Soft Edge Smoothness Prior for Alpha Channel Super Resolution

Shengyang Dai[†], Mei Han[‡], Wei Xu[‡], Ying Wu[†], and Yihong Gong[‡]

[†]EECS Department, Northwestern University, Evanston, IL 60208
[‡]NEC Laboratories America, Inc., Cupertino, CA 95014

Abstract

Effective image prior is necessary for image super resolution, due to its severely under-determined nature. Although the edge smoothness prior can be effective, it is generally difficult to have analytical forms to evaluate the edge smoothness, especially for soft edges that exhibit gradual intensity transitions. This paper finds the connection between the soft edge smoothness and a soft cut metric on an image grid by generalizing the Geocuts method [5], and proves that the soft edge smoothness measure approximates the average length of all level lines in an intensity image. This new finding not only leads to an analytical characterization of the soft edge smoothness prior, but also gives an intuitive geometric explanation. Regularizing the super resolution problem by this new form of prior can simultaneously minimize the length of all level lines, and thus resulting in visually appealing results. 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 normalizing edge segments with their alpha channel description, to achieve a unified treatment of edges with different contrast and scale.

1. Introduction

The objective of image super resolution (SR) [18] is to obtain high quality images from low resolution inputs. It is widely applicable in video communication, object recognition, HDTV, image compression, *et al*. There are many cases that only one low resolution image is available. In this paper, we mainly focus on super resolution (or image hallucination) from one single low resolution input image.

Generally speaking, low resolution images are generated by smoothing and down-sampling target scenes with low-quality image sensors. The task of recovering the original high resolution (HR) image from a single low resolution (LR) input is an inverse problem of this generation procedure. One criterion of solving this inverse problem is to minimize the reconstruction error. In other words, the result which can produce the same low resolution image as the input one is preferred. Back-projection [17] is proposed to minimize the reconstrution error efficiently by an iterative algorithm. However, since a lot of information is lost in the generation process, this problem is severely under-determined. There might be multiple solutions to minimize this error, even for multiple LR input images [2, 22]. To overcome this difficulty, image priors need to be incorporated for regularizing the inverse problem.

One of the most widely used priors is the *edge smoothness prior* that prefers a HR image with smooth edges. This is reasonable, because the human perception seems to favor this choice. Given the severely under-determined nature of super resolution, such a prior is especially important for getting rid of chessboard effect at region boundaries, which is a common drawback of simple interpolation-based method, like bilinear or bicubic. However, in practice, there are two main difficulties in incorporating this prior:

1. An edge in an intensity image is much more complex than a single geometric curve. In reality, an image edge exhibits a gradual intensity transition. We call it a *soft edge* in our paper (see Fig. 1(b) for an example). Working on soft edges is more meaningful for super resolution because they are more realistic in practice. But it is difficult to quantify such a prior and evaluate the smoothness of soft edges.

2. Natural color images show a large variety of edges with different contrast and scale. Besides, the edges are determined simultaneously by information from all three color channels. How to explore the 3D color information and treat those edges uniformly is important for color image SR.

This paper mainly addresses the two issues mentioned above. The main contributions are as follows:

1. For an ideal curve, or a *hard edge*, Geocuts [5] method can approximate its Euclidean length with a *cut metric* on image grid. We are not only simply introducing this to super resolution, but more importantly, we extend this idea to intensity images by defining a *soft edge cut metric* to measure the smoothness of soft edges. We prove that this metric can approximate the average length of all level lines in the

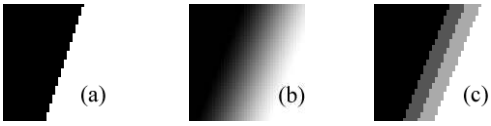


Figure 1. Examples of (a) hard edge, (b) soft edge, and (c) level lines (boundaries between different intensities) for (b) with quantization step 64 on image intensity.

intensity image, where level line means boundary between pixels with intensity smaller and larger than a given value (an example is shown in Fig. 1(c)). This extension is significant because it leads to a new analytical form for the soft edge smoothness prior. The new smoothness measure can be used to regularize the objective function of the SR task, and results in impressive results.

2. To apply the soft edge smoothness prior on natural color images, a novel approach is proposed based on alpha matting technique. We show that the problem of color image super resolution can be transformed to a combination of alpha matting and alpha channel super resolution. A closed form alpha matting solution [20] can help to describe each edge segment in a unified way by alpha channel.

The benefits of our SR algorithm are three-fold: (1) By integrating the soft edge smoothness measure into an objective function, the length of all image level lines can be minimized simultaneously. Thus result with smooth edges can be obtained. The resulting edges are also sharp due to the edge-preserving property of the proposed prior term. (2) Both image likelihood and image prior terms are integrated together in a single objective function, which can be optimized efficiently. (3) Alpha channel SR utilizes color information from all three channels simultaneously, and the edge description by an alpha channel provides a unified treatment of edges with different contrast and scale.

The related work is briefly summarized in Sec. 2. The main theoretical results are presented in Sec. 3, and its application on alpha channel color image SR is described in Sec. 4. Experiments are shown in Sec. 5 and Sec. 6 concludes this paper.

2. Related work

Various image priors have been considered in the literature of super resolution. Two of the most extensively studied image modeling priors are image smoothness prior and edge smoothness prior. Neighboring pixels are likely to have the same color. Various filtering/interpolation algorithms (*e.g.*, bilinear and bicubic interpolation) tend to produce smooth HR images. Another way is trying to minimize the image derivatives [11, 13]. For one dimensional case, a linear closed form solution is derived in [8]. However, the image smoothness prior is not valid at region boundaries. To preserve edge sharpness, edge directed interpolation [1, 21] is proposed to fit smooth sub-pixel edges to the image and use these edges to prevent cross-edge interpolation. Locating high precision edge position is also necessary for removing the chessboard effect, which is another common problem for interpolation-based method. The edge smoothness prior is usually used to handle this problem. Smooth curves are preferred, which is consistent with human perception. Level-set method [24] can reconstruct smooth approximation of all of the image level-set contours simultaneously to refine the edges. To avoid over-smoothness, hard constraints are introduced to model the image likelihood. The HR curves can also be inferred by multi-scale tensor voting method [28], and the HR image is recovered accordingly by a modified back-projection iteration. All three color channels are considered together. Snake-based vectorization is used in [25] to achieve smooth boundaries for icon image SR. There are also some other image modeling priors for the SR taks, such as the two color image prior [3] and the sparse derivative prior [29].

Instead of image prior modeling, many researchers use image exemplars directly. Candidates for each position are selected based on the middle frequency information. Spatial consistency is enforced by pairwise interaction between neighboring positions under a Markov Random Fields framework [14, 15, 23]. The final discrete optimization problem is usually solved by belief propagation. This method is extended to video SR in [4]. The domain-specific case is discussed in [19]. Two key issues usually need to be addressed for exemplar-based method: one is how to find HR candidate patches efficiently, Locality Sensitive Hashing [30] and KD-tree [19] are applied to speed up the searching; the other is how to solve the optimization problem efficiently, image primal sketch [27] method can simplify the problem to a chain structure. Learning algorithms, such as locally linear embedding [10], can also be used to infer the high frequency information.

3. Soft edge smoothness prior

3.1. Geocuts

Our work is motivated by [5]. To make this paper self-contained, we briefly summary the basic idea of [5].

Given a weighted grid-graph $\mathcal{G} = \langle V, E \rangle$, and a curve C in $\mathbb{R}^2$, assume E_C is the set of edges intersecting with this curve. The *cut metric* of C is defined as

$$|C|_{\mathcal{G}} = \sum_{e \in E_C} w_e, \tag{1}$$

where w_e is the edge weight. $|C|_{\mathcal{G}}$ is the weight summation of the edges intersecting with C.

Define the neighborhood system of a regular grid $\mathcal{G}$ as a set of vectors $\mathcal{N}_{\mathcal{G}} = \{e_k \mid 1 \leq k \leq n_{\mathcal{G}}\}$, where e_k are ordered by their corresponding angle ϕ_k w.r.t. the $+x$ axis, such that $0 \leq \phi_1 < \phi_2 < ... < \phi_{n_{\mathcal{G}}} < \pi$. Besides, e_k is chosen as the k-th nearest neighbor group in $\mathcal{G}$. Some examples are shown in Fig. 2.

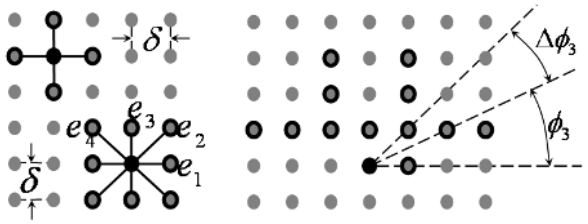


Figure 2. Neighborhood system for $n_{\mathcal{G}} = 2, 4$ (left) and $n_{\mathcal{G}} = 12$ (right, only the neighbors on the upper plane are shown).

Assume $|C|_{\mathcal{E}}$ is the Euclidean length of curve C, $\Delta\phi_k = \phi_{k+1} - \phi_k$ (set $\phi_{n_{\mathcal{G}}+1} = \pi$), then by setting

$$w_k = \frac{\delta^2 \Delta\phi_k}{2|e_k|}, \tag{2}$$

The follow theorem is derived in [5]:

Theorem 1 *[5] If C is a continuously differentiable regular curve in $\mathbb{R}^2$ intersecting each straight line a finite number of times then*

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

as δ, $\sup_k |\Delta\phi_k|$, and $\sup_k |e_k|$ get to zero.

In other words, the length of a curve can be approximated by its cut metric. This method can be generalized to 3D, and arbitrary Riemannian metric. The global minimum can be found in a close linear time by the Graph Cuts [6, 7] method. As its name suggested, Geocuts reveals the underlining relationship between two well-known segmentation algorithms, *i.e.*, Geodesic active contours and Graph Cuts. Geocuts also provides a principled solution to choose the edge weights for using higher order neighborhood.

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

3.2. Smoothness measure for soft edges

Now, we present our generalization of Geocuts method.

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

$$|C|_{\mathcal{G}} = |F_C|_{\mathcal{G}} = \sum_{1 \le k \le n_{\mathcal{G}}} \big(w_k \sum_{e_{pq} \in N_k} |F_C(p) - F_C(q)| \big), \tag{3}$$

where N_k contains all node pairs in the k-th group of neighborhood. Although Eqn. 3 is equivalent to Eqn. 1, it 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. grid-graph $\mathcal{G}$ as follows

$$|S|_{\mathcal{G}} = \sum_{1 \le k \le n_{\mathcal{G}}} \big(w_k \sum_{e_{pq} \in N_k} |S(p) - S(q)| \big). \tag{4}$$

By uniformly quantizing the function values with step $\frac{1}{n}$, S can be approximately by S_d, which takes values from $\{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$. The soft cut metric of S_d can be similarly defined by Eqn. 4, by replacing S with S_d. Moreover, S_d can be equivalently described by a set of level lines $l_1, l_2, ..., l_n$, where l_i is the boundary between points with S_d values $<$ and $\ge \frac{i}{n}$, 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, the following theorem can be proved.

Theorem 2 *Assume S is a continuous differentiable regular function on $\mathbb{R}^2$, which ranges in $[0, 1]$, and S_d discretize 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, or*

$$|S_d|_{\mathcal{G}} \to \frac{1}{n} \sum_{1 \le i \le n} |l_i|_{\mathcal{E}} \tag{5}$$

under the same conditions of Theorem. 1

The proof is presented in Appendix. A. Theorem. 2 generalizes Theorem. 1 to be applicable to soft edges instead of hard boundaries. It implies that by minimizing the soft cut metric in Eqn. 4, the length summation of discrete level lines can be minimized. So adding this metric as a regularity term can help to obtain result with smaller length of image level lines, thus the soft smoothness prior can be integrated.

3.3. Application on super resolution

In theory, the generation process of LR image can be modeled by a combination of atmosphere blur, motion blur, camera blur, and down-sampling. By simplify the first 3 factors with a single filter G for the entire image, the generation process can be formulated as follows

$$I^l = (I^h * G) \downarrow, \tag{6}$$

where I^h and I^l are the HR and LR images respectively, G is a spatial filter, $*$ is the convolution operator, and $\downarrow$ is the down-sampling operator. Noise is not considered here.

The soft cut metric is directly applicable to the problem of SR, by defining the objective function as

$$I^h = \arg\min_I \big(d(I^l, I) + \lambda |I|_{\mathcal{G}} \big), \tag{7}$$

where $d(I^l, I) = ||I^l - (I * G) \downarrow||_2^2$ is the likelihood term. It is the square of L_2 distance between the given LR image I^l and synthesized LR image by I. $|I|_{\mathcal{G}}$ is the smoothness prior term for soft edge defined by Eqn. 4. λ is a parameter to balance these two terms.

The reasons that different norms are used for likelihood and prior terms are as follows: (1) The L_2 norm is used for

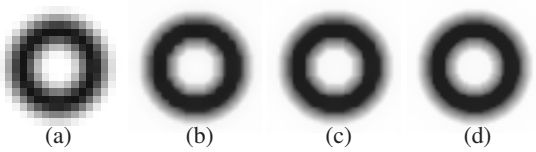


Figure 3. (a) LR input image, (b), (c), (d) are the SR results ($\times 3$) with soft edge smoothness prior when $n_{\mathcal{G}} = 2, 4, 12$ respectively ($\lambda = 0.01$).

likelihood term since it punishes more on large reconstruction error than L_1. (2) Minimizing the L_1 norm of gradient is edge-preserving, it does not severely penalize steep local gradients, while minimizing the L_2 norm of gradient usually leads to a graduate transition across edges, and the geometrical explanation in Theorem 2 does not hold for it.

The objective function is optimized by steepest descent algorithm. By putting the same group of neighborhood together, it can be implemented efficiently as follows:

$$I^{t+1} = I^t - \beta(p_{re} + p_s), \tag{8}$$

where

$$p_{re} = \big((I^t * G) \downarrow - I^l\big) \uparrow * G, \tag{9}$$

$$p_s = \lambda \sum_k w_k \text{sgn}\big(I^t - I^t D_{e_k}\big)(1 - D_{e_k}), \tag{10}$$

β is the descent step, p_{re} is similar to the update function of back-projection [17], except the back-projection kernel is chosen the same as the blur filter. $\uparrow$ is the up-sampling operator. D_{e_k} is the displacement operator, which translates the entire image by e_k, and sgn is the sign indication function. p_s is the derivative of the soft edge smoothness measure defined by Eqn. 4. In fact, each term in Eqn. 4 will produce a $+w_k$ or $-w_k$ change for the two corresponding pixels. In Eqn. 10, the sgn function determines the $+/-$, and $1 - D_{e_k}$ can apply the changes to the two corresponding pixels. This updating strategy is the same as in [13]. I^0 is set to the bicubic interpolation result in our experiments.

For color images, in this section, we simply apply our method on three color channels separately. Our novel treatment of color images will be presented in Sec. 4.

3.4. Results and discussions

Figure 3 illustrates the necessity of using higher order neighborhood. Metrication effect can be observed for small $n_{\mathcal{G}}$. There are some 45^o artifacts in Fig. 3(c), since 8-neighborhood system is used for it.

Figure 4 shows the result comparison of different parameter settings with a LR icon image (icon image SR is also studied in [25]). Larger $n_{\mathcal{G}}$ is applied in (b) than in (e),thus more smooth edges are produced. In (c), smaller λ is used than in (b), thus less weight is put on the smoothness prior, this makes the result look over-sharpened on high contrast edges, while better result is archived at low contrast part (such as the foot). In (f), larger λ is used than in (b), the edge smoothness prior is over stressed, all boundaries are

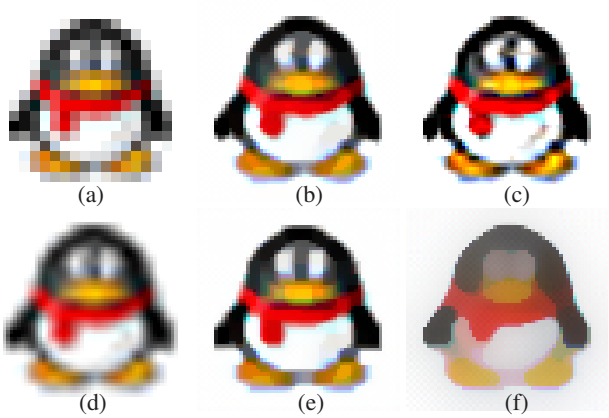


Figure 4. Result comparison of SR by soft edge smoothness prior with different parameters ($\times 3$), (a) LR input image (20×20), (b) $\lambda = 0.01, n_{\mathcal{G}} = 12$, (c) $\lambda = 0.001, n_{\mathcal{G}} = 12$, (d) bicubic interpolation, (e) $\lambda = 0.01, n_{\mathcal{G}} = 2$, (f) $\lambda = 0.1, n_{\mathcal{G}} = 12$.

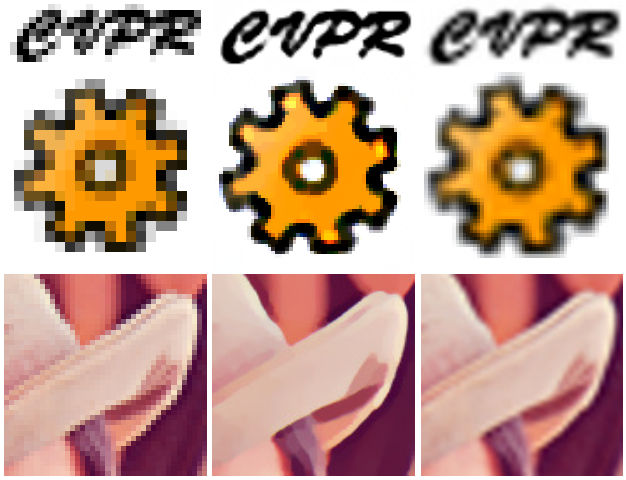


Figure 5. More SR results ($\times 3$) with soft edge smoothness prior, 1^{st} column: LR inputs, 2^{nd} column: SR results ($\lambda = 0.01, n_{\mathcal{G}} = 12$), 3^{rd} column: bicubic interpolation.

very smooth, but the result is blurry. Generally speaking, the effect of the parameters can be summarized as follows: (1) larger $n_{\mathcal{G}}$ will produce smoother boundary, but more computational demanding. In all of the later experiments, $n_{\mathcal{G}}$ is set to 12, the corresponding neighborhood system is shown in the right part of Fig. 2. (2) The value of λ is critical, small λ is suitable for low contrast edges, while large λ is suitable for high contrast edges. In fact, the filter G in the generation model (Eqn. 6) also influences the result. However, estimating G is out of the scope of this paper. We fix it as a Gaussian filter with $\sigma = 2$ throughout this paper. More results are shown in Fig. 5, good results can be achieved by the proposed algorithm even when the quality of the LR input images is very low.

There is some related work in the literature. Level-set method is used in [24] to incorporate edge smoothness prior. To avoid over-smoothness, hard constraints are enforced as

image likelihood terms . In [28], the smoothness prior is integrated by multiple-scale tensor voting, where the edge tokens can interact with each other to get smooth curves. Image gradient on a large neighborhood is also used in [13] as a regularity term. Comparing with these existing works, the benefit of our algorithm is that we have an explicit objective function which integrates both the prior and the likelihood terms, and there is an exact geometric explanation for it. When $n_{\mathcal{G}} = 2$, Eqn. 4 becomes an approximation of the total variation (TV) regularity term [9, 26], which is very powerful in edge-preserving image reconstruction.

4. Color image super resolution

For natural color image SR, three reasons limit the performance of applying soft edge smoothness prior directly by simply processing each color channels separately on the entire image.

1. Exact edge position is determined by the color information from all three channels. Decisions made on each channel separately might be wrong and inconsistent.

2. SR by soft edge smoothness prior is sensitive to the value of λ, which is related to the local contrast. Take the 3^{rd} image in Fig. 5 as an example, some subtle edges are smoothed out with this set of parameters, while in fact, they can be perfectly extracted by smaller λ in our experiments. Some edge strength normalization mechanism is needed to make possible a unified treatment for all edges.

3. Enforcing soft edge smoothness prior on regions near corner points will produce undesired smoothed curves, which is also observable in Fig. 5.

All these problems motivate our natural color image SR approach as follows.

4.1. The proposed approach

The entire system is illustrated in Fig. 6. The standard canny edge detection algorithm is used to extract continues edges. A robust corner detection algorithm based on curvature scale space [16] is applied. These corner points can break the edges into segments. Each edge segment c_i is a continuous curve (may be closed), and a nearby patch P_i is assigned to it by morphological operations.

We process each edge segment at P_i separately. For each segment, if we consider the two sides of this edge as foreground and background, the problem can be reduced to the alpha matting problem. Thus the true colors for two sides of the edge can be recovered by a recently proposed closed form solution [20]. The LR input is a blending of these two patches through an alpha channel, which ranges in $[0, 1]$. All the alpha matting parts are processed on low resolution. After that, SR based on soft edge smoothness prior (Sec. 3.3) is used to generate the HR alpha channel given the LR alpha channel extracted by alpha matting. The HR alpha channel is combined with the LR patches of two sides of the edge to generate the HR edge. At the end, Eqn. 8 is applied on the entire image to reinforce the reconstruction constraint for region without salient edge segment. A small λ ($= 0.002$ in our experiments) is used with a fixed number of iterations (15 in our experiments). Back-projection is also used in [27] as a post-processing step.

Input LR image I^l and scale factor s.
Output HR image I^h

1. Edge segment extraction and region assignment to get $\{c_i\}$ and $\{P_i\}$.
2. For each segment c_i, process P_i as follows
   (a) Compute $I^l_{L,i}$, $I^l_{R,i}$, and α^l_i from I^l by a closed form alpha matting solution.
   (b) Alpha channel SR to get α^h_i from α^l_i by single channel SR with soft edge smoothness prior.
   (c) Synthesis the HR patch by $I^l_{L,i}$, $I^l_{R,i}$ and α^h_i.
3. Reinforce the reconstruction constraint for the entire image by Eqn. 8 with small λ.

Figure 6. The overview of alpha channel color image SR approach.

The proposed approach has the following benefits: (1) Alpha matting technique can extract the edge by combining color information from all three channels, thus more precise result can be obtained. (2) In the meanwhile, describing each edge segment by the alpha channel can normalizes it into a unified scale, the problem of parameter selection for soft edge smoothness prior can be avoided. (3) The corner point detection algorithm can help to avoid oversmoothness at corner points. So all of the three problems presented in the beginning of this section can be addressed.

4.2. Edge decomposition by alpha matting

Alpha matting is a technique to decompose an image into a linear combination of foreground image and background image through an alpha channel. It is an important problem in computer graphics to extract the foreground object for image editing. Ideally, the influence of the neighboring background color should be removed. Assume the foreground and background images are F and B, then the following equation should hold for each pixel p,

$$I_p = \alpha_p F_p + (1 - \alpha_p) B_p, \tag{11}$$

where $\alpha_p \in [0, 1]$ is the foreground opacity of pixel p. Given the blended image I, solving for F, B, and α is also an under-determined inverse problem. In [20], by assuming that both F and B satisfy a locally linear model approximately, a closed form solution is derived. Hard constraint can be easily enforced into the cost function.

Similar to the idea of alpha matting, an HR step edge can also be considered as a combination of two smooth patches through a weight channel α as follows

$$I^h = \alpha^h I^h_L + (1 - \alpha^h) I^h_R, \tag{12}$$

where I^h_L and I^h_R represent the actual image colors for two sides of the edge at HR. Then by Eqn. 6, the corresponding

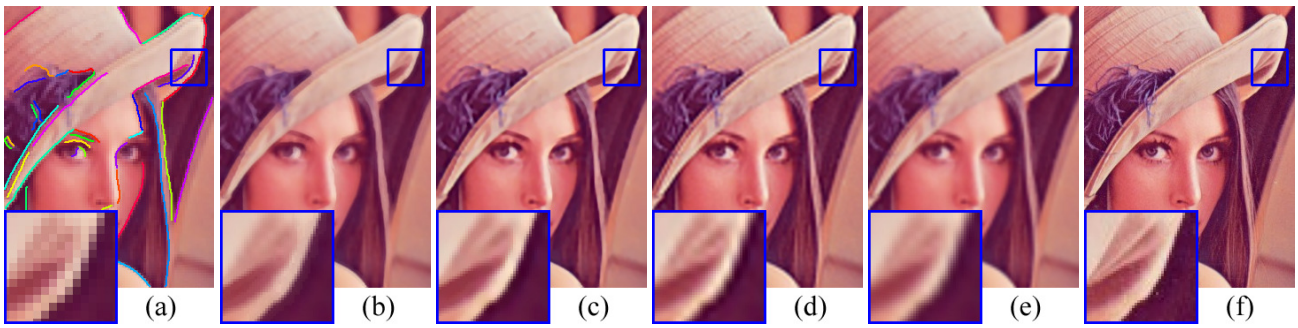


Figure 7. An example of the process of alpha channel super resolution and result comparison($\times 3$), (a) LR input & extracted edge segments, (b) result after alpha channel SR (Fig. 6, step 2), (c) final result (Fig. 6, step 3), (d) back-projection [17], (e) bicubic, and (f) ground truth.

LR image can be expressed as follows

$$I^l = \left(\alpha^h I_L^h + (1-\alpha^h) I_R^h\right) * G \downarrow \quad (13)$$

$$\simeq (\alpha^h * G) \downarrow I_L^h \downarrow + (1 - (\alpha^h * G) \downarrow) I_R^h \downarrow . \quad (14)$$

The approximate equality can be taken if we assume that both I_L^h and I_R^h are locally smooth, which is reasonable for the SR task. By assuming $\alpha = (\alpha^h * G) \downarrow$, $F = I_L^h \downarrow$, and $B = I_R^h \downarrow$, Eqn. 14 is exactly the same as Eqn. 11. It means that we can do alpha matting for I^l, to get $(\alpha^h * G) \downarrow$, $I_L^h \downarrow$, and $I_R^h \downarrow$, then α^h, I_L^h, and I_R^h can be recovered accordingly from them. Recovering α^h from $\alpha^l = (\alpha^h * G) \downarrow$ is exactly the problem discussed in Sec. 3.3, while I_L^h and I_R^h can be interpolated with bicubic method given their down-sampled version due to the smoothness assumption for them. $\lambda = 0.03$ is used for recovering the HR alpha channel in our experiments.

When applying the close form solution of alpha matting [20] on an image region R_i, the hard constraints for both sides are chosen according to the local topology and image gradient. Low contrast pixels are selected, since they correspond to pure color of one side. The matting algorithm in [20] is very robust in our experiments, even for very limited quantity of hard constraints.

Alpha matting is also used in [28], where the α value is extracted to get the sub-pixel location of the curve. A two color image prior is used in [3] for demosaicing, which assumes that each pixel within a local neighborhood is either one of two representative colors or a linear combination of them. This assumption is in essential quite similar to the idea of using alpha matting for SR. In [12], various aspects of color information is combined by a linear summation.

5. Experiments

Figure 7 shows an example of the entire process. After doing alpha channel SR, sharp and smooth edges can be obtained for salient edge segments. Fig. 7(c) shows the final result after global updating procedure. It can recover some subtle structures, thus enhance the image quality for the entire image. Compare to the results of bicubic interpolation and back-projection, the chessboard artifact is greatly reduced without introducing blur or ringing effect. It is also more natural compared with the result obtained by simply applying the proposed soft edge smoothness prior on three color channel separately (shown in Fig. 5).

Figure 8 illustrates the idea of alpha channel SR by one edge segment in Fig. 7(a). The LR patch is decomposed into two image patches and a LR alpha channel. Fig. 8(e) is the recovered HR alpha channel by the proposed soft edge smoothness prior in Sec. 3.3. Combining Fig. 8 (b) (c) (e) together by Eqn. 11 will produce a sharp and smooth edge, which is shown in Fig. 7 (b).

Figure 9 compares our method with some exemplar-based algorithms. The existing exemplar-based methods can produce very sharp edges. Compare to them, more smooth boundaries can be archived by our method, this makes the result looks natural. Fig. 10 shows another example for comparison with other reconstruction-based algorithms. Visually appealing result is obtained by our algorithm, even for very fine image structure, the chessboard artifact is largely suppressed. More results are shown in Fig. 11. Please refer the electronic version for better visualization of our experimental results.

Some quantitative results are shown in Table. 1. The RMS and edge RMS (ERMS) error per pixel for some example images are listed. Compare to bicubic interpolation and back-projection [17], although the visual effect is greatly improved, the error is just reduced a little bit. We run our experiments on a PIV3.4G PC with 2G RAM by Matlab implementation. The computation time for some example images are also shown in Table. 1. It greatly depends on the density of salient edge segments.

Table 1. Error reduction and computation time (for each box with two numbers, the 1^{st} is the RMS error, the 2^{nd} is the ERMS error).

	Lena	Head	Zebra	Temple
LR size	80×107	70×70	100×170	161×107
Bicubic	8.5 14.5	9.7 13.7	9.6 29.3	11.4 22.6
BP [17]	7.5 11.3	9.4 12.8	8.5 20.4	10.9 20.5
Ours	7.5 10.7	9.3 12.6	8.3 19.3	10.7 20.3
Time (s)	40	25	135	106

(a) (b) (c) (d) (e)

Figure 8. An example of edge decomposition and alpha channel SR, (a) LR input, (b) (c) true color on two sides of this edge segment, (d) LR alpha channel, (e) HR alpha channel by soft edge smoothness prior (please refer Fig. 7 for the edge position, and result comparison with other reconstruction-based methods).

6. Conclusion

In this paper, a novel single image super resolution algorithm is proposed. A soft edge smoothness measure is defined on a large neighbored 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, a novel approach is proposed. A closed form alpha matting algorithm is applied to decompose each edge segment. It makes possible a unified treatment of them. Visually appealing results for a large variety of images are obtained by this algorithm.

References

[1] J. Allebach and P. W. Wong. Edge-directed interpolation. In *ICIP*, 1996. 2

[2] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Trans. on PAMI*, 24(9):1167–1183, 2002. 1

[3] 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. 2, 6

[4] C. M. Bishop, A. Blake, and B. Marthi. Super-resolution enhancement of video. In *Proc. Artificial Intelligence and Statistics*, 2003. 2

[5] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *ICCV*, 2003. 1, 2, 3

[6] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. on PAMI*, 26(9):1124–1137, 2004. 3

[7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. on PAMI*, 23(11):1222–1239, 2001. 3

[8] F. Champagnat, G. L. Besnerais, and C. Kulcsar. Continuous super-resolution for recovery of 1-d image features: Algorithm and performance modeling. In *CVPR*, 2006. 2

[9] T. F. Chan, S. Osher, and J. Shen. The digital tv filter and nonlinear denoising. *IEEE Trans. on Image Processing*, 10(2):231–241, 2001. 5

[10] H. Chang, D. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *CVPR*, 2004. 2, 8

[11] M. Elad and A. Feuer. Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images. *IEEE Trans. on Image Processing*, 6(12):1646–1658, 1997. 2

[12] S. Farsiu, M. Elad, and P. Milanfar. Multi-frame demosaicing and super-resolution of color images. *IEEE Trans. on Image Processing*, 15(1):141–159, 2006. 6

[13] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Trans. on Image Processing*, 13(10):1327–1344, 2004. 2, 4, 5

[14] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 2002. 2

[15] W. T. Freeman, E. Pasztor, and O. Carmichael. Learning low-level vision. *IJCV*, 40(1):25–47, 2000. 2, 8

[16] X. C. He and N. Yung. Curvature scale space corner detector with adaptive threshold and dynamic region of support. In *ICPR*, 2004. 5

[17] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion and transparency. *JVCIP*, 1993. 1, 4, 6, 8

[18] A. Katsaggelos, R. Molina, and J. Mateos. *Super resolution of images and video*. Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool, 2007. 1

[19] D. Kong, M. Han, W. Xu, H. Tao, and Y. Gong. Video super-resolution with scene-specific priors. In *BMVC*, 2006. 2

[20] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *CVPR*, 2006. 2, 5, 6

[21] X. Li and M. Orchard. New edge-directed interpolation. *IEEE Trans. on Image Processing*, 10(10):1521–1527, 2001. 2

[22] Z. Lin and H.-Y. Shum. Fundamental limits of reconstruction based super-resolution algorithms under local translation. *IEEE Trans. on PAMI*, 26(1):83–97, 2004. 1

[23] 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. 2

[24] B. S. Morse and D. Schwartzwald. Image magnification using level set reconstruction. In *CVPR*, 2001. 2, 4

[25] V. Rabaud and S. Belongie. Big little icons. In *CVAVI*, 2005. 2, 4

[26] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992. 5

[27] J. Sun, N. Zheng, H. Tao, and H. Shum. Image hallucination with primal sketch priors. In *CVPR*, 2003. 2, 5

[28] Y.-W. Tai, W.-S. Tong, and C.-K. Tang. Perceptually-inspired and edge-directed color image super-resolution. In *CVPR*, 2006. 2, 5, 6

[29] 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. 2

[30] Q. Wang, X. Tang, and H. Shum. Patch based blind image super resolution. In *CVPR*, 2005. 2

A. Proof of Theorem. 2

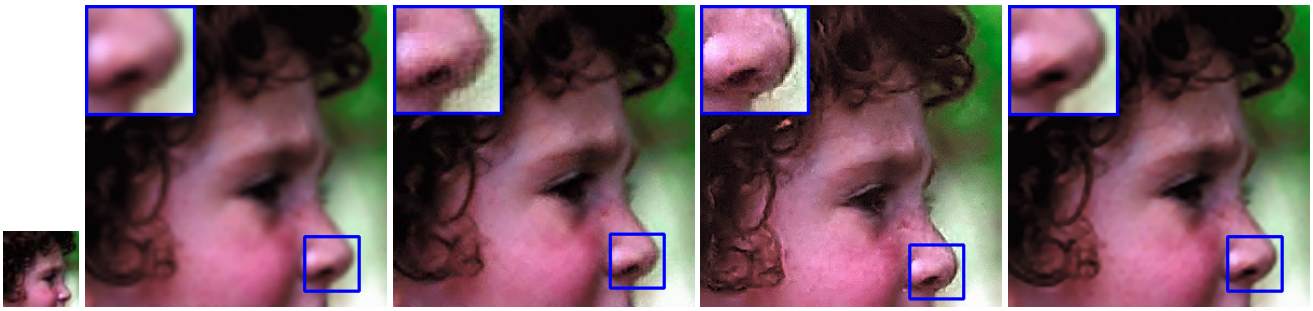
From Eqn. 3, we have

$$|l_i|_{\mathcal{G}} = \sum_k \big(w_k \sum_{N_k} |F_{l_i}(p) - F_{l_i}(q)|\big),$$

$$\begin{aligned}
\text{so} \quad \sum_i |l_i|_{\mathcal{G}} &= \sum_i \sum_k \big(w_k \sum_{N_k} |F_{l_i}(p) - F_{l_i}(q)|\big) \\
&= \sum_k \big(w_k \sum_{N_k} \sum_i |F_{l_i}(p) - F_{l_i}(q)|\big), \\
\text{and} \sum_i |F_{l_i}(p) - F_{l_i}(q)| &= \#\{i \mid F_{l_i}(p) \neq F_{l_i}(q)\} \\
&= n \cdot |S_d(p) - S_d(q)|, \\
\text{so} \quad \sum_i |l_i|_{\mathcal{G}} &= \sum_k \big(w_k \sum_{N_k} n \cdot |S_d(p) - S_d(q)|\big) \\
&= n \cdot \sum_k \big(w_k \sum_{N_k} |S_d(p) - S_d(q)|\big) \\
&= n \cdot |S_d|_{\mathcal{G}} \quad \text{(from Eqn. 4)}.
\end{aligned}$$

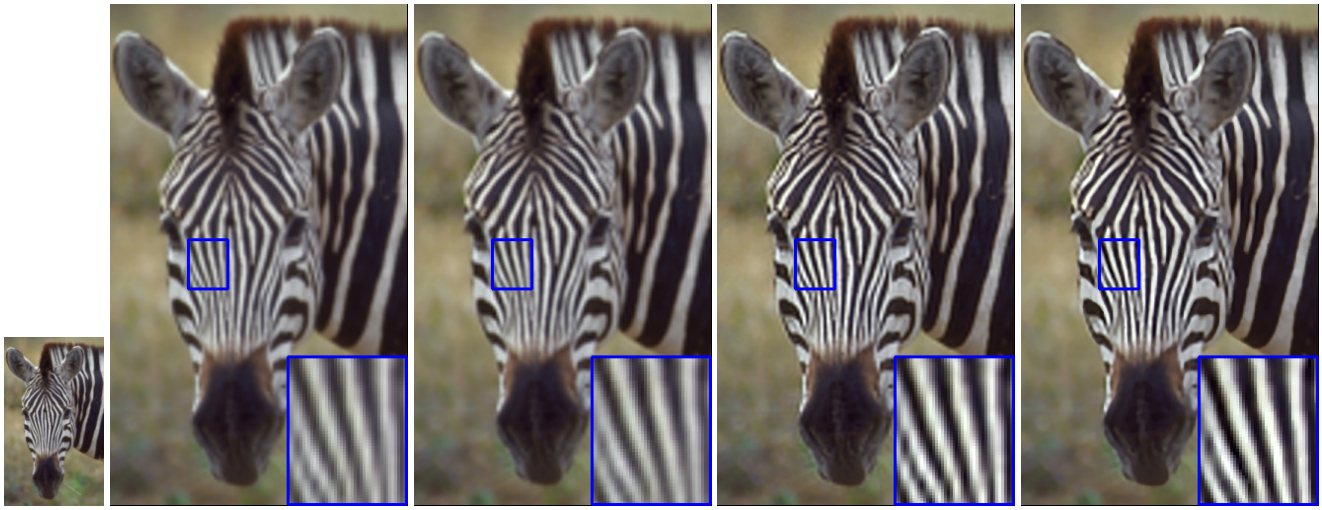
Thus from Theorem. 1, we have

$$|S_d|_{\mathcal{G}} = \frac{1}{n} \sum_i |l_i|_{\mathcal{G}} \to \frac{1}{n} \sum_i |l_i|_{\mathcal{E}}.$$



(a) LR input (b) bicubic (c) result in [15] (d) result in [10] (e) our result

Figure 9. Comparison results with exemplar-based methods ($\times 4$).



(a) LR input (b) bicubic (c) sharpened bicubic (d) back-projection [17] (e) our result

Figure 10. Comparison results with reconstruction-based methods ($\times 3$).

Figure 11. More results. For each pair of images, the upper one is the LR input, and the lower one is our result ($\times 3$).