

A Bayesian Approach to Alignment-Based Image Hallucination

Marshall F. Tappen¹ and Ce Liu²

¹ University of Central Florida
mtappen@eecs.ucf.edu

² Microsoft Research New England
celiu@microsoft.com

Abstract. In most image hallucination work, a strong assumption is held that images can be aligned to a template on which the prior of high-res images is formulated and learned. Realizing that one template can hardly generalize to all images of an object such as faces due to pose and viewpoint variation as well as occlusion, we propose an example-based prior distribution via dense image correspondences. We introduce a Bayesian formulation based on an image prior that can implement different effective behaviors based on the value of a single parameter. Using faces as examples, we show that our system outperforms the prior state of art.

1 Introduction

In this paper, we propose an alignment-based approach to hallucinating image detail for single-image super-resolution.

In [2], Baker and Kanade showed that smoothness constraints alone cannot produce the image content necessary to create a high-quality image. Since the introduction of this work, a number of different methods and models have been introduced for estimating the image information lost in the downsampling process. These systems differ in how they model the high resolution image.

Systems such as [1,3,4] have found success by focusing on modeling edge statistics. These systems produce a higher-resolution image that has sharp edges and is generally smooth in other regions. Given an input image where the important edges can be identified, such as the input image shown in Figure 1(a), these methods can produce a high-quality result with pleasing, sharp edges, such as the result shown in Figure 1(b), produced using the Gaussian Profile Prior from [1].

However, the performance of these methods will degrade quickly on low-resolution images. Given a reduced version of the same input, shown in Figure 1(c), the result produced by the Gaussian Profile Prior, shown in Figure 1(d), becomes significantly worse because the perceptually-important edges in the input image are no longer present.

A possible solution to this would be to synthesize the image content using examples of high-resolution images, such as [5,6,7,8], though this approach must

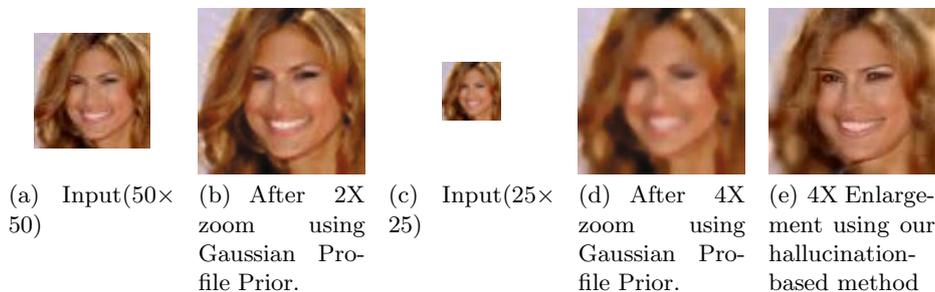


Fig. 1. Super-resolution systems perform their best when image content and key edges are visible in the low-resolution input. Because edges are clearly visible in the input in (a), a super-resolution approach like the Gaussian Profile Prior [1] is able to produce good output. However, as the result in (d) shows, as the resolution becomes lower, it is difficult for the system to enhance edges well enough. In cases such as this, image hallucination can provide the necessary information. In this paper, we present a novel image hallucination framework that can produce improved results, as shown in (e), on difficult, low-resolution imagery.

confront the difficult technical problem: *How should the example patches be chosen?* Systems have used a number of different approaches, including drawing examples on a patch-by-patch basis, as in [5,6], using self-similarity, as in [9], using user-input, as in [8], or recognizing textural-properties, as in [7,10].

While these systems can perform well on general images, both the original work in [2] and later work by Liu et al. in [11] have shown that knowing that the image contains a face makes it possible for a super-resolution system to perform much better because the system can leverage regularities in face appearance to hallucinate more detail than could be created from a general image model. The face-based super-resolution systems in [2] and [11] showed promising results, but both systems are limited to frontal facial images that can be warped to the template, and cannot handle large pose and viewpoint variations.

1.1 Paper Contributions

In this paper, we propose an **alignment-based** image hallucination approach to super-resolution. Our system leverages large databases of face images, recent advances in image matching, such as SIFT flow [12] and Patch Match [13], and a novel MAP estimation framework to make it possible to enhance facial images with a wide variety of poses and expressions. As shown in Figure 1(e), this approach is able to synthesize detail in areas where systems designed around edge statistics, like the Gaussian Profile Prior from [1], fail.

Using a large set of images from the PubFig 83 database, Section 4 will show that our method outperforms the recent work of Sun et al. [7] in both visual quality and using quantitative image metrics. We use [7] as a reference because it utilizes both edge statistics and texture statistics used in previous work. In

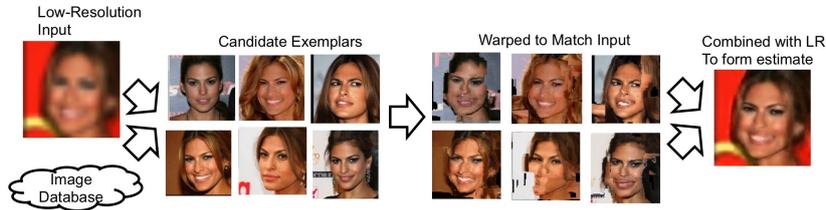


Fig. 2. In this paper, we propose alignment-based hallucination of images. This approach produces a high-resolution estimate from a low-resolution input by drawing exemplars from a database, computing a dense correspondence between the input and those exemplars, then computing the estimated high-resolution image through MAP estimation.

addition, [7] shows that this system performs well when compared to recent super-resolution systems.

More recent work, such as [14,15], uses face alignment techniques to improve images. The work presented here is unique in that the matching is performed on low-resolution images, which is more difficult. In addition, the exemplar images needed to enhance the input image are chosen automatically.

2 System Overview

Figure 2 shows an overview of how images are enhanced. Given a low-resolution image, the estimated high-resolution image is produced with the following steps:

1. Given the low-resolution input, the database of candidate exemplars is searched to find images that can be downsampled and warped to match the input well.
In practice, we have found that seven candidates is suitable to generate good results.
2. The SIFT Flow algorithm is used to warp each of the candidates to match the low-resolution input.
3. The MAP estimate of the high-resolution image is computed using a Bayesian framework that will be introduced in Section 3.

3 A Bayesian Framework for Example-Based Super Resolution

We will express the super-resolution process in a Bayesian MAP inference framework. The variable I_L will be used to denote the input low-res image and I_H will denote the high-res image to be inferred. Under Bayesian MAP inference framework, our goal is to find the image I_H^* that maximizes the posterior

$$p(I_H|I_L) \propto p(I_L|I_H)p(I_H). \quad (1)$$

3.1 Likelihood Function

The likelihood function, $p(I_L|I_H)$, expresses the compatibility between the low-resolution observed image and some high-resolution image I_H . We expect that I_L should be similar to the result of filtering I_H with an anti-aliasing filter, then down-sampling. This is expressed formally as

$$p(I_L|I_H) = \frac{1}{Z_r} \exp \left\{ -\lambda_r (\mathbf{M}I_H - I_L)^T (\mathbf{M}I_H - I_L) \right\}, \quad (2)$$

where \mathbf{M} is the product of a sub-sampling matrix and a 2D convolution matrix and Z_r is the normalization constant. Because this is a conditional Gaussian distribution, the coefficient λ_r can be thought of as an inverse variance. It can also be viewed as effectively controlling how strongly this reconstruction constraint is imposed: the larger λ_r , the stronger the constraint.

3.2 Image Prior

The key to success on this problem is the construction of an effective image prior, $p(I_H)$. A basic prior that focuses on edges, similar to work discussed previously, can be formed by focusing on image gradients :

$$p(I_H) = \frac{1}{Z_g} \exp \left\{ -\lambda_g \sum_{\mathbf{q}} \left| \nabla I_H(\mathbf{q}) \right| \right\} \quad (3)$$

where \mathbf{q} indexes an image lattice and $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$ is a gradient operator. The operator $|\cdot|$ is the L1 norm, λ_g scales the regularization, and Z_g is a normalization constant.

Using this prior in a MAP inference formulation leads to the following optimization problem

$$I_H^* = \arg \min \lambda_r (\mathbf{M}I_H - I_L)^T (\mathbf{M}I_H - I_L) + \lambda_g \sum_{\mathbf{q}} \left| \nabla I_H(\mathbf{q}) \right|, \quad (4)$$

which is similar to super-resolution based on the total-variation model and has been shown to sharpen edges better than bicubic sampling.

3.3 Example-Based Prior

Unfortunately, the gradient prior in Equation (3) is too simple to model the complexity of high-resolution images well. In this work, instead of relying on a simple parametric prior on image gradients, our prior will be formed from a large database of images, which will be denoted as the set $\{I_i\}_{i=1}^N$. It can be challenging to derive parametric forms of priors for natural images or for a particular class of images such as faces, so we instead write the distribution in a form of Parzen window (or kernel density):

$$p(I_H) = \frac{1}{N} \sum_{i=1}^N \frac{1}{Z_i} \exp \{ -\mathbf{K}(I_H, I_i) \} \quad (5)$$

where $\mathbf{K}(I_H, I_i)$ is a kernel function measuring image similarities, and we will introduce $\mathbf{K}(\cdot, \cdot)$ below. As we enforce $\int \frac{1}{Z_i} \exp\{-\mathbf{K}(I_H, I_i)\} dI_H = 1$, it is obvious that the probability distribution defined in Eqn (5) is valid.

Defining the Kernel to Measure Image Similarities. Following [12], we propose that the first step in measuring the similarity between two images is to establish a dense correspondence between the images using the SIFT flow algorithm. If s_1 and s_2 hold the dense SIFT features for two images, the SIFT flow is obtained by minimizing the following function [12]:

$$E(\mathbf{w}) = \sum_{\mathbf{q}} \min \left(\|s_1(\mathbf{q}) - s_2(\mathbf{q} + \mathbf{w}(\mathbf{q}))\|_1, t \right) + \sum_{\mathbf{q}} \eta (|u(\mathbf{q})| + |v(\mathbf{q})|) + \sum_{(\mathbf{q}, \mathbf{p}) \in \varepsilon} \min \left(\alpha \|u(\mathbf{p}) - u(\mathbf{q})\|, d \right) + \min \left(\alpha \|v(\mathbf{p}) - v(\mathbf{q})\|, d \right) \quad (6)$$

where ε are the sets of edges between adjacent pixels and \mathbf{w}_j defines a dense correspondence between I_H and I_j .

Once the dense correspondence is established, the distance $\mathbf{K}(I_H, I_j)$ can be computed by effectively warping one image to match the other, then measuring the difference. Because we are only interested in transferring high frequencies to the estimate, as the low-frequencies are visible in the observation, the difference is measured on high-pass filtered versions of the image, which we denote using \hat{I} .

This is expressed formally as

$$\begin{aligned} \mathbf{K}(I_H, I_j) &= \sum_{\mathbf{q}} \sum_{\mathbf{p} \in \mathcal{N}} \left(\hat{I}_H(\mathbf{q} + \mathbf{p}) - \hat{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q}) + \mathbf{p}) \right)^2 \\ &= \sum_{\mathbf{q}} \left(\mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \right)^T \left(\mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \right) \\ &= \sum_{\mathbf{q}} \left\| \mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \right\|_2^2, \end{aligned} \quad (7)$$

where $\mathcal{N} = [-s, \dots, s] \times [-s, \dots, s]$ defines a $(2s + 1) \times (2s + 1)$ patch.

We use $\mathbf{I}_H(\mathbf{q})$ to denote the image patch of \hat{I}_H centered at \mathbf{q} . This causes the kernel similarity to be based on aggregated patch distances to help avoid matching discontinuities.

Choosing a Subset of Images. When there are many samples, evaluating Equation (5) can be expensive. To alleviate this, it is possible to eliminate examples that are dissimilar to the input I_L because the likelihood will be low and the contribution of that example will be negligible.

In practice, we use the PatchMatch algorithm to quickly estimate eliminate examples that are very dissimilar from the input. Before the difference is computed, the image bilinear interpolation is used to downsample and upsample I_i by the same scale factor that will be used to enlarge the I_i . In our experiments, all of the images are roughly the same size, but the example images could be reduced to match the size of I_L after it is upsampled.

3.4 Understanding the Behavior of the Prior

Using this non-parametric prior, the log posterior can be computed as

$$\begin{aligned}
 & -\log p(I_L|I_H)p(I_H) \\
 &= -\log p(I_L|I_H) \\
 & -\log \frac{1}{M} \sum_{I_j \in \Omega_{I_L}} \frac{1}{Z_i} \exp\{-\mathbf{K}(I_H, I_j)\}
 \end{aligned} \tag{8}$$

To understand the function of this prior, it is helpful to consider the soft approximation of the min operator based on the log operation. The min operator over a collection of variables x_1, \dots, x_n can be approximated as

$$\min(x_1, \dots, x_n) \approx -\frac{1}{\eta} \log \sum_{i=1}^n e^{-\eta x_i}, \tag{9}$$

where $\eta \geq 0$ controls softness. The approximation gets tighter as $\eta \rightarrow \infty$:

$$\lim_{\eta \rightarrow \infty} -\frac{1}{\eta} \log \sum_{i=1}^n e^{-\eta x_i} = \min_i x_i. \tag{10}$$

In the context of this approximation, the prior can be seen as encouraging I_H to resemble one of the exemplar images in the non-parametric prior distribution. It has the unfortunate side-effect of effectively choosing one of the example images in the prior, then encouraging every patch in I_H to match the corresponding patch in the chosen example.

This will be problematic when patches in some exemplars do not match well with the patch in the I_H . The ideal behavior is to enable different patches in I_H to match different example images, thus giving the system flexibility to draw information from a wide variety of images.

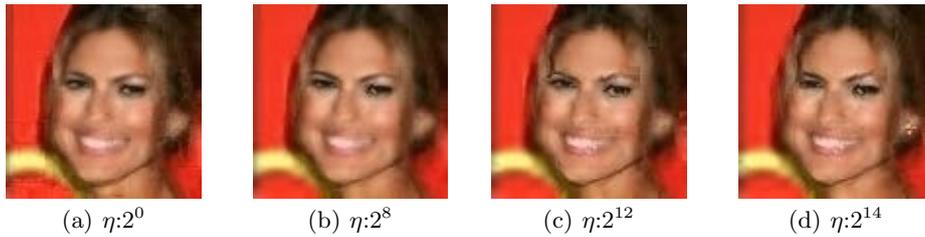


Fig. 3. These figures show how different values of η affect results. When η has a low value, the optimization tends to average patches. As it rises, the optimization more closely approximates the min operation, so the result looks sharper. However, this also introduces artifacts in the image.

3.5 Flexibly Drawing Patches From Multiple Images

Going back to the soft approximation of the min operation, finding the MAP estimate of the posterior in Equation (8) is closely related to the optimization problem

$$I_H^* = \arg \min \lambda_r \| \mathbf{M}I_H - I_L \|_2^2 + \min_{\mathbf{q}} \sum_j \| \mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \|_2^2. \quad (11)$$

To make it possible for patches to be drawn from different images, Equation (11) can be modified to move the min operation inside the summation, resulting in

$$I_H^* = \arg \min \lambda_r \| \mathbf{M}I_H - I_L \|_2^2 + \sum_{\mathbf{q}} \min_j \| \mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \|_2^2. \quad (12)$$

This optimization can be mirrored by modifying the prior to form the posterior

$$-\log p(\mathbf{I}_L | I_H) p(I_H) + \dots - \frac{1}{\eta} \sum_{\mathbf{q}} \log \sum_j \frac{1}{M} \exp \left\{ -\eta \| \mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \|_2^2 \right\}, \quad (13)$$

where Z is the normalization constant of the prior and there are M exemplars. Because the patches overlap, this prior is a Markov Random Field with loops in the graph representing this distribution.

This makes the MAP inference procedure become the optimization problem

$$I_H^* = \arg \min \lambda_r \| \mathbf{M}I_H - I_L \|_2^2 + \lambda_g \sum_{\mathbf{q}} \left| \nabla I_H(\mathbf{q}) \right| - \frac{1}{\eta} \sum_{\mathbf{q}} \log \sum_j \frac{1}{M} \exp \left\{ -\eta \| \mathbf{I}_H(\mathbf{q}) - \mathbf{I}_j(\mathbf{q} + \mathbf{w}_j(\mathbf{q})) \|_2^2 \right\}, \quad (14)$$

The optimization is implemented by using Jensen's inequality to compute a quadratic upper-bound to the second term in Equation (14). The complete energy-function is optimized by alternating steps of fitting the upper-bound, optimizing the upper-bound, then fitting the bound again.

Controlling System Behavior Through η . An advantage of this formulation is that it produces a family of systems with different behaviors that are indexed by the parameter η . As η approaches 0, the behavior this formulation approaches averaging because

$$\lim_{\eta \rightarrow 0} -\frac{1}{\eta} \log \frac{1}{n} \sum_{i=1}^n e^{-\eta x_i} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (15)$$

Interestingly, it is also possible to show that as η approaches zero, the negative log posterior in Equation (13) is an upper bound on the negative log posterior in Equation (8)

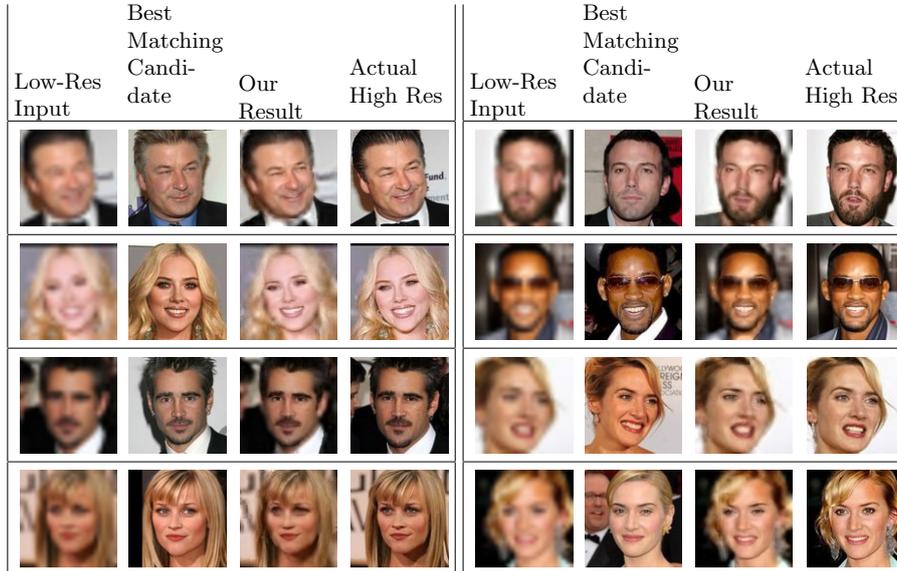


Fig. 4. This figure shows results produced by our system on images from the Pub-Fig83 dataset. In each of the images, the original is reduced by a factor of four, then enlarged by the same factor. The first column shows this low-resolution image after it has been enlarged with bicubic interpolation. The second column shows the most similar candidate (out of seven possible candidates) chosen to enhance the input. The third column shows our result, containing detail hallucinated by our algorithm, which can be compared with the actual high-resolution image in the fourth columns.

In contrast, as η rises, the min operation is more closely approximated. In practice, we have found that this tends to create sharper images, though the images tend to contain more artifacts. Figure 3 shows an example of how results change with different values of η .

An advantage of this approach is that it makes it possible to choose behavior that is in between pure averaging and a hard min operation. We have empirically chosen the value of 2^8 as a value that produces pleasing results.

Similarities to other Image Optimization Approaches. At a high-level, our approach is similar to the approach used by Simakov et al. in [16], which was also used in [13]. Both our system and [16] operate by combining sets of overlapping patches, but there are key differences in the actual process used to reconstruct the image between our approach and the approach in [16].

In [16], overlapping patches are combined through a patch-voting process where each patch contributes a color vote to the value of a pixel. The process optimizes a cost function based on finding patches that cohere and capture the source image as completely as possible.

In contrast, our estimation process optimizes a cost function that is focused on modeling the image formation process. In super-resolution, this makes it possible

to incorporate a reconstruction constraint that forces the estimated image to match the observation. This will also be important for other applications where an observation enforces a strong constraint on the final result.

The optimization used in this work is similar to that used in [7], with a key difference. Our optimization is based around patches rather than single pixels, as in [7].

4 Experiments

Our experiments will rely on the PubFig83 dataset presented in [17], which is a refinement of the PubFig database in [18]. The PubFig83 database is made up of a collection of over 14,000 images of 83 public figures. The images are all cropped to include just the face and resized to a resolution of 100×100 , but no alignment of features is performed.

Figure 4 shows examples of results produced by our system for several individuals in the PubFig83 database. All of these results are generated in the same fashion. The input is a 100×100 -pixel image of the individual. This image is downsampled by a factor of four, then used as input to our system. Our system then estimates an image that is the same size as the original 100×100 -pixel input. In these experiments, we only enhance the luminance channel of a YUV image. The chrominance channels are estimated using bicubic upsampling of the chrominance channels of the low-resolution input.

To simulate the proposed usage scenario, where an accurate face recognition system has already been applied to determine the identity of the individual in the picture, the system is provided with all of the images of that same individual, except the input image, to serve as potential exemplars. To best preserve the integrity of the experiments, we made no modifications to the PubFig83 dataset. Because of this decision, some images in the dataset are cropped, rescaled versions of another image in the same set – though this only affects a small minority of images.

There are also images, such as the first row of Figure 5, where the best candidate is a distinct image that appears to have been taken at a similar event. These cases demonstrate the power of our approach. Photos taken at the same event are ideal candidates for improving photos that did not turn out well. Each row in Figure 4 contains, in order, the low-resolution input that has been upsampled using bicubic interpolation, the exemplar in the database deemed to be the best match, and the result produced by our system.

4.1 Comparison with Other Approaches

As these results show, our system is able to synthesize realistic detail from very low-resolution results that lack both texture and image structure. To validate our approach, we compare against two different approaches, the super-resolution approach proposed by Sun et al. in [7] and the VISTA super-resolution approach proposed by Freeman et al. in [5]. The comparisons with [7] are useful because this approach was shown to perform well relative to a number of recent super-resolution algorithms in both qualitative and quantitative results. As mentioned

Table 1. This table contains a quantitative comparison of our system, on images of Eva Mendes, with the VISTA system proposed by Freeman et al. [5] and the texture hallucination system presented by Sun et al. in [7]. For both measures, higher values are better. The results produced by our approach are superior using either measure.

Algorithm	PSNR (dB)	SSIM Index
VISTA	23.47	0.669
Sun et al. [7]	23.82	0.741
Our Approach	24.05	0.748

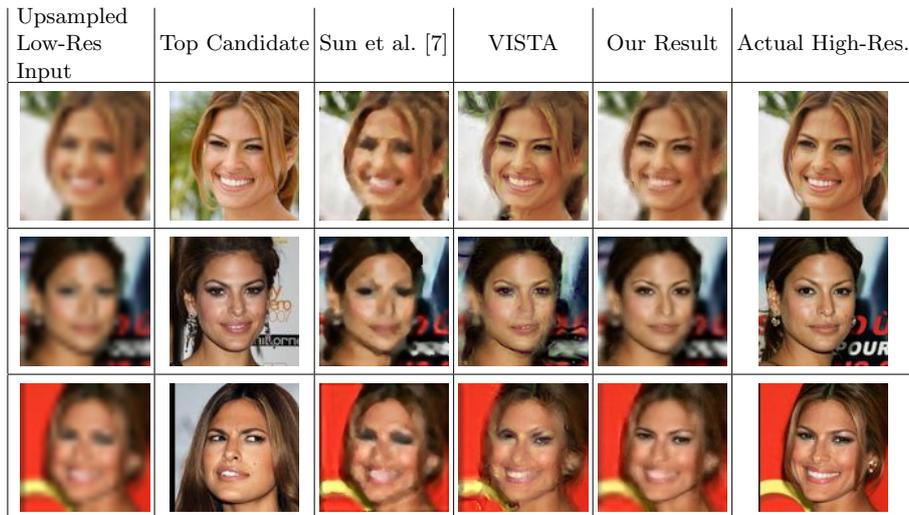


Fig. 5. These images compare the results from our system with the results from [5] and [7]. In addition to providing quantitatively improved results, our system also produces qualitatively better results.

in the introduction, the super-resolution system in [7] is a good system for baseline comparisons because it incorporates multiple super-resolution strategies, including models of edge appearance and patch-based texture models.

In all experiments with VISTA, we use the implementation created as part of [19]. To make the comparisons with VISTA as fair as possible, the algorithm was provided with the same candidate images chosen to serve as exemplars in our system. The results from the Sun et al. algorithm were produced courtesy of the authors of [7].

Table 1 presents quantitative results on the Eva Mendes images in the Pub-Fig83 database, using both PSNR and the SSIM index [20]. Our method produces both improved PSNR and SSIM scores over previous approaches. We also experimented with providing the VISTA algorithm with the warped exemplars, however we found that it did not significantly impact the resulting scores.

Figure 5 shows qualitative comparisons of our approach and these other approaches. Compared with the texture hallucination results, which relies strongly on a gradient-based image prior, our results contain more detail, while the results

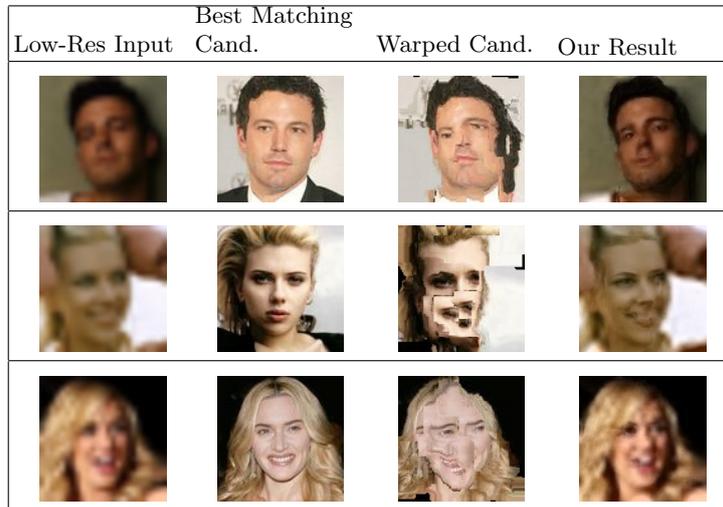


Fig. 6. This figure shows some examples of our system failing. Most failures can be attributed to the failure to find an image that can be warped to match the input well.

from [7] appear flatter. Compared with the results from VISTA, the artifacts in the results produced by our approach are greatly reduced.

4.2 System Limitations

As should be expected, the results are best when a candidate that is similar to the original image can be found. Figure 6 shows examples when the warped exemplars do not match the input image well. In general, if a good match cannot be found, then the system cannot be expected to perform well. The primary image artifact created by a poor match will be blurriness in areas where a well-matched high-resolution patch is unavailable. We have also found that some undesirable smoothing of edges is caused by the difficulty of matching content drawn from the database with the constraint that the output must downsample to match the input.

In a production environment, these limitations could be compensated for by falling back to an edge-based system, like [3,1], when matching costs become too high. It should be noted, though, that our quantitative and qualitative results show that it is much more common for our approach to find a good match and produce a higher quality image.

4.3 Necessity of SIFTFlow Step

As noted in Section 3.3, the PatchMatch algorithm is used first to quickly match the query to the database of exemplars to find a small set of exemplars that match well. The SIFTFlow algorithm is then used to find the correspondence

between the query and each of the selected exemplars. As can be seen in Figure 7, the spatial constraints in this SIFTFlow step improve results. Figure 7(a) shows a candidate estimated using the reconstruction routines in PatchMatch package. Because they lose much of the detail in the image, these exemplars lead to the blurry estimate shown in Figure 7(b).

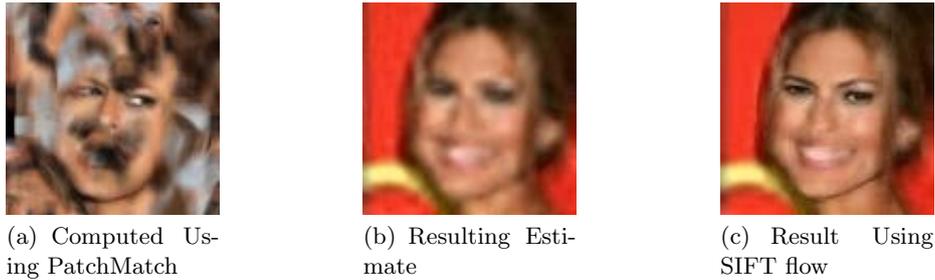


Fig. 7. We found that the spatial constraints in SIFTFlow to be important for a high-quality result. The image in (a) shows a candidate produced using the PatchMatch [13] algorithm. Because of the quality of this result, the result in (b) is not significantly improved.

4.4 Results Without Recognition

To measure the importance of recognition and finding exemplars of the same individual, we conducted experiments where the candidate exemplars all consisted of other individuals. Column (a) in Figure 8 shows the images obtained when

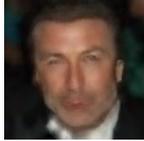
Low-Res Input	Best Matching Cand. (a)	Result	Result from Same Person(b)
			
			

Fig. 8. These examples shows the benefit of using exemplars of the same individual pictured in the image. The result in the third column is created from thousands of images drawn from 30 individuals, none of whom is the person shown in this image. For comparison, the fourth column shows the result of just using exemplars of the same person, which leads to a significantly improved results.

the first 30 individuals in the PubFig83 dataset, excluding the person shown in the input image, are used to produce a high-resolution estimate. For comparison, column (b) in Figure 8 shows the result using just images of the same person. While the result produced using the larger set of individuals contains realistic details, the result produced using images of the same person is more realistic and contains fewer artifacts.

5 Conclusion

In this paper we have proposed a image hallucination system based on new image alignment technologies. Our system is flexible enough to work with differences in pose and illumination, while still able to produce state-of-the-art results.

Acknowledgments. This work was initiated when M.F.T was a Consulting Researcher at Microsoft Research New England. M.F.T. was also supported by NSF grants IIS-0905387 and IIS-0916868.

References

1. Sun, J., Sun, J., Xu, Z.B., Shum, H.Y.: Image super-resolution using gradient profile prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2008)
2. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 1167–1183 (2002)
3. Fattal, R.: Upsampling via imposed edges statistics. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007)* 26 (2007)
4. Greenspan, H., Anderson, C.H., Akber, S.: Image enhancement by nonlinear extrapolation in frequency space. *IEEE Transactions on Image Processing* 9, 1035–1048 (2000)
5. Freeman, W.T., Pasztor, E.C., Carmichael, O.T.: Learning low-level vision. *International Journal of Computer Vision* 40, 25–47 (2000)
6. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. *IEEE Trans. Image Processing* 19, 2861–2873 (2010)
7. Sun, J., Zhu, J., Tappen, M.F.: Context-constrained hallucination for image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 231–238 (2010)
8. Tai, Y.W., Liu, S., Brown, M.S., Lin, S.: Super resolution using edge prior and single image detail synthesis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2400–2407 (2010)
9. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: ICCV (2009)
10. HaCohen, Y., Fattal, R., Lischinski, D.: Image upsampling via texture hallucination. In: Proceedings of the International Conference on Computational Photography (2010)
11. Liu, C., Shum, H.Y., Freeman, W.T.: Face hallucination: Theory and practice. *International Journal of Computer Vision* 75, 115–134 (2007)

12. Liu, C., Yuen, J., Torralba, A.: Sift flow: Dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 978–994 (2011)
13. Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B.: PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 28 (2009)
14. HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D.: Non-rigid dense correspondence with applications for image enhancement. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2011)* 30, 70:1–70:9 (2011)
15. Yang, F., Wang, J., Shechtman, E., Bourdev, L.D., Metaxas, D.N.: Expression flow for 3d-aware face component transfer. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2011)* 30, 60 (2011)
16. Simakov, D., Caspi, Y., Shechtman, E., Irani, M.: Summarizing visual data using bidirectional similarity. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2008)
17. Pinto, N., Stone, Z., Zickler, T., Cox, D.D.: Scaling-up Biologically-Inspired Computer Vision: A Case-Study on Facebook. In: *IEEE Computer Vision and Pattern Recognition, Workshop on Biologically Consistent Vision* (2011)
18. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Attribute and simile classifiers for face verification. In: *Proceedings of the IEEE International Conference on Computer Vision* (2009)
19. Freeman, W.T., Liu, C.: Markov random fields for super-resolution and texture synthesis. In: *Advances in Markov Random Fields for Vision and Image Processing*. MIT Press (2011)
20. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 600–612 (2004)