Sub-Band Energy Constraints For Self-Similarity Based Super-Resolution

Abhishek Singh and Narendra Ahuja
University of Illinois at Urbana-Champaign
Email: asingh18@illinois.edu, n-ahuja@illinois.edu

***Abstract*—In this paper, we propose a new self-similarity based single image super-resolution (SR) algorithm that is able to better synthesize fine textural details of the image. Conventional self-similarity based SR typically uses scaled down version(s) of the given image to first build a dictionary of low-resolution (LR) and high-resolution (HR) image patches, which is then used to predict the HR patches for each LR patch of the given image. However, metrics like pixelwise sum of squared differences (L_2 distance) make it difficult to find matches for high frequency textured patches in the dictionary. Textural details are thus often smoothed out in the final image. In this paper, we propose a method to compensate for this loss of textural detail. Our algorithm uses the responses of a bank of orientation selective bandpass filters to represent texture instead of the spatial variation of intensity values directly. Specifically, we use the energies contained in different sub-bands of an image patch to separate different types of details of a texture, which we then impose as additional priors on the patches of the super-resolved image. Our experiments show that for each patch, the low energy sub-bands (which correspond to fine textural details) get severely attenuated during conventional L_2 distance based SR. We propose a method to learn this attenuation of sub-band energies in the patches, using scaled down version(s) of the given image itself (without requiring external training databases), and thus propose a way of compensating for the energy loss in these sub-bands. We demonstrate that as a consequence, our SR results appear richer in texture and closer to the ground truth as compared to several other state-of-the-art methods.**

I. Introduction

The single image super-resolution (SR) problem involves estimating pixels in a high-resolution (HR) image from the smaller number of pixels available in its given, low-resolution (LR) version. Being fundamentally ill-posed, priors or regularizers are a key component in addressing this problem.

Over the last few years, learning based priors have demonstrated considerable success on this problem [5], [13], [4], [6], [12], [2]. In general, these priors exploit some form of statistical regularity in properties (such as gradient distributions, patch recurrence, etc.) of natural images, which is learnt from training data. A popular class of such methods seeks similar patches across scales of the given image to build a database of LR-HR patch pairs. This database or dictionary is then used to predict the HR patch corresponding to each patch of the given image [2], [6], [4]. Such self-similarity methods find their roots in fractal image coding from the 1990s [1]. More recently, such methods have also been justified as priors by studies on the *internal* statistics of natural images which suggest that patches from natural images tend to recur within and across scales in the same image [15].

These recent studies have also shown that the likelihood of finding a good match for a patch falls, as the gradient content of the image increases [15]. This suggests that textural details like hair, animal fur etc, often find suboptimal matches, using a self-similarity approach. This problem can be partly attributed to the limitation in using distance metrics such as pixel-wise sum of squared difference (L_2 distance) for matching textural patches. The L_2 distance between two patches is largely determined by the high contrast and prominent structures (macrostructures) in the patch, and is less sensitive to the fine details (microstructures) of the patch. Indeed, this problem manifests itself in the final results of patch based SR reconstruction methods - poor patch matches lead to inconsistent explanations of pixels in textural regions, and fine textural details or microstructures are thus averaged out.

In this paper, we propose a solution to the above problem. We argue that the L_2 distance by itself is not a sufficient criterion to find suitable matches for textural patches. Indeed, metrics based on pixelwise differences have been rather unsuccessful in applications such as texture classification or texture retrieval. On the other hand, texture descriptors based on responses to a multi-orientation bank of bandpass filters have been effective for such tasks [8], [14], [11]. In the SR application at hand, we therefore combine the conventional L_2 distance based patch matching procedure with additional prior constraints on the energies of the different orientation selective sub-bands of the patch. We observe through experiments in this paper that for each patch, the low energy sub-bands (which correspond to fine textural details) get severely attenuated during conventional L_2 distance based SR. Based on this observation, we propose a method to learn this attenuation of sub-band energies in the patches, using scaled down version(s) of the given image itself (without requiring external training databases), and thus propose a way of compensating for the energy loss in these sub-bands of each patch. More specifically, we propose the use of scaling coefficients to boost the sub-bands of the patch that constitute the fine textural details (microstructures). As a consequence, our SR results appear richer in texture and more natural as compared to state-of-the-art methods, as shown by our experiments.

In the next section, we present a stepwise summary of the proposed algorithm. The subsequent sections present details of the steps involved.

II. Algorithm Overview

Notation. We denote the given image to be super-resolved as I_0. By I_1 we denote the HR version of I_0, whose linear dimension, or scale, is larger by a factor of s. Similarly, we denote by I_{-1}, I_{-2} etc., the smaller versions of I_0, by scaling factors of $1/s, 1/2s$ etc., respectively. We denote the super-resolved image(s) obtained using our algorithm using a *hat*

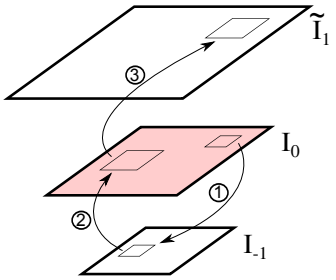


Fig. 1. Conventional self-similarity based SR. Given LR image I_0 is shown in red. Each patch of I_0 is matched to $k = 5$ most similar patches in I_{-1} in step 1. For simplicity we show only $k = 1$ most similar patch in this figure. The corresponding patch (in the same location) in I_0 serves as the HR predictor (step 2). This patch is then pasted in the HR image $\tilde{I}_1$ (step 3). See Section III for details.

($\hat{}$) symbol. Therefore, our objective is to super-resolve I_0 to obtain an HR image $\hat{I}_1$, that best approximates the true HR image I_1. We use scripted letters to denote sets, we use boldface lowercase letters to denote image patches, and lowercase italicized letters to denote scalars and indices.

Algorithm Summary. To obtain $\hat{I}_1$, from I_0, our proposed algorithm involves the following steps:

1) Using I_0, we first compute an intermediate HR image $\tilde{I}_1$ that is obtained by the conventional patch-similarity based SR approach, along the lines proposed earlier [15], [6], [4]. We present the general framework of such an algorithm in Fig. 1, and its details in Section III.

2) For each patch $\tilde{\mathbf{p}}$ of $\tilde{I}_1$, we compute the response of bank of R orientation selective bandpass filters, yielding the sub-bands $\{\tilde{\mathbf{p}}^{(1)}, \tilde{\mathbf{p}}^{(2)}, ..., \tilde{\mathbf{p}}^{(R)}\}$. To selectively amplify the patch macrostructure vs. microstructure, we differentially scale the patch's energy contents in different sub-bands by using the coefficients $\{\alpha^{(j)}\}_{j=1}^{R}$, to yield a transformed set of bandpass patches,

$$\hat{\mathbf{p}}^{(j)} = \alpha^{(j)} \tilde{\mathbf{p}}^{(j)} \qquad j = 1, 2, ..., R. \tag{1}$$

We discuss our algorithm that learns these coefficients in Sections IV and V. The scaling coefficients $\alpha^{(j)}$ allow us to impose sub-band energy constraints on each patch of the super-resolved image $\tilde{I}_1$, to minimize the loss of textural detail in the patch.

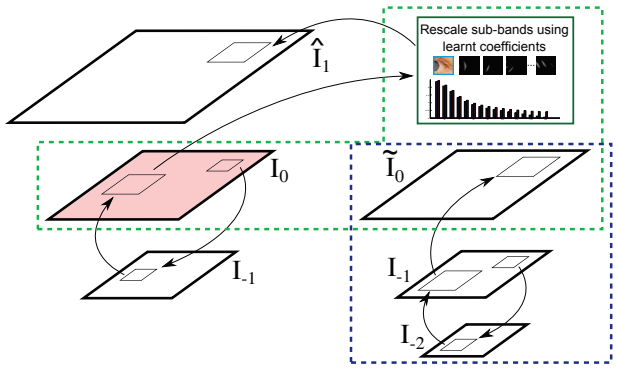


Fig. 2. Schematic summary of proposed algorithm. The left part of the figure depicts the conventional patch similarity based SR procedure of Fig. 1. Each HR patch obtained using the conventional SR algorithm is further enhanced by amplifying its sub-bands using scaling coefficients. These scaling coefficients are obtained by learning the attenuation caused in the sub-bands of the patches of $\tilde{I}_0$, compared to those of I_0. The image pair I_0 and $\tilde{I}_0$ (green dotted box) therefore serves as a source of training patches. $\tilde{I}_0$ is obtained by using the conventional patch-similarity based SR (blue dotted box) with I_{-1} as input. See text for details.

3) The rescaled sub-bands $\{\hat{\mathbf{p}}^{(j)}\}_{j=1}^{R}$ of each patch are recombined to yield the texture-enhanced patch $\hat{\mathbf{p}}$, and all such enhanced patches constitute the super-resolved image $\hat{I}_1$. Finally, we also then run a few iterations of the classical backprojection constraint [7], to ensure that the blurred and downsampled version of $\hat{I}_1$ matches the given LR image I_0. We elaborate on this step in Section VI.

A schematic summary of our algorithm is presented in Fig. 2. The details of each step follow.

III. Patch-Similarity Based SR

The conventional patch-similarity based SR approach that we adopt to obtain $\tilde{I}_1$ from I_0 follows similar steps as done in existing work [15], [6], [4], and is summarized in Fig. 1. Given the LR image I_0, we first obtain its downsampled version,

$$I_{-1} = (I_0 * f_{psf}) \downarrow \tag{2}$$

where f_{psf} is an assumed point spread function. We then create two sets of image patches $\mathcal{L}$ and $\mathcal{H}$, that contain patches from I_{-1} and their corresponding (bigger) patches extracted from I_0, respectively. The sets $\mathcal{L}$ and $\mathcal{H}$ serve as our database of LR-HR training patches. To super-resolve the given image I_0 to $\tilde{I}_0$, for every patch $\mathbf{l}$ of I_0, we look for its $k = 5$ most similar patches $\{\mathbf{l}_i\}_{i=1}^{k}$ in the LR set $\mathcal{L}$, based on L_2 distances. Their corresponding HR patches $\{\mathbf{h}_i\}_{i=1}^{k}$ from the set $\mathcal{H}$ serve as individual predictors for the patch $\mathbf{l}$. We average $\{\mathbf{h}_i\}_{i=1}^{k}$ to estimate the HR patch $\tilde{\mathbf{h}}$ of $\mathbf{l}$ as follows,

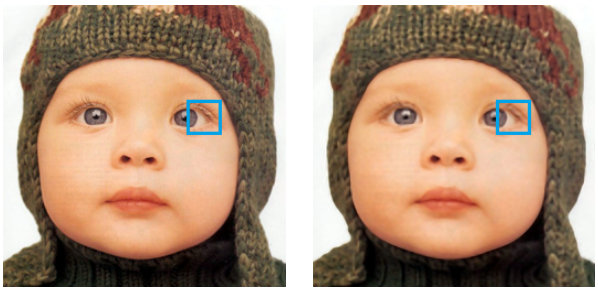
$$\tilde{\mathbf{h}} = \frac{\sum w_i \cdot \mathbf{h}_i}{\sum w_i}, \text{ where, } w_i = exp\left(\frac{-||\mathbf{l} - \mathbf{l}_i||_2^2}{2\sigma^2}\right). \tag{3}$$

We repeat the above procedure for every patch $\mathbf{l}$ of I_0, and get their corresponding HR patches, which constitute the HR image $\tilde{I}_1$.

IV. Analysis of Sub-band Energies

We argue that the patch similarity based SR algorithm described in Section III tends to smooth out fine textural details, due to the limitation of the L_2 distance in capturing textural similarity between patches. To quantify this loss of textural detail, we now perform a simple experiment. We use the *baby* image of Fig. 3(a) as our example. We denote I_1 to be the ground truth HR version of this image, as shown in Fig. 3(a). We compute its LR version I_0 by blurring and downsampling. We then use the SR algorithm described in Section III to super-resolve this LR image I_0 to obtain the image $\tilde{I}_1$ as shown in Fig. 3(b). We now examine the textural loss in $\tilde{I}_1$ when compared to the ground truth image I_1.

Let $\tilde{\mathbf{p}}$ and $\mathbf{p}$ represent corresponding patches from the super-resolved image $\tilde{I}_1$ and the ground truth image I_1, as illustrated by the blue box in Fig. 3. Let $\{\tilde{\mathbf{p}}^{(j)}\}_{j=1}^{R}$ and $\{\mathbf{p}^{(j)}\}_{j=1}^{R}$ denote the decomposition of these patches into R orientation sub-bands, as illustrated in Figs. 3(d) and 3(c). We use the steerable pyramid decomposition [9], [10] to obtain the orientation selective sub-bands. The steerable pyramid provides jointly-localized (space/frequency) representation of images using an invertible multi-scale, multi-orientation image decomposition [9], [10], as shown in Fig. 4. We use $R = 16$ orientations (and just a single scale) in our algorithm.



(a) Ground-truth I_1 (b) Super-resolved image $\tilde{I}_1$

(c) Patch $\mathbf{p}$ from I_1 and its sub-bands $\{\mathbf{p}^{(j)}\}_{j=1}^{R}$.

(d) Patch $\tilde{\mathbf{p}}$ from $\tilde{I}_1$ and its sub-bands $\{\tilde{\mathbf{p}}^{(j)}\}_{j=1}^{R}$.

Fig. 3. (a) Ground truth HR image I_1. (b) Image $\tilde{I}_1$ obtained after super-resolution using conventional approach of Section III. (c) An example patch from the ground truth image I_1, along with its decomposition into orientation selective sub-bands. (d) Similar decomposition for the corresponding patch from $\tilde{I}_1$. We analyze the loss of energy in the sub-bands of patches from $\tilde{I}_1$, compared to those from I_1. See text for details. The patches shown here are chosen large for illustration purpose.

Let $\tilde{e}^{(j)}$ and $e^{(j)}$ be the energies of the j^{th} sub-bands $\tilde{\mathbf{p}}^{(j)}$ and $\mathbf{p}^{(j)}$ respectively.

$$\tilde{e}^{(j)} = ||\tilde{\mathbf{p}}^{(j)}||_2^2, \text{ and } e^{(j)} = ||\mathbf{p}||_2^2 \tag{4}$$

We now sort the sub-band energies $\{\tilde{e}_i^{(j)}\}_{j=1}^{R}$ and $\{e_i^{(j)}\}_{j=1}^{R}$ according to decreasing values of $\tilde{e}_i^{(j)}$. The sorted set of energy values helps us observe the relative energy distribution between the macrostructure (high energy sub-bands) and the microstructures (low energy sub-bands) in the patch, irrespective of their orientations. The sorting helps us achieve this rotation invariance. Therefore, if an image patch recurs in the image in a rotated form, both the patches would yield the same sorted set of sub-band energy values.

We repeat the above procedure for all patch pairs $\tilde{\mathbf{p}}$ and $\mathbf{p}$ from the images $\tilde{I}_1$ and I_1, and obtain a sorted array of sub-band energy values for each patch. We then compute an average of these sorted arrays or sets, across all the patches. Fig. 5(a) shows this average set of sorted energy values, for patches from the super-resolved image $\tilde{I}_1$ (blue bars) and from the ground truth I_1 (red bars).

We make the following two interesting observations: 1) The energy in the high energy bands of the super-resolved image $\tilde{I}_1$ is much closer to those of the ground truth image I_1. This shows that the patch-similarity based SR algorithm using L_2 distances is able to preserve the macrostructures quite well.

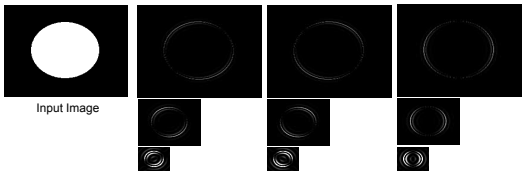
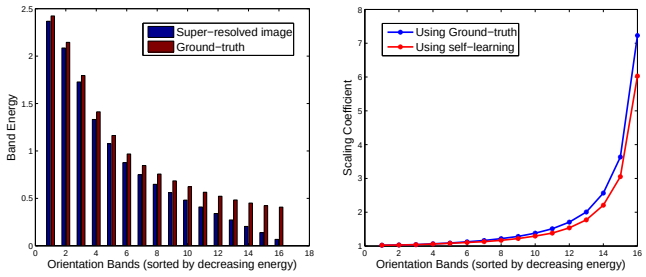


Fig. 4. An example showing the multi-orientation image decomposition yielded by the steerable pyramid [9], [10], on a synthetic image.



(a) (b)

Fig. 5. (a) Average sub-band energy value of all patches in the *Baby* image, sorted in decreasing order. Clearly, high energy sub-bands (macrostructures) are reasonably well recovered, whereas low energy sub-bands (microstructures) get severely attenuated using conventional patch-similarity based SR. (b) Blue plot shows the sub-band scaling coefficients obtained using the ground truth image I_1 (i.e. by comparing patches from I_1 and $\tilde{I}_1$). Red plot shows the coefficients obtained using the proposed self-learning scheme (i.e. by comparing patches from I_0 and $\tilde{I}_0$).

2) Relatively, the low energy sub-bands suffer from severe attenuation, confirming our hypothesis stated earlier that fine textures (microstructures) are much less preserved by such an SR algorithm.

Can we recover or compensate for this loss? Based on examining the bar plot of Fig. 5(a), a possible way to 'optimally' compensate for the sub-band attenuation is to amplify each sub-band $\tilde{\mathbf{p}}^{(j)}$ of the patch $\tilde{\mathbf{p}}$ by multiplying with scaling factors $\alpha^{(j)}$, where,

$$\alpha^{(j)} = \frac{e^{(j)}}{\tilde{e}^{(j)}}, \qquad j = 1, 2, ..., R. \tag{5}$$

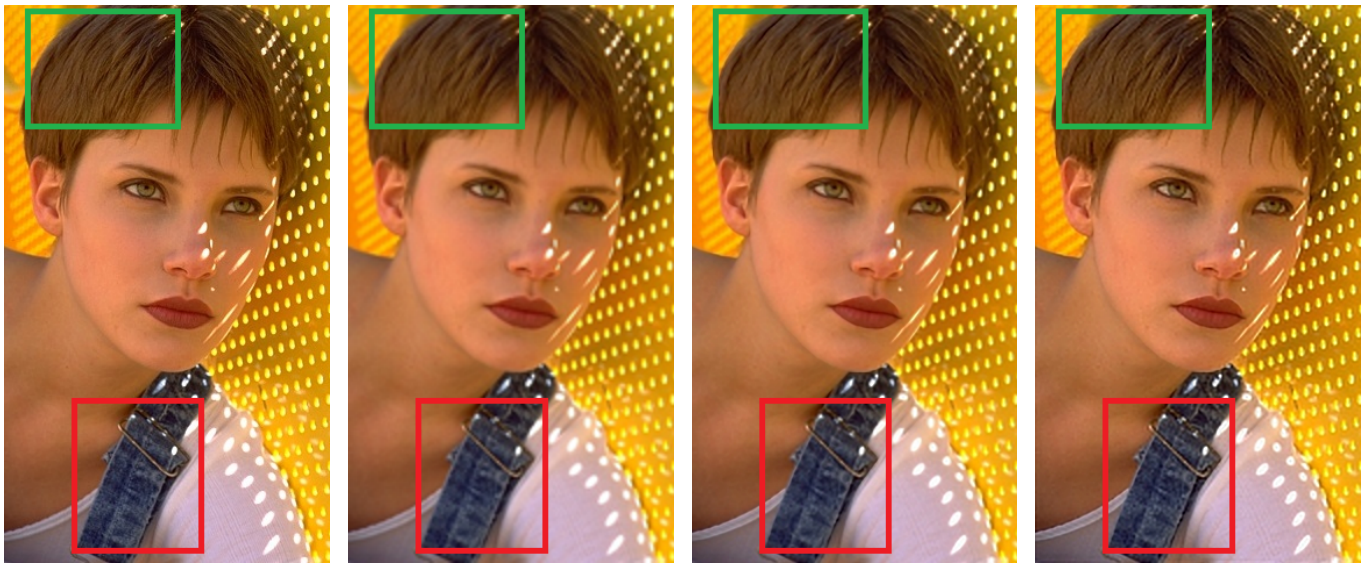
Using the coefficients $\alpha^{(j)}$, the sub-bands can be amplified such that their energies match those of the ground truth. The blue curve in Fig. 5(b) shows the values of these coefficients computed using Eq. (5) for the *baby* image. As expected, the lower energy sub-bands have higher scaling coefficients as they are more severely attenuated.

An obvious problem in using Eq. (5) is that the ground truth image I_1 is never available in any practical SR problem. Therefore, the sub-band energies $\{e^{(j)}\}_{j=1}^{R}$ of the ground truth image patches are never available, and the coefficients $\alpha^{(j)}$ of (5) cannot be determined. In the next section, we propose a method to learn these coefficients.

V. Self-Learning of Sub-band Constraints

Given an input image I_0, our analysis in the previous section showed that patches of the super-resolved image $\tilde{I}_1$, obtained using the conventional patch-similarity approach of Section III, suffer attenuation of the low energy sub-bands. We saw that the scaling coefficients α_j of Eq. (5) could compensate for this attenuation by appropriately boosting the sub-bands of each patch. However, computing these coefficients required knowledge of the ground truth HR image I_1, which is not available in practical scenarios.

A solution to the above problem is to estimate these coefficients from training patches extracted from natural images and treat these learned coefficients as a statistical prior. Such a prior would indicate the relative amplifications required for different sub-bands of the super-resolved image patch.



(a) Ground-truth HR (b) Bicubic (0.9167) (c) Conventional self-similarity (0.9360) (d) Ours (**0.9378**)

Fig. 6. *Sunlight* (2X): Best viewed when zoomed in. Our result shows much richer texture in the hair, facial features and the blue shoulder strap etc., as compared to the conventional patch-similarity based SR. Our result appears almost indistinguishable from the ground truth. Numbers in brackets denote SSIM [16] values.

(a) Ground-truth HR (b) Bicubic (0.6183) (c) Conventional self-similarity (0.6467) (d) Ours (**0.6546**)

Fig. 7. *Fur* (4X): The fur is reconstructed better in our result, and it appears sharper and richer in texture. Numbers in brackets denote SSIM [16] values.

In this paper, instead of resorting to an external database of image patches for learning such a prior, we propose a self-learning scheme, that operates as follows: We utilize scaled down versions of the given image I_0, to generate training data for learning the scaling coefficients. More specifically, we first obtain I_{-1} from I_0 by a blurring and downsampling operation. We then compute a super-resolved image $\tilde{I}_0$, by using the patch-similarity based SR algorithm of Section III with I_{-1} as the input image. The computation of $\tilde{I}_0$ is schematically illustrated in the blue dotted box of Fig. 2.

Our training image pair consists of the super-resolved image $\tilde{I}_0$, and its corresponding 'ground-truth' I_0, which is available to us. Our objective is now to learn the attenuation in the sub-bands of the patches of $\tilde{I}_0$, when compared to those from I_0. We extract around 1000 randomly sampled patches from $\tilde{I}_0$ along with their corresponding ground-truth patches from I_0. Using these two sets of patches, we repeat the analysis presented in Section IV to obtain the scaling coefficients $\alpha^{(j)}$ using Eq. (5).

The red plot in Fig. 5(b) shows the coefficients thus obtained using the proposed self-learning scheme (using $\tilde{I}_0$ and I_0) for the *Baby* image. We can see that these coefficients closely approximate the 'optimal' coefficients learnt with knowledge of the ground truth image I_1 (blue plot), as described in the previous section.

VI. Backprojection Constraint

Once the coefficients $\{\alpha^{(j)}\}_{j=1}^{R}$ have been determined, we use it to amplify or boost the respective sub-bands of each patch from the image $\tilde{I}_1$, using Eq. (1). The enhanced patches thus obtained form the super-resolved image $\hat{I}_1$ that we set out to achieve. However, we must also ensure that the image $\hat{I}_1$ on blurring and downsampling, yields the LR image I_0. We therefore need to minimize the cost function,

$$J(\hat{I}_1) = || \left(\hat{I}_1 * f_{psf} \right) \downarrow - I_0||_2^2 \tag{6}$$

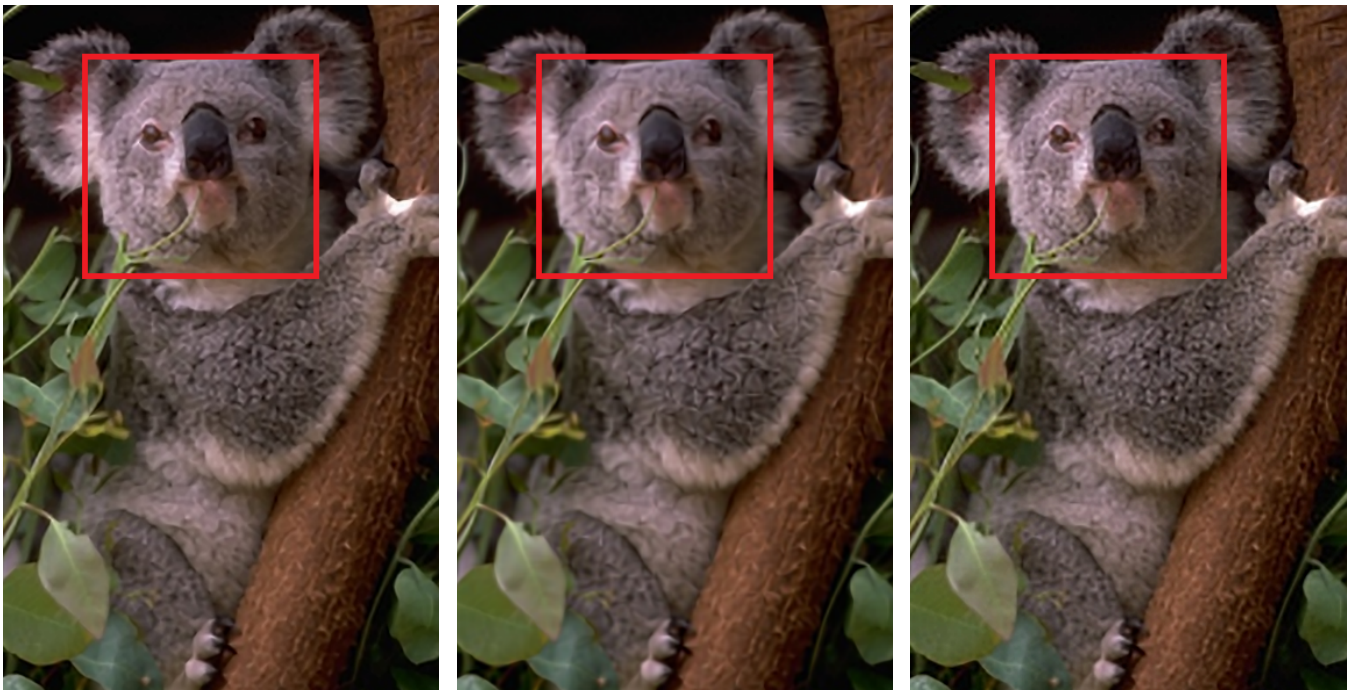
To satisfy this constraint, we run around 10 iterations of the following gradient based update rule,

$$\hat{I}_1^+ = \hat{I}_1 - \mu \nabla J(\hat{I}_1) \tag{7}$$

where we choose the stepsize $\mu = 1$. The above procedure is called the iterative backprojection algorithm [7].

VII. Results

Implementation Details. We use the proposed algorithm for upscaling images with a relatively small scaling factor, not exceeding $s = 2$. Therefore, for super-resolving images to $4X$ resolution, we apply the proposed algorithm twice, each time with scaling factor $s = 2$. Similarly, for an overall super-resolution of $3X$, we apply our algorithm twice with scaling factor $s = \sqrt{3}$ each time. For super-resolving color images, we use our algorithm only on the luminance channel. The chroma channels are upscaled using simpler methods such as bicubic interpolation, and then recombined to obtain the color image.



(a) Glasner *et. al.* [6] (b) Freedman *et. al.* [4] (c) Ours

Fig. 8. *Koala* (3X): Best viewed if zoomed in. Textural details of the fur and the tree trunk are better recovered in our result as compared to competing methods.

(a) Glasner *et. al.* [6] (b) Freedman *et. al.* [4] (c) Ours

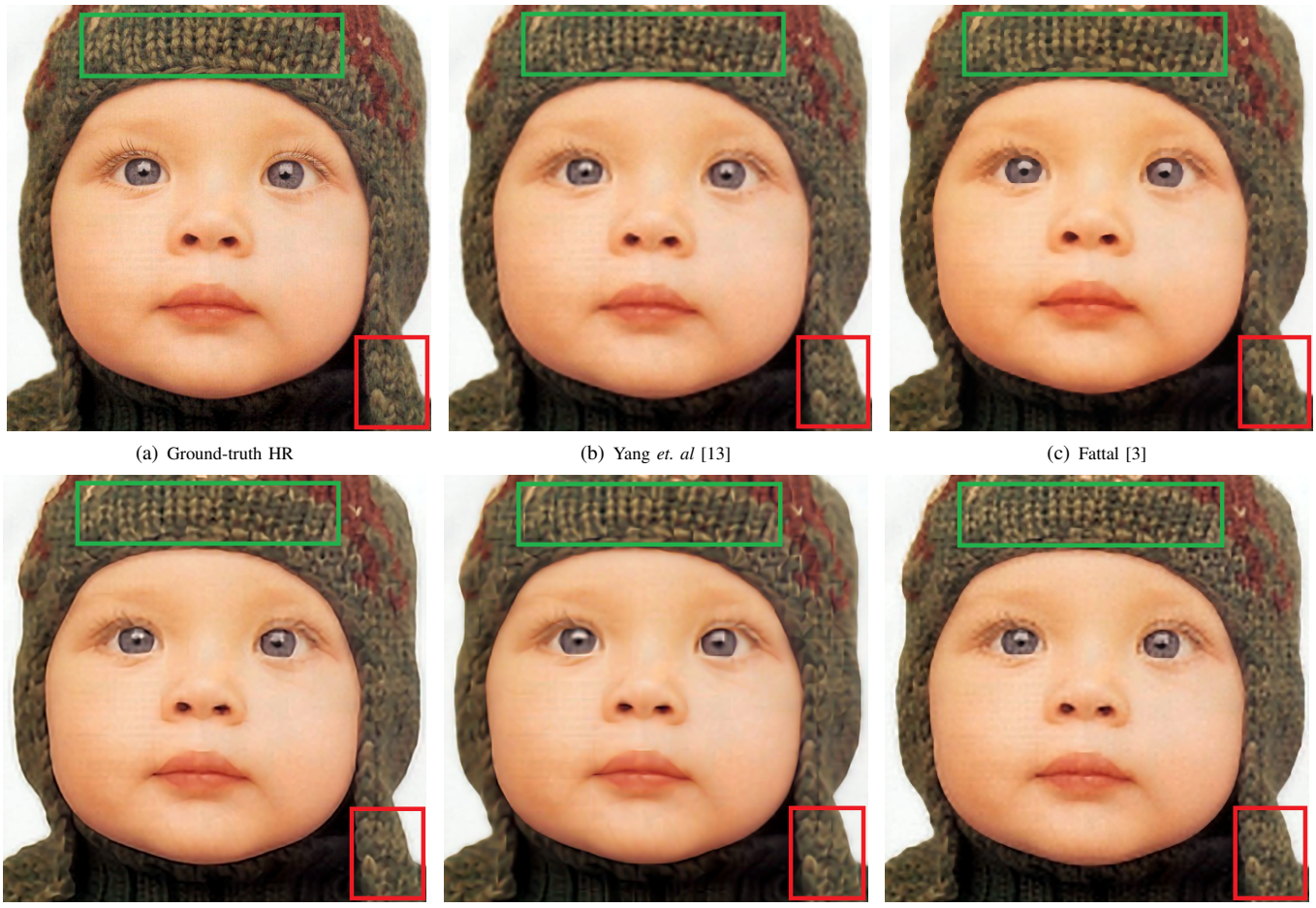
Fig. 9. *Girl* (3X): Textural details of the hair are more enhanced in our result. The freckles on the face also are clearer if seen while zoomed in.

We first run our algorithm on images that have known ground truth HR versions. We compare our approach with the conventional patch similarity based method as described in Section III and see the improvement in results our algorithm brings. Fig. 6 shows our result on the *Sunlight* image. Clearly, our result shows much richer texture in the hair, facial features, the blue shoulder strap etc. Visually, our result appears almost indistinguishable from the ground truth in this example. We report the structural similarity measure (SSIM) [16] below each result, although the correlation of numerical metrics with human perception of image quality is debatable.

Fig. 7 shows our result on the *Fur* image. In this case as well, our result looks visually more appealing and bears closer visual resemblance to the ground truth.

We now compare our results to those obtained in the past work. Specifically, we compare our results to those of two state-of-the-art methods, the self-similarity based methods of Glasner *et. al.* [6] and Freedman *et. al* [4], taken from the respective authors' websites. Fig. 8 shows the results on the *Koala* image. We can see that our result better shows the fine details in the animal fur and the tree trunk than the other two methods. Fig. 9 shows another set of results on the *Girl* image, where fine details of the hair are more clearly visible in our result. These images do not have ground truth HR available.

Finally, in Fig. 10, we also compare against two more methods, that are based on learning from external databases - the dictionary learning based method of Yang *et. al* [13] and the edge statistics based method of Fattal [3]. Yang *et. al.* [13] is not able to produce sufficiently sharp edges (e.g. the lips).



(a) Ground-truth HR (b) Yang *et. al* [13] (c) Fattal [3]

(d) Glasner *et. al.* [6] (e) Freedman *et. al.* [4] (f) Ours

Fig. 10. *Baby* (4X): Textures in the woolen cap appear richer in our result as compared to other methods. Our edges are also sharp. Details in the eyes are slightly better.

The textures produced by the edge based method of Fattal [3] tend to appear un-natural, e.g., as in the green box. Our result appears richer in texture, and also yields sharper edges.

VIII. Conclusion

We have presented an SR algorithm that delivers better super-resolved texture. Our algorithm is based on an observation we have made, that the conventional L_2 distance based patch-matching does not sufficiently characterize fine textures. Additional criteria are needed to ensure that the subtle textural elements are super-resolved better. To take advantage of oriented bandpass filters in characterizing textures, we have presented an algorithm that additionally constrains the energies of the sub-bands of the super-resolved patches. We have proposed a self-learning scheme that determines an optimal set of scaling coefficients, to balance the energies in the sub-bands to mimic their distributions in the natural images. Our algorithm does not use any external training database.

IX. Acknowledgements

This work was supported by US Office of Naval Research grant N00014-12-0259. Abhishek Singh was also supported by the Joan & Lalit Bahl Fellowship and the Computational Science & Engineering Fellowship at the University of Illinois.

References

[1] M. Barnsley. *Fractals Everywhere*. Academic Press Professional, Inc., 1988.
[2] M. Ebrahimi and E. Vrscay. Solving the inverse problem of image zooming using "self-examples". In *International Conference on Image Analysis and Recognition*, 2007.
[3] R. Fattal. Image upsampling via imposed edge statistics. *ACM Trans. Graph.*, 2007.
[4] G. Freedman and R. Fattal. Image and video upscaling from local self-examples. *ACM Trans. Graph.*, 2010.
[5] W. T. Freeman and E. C. Pasztor. Learning low-level vision. *IJCV*, 2000.
[6] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *ICCV*, 2009.
[7] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP*, 1991.
[8] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *IJCV*, 2000.
[9] E. Simoncelli and W. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *ICIP*, 1995.
[10] E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger. Shiftable multiscale transforms. *IEEE Trans. Info. Theory*, 1992.
[11] A. Singh, F. Porikli, and N. Ahuja. Super-resolving noisy images. In *CVPR*, 2014.
[12] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum. Gradient profile prior and its applications in image super-resolution and enhancement. *IEEE Trans. Image Proc.*, 2011.
[13] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE Trans. Image Proc.*, 2010.
[14] S. C. Zhu, Y. Wu, and D. Mumford. Filters, random fields and maximum entropy: Towards a unified theory for texture modeling. *IJCV*, 1998.
[15] M. Zontak and M. Irani. Internal statistics of a single natural image. In *CVPR*, 2011.
[16] Z.Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Proc.*, 2004.