Colorization for Single Image Super Resolution

Shuaicheng Liu[1], Michael S. Brown[1], Seon Joo Kim[1], Yu-Wing Tai[2]

[1] National University of Singapore
[2] Korea Advanced Institute of Science and Technology

Abstract. This paper introduces a new procedure to handle color in single image super resolution (SR). Most existing SR techniques focus primarily on enforcing image priors or synthesizing image details; less attention is paid to the final color assignment. As a result, many existing SR techniques exhibit some form of color aberration in the final upsampled image. In this paper, we outline a procedure based on image colorization and back-projection to perform color assignment guided by the super-resolution luminance channel. We have found that our procedure produces better results both quantitatively and qualitatively than existing approaches. In addition, our approach is generic and can be incorporated into any existing SR techniques.

Keywords: Super resolution, colorization, image upsampling

1 Introduction and Related Work

Image super resolution (SR) refers to techniques that estimate a high-resolution (HR) image from a single low-resolution (LR) image input. Strategies to address the image SR problem are typically categorized into three broad methods: interpolation based methods, reconstruction based methods, and learning based methods.

Interpolation based techniques (e.g., [1–4]) have their roots in sampling theory and interpolate the HR image directly from the LR input. While these approaches tend to blur high frequency details resulting in noticeable aliasing artifacts along edges, they remain popular due to their computational simplicity. Reconstruction based approaches (e.g., [5–13]) estimate an HR image by enforcing priors in the upsampling process. Such priors are commonly incorporated into a back-projection framework to reduce artifacts around edges while constraining the estimated HR image against the LR input. Learning based techniques estimate high frequency details from a training set of HR images that encode the relationship between HR and LR images (e.g., [14–21]). These approaches synthesize missing details based on similarities between the input LR image and the examples in the training set based on patch similarities. Hybrid approaches that combine elements of reconstruction and learning based methods have also been proposed (e.g., [22, 23]).

While these existing SR techniques have successfully demonstrated ways to enhance image quality through priors or detail hallucination – how to han-

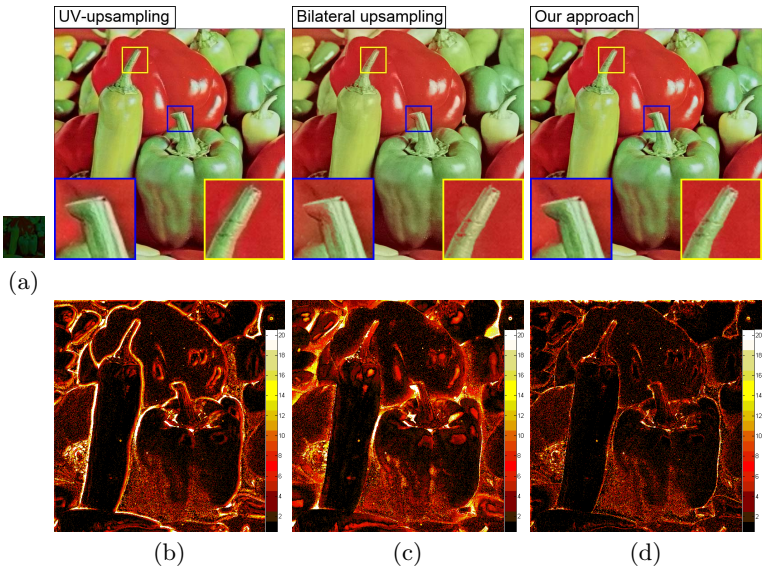

Fig. 1. (a) LR chrominance input. Results using bicubic interpolation of the UV channels (b), using joint-bilateral upsampling [25] (c), and our result (d). Color difference maps (bottom) are computed based on the CIEDE2000 color difference formula ([26, 27]).

dle color in the SR process has received far less attention. Instead, two simple approaches are commonly used to assign color. The first approach is to perform color assignment using simple upsampling of the chrominance values. This approach, used extensively in both reconstruction-based and learning-based SR (e.g. [12, 13, 19, 24]), first transforms the input image from RGB to another color space, most notably YUV. Super resolution is applied only to the luminance channel, Y. The chrominance channels, U and V, are then upsampled using interpolation methods (e.g. bilinear, bicubic) and the final RGB is computed by recombining the new SR luminance image with the interpolated chrominance to RGB. The second approach, used primarily in learning-based techniques (e.g. [14–16]), is to use the full RGB channels in patch matching for detail synthesis, thus directly computing an RGB output.

These two existing approaches for SR color assignment have drawbacks. The basis for the UV-upsampling approach is that the human visual system is more sensitive to intensities than color and can therefore tolerate the color inaccuracies in this type of approximation. However, color artifacts along the edges, are still observable, especially under large magnification factors as shown in Fig. 1. Performing better upsampling of the chrominance, by weighted average [28] or joint-bilateral filtering [25], can reduce these artifacts as shown in Fig. 1(c), but not to the same extent as our algorithm (Fig. 1(d)). In addition, techniques such as joint-bilateral upsampling requires parameter-tuning to adjust the Gaussian

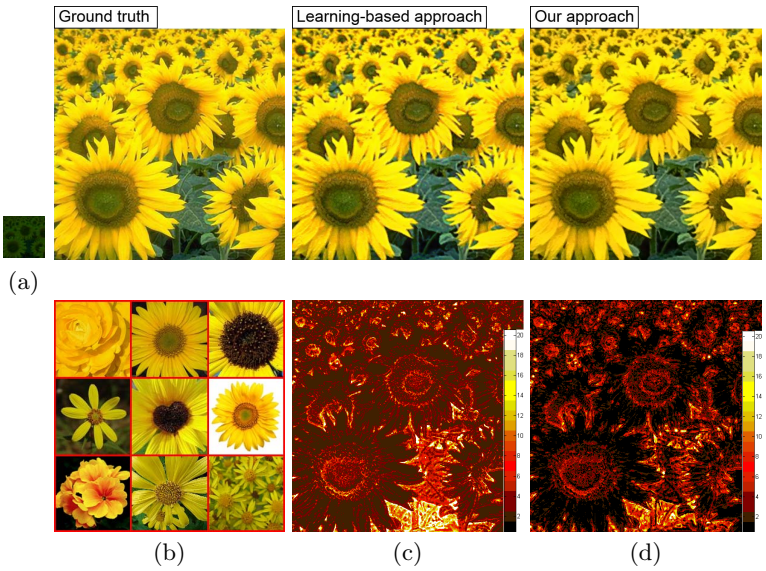

Fig. 2. (a) LR chrominance input, (b) ground truth image (top) and training images (bottom), (c) result using learning based SR [16], (d) our result. Color difference maps are computed based on the CIEDE2000 color difference formula ([26, 27]).

window size and parameters of the bi-lateral filter's spatial and range components to obtain optimal results.

For learning-based techniques, the quality of the final color assignment depends heavily on the similarity between the training data and the input image. The techniques that perform full RGB learning can exhibit various color artifacts when suitable patches cannot be found in the the training data. Approaches that apply learning-based on the luminance channel in tandem with UV-upsampling can still exhibit errors when the estimated SR luminance images contains contrast shifts due to training set mismatches. Since back-projection is often not used in learning-based techniques, this error in the SR luminance image can lead to color shifts in the final RGB assignment. Fig. 2 shows examples of the color problems often found in learning-based approaches.

In this paper, we propose a new approach to reconstruct colors when performing single image super resolution. As with chrominance upsampling, our approach applies super resolution only to the luminance channel. Unique to our approach, however, is the use of image colorization [29, 30] to assign the chrominance values. To do this, we first compute a chrominance map that adjusts the spatial locations of the chrominance samples supplied by the LR input image. The chrominance map is then used to colorize the final result based on the SR luminance channel. When applying our approach to learning-based SR techniques, we also introduce a back-projection step to first normalize the luminance channel before image colorization. We show that this back-projection procedure has

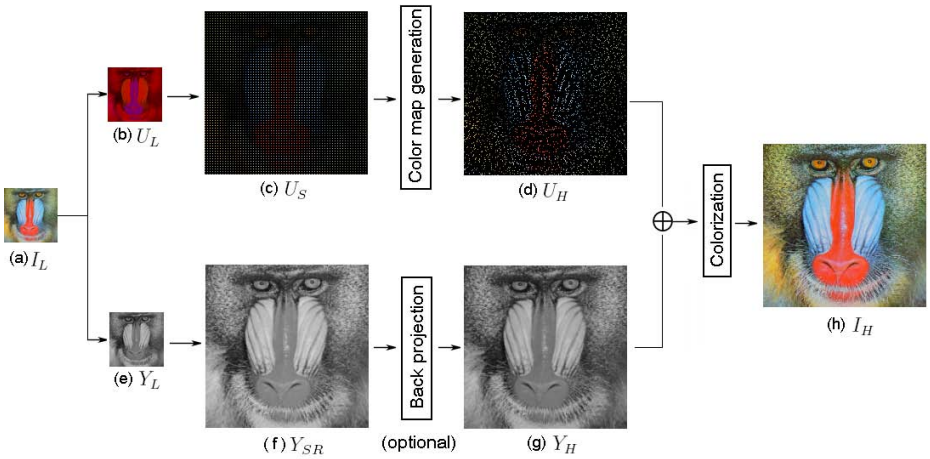

Fig. 3. The pipeline of our algorithm. (a) LR input image. (b) The chrominance component of input image. (c) Initial chrominance map produced by expanding (b) with the desired scale without any interpolation. (d) Adjusted chrominance map. (e) The luminance component of input image. (f) Upsampled image using any SR algorithm. (g) Upsampled image produced by adding the back-projection constraint (if necessary). (h) Final color SR image obtained by combining the color map (d) and the SR luminance image (g) using colorization.

little adverse impact on the synthesized details. Our approach not only shows improvements both visually and quantitatively, but is straight-forward to implement and requires no parameter tuning. Moreover, our approach is generic and can be used with any existing SR technique.

The remainder of this paper discusses our SR color assignment procedure and demonstrates results on several examples using both reconstruction and learning-based techniques. The paper is concluded with a short discussion and summary.

2 Colorization Framework for Super Resolution

The pipeline of our approach is summarized in Fig. 3. Given a LR color image (Fig.3 (a)), our goal is to produce a SR color image (Fig.3 (h)). To achieve this goal, the input LR image is first decomposed into the luminance channel Y_L and the chrominance channels U_L and V_L . For simplicity, we use only the U channel to represent chrominance since the operations on the U and V channels are identical. For the luminance, the HR luminance channel Y_H is constructed from Y_L by using any preferred SR algorithm. To assign the RGB colors to the final SR image I_H, we use the colorization framework introduced by Levin et al. [29]. For the colorization, we introduce a method to generate chrominance

samples which act as the seeds for propagating color to the neighboring pixels. The chrominance samples are obtained from the low resolution input, U_L, however the spatial arrangement of these chrominance values are generated automatically from the relationships between intensities in Y_L and Y_H.

Before we explain the colorization scheme, we note that we apply back-projection for computing Y_H from Y_L when the selected SR algorithm does not already include the back-projection procedure. We explain the reason for this first, before describing the colorization procedure.

2.1 Luminance Back-projection

Enforcing the reconstruction constraint is a standard method which is used in many reconstruction based algorithms [9–13]. The difference among these various approaches is the prior imposed on the SR image. In our framework, the reconstruction constraint is enforced by minimizing the back-projection error of the reconstructed HR image Y_H against the LR image Y_L without introducing extra priors. This can be expressed as as:

$$Y_H = \arg\min_{Y_H} \|Y_L - (Y_H \otimes h) \downarrow \|^2, \tag{1}$$

where $\downarrow$ is the downsampling operator and $\otimes$ represents convolution with filter h with proportional to the magnification factor.

Assuming the term $Y_L - (Y_H \otimes h) \downarrow$ follows a Gaussian distribution, this objective equation can be cast as a least squares minimization problem with an optimal solution Y_H obtained by the iterative gradient descent method [5].

The reason to incorporate the reconstruction constraint is that the desired output should have the similar intensity values as the input image. As discussed in Section 1, learning-based techniques often suffer from luminance shifts due to training example mismatches. Conventional wisdom is that back-projection may remove hallucinated details, however, we found that adding this procedure had little effect on the synthesized details. Fig. 4 shows an example of the gradient histogram of the original Y_{SR} as more iterations of back-projection are applied. We can see that the gradient profiles exhibit virtually no change, while the color errors measured using the CIEDE200 metric against the ground truth are significantly reduced. This is not too surprising given that the estimated luminance image is downsampled in the back-projection process described in Eq. (1). Thus, back-projection is correcting luminance mismatches on the low-pass filtered image, allowing the fine details to remain. For SR techniques that already includes back-projection, this step can be omitted.

3 Colorization Scheme

The core of our approach lies in using image colorization to propagate the chrominance values from the LR input in order to add color to the upsampled SR luminance image. In [29], a gray-scale image is colorized by propagating chrominance values which are assigned via scribbles drawn on the image by the user.

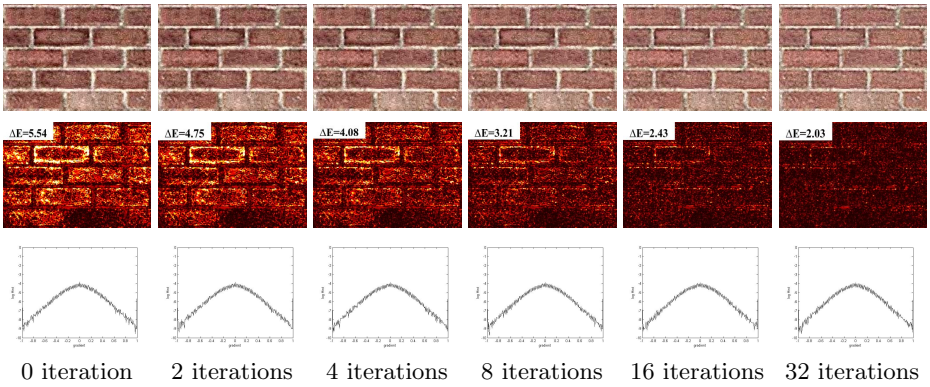

Fig. 4. Illustration of the benefits of back-projection. Estimated HR images (top), their CIEDE2000 color difference maps (middle), and gradient magnitude profiles (bottom) are shown at different iterations based on Eq. (1).

In our approach, the initial chrominance assignment comes from the LR image. The positions of these assignments are adjusted to better fit the HR luminance channel. We first review the image colorization and then describe our procedure to build the chrominance map.

3.1 Image Colorization

Image colorization [29] computes a color image from a luminance image and a set of sparse chrominance constraints. The unassigned chrominance values are interpolated based on the assumption that neighboring pixels $\mathbf{r}$ and $\mathbf{s}$ should have similar chrominance values if their intensities are similar. Thus, the goal is to minimize the difference between the chrominance $U_H(\mathbf{r})$ at pixel $\mathbf{r}$ and the weighted average of the chrominance at neighboring pixels:

$$E = \sum_{\mathbf{r}} (U_H(\mathbf{r}) - \sum_{\mathbf{s} \in N(\mathbf{r})} w_{\mathbf{rs}} U_H(\mathbf{s})) \tag{2}$$

where $w_{\mathbf{rs}}$ is a weighting function that sums to unity. The weight $w_{\mathbf{rs}}$ should be large when $Y_H(\mathbf{r})$ is similar to $Y_H(\mathbf{s})$, and small when the two luminance values are different. This can be achieved with the affinity function [29]:

$$w_{\mathbf{rs}} \propto e^{-(Y_H(\mathbf{r}) - Y_H(\mathbf{s}))^2 / 2\sigma_r^2} \tag{3}$$

where σ_r is the variance of the intensities in a 3×3 window around $\mathbf{r}$. The final chrominance image is obtained by minimizing Eq. 2 based on the input luminance image and chrominance constraints. The final RGB image is computed by recombining the luminance and the estimated chrominance.

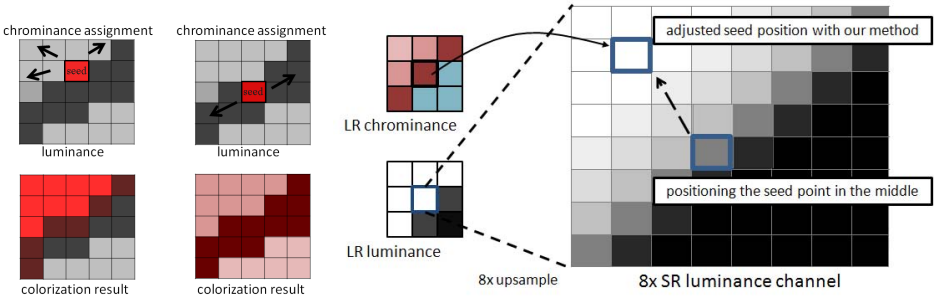

Fig. 5. (a) The effect of the chrominance seed position on the final colorization result are shown. The arrows indicate the chrominance propagation based on the intensity affinity based on the seed location. (b) Our aim is to adjust the seed point to be located at a position in the HR luminance result that is more similar the LR image luminance. This will produce a better colorization result.

3.2 Chrominance map generation

To perform image colorization, chrominance values must be assigned to a set of pixels, or seed points, from which the color is propagated. In [29], scribbles from the user-input are used as the initial assignment of color. In this paper, the chrominance from the LR image is used for the initial color assignment. For example, for an $8\times$ upsampling, a pixel in the LR image can be mapped to any of the pixels in the corresponding 8×8 block of corresponding HR pixels. The key in our colorization scheme lies in the positioning of the seed points in the upsampled image since blindly assigning the chrominance value to the middle of the patch may not produce the best result and can likely result in undesired color bleeding. This is illustrated in Fig. 5(a), where the we see that the estimated chrominance values are sensitive to the position of the seed point (i.e. hard constraint), especially on the edges.

Our strategy is to place the chrominance value in a position in the upsampled patch where the luminance value of the computed SR (Y_H) is closest to the original LR pixel's intensity (Y_L) as shown in Fig. 5(b). This approach, however, can be sensitive to noise and we therefore introduce a simple Markov Random Field (MRF) formulation to regularize the search direction for assigning the seed point. The idea is that the neighboring seed points are likely to share the same search direction in the HR image. Fig. 6 outlines the approach using an example with $8\times$ upsampling.

The search directions are discretized into four regions (Fig. 6 (a)) which serve as the four labels of the MRF, i.e. $l_x \in \{0, 1, 2, 3\}$. Let $\mathbf{x}$ be a pixel coordinate in the LR image and $\mathbf{X}$ be the upsampled coordinate of the point $\mathbf{x}$. Let $N_i(\mathbf{X})$ be the neighborhood of $\mathbf{X}$ in the direction i, where $i \in \{0, 1, 2, 3\}$. A standard MRF formulation is derived as:

$$E = E_d + \lambda E_s, \tag{4}$$

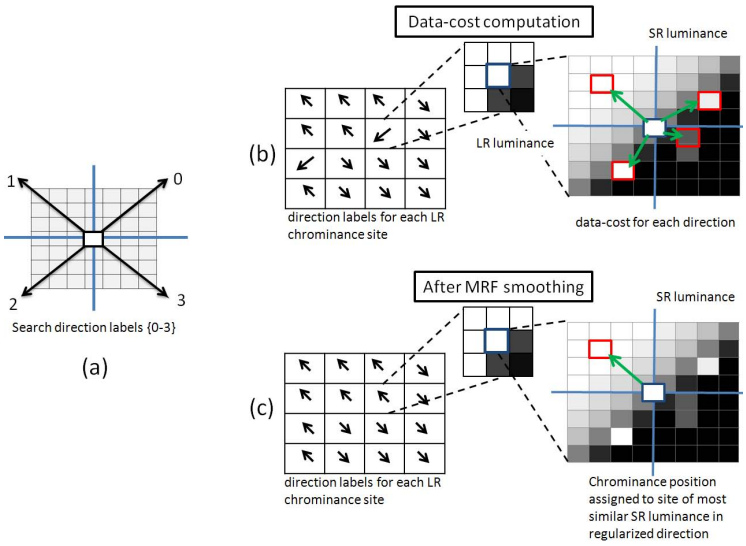

Fig. 6. The MRF example: (a) Discretized search directions. (b) Data cost computation in each search direction. (c) Smoothness constraint to regularize results. The MRF smoothness prior regularizes the search direction to be similar to the search directions of neighboring LR pixels.

where E_d is the data cost of assigning a label to each point $\mathbf{x}$ and E_s is the smoothness term representing the cost of assigning different labels to adjacent pixels. The term λ serves as the typical balancing weight between the data cost and the smoothness cost. Each cost is computed as follows :

$$E_d(l_{\mathbf{x}} = i) = \min_{\mathbf{Z} \in N_i(\mathbf{X})} |Y_L(\mathbf{x}) - Y_H(\mathbf{Z})|\,, \tag{5}$$

and

$$E_s(l_p, l_q) = f(l_p, l_q) \cdot g(Y_{pq}), \tag{6}$$

where $f(l_p, l_q) = 0$ if $l_p = l_q$ and $f(l_p, l_q) = 1$ otherwise. The term $g(\xi) = \frac{1}{\xi+1}$ with $Y_{pq} = \|Y_L(p) - Y_L(q)\|^2$, where $\mathbf{p}$ and $\mathbf{q}$ are neighboring pixels. This weighting term encourages pixels with similar LR luminance intensity values to share the same directional label. The MRF labels are assigned using the belief propagation (BP) algorithm [31].

After computing the search direction using the MRF regularization, the chrominance value from the LR image is placed on the pixel with the most similar luminance value in the regularized search direction. Fig. 7 shows an example of the results obtained before and after applying the chrominance map adjustment. Bleeding is present without the adjustment, however, the results is much closer to the ground truth with the adjustment .

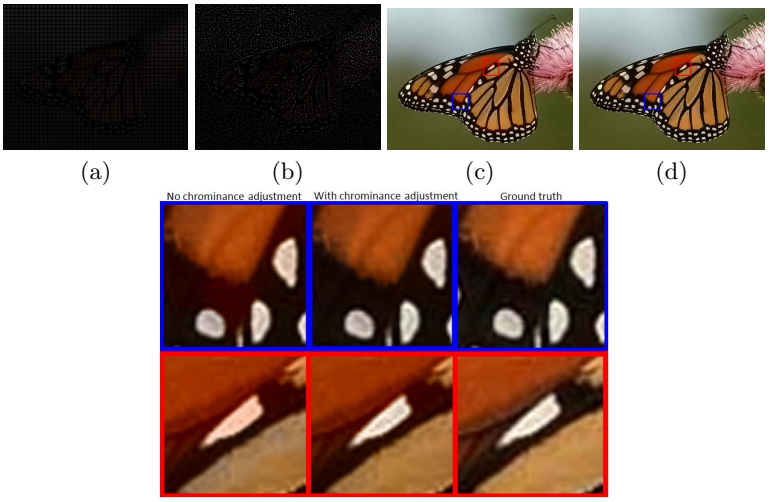

Fig. 7. (a) Initial color map U_S. (b) Color map U_H. (c) Colorization result using (a). (d) Colorization result using (b). Color map (b) produce better results without leakage at boundaries since the chrominance points are well located.

Fig. 8. (Top) Images used for our experiments. (Bottom) Images used as the training examples for the learning-based SR.

4 Experimental Results

Here we show results of our colorization scheme on 4 representative images shown in Fig. 8. For brevity, we only show the error maps and selected zoomed regions. Full resolution images of our results, together with additional examples, are available online. For the color difference measure, we use the CIEDE2000 metric [26, 27] together with a "hot" color-map. The mean color errors, ΔE, for all pixels as defined by the CIEDE2000 metric are provided.

The first two results are shown in Fig. 9 and Fig. 10. The images have been upsampled using $4\times$ magnification using the recent reconstruction based SR algorithm in [13]. The results were produced with executable code available on the author's project webpage. Our colorization results are compared with the de facto UV-upsampling technique (also used in [13]). As can be seen, the overall

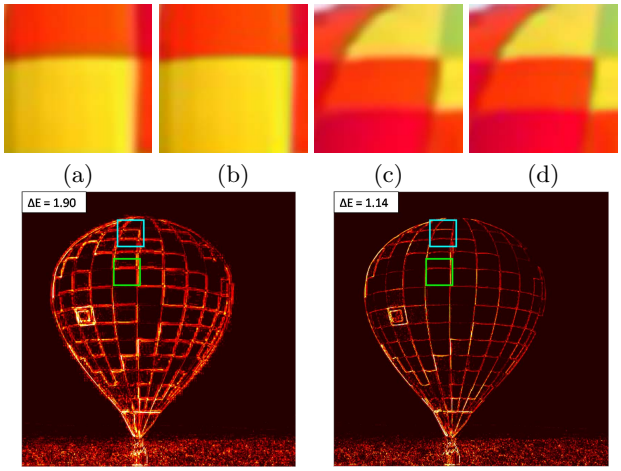

Fig. 9. Example 1 (Ballon): 4× reconstruction-based upsampling has been applied to the "ballon" image. UV-upsampling (a,c) is compared with our result (b,d).

error maps for our results are better. For the zoomed regions, we can see that artifacts about edges are less noticeable using our technique.

The next two results are shown in Fig. 11 and Fig. 12. Fig. 8 (bottom) shows the training images used for the learning examples, which are the the same images used in the [16]. We use our own implementation of the full RGB learning method using the one-pass algorithm described in [16]. For our results, we first apply back-projection on the SR luminance channel before performing the colorization step. Learning-based techniques exhibit more random types of color artifacts, however, our approach is still able to improve the results as shown in the errors maps and zoomed regions.

The final example demonstrates the benefits of the optional back-projection procedure when the SR luminance image exhibits significant intensity shifting. In this example, only two of the training images are used to produce the SR image. Fig. 13(a) shows the result and the associated error. Fig. 13(b) shows our results obtained by only applying the colorization step and Fig. 13(c) shows the results when back-projection is used followed by our colorization method. We can see the error is significantly reduced when the back-projection procedure is incorporated.

5 Discussion and Summary

The focus of this paper is on assigning the final color values in the super resolution pipeline, and not how to perform SR itself. Therefore, our results are affected by the quality of the SR technique used, which is evident in the learning-based examples which tend to produce a higher overall error. However, even in these examples, our approach is able to offer a better final color assignment when com-

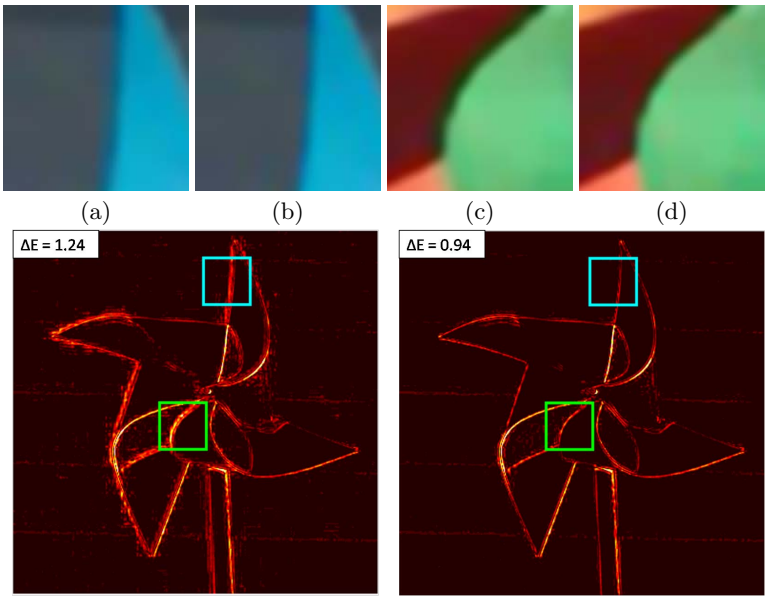

Fig. 10. Example 2 (Pinwheel): 4× reconstruction-based upsampling has been applied to the "pinwheel" image. UV-upsampling (a,c) is compared with our result (b,d).

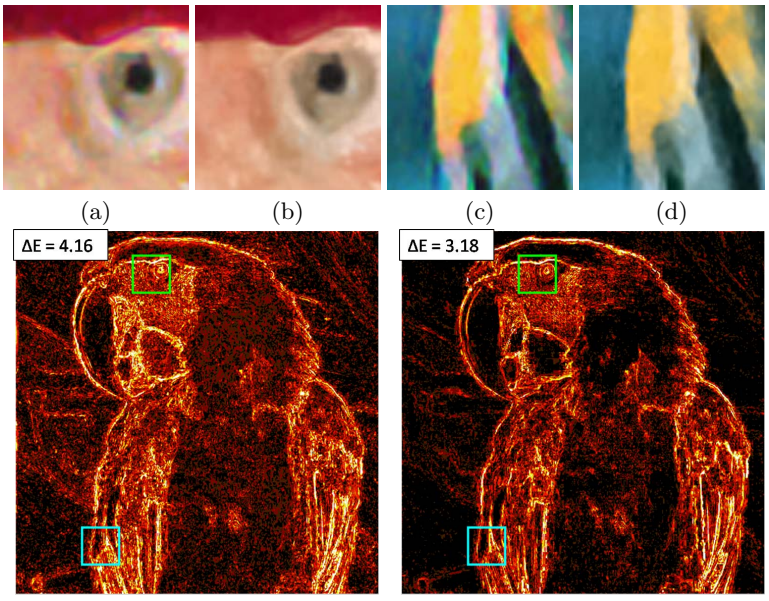

Fig. 11. Example 3 (Parrot): 4× learning-based upsampling (a,c) has been applied to the the "parrot" image. Full RGB SR is compared with our result (b,d).

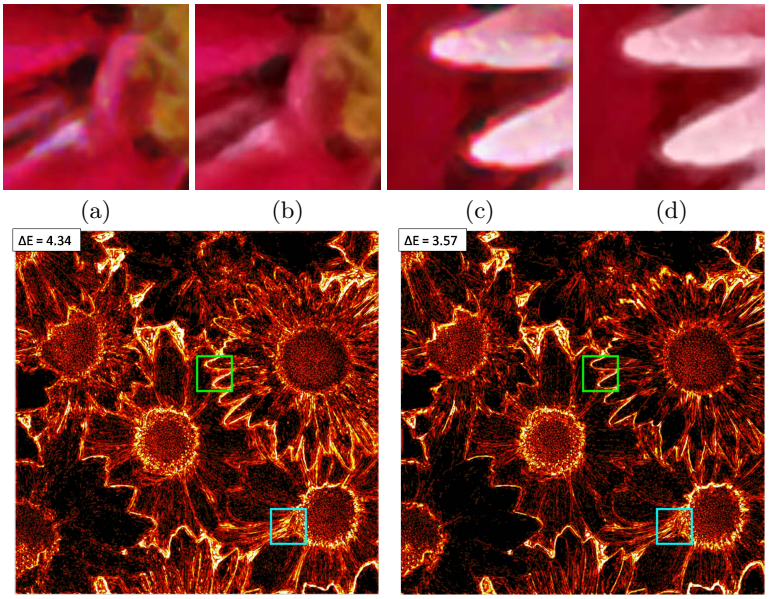

Fig. 12. Example 4 (Flowers): Example 2 (Parrot): 4× learning-based upsampling (a,c) has been applied to the the "parrot" image. Full RGB SR is compared with our result (b,d).

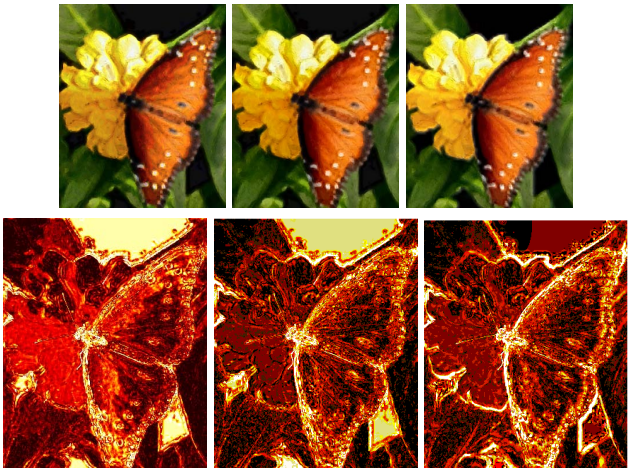

Fig. 13. Example showing the benefits of back-projection. (a) learning-based result; (b) our approach without back-projection; (c) our approach with back-projection.

pared with the ground truth. For reconstruction-based approaches, our overall edges appear sharper compared to basic UV-upsampling. We note that our approach inherits the limitations of image colorization. In particular, color bleeding may occur in regions with different chrominance but similar luminance values. However, the reasonably dense chrominance sampling from the LR image helps to keep such artifacts localized.

While we introduce an MRF regularization to aid in the chrominance map assignment, poor assignment of chrominance values can obviously result in undesired artifacts. Our quantitative measurements suggest our current approach is reasonable. We envision that better results could be obtained in the future with more sophisticated strategies for the chrominance placement.

In summary, we have introduced a new approach for assigning colors to SR images based on image colorization. Our approach advocates using back-projection with learning-based techniques and describes a method to adjust the chrominance values before performing image colorization. Our approach is generic and can be used with any existing SR algorithms.

Acknowledgements

This work was supported by Singapore MOE AcRF Tier 1 grant (Project No: R-252-000-333-133).

References

1. Allebach, J., Wong, P.: Edge-directed interpolation. In: Proc. IEEE International Conf. on Image Processing. (1996) 707–710
2. Li, X., Orchard, M.T.: New edge-directed interpolation. In: Proc. IEEE International Conf. on Image Processing. (2000) 311 – 314
3. Caselles, V., Morel, J.M., Sbert, C.: An axiomatic approach to image interpolation. IEEE Trans. on Image Processing **7** (1998) 376–386
4. Thevenaz, P., Blu, T., Unser, M.: Image Interpolation and Resampling. Academic Press, USA (2000)
5. Irani, M., Peleg, S.: Motion analysis for image enhancement: Resolution, occlusion, and transparency. Journal of Visual Communication and Image Representation **4** (1993) 324–335
6. Morse, B., Schwartzwald, D.: Image magnification using level-set reconstruction. In: Proc. IEEE International Conf. Computer Vision. (2001) 333–341
7. Tappen, M.F., Russell, B.C., Freeman, W.T.: Exploiting the sparse derivative prior for super-resolution and image demosaicing. In: IEEE Workshop on Statistical and Computational Theories of Vision. (2003) 2074 – 2081
8. Lin, Z., Shum, H.: Fundamental limits of reconstruction-based superresolution algorithms under local translation. IEEE Trans. on Pattern Analysis and Machine Intelligence **26** (2004) 83–97
9. Tai, Y.W., Tong, W.S., Tang, C.K.: Perceptually-inspired and edge directed color image super-resolution. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2006) 1948–1955
10. Dai, S., Han, M., Xu, W., Wu, Y., Gong, Y.: Soft edge smoothness prior for alpha channel super resolution. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2007) 1–8

11. Ben-Ezra, M., Lin, Z., Wilburn, B.: Penrose pixels: Super-resolution in the detector layout domain. In: Proc. IEEE International Conf. Computer Vision. (2007) 1–8
12. Sun, J., Sun, J., Xu, Z., Shum, H.: Image super-resolution using gradient profile prior. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2008) 1–8
13. Shan, Q., Li, Z., Jia, J., Tang, C.K.: Fast image/video upsampling. ACM Trans. Graph. (Proc. of SIGGRAPH ASIA) **27** (2008) 1–7
14. Freeman, W.T., Pasztor, E.C., Carmichael, O.T.: Learning low-level vision. International Journal of Computer Vision **40** (2000) 25 – 47
15. Liu, C., Shum, H.Y., Zhang, C.S.: Two-step approach to hallucinating faces: global parametric model and local nonparametric model. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2001) 192–198
16. Freeman, W.T., Jones, T., Pasztor, E.C.: Example-based super-resolution. IEEE Computer Graphics and Applications **22** (2002) 56–65
17. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. on Pattern Analysis and Machine Intelligence **24** (2002) 1167–1183
18. Sun, J., Zheng, N.N., Tao, H., Shum, H.Y.: Image hallucination with primal sketch prior. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2003) 729–736
19. Yeung, D., Chang, H., Xiong, Y.: Super resolution through neighbor embedding. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2004) 275–282
20. Wang, Q., Tang, X., Shum, H.Y.: Patch based blind image super resolution. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2005) 709–716
21. Liu, C., Shum, H.Y., Freeman, W.T.: Face hallucination: Theory and practice. International Journal of Computer Vision **75** (2007) 115–134
22. Tai, Y.W., Liu, S., Brown, M.S., Lin, S.: Super resolution using edge prior and single image detail synthesis. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2010) 1–8
23. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: Proc. IEEE International Conf. Computer Vision. (2009) 349 – 356
24. Jianchao, Y., John, W., Thomas, H., Yi, M.: Image super-resolution as sparse representation of raw image patches. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2008) 1–8
25. Kopf, J., Cohen, M., Lischinski, D., Uyttendaele, M.: Joint bilateral upsampling. ACM Transactions on Graphics (Proc. of SIGGRAPH) **26** (2007)
26. Johnson, G.M., Fairchild, M.D.: A top down description of S-CIELAB and CIEDE2000. Color Research and Application **28** (2002) 425 – 435
27. Sharma, G., Wu, W., Dalal, E.D.: The CIEDE2000 color difference formula: Implementations notes, supplementary test data and mathematical observations. Color Research and Application **30** (2005) 21 – 30
28. Fattal, R.: Upsampling via imposed edges statistics. ACM Trans. Graph. (Proc. of SIGGRAPH) **26** (2007)
29. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. ACM Trans. Graph. (Proc. of SIGGRAPH) **23** (2004) 689–694
30. Liu, X., Wan, L., Qu, Y., Wong, T.T., Lin, S., Leung, C.S., Heng, P.A.: Intrinsic colorization. ACM Transactions on Graphics (Proc. of SIGGRAPH Asia) **27** (2008) 152:1–152:9
31. Tappen, M.F., Freeman, W.T.: Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters. In: Proc. IEEE International Conf. Computer Vision. (2003) 900–907