Efficient Self-learning for Single Image Upsampling

Nilay Khatri*, Manjunath V. Joshi
DA-IICT, Gandhinagar, India
{khatri_nilay, mv_joshi}@daiict.ac.in

ABSTRACT

Exploiting similarity of patches within multiple resolution versions of an image is often utilized to solve many vision problems. Particularly, for image upsampling, recently, there has been a slew of algorithms exploiting patch repetitions within- and across- different scales of an image, along with some priors to preserve the scene structure of the reconstructed image. One such method, self-learning algorithm [1], uses only one image to achieve high magnification factors. But, as the image resolution increases, the number of patches in dictionary increases dramatically, and makes the reconstruction computationally prohibitive. In this paper, we propose a method that removes the redundancies inherent in large self-learned dictionaries to upsample an image without using any regularization methods or priors, and drastically reduces time complexity. We further prove that any low-variance (low details) patch that does not find any match can be represented as a linear combination of only low-variance patches from dictionary. The same principle applies to high-variance (high details) patches. Images with high scaling factors can be obtained with this method without *any regularization or prior information*, which can be subjected to further regularization with necessary prior(s) to refine the reconstruction.

Keywords

Self-learning, image upsampling, super resolution, dictionary learning

1 INTRODUCTION AND RELATED WORK

Inherent redundancies present within natural images can be utilized to solve many vision problems such as image denoising, texture synthesis [2], image compression [3], super resolution. A small patch [e.g. 3x3] is highly likely to recur within- and across different image resolution scales [4]. Zontak et al. [5] exploited the recurrence of patches for image denoising. The fact that a coarser scale of any noisy image is less noisier than the original noisy image is exploited in their work across different coarser scales to extract a clean patch for almost any noisy patch.

Freeman et al. [6] introduced example-based image upsampling, wherein high-resolution (HR) image patches are generated from the HR image patches learned from external database. Fattal [7] learned a parametric model of edges from a large database of natural images to achieve single image upsampling. The literature concerning single image upsampling using image database is growing rich for sometime now. However, these approaches being patch-based, finding a satisfactory match for all the image patches requires very large external database which increases the computational complexity. According to Ebrahimi and Vrscay [8], the patches learned from image itself form much more *relevant* database than any other external database. Glasner et al. [4] empirically proved the repetition of a small patch across different scales of an image and proposed a single image super resolution algorithm that uses multiple coarser scale versions of an image. Zontak and Irani [9] quantified the recurrence property of natural images and established the superiority of internal image database over external image database. Freedman and Fattal [10] proposed a local self-similarity approach which reduced the search region for the nearest patch and employed non-dyadic filters to achieve high scaling factors. Shan et al. [11] employed feedback-control mechanism to match the upsampled image to that of imaging model through a loop of deconvolution and reconvolution. Khatri and Joshi [1] drew their inspiration from the work of Glasner et al. [4] and proposed a self-learning algorithm. However, they built

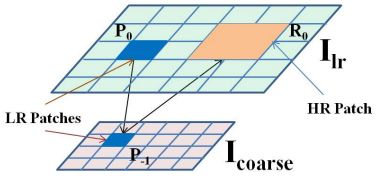


Figure 1: **Self-learning**: illustration for forming LR-HR pairs using LR image and its coarser resolution version.

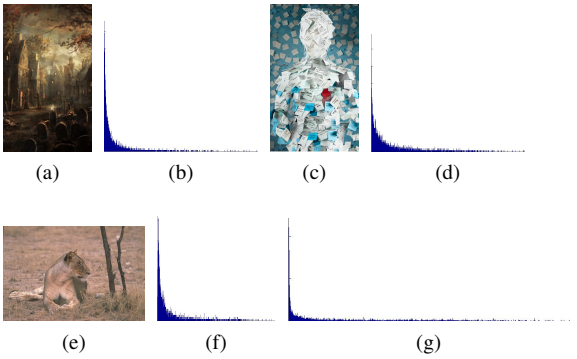


Figure 2: **Variance histogram of LR dictionary patches**: (a) painting, (c) abstract art, (e) natural scene, (b)-(d)-(f) histograms of variances, (g) combined variance histogram of all 1300 images. Majority of the image dictionaries have a long-tailed variance distribution. The concentration of histograms near the origin indicates a large number of low-variance patches. In the plots, x- and y- axes indicate variances and their frequency, respectively.

low-resolution (LR) and HR dictionaries comprising patches having exact match with only one coarser-scale version and generated the remaining patch pairs from learnt patch pairs using l_1-minimization. They further improved upon the solution by introducing Gabor prior which forced the similarity of details between LR and downsampled HR patches at various frequencies.

Majority of the algorithms mentioned above fall in the category of image super resolution (SR). Image SR is an ill-posed problem that requires accurate aliasing and image formation models to correctly reconstruct an HR image. There have been no models so far (to our knowledge) which incorporate aliasing to faithfully super resolve an image. The learning methods employed (e.g. K-SVD) for dictionary reduction are computationally taxing. Also, all the SR algorithms require costly optimization methods that render them ineffective for efficient and fast image upsampling. On the other side, traditional image interpolation methods (bicubic, splines etc.) use only neighborhood pixels to interpolate a pixel value ignoring the redundancies within images. Motivated by these limitations of SR and interpolation algorithms, and drawing our inspiration from self-learning [1], in this paper, we propose a smart, and computationally efficient, *image upsampling* approach that not only works successfully on various kinds of images (e.g. natural scenes, abstract arts, paintings), but for very high upsampling factors as well. The proposed method, because of its simple operations, can be a useful tool for quick upsampling of the images captured on mobile phones. Unlike self-learning that builds a large LR-HR dictionaries (for a reasonably large image, thereby increasing the time complexity), our approach reduces the dictionary size by taking only a small percentage of the total dictionary patches, and further segments it into three smaller dictionaries, namely, low-variance, high-variance, and medium-variance dictionaries. The non-dictionary patches (i.e. image patches that do not find any best match) are obtained from these reduced dictionaries. The proposed approach upsamples images without any prior information, which can be subjected for further refinements.

The rest of the paper is organized as follows: section 2 briefly reviews self-learning approach. Section 3 introduces the proposed approach. In section 4, the proposed algorithm is evaluated experimentally. Section 5 concludes the paper.

2 SELF-LEARNING - A SUMMARY

Self-learning [1] uses two images (LR image, I_{lr}, and its downsampled coarser version, I_{coarse}) to achieve very high magnification factors. Contrary to Glasner et al. approach [4], LR-HR patch pair dictionaries containing only patches of I_{lr} that find an exact match in I_{coarse} are created (also called *best-mapped* patches). Figure 1 graphically explains the process of dictionary generation by self-learning. In Fig. 1, P_0 and P_{-1} form the best-mapped patch pair, and R_0 is the corresponding HR patch. P_0 and R_0 form an LR-HR patch pair. This process of self-learning is repeated to search for all such LR-HR patch pairs. All such patch pairs form LR-HR dictionaries. Non-dictionary patches are obtained as a linear combination of patches from dictionaries using l_1-minimization under sparsity assumption, by formulating the problem to estimate the coefficients as in Eq. 1 [12] .

$$\min_{\mathbf{x}\in\mathbb{R}^N} ||\mathbf{x}||_{l_1}, \text{subj. to } \mathbf{y} = \phi\mathbf{x}; \text{where } ||\mathbf{x}||_{l_1} = \sum_{i=1}^{N} |\mathbf{x}_i| \quad (1)$$

Here, $\mathbf{y}$ and ϕ represent LR patch vectors and the LR dictionary, respectively. The estimated $\mathbf{x}$ is used to reconstruct non-dictionary HR patches from the HR dictionary.

A degradation kernel is estimated for each LR-HR patch pair from the learned dictionaries. These estimated degradation kernels are subsequently used in conjunction with Gabor prior to obtain final upsampled image. All these operation takes hours (see Table 2) to complete and render self-learning ineffective for fast and quick image upsampling.

3 THE PROPOSED FRAMEWORK

Self-learning [1] considers exact matching patches as the candidates for LR-HR dictionaries; though, for large images, these dictionaries become prohibitively large. The proposed approach begins by creating such LR-HR dictionaries from images captured using

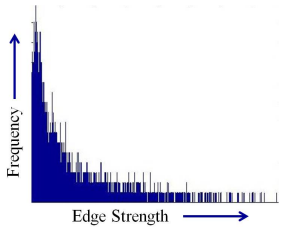


Figure 3: **Histogram of edge strengths of LR dictionary patches**: The shape of edge-strengths distribution resembles that of variance distributions of dictionary patches in Fig. 2. This observation allows reconstruction of a patch of a particular variance class as a linear combination of patches from the same variance class.

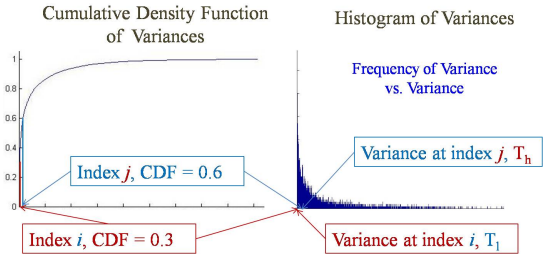


Figure 4: **Selection of thresholds for dictionary division**: A sharp rise in the CDF curve near the origin indicates a very high density of low-variance patches. Indices i and j in CDF are used to select T_l and T_h from histogram of variances. T_l and T_h are variables.

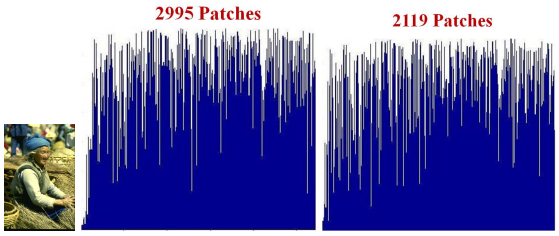
camera in the same fashion as self-learning [1] with spatial resolution of X and $0.5X$. Our experimental observation of variance distribution (histogram of variances) of patches of LR dictionaries of around 1300 images (including paintings, abstract art, and natural images) indicates that the majority of LR dictionary patches have very low-variance. Figure 2 shows some of the sample images along with the variance histogram of their corresponding LR dictionary patches. We use this evidence to create dictionaries.

Dictionary creation

Self-learning employs l_1-minimization to solve for non-dictionary patches, which not only has very high time complexity, but the LR dictionary itself has inherent redundancies of the patches. The proposed approach eliminates these redundancies in a simple and efficient manner. In the proposed approach, instead of using computationally taxing l_1-minimization, we opt for much faster l_2-minimization. In our experiments we have observed that, based on the variance distribution of LR dictionary patches, any low-variance non-dictionary patch can be reconstructed by a linear combination of only low-variance dictionary patches. Similarly, for high- or medium-variance non-dictionary patches, linear combination of patches from their respective dictionaries, is sufficient. This observation can be supported by considering low-variance patches as patches lacking in details and high-variance patches as patches with high-details (edges). For this assumption to be true, the histogram of edge-strengths (gradient value for each dictionary patch) of an LR dictionary patches should resemble that of the histograms in Fig. 2. Figure 3 shows histogram of edge-strengths of LR dictionary patches for one of the sample images. It indeed validates our assumption that high-variance patches contain details and vice versa. As a consequence of this observation, in the proposed approach, the self-learned LR dictionary is divided into three separate dictionaries comprising low-, medium-, and high-variance patches, separately. *Cumulative Density Function* (CDF) of LR dictionary variances (Figure 4) provides a guideline for the division of LR dictionary. It represents the process of threshold selection for LR dictionary division. The variance value in the histogram of variances plot that has a cumulative value of 0.3 with all the preceding variance values is chosen as low-variance threshold, T_l (index i in Fig. 4). Similarly, the variance value in the histogram that has a cumulative value of 0.6 with all the preceding variance values is chosen as high-variance threshold, T_h (index j in Fig. 4). Any LR dictionary patch with variance less than T_l is classified as low-variance patch, and any LR dictionary patch with variance greater than T_h is classified as high-variance patch; while the rest of the dictionary patches are termed as medium-variance patches. It is to be noted that T_l and T_h are variables and can have values other than 0.3 and 0.6, respectively.

Dictionary redundancy removal

Even after the division, low-variance dictionary still exhibits redundancies among its patches. These redundancies are eliminated by a simple *k-nearest neighbor* (k=1) operation. In the proposed approach, for every patch in low-variance dictionary (unlike self-learning that searches for the exact match), its most similar patch is searched within the low-variance dictionary itself. For every such LR patch pair found, the patch with a higher variance value is retained in the dictionary, while the other patch of the pair is discarded from the dictionary. Since medium- and high-variance represent details of the image, this operation is performed only on low-variance dictionary. Figure 5 shows a dictionary variance distribution *before* and *after* redundancy removal. HR dictionary is formed by extracting for every LR patch its corresponding HR patch from the HR dictionary. We thus have three LR dictionaries and their corresponding HR dictionaries. After obtaining LR dictionaries in the aforementioned manner, non-dictionary LR patches are represented as a linear combination of patches from one of the three LR dictionaries depending upon their variances, and the corresponding coefficients for each non-dictionary LR patch are obtained



(a) (b) (c)

Figure 5: **Redundancy elimination for low-variance dictionary**: (a) Input image; (b) Distribution *before* redundancy removal; (c) Distribution *after* redundancy removal. The shape of the distribution is preserved pre- and post- redundancy removal with 30% reduction in low-variance dictionary patches.

using a *least-norm* minimization, as in Eq. 2, where $\mathbf{y}$ is an LR patch vector; L_{lv}, L_{mv}, L_{hv} are low-variance, medium-variance and high-variance LR dictionaries, respectively; $\mathbf{x}$ is a coefficient vector. Equation 2 represents an *under-determined* system of equations. Hence, there is no unique solution in the least-square sense.

$$min.\|\mathbf{x}\|_2;\ \text{s.t. } \mathbf{y} = \begin{cases} L_{lv}\mathbf{x} & \text{if} Var(\mathbf{y}) \leq T_l \\ L_{mv}\mathbf{x} & \text{if} T_l < Var(\mathbf{y}) \leq T_h \\ L_{hv}\mathbf{x} & \text{if } Var(\mathbf{y}) > T_h \end{cases} \quad (2)$$

For every LR patch, its corresponding HR patch (based on its variance) can be reconstructed from one of the three HR dictionaries and the estimated coefficient vector $\mathbf{x}$, as shown in Eq. 3. Here $\mathbf{z}$ is a reconstructed HR patch; H indicates HR dictionary and the subscript indicates the dictionary class.

$$\mathbf{z} = \begin{cases} H_{lv}\mathbf{x} & \text{if } Var(\mathbf{y}) \leq T_l \\ H_{mv}\mathbf{x} & \text{if } T_l < Var(\mathbf{y}) \leq T_h \\ H_{hv}\mathbf{x} & \text{if } Var(\mathbf{y}) > T_h \end{cases} \quad (3)$$

The histogram plot of the difference of variances between LR and their corresponding HR patches in Figure 6 validates the use of estimated coefficients from LR patches to reconstruct corresponding HR patches. A spike at the origin hints at the sameness of low-variance and medium-variance patches; while the long tail of the distribution points at the similarity between the high-variance LR and HR patches. Most of the image dictionaries exhibit such behavior. The division of dictionary into smaller dictionaries, along with the use of least-norm, makes the proposed approach smarter and computationally efficient from self-learning [1]. It is to be noted that the procedure explained above upsamples the image by a factor of 2. Higher magnification factors can be achieved by considering the current upsampled image as an LR image and repeating the above procedure.

4 EXPERIMENTAL RESULTS

In this section, the proposed approach is evaluated both, perceptually and quantitatively, for upsampling factors

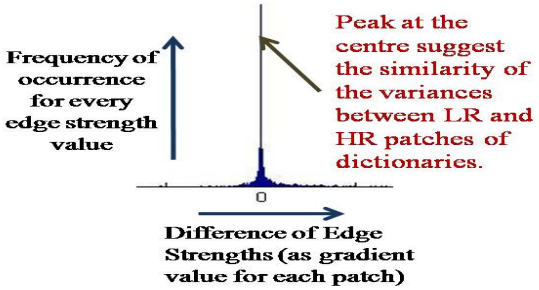


Figure 6: **Distribution of difference of variances between LR-HR patches of dictionary**: Gaussian shape of the distribution along with a long tail indicates the similarity of variances between LR and their corresponding HR patches.

up to 8. Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE) [13] is used for quantitative evaluation. We work in YC_bC_r color space. Only the luminance (Y) channel is upsampled, while C_b and C_r channels are bicubically interpolated. The images are either downloaded from the Internet or captured using a mobile phone (XOLO A500, 5MP camera). I_{coarse} is obtained from I_{lr} by a blur & downsample process. Patches of size 3x3 are used. For self-learning [1], nearest neighbor search is performed using *Approximate Nearest Neighbor* [14]. The threshold values, T_l and T_h, for dictionary division, after many trials, are set to *0.3* and *0.6*, respectively.

Perceptual and quantitative evaluation

Figures 7−10 show the results for upsampling factor of 2. Self-learned dictionary of Fig. 7(c) (top row) comprises around 28,000 patches. The combined dictionary using our approach comprises 4,744 patches (a reduction of 83%). It would be of interest to note that only a small percentage of the patches in self-learned dictionary are used in our approach for reconstruction (see figure captions for relevant statistics). Insets in figures prove the superiority of our approach over self-learning, and its comparability to Shan et al. [11]. While for edges with smooth nearby regions Shan et al. [11] produces visually pleasant results (insets in Figs. 7b-top row, 9d, 11a), for images with very fine textures, it seems to produce a cartoonish feel (insets in Figs. 8b, 10d).

Figures 11 and 12 present the results for upsampling factors of 4 and 8, respectively. For $4X$, reduction of 94% is achieved in Fig. 11(c) (113,218 to 6,602 patches); while Fig. 11(f) has 93% reduction (108,413 to 7,867 patches). Final spatial resolution of $8X$ images is 3736x2492. For both the results of $8X$ in Figure 12, the dictionary reduction of 99% is achieved. This means only 1% of the total dictionary patches are sufficient for $8X$ reconstruction, and thus establishing the computational efficiency of our approach for higher

scaling factors. Figure 13 shows the upsampling results for images captured using mobile phones. Quantitative evaluation is done with Bicubic interpolation, along with image upsampling methods. Table 1 lists the obtained BRISQUE scores for Bicubic, self-learning [1], Shan et al. [11] and proposed approach. Kindly see the supplementary material for more upsampling results of different types of images. Figures 14-15 show some more results for $2X$ and $4X$. All the input images had a spatial resolution of $\sim$1-1.5MP. Shan et al. [11] has the tendency either to oversharp the edges or oversmooth the fine details.

A word on time complexity

The proposed approach does not use any type of optimization methods to upsample the image. So, the processes that incur significant time cost in the proposed approach are, nearest-neighbor search, and estimation of coefficient vector $\mathbf{x}$. The proposed approach is implemented in MATLAB on a system with 4GB RAM and 2.50 GHz intel-i5 processor system, and although it is an unoptimized and proof-of-concept version, a fair idea about its efficiency can be had. MATLAB in-built optimized functions were used wherever possible to speed up the upsampling process. For nearest-neighbor search an optimized version of ANN was used. For l_1-minimization used in self-learning, l_1-magic [15] package by Justin Romberg was used. Results of Shan et al. [11] were produced by an optimized executable obtained from their online resources and the best result was chosen for the comparison. Table 2 compares the proposed approach with self-learning [1] and Shan et al. [11]. While clearly outrunning self-learning [1], the proposed approach performs comparably with Shan et al. [11] indicating at its viability for practical implementation when used in its optimized form.

5 CONCLUSION

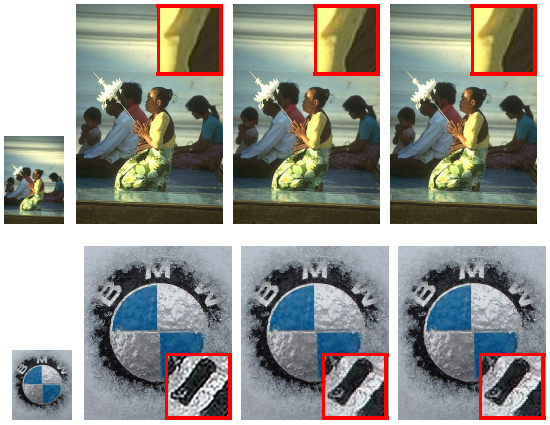
We have proposed a single image upscaling approach which uses only a small percentage of the total dictionary patches to reconstruct high resolution images in an efficient manner. The time performance of the proposed approach hints at its potential ability to be utilized in (an optimized form) the current handheld devices.

Acknowledgement This research has been carried out as a part of the project *Indian Digital Heritage (IDH) - Hampi*, sponsored by Department of Science and Technology, Govt. of India (Grant No: NRDMS/11/1586/2009/Phase-II).

6 REFERENCES

[1] N. Khatri and M. V. Joshi, "Image super-resolution: Use of self-learning and gabor prior," in *ACCV*, 2012.

[2] S. Lefebvre and H. Hoppe, "Appearance-space texture synthesis," *ACM Trans. Graph.*, vol. 25, no. 3, Jul. 2006.



(a) LR (b) Shan et al.[11] (c) Self-learn. [1] (d) Our Approach

Figure 7: Upsampling by **2X** for the Internet images [Woman (480x320) and Logo (416x352)]: (a) LR images; (b) Shan et al. [11]; (c) Self-learning; (d) Our Result. Self-learned dictionaries in (c) are of size $28,000$ (top-row) and $32,000$ (bottom row), respectively. Our approach has dictionary size of $4,744$ (83% reduction) and $6,096$ (81% reduction), respectively.

[3] E. R. Vrscay, "From fractal image compression to fractal-based methods in mathematics," in *Fractals in Multimedia*. Springer New York, 2002, vol. 132.

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

[5] M. Zontak, I. Mosseri, and M. Irani, "Separating signal from noise using patch recurrence across scales," in *CVPR*, 2013.

[6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, 2002.

[7] R. Fattal, "Image upsampling via imposed edge statistics," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.

[8] M. Ebrahimi and E. R. Vrscay, "Solving the inverse problem of image zooming using "self-examples"," in *ICIAR*, 2007.

[9] M. Zontak and M. Irani, "Internal statistics of a single natural image," in *CVPR*, 2011.

[10] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Trans. Graph.*, 2011.

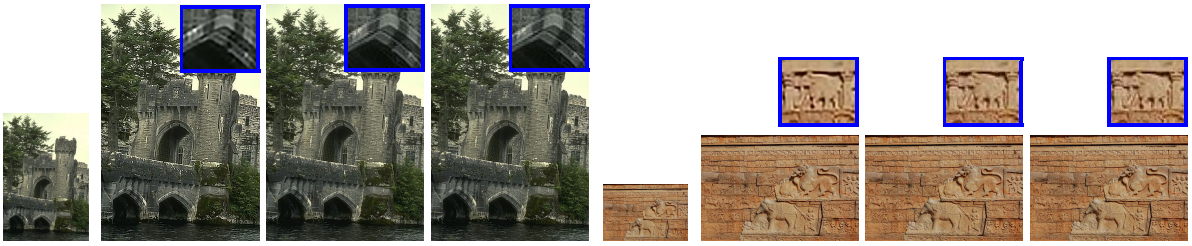
[11] Q. Shan, Z. Li, J. Jia, and C.-K. Tang, "Fast image/video upsampling," in *ACM SIGGRAPH Asia 2008 Papers*, ser. SIGGRAPH Asia '08, 2008, pp. 153:1–153:7.

[12] E. Candes, "The restricted isometry property and its implications for compressed sensing," *C. R. Math*, 2008.

[13] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE T-IP*, vol. 21, no. 12, 2012.

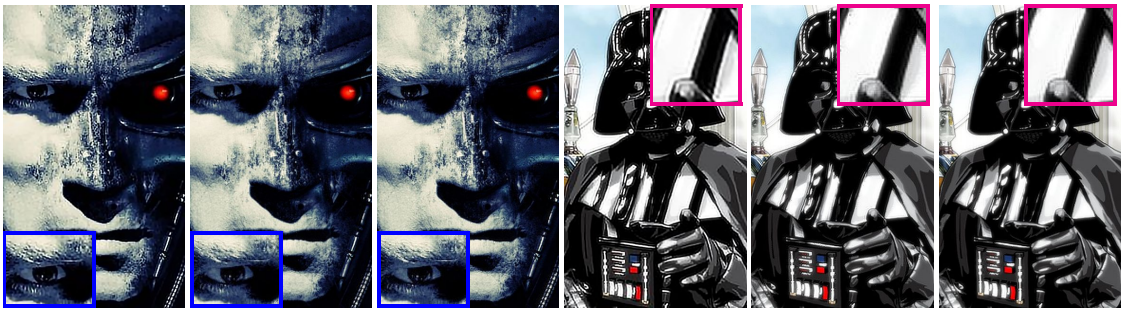
[14] S. Arya and D. Mount, "Approximate nearest neighbor queries in fixed dimensions," in *SODA*, 1993.

[15] J. Romberg, "ℓ_1-magic." [Online]. Available: http://users.ece.gatech.edu/~justin/l1magic/#code



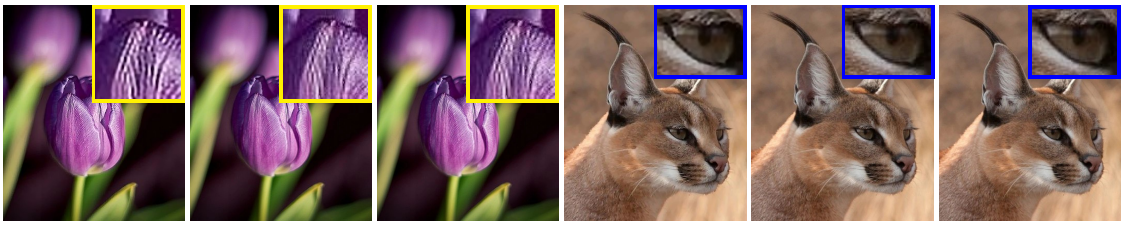
(a) LR (b) Shan et al.[11] (c) Self-learn. [1] (d) Our Approach (e) LR (f) Shan et al.[11] (g) Self-learn. [1] (h) Our Approach

Figure 8: Upsampling by **2X** for architectural images [Castle (480x320) and Mural (480x720)]: (a), (e) LR images; (b), (f) Shan et al. [11]; (c), (g) Self-learning; (d), (h) Our Result. Self-learned dictionaries in (c) and (g) are of size $30,680$ and $76,797$, respectively. Our approach has dictionary size of $5,640$ (82% reduction) and $7,081$ (91% reduction), respectively.



(a) Shan et al.[11] (b) Self-learning [1] (c) Our Approach (d) Shan et al.[11] (e) Self-learning [1] (f) Our Approach

Figure 9: Upsampling by **2X** for abstract/paintings [Arnie (320x192) and Vader (320x192)]: (a), (d) Shan et al. [11]; (b), (e) Self-learning; (c), (f) Our Result. Self-learned dictionaries in (b) and (e) are of size $11,581$ and $9,504$, respectively. Our approach has dictionary size of 982 and 794 (92% reduction for both the images), respectively.



(a) Shan et al.[11] (b) Self-learning [1] (c) Our Approach (d) Shan et al.[11] (e) Self-learning [1] (f) Our Approach

Figure 10: Upsampling by **2X** for natural images [Tulip (416x352) and Cat (416x352)]: (a), (d) Shan et al. [11]; (b), (e) Self-learning; (c), (f) Our Result. Self-learned dictionaries in (b) and (e) are of size $17,993$ and $25,886$, respectively. Our approach has dictionary size of $4,314$ (76% reduction) and $6,106$ (76% reduction), respectively.

(a) Shan et al.[11] (b) Self-learning [1] (c) Our Approach (d) Shan et al.[11] (e) Self-learning [1] (f) Our Approach

Figure 11: Upsampling by **4X** [Baboon (320x480) and Lady (480x312)]: (a), (d) Shan et al. [11]; (b), (e) Self-learning; (c), (f) Our Result.

(a) Shan et al.[11] (b) Self-learning [1] (c) Our Approach (d) Shan et al.[11] (e) Self-learning [1] (f) Our Approach

Figure 12: Upsampling by **8X** [Lady (480x320) and Boat (466x310)]: (a), (d) Shan et al. [11]; (b), (e) Self-learning; (c), (f) Our Result. For (c), the dictionary is reduced from $356,175$ to $7,977$ patches. For (f), the dictionary is reduced from $419,623$ to $2,464$ patches.

(a) Shan et al.[11] (b) Our Approach (c) Shan et al.[11] (d) Our Approach (e) Shan et al.[11] (f) Our Approach

Figure 13: Upsampling by **2X** for mobile phone images: (a), (c), (e) Shan et al. [11]; (b), (d), (f) Our Result. Shan et al. [11] images have BRISQUE score of 40.72, 42.20 and 46.88, respectively. In comparison, our approach has BRISQUE scores of 54.78, 42.33 and 47.03, respectively. For all three images, 77% of reduction in dictionaries is achieved. Input images are of resolution 480x640, 640x480, and 640x480, respectively.

	2X								**4X**		**8X**		**Avg. Score**
	Woman	Logo	Castle	Mural	Arnie	Vader	Tulip	Cat	Baboon	Lady	Lady	Boat	
Bicubic	26.02	37.04	43.80	41.64	19.63	22.00	59.02	36.74	55.48	38.31	45.82	52.99	39.87
Self-learning [1]	26.14	60.74	43.06	43.63	24.02	26.66	62.04	37.85	55.48	46.35	55.27	54.78	44.67
Shan et al. [11]	34.50	61.65	51.35	47.92	27.38	34.40	68.87	38.43	60.19	41.68	66.67	68.11	50.09
Our Approach	34.21	60.79	**51.99**	47.81	**28.51**	**38.80**	68.82	38.13	58.83	**49.97**	**68.50**	67.89	**51.19**

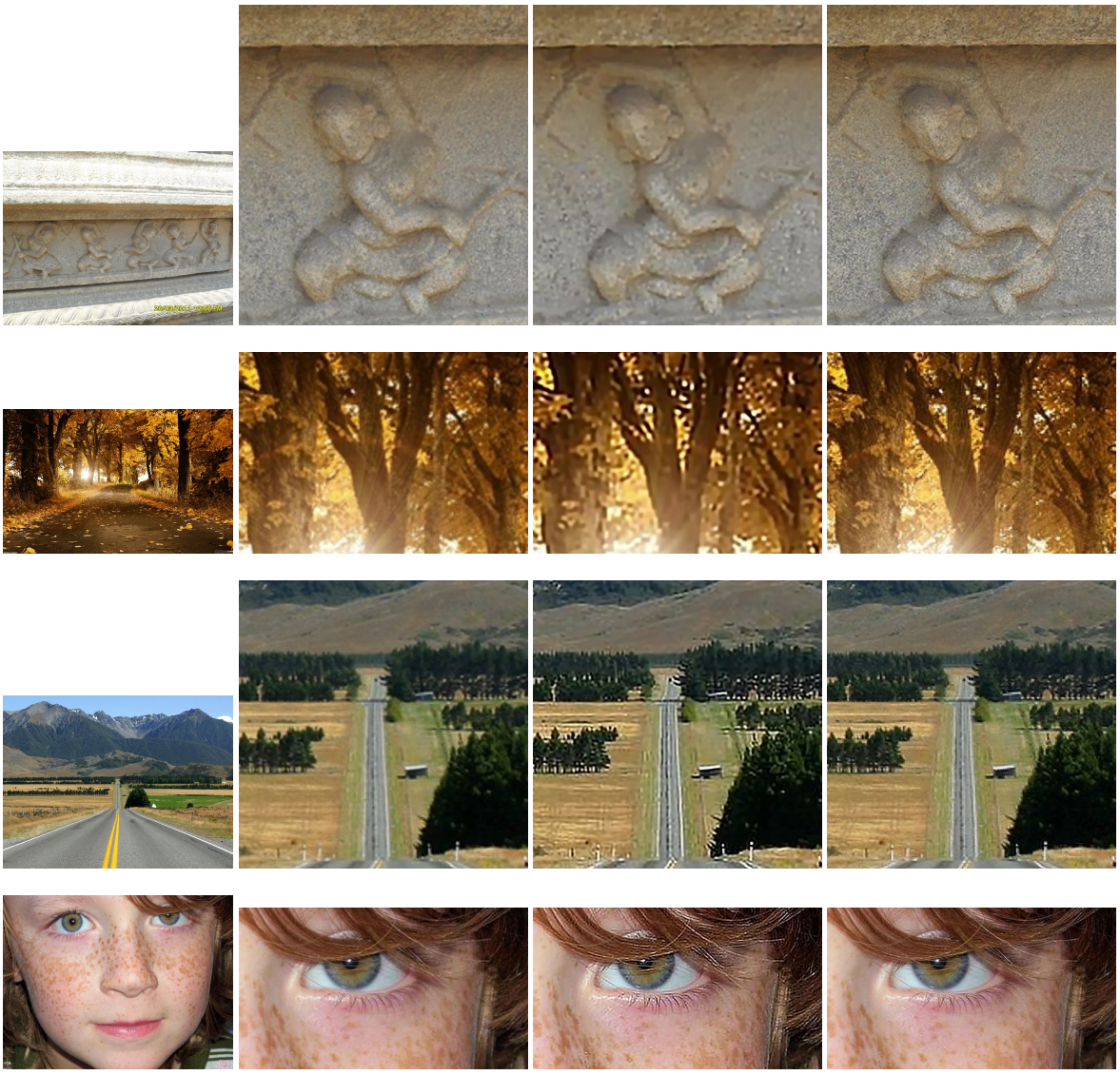
Table 1: BRISQUE scores for the results in Figs. 7−12. The average BRISUQE score is better for our approach than those of self-learning [1] and Shan et al. [11]. (Higher score indicates better image quality.)

	2X								**4X**		**8X**	
	Woman	Logo	Castle	Mural	Arnie	Vader	Tulip	Cat	Ba-boon	Lady	Lady	Boat
Shan et al. [11]	15.00	12.82	23.89	46.06	7.11	6.58	48.14	13.18	70.18	127.83	348.49	342.52
Self-learning [1]	144.57	150.00	136.67	183.33	155.85	109.95	123.88	96.98	3.58 Hrs.	3.79 Hrs.	7.88 Hrs.	3.79 Hrs.
Our Approach	10.69	67.45	72.53	153.74	16.37	14.50	61.52	67.12	144.32	168.86	340.52	380.12

Table 2: Time complexity (in *seconds*) for the results in Figs. 7−12. Note that for 8*X* the time complexity performance of our approach is comparable to that of Shan et al. [11].

(a) Shan et al.[11] (b) Bicubic Interpolation (c) Our Approach (d) Shan et al.[11] (e) Bicubic Interpolation (f) Our Approach

Figure 14: Upsampling by **4X**. Input images are of size 468x312 and 480x720, respectively



(a) LR (b) Bicubic Interpolation (c) Shan et al.[11] (d) Our Approach

Figure 15: Upsampling by **2X**: Each LR image has a resolution of $\sim$ 1-1.5 MP (630x840, 600x960, 538x717, and 674x899, respectively). Shan et al. [11] often produces results with over-sharpened edges (3^{nd} row) or fine details smoothened (2^{nd} row)

.