcal{G}} = \sum_{(p,q) \in \mathcal{N}} w_{pq} |\chi_p^C - \chi_q^C|, \quad (5)$$

where w_{pq} is the weight for neighboring pixel pair (p, q) , which is determined by the relative position of p and q . In other words, for $(p, q) \in \mathcal{N}_k$, $w_{pq} = w_k$. Although Eqn. 5 is equivalent to Eqn. 2, the former is easier to be generalized to a real valued function S on \mathbb{R}^2 . We define the *soft cut metric* for S w.r.t. the grid-graph \mathcal{G} as follows

$$|S|_{\mathcal{G}} = \sum_{(p,q) \in \mathcal{N}} w_{pq} |S_p - S_q|. \quad (6)$$

Equation 6 takes the same form as Eqn. 5, with the only difference that continuous values are allowed for S in Eqn. 6, instead of binary value for χ^C in Eqn. 5.

By uniformly quantizing the function values with step $\frac{1}{n}$, S can be approximately by S^d , which takes values from $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$. The soft cut metric of S^d can be similarly defined by Eqn. 6, by replacing S with S^d . Moreover, S^d can be equivalently described by a set of level lines l_1, l_2, \dots, l_n , where l_i is the boundary between regions with S^d values that $S^d < \frac{i}{n}$ and $S^d \geq \frac{i}{n}$ in \mathbb{R}^2 . From Theorem 1, we know that the length of l_i can be approximated by its cut metric $|l_i|_{\mathcal{G}}$. Based on this, we have the following theorem (the proof is given in Appendix A).

Theorem 2: Assume that S is a continuous differentiable regular function on \mathbb{R}^2 , which ranges in $[0, 1]$, and S^d discretizes S with step $\frac{1}{n}$. Then the average length of all level lines of S^d w.r.t. $\frac{1}{n}$ can be approximated by the soft cut metric of S^d , i.e.,

$$|S^d|_{\mathcal{G}} \rightarrow \frac{1}{n} \sum_{1 \leq i \leq n} |l_i|_{\mathcal{E}} \quad (7)$$

under the same conditions of Theorem. 1

Theorem 2 generalizes Theorem 1 to soft edges from hard boundaries. It implies that by minimizing the soft cut metric in Eqn. 6, the sum of lengths of discrete level lines can be minimized. So adding this metric as a regularization term can help us obtain results with smaller length of image level lines. Thus the soft smoothness prior can be easily incorporated into the SR objective function.

There are several related works in the literature. The level-set method [16] and multiple-scale tensor voting [17] has been used to incorporate the edge smoothness prior. Image gradient on a large neighborhood is also used in [9] as a regularization term. Compared with these existing works, the benefit of the proposed *SoftCuts* method is that we have an explicit analytic term to characterize the edge smoothness, with a specific geometric explanation. When $n_G = 2$, Eqn. 6 becomes an approximation to the total variation (TV) regularization term [31], [32], [33], which is very powerful in edge-preserving image reconstruction.

IV. α CHANNEL IMAGE DECOMPOSITION

For a color image \mathbf{I} , matting is the technique that separates its foreground objects from the background. As the pixels on

the object boundary tend to be mixtures of both foreground and background color components, the separation needs to be done softly. A color image can be treated as a linear combination of the foreground and background images. Specifically, for each pixel p , we have

$$\mathbf{I}_p = \alpha_p \mathbf{F}_p + (1 - \alpha_p) \mathbf{B}_p, \quad (8)$$

where \mathbf{I}_p , \mathbf{F}_p , and \mathbf{B}_p are the pixel color vectors of the input, foreground, and background images, respectively, and $\alpha_p \in [0, 1]$ determines the weights of the convex combination of foreground and background at position p . Alpha matting tries to recover \mathbf{F} , \mathbf{B} , and α simultaneously for each pixel, given an input color image \mathbf{I} . An example is shown in Fig. 3.

For a color image edge, if we consider the two sides of this edge one as the foreground layer \mathbf{F} and the other as the background layer \mathbf{B} , then the region close to this edge can also be decomposed by alpha matting. Once having the decomposition, instead of using the original pixel colors, we use α to represent this edge, and refer to it as *α -channel edge description*.

The idea of using the α -channel is attractive due to its special properties as described next. First of all, the α -channel provides a normalization of the edge, that is, they transition from 0 to 1 between the two sides of the edge, instead of a large range of possible values in the color channels. More importantly, the α -channel actually preserves the edge information in a single channel as we will see in Sec. V-C. So it can be considered as a normalized version of the image edge. In addition, using the α -channel can explore the relationship among three color channels, and naturally combine all color information from three channels. In fact, color channels are closely related to each other. One example is the linear color model [34], which means that the pixels colors in a local image patch tend to form a line in the 3D color space. Such information is totally ignored by methods that process three color channels separately. The α -channel is extracted in a way such that all color information is taken into account simultaneously. Thus, the color channel relationship information is implicitly integrated.

The idea of using the α -channel has been successfully applied to image deblurring [35], [36], [37]. Matting is also used in [17] to extract sub-pixel location of the curve for super-resolution. A two color image prior, which in essence very similar to the matting decomposition, is used in [38] for demosaicing.

Given an input image \mathbf{I} , solving for α , \mathbf{F} , and \mathbf{B} simultaneously is obviously an under-determined problem. Image priors are needed to regularize the problem. One commonly used prior is the image smoothness prior, which assumes local color smoothness for both foreground and background. Figure 3 shows an example of a matting solution with this prior, where we can clearly see that \mathbf{F} and \mathbf{B} are locally smooth along the matting boundary, and α retains most of the edge information. This smoothness prior is extended to the local linear color model in [34]. Based on this assumption, a closed-form solution is derived. User interaction can be considered as another kind of prior, where some pure foreground / background pixels

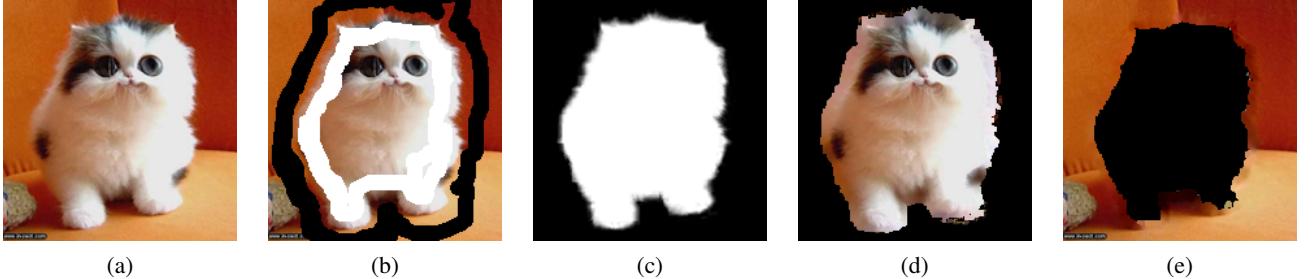


Fig. 3. An example of the alpha matting technique: (a) input color image \mathbf{I} , (b) human interaction specifying pure foreground and background pixels with black and white strokes, (c) extracted α channel (white represents 1, and black represents 0), (d) foreground patch \mathbf{F} , (e) background patch \mathbf{B} .

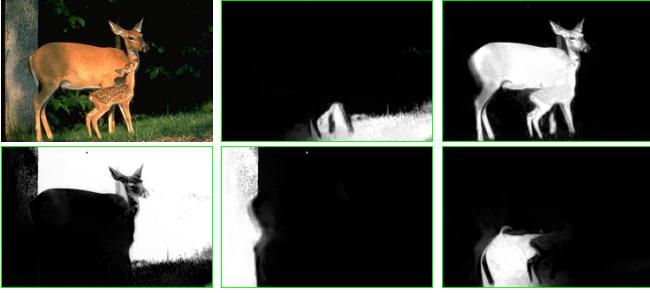


Fig. 4. Multiple layer image representation. The top left image is the input, the other five images represent the α value for the five layers.

are indicated manually by providing a trimap or using a brush tool (Fig. 3(b)) [39].

Traditionally, the alpha matting technique decomposes an input image into two components, *i.e.*, a foreground layer and a background layer. Very recently, the spectral matting technique [7] was proposed to generalize this concept by decomposing a color image into a linear combination of multiple image layers as follows

$$\mathbf{I} = \sum_{i=1}^n \alpha_i \mathbf{L}_i, \quad (9)$$

where \mathbf{L}_i is the i -th image layer, α_i is the corresponding combination weight, satisfying $\sum_i \alpha_i = 1$ for each pixel, and n is the total number of layers. This method not only enables multiple layer image representation, but also leads to a fully automatic and unsupervised solution for α_i by the spectral clustering technique. Thus, human interaction that specifies partial foreground or background pixels is not necessary anymore. An example is shown in Fig. 4. Each layer roughly corresponds to a homogeneous image region, and the edge information is captured by at least one of those α channels.

The definition of α -channel edge description can be easily extended to an α -channel image description as $\mathbf{a} = (\alpha_1, \alpha_2, \dots, \alpha_n)$, given the multiple-layered image decomposition. Similar to the α -channel edge description, image edges in the α -channel image description are also normalized in the range $[0, 1]$ over one single channel. Pixels along the border of two neighboring image layers have two non-zeros components in their α -channel image description, and these two values should sum up to 1 for each pixel. The α -channel edge description and the α -channel image description contain the same edge transition information for those pixels. Thus

all the properties of the α -channel edge description can be directly applied to the α -channel image description scheme in those regions. Although this property is not satisfied for pixels corresponding to more than two layers, such pixels only represent a very small percentage of the total number of pixels [7]. Moreover, the α -channel image description integrates the information of all image edges simultaneously, instead of one single edge for the α -channel edge description. This property makes the α -channel image description more suitable since it enables a unified image processing solution, as shown later in Sec. V-C.

V. SoftCuts FOR COLOR IMAGE SUPER-RESOLUTION

In this section, we first apply the *SoftCuts* method directly to single channel image super-resolution. Then we use the *SoftCuts* on α channel to super-resolve color image edges. Finally, the adaptive *SoftCuts* method is presented to process the entire color image in a unified framework.

A. Single channel super-resolution by SoftCuts

Theoretically, the generation process of an LR image can be modeled by the combination of various blurs, such as atmospheric blur, motion blur, camera blur, out-of-focus blur, and down-sampling. By combining the various blurs into a single filter G for the entire image, the generation process of a single channel LR image can be formulated as follows

$$I^l = (I^h * G) \downarrow + n, \quad (10)$$

where I^h and I^l represent the HR and LR single channel images respectively, G the impulse response of the blurring filter, $*$ the convolution operator, \downarrow the down-sampling operator, and n the additive noise.

We propose the *SoftCuts* method by utilizing the soft cut metric for single channel SR as the regularization term. The objective function is now defined as follows

$$I^h = \arg \min_I \left(\|I^l - (I * G) \downarrow\|_2^2 + \lambda |I|_G \right), \quad (11)$$

where the first term on the right side represents the fidelity to the data, the second term represents our prior knowledge about the original image defined by Eqn. 6, and λ is the regularization parameter controlling the contribution of the two terms.

While the l_2 norm is used for the data fidelity term (stemming also from a Gaussian assumption for the noise term in a stochastic formulation of the problem), the l_1 norm is used for

the regularization term. Its geometric meaning is explained by Theorem 2. Besides, the l_1 norm does not severely penalize large local gradients in general, thus allowing for sharp edges in the restored image (the l_2 norm results in gradual intensity transition across edges).

The objective function is optimized by the steepest descent algorithm. Denoting by p_{re} and p_s the gradients of the first and second term in Eqn. 11, we obtain the following iteration

$$I^{t+1} = I^t - \beta(p_{re} + p_s), \quad (12)$$

where

$$p_{re} = ((I^t * G) \downarrow - I^l) \uparrow * G, \quad (13)$$

$$p_s = \lambda \sum_{1 \leq k \leq n_G} w_k \{ \text{sgn}(I^t - D_{e_k} I^t) - D_{e_k} \text{sgn}(I^t - D_{e_k} I^t) \}, \quad (14)$$

β is the descent step size and \uparrow denotes the up-sampling operator. p_{re} is similar to the updating function of the back-projection [4] method, except that the back-projection kernel is the same as the blur filter. D_{e_k} is the displacement operator, which translates the entire image by e_k (an integer valued vector defined in Sec. III-A), and sgn is the sign indication function. p_s is the derivative of the soft cut metric defined by Eqn. 6. In fact, each term in Eqn. 6 will produce a $(+w_k)$ or $(-w_k)$ change for the two corresponding pixels. This updating strategy is the same as in [9]. In our experiments, I^0 is equal to the bicubic interpolation result.

B. Color edge SR by SoftCuts

For natural color image SR, a naive solution is to apply the above single channel SR algorithm to all three color channels independently. However, this approach tends to fail for a couple of reasons. First, the *SoftCuts* method is sensitive to the value of λ . The selection of this parameter is related to the local edge properties. Some edge strength normalization mechanism is needed to make possible a unified treatment for all edges. Second, in order to determine the exact edge position, information from all three color channels is required. Decisions made on each channel separately might be erroneous and inconsistent.

To address the above issues for color image super-resolution, in this section, we propose to super-resolve color edges by super-resolving its α channel.

Assume that an LR image edge \mathbf{I}^l can be decomposed as $\mathbf{I}^l = \alpha^l \mathbf{F}^l + (1 - \alpha^l) \mathbf{B}^l$. Based on the local smoothness assumption of \mathbf{F}^l and \mathbf{B}^l mentioned before, it is easy to see that super-resolving the edge \mathbf{I}^l can be achieved by super-resolving its α channel α^l . Thus the image super-resolution can be performed by processing each image edge separately. To be more specific, for each edge, we first perform alpha matting to a nearby region \mathbf{I}^l , to obtain α^l , \mathbf{F}^l , and \mathbf{B}^l . Then their HR counterparts α^h , \mathbf{F}^h , and \mathbf{B}^h can be recovered accordingly from them. Recovering α^h from $\alpha^l = (\alpha^h * G) \downarrow$ is exactly the single channel SR problem discussed in Sec. V-A, while \mathbf{F}^h and \mathbf{B}^h can be interpolated using for example the bicubic method given their down-sampled versions due to the smoothness assumption.

Figure 5 illustrates the idea of α -channel SR for one image edge. The LR patch is decomposed into two image patches and an LR α channel. Fig. 5(e) shows the recovered HR α channel by the proposed *SoftCuts* method. Combining the images in Fig. 5(b)(c)(e) using Eqn. 8 will produce the sharp and smooth edge, shown in Fig. 5(f).

The *SoftCuts* edge SR method has the following benefits. First, with α -channel edge description, each edge is normalized to a unified scale. The same value of λ is applied to super-resolve the α -channel descriptions of all edges, thus the problem of parameter selection for the *SoftCuts* method can be avoided. The underlying rationale is that we want to apply equal weights to all level lines over the α channel of different edges. Besides, the alpha matting technique can extract an edge by combining color information from all three channels, thus more precise results can be obtained.

However, applying this edge SR method on an entire color image is not easy [28]. It heavily relies on successful extraction of image edges for the low resolution input image. In addition, to enable edge decomposition by alpha matting, hard constraints for foreground and background pixels need to be automatically specified, which is also quite challenging, and sometimes even impossible.

C. Adaptive SoftCuts for color image SR

In this section, based on the α -channel image description, an adaptive method is proposed to overcome the difficulties of applying the α -channel super-resolution strategy mentioned above. Based on that, the entire image can be super-resolved in a unified way by implicitly processing all image edges simultaneously.

We start by considering an image edge, and rewriting Eqn. 8 as follows

$$\alpha = \frac{1}{F^c - B^c} I^c - \frac{B^c}{F^c - B^c}, \quad (15)$$

where $c \in \{r, g, b\}$ is the index for the color channels, and the RGB color space is used in this work. The subscript p in Eqn. 8 is omitted for simplicity. For a single image edge, the assumption is made that both \mathbf{F} and \mathbf{B} are locally smooth (Please refer to Sec. IV for more details). Thus, Eqn. 15 shows that the α -channel edge description is a linear function of the original image intensity, with scaling factor $\frac{1}{F^c - B^c}$, and an approximately constant shift $\frac{B^c}{F^c - B^c}$. This also explains how the edge information is preserved in the α -channel, and why the α -channel is a normalized description of the image edge. $F^c - B^c$ in the denominator serves as the normalization factor. More importantly, from Eqn. 15, we have that the following equation holds locally:

$$|\alpha|_{\mathcal{G}} = \frac{1}{|F^c - B^c|} |I^c|_{\mathcal{G}}, \quad (16)$$

where $|\alpha|_{\mathcal{G}}$ and $|I^c|_{\mathcal{G}}$ are the *SoftCuts* terms defined by Eqn. 6 on the α -channel and the single color channel, and the constant term is removed due to the smoothness assumption. Eqn. 16 suggests a *locally adaptive* strategy for applying the *SoftCuts* regularization term to the entire image instead of a single edge. More specifically, for different edges, applying the

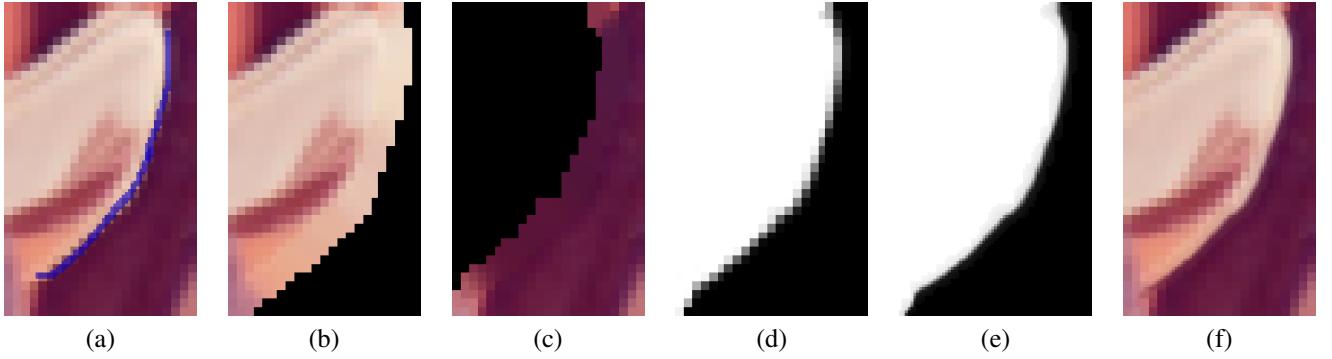


Fig. 5. An example of α -channel edge SR, (a) LR input and the edge needed to be processed, (b) (c) \mathbf{F} and \mathbf{B} on two sides of this edge segment, (d) LR α channel, (e) HR α channel using *SoftCuts*, (f) Smooth edge generated with matting equation (Eqn 8).

edge adaptation factor $\frac{1}{|F^c - B^c|}$ onto the color channel c can normalize the different edges to the α -channel, which always has a fixed contrast value of 1, instead of a large number of possible values in the original color channels. Thus by applying this adaptation scheme, the same weights are placed on the soft edge smoothness prior for different edges, *i.e.*, the smoothness measures of all the level lines on the α channel are equally weighted.

Now we consider the entire input image with its multiple layer decomposition defined in Eqn. 9. Based on the above discussion, for each pixel p with two non-zero components in its α -channel image description \mathbf{a} , the local adaptive factor on the color channel c is

$$\mu_p^c = \frac{1}{|L_{i,p}^c - L_{j,p}^c|}, \quad (17)$$

where i and j are the layer indices for those two non-zeros components. Since the alpha matting technique does not provide a solution for image layers L_i , alternatively, from Eqn. 16, we use the following equation as an estimator of the above adaptive factor

$$\mu_p^c = \frac{|\nabla \mathbf{a}_p|_\infty}{\sqrt{|\nabla I_p^c|^2 + \gamma}}, \quad (18)$$

where γ is a small positive number to avoid numerical problems, $|\cdot|_\infty$ is the l_∞ norm, which returns the largest absolute value of the α -channel gradient among all layers. Since most pixels do have one or two non-zero α components [7], the above normalization factor can successfully normalize most part of the image in the same way as in Eqn. 16, and also gives reasonable results for other parts, as shown in the experiments. Due to the smoothness assumption of image layers, μ^c should also be smooth. To suppress image noise, in practice, we convolve the weight map given by Eqn. 18 with a Gaussian kernel ($\sigma = 1.8$ in our experiments). Please notice that given an LR input image, the resulting alpha description is also in LR. The LR adaptive factor obtained by Eqn. 18 is then up-sampled with bicubic interpolation to get the HR adaptive factor. Here, bicubic interpolation does not introduce artifacts due to the smoothness property of the adaptive factor. Besides, since different sampling rates do not change the value of the RHS of Eqn. 18 at the same position, so we do not need to scale the value during up-sampling.

Input LR image \mathbf{I}^l and the scaling factor s .

Output HR image \mathbf{I}^h

- 1) Decompose \mathbf{I}^l by using the spectral matting algorithm [7] to get its α -channel image description.
- 2) Compute the LR adaptive factor based on Eqn. 18 for each pixel in each color channel.
- 3) Compute the HR adaptive factor by up-sampling with the bicubic interpolation and the scaling factor s .
- 4) For each color channel, optimize the adaptive *SoftCuts* objective function in Eqn. 20, to obtain the high resolution result \mathbf{I}^h .

Fig. 6. Adaptive *SoftCuts* algorithm for color image super-resolution.

Finally, the adaptive *SoftCuts* regularization term for color channel $c \in \{r, g, b\}$ is defined as follows

$$|I^c|'_G = \sum_{(p,q) \in \mathcal{N}} (w_{pq} \mu_{pq} |I_p^c - I_q^c|), \quad (19)$$

where $\mu_{pq} = \frac{1}{2}(\mu_p^c + \mu_q^c)$, and μ_p^c and μ_q^c are the local adaptive factors defined in Eqn. 18. The final objective function for adaptive *SoftCuts* on each channel c is

$$I^{h,c} = \arg \min_{I^c} (||I^{l,c} - (I^c * G) \downarrow ||_2^2 + \lambda |I^c|'_G). \quad (20)$$

Similar iterative method as in Sec. V-A is applied for optimization.

The entire algorithm is summarized in Fig. 6. Applying the adaptive *SoftCuts* regularization term to each color channel provides a unified solution by processing all image edges simultaneously. The reliance on the success of the edge extraction algorithms and hard constraint selection for matting are avoided. The benefits, however, of performing SR with the α -channel edge description is still materialized.

VI. EXPERIMENTS

In this section, we first present the SR results of applying *SoftCuts* on single channel to demonstrate its ability to generate smooth edges, and then show the results by using the adaptive *SoftCuts* method for single color image SR.

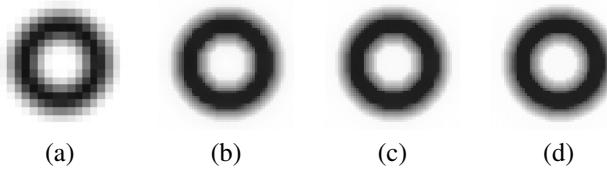


Fig. 7. (a) LR input image, (b), (c), (d) are the SR results ($\times 3$) with soft edge smoothness prior when $n_G = 2, 4, 12$ respectively ($\lambda = 0.01$).

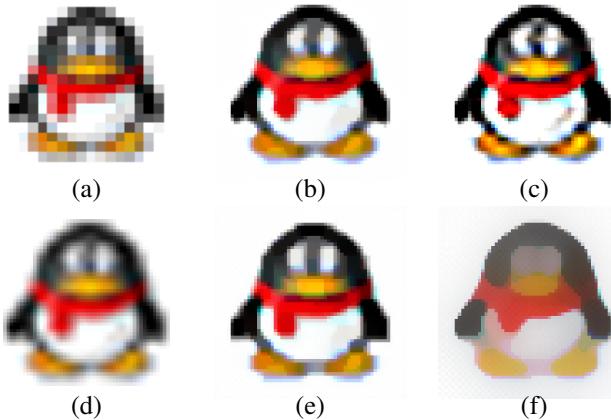


Fig. 8. Comparison of SR results by soft edge smoothness prior with different parameters ($\times 3$), (a) LR input image (20×20), (b) $\lambda = 0.01, n_G = 12$, (c) $\lambda = 0.001, n_G = 12$, (d) bicubic interpolation, (e) $\lambda = 0.01, n_G = 2$, (f) $\lambda = 0.1, n_G = 12$.

A. Single channel SR by SoftCuts

Figure 7 shows a proof-of-concept experiment, illustrating the necessity of using higher order neighborhood. Jaggy effects can be observed for small n_G , especially in (b) for a 4-neighborhood system. There are some 45° artifacts in (c), since 8-neighborhood system is used for it. The soft edge is much smoother in (d) with $n_G = 12$.

Figure 8 shows the result comparison of different parameter settings with an LR icon image (icon image SR is also studied in [18]). Larger n_G is applied in (b) than in (e), thus smoother edges are produced. In (c), a smaller λ is used than in (b), thus a smaller weight is placed on the smoothness prior. This makes the result look over-sharpened on high contrast edges, while a better result is obtained in other parts (such as the foot). In (f), a larger λ is used than in (b), the edge smoothness prior is over-weighted. All boundaries are very smooth, but the result is very blurry.

The effect of the parameters can be summarized as follows: (1) a larger n_G will produce smoother boundaries, and is also more computational expensive. In all of the later experiments, n_G is set equal to 12, with the neighborhood system shown in Fig. 2. (2) The value of λ is critical. As revealed by Eqn. 18, the desirable weight should consider both image gradient and α -channel gradient. Besides, the filter G in the generation model (Eqn. 10) also influences the quality of the result. However, estimating G is beyond the scope of this paper. We fix it as a Gaussian filter with $\sigma = 1.4$ throughout this paper.

Figure 9 shows an example of how each term in the objective function (in Eqn 11) changes during the iteration (Eqn. 12). In the first several steps (about 20 in this example),

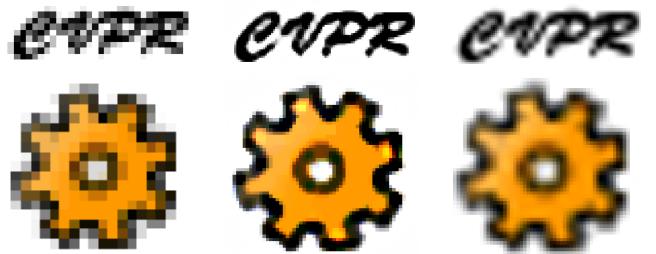


Fig. 10. SR results ($\times 3$) with soft edge smoothness prior; 1st column: LR inputs; 2nd column: SR results ($\lambda = 0.01, n_G = 12$); 3rd column: bicubic interpolation.

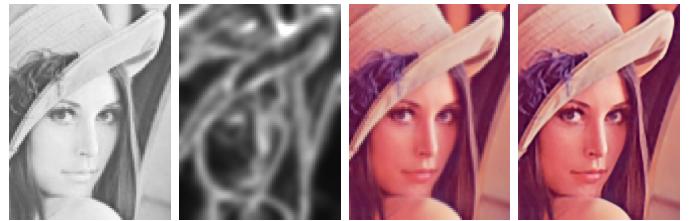


Fig. 11. Visualization of the adaptation weight. From left to right: one LR color channel, LR adaptation weights for this channel (large values of weights are represented by brighter intensity value), color LR input image, and result of the proposed adaptive *SoftCuts* method.

the change of the reconstruction error dominates, and the Soft cut metric may increase. After that, although the reconstruction error can hardly be further reduced, the result still looks jaggy as shown in (e). The Soft cut metric is further reduced by the iteration, and results in the final result with smooth and sharp edges in (h).

Additional results are shown in Fig. 10. The resulting edges are smooth and with sharp transitions, even when the quality of the LR input images is very low. However, for natural color images like the one in the third row of Fig. 10, although much smoother edges are obtained, the image looks unnatural. Some subtle edges are smoothed out. As discussed, the reason is that it is difficult, or even impossible, to find a single value of λ to address all edges in an image. That is why we need to work on the α channel to normalize different edges, and further develop the proposed adaptive *SoftCuts* method.

B. Adaptive SoftCuts for color image SR

Figure 11 displays the adaptation weights for one color channel. Heavier weights are placed on low contrast edges (e.g., those in the regions of the hat and nose), and smaller weights are placed on high contrast edges. Thus the influence of the smoothness prior on different edges is balanced. Please also notice that a zero weight is assigned to smooth regions, thus some subtle image fluctuation in those regions could still

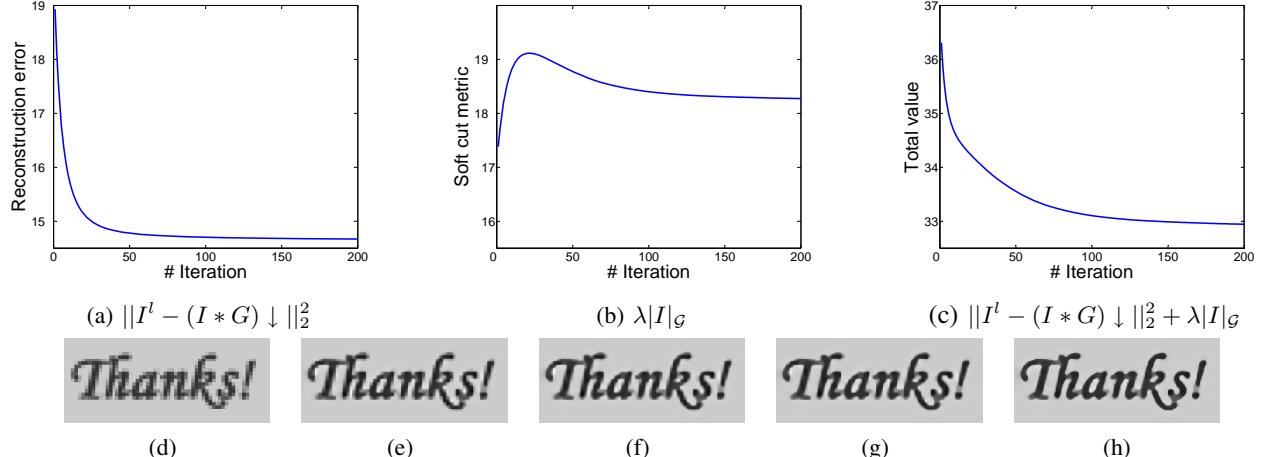


Fig. 9. An example of the iterative optimization process in Sec. V-A. The value of the reconstruction error, the Soft cut metric, and the entire objective function during the iteration process are plotted in (a), (b), and (c) respectively. The LR input image is shown in (d), and the result after 20, 40, 60, and 200 iterations are shown in (e), (f), (g), and (h) respectively.

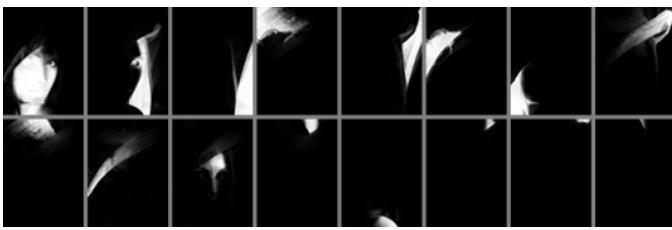


Fig. 12. Matting components extracted by the spectral matting algorithm. The original image is shown in Fig. 11.

be present. This makes the result look much more natural than the result with the non-adaptive *SoftCuts* method shown in Fig. 10. The intermediate result from the spectral matting decomposition is shown in Figure 12, where the image edge information is clearly extracted.

Figure 13 shows an example for comparison with other widely used algorithms, including bicubic interpolation, bicubic followed by unsharp masking (Photoshop), and the back-projection method in [4]. Figure 14 shows a zoomed in image patch. Blurry edges can be observed with the result of bicubic interpolation. The unsharp masking (a large radius of 10, is used to better improve the image contrast) method and the back-projection method can increase the image contrast, but the results still look jaggy. Both blurry and jaggy artifacts are successfully removed by the proposed adaptive *SoftCuts* method, even for very fine image structures.

Figure 15 compares the proposed method with an exemplar-based algorithm [21]. The existing exemplar-based methods can produce very sharp edges, but rely heavily on effective training data. Compared with it, smoother boundaries can be archived by our method, thus making the result look natural. Figure 16 shows a zoomed in patch to better illustrate the effectiveness of the proposed method. The image boundaries of our result are both smooth and with sharp transitions at the same time.

Figure 17 compares the proposed method with some existing commercial / free softwares designed specifically for

image resizing, including IrfanView [40] Lanczos Interpolator, VSO [41] Image Resizer, Genuine Fractals 5.0 [42], and PhotoZoom Pro 2 [43]. A close-up view is also shown. Image edges in our result looks smooth, sharp, and natural.

Additional results are shown in Fig. 18¹. Various images are tested, and visually appealing results are obtained by the adaptive *SoftCuts* algorithm. Please notice that the same set of parameters are applied for all of these images, except that the weight of the regularization term λ is set equal to 0.01 for natural images, and 0.02 for graphic images. The reason is that the graphic images usually have much smoother edges, thus a larger weight should be placed on the regularization.

Quantitatively, the RMS error (root-of-mean-square error) for bicubic interpolation, back-projection, and the proposed *SoftCuts* method is shown in Table I. The back-projection method [4] is effective in reducing the RMS error by enforcing the reconstruction constraint. The RMS error for our algorithm is roughly the same as the back-projection method, since a similar reconstruction constraint is enforced. However, the back-projection algorithm aims at minimizing only the data fidelity term, which is reflected in the RMS error. Therefore, in certain cases, the back-projection algorithm results in a smaller RMS error, although the visual quality of the results is not as good as the one provided by the proposed algorithm. Due to the additional soft edge smoothness prior, some improvement is observed for images with smooth and sharp edges, such as Zebra, Fonts, and Mickey. The ERMS (edge RMS) errors are also presented, since image edges are more important for visual perception. Similar observation can be made with the ERMS error.

For complexity, we run our experiments on a PIV3.4G PC with 1G RAM by Matlab implementation. For an LR input image of size 107×160 pixels, the spectral matting algorithm for image decomposition takes 120 seconds, and the adaptive *SoftCuts* method takes 35 seconds for 30 iterations.

¹More on <http://vision.eecs.northwestern.edu/research/IP/SR/index.html>.

TABLE I
ERROR COMPARISON FOR BICUBIC INTERPOLATION, BACK-PROJECTION [4], AND THE PROPOSED ALGORITHM (FOR EACH BOX WITH TWO NUMBERS, THE FIRST IS THE RMS ERROR, THE SECOND IS THE ERMS ERROR).

	Lena	Boy	Zebra	Monarch	Peppers	Fonts	Mickey
Bicubic	9.14 17.99	10.12 14.92	9.59 29.82	9.43 32.60	9.44 18.25	12.47 66.85	9.74 29.89
BP [4]	7.52 11.87	9.36 12.99	7.99 19.48	7.65 17.93	8.69 12.95	11.10 53.78	9.87 27.08
Ours	7.88 12.73	9.54 13.41	7.62 19.97	7.64 18.53	8.70 12.88	8.71 45.50	9.28 26.30

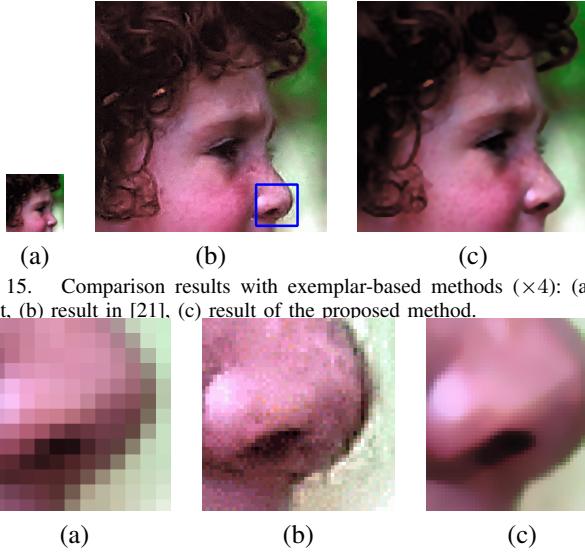


Fig. 15. Comparison results with exemplar-based methods ($\times 4$): (a) LR input, (b) result in [21], (c) result of the proposed method.

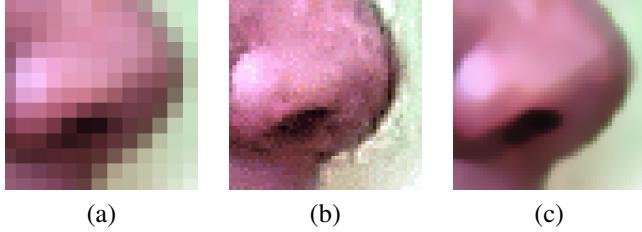


Fig. 16. A close-up view of the comparison results in Fig. 15: (a) LR input, (b) result in [21], (c) result of the proposed method. The selected image patch is highlighted in Fig. 15(b) by a blue box.

VII. CONCLUSION

In this paper, a novel single image super-resolution algorithm is proposed. A soft edge smoothness measure is defined on a large neighborhood system, which is an approximation of the average length of all level lines in the image. To extend this method to natural color image SR, an adaptive *SoftCuts* method is proposed based on a novel α -channel image description. It enables a unified treatment of edges with different contrasts on the α channel. Promising results for a large variety of images are obtained by this algorithm.

APPENDIX A PROOF OF THEOREM 2

From Eqn. 5, we have

$$|l_i|_{\mathcal{G}} = \sum_k (w_k \sum_{N_k} |\chi_p^{l_i} - \chi_q^{l_i}|),$$

thus

$$\begin{aligned} \sum_i |l_i|_{\mathcal{G}} &= \sum_i \sum_k (w_k \sum_{N_k} |\chi_p^{l_i} - \chi_q^{l_i}|) \\ &= \sum_k (w_k \sum_{N_k} \sum_i |\chi_p^{l_i} - \chi_q^{l_i}|), \end{aligned}$$

and

$$\begin{aligned} \sum_i |\chi_p^{l_i} - \chi_q^{l_i}| &= \#\{i \mid \chi_p^{l_i} \neq \chi_q^{l_i}\} \\ &= n \cdot |S_p^d - S_q^d|, \end{aligned}$$

resulting in

$$\begin{aligned} \sum_i |l_i|_{\mathcal{G}} &= \sum_k (w_k \sum_{N_k} n \cdot |S_p^d - S_q^d|) \\ &= n \cdot \sum_k (w_k \sum_{N_k} |S_p^d - S_q^d|) \\ &= n \cdot |S^d|_{\mathcal{G}} \quad (\text{from Eqn. 6}). \end{aligned}$$

Thus from Theorem 1, we have

$$|S^d|_{\mathcal{G}} = \frac{1}{n} \sum_i |l_i|_{\mathcal{G}} \rightarrow \frac{1}{n} \sum_i |l_i|_{\mathcal{E}}.$$

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation Grants IIS-0347877.

REFERENCES

- [1] Y. Boykov and V. Kolmogorov, “Computing geodesics and minimal surfaces via graph cuts,” in *ICCV*, 2003.
- [2] A. Katsaggelos, R. Molina, and J. Mateos, *Super resolution of images and video, Synthesis Lectures on Image, Video, and Multimedia Processing*. Morgan & Claypool, 2007.
- [3] S. Chaudhuri and J. Manjunath, *Motion-Free Super-Resolution*. Springer, 2005.
- [4] M. Irani and S. Peleg, “Motion analysis for image enhancement: resolution, occlusion and transparency,” *JVCIP*, 1993.
- [5] 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.
- [6] Z. Lin and H.-Y. Shum, “Fundamental limits of reconstruction based super-resolution algorithms under local translation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, 2004.
- [7] A. Levin, A. Rav-Acha, and D. Lischinski, “Spectral matting,” in *CVPR*, 2007.
- [8] M. Elad and A. Feuer, “Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images,” *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, 1997.
- [9] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super resolution,” *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, 2004.
- [10] M. F. Tappen, B. Russell, and W. T. Freeman, “Exploiting the sparse derivative prior for super-resolution and image demosaicing,” in *IEEE Workshop on Statistical and Computational Theories of Vision*, 2003.
- [11] S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Total variation super resolution using a variational approach,” in *ICIP*, 2008.
- [12] S. Farsiu, M. Elad, and P. Milanfar, “Multi-frame demosaicing and super-resolution of color images,” *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 141–159, 2006.
- [13] J. Allebach and P. W. Wong, “Edge-directed interpolation,” in *ICIP*, 1996.
- [14] X. Li and M. Orchard, “New edge-directed interpolation,” *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, 2001.
- [15] D. D. Muresan, “Fast edge directed polynomial interpolation,” in *ICIP*, 2005.
- [16] B. S. Morse and D. Schwartzwald, “Image magnification using level set reconstruction,” in *CVPR*, 2001.
- [17] Y.-W. Tai, W.-S. Tong, and C.-K. Tang, “Perceptually-inspired and edge-directed color image super-resolution,” in *CVPR*, 2006.
- [18] V. Rabaud and S. Belongie, “Big little icons,” in *CAVI*, 2005.

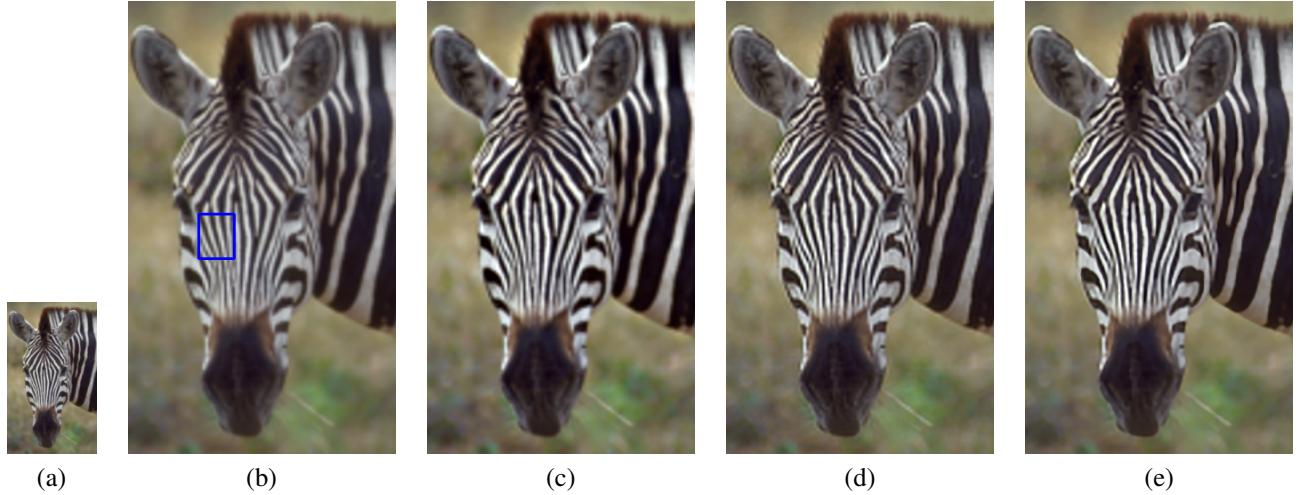


Fig. 13. Comparison results: (a) LR input, (b) bicubic, (c) bicubic followed by unsharp masking, (d) back-projection [4], (e) proposed method ($\times 3$).

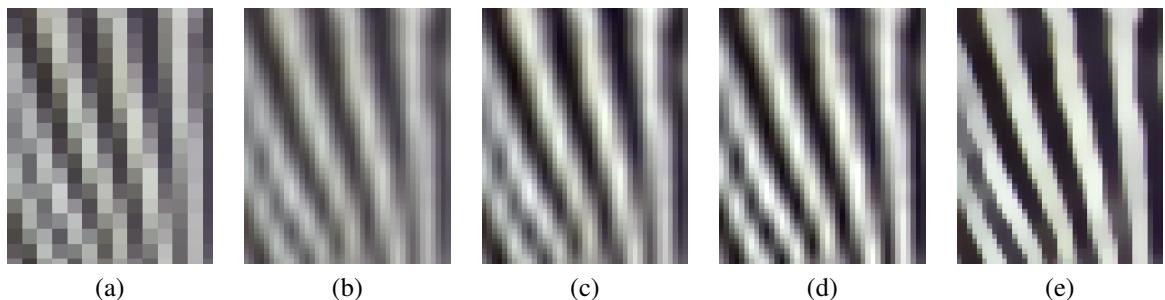


Fig. 14. A close-up view of the comparison results in Fig. 13: (a) LR input, (b) bicubic, (c) bicubic followed by unsharp masking, (d) back-projection [4], (e) proposed method. The selected image patch is highlighted in Fig. 13(b) by a blue box.



Fig. 17. Comparison with some existing commercial / free softwares designed specifically for image resizing.

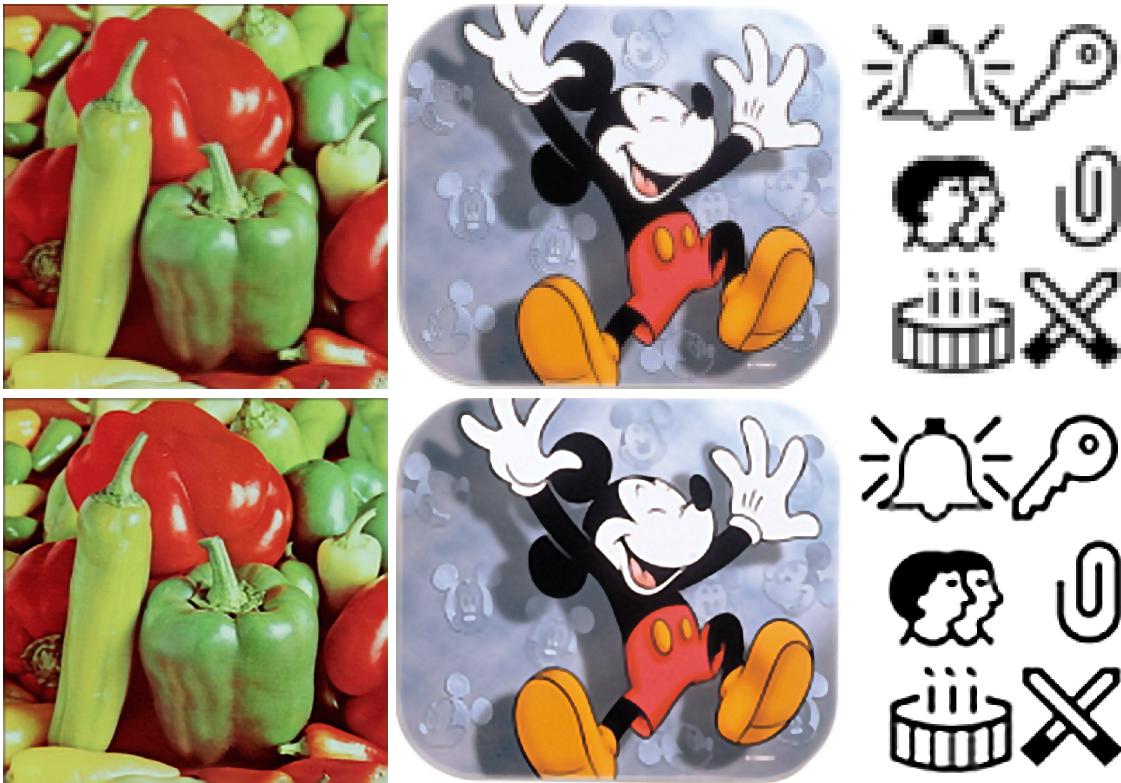
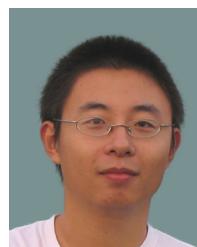


Fig. 18. Additional results with the proposed adaptive *SoftCuts* method. For each pair of images, the upper one is the LR input, and the lower one is our result ($\times 3$).

- [19] R. Nakagaki and A. K. Katsaggelos, "A vq-based blind image restoration algorithm," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1044 – 1053, 2003.
- [20] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, 2002.
- [21] W. T. Freeman, E. Pasztor, and O. Carmichael, "Learning low-level vision," *IJCV*, vol. 40, no. 1, pp. 25–47, 2000.
- [22] C. Liu, H.-Y. Shum, and C.-S. Zhang, "A two-step approach to hallucinating faces: Global parametric model and local nonparametric model," in *CVPR*, 2001.
- [23] C. M. Bishop, A. Blake, and B. Marthi, "Super-resolution enhancement of video," in *Proc. Artificial Intelligence and Statistics*, 2003.
- [24] D. Kong, M. Han, W. Xu, H. Tao, and Y. Gong, "Video super-resolution with scene-specific priors," in *BMVC*, 2006.
- [25] Q. Wang, X. Tang, and H. Shum, "Patch based blind image super resolution," in *CVPR*, 2005.
- [26] J. Sun, N. Zheng, H. Tao, and H. Shum, "Image hallucination with primal sketch priors," in *CVPR*, 2003.
- [27] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *CVPR*, 2004.
- [28] S. Dai, M. Han, W. Xu, Y. Wu, and Y. Gong, "Soft edge smoothness prior for alpha channel super resolution," in *CVPR*, 2007.
- [29] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [30] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [31] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.
- [32] T. F. Chan, S. Osher, and J. Shen, "The digital tv filter and nonlinear denoising," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 231–241, 2001.
- [33] S. Babacan, R. Molina, and A. Katsaggelos, "Parameter estimation in tv image restoration using variational distribution approximation," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 326 – 339, 2008.
- [34] A. Levin, D. Lischinski, and Y. Weiss, "A closed form solution to natural image matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007.
- [35] J. Jia, "Single image motion deblurring using transparency," in *CVPR*, 2007.
- [36] Q. Shan, W. Xiong, and J. Jia, "Rotational motion deblurring of a rigid object from a single image," in *ICCV*, 2007.
- [37] S. Dai and Y. Wu, "Motion from blur," in *CVPR*, 2008.
- [38] E. P. Bennett, M. Uyttendaele, C. L. Zitnick, R. Szeliski, and S. B. Kang, "Video and image bayesian demosaicing with a two color image prior," in *ECCV*, 2006.
- [39] J. Wang, M. Agrawala, and M. F. Cohen, "Soft scissors : An interactive tool for realtime high quality matting," *ACM Trans. on Graphics (Proc. SIGGRAPH)*, 2007.
- [40] "Irfanview," <http://www.irfanview.com/>.
- [41] "Vso image resizer," <http://www.vso-software.fr>.
- [42] "Genuine fractals 5.0," <http://www.ononesoftware.com>.
- [43] "Photozoom pro 2," <http://www.benvista.com>.



Shengyang Dai (S'05) is currently a Ph.D. candidate in the Electrical Engineering and Computer Science department of Northwestern University. His research interests include image/video processing, computer vision, and machine learning. He did summer interns at NEC Laboratories America (Cupertino, CA), Microsoft Research (Redmond, WA), and Google Research (Mountain View, CA) in 2006, 2007, and 2008, respectively. He received his B.S. and M.S. degrees from the Electrical Engineering department of Tsinghua Univresity, Beijing, China, in 2001 and 2004 respectively. He received the Outstanding Graduate Student Fellowship at Tsinghua University in 2004, and the Everly Fellowship at Northwestern University in 2008. He is a Student member of the IEEE.



Mei Han received her B.S. and Ph.D. in Computer Science from Tsinghua University, China, in 1992 and 1995, and Ph.D. in Robotics from Carnegie Mellon University in 2001. From 2001 to 2007, she was a research staff member at NEC Laboratories America. She has been a research scientist at Google since August 2007. Her research interests include computer vision, computer graphics, image and video analysis, video surveillance, multimedia data mining, and machine learning. She served as a committee member on ECCV 2002 workshop on

Vision and Modelling of Dynamic Scenes and a session chair on International Conference on Image Processing 2004.



Aggelos K. Katsaggelos (S'80-M'85-SM'92-F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively.

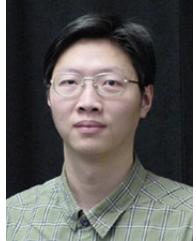
In 1985, he joined the Department of Electrical and Computer Engineering at Northwestern University, Evanston, IL, where he is currently a Professor. He was the holder of the Ameritech Chair of Information Technology (1997-2003). He is also the Director of the Motorola Center for Communications and a member of the Academic Affiliate Staff, Department of Medicine, Evanston Hospital.

He is the editor of Digital Image Restoration (Springer-Verlag, 1991), co-author of Rate-Distortion Based Video Compression (Kluwer, 1997), co-editor of Recovery Techniques for Image and Video Compression and Transmission (Kluwer, 1998), and co-author of Super-Resolution for Images and Video (Claypool, 2007) and Joint Source-Channel Video Transmission (Claypool, 2007), and the co-inventor of 12 patents.

Dr. Katsaggelos has served the IEEE and other Professional Societies in many capacities; he was, for example, Editor-in-Chief of the IEEE Signal Processing Magazine (1997-2002), a member of the Board of Governors of the IEEE Signal Processing Society (1999-2001), and a member of the Publication Board of the IEEE PROCEEDINGS (2003-2007). He is the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), an IEEE Signal Processing Society Best Paper Award (2001), an IEEE International Conference on Multimedia and Expo Paper Award (2006), and an IEEE International Conference on Image Processing Paper Award (2007). He is a Distinguished Lecturer of the IEEE Signal Processing Society (2007-2008).



Wei Xu received his B.S. degree from Tsinghua University, Beijing, China, in 1998, and M.S. degree from Carnegie Mellon University (CMU), Pittsburgh, USA, in 2000. From 1998 to 2001, he was a research assistant at the Language Technology Institute at CMU. In 2001, he joined NEC Laboratories America working on intelligent video analysis. His research interests include computer vision, image and video understanding, machine learning and data mining.



Ying Wu (SM'06) received the B.S. from Huazhong University of Science and Technology, Wuhan, China, in 1994, the M.S. from Tsinghua University, Beijing, China, in 1997, and the Ph.D. in electrical and computer engineering from the University of Illinois at Urbana-Champaign (UIUC), Urbana, Illinois, in 2001.

From 1997 to 2001, he was a research assistant at the Beckman Institute for Advanced Science and Technology at UIUC. During summer 1999 and 2000, he was a research intern with Microsoft Research, Redmond, Washington. In 2001, he joined the Department of Electrical and Computer Engineering at Northwestern University, Evanston, Illinois, as an assistant professor. He is currently an associate professor of Electrical Engineering and Computer Science at Northwestern University. His current research interests include computer vision, image and video analysis, pattern recognition, machine learning, multimedia data mining, and human-computer interaction. He serves as associate editors for IEEE Transactions on Image Processing, SPIE Journal of Electronic Imaging, and IAPR Journal of Machine Vision and Applications. He received the Robert T. Chien Award at UIUC in 2001, and the NSF CAREER award in 2003. He is a senior member of the IEEE.



Yihong Gong received his B.S., M.S., and Ph.D. degrees in Electronic Engineering from University of Tokyo in 1987, 1989, and 1992, respectively. He then joined the Nanyang Technological University (NTU) of Singapore, where he worked as an assistant professor in the School of Electrical and Electronic Engineering. Between June 1996 and December 1998, he worked for Robotics Institute, Carnegie Mellon University as a project scientist, and was a principal member of both the Informedia Digital Video Library project and the Experience-

On-Demand project funded by NSF, DARPA and other governmental agencies. In 1999, he joined NEC Laboratories America, where he built the multimedia content analysis team from scratch. In 2006, he became the head of the silicon valley branch of the labs. His research interests include computer vision, multimedia content analysis, and machine learning.