

 Open access • Proceedings Article • DOI:10.1109/CVPR.2010.5540206

## **Context-constrained hallucination for image super-resolution** — [Source link](#)

Jian Sun, Jiejie Zhu, Marshall F. Tappen

**Institutions:** Xi'an Jiaotong University, University of Central Florida

**Published on:** 13 Jun 2010 - Computer Vision and Pattern Recognition

**Topics:** Image texture, Image segmentation, Image resolution and Pixel

Related papers:

- [Image Super-Resolution Via Sparse Representation](#)
- [Super-resolution from a single image](#)
- [Example-based super-resolution](#)
- [Super-resolution through neighbor embedding](#)
- [Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/context-constrained-hallucination-for-image-super-resolution-3sz7wzknbv>

# Context-Constrained Hallucination for Image Super-Resolution

Jian Sun  
Xi'an Jiaotong University  
Xi'an, P. R. China  
jiansun@mail.xjtu.edu.cn

Jiejie Zhu  
EECS, University of Central Florida  
Orlando, FL, USA  
{jjzhu,mtappen}@eecs.ucf.edu

Marshall F. Tappen

EECS, University of Central Florida

Orlando, FL, USA

{jjzhu,mtappen}@eecs.ucf.edu

## Abstract

*This paper proposes a context-constrained hallucination approach for image super-resolution. Through building a training set of high-resolution/low-resolution image segment pairs, the high-resolution pixel is hallucinated from its texturally similar segments which are retrieved from the training set by texture similarity. Given the discrete hallucinated examples, a continuous energy function is designed to enforce the fidelity of high-resolution image to low-resolution input and the constraints imposed by the hallucinated examples and the edge smoothness prior. The reconstructed high-resolution image is sharp with minimal artifacts both along the edges and in the textural regions.*

## 1. Introduction

The goal of single image super-resolution system is to estimate a sharp high-resolution image from low-resolution input. The large body of work on this problem can be roughly divided into two general categories:

- Edge-focused methods that focus on sharpening edges, while staying faithful to the low-resolution input [20, 15, 16, 5, 23, 7, 24].
- Patch-based hallucination methods that seek to create a high-resolution image by learning from patches of example images [10, 9, 25, 4, 27, 1, 17, 29, 33, 34].

Both approaches have different advantages and weaknesses. Edge-focused methods are able to produce sharp edges with minimal jaggy or ringing artifacts. However, they cannot introduce novel high frequency details which do not appear in the low-resolution image. Patch-based hallucination methods can synthesize new details by constructing the high-resolution estimate from example images, but the mathematical models used to construct the image from patches make it difficult to control artifacts introduced into the estimate.

In this paper, we introduce a single-image super-resolution method that makes two significant advances:

1. Our approach integrates edge-focused and patch-based hallucination methods into a single, easily-optimized energy-function framework. This framework makes it possible to leverage edge-sharpness cues, while using patches to hallucinate texture.
2. We present an improved approach for finding the patches that are used to construct the high-resolution image. While previous methods have searched for high-resolution patches that can be down-sampled to match the input, we formulate a patch-selection process by analyzing the textural characteristics of a region surrounding each pixel to find better candidate image patches for the high-resolution estimate.

The new patch-selection process is designed to utilize the texture appearance around image pixel, which is called the textural context, to search a database of high-resolution image segments for segments with similar low-resolution textural properties. These texturally-similar segments are then used to hallucinate the candidate examples for each pixel of high-resolution image. The texturally-similar segments help to introduce reasonable high-frequency details into the high-resolution result. The texturally-constrained approach can be thought of as bridging the gap between image hallucination and texture synthesis [6, 14] by constraining the hallucinated patches to match the low-resolution input.

Given the discrete hallucinated candidate examples, we cast the image super-resolution problem into a continuous energy minimization framework. The energy function enforces the fidelity of high-resolution image to the input low-resolution image, the fidelity of pixel to their discrete candidate examples, and the edge smoothness. Compared with the previous methods, our method not only produces sharp edges but also hallucinates results with reasonable textures.

The rest of this paper is organized as follows. Related works are discussed in Section 2. In Section 3, we introduce how to perform context-constrained hallucination. Section 4 presents how to utilize the hallucinated examples to design a continuous energy model for super-resolution. Comparisons and Experiments are performed in Section 5. Finally, this paper is concluded in Section 6.

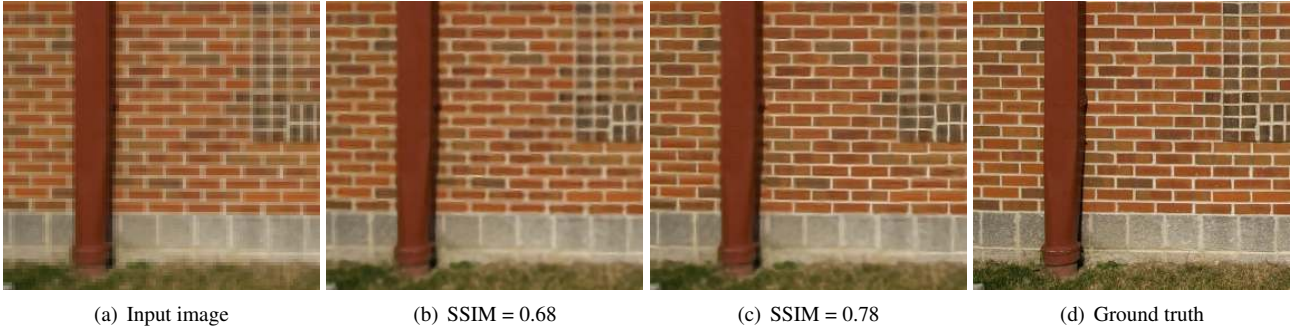


Figure 1. Comparison of super-resolution results ( $\times 3$ ). (b) Without context constraint, examples are searched from randomly sampled pairs of high-resolution/low-resolution patches and the wall textures are not well hallucinated. (c) With context constraint, the wall textures are successfully hallucinated by learning examples from the texturally similar segments. The super-resolution model in Section 4 is used to produce both of the results. Compared with the ground-truth image, our result achieves significantly higher SSIM value.

## 2. Related work

Most of the edge-focused methods try to impose some prior knowledge on the edges. Dai [5] propose a soft edge smoothness prior and apply it to the alpha-channel of salient edges. In [7, 24], the statistical relationship between high-resolution and low-resolution edge features is studied and imposed on high-resolution image. More edge-focused methods can be found in [15, 20, 23, 26]. These methods have produced excellent results with sharp edges. However, they can hardly introduce novel high-frequency details in the textural regions.

The hallucination method, which is also called learning based method, advances the edge-focused method by learning high-resolution examples from a training database of high-resolution/low-resolution patch pairs. Then the examples are synthesized to produce the high-resolution image using Markov network [10, 27, 25] or combined to extend the generalization ability of examples [4, 33, 28]. Due to the difficulty in texture hallucination, Sun et al. [25] propose to only hallucinate the primal sketches (e.g. edges, ridges and corners). Xiong et al. [32] then propose to enhance the image features before hallucination. To hallucinate correct textures, the domain specific texture priors are proposed [21, 13], in which the within-class texture example should be given. Recently, Glasner et al. [11] combine the hallucination method and multi-image super-resolution method in an unified framework. It learns the patch examples from the single input low-resolution image and produce state-of-the-art results. However, it might not produce the desired high-frequency details when no high resolution examples exist in the input image. These above hallucination methods have achieved excellent results in primal sketches or textural regions. However, it is still hard to hallucinate both the edges and textures automatically and simultaneously.

Unlike the previous work, the proposed method in this paper takes advantage of the edge-focused method and hallucination method in an unified energy minimization framework and performs well both in textural regions and along

edges simultaneously.

## 3. Context-constrained hallucination

The first step in estimating the high-resolution image is the selection of high-resolution examples that will be used to create the high-resolution estimate. The goal is to produce a high-resolution image with rich and natural texture, therefore the examples themselves must have the correct texture.

We refer to hallucination under the constraint of textural context as *context-constrained hallucination*. The textural context of a pixel is defined as the textural appearance of the segment surrounding the pixel. In this method, the pixels in the high-resolution result are constrained by the overall textural context of the corresponding pixels in low-resolution input. This type of hallucination cannot be achieved by directly finding the examples from the training set of high-resolution/low-resolution patch pairs as in the traditional learning-based image hallucination [10, 25]. This is because just comparing the patch example with the low-resolution observed patch makes it impossible to judge if the texture in the high-resolution example, but not visible in the corresponding low-resolution example, is correct.

We implement context-constrained image hallucination by dividing all of the images in the training database into segments with roughly uniform texture. The textural properties of example are controlled by drawing from the segments with textural properties that match the textural context of the low-resolution pixel for which the example is being generated.

Figure 1 presents an example for comparison between the results with and without the textural context constraint respectively. Using the same super-resolution model that will be discussed in Section 4, the wall textures are successfully hallucinated in Figure 1(c) by learning examples from the texturally similar segments. However, the traditional hallucination methods, which directly find examples from training set of high-resolution/low-resolution patch pairs,

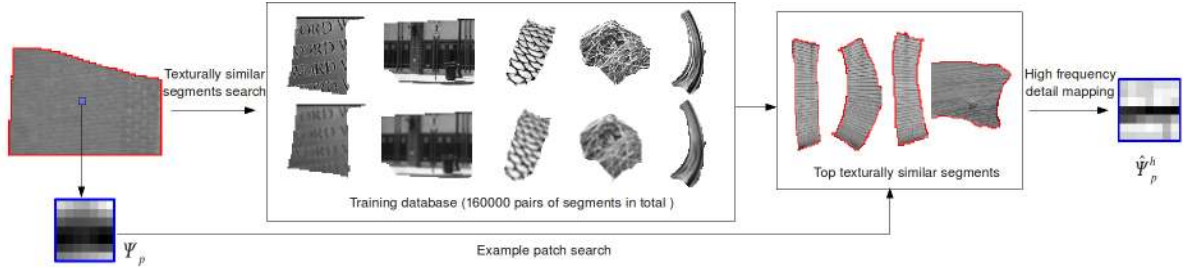


Figure 2. The overview of context-constrained patch example search. The training set is composed of high-resolution/low-resolution segment pairs. First, the input low-resolution image is up-sampled and segmented into texturally consistent segments [8, 19]. For each pixel  $p$ , we utilize the segment surrounding the pixel to search ten texturally similar segment pairs from the training set by matching the texture appearance in the low resolution. Then the texturally similar segments are searched to find the best example for the patch  $\Psi_p$  surrounding  $p$ . Finally, the high-resolution example patch  $\hat{\Psi}_p^h$  is hallucinated by high-frequency detail mapping from the example to the observed patch (patches are normalized for better illustration).

similar to [10, 9], fail to hallucinate the correct textures. The key to success in this approach is the ability to find image segments in the training database that match the textual context of each pixel in the low-resolution observation.

### 3.1. Overview of example generation

The example generation process is structured so that one  $m \times m$  high-resolution example patch is chosen for each pixel in an up-sampled version of the low-resolution input. Excluding the borders, this leads to  $m^2$  candidate examples for each pixel in the estimated high-resolution image. These candidate examples for each pixel will be utilized to constrain the super-resolution model discussed in Section 4.

An example patch is found for a pixel  $p$  by two steps. First, we search the database to find the top ten low-resolution segments that are texturally similar to the image segment that pixel  $p$  lies in. Then these ten segments are searched to find the patch that best matches the patch centered at  $p$ . For efficiency reasons, this search is limited to 100,000 randomly sampled locations within the ten segments that are chosen for their textural similarity. Once the best matching low-resolution patch is found, the high-resolution example patch is extracted from the corresponding high-resolution segments in the database.

Please refer to Figure 2 for the procedures of patch example generation. The following paragraphs of Section 3 will discuss the components in detail.

### 3.2. Training database

The training database is constructed from 4000 natural images of indoor and outdoor scenes, drawn from the internet photography forums (e.g. Flickr space) and LabelMe database<sup>1</sup>. We down-sample the images and then up-sample the derived low-resolution images by bicubic interpolation to construct a set of high-resolution/low-resolution (i.e. the up-sampled low-resolution) image pairs. Each pair is seg-

mented into roughly 40 segments using the Berkeley Segmentation Engine<sup>2</sup> (BSE) [8, 19], which produces segments that are generally consistent texturally. In this way, we construct a large training database with  $1.6 \times 10^5$  pairs of high-resolution/low-resolution segments. We denote the training database as  $\{\Omega_f^h, \Omega_f^l\}$ , in which  $f$  is the down-sampling factor.

### 3.3. Texturally similar segments search

We now discuss the definition of textural context and the method to find texturally similar segments from the training database.

Visual context is a general term, and it has different definitions in object recognition [2], motion estimation [31], etc. In this application, the *textural context* of a pixel is defined as the textural appearance of segment surrounding it. Inspired by biological vision research on texture perception [3], we define the texture appearance of an image segment by the distributions of its responses to the derivative filter banks with 6 orientations and 3 scales. Each distribution of responses to a derivative filter at scale  $s$  and orientation  $o$  is quantized by a histogram with 20 bins  $H_{\{s,o\}} = \{h_{\{s,o,b\}}\}_{b=1}^{20}$ . Then the texture similarity between two segments  $S^1$  and  $S^2$  is measured by the distance between the distributions across different scales and orientations:

$$\chi^2(S^1, S^2) = \sum_s \sum_o \sum_b \frac{(h_{\{s,o,b\}}^1 - h_{\{s,o,b\}}^2)^2}{h_{\{s,o,b\}}^1 + h_{\{s,o,b\}}^2}, \quad (1)$$

where the  $\chi^2$  distance is the first-order approximation of the Kullback-Leibler (KL) divergence distance [18].

For each segment  $S_k$  in the up-sampled low-resolution image, we search the top 10 most similar low-resolution segments in the training database, then their corresponding high-resolution segments provide the textural context for the hallucination of pixels in segment  $S_k$ . Figure 3 shows three results of segment search. As we can observe, the

<sup>1</sup><http://labelme.csail.mit.edu>

<sup>2</sup><http://www.cs.berkeley.edu/~fowlkes/BSE/>



Figure 3. Texturally similar segments search (segments are enclosed in the red curves). The first column shows the upsampled low-resolution segments. Two of the top similar segments in high resolution are shown in the second to the third columns.

most similar segments have the texture appearance similar to the query segment. Instead of providing the high-level context information coming from the within-class objects, the searched segments provide a middle-level textural context for the pixels in the query segment.

### 3.4. High frequency detail mapping

Once the segments with the most similar textural appearance have been found, the next step is to pick the actual example patch. To extend the generalization ability of patch representation, we normalize the low-resolution patch, both in the query image and training set, by its mean and variance as in [25]. Patches are then compared using their Mean Squared Error (MSE). For the observed patch  $\Psi_p$  centered at  $p$ , denote the best sampled high-resolution/low-resolution patch examples as  $\{\Psi^l, \Psi^h\}$ , then the actual high-resolution example is produced by modifying the mean and norm of  $\Psi^h$  to match the observed patch  $\Psi_p$ :

$$\hat{\Psi}_p^h = \frac{(\Psi^h - \mu^l)}{\sigma^l} \sigma_p^l + \mu_p^l, \quad (2)$$

where  $\{\mu^l, \sigma^l\}$  and  $\{\mu_p^l, \sigma_p^l\}$  are the means and standard deviations of patches  $\Psi^l$  and  $\Psi_p$ .

## 4. Image super-resolution model

In this section, we apply the hallucinated candidate examples to image super-resolution. Instead of modeling super-resolution as a discrete-valued Markov network, as in [10, 9, 27, 25], to choose among the candidate examples, we cast the super-resolution problem into a continuous energy minimization problem, which is natural to combine the high-resolution image reconstruction constraint and natural image sparseness constraint [27].

Given the low-resolution image  $I^l$ , in order to reconstruct the high-resolution image  $I^h$ , we minimize the fol-

lowing energy function:

$$E(I^h) = E(I^h|I^l) + \beta_1 E_h(I^h) + \beta_2 E_e(I^h) \quad (3)$$

The first energy term is the **high-resolution image reconstruction term**:  $E(I^h|I^l) = \|(I^h * G) \downarrow - I^l\|^2$ , which forces the down-sampled version (denoted by operator  $\downarrow$ ) of the high-resolution image to be close to the input low-resolution image  $I^l$ . The down-sampling of the high-resolution image is approximated through the convolution with a Gaussian kernel (with bandwidth 0.8, 1.2, 1.5 for down-sampling factor of 2, 3, 4 in our implementation) and then down-sampling by nearest neighbor interpolation as in [5, 24].

The second and third energy term will be introduced in the following subsections.

### 4.1. Hallucination term

The energy term  $E_h(I^h)$  forces the value of pixel  $p$  in high-resolution image  $I^h$  to be close to the hallucinated candidate examples learned in Section 3.

For each pixel  $p$ , the patch examples  $\hat{\Psi}_q^h$  of size  $m \times m$  (in Equation (2)) centered at pixel  $q$  in  $m \times m$  neighborhood of  $p$  provide  $m^2$  candidate examples  $\{c_i\}_{i=1}^{m^2}$  for pixel  $p$ . Since the example pixels are discrete values, we should develop a continuous energy term with local minima located at these candidate examples. This can be modeled as:

$$E_h(I^h) = \sum_p \min_i \{(I^h(p) - c_i(p))^2\} \quad (4)$$

which achieves the minimum value 0 when  $I^h(p)$  equals one of the examples  $c_i(p)$  for each pixel  $p$ .

However, this energy term is not differentiable and introduces the difficulty in energy minimization. Therefore, we propose to approximate Equation (4) by a continuous energy function:

$$E_h(I^h) = -\frac{1}{\lambda} \sum_p \log \left( \sum_{i=1}^{m^2} \exp(-\lambda d_i(p)) \right) \quad (5)$$

where  $d_i(p) = (I^h(p) - c_i(p))^2$ .

This term is effectively a smooth approximation to the minimum operator in Equation (4), where

$$\min\{d_1, d_2, \dots, d_{m^2}\} \approx -\frac{1}{\lambda} \log \sum_{i=1}^{m^2} \exp(-\lambda d_i). \quad (6)$$

This is reasonable because the exponential of the minimum among  $\{\lambda d_i\}$  will dominate the summation in the right formula, then the logarithm of the summation will approximate the minimum among  $\{d_i\}$ . The  $\lambda$  is set to scale the differences among  $\{d_i\}$ , which is set to 10 in order to make the minimum  $d_i$  sufficiently dominate the summation.

The energy term in Equation (5) is called the **hallucination term**. It forces  $I^h(p)$  to be close to one of the candidate examples when the local minimum is achieved.



Figure 4. Comparison of super-resolution results ( $\times 4$ ). (b) shows the result of bicubic interpolation, and it is blurry both in textural regions and along edges. Combining the edge smoothness term and reconstruction term, the result shown in (c) is still not sharp in the hat region, and artifacts along edges are introduced by the reconstruction term. Combining the hallucination term and reconstruction term, the result shown in (d) is sharp in the hat region. However, there are still artifacts along the salient edges. The model combining reconstruction term, hallucination term and edge smoothness term produces sharp edges with minimum artifacts and reasonable textures as shown in (e).

## 4.2. Edge smoothness term

The third energy term  $E_e(I^h)$  is the edge smoothness term which forces the edges of the high-resolution image to be sharp, with minimal ringing or jaggy artifacts. This energy term is designed to compensate the possible artifacts along the salient edges introduced when optimizing the hallucination term or reconstruction term.

We regularize the high-resolution image edges by natural image statistics. The student’s t-distribution [22] is used to model the heavy-tailed distribution of image responses to the first-order derivative filters  $[-0.25, 0.25]$  in four directions (i.e.  $0^\circ, 45^\circ, 90^\circ$  and  $135^\circ$ ). By performing the negative logarithm of this sparseness prior, it imposes an *edge smoothness term* on the high-resolution image:

$$E_e(I^h) = - \sum_p \{p_b(p) \cdot \sum_{k=1}^4 \alpha_k \log[1 + \frac{1}{2}(f_k * I^h(p))^2]\},$$

where  $f_k$  is the  $k$ -th filter and  $p_b$  measures the probability of boundary computed by color gradient and texture gradient [19]. Due to the multi-orientation nature of natural images, the distribution parameters  $\alpha_k$  ( $k = 1, 2, 3, 4$ ) can be reasonably assumed to be same, and then they will be incorporated into the coefficient  $\beta_2$  of the energy term in Equation (3).

The edge smoothness energy term mainly performs edge enhancement on the salient edges weighted by  $p_b$ , which is always aligned with the segment boundaries produced by the BSE [8, 19] algorithm when performing context-constrained hallucination. Due to the high repetition of image patches [11], the image pixels on the boundary of segments can be mostly hallucinated by the context-constrained hallucination. This edge smoothness term helps to produce clean and sharp edges by suppressing the possible artifacts along the salient edges.

## 4.3. Optimization

The energy function in Equation (3) for super-resolution can be optimized by gradient descent method:

$$I_{t+1}^h = I_t^h - \tau \nabla E(I_t^h), \quad (7)$$

where  $t$  is iteration index and  $\tau$  is the step size, which is set to 0.25. The gradient of the energy function is:

$$\nabla E(I^h) = ((I^h * G) \downarrow - I^l) \uparrow * G + \beta_1 \nabla E_h(I^h) + \beta_2 \nabla E_e(I^h), \quad (8)$$

where  $\nabla E_h(I^h) = \sum_p \sum_{i=1}^{m^2} \frac{\exp(-\lambda d_i(p))}{\sum_k \exp(-\lambda d_k(p))} (I^h(p) - c_i(p))$ , and  $\nabla E_e(I^h) = -p_b \sum_{k=1}^4 f_k^- * \frac{(f_k * I^h)}{1 + 1/2(f_k * I^h)^2}$ .

When optimizing the energy function, we initialize the high-resolution pixel  $I^h(p)$  by the median of hallucinated examples. The parameters  $\beta_1$  and  $\beta_2$  are set to 1 and 2 respectively by experimental justification, and the size of patch  $m$  is set to 7, which produce all of the results in this paper.

## 5. Experiments

We test our approach on a variety of natural images with rich textures and edges. For color images, we only perform image super-resolution on the illuminance channel in YUV color space, and the color channels are up-sampled by the bicubic interpolation.

To show the quality of results produced by the proposed super-resolution model, Figure 4 compares the results of models by combining the reconstruction term, edge smoothness term and hallucination term in different ways. Figure 4 (c) is the result of the model combining construction term and edge smoothness term, in which the strength of edge smoothness term (i.e.  $\beta_2$ ) is set to 2 in order to better suppress artifacts while preserving textures. The result

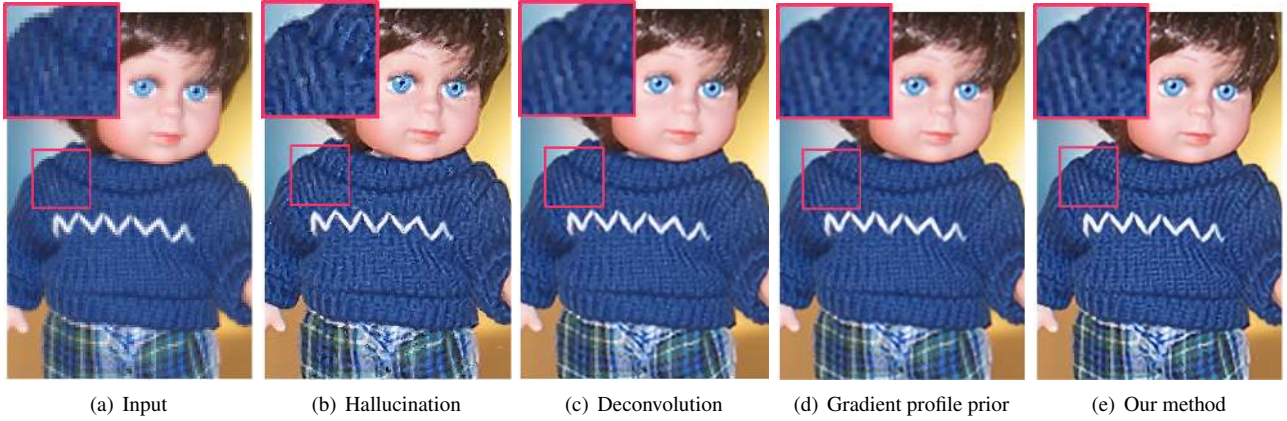


Figure 5. Comparison ( $\times 4$ ) with hallucination [10], deconvolution [23], gradient profile prior [24] methods. Our method produces sharp edges and reasonable textures in the result.

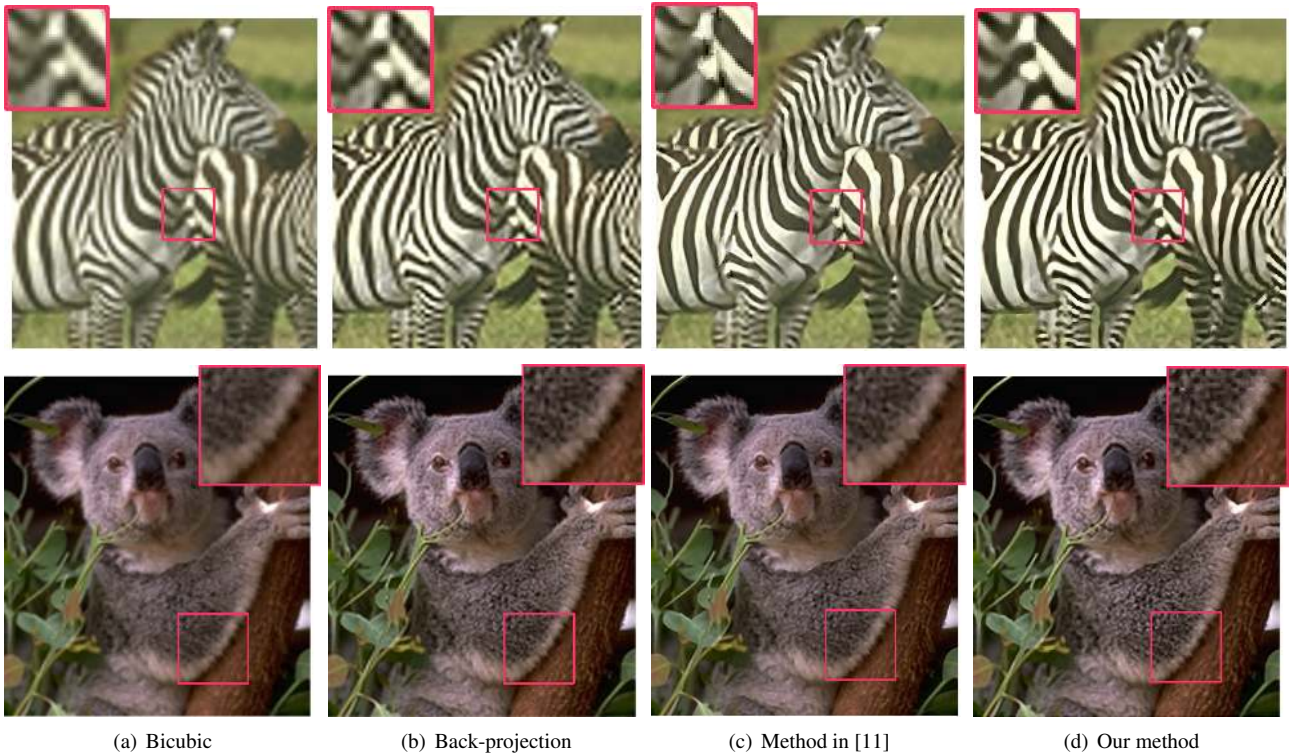
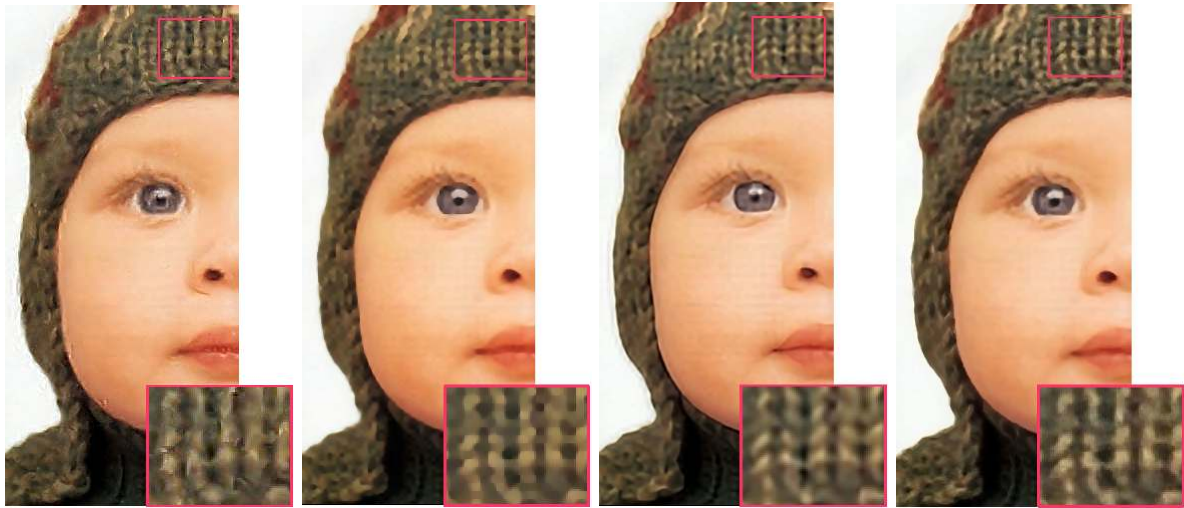


Figure 6. Comparison ( $\times 3$ ) on zebra and koala images. The results of back-projection [12] have ringing and jaggy artifacts along edges. The method in [11] produces excellent results. However, the stripes of horse and fur of koala shown in rectangles are not perfect. Our method produces both sharp stripes with little artifacts on the body of horse, and sharp textures on the body of koala.

is still not sharp in the hat textural region, and there are jaggy artifacts along the edges. Figure 4 (d) shows the result of model combining hallucination term and reconstruction term. In this result, the hat texture is sharp. However, there are still jaggy artifacts along the edges as shown in the rectangle. The model combining reconstruction term, hallucination term and edge smoothness term produces the best result with sharp edges and reasonable textures as shown in Figure 4 (e).

In Figure 5, we compare our method with the edge-

focused approach, including deconvolution [23] and gradient profile prior [24] methods, and the traditional hallucination method [10]. The result of hallucination method in Figure 5 (b) is sharp in the textures. However, unpleasant artifacts are also introduced. The deconvolution-based method uses the sparseness prior to regularize the high-resolution image, which suppresses the high frequency details in the texture region as shown in Figure 5 (c). The result of gradient profile prior method shown in Figure 5 (d) is sharp along salient edges. However, the sweater texture is blurry.



(a) SSIM = 0.871

(b) SSIM = 0.903

(c) SSIM = 0.914

(d) SSIM = 0.919

Figure 7. Comparison ( $\times 4$ ) of results of (a) hallucination [10], (b) edge statistics [7], (c) method in [11] and (d) our method. Our method produces comparable sharp edges as (b) (c) and hallucinates reasonable textures shown in the rectangle. Our method also produces the highest SSIM score.

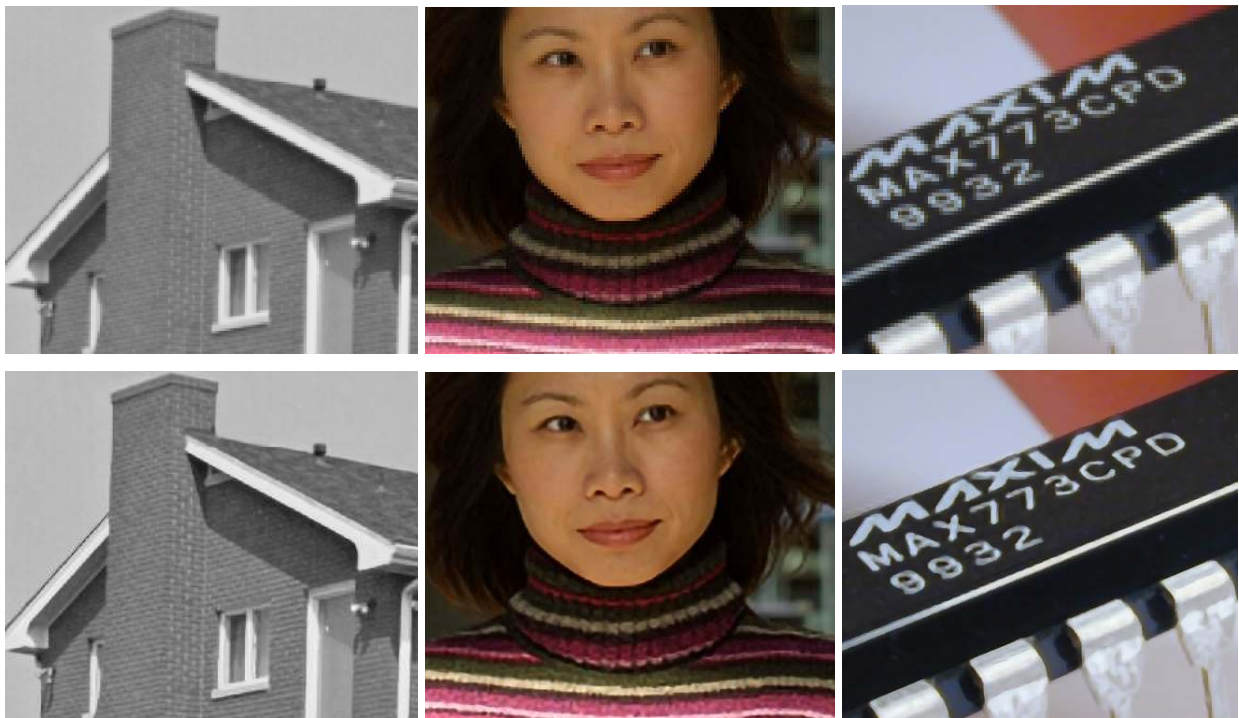


Figure 8. More results ( $\times 3$  in the left and middle examples and  $\times 4$  in the right example).

Our result in Figure 5 (e) is sharp both along edges and in the textural regions as shown in the rectangle.

In Figure 6, we compare our result with back-projection method [12] and the method in [11] on the images of zebra and koala. The results of back-projection shown in Figure 6 (b) have ringing and jaggy artifacts along edges. The method in [11] performs well through self-learning. However, the results are not sharp in the fur texture of koala and there are artifacts in the stripes of zebra as shown in the

rectangles. Our method produces both sharp stripes with little artifacts on the body of horse and reasonable textures on the body of koala through context-constrained hallucination and edge smoothness constraint.

In Figure 7, we apply our method to the test image in [7] and compare with the results of hallucination method [10], edge statistics method [7], and the method in [11]. Our method produces comparable sharp edges as Figure 7 (b) (c), and hallucinates sharp and reasonable textures in the



hat region enclosed by rectangle. We also quantitatively compare the Structural Similarity [30] (SSIM) index of results with the original high-resolution image, and our result produces the highest SSIM index.

Figure 8 shows more results which are magnified in factors of 3 and 4. All of these results show that our algorithm performs well both in textural regions and along edges.

## 6. Conclusion

This paper proposed a context-constrained hallucination approach by learning high-resolution examples from the texturally similar training segments. This hallucination approach helps to introduce reasonable high frequency details in results. Then an easily-optimized energy function combining the hallucinated examples, edge smoothness constraint and high-resolution image reconstruction constraint was proposed. We have shown that our method produces better textures and comparable sharp edges compared with the other state-of-the-art super-resolution methods.

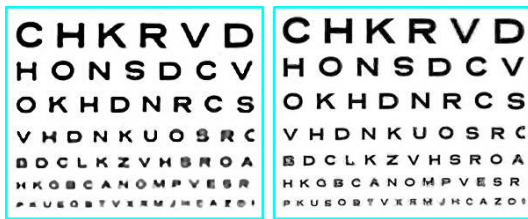


Figure 9. Left: our result. Right: result in [11].

In Figure 9, we compare with the method in [11] on an image with repeated characters across different scales. For this type of image with repeated patterns in different scales, the method in [11] is able to hallucinate the smaller characters guided by their high resolution examples in the same image. Our method also produces sharp and clear high-resolution result. However, it does not recover the smaller characters, as it is hard to recognize the same high-resolution characters from the training database by texture similarity. In the future, we are interested in defining high-level context of image pixels and learning examples by recognizing the textures from the same object, in order to further improve the hallucination term in our super-resolution framework.

## Acknowledgement

The authors thank the Central Florida Division of Applied Research Associates for helpful discussions and support. This work was partially supported by a grant from the NGA NURI program (HM1582-08-1-0021). This work was done when the first author worked as a postdoctoral research associate at University of Central Florida.

## References

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

[2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI*, 24, 2002.

[3] J. R. Bergen and E. H. Adelson. Early vision and texture perception. *Nature*, 333:363–364, 1988.

[4] H. Chang, D. Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *CVPR*, volume 1, pages 275–282, 2004.

[5] S. Y. Dai, M. Han, W. Xu, Y. Wu, and Y. H. Gong. Soft edge smoothness prior for alpha channel super resolution. In *CVPR*, 2007.

[6] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *ICCV*, 1999.

[7] R. Fattal. Image upsampling via imposed edge statistics. *ACM Transactions on Graphics*, 26(3):95:1–95:8, 2007.

[8] C. Fowlkes, D. Martin, and J. Malik. Learning affinity functions for image segmentation: Combining patch-based and gradient-based approaches. In *CVPR*, 2003.

[9] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 22(2):56–65, 2002.

[10] W. T. Freeman, E. Pasztor, and O. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, 2000.

[11] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *ICCV*, 2009.

[12] M. Irani and S. Peleg. Motion analysis for image enhancement: Resolution, occlusion and transparency. *Journal of Visual Communication and Image Representation*, 4(4):324–335, 1993.

[13] S. Lefebvre and H. Hoppe. Parallel controllable texture synthesis. *ACM Transactions on Graphics*, 24(3):777–786, 2005.

[14] S. Lefebvre and H. Hoppe. Appearance-space texture synthesis. *ACM Transactions on Graphics*, 25(3):541–548, 2006.

[15] X. Li and M. T. Orchard. New edge-directed interpolation. *IEEE Trans. on IP*, 10(10):1521–1527, 2001.

[16] Z. C. 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.

[17] C. Liu, H. Y. Shum, and W. T. Freeman. Face hallucination: Theory and practice. *International Journal of Computer Vision*, 75(1):115–134, 2007.

[18] X. Liu and D. Wang. Texture classification using spectral histograms. *IEEE Trans. on Image Processing*, 12(6):661–670, 2003.

[19] D. Martin, C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color and texture cues. *IEEE Trans. on PAMI*, 26(5):530–549, 1988.

[20] B. S. Morse and D. Schwartzwald. Image magnification using level-set reconstruction. In *CVPR*, volume 1, pages 333–340, 2001.

[21] L. C. Pickup, S. J. Roberts, and A. Zisserman. A sampled texture prior for image super-resolution. In *NIPS*, 2003.

[22] S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *CVPR*, 2005.

[23] Q. Shan, Z. Li, J. Jia, and C. K. Tang. Fast image/video upsampling. In *SIGGRAPH ASIA*, 2008.

[24] J. Sun, J. Sun, Z. B. Xu, and H. Y. Shum. Image super-resolution using gradient profile prior. In *CVPR*, 2008.

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

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

[27] M. F. Tappen, B. C. 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.

[28] J. Wang, S. Zhu, and Y. Gong. Resolution-invariant image representation and its applications. In *CVPR*, 2008.

[29] Q. Wang, X. Tang, and H. Y. Shum. Patch based blind image super resolution. In *ICCV*, 2005.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Quality assessment: from error measurement to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612, 2004.

[31] Y. Wu and J. Fan. Contextual flow. In *CVPR*, 2009.

[32] Z. Xiong, X. Sun, and F. Wu. Image hallucination with feature enhancement. In *CVPR*, 2009.

[33] J. C. Yang, J. Wright, Y. Ma, and T. Huang. Image super-resolution as sparse representation of raw image patches. In *CVPR*, 2008.

[34] W. Zhang and W. K. Cham. Learning-based face hallucination in dct domain. In *CVPR*, 2008.