Fast image super resolution via local regression

Shuhang Gu
The Hong Kong Polytechnic University

Nong Sang
Huazhong University of Science and Technology

21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Fast Image Super Resolution via Local Regression

Shuhang Gu Nong Sang Fan Ma
Institute for PR&AI, Huazhong University of Science & Technology, Wuhan 430074, PR China
shuhanggu@gmail.com

Abstract

In this paper, we propose a super resolution method based on linear regression in different middle-frequency texture categories. We benefit from the hypothesis that the mapping from middle-frequency manifold to high-frequency manifold is similar locally, and use simple linear regression method to learn mapping functions in different area of middle-frequency manifold. Different from previous works, our method only uses the database to learn the mapping functions in different categories in the training phase, then we just need to save these mapping functions instead of a huge external database to get the missing details. Some experiments are used to confirm the effectiveness and efficiency of our method as well as our hypothesis.

1. Introduction

The purpose of super resolution(SR) is to recover a high resolution image from one or more low resolution images. Reconstruction based SR is a classical approach which uses several low resolution images to reconstruct the high resolution output. If we can get the relationship of these low resolution images in sub-pixel accuracy, we can fuse details from different images to generate the high resolution image. However, this approach is a severely ill-posed problem and limited only to small increases in resolution [1,2].

Another category of SR approach uses a single image as input to recover the high resolution image, these approaches can`t benefit from details of different images, so it often needs some other information (e.g. hypothesis or external database). Interpolation methods use smooth prior to predict the new pixels using analytical interpolation formulae, they are very fast, but tend to generate artifacts alone the edges and get over smoothed image. Some researchers have used more complex priors to improve interpolation [3,4].

Example based SR was firstly suggested by Freeman et al.[5], they used prior indirectly with an external database. Low and high resolution patch pairs in the database are used to indicate correspondence between low and high images with different machine learning algorithms. Example-based SR is an ill-posed problem, because different high resolution patches may correspond to the same low resolution patch. Nevertheless it had been a very active research area these years and had lead to the states of art .

In this paper, we propose a new example-based SR method. We assume the mapping from middle-frequency manifold to high-frequency manifold is similar locally, and learn mapping functions in different areas of middle-frequency manifold. Different from previous works, we use a database to learn the mapping functions in training phase, then we just need to save these mapping functions instead of a huge database in SR phase. Our method can generate comparable high resolution results to the state of art algorithms efficiently. The flowchart of our method is shown in figure 1.

The rest of the paper is organized as follow, in section 2 we give a briefly review of example-based SR method, in section 3, we introduce our hypothesis about local mapping similarity and details of our SR algorithm. Experiments and conclusion is in section 4.

2. Previous work

In this part, we review some other example-based SR approaches. Example-based SR uses a database of example co-patches to predict high resolution image. It was first proposed by Freeman et al [5], they use k-nearest neighbors algorithm to search similar low resolution patches in the database, the corresponding high resolution patches compose candidate patches of

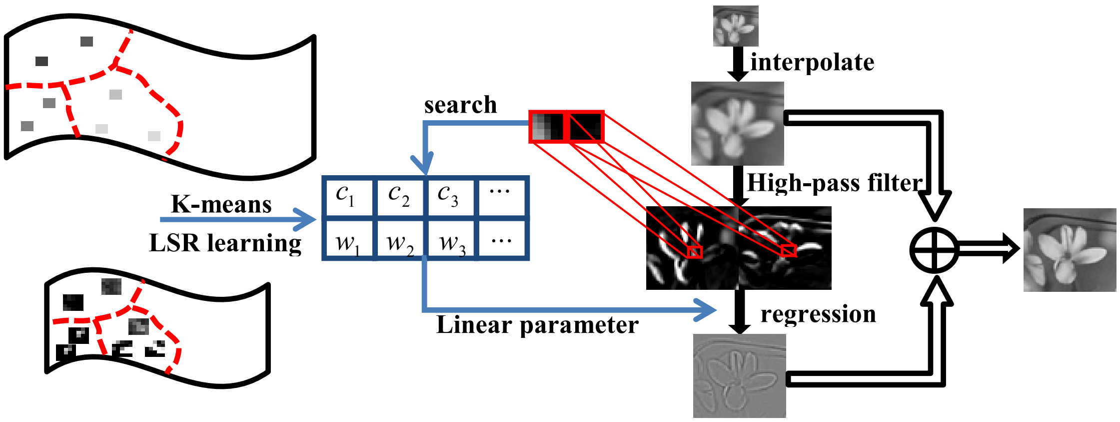


Figure 1. Flowchart of our SR framework. Build a database with middle frequency patches and high frequency value at the center of these patches. Categorize the database and learn linear functions in each category. Then we can use these specific regression functions to predict missing high frequency information

the high resolution output, then a MRF is used to generate the final high resolution image. Chang et al.[6] assumed small image patches in low and high resolution images form manifolds with similar local geometry, they used locally linear embedding (LLE) algorithm to compute reconstruction weights of k-nearest neighbors of each patch. Yang et al.[7] seek sparse representation of input low resolution patches, then use the coefficients of these representations to generate the high-resolution output. Glaneser et al.[8] benefited from the observation that patches in a nature image tend to redundantly recur in different scales inside the image, proposed a SR from single image approach.

Recently, Yang et al.[9] proposed a method in which they constructed an image segment dataset with different context categories, then the sparse representation of the interpolated image patches in specific categories are used to generate their ground truth pixel values via SVR. Tang et al.[10] assumed any manifold structure can be well approximated by a linear structure in sufficiently small range, they used kernel ridge regression locally in low resolution space to generate local high frequency information. Ref[9] and [10] both improve the super resolution results by learning the mapping from low resolution manifold to high resolution manifold in a small range. Motivated from these approaches, we assume the mapping from middle frequency manifold to high frequency manifold is similar locally, and use simple least square regression(LSR) method to learn mapping functions in different areas of the middle frequency manifold. Given an input low resolution image, we use specific mapping functions to predict missing high-frequency information, helping avoid searching in the huge data base so that we can get the high resolution output in shorter time.

3. fast image super resolution

3.1.super resolution via regression

In this part, we introduce how to build our database and model the relationship between the middle frequency information and high frequency information.

Low resolution observations are obtained by blurring and subsampling from the high resolution image. This process can be modeled as follow:

$$Y = DBX$$

where X represents the high resolution image and Y represents the low resolution image, B denotes the blurring operator and D denotes the subsampling operator. With this model, we can get the corresponding low resolution image Y given a high resolution training image X. A bigger image I can be obtained from Y via interpolate, sequentially, X sub I is the missing high frequency information H caused by interpolate enhancement. To estimate H given Y, we apply horizontal and vertical high-pass filters to extract middle frequency features M from I, patches m^j from M are used as inputs of regression function to predict the missing high frequency information h^j at the center of these patches.

In this paper, we want to indicate the validity of local mapping similarity hypothesis and focus on the framework of fast SR based on this hypothesis. So we just use a simple liner model to predict missing high frequency details. To learn the linear parameter w, we minimize the error function :

$$\min \quad \frac{1}{2}\sum_{j=1}^{n}(h^j - w^T m^j)^2 + \frac{\lambda}{2} w^T w$$

$$\Rightarrow w = (\lambda I + m^T m)^{-1} m^T h$$

Given a test image $\tilde{Y}$, we can get the corresponding middle frequency information from interpolation image $\tilde{I}$, then these information can be used to predict the missing high resolution details $\tilde{H}$ and generate the final high resolution image $\hat{X}$.

3.2.local mapping similarity

The method introduced in 3.1 can predict high resolution details, however, a single linear function can`t model the mapping function precisely. Our tests have shown that a single regression function improve interpolation about 1.4 in terms of RMSE(Root mean square error) on monarch image with zooming factor 2. We assume the mapping from middle frequency manifold to high-frequency manifold is similar locally, and learn different regression parameters in different areas of middle frequency manifold to predict the missing high-frequency information more precisely.

To confirm our hypothesis, we take 50 images in Berkley segmentation database to build a 5 million patches database and use k-means cluster algorithm to divide the database into k categories, then linear function parameters w_i are learned in each category. Fig.2 illustrates the RMSE value with different class number k, the RMSE of test image decreases rapidly with the class number increasing in a small range. However, further increasing of class number bring the RMSE begin to increase, this trend is obviously in small database and can be explained as follow: learning regression function in different categories can improve the accuracy of the mapping model, while with the increasing of class number k, samples in each categories getting less, and some categories can`t get enough samples to model the local mapping function and begin to suffer from over-fitting. That`s why with a larger database, we can keep the RMSE value decreasing in a bigger range and get better results.

3.3.fast image super resolution

In this part, we introduce details of our SR framework. In training phase, we build a database which contains middle frequency patches and corresponding high resolution values at the center of these patches, then categorize the database and learn linear function parameter w_i in each category with a simple LSR approach. In SR phase, given a low resolution image Y, horizontal and vertical high-pass filters are used to extract middle frequency information M in interpolate image. For each of patches in M, we search for its

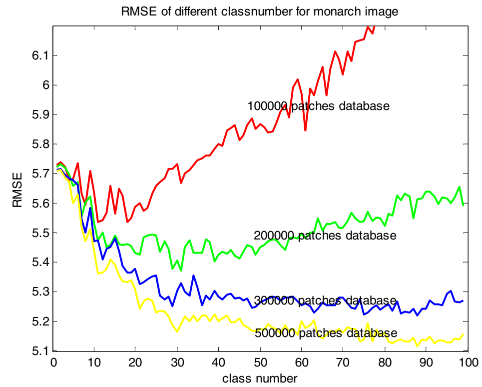


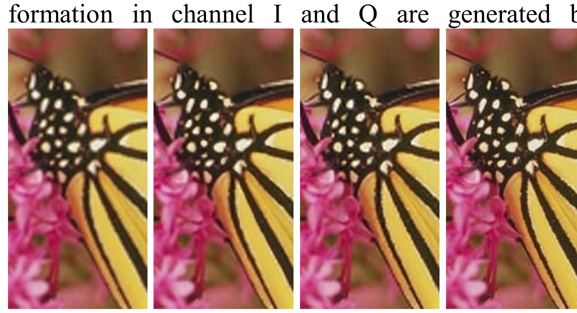
Figure 2. RMSE of different class numbers for monarch image with zooming factor 2.

nearest class center and apply the corresponding regression parameter to predict the high frequency information. Finally, the high resolution image $\hat{X}$ is obtained by adding high resolution information into interpolated image I.

Nevertheless, we tested on 1.5 million patches randomly extracted from nature image and found that high frequency value is very small(less than 1) at center of patches which have low middle frequency energy(less than 100 in terms of Frobenius norm in 5X5 patch). In practice, we only perform our method on patches which have large middle frequency, tests show this simplified method has little effect on the final high resolution output while allowing fast computation.

4. Experiments and conclusion

We tested our algorithm on different images and compared the SR results with other methods. We took 50 images in Berkley segmentation database to build a 500000 patches database. In practice, we chose the class number 30 as a trade-off between SR quality and running time. When working with color image, SR algorithm is only performed on Y channel and the information in channel I and Q are generated by

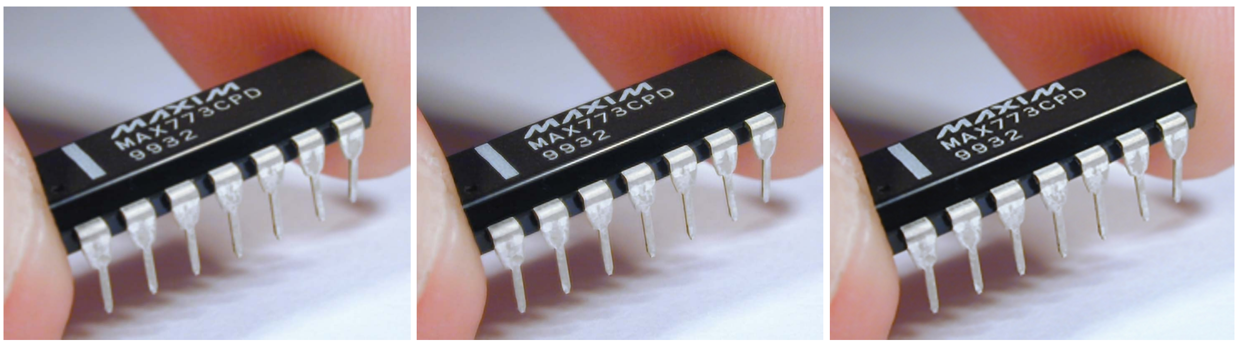


(a) (b) (c) (d)

Figure 3. SR results of monarch image (magnified x2), (a)Bicubic (b) Yang et al[9] (c) our method (d) ground truth

(a) Bicubic (b) Fattal et al. [11] (c) Glasner et al. [8] (d) Our method

Figure 4. SR results with different methods (magnified x4)



(a) Bicubic (b) Glasner et al. [8] (c) Our method

Figure 5. SR results with different methods (magnified x4)

interpolate. It takes about 0.3 seconds to upscale an image of 256-by-256 pixels by a factor of 2 along each axis using my Core i5 CPU laptop with single core.

We only compared our algorithm with [9] by a factor of 2 in quantity because we haven`t got the SR results of test images from other authors. The PSNR of the results are shown in table1.We also compared our SR results with the images on http://www.wisdom.weizmann.ac.il/~vision/SingleImageSR.html which doesn`t have ground truth images as shown in Fig.3,4,5.

Compared with existing SR methods, our method can save a lot of running time as well as generate high quality SR results (comparable results with state of art methods). Furthermore, our method predicts the missing high frequency information pixel by pixel, so our method is trivially parallel and can be optimization for speed conveniently.

Table 1. PSNR values of SR images

	lena	boat	fruit	childface	monarch
Bicubic	33.93	27.47	35.34	35.56	31.59
Method in[9]	34.49	27.15	35.74	35.78	33.28
Our method	**35.17**	**29.12**	**37.11**	**36.37**	**33.76**

References

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

[2] H.Y. Shum, Z.C. Lin. Fundamental limits of reconstruction based superresolution algorithms under local translation.*PAMI*,2006.

[3] J. Sun, Z. Xu, H. Y. Shum. Image Super-Resolution using Gradient Profile Prior. *CVPR*,2008.

[4] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong. Soft Edge Smoothness Prior for Alpha Channel Super Resolution. *CVPR*,2007.

[5] W. Freeman, E.Pasztor, O.Carmichael. Learning in low-level vision. *IJCV*,(1),2002.

[6] H. Chang, D.Y. Yeung, Y. Xiong. Super-Resolution Through Neighbor Embedding. *CVPR*,2004.

[7] J. Yang, J. Wright, T. Huang, Y. Ma. Image Super-Resolution via Sparse Representation. *TIP*,2010.

[8] D. Glasner, S. Bagon, M. Irani. SuperResolution from a signle Image. *ICCV*,2009.

[9] M.C. Yang , C.H. Wang, T.Y Hu, Y.C. Frank Wang. Learning context-aware sparse representation for single image super-resolution.*ICIP*,2011.

[10] Y. Tang , P. Yan, Y. Yuan, X. Li. Single-image super-resolution via local learning. *IJML* (2),2011.

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