Super-Resolution from a Single Image

Daniel Glasner   Shai Bagon   Michal Irani
Dept. of Computer Science and Applied Mathematics
The Weizmann Institute of Science
Rehovot 76100, Israel

Abstract

Methods for super-resolution can be broadly classified into two families of methods: (i) The classical multi-image super-resolution (combining images obtained at subpixel misalignments), and (ii) Example-Based super-resolution (learning correspondence between low and high resolution image patches from a database). In this paper we propose a unified framework for combining these two families of methods. We further show how this combined approach can be applied to obtain super resolution from as little as a single image (with no database or prior examples). Our approach is based on the observation that patches in a natural image tend to redundantly recur many times inside the image, both within the same scale, as well as across different scales. Recurrence of patches within the same image scale (at subpixel misalignments) gives rise to the classical super-resolution, whereas recurrence of patches across different scales of the same image gives rise to example-based super-resolution. Our approach attempts to recover at each pixel its best possible resolution increase based on its patch redundancy within and across scales.

1. Introduction

The goal of Super-Resolution (**SR**) methods is to recover a high resolution image from one or more low resolution input images. Methods for SR can be broadly classified into two families of methods: (i) The classical multi-image super-resolution, and (ii) Example-Based super-resolution. In the classical multi-image SR (e.g., [12, 5, 8] to name just a few) a set of low-resolution images of the same scene are taken (at subpixel misalignments). Each low resolution image imposes a set of linear constraints on the unknown high-resolution intensity values. If enough low-resolution images are available (at subpixel shifts), then the set of equations becomes determined and can be solved to recover the high-resolution image. Practically, however, this approach is numerically limited only to small increases in resolution [3, 14] (by factors smaller than 2).
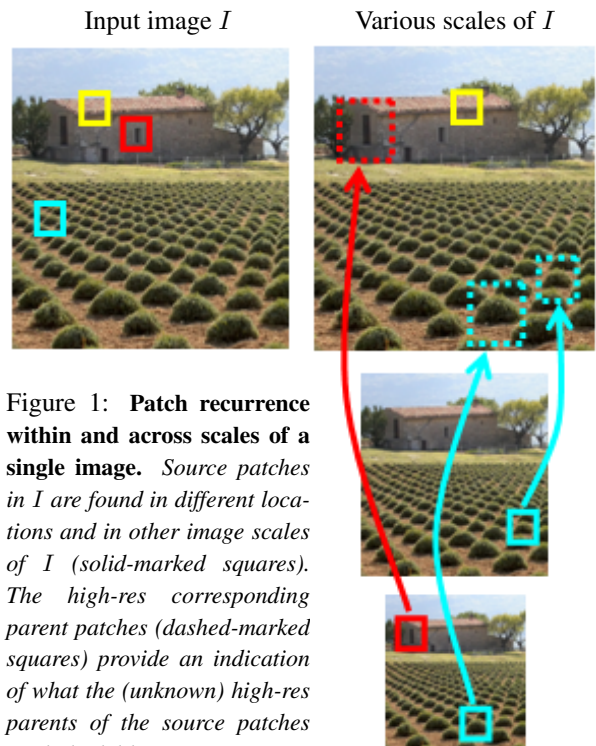
These limitations have lead to the development of

Input image I Various scales of I

Figure 1: **Patch recurrence within and across scales of a single image.** *Source patches in I are found in different locations and in other image scales of I (solid-marked squares). The high-res corresponding parent patches (dashed-marked squares) provide an indication of what the (unknown) high-res parents of the source patches might look like.*

"Example-Based Super-Resolution" also termed "image hallucination" (introduced by [10, 11, 2] and extended later by others e.g. [13]). In example-based SR, correspondences between low and high resolution image patches are learned from a database of low and high resolution image pairs (usually with a relative scale factor of 2), and then applied to a new low-resolution image to recover its most likely high-resolution version. Higher SR factors have often been obtained by repeated applications of this process. Example-based SR has been shown to exceed the limits of classical SR. However, unlike classical SR, the high resolution details reconstructed ("hallucinated") by example-based SR are not guaranteed to provide the true (unknown) high resolution details.

Sophisticated methods for image up-scaling based on learning edge models have also been proposed (e.g., [9,

19]). The goal of these methods is to magnify (up-scale) an image while maintaining the sharpness of the edges and the details in the image. In contrast, in SR (example-based as well as classical) the goal is to recover new *missing high-resolution details* that are not explicitly found in any individual low-resolution image (details beyond the Nyquist frequency of the low-resolution image). In the classical SR, this high-frequency information is assumed to be split across multiple low-resolution images, implicitly found there in aliased form. In example-based SR, this missing high-resolution information is assumed to be available in the high-resolution database patches, and learned from the low-res/high-res pairs of examples in the database.

In this paper we propose a framework to combine the power of both SR approaches (Classical SR and Example-based SR), and show how this combined framework can be applied to obtain SR *from as little as a single low-resolution image*, without any additional external information. Our approach is based on an observation (justified statistically in the paper) that patches in a single natural image tend to redundantly recur many times inside the image, both within the same scale, as well as across different scales. Recurrence of patches within the same image scale (at subpixel misalignments) forms the basis for applying the classical SR constraints to information from a single image. Recurrence of patches across different (coarser) image scales implicitly provides examples of low-res/high-res pairs of patches, thus giving rise to example-based super-resolution from a single image (without any external database or any prior examples). Moreover, we show how these two different approaches to SR can be combined in a single unified computational framework.

Patch repetitions within an image were previously exploited for noise-cleaning using 'Non-Local Means' [4], as well as a regularization prior for inverse problems [15]. A related SR approach was proposed by [16] for obtaining higher-resolution video frames, by applying the classical SR constraints to similar patches across consecutive video frames and within a small local spatial neighborhood. Their algorithm relied on having multiple image frames, and did not exploit the power of patch redundancy *across* different image scales. The power of patch repetitions across scales (although restricted to a fixed scale-factor of 2) was previously alluded to in the papers [10, 18, 6]. In contrast to all the above, we propose a single unified approach which combines the classical SR constraints with the example-based constraints, while exploiting (for each pixel) patch redundancies across all image scales and at varying scale gaps, thus obtaining adaptive SR with as little as a single low-resolution image.

The rest of this paper is organized as follows: In Sec. 2 we statistically examine the observation that small patches in a single natural image tend to recur many times within
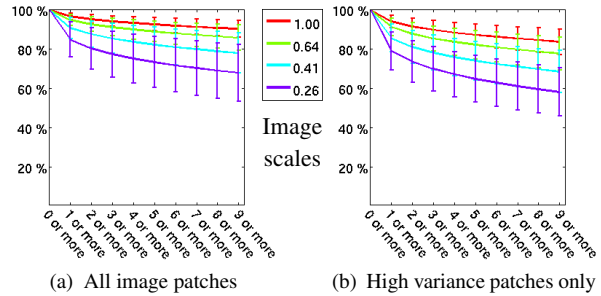

(a) All image patches   (b) High variance patches only

Figure 2: *Average patch recurrence within and across scales of a single image (averaged over hundreds of natural images – see text for more details). (a) The percent of image patches for which there exist n or more similar patches ($n = 1, 2, 3, ..., 9$), measured at several different image scales. (b) The same statistics, but this time measured only for image patches with the highest intensity variances (top 25%). These patches correspond to patches of edges, corners, and texture.*

and across scales of the same image. Sec. 3 presents our unified SR framework (unifying classical SR and example-based SR), and shows how it can be applied to as little as a single image. Results are provided in Sec. 4, as well as the url of the paper's website where more results can be found.

2. Patch Redundancy in a Single Image

Natural images tend to contain repetitive visual content. In particular, small (e.g., 5×5) image patches in a natural image tend to redundantly recur many times inside the image, both within the same scale, as well as across different scales. This observation forms the basis for our single-image super-resolution framework as well as for other algorithms in computer vision (e.g., image completion [7], image re-targeting [17], image denoising [4], etc.) In this section we try to empirically quantify this notion of patch redundancy (within a single image).

Fig. 1 schematically illustrates what we mean by "patch recurrence" *within and across* scales of a single image. An input patch "recurs" in another scale if it appears 'as is' (without blurring, subsampling, or scaling down) in a scaled-down version of the image. Having found a similar patch in a smaller image scale, we can extract its high-resolution parent from the input image (see Fig. 1). Each low-res patch with its high-res parent form a "low-res/high-res pair of patches" (marked by arrows in the figure). The high-res parent of a found low-res patch provides an indication to what the (unknown) high-res parent of the source patch might look like. This forms the basis for Example-Based SR, even *without* an external database. For this approach to be effective, however, enough such recurring patches must exist in different scales of the same image.

The patches displayed in Fig. 1 were chosen large for illustration purpose, and were displayed on clear repetitive structure in the image. However, when much smaller image

patches are used, e.g., 5×5, such patch repetitions occur abundantly within and across image scales, even when we do not visually perceive any obvious repetitive structure in the image. This is due to the fact that very small patches often contain only an edge, a corner, etc. such patches are found abundantly in multiple image scales of almost any natural image.

Moreover, due to the perspective projection of cameras, images tend to contain scene-specific information in diminishing sizes (diminishing toward the horizon), thus recurring in multiple scales of the same image.

We statistically tested this observation on the *Berkeley Segmentation Database*[1] (Fig. 2). More specifically, we tested the hypothesis that small 5×5 patches in a single natural *grayscale* image, when removing their DC (their average grayscale), tend to recur many times within and across scales of the same image. The test was performed as follows: Each image I in the Berkeley database was first converted to a grayscale image. We then generated from I a cascade of images of decreasing resolutions $\{I_s\}$, scaled (down) by scale factors of 1.25^s for $s = 0, -1, .., -6$ ($I_0 = I$). The size of the smallest resolution image was $1.25^{-6} = 0.26$ of the size of the source image I (in each dimension). Each 5×5 patch in the source image I was compared against the 5×5 patches in all the images $\{I_s\}$ (without their DC), measuring how many similar[2] patches it has in each image scale. This intra-image patch statistics was computed separately for each image. The resulting independent statistics were then averaged across all the images in the database (300 images), and are shown in Fig. 2a. Note that, on the average, more than 90% of the patches in an image have 9 or more other similar patches in the same image at the original image scale ('within scale'). Moreover, more than 80% of the input patches have 9 or more similar patches in $0.41 = 1.25^{-4}$ of the input scale, and 70% of them have 9 or more similar patches in $0.26 = 1.25^{-6}$ of the input scale.

Recurrence of patches forms the basis for our single-image super-resolution approach. Since the impact of super-resolution is expressed mostly in highly detailed image regions (edges, corners, texture, etc.), we wish to eliminate the effect of uniform patches on the above statistics. Therefore, we repeated the same experiment using only 25% of the source patches with the highest intensity variance. This excludes the uniform and low-frequency patches,

---

[1]www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench

[2]Distances between patches were measured using gaussian-weighted SSD. Note that textured patches tend to have much larger SSD errors than smooth (low-variance) patches when compared to other very similar-looking patches (especially in the presence of inevitable sub-pixel misalignments). Thus, for each patch we compute a patch-specific 'good distance', by measuring its (gaussian-weighted) SSD with a slightly-misaligned copy of itself (by 0.5 pixel). This forms our distance threshold for each patch: Patches with distance below this threshold are considered similar to the source patch.

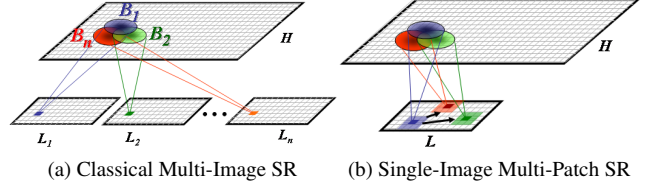(a) Classical Multi-Image SR   (b) Single-Image Multi-Patch SR

Figure 3: *(a) Low-res pixels in multiple low-res images impose multiple linear constraints on the high-res unknowns within the support of their blur kernels. (b) Recurring patches within a single low-res image can be regarded as if extracted from multiple different low-res images of the same high resolution scene, thus inducing multiple linear constraints on the high-res unknowns.*

maintaining mostly patches of edges, corners, and texture. The resulting graphs are displayed in Fig. 2b. Although there is a slight drop in patch recurrence, the basic observation still holds even for the high-frequency patches: Most of them recur several times within and across scales of the same image (more than 80% of the patches recur 9 or more times in the original image scale; more than 70% recur 9 or more times at 0.41 of the input scale, and 60% of them recur 9 or more times in 0.26 of the input scale.)

In principle, the lowest image scale in which we can still find recurrence of a source patch, provides an indication of its maximal potential resolution increase using our approach (when the only available information is the image itself). This is pixel-dependent, and can be estimated at every pixel in the image.

3. Single Image SR – A Unified Framework

Recurrence of patches within the same image scale forms the basis for applying the Classical SR constraints to information from a single image (Sec. 3.1). Recurrence of patches across different scales gives rise to Example-Based SR from a single image, with no prior examples (Sec. 3.2). Moreover, these two different approaches to SR can be combined into a single unified computational framework (Sec. 3.3).

3.1. Employing in-scale patch redundancy

In the classical Multi-Image Super-resolution (e.g., [12, 5, 8]), a set of low-resolution images $\{L_1, ..., L_n\}$ of the same scene (at subpixel misalignments) is given, and the goal is to recover their mutual high-resolution source image H. Each low resolution image L_j $(j = 1, \ldots, n)$ is assumed to have been generated from H by a blur and sub-sampling process: $L_j = (H * B_j) \downarrow_{s_j}$, where $\downarrow$ denotes a subsampling operation, s_j is the scale reduction factor (the subsampling rate) between H and L_j, and $B_j(q)$ is the corresponding blur kernel (the Point Spread Function – PSF), represented in the high-resolution coordinate system – see Fig. 3a. Thus, each low-resolution pixel $p = (x, y)$ in each low-resolution image L_j induces one linear constraint on

the unknown high-resolution intensity values within the local neighborhood around its corresponding high-resolution pixel $q \in H$ (the size of the neighborhood is determined by the support of the blur kernel B_j):

$$L_j(p) = (H * B_j)(q) = \Sigma_{q_i \in Support(B_j)} H(q_i) B_j(q_i - q) \tag{1}$$

where $\{H(q_i)\}$ are the unknown high-resolution intensity value. If enough low-resolution images are available (at sub-pixel shifts), then the number of independent equations exceeds the number of unknowns. Such super-resolution schemes have been shown to provide reasonably stable super resolution results up to a factor of ≈ 2 (a limit of 1.6 is shown in [14] when noise removal and registration are not good enough).

In principle, when there is only a single low-resolution image $L = (H * B) \downarrow_s$, the problem of recovering H becomes under-determined, as the number of constraints induced by L is smaller than the number of unknowns in H. Nevertheless, as observed in Sec. 2, there is plenty of patch redundancy within a single image L. Let p be a pixel in L, and $\mathcal{P}$ be its surrounding patch (e.g., 5×5), then there exist multiple similar patches $\mathcal{P}_1, ... \mathcal{P}_k$ in L (inevitably, at sub-pixel shifts). These patches can be treated as if taken from k different low-resolution images of the same high resolution "scene", thus inducing k times more linear constraints (Eq. (1)) on the high-resolution intensities of pixels within the neighborhood of $q \in H$ (see Fig. 3b). For increased numerical stability, each equation induced by a patch $\mathcal{P}_i$ is globally scaled by the degree of similarity of $\mathcal{P}_i$ to its source patch $\mathcal{P}$. Thus, patches of higher similarity to $\mathcal{P}$ will have a stronger influence on the recovered high-resolution pixel values than patches of lower similarity.

These ideas can be translated to the following simple algorithm: For each pixel in L find its k nearest patch neighbors in the same image L (e.g., using an Approximate Nearest Neighbor algorithm [1]; we typically use k=9) and compute their subpixel alignment (at $\frac{1}{s}$ pixel shifts, where s is the scale factor.) Assuming sufficient neighbors are found, this process results in a *determined* set of linear equations on the unknown pixel values in H. Globally scale each equation by its reliability (determined by its patch similarity score), and solve the linear set of equations to obtain H. An example of such a result can be found in Fig. 5c.

3.2. Employing cross-scale patch redundancy

The above process allows to extend the applicability of the classical Super-Resolution (SR) to a *single image*. However, even if we disregard additional difficulties which arise in the single image case (e.g., the limited accuracy of our patch registration; image patches with insufficient matches), this process still suffers from the same inherent limitations of the classical multi-image SR (see [3, 14]).

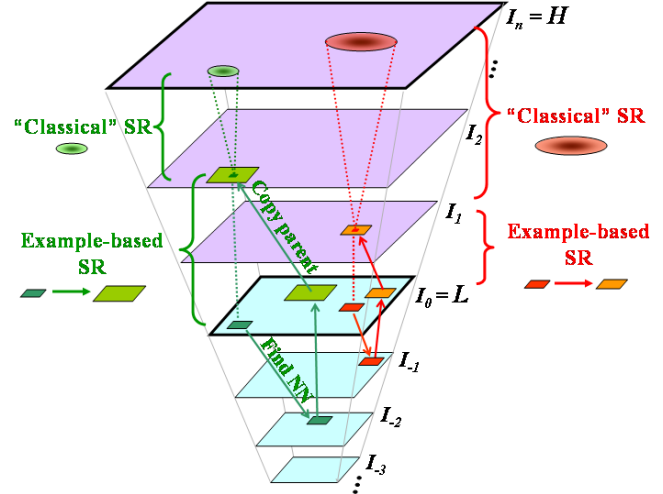The limitations of the classical SR have lead to the devel-

Figure 4: **Combining Example-based SR constraints with Classical SR constraints in a single unified computational framework.** *Patches in the input low-res image L (dark red and dark green patches) are searched for in the down-scaled versions of L (blue-marked images). When a similar patch is found, its parent patch (light red and light green) is copied to the appropriate location in the unknown high-resolution image (purple images) with the appropriate gap in scale. A 'learned' (copied) high-res patch induces classical SR linear constraints on the unknown high-res intensities in the target high-res H. The support of the corresponding blur kernels (red and green ellipses) are determined by the residual gaps in scale between the resolution levels of the 'learned' high-res patches and the target resolution level of H. Note that for different patches found in different scale gaps, the corresponding blur kernels (red and green ellipses) will accordingly have different supports. (See text for more details.)*

opment of "Example-Based Super-Resolution" (e.g., [11, 2]). In example-based SR, correspondences between low and high resolution image patches are learned from a database of low and high resolution image pairs, and then applied to a new low-resolution image to recover its most likely high-resolution version. Example-based SR has been shown to exceed the limits of classical SR. In this section we show how similar ideas can be exploited within our single image SR framework, without any external database or any prior example images. The low-res/high-res patch correspondences can be learned directly from the image itself, by employing patch repetitions across multiple image scales.

Let B be the blur kernel (camera PSF) relating the low-res input image L with the unknown high-res image H: $L = (H * B) \downarrow_s$. Let $I_0, I_1, ..., I_n$ denote a cascade of *unknown* images of *increasing* resolutions (scales) ranging from the low-res L to the target high-res H ($I_0 = L$ and $I_n = H$), with a corresponding cascade of blur functions $B_0, B_1, ..., B_n$ (where $B_n = B$ is the PSF relating H to L, and B_0 is the δ function), such that every I_l satisfies: $L = (I_l * B_l) \downarrow_{s_l}$ (s_l denotes the relative scaling factor).

(a) Input image (scaled for display).   (b) Bicubic interpolation (×2).   (c) Within image repetitions (×2).   (d) Unified single-image SR (×2).

Figure 5: *Comparing single-image SR with the 'classical' SR constraints only, to the unified single-image SR (Classical + Example-based constraints). Note that the 'classical' SR constraints, when applied to similar within-scale patches, results in a high-resolution image (c) which is sharper and cleaner than the interpolated image (b), but is not able to recover the fine rail in the intermediate arched windows. In contrast, the high-resolution image (d) produced using the unified Classical + Example-based constraints recovers these fine rails.*

The resulting cascade of images is illustrated in Fig. 4 (the purple images).

Note that although the images $\{I_l\}_{l=0}^n$ are unknown, the cascade of blur kernels $\{B_l\}_{l=0}^n$ can be assumed to be known. When the PSF B is unknown (which is often the case), then B can be approximated with a gaussian, in which case $B_l = B(s_l)$ are simply a cascade of gaussians whose variances are determined by s_l. Moreover, when the scale factors s_l are chosen such that $s_l = \alpha^l$ for a fixed α, then the following constraint will also hold for all $\{I_l\}_{l=1}^n$: $I_l = (H * B_{n-l}) \downarrow_{s_{n-l}}$. (The uniform scale factor guarantees that if two images in this cascade are found m levels apart (e.g., , I_l and I_{l+m}), they will be related by the same blur kernel B_m, regardless of l.)

Let $L = I_0, I_{-1}, ..., I_{-m}$ denote a cascade of images of *decreasing* resolutions (scales) obtained from L using the same blur functions $\{B_l\}$: $I_{-l} = (L * B_l) \downarrow_{s_l}$ ($l = 0, .., m$). Note that unlike the high-res image cascade, these low-resolution images are known (computed from L). The resulting cascade of images is also illustrated in Fig. 4 (the blue images).

Let $\mathcal{P}_l(p)$ denote a patch in the image I_l at pixel location p. For any pixel in the input image $p \in L$ ($L = I_0$) and its surrounding patch $\mathcal{P}_0(p)$, we can search for similar patches within the cascade of low resolution images $\{I_{-l}\}, l > 0$ (e.g., using Approximate Nearest Neighbor search [1]). Let $\mathcal{P}_{-l}(\tilde{p})$ be such a matching patch found in the low-res image I_{-l}. Then its higher-res 'parent' patch, $\mathcal{Q}_0(s_l \cdot \tilde{p})$, can be extracted from the input image $I_0 = L$ (or from any intermediate resolution level between I_{-l} and L, if desired). This provides a low-res/high-res patch pair $[\mathcal{P}, \mathcal{Q}]$, which provides a prior on the appearance of the high-res parent of the low-res input patch $\mathcal{P}_0(p)$, namely patch $\mathcal{Q}_l(s_l \cdot p)$ in the high-res unknown image I_l (or in any intermediate resolution level between L and I_l, if desired). The basic step is therefore as follows (schematically illustrated in Fig. 4):

$$\mathcal{P}_0(p) \xrightarrow{findNN} \mathcal{P}_{-l}(\tilde{p}) \xrightarrow{parent} \mathcal{Q}_0(s_l \cdot \tilde{p}) \xrightarrow{copy} \mathcal{Q}_l(s_l \cdot p)$$

3.3. Combining Classical and Example-Based SR

The process described in Sec 3.2, when repeated for all pixels in L, will yield a large collection of (possibly overlapping) suggested high-res patches $\{Q_l\}$ at the range of resolution levels $l = 1, .., n$ between L and H. Each such 'learned' high-res patch Q_l induces *linear constraints* on the unknown target resolution H. These constraints are in the form of the classical SR constraints of Eq. (1), but with a *more compactly supported blur kernel* than $B = PSF$. These constraints are induced by a smaller blur kernel B_{n-l} which needs to compensate only for the *residual gap in scale* $(n - l)$ between the resolution level l of the 'learned' patch and the final resolution level n of the target high-res H. This is illustrated in Fig. 4.

The closer the learned patches are to the target resolution H, the better conditioned the resulting set of equations is (since the blur kernel gradually approaches the δ function, and accordingly, the coefficient matrix gradually approaches the identity matrix). Note that the constraints in Eq. 1 are of the same form, with $l = 0$ and $B = PSF$. As in Sec. 3.1, each such linear constraint is globally scaled by its reliability (determined by its patch similarity score). Note that if, for a particular pixel, the only similar patches found are within the input scale L, then this scheme reduces to the 'classical' single-image SR of Sec. 3.1 at that pixel; and if no similar patches are found, this scheme reduces to simple deblurring at that pixel. Thus, the above scheme guarantees to provide the *best possible resolution increase at each pixel* (according to its patch redundancy within and across scales of L), *but never worse than simple upscaling (interpolation)* of L.

Solving Coarse-to-Fine: In most of our experiments we used the constant scale factor $\alpha = 1.25$ (namely, $s_l = 1.25^l$). When integer magnification factors were desired this value was adjusted (e.g. for factors 2 and 4 we used $\alpha = 2^{(1/3)}$). In our current implementation the above set of linear equations was not solved at once to produce H,

but rather gradually, coarse-to-fine, from the lowest to the highest resolution. When solving the equations for image I_{l+1}, we employed not only the low-res/high-res patch correspondences found in the input image L, but also all newly learned patch correspondences from the newly recovered high-res images so far: $I_0, ..., I_l$. This process is repeated until the resolution level of H is reached. We found this gradual scheme to provide numerically more stable results.

To further guarantee consistency of the recovered high-res results, when a new high-res image I_l is obtained, it is projected onto the low-res image L (by blurring and subsampling) and compared to L. Large differences indicate errors in the corresponding high-res pixels, and are thus 'back-projected' [12] onto I_l to correct those high-res pixels. This process verifies that each newly recovered I_l is consistent with the input low resolution image.

Observations: Our experiments show that the main improvement in resolution comes from the Example-Based SR component in our combined framework. However, the Classical-SR component (apart from providing small resolution increase - see Fig. 5c), plays a central role in preventing the Example-Based SR component from hallucinating erroneous high-res details (a problem alluded to by [11]). Our combined Classical + Example-Based SR framework can be equivalently posed as optimizing an objective function with a 'data-term' and two types of 'prior-terms': The data-term stems from the blur+subsample relation (of the Classical SR) between the high-res image H and low-res image L. The Example-Based SR constraints form one type of prior, whereas the use of multiple patches in the Classical SR constraints form another type of prior (at sub-pixel accuracy). The high-res image H which optimizes this objective function must satisfy both the Example-Based SR and the Classical SR constrains simultaneously, which is the result of our combined framework.

Although presented here in the context of single-image SR, the proposed unified framework (classical + example-based) can be applied also in other contexts of SR. It can extend classical SR of multiple low-res images of the same scene by adding the example-based cross-scale constraints. Similarly, existing example-based SR methods which work with an external database can be extended by adding our unified SR constraints.

4. Experimental Results

Figs. 5,6,7,8 and 9 show results of our SR method. **Full scale images, comparisons with other methods and many more results can be found in** http://www.wisdom.weizmann.ac.il/∼vision/SingleImageSR.html. When working with color images, the image is first transformed from RGB to YIQ. The SR algorithm is then applied to the Y (intensity) channel. The I and Q chromatic channels (which are characterized by low frequency information) are only interpolated (bi-cubic). The three channels are then combined to form our SR result.

Fig. 6 compares our unified SR result against ground truth. In Fig. 7 we compare our method to results from [11, 13] and [9]. Note that our results are comparable, even though we do not use any external database of low-res/high-res pairs of patches [11, 13], nor a parametric learned edge model [9]. Other examples with more comparisons to other methods can be found in the paper's website.

Fig. 5 displays an example of the different obtainable resolution improvements by using only within-scale classical SR constraints (Sec. 3.1), versus adding also cross-scale example-based constraints (Sec. 3.3). Further results on various types of natural images can be seen in Figs. 8 and 9, as well as in the paper's website.

References

[1] S. Arya and D. M. Mount. Approximate nearest neighbor queries in fixed dimensions. In *SODA*, 1993.

[2] S. Baker and T. Kanade. Hallucinating faces. In *Automatic Face and Gesture Recognition*, 2000.

[3] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *PAMI*, (9), 2002.

[4] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *SIAM MMS*, (2), 2005.

[5] D. Capel. *Image Mosaicing and Super-Resolution*. Springer–Verlag, 2004.

[6] M. Ebrahimi and E. Vrscay. Solving the inverse problem of image zooming using "self-examples". In *Image Analysis and Recognition*, 2007.

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

[8] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *T-IP*, (10), 2004.

[9] R. Fattal. Image upsampling via imposed edge statistics. In *SIGGRAPH*, 2007.

[10] W. Freeman, E. Pasztor, and O. Carmichael. Learning low-level vision. *IJCV*, (1), 2000.

[11] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *Comp. Graph. Appl.*, (2), 2002.

[12] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP*, (3), 1991.

[13] K. Kim and Y. Kwon. Example-based learning for single-image SR and JPEG artifact removal. *MPI-TR*, (173), 08.

[14] Z. Lin and H. Shum. Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation. *PAMI*, (1), 04.

[15] G. Peyré, S. Bougleux, and L. D. Cohen. Non-local regularization of inverse problems. In *ECCV*, 2008.

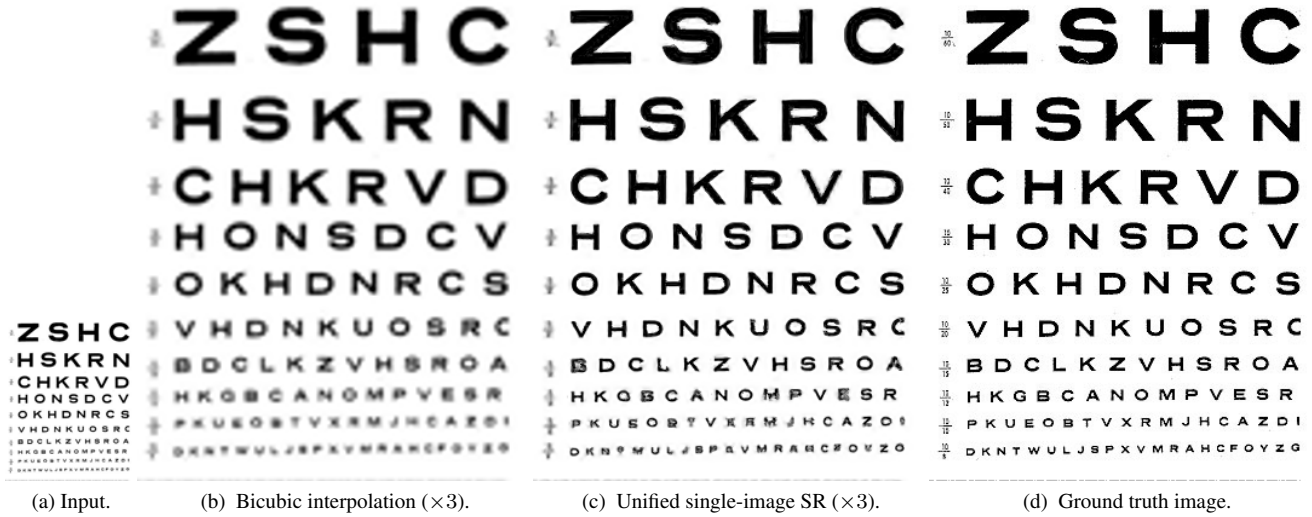| (a) Input. | (b) Bicubic interpolation (×3). | (c) Unified single-image SR (×3). | (d) Ground truth image. |

Figure 6: **Comparison against ground truth:** *The input image (a) was down-scaled (blurred and subsampled) by a factor of 3 from the ground-truth image (d). (b) shows bicubic interpolation of the input image (a) and (c) is the result of our unfied single-image SR algorithm. Note the bottom part of the image. The lines of letters have been recovered quite well due to the existence of cross-scale patch recurrence in those image areas. However, the small digits on the left margin of the image could not be recovered, since their patches recurrence occurs only within the same (input) scale. Thus their resulting unified SR constraints reduce to the "classical" SR constraints (imposed on multiple patches within the input image). The resulting resolution of the digits is better than the bicubic interpolation, but suffers from the inherent limits of classical SR [3, 14].*
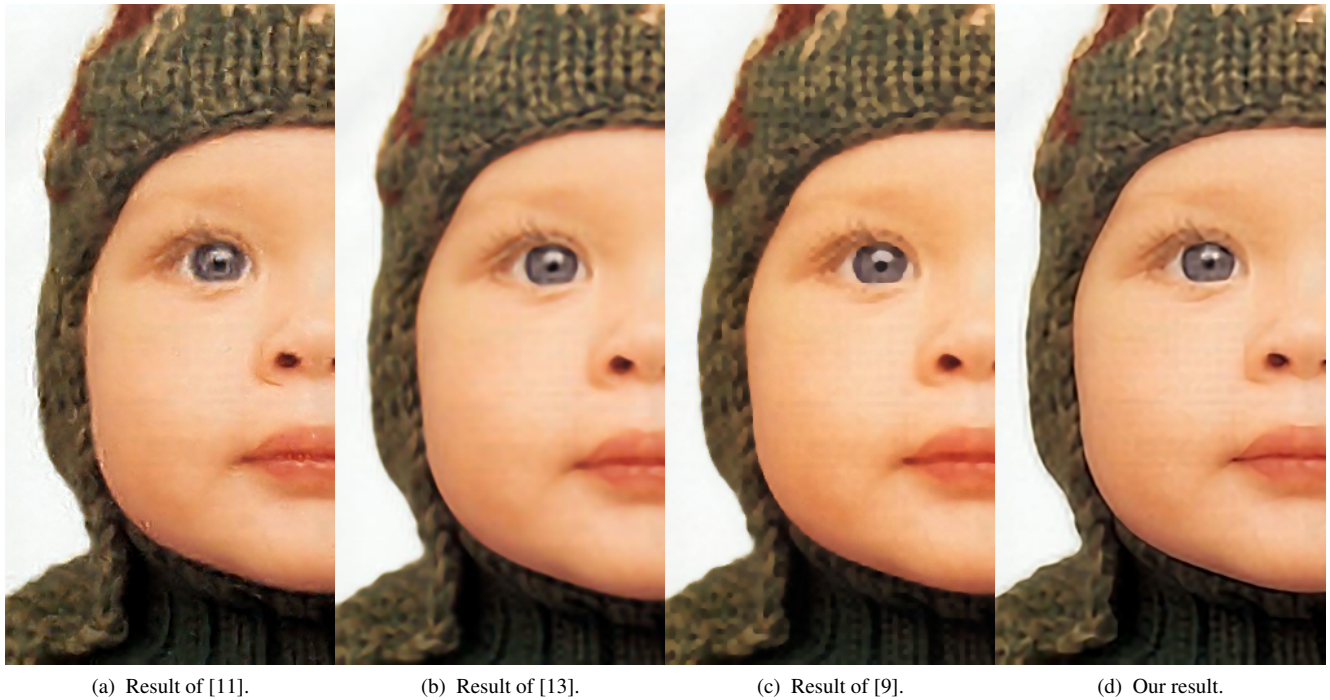

| (a) Result of [11]. | (b) Result of [13]. | (c) Result of [9]. | (d) Our result. |

Figure 7: *Comparison with [11, 13] and [9] of a SR result on the child image (magnified ×4). Our method (d) shows comparable or better results. Note that in contrast to [11, 13] our method does not use any external database of low-res/high-res pairs of patches, nor a parametric learned edge model as in [9]. Images (a) and (c) were taken from [9] and image (b) was taken from [13].*

[16] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *T-IP*, 09.

[17] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani. Summarizing visual data using bidirectional similarity. *CVPR*, 2008.

[18] N. Suetake, M. Sakano, and E. Uchino. Image super-resolution based on local self-similarity. *Opt.Rev.*, (1), 2008.

Figure 8: *Input image, scaled for display (overlay top left), bicubic interpolation ×3 (left) and our SR ×3 result (right).*

Figure 9: *Input image, scaled for display (overlay top left), bicubic interpolation ×3 (left) and our SR ×3 result (right).*

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