

Image Super Resolution Using Multiple Kernel Learning

Dharani Sampath, Sampath Kumar.P

Abstract— Learning-based approaches for image super-resolution (SR) have attracted the attention of researchers in the past few years. We present a novel self-learning approach with multiple kernel learning for adaptive kernel selection for SR. The Multiple Kernel Learning is theoretically and technically very attractive, because it learns the kernel weights and the classifier simultaneously based on the margin criterion. With theoretical supports of kernel matching search method and Optimization approach (Gradient) are proposed in SR framework learns and selects the optimal Kernel ridge regression model when producing an SR image, which results in the minimum SR reconstruction error. Evaluate this method on a variety of images, and obtain very promising SR results. In most cases, this method quantitatively and qualitatively outperforms bi-cubic interpolation and state-of-the-art learning based SR approaches.

Index Terms— Kernel ridge regression, Multiple kernel learning, Self-learning, Super-resolution, Sparse Representation.

1 INTRODUCTION

Super Resolution (SR) has been an active research topic in the areas of image processing and computer vision. It is a process to produce a high-resolution (HR) image from one or several low-resolution (LR) images. Conventional methods are based on the reconstruction of multiple LR images, and they approach SR as solving an inverse problem, i.e., they recover the HR image as a linear operation of multiple LR patches. Recently, learning-based SR approaches which focus on modeling the relationship between training low and high-resolution images have also attracted researchers, while the existence of the aforementioned relationship is typically seen in natural images [1], [2]. However, the difficulty of learning-based SR methods lies on the selection of proper training data and proper learning models for SR from an unseen target image.

In machine learning, support vector regression (SVR) [3] is considered as an extension of support vector machine (SVM), which exhibits excellent generalization ability in predicting functional outputs without any prior knowledge or assumption on the training data (e.g., data distribution, etc.). SVR is capable of fitting data via either linear or nonlinear mapping, and the use of SVR has been applied in applications of data mining, bioinformatics, financial forecasting, etc. Previously, SVR has been shown to address SR problems in [1], [4]; however, these SVR-based SR approaches require the collection of training low and high-resolution image pairs in advance, and this might limit their practical uses. In this paper, we propose a self-learning framework for SR. We not only present quantitative and qualitative SR results to support our method, we will also provide theoretical backgrounds to verify the effectiveness of our learning framework.

- Dharani Sampath is currently pursuing masters of Engineering in Computer Science and Engineering from Sri Krishna College of Engineering and Technology, Coimbatore, India, PH-9677396888, E-mail: dharanis012@gmail.com
- P.Sampath Kumar is currently working as a Associate Professor in a Department of Computer Science and Engineering from Sri Krishna College of Engineering and Technology, Coimbatore, India, PH-9600955088, Email:sampathkumarp@skcet. ac.in

The remainder of this paper is organized as follows. Prior SR works are discussed in Section 2. Section 3 details the novel self-learning approach with multiple kernel learning for adaptive kernel selection for SR. Section 4 provides empirical results on a variety of images with different magnification factors, including comparisons with several SR methods. Finally, Section 5 concludes this paper.

2 RELATED WORK

2.1 Reconstruction-Based SR

Typically, reconstruction-based SR algorithms require image patches from one or several images (frames) when synthesizing the SR output. This is achieved by registration and alignment of multiple LR image patches of the same scene with sub-pixel level accuracy [5]–[7]. For single-image reconstruction-based SR methods, one needs to exploit self similarity of patches within the target LR image. With this property, one can thus synthesize each patch of the SR image by similar patches in the LR version. However, reconstruction-based methods are known to suffer from ill-conditioned image registration and inappropriate blurring operator assumptions (due to an insufficient number of LR images) [8]. Moreover, when an image does not provide sufficient patch self-similarity, single-image reconstruction based methods are not able to produce satisfying SR results [9]. Although some regularization based approaches [5], [7], [10] were proposed to alleviate the above problems, their SR results will still be degraded if only a limited number of low-resolution images/patches are available or if a larger image magnification factor is needed. According to [8], [11], the magnification factor of reconstruction-based SR approaches is limited to be less than 2 for practical applications.

A recent approach proposed in [12] alleviates this limitation by learning image prior models via kernel principal component analysis from multiple image frames. Since single-image SR does not require multiple LR images as inputs, it

attracts the interest from researchers and engineers due to practical applications. As discussed above, methods assuming the existence of image patch self-similarity need to search for similar patches from an input image when synthesizing the SR output. However, the assumptions of self-similarity might not always hold, and the associated SR performance varies with the similarity between different categories of image patches. The nonlocal means (NLM) method [13] is one of the representatives which advocates such a property in image related applications.

2.2 Learning-Based SR

In the past few years, much attention has been directed to learning (or example) based SR methods (e.g., [1], [14],[15]), which can be considered as single image SR approaches utilizing the information learned/observed from training imagedata. With the aid of training data consisting of low and high resolution image pairs, learning-based methods focus on modeling the relationship between the images with different resolutions (by observing priors of specific image or context categories). For example, Chang et al. [15] applied the technique of locally linear embedding (LLE) for SR purposes. They collected a training data set with multiple low and high-resolution image pairs. For each patch in an input LR image, they proposed to search for similar patches from LR training images, and they used the corresponding training HR patches to linearly reconstruct the SR output (using the weights determined by LLE). Ni et al. [1] proposed to use support vector regression (SVR) to fit LR image patches and the pixel value of the corresponding HR images in spatial and DCT domains. It is not surprisingly that, however, the performance of typical learning-based methods varies significantly on the training data collected. As a result, in order to achieve better SR results, one needs to carefully/manually select the training data. In such cases, the computation complexity of training and difficulty of training data selection should be both taken into consideration.

Recently, Glasner et al. [2] proposed to integrate both classical and example-based SR approaches for single image SR. Instead of collecting training image data beforehand, they searched for similar image patches across multiple down-scaled versions of the image of interest. It is worth noting that this single image SR method advocates the reoccurrence of similar patches across scales in natural images, so that their approach simply down grades the resolution of the input image and perform example-based SR. In other words, once similar patches are found in different scaled versions, classical SR methods such as [7], [16], [17] can be applied to synthesize the SR output. Although very promising SR examples were shown in [2], there is no guarantee that self-similarity always exists within or across image scales, and thus this prohibits the generalization of their SR framework for practical problems. Moreover, it is not clear what is the preferable magnification factor when applying their approach (SR images with different magnification factors were presented in [2]).

2.3 Sparse Representation for SR

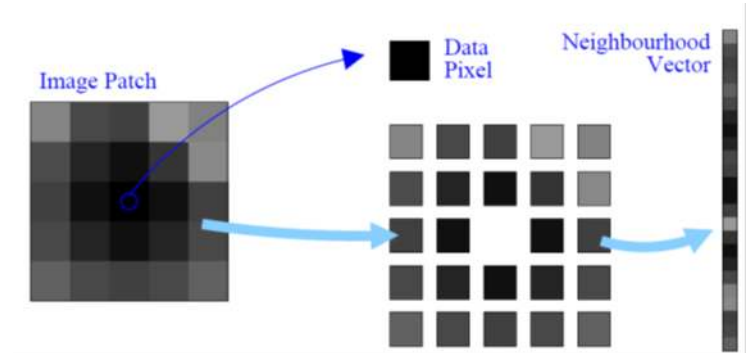
Originally applied to signal recovery, sparse coding [18] has shown its success in image related applications such as image denoising [19], and it was first applied to SR by Yang et

al.[20], [21]. They considered the image patch from HR images as a sparse representation with respect to an over-complete dictionary composed of signal-atoms. They suggested that, under mild conditions, the sparse representation of high-resolution images can be recovered from the low-resolution image patches [20], [21]. They used a small set of randomly chosen image patches for training, and implied that their SR method only applies to images with similar statistical nature. Kim and Kwon [22], [23] proposed an example-based single image SR for learning the mapping function between the low and high-resolution images by using sparse regression and natural image priors. However, blurring and ringing effects near the edges exist in their SR results, and additional post-processing techniques are still needed to alleviate this problem. Recently, Yang et al. [24] extended the framework of [2]. Based on the assumption of image patch self-similarity, they concatenated high and low-resolution image pairs from the image pyramid and jointly learned their sparse representation. When super-resolve an LR input patch, they searched for similar patches from the image pyramid and use the associated sparse representation (the HR part) to predict its final SR version.

3. THE PROPOSED SELF-LEARNING APPROACH TO SINGLE IMAGE SUPER-RESOLUTION

3.1 Image patch Categorization

In a First step, the input image size is up-scaled into Magnitude factor. The up-scaling can be done by bi-cubic interpolation. To alleviate edge artifacts due to the bi-cubic interpolation, the input image is padded by four pixels on each side by replicating border pixels. The Image patch-based approach (developed at the same time as the texture-based prior presented here) for image-based rendering. For their application the likelihood function was multi-modal, and a prior was needed to help ensure consistency across the output image. For our application, a small image patch around each high-resolution image pixel is selected. We can learn a distribution for this central pixel's intensity value by examining the values at the centers of similar patches from other images.



The central pixel is held back, so that when a match is made between this neighbourhood vector and a new trial patch, the central pixel value can be returned as a prototype for what value we might expect to find associated with the centre of the new patch. The patches in the image are normalized to sum to unity, and centre weighted as in 2-dimensional Gaussian kernel. For simplicity, the patches are always chosen to be squares with an odd number of pixels to a side, and the Gaussian kernel standard deviations are chosen to be equal to the floor of half the width of the square, i.e. a 3x3 patch would use a Gaussian with a standard deviation of 3 pixels.

3.2 Sparse Representation

Sparse representations are representations that account for most or all information of a signal with a linear combination of a small number of elementary signals called atoms. Super-resolution as sparse representation in dictionary of raw image patches. The technique of finding a representation with a small number of significant coefficients is often referred to as Sparse Coding. Decoding merely requires the summation of the relevant atoms, appropriately weighted, however, unlike a transform coder with its invertible transform, the generation of the sparse representation with an over-complete dictionary is non-trivial. The image patch self-similarity, they concatenated high and low-resolution image pairs from the image pyramid and jointly learned their sparse representation. The super-resolve an LR input patch, they searched for similar patches from the image pyramid and use the associated sparse representation (the HR part) to predict its final SR version.

3.3 Support Vector Regression

Support vector regression (SVR) is an extension of support vector machine, which is able to fit the data in a high-dimensional feature space without assumptions of data distribution. Similar to SVM, the generalization ability makes the SVR very powerful in predicting unknown outputs, and the use of SVR has been shown to produce effective SR outputs. The observed SVR will be applied to predict the final SR output for a test LR input. Our work in first synthesizes the HR version from the test input using bi-cubic interpolation. To refine this HR image into the final SR output, we derive the sparse representation of each patch and update the centered pixel value for each using the learned SVR models. While this refinement process exhibits impressive capability in producing SR images, its need to collect training LR/HR image data in advance is not desirable.

3.4 Kernel ridge regression of Optimization model

A parameterized combination functions and learns the parameters by solving an optimization problem. This optimization can be integrated to a kernel-based learner or formulated as a different mathematical model for obtaining only the combination parameters. In order to incorporate SVR with KRR (Kernel Ridge Regression), mixture probability should be formulated with SVR outputs Probabilistic output from SVR was done in phase I researches, but they didn't consider optimization mixture probability. To find mixture probability with SVR in this research distance from Gradient, histograms and interpolation were used in this module.

A Gradient Pyramid approach is obtained by applying a set of 4 directional gradient filters (horizontal, vertical and 2 diagonal) to the Gaussian Pyramid at each level. At each level, these 4 directional Gradient Pyramids are combined together to obtain a combined gradient pyramid that is similar to a Laplacian pyramid. The gradient pyramid solution is therefore the same as the image angle using the Laplacian pyramid method except replacing the Laplacian pyramid with the combined Gradient Pyramid.

A good solution to suppress halos is to apply the scene gradients to adjust the gradient of the synthesized SDR image. The scene gradient information is adaptively captured by setting the different exposure levels, i.e., the scene gradients are captured through the local adaptation to the scene luminance for an window $M \times M$ centered at (x, y) . Technically, the scene gradient of a point is reflected by the gradient that is perceivable by human eyes, called visible gradient, and that can be measured by counting the number of visible differences of luminance's between neighboring pixels in the window.

To compute the quantity of the visible gradient $\psi(x, y)$ by

$$\psi(x, y) = \sum_{i=x-\frac{M}{2}}^{x+\frac{M}{2}} \sum_{j=y-\frac{M}{2}}^{y+\frac{M}{2}} T(|c(I_H(x, y)) \cdot VMAX(\nabla I_H(x, y))|) \quad (1)$$

These exposure levels lead to different gradient magnitudes because the gradient magnitude depends on the image luminances and the image luminance depends on the exposure level. The scene gradient extraction is a process to find gradient $G(x, y)$ which maximizes the quantity of the visible gradient,

$$G(x, y) = \arg \max_{\nabla I_H(x, y)} \psi(x, y) \quad (2)$$

In this gradient extraction, the pixels positions are estimated along with x-coordinate and y-coordinate. In vector calculus, the gradient of a scalar field is a vector field that points in the direction of the greatest increase rate in the scalar, and in the magnitude. The variation in space of any quantity can be represented (e.g. graphically) by a slope in general. The gradient represents the steepness and direction of the slope.

4. EXPERIMENTAL RESULTS

For companding color images we first convert RGB to the HSV space. The value (V) is then run through the companding loop and a compressed V is obtained when the iterations stop. This compressed V is combined with the original hue (H) and the original saturation divided by a factor, and converted back to RGB to get the compressed color image. This is the same as what we did for color HDR image compression. Similarly when we're going to expand a compressed color image up to one-step range expansion is done on its V channel. The saturation is multiplied by the same the hue is kept the same, and they are combined with the expanded V to get the HDR color

image back.

The assessment method is the recently proposed objective assessment especially designed for tone mapped images. It combines a multi-scale structural fidelity measure and a measure of image naturalness. The structural fidelity measure is a full-reference assessment based on the structural similarity (SSIM) index, and the naturalness measure is a no-reference assessment based on statistics of good-quality natural images. This method provides a single quality score of an entire image. The combined single quality is represented by 'Quality' in this

study.

The test image is matched with matched database to identify high frequency regions. The PSNR fraction measure of quality of reconstruction of lossy compression codecs (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression.

$$PSNR = 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE)$$

$$MSE = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

TABLE 1
 STRUCTURAL SIMILARITY RATIO FOR DIFFERENT METHODS COMPARISON

Images		Yang et al	LLE	Proposed System
Child	SSIM	0.8392	0.8400	0.9165
fruit	SSIM	0.9751	0.9766	0.9899
lena	SSIM	0.9321	0.9653	0.9997

TABLE 2
 COMPARISONS OF SR IMAGES PRODUCED BY DIFFERENT METHODS WITH A MAGNIFICATION FACTOR OF 2.

	Boat	cars	skyview	lena	Child	Fruit	station	airplane	tree	Susan
Bicubic	30.21	30.87	27.24	34.60	27.45	31.97	25.12	25.64	26.87	33.74
LLE	25.61	26.87	22.73	28.63	24.05	27.52	22.46	23.24	22.16	28.61
Yang et al.	29.66	30.34	27.84	32.92	27.83	32.09	25.71	25.25	27.59	33.44
Wang et al.	30.36	30.89	27.52	34.54	27.60	32.48	25.44	25.91	27.50	34.38
Glasner et al.	31.23	30.10	28.06	35.62	27.92	35.39	25.61	26.35	26.28	33.86
SVR	30.46	31.32	27.56	35.34	42.08	33.06	25.63	25.93	27.22	33.77
Ours	31.21	32.52	30.20	45.94	48.05	32.98	30.76	26.22	28.85	35.26



Fig 1.

Example SR results (with a magnification factor 4) and the corresponding PSNR values. Top row: Ground truth image, SR image produced by Genuine Fractals (software available at www.ononesoftware.com), Freeman et al. [14] (PSNR: 21.34), Fattal et al. [31] (PSNR: 21.36), Wang et al. [4] (PSNR: 23.86). Bottom row: Kim and Kwon [22], [23] (PSNR: 20.93), Glasner et al. [2] (PSNR:

20.94), LLE [15] (PSNR: 29.8), Yang et al. [21] (PSNR: 24.43), SVR(PSNR: 31.14)and ours(PSNR:48.05).



Fig 2.

Example SR results (with a magnification factor 4) and their PSNR values. Top row: ground truth HR image, bicubic interpolation (PSNR: 28.19), LLE [15] (PSNR: 26.85). Bottom row: Yang et al. [21] (PSNR: 26.12), Glasner et al. [2] (PSNR: 24.87), SVR (PSNR:28.66)and ours (PSNR: 32.63s). Note that nearest neighbor interpolation is applied to degrade the image resolution in this example.



Fig 3.

Example SR results (with a magnification factor 2) and their PSNR values. Top row: ground truth HR image, bicubic interpolation (PSNR: 29.70), LLE [15] (PSNR: 27.43). Bottom row: Yang et al. [21] (PSNR: 28.86), Glasner et al. [2] (PSNR: 28.81),SVR (PSNR: 32.15) and ours (PSNR: 32.63). Note that nearest neighbor interpolation is applied to degrade the image resolution in this example.

5. CONCLUSION

This paper proposed a novel in-scale self-learning framework

for single image SR and the system is tested accurately with all testing methods. Since the project is heavily used to view the

detail about image edge preserving and concerned with the color features. This project is highly concerned in the organization and it has been successfully implemented.

The analysis-synthesis sub band architectures and smooth gain control, gives good range compression without disturbing halos. We describe some simple implementations of sub band range compression, and show that the results are competitive with the leading techniques

A novel definition for image detail as fluctuations between local minima and maxima. We proposed a simple algorithm to smooth an input image. By recursively performing the smoothing with extrema detection at single scales, we performed a decomposition of the input image into multiple-scale layers of detail and a coarse residual. Our algorithm smoothes high-contrast surface while preserving significant edges. Finally, we exploited this ability by applying our decomposition in a variety of applications.

In the future, it is likely that hardware wavelet processing will be common in image processing systems, and it will be straightforward to utilize this hardware for range compression. Adding features such as robustness, hardware memory and computation efficiency, color consideration, and automatic selection of parameters in super-resolution methods will be the ultimate goal for the Super-Resolution researchers and practitioners in the future.

6. REFERENCES

- [1] K. S. Ni and T. Q. Nguyen, "Image superresolution using support vector regression," *IEEE Trans. Image Process.*, vol. 16, no. 6, pp.1596–1610, Jun. 2007.
- [2] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE Int. Conf. Comput. Vision*, 2009.
- [3] A. J. Smola and B. Schölkopf, *A Tutorial on Support Vector Regression Statistics and Computing*, 2003.
- [4] M.-C. Yang, C.-T.Chu, and Y.-C. F. Wang, "Learning sparse image representation with support vector regression for single-image superresolution," in *Proc. IEEE Int. Conf. Image Processing*, 2010.
- [5] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint map registration and high-resolution image estimation using a sequence of under-sampled images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp.1621–1633, 1997.
- [6] N. Nguyen, P. Milanfar, and G. H. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 573–583, Apr. 2001.
- [7] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [8] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1167–1183, Sep. 2002.
- [9] J. Sun, J. Zhu, and M. F. Tappen, "Context-constrained hallucination for image super-resolution," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2010.
- [10] M. E. Tipping and C. M. Bishop, "Bayesian image super-resolution," *Adv. Neural Inf. Processing Syst.*, 2002.
- [11] H. Y. Shum and Z. C. Lin, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 83–97, Jan. 2004.
- [12] A. Chakrabarti, A. N. Rajagopalan, and R. Chellappa, "Super-resolution of face images using kernel PCA-based prior," *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 888–892, 2007.
- [13] M. Protter and M. Elad, "Image sequence denoising via sparse and redundant representations," *IEEE Trans. Image Process.*, vol. 18, no.1, pp. 27–35, Jan. 2009.
- [14] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, 2002.
- [15] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2004.
- [16] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Process.*, vol. 53, no. 3, 1991.
- [17] D. Capel, *Image Mosaicing and Super-Resolution (Cphc/Bcs Distinguished Dissertations)*. New York, NY, USA: Springer-Verlag, 2004.
- [18] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [19] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. ICCV*, , pp. 2272–2279.
- [20] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2008.
- [21] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [22] K. I. Kim and Y. Kwon, Example-Based Learning for Single-Image Super-Resolution and jpeg Artifact Removal, Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany, 2008.
- [23] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, Jun. 2010.
- [24] C.-Y. Yang, J.-B. Huang, and M.-H. Yang, "Exploiting self-similarities for single frame super-resolution," in *Proc. Asian Conf. Comput. Vision*, 2010, pp. 497–510.