Postprint of Single-Image Super-Resolution Algorithm Based on Sparse Bayesian Estimation

Authors: Yuan Guixia, Zhou Xianchun

Date: 2018-05-18T00:00:00+00:00

Abstract

To address the problem of significant performance variation among existing super-resolution methods when applied to different low-resolution images, we propose a single-image super-resolution approach based on sparse Bayesian estimation. This method formulates single-image super-resolution as a regression problem, utilizing Kronecker pulse functions as regression basis functions. It comprehensively exploits both local and global image information to seek the optimal sparse solution for specific predictions, estimates weights through Bayesian methods, and reconstructs the super-resolution image accordingly. Experimental results on 14 test images demonstrate that the proposed method achieves a high average peak signal-to-noise ratio (PSNR) with low variance and reduced computational time, thereby confirming its effective super-resolution performance, strong adaptability, and high computational efficiency.

Full Text

Single Image Super-Resolution Method Based on Sparse Bayesian Estimation

Yuan Guixia[1], Zhou Xianchun[2]
[1]School of Information, Mechanical & Electrical Engineering, Jiangsu Open University, Nanjing 210017, China
[2]School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

Abstract

Existing super-resolution methods exhibit significant performance variations across different low-resolution images. To address this issue, we propose a novel single image super-resolution approach based on sparse Bayesian estimation.

This method frames single image super-resolution as a regression problem, employing Kronecker pulse functions as regression basis functions. By integrating both local and global image information, it seeks the optimal sparse solution for specific predictions, with weights estimated through Bayesian methods to reconstruct the super-resolution image. Experimental results demonstrate that when applied to 14 test images, our method achieves high average peak signal-to-noise ratio (PSNR), low variance, and reduced computational time, confirming its superior super-resolution performance, strong adaptability, and high computational efficiency.

Keywords: single image super-resolution; super-resolution; Bayesian estimation; regression; sparse representation

0 Introduction

Single image super-resolution (SISR) is a classic problem in computer vision that has attracted considerable attention in recent years. The task of SISR is to recover a high-resolution image from a corresponding low-resolution input [**?**]. Current SISR methods typically fall into two categories: blind super-resolution and non-blind super-resolution. Blind super-resolution assumes that both the high-resolution image and the blur kernel are unknown, whereas non-blind methods assume only the high-resolution image is unknown while the blur kernel is known. Non-blind super-resolution methods can effectively handle low-resolution images with different blur kernels and generally require no training phase, yielding favorable results. However, blur kernels in natural images are typically unknown or fail to satisfy ideal assumptions, limiting the applicability of such methods.

Blind super-resolution methods offer strong universality for single image super-resolution. These approaches learn mapping rules from low-resolution to high-resolution images from extensive training datasets [**?**, **?**, **?**, **?**, **?**], operating under the assumption that training and test samples share the same blur kernel, making the learned model applicable to test data. Most contemporary super-resolution methods employ machine learning for training and inference. However, these methods are largely tailored for image deconvolution, exhibiting excessive dependence on blur kernel selection and poor adaptability, which manifests as significant performance variations across different low-resolution images. To overcome this limitation, we introduce sparse Bayesian estimation into the super-resolution framework, treating single image super-resolution as a regression problem and solving it through sparse Bayesian estimation to achieve image enhancement. Our approach demonstrates stronger adaptability as its performance does not depend on blur kernel selection.

1 Related Work

Image super-resolution is a fundamental research topic in computer vision, comprising two main branches. The first branch reconstructs high-resolution images from sequences of low-resolution images captured by multiple or single cameras at different angles. These methods rely on precise sub-pixel alignment (typically using SIFT for keypoint matching) to infer missing pixel values in the high-resolution domain [?, ?, ?]. The second branch reconstructs high-resolution images from a single low-resolution input—single image super-resolution—which is the focus of this paper.

Single image super-resolution methods can be further categorized into blind and non-blind approaches, as discussed in the introduction. Blind methods attempt to estimate both the unknown blur kernel and the high-resolution image simultaneously. State-of-the-art SISR methods are based on various machine learning principles, directly learning mappings between low-resolution and high-resolution domains from large training image collections. For instance, dictionary learning methods train coupled dictionaries in both high-resolution and low-resolution domains to effectively represent training data. During inference, low-resolution images are encoded using the low-resolution dictionary (e.g., via sparse coding), and high-resolution data is reconstructed using the high-resolution dictionary [?, ?, ?, ?]. Neighborhood embedding methods, in contrast, encode directly using all training data without dictionary learning [?, ?, ?]. Reference [?] proposed a super-resolution method combining sparse learning and neighborhood embedding, which uses correlation with dictionary atoms rather than Euclidean distance for neighbor computation, and introduced anchored neighborhood regression to embed low-resolution patches to the nearest dictionary atoms with precomputed embedding matrices, improving efficiency. Reference [?] presented a random forest-based super-resolution method that optimizes a novel regularized objective during tree training, operating on both output and input spaces, making it particularly suitable for SISR regression tasks. The method in [?] also employs sparse representation, obtaining high-resolution patch predictions through MMSE estimation and using neural networks for learning, balancing quality and computational complexity. Reference [?] applied deep learning to SISR, using deep convolutional neural networks to learn end-to-end mappings from low-resolution to high-resolution for image enhancement.

However, existing blind super-resolution methods are typically customized for image deconvolution, showing significant performance variations and poor adaptability across different low-resolution images. Our introduction of sparse Bayesian estimation into super-resolution eliminates dependence on blur kernel selection, offering stronger adaptability.

2 Proposed Method

Our sparse Bayesian estimation-based super-resolution method employs Kronecker pulse functions as regression basis functions to formulate the SISR regression problem. By integrating local and global image information, it seeks optimal sparse solutions for specific predictions, with weights estimated through Bayesian methods to transform low-resolution inputs into high-resolution outputs.

2.1 Super-Resolution Reconstruction Model

Given a dataset of the form $\{(\mathbf{x}_i, I_i)\}_{i=1}^N$, where $\mathbf{x}_i$ are known input vectors and I_i are output targets, the goal is to make accurate predictions for new input vectors $\mathbf{x}$. Let the set of observed samples be $\mathcal{I} = \{I_i^l\}_{i=1}^N$, where these images are low-resolution. The objective of super-resolution algorithms is to obtain a high-resolution image set from the observed low-resolution collection.

In the super-resolution context, $\mathbf{x}$ is associated with positions in the image sampling grid, where $\mathbf{x}_l$ corresponds to low-resolution sampling positions and $\mathbf{x}_h$ to high-resolution positions. Given $\mathbf{x}$, super-resolution reconstruction can be expressed as $I(\mathbf{x}) = f(\mathbf{x})$, where $I(\mathbf{x})$ represents brightness or color values at spatial location $\mathbf{x}$, and $f(\cdot)$ is a continuous function.

Thus, the super-resolution problem can be viewed as an optimal search for continuous information from a discrete, irregular sample grid—a classic regression problem. For digital images, the signal is discrete, meaning the image I is represented by a continuous function. During regression, the image signal is typically approximated as:

$$\hat{f}(\mathbf{x}_h) = \sum_{j=1}^{J} I_{h_j} \delta(\mathbf{x}_h - \mathbf{x}_{h_j})$$

where $\hat{f}$ denotes the discrete function of f at high-resolution position $\mathbf{x}_h$, and J is the number of sampling points.

The regression model can be expressed as:

$$\hat{f} = \Phi \cdot \mathbf{w}$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_J]^T$ is the weight vector to be learned, and $\Phi = [\phi_1, \phi_2, \ldots, \phi_J]$ with $\phi_j = \delta(\mathbf{x}_h - \mathbf{x}_{h_j})$ represents the basis functions.

Given a regression basis Φ, Bayesian treatment yields the predictive distribution, typically expressed as the product of likelihood and prior terms:

$$p(\mathbf{w}|\mathbf{I}^l, \Phi) \propto p(\mathbf{I}^l|\Phi, \mathbf{w}) p(\mathbf{w})$$

The likelihood model follows a Gaussian distribution:

$$p(\mathbf{I}^l|\Phi, \mathbf{w}) = \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\|I_i^l - \mathbf{g}_i^T \Phi \mathbf{w}\|^2}{2\sigma^2}\right)$$

where M is the number of image patches (with patch size 3×3), and $\mathbf{g}_i$ maps high-resolution to low-resolution through subsampling.

The prior model follows a Gibbs distribution, providing a convex cost function with a global minimum that is easier to optimize. The model can be rewritten as:

$$p(\mathbf{w}) = \frac{1}{Z} \exp\left(-\frac{1}{2}\mathbf{w}^T\mathbf{K}^{-1}\mathbf{w}\right)$$

where $\mathbf{K}$ is an $M \times M$ kernel matrix to be learned, representing prior assumptions about local relationships between $\mathbf{w}$ values. The kernel matrix is learned using a Gibbs distribution with predefined neighborhoods.

2.2 Sparse Bayesian Estimation

The challenge in super-resolution reconstruction lies in the uncertainty of the sampling grid, which we address using a sparse Bayesian framework. From a sparsity perspective, super-resolution can be viewed as estimating the weight vector $\mathbf{w}$ to find optimal weights that minimize the difference between regression values and observations:

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \|\mathbf{I}^l - \Phi\mathbf{w}\|^2 + \lambda\|\mathbf{w}\|_0$$

Unlike traditional methods, our approach constrains sparsity by assuming weights are set within predefined neighborhoods, aiming to maximize covariance between weights. The regression basis corresponds to unit pulse functions placed on a predefined high-resolution grid, while weights correspond to kernel functions that transform image color space into a Gibbs distribution. This kernel modulates the likelihood function, approximated by a Gaussian function of the difference between observations and the linear model.

Maximum a posteriori estimation can be solved through an expectation-maximization framework, treating the high-resolution image $\mathbf{I}^h$ as a latent variable. The EM framework consists of two steps: the E-step estimates the high-resolution image and its covariance using the current kernel, while the M-step determines the optimal kernel using the high-resolution image from the E-step.

The E-step computes the mean estimate of the high-resolution image $\mathbf{I}^h$ and its covariance. The M-step minimizes:

$$\mathcal{L}(\mathbf{w}) = \|\mathbf{I}^l - \Phi\mathbf{w}\|^2 + \mathbf{w}^T\mathbf{K}^{-1}\mathbf{w}$$

using gradient descent:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \epsilon \nabla_{\mathbf{w}} \mathcal{L}$$

where the basis functions Φ are approximated using Kronecker pulse functions.

The above steps consider only local information from different sampling points, which is incomplete. We incorporate global smoothness estimation from the low-resolution image to compensate the prior term and improve regression accuracy. Specifically, we smooth the low-resolution image to obtain a smoothed estimate $\tilde{\mathbf{I}}^l$ using a 3×3 window with uniform weights. Global information from the low-resolution image is compensated by making the data vary smoothly around this estimate:

$$\mathcal{L}_{\text{global}} = \|\mathbf{I}^l - \tilde{\mathbf{I}}^l\|^2$$

We solve the complete Bayesian formula using a likelihood cost function, allowing local and global prior terms to be defined as:

$$\mathcal{L}(\mathbf{w}) = \|\mathbf{I}^l - \Phi\mathbf{w}\|^2 + \alpha \mathbf{w}^T \mathbf{K}^{-1} \mathbf{w} + \beta \|\mathbf{I}^l - \tilde{\mathbf{I}}^l\|^2$$

2.3 Image Super-Resolution

Learning-based super-resolution methods typically require a training phase to learn the weight vector, followed by reconstruction. We select the publicly available Set5 subset [?] for training, which contains five images [Figure 1: see original paper]. All training images are color images. Since human visual perception is more sensitive to intensity than color variations, we convert images to YUV color space and apply our method only to the Y component for training and reconstruction. During training, original images serve as high-resolution targets, while bilinearly interpolated subsampled versions serve as low-resolution inputs, yielding the weight vector $\mathbf{w}$ through the process described above.

For super-resolution, grayscale images are reconstructed directly using the learned weight vector. For color images, we convert to YUV space and apply our method only to the Y component. The U and V components are upsampled using standard bicubic interpolation to match the high-resolution dimensions. The final result is obtained by converting the high-resolution YUV components back to RGB color space.

3 Experimental Results and Analysis

To validate our method, we test on the publicly available Set14 subset [?], shown in [Figure 2: see original paper]. Algorithm evaluation employs both objective and subjective assessment. For objective evaluation, we use peak signal-to-noise ratio (PSNR), the standard metric for image quality assessment. Performance is evaluated through PSNR values of super-resolved images, with variance across different images measuring adaptability. Computational efficiency is assessed through runtime comparison. For subjective evaluation, we visually compare results from different methods. We compare against four state-of-the-art methods [?, ?, ?, ?], all trained on Set5 and tested on Set14 following the protocol in [?].

3.1 Objective Evaluation

presents PSNR values for 14 images from Set14 processed by all five methods. Bold values indicate the highest PSNR for each image. Our method achieves the highest PSNR for 12 images, with only two images slightly lower than [?], demonstrating superior enhancement performance.

[Figure 3: see original paper] shows the mean and variance of PSNR values across all 14 images. Our method not only achieves the highest mean PSNR but also the lowest variance, indicating strong adaptability and consistent performance across different low-resolution images with minimal dependence on blur kernels.

reports average super-resolution runtime (excluding training). All methods were tested on identical hardware: Intel i7 CPU, 16 GB DDR3 memory, 64-bit Windows 7, Visual Studio 2010. Our method exhibits the shortest runtime, confirming highest computational efficiency.

3.2 Subjective Evaluation

Due to space limitations, we present local super-resolution results for the "baboon" image in [Figure 4: see original paper]. Figure 4: see original paper shows the ground truth, while Figure 4: see original paper show results from methods [?, ?, ?, ?] and our method, respectively. Our method produces sharper details, indicating superior visual enhancement.

In summary, both objective and subjective evaluations demonstrate that our method outperforms the four comparison methods, achieving higher PSNR values across different low-resolution images with better enhancement quality, stronger adaptability, and higher computational efficiency.

4 Conclusion

The proposed single image super-resolution method based on sparse Bayesian estimation is a blind super-resolution approach that frames SISR as a regression problem. Using Kronecker pulse functions as regression basis functions, it integrates local and global image information to find optimal sparse solutions, with weights estimated through Bayesian methods. Our method' s key advantage is its independence from blur kernel selection, yielding consistent performance across different low-resolution images with strong adaptability. It achieves higher PSNR values and superior computational efficiency in single image super-resolution tasks.

References

[1] Ding L, Wang Z, Wen B, et al. Robust single image super-resolution via deep networks with sparse prior [J]. IEEE Trans on Image Processing, 2016, 25(7): 3194-3201.

[2] Yang M C, Wang Y C F. A self-learning approach to single image super-resolution [J]. IEEE Trans on Multimedia, 2013, 15(3): 498-508.

[3] Wang L, Xiang S, Meng G, et al. Edge-Directed single-image super-resolution via adaptive gradient magnitude self-interpolation [J]. IEEE Trans on Circuits & Systems for Video Technology, 2013, 23(8): 1289-1299.

[4] Zhang K, Gao X, Tao D, et al. Single image super-resolution with multiscale similarity learning [J]. IEEE Trans on Neural Networks & Learning Systems, 2013, 24(10): 1648-1659.

[5] Rueda A, Malpica N, Romero E. Single-image super-resolution of brain MR images using overcomplete dictionaries [J]. Medical Image Analysis, 2013, 17(1): 113-132.

[6] Zhang K, Tao D, Gao X, et al. Learning multiple linear mappings for efficient single image super-resolution [J]. IEEE Trans on Image Processing, 2015, 24(3): 846-860.

[7] Li Y, Lv X, Yu H. Sequence image super-resolution reconstruction based on mixed L1 and L2 norms [J]. Computer Applications, 2015, 35(3): 840-843.

[8] Guo L, Ye B, Chen Q, et al. Sequence image super-resolution reconstruction fusing geometric transformation similar blocks [J]. Opto-Electronic Engineering, 2015, 42(8): 79-85.

[9] Ding L, Fu Z, Li X, et al. Hybrid resolution multi-view image super-resolution for crack repair using non-uniform sampling restoration theory [J]. Opto-Electronic Engineering, 2013, 40(11): 76-82.

[10] Tang Y, Yuan Y, Yan P, et al. Greedy regression in sparse coding space for single-image super-resolution [J]. Journal of Visual Communication & Image Representation, 2013, 24(2): 148-159.

[11] Xu H, Zhai G, Yang X. Single image super-resolution with detail enhancement based on local fractal analysis of gradient [J]. IEEE Trans on Circuits & Systems for Video Technology, 2013, 23(10): 1740-1754.

[12] Jiang J, Hu R, Zhen H, et al. Efficient single image super-resolution via graph-constrained least squares regression [J]. Multimedia Tools & Applications, 2014, 72(3): 2573-2596.

[13] Zhu Z, Guo F, Yu H, et al. Fast single image super-resolution via self-example learning and sparse representation [J]. IEEE Trans on Multimedia, 2014, 16(8): 2178-2190.

[14] Chen X, Qi C. Low-rank neighbor embedding for single image super-resolution [J]. Signal Processing, 2014, 94(1): 6-22.

[15] Bevilacqua M, Roumy A, Guillemot C, et al. Single-image super-resolution via linear mapping of interpolated self-examples [J]. IEEE Trans on Image Processing, 2014, 23(12): 5334-4537.

[16] Li J, Gong W, Li W, et al. Single-image super-resolution reconstruction based on global non-zero gradient penalty and non-local Laplacian sparse coding [J]. Digital Signal Processing, 2014, 26(1): 101-112.

[17] Timofte R, De V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution [C]// Proc of IEEE International Conference on Computer Vision. 2013: 1920-1927.

[18] Schulter S, Leistner C, Bischof H. Fast and accurate image upscaling with super-resolution forests [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3791-3799.

[19] Peleg T, Elad M. A statistical prediction model based on sparse representations for single image super-resolution [J]. IEEE Trans on Image Processing, 2014, 23(6): 2569-2582.

[20] Dong Chao, Loy C C, He Kaiming, et al. Learning a deep convolutional network for image super-resolution [C]// Proc of European Conference on Computer Vision. Springer International Publishing, 2014: 184-199.

Note: Figure translations are in progress. See original paper for figures.