Fast Single Image Super-resolution by Self-trained Filtering

Dalong Li, Steven Simske

HP Laboratories
HPL-2011-94

Keyword(s):

Super-resolution; PSNR; filter; image restoration; image enhancement

Abstract:

This paper introduces an algorithm to super-resolve an image based on a self-training filter (STF). As in other methods, we first increase the resolution by interpolation. The interpolated image has higher resolution, but is blurry because of the interpolation. Then, unlike other methods, we simply filter this interpolated image to recover some missing high frequency details by STF. The input image is first downsized at the same ratio used in super-resolution, then upsized. The super-resolution filters are obtained by minimizing the mean square error between the upsized image and the input image at different levels of the image pyramid. The best STF is chosen as the one with minimal error in the training phase. We have shown that STF is more effective than a generic unsharp mask filter. By combining interpolation and filtering, we achieved competitive results when compared to support vector regression methods and the kernel regression method.

External Posting Date: July 06, 2011 [Fulltext] Approved for External Publication
Internal Posting Date: July 06, 2011 [Fulltext]

Fast Single Image Super-resolution by Self-trained Filtering

Dalong Li, Steven Simske

Hewlett Packard Company
{dalong.li,steven.simske}@hp.com

Abstract. This paper introduces an algorithm to super-resolve an image based on a self-training filter (STF). As in other methods, we first increase the resolution by interpolation. The interpolated image has higher resolution, but is blurry because of the interpolation. Then, unlike other methods, we simply filter this interpolated image to recover some missing high frequency details by STF. The input image is first downsized at the same ratio used in super-resolution, then upsized. The super-resolution filters are obtained by minimizing the mean square error between the upsized image and the input image at different levels of the image pyramid. The best STF is chosen as the one with minimal error in the training phase. We have shown that STF is more effective than a generic unsharp mask filter. By combining interpolation and filtering, we achieved competitive results when compared to support vector regression methods and the kernel regression method.

Keywords: Super-resolution, PSNR, filter, image restoration, image enhancement

1 Introduction

Super-resolution techniques estimate an image at higher resolution based on its low-resolution observations. It is very useful in many functional imaging applications, such as face recognition [1]. In multi-frame super-resolution, multiple low-resolution images are provided. A high-resolution image is obtained by combining the non-redundant information in the low-resolution images. It normally involves image registration and image reconstruction steps. In this paper, we are concerned with single frame super-resolution (i.e. image upscaling), which like other inverse problems is heavily ill-posed.

A straightforward approach in image upscaling is interpolation: from a low-resolution image, interpolation algorithms can be used to fill in the missing pixels values on a finer grid. Commonly used interpolation methods are bilinear, bicubic, or B-spline kernels [2]. Kernel Regression (KR) [3] was recently introduced as a tool for image denoising, upscaling, interpolation, fusion etc. These methods all implicitly assume image smoothness. In this work, the kernel regression method is compared with our up-scaling method. Interpolated images are commonly used as the input for super-resolution methods. In the example-based super-resolution method [4], an image is first interpolated into a higher resolution. In the training phase, the mapping between patches from the interpolated images and the patches from the ground-truth images are learnt. The authors explored a Markov network in learning the mapping. In the testing phase, the learnt model is able to super-resolve on a patch basis. Our previous work has introduced support vector regression [5] in a number of image processing tasks including blind image deconvolution [6], image denoising [7] and super-resolution [8] and [9]. In our previous Support Vector Regression (SVR) based super-resolution method, we learned a mapping from a patch in the interpolated image to the corresponding central pixel in the ground truth image. The mapping is modeled by SVR.

In this paper, we explore a new idea in single image super-resolution. The interpolated image is normally blurry as a result of the interpolation methods which assume the smoothness of an image. Some higher frequency details are missing. Intuitively, a high frequency emphasizing filter such as a sharpening filter will improve the quality of the image. Sharpening filters are implemented in, for example, Abode Photoshop and the Gimp open source imaging applications. Though sharpened images look sharper, there are usually artifacts in the images. In terms of *peak signal-to-noise ratio (*PSNR), the quality of the processed image is actually degraded, this is confirmed in our experiments. Instead of a fixed filter such as an u*nsharp* mask *filter (UF)*, in this work we learn a high frequency emphasis filter from the input image itself. By downsizing and upsizing the input low resolution image, we created a blurry interpolated image. We then use it as the input and the output (ground-truth) image is the given low resolution image. By minimizing the mean square error, a filter is readily available. In this way, we learn a filter by self-training that can always improve the interpolation result.

The rest of the paper is organized as following: the detail of the proposed method is presented in Section 2. Comparative results are in Section 3, followed by a concluding remark.

2 Self-trained filtering based supre-resolution

The basic steps involved (shown in Fig. 1) are as follows:

1. A lower resolution image (L) is created by down-sampling (the image is anti-aliased by using one of several standard techniques—Hermite, cubic, nearest neighbor, bilinear interpolation, wavelets, etc.) the input image (I).

2. The lower resolution image is then up-sampled. Here, the missing pixel values are filled by interpolation. This image is denoted as U. In this work, we use bi-cubic interpolation.

3. Optimal MSE filters (f) with different PSF support are found between the interpolated high resolution image and the input image: I=U*f, where * denotes

convolution. Optionally, we apply it at a few more levels of the image pyramid. The filter that has minimal error is selected as the final filter.

4. The input image is initially up-sampled by the same interpolation (i.e. bi-cubic interpolation in our work) used in step 2. The interpolated high resolution image is denoted as IU. It is very important for the interpolation method to be the same one as used in the training phase. Otherwise, the learnt filter is not the most effective one.

5. IU is filtered by the trained filter f to get the final super-resolved image O.

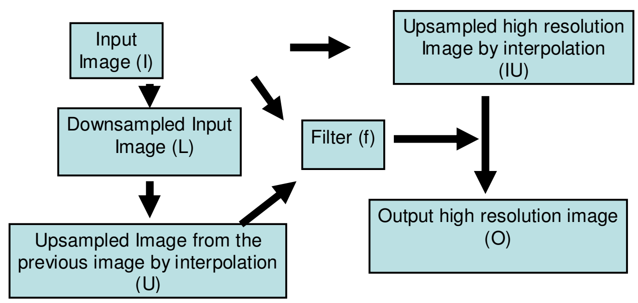

Fig. 1. Flow chart of the proposed Self-trained Filtering method.

In the above steps, step 3 is the most important one since it is where the optimal filter is learnt. We can try different PSF supports, e.g. 3x3, 5x5, and 7x7. We can also train the filter on different level of the image pyramid, meaning we further downsize the image and train the filter on that level also. However, the filters learnt from smaller images tend to be less useful. For each filter, we compute the error between the filtered image (predicted output) and the ground-truth. The error is used in selecting the final filter. The final filter is associated with the minimal prediction error. We can achieve spatial adaptive super-resolution by applying the above schema to a segmented image.

3 Experiments

Several images from the USC-SIPI [10] image database were evaluated to show the utility of our approach. A document image is also added. The upscaling ratio is 2, i.e. the image is to be doubled in size. To be consistent with the comparing methods, the interpolation method is bi-cubic. LibSVM [11] is used in our SVR method. The size of patch is 7x7 and the cross validation tool in LibSVM is used to find the optimal parameters for the default Gaussian kernel. In addition to visual inspection of the results, PSNR and ISNR are used to evaluate the performance quantitatively. PSNR is defined as:

$$PSNR(\hat{f}) = 10\log_{10}\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}255^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(f(i,j)-\hat{f}(i,j))^2} \quad (1)$$

where $\hat{f}(i,j)$ is the super-resolved image, and $f(i,j)$ is the original high-resolution image. The size of the images are M × N. From PSNR, ISNR (improvement in signal-to-noise ratio) can be computed as ISNR = PSNR($\hat{f}$) − PSNR(g) where g is the interpolated image. Thus, ISNR can reflect the improvement in terms of signal-to-noise ratio.

Table 1 lists the comparative results of various methods: interpolation by Bi-cubic, interpolation + STF, interpolation + UF (unsharp mask filter) and our SVR method [8]. Table 2 lists the results when Kernel Regression (KR) [3] is used as the interpolation method. The mean and standard deviation for each method are reported in the table also. Analysis of Variance (ANOVA) statistical tests report an F-ratio of 31.57 for Table 1 and 19.52 for Table 2, indicating that the performance of each method is statistically significantly different. In terms of PSNR/ISNR, the top two methods are our previous SVR method and the STF method proposed in this paper.
The result of interpolation+UF is always negative in terms of ISNR, meaning the processed images are degraded actually though the images might look sharper. On the other hand, the proposed STF always increase PSNR. STF is fundamentally different from UF since STF is a restoration filter, rather than a generic high frequency emphasizing filter. Though the STF method results in a lower ISNR on several of the images than those SVR results, it is much more computationally efficient than the SVR method.

Figures 2-9 shows the results on the TREE image. Both STF and UF are high frequency emphasizing filters. Since STF is adaptive to interpolation kernels/methods, it works better than UF. UF looks sharper but more artifacts are observed. Both the kernel regression result and the bi-cubic interpolated results are smooth. The kernel regression result is smoother than the interpolated one by the bi-cubic. Very little high frequency information is left in the result by kernel regression and as a result, neither UF nor STF can improve the result much. Both KF+UF and KF+STF results have very strong artifacts that make the image look very unnatural.

Table 1. Comparative super-resolution results when interpolated by bi-cubic.

Image	Interp PSNR	Interp+STF ISNR	Interp+UF ISNR	**SVR ISNR**
Baboon	21.51	0.31	-0.59	**0.53**
Boat	25.80	**0.52**	-1.33	0.39
Cameraman	26.32	**0.75**	-1.10	0.19
Doc	20.71	1.00	-0.36	**1.20**
House	31.69	0.77	-2.87	**2.00**
Lena	28.91	**0.63**	-2.12	0.15

Peppers	29.63	**0.64**	-2.26	0.32
Tree	27.59	0.92	-2.93	**1.41**
Mean ISNR	26.52	0.69	-1.70	**0.77**
Stand. Dev.	3.83	0.22	0.99	0.68

Table 2. ISNR Comparative super-resolution results when interpolated by KR.

Image	KR	**KR+STF**	KR+UF
Baboon	-1.18	**-0.01**	-0.97
Boat	-2.05	**-0.18**	-1.64
cameraman	-1.94	**-0.25**	-1.66
Doc	-1.87	**-0.20**	-1.22
House	-3.29	**-0.01**	-3.07
Lena	-2.81	**-0.11**	-2.12
Peppers	-3.41	**-0.01**	-0.17
Tree	-4.03	**-0.42**	-2.78
Mean ISNR	-2.57	-0.15	-1.70
Stand. Dev.	0.96	0.14	0.95

4 Conclusion

In this paper, we presented a fast and robust single image super-resolution algorithm by STF. The main major advantages of the proposed method are that it does not require other training images and that, since it is done by filtering, it is very computational efficient. In addition to its use in super-resolution, STF can also be used as a high frequency emphasizing filter and it works better than a generic filter such as Unsharp mask filter (UF) since it is adaptive rather than fixed. STF works best with the image interpolated by the same interpolation kernel used in the training phase, e.g. bi-cubic used in our work since that is how the STF is trained. When a different interpolation method is used, the STF shall be trained again to learn a filter than is specific to the method.

Our experiments show that the proposed method is effective, in terms of PSNR/ISNR, the result images are quantitatively closer to the ground-truth images. Visual inspections on the images also indicate that some of the missing high frequency details were recovered, though the filter cannot completely remove the artifacts introduced in the interpolation phase.

Our future work aims to improve it by applying it locally along with other enhancements in the learning phase. We will also work on how to improve the computation performance of the method.

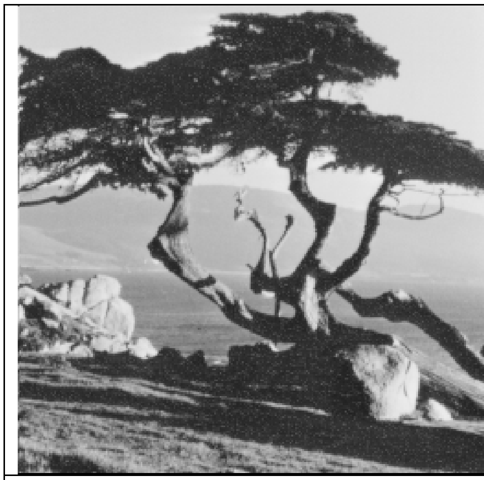

Fig. 2. The Ground Truth Tree image.

Fig. 3. The SVR result

Fig. 4. Interpolated by Bi-Cubic.

Fig. 5. Bi-Cubic Interpolated + STF.

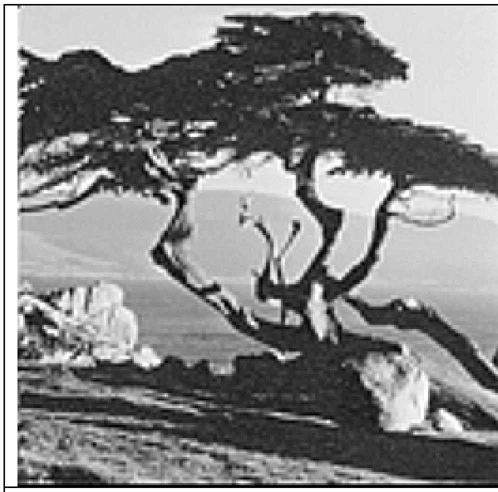

Fig. 6. Bi-Cubic Interpolated +UF.

Fig. 7. Interpolated by Kernel Regression (KR).

Fig. 8. KR+STF

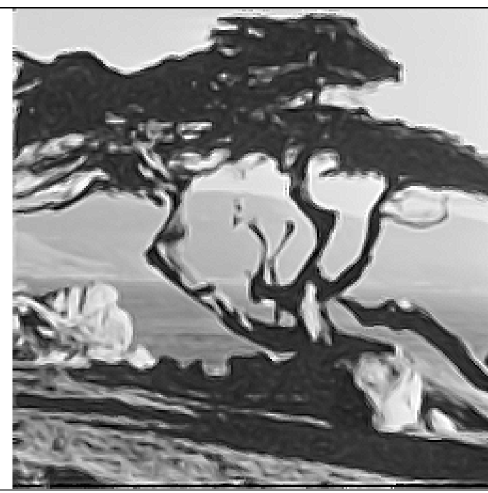

Fig. 9. KR+UF.

References

1. Gunturk, B. K., Batur, A.U., Altunbasak, Y., Hayes, M. H., Mersereau, R. M.: Eigenface-domain Super-resolution for Face Recognition. IEEE Trans. Image Process, pp. 1597--606, 2003.
2. Unser, M., Aldroubi, A., Eden, M.: Fast B-spline Transforms for Continuous Image Representation and Interpolation. IEEE Trans. Pattern Anal. Mach. Intell., vol. 13, no. 3, pp. 277--285, March 1991.
3. Takeda, H., Farsiu, S., Milanfar, P.: Kernel Regression for Image Processing and Reconstruction. IEEE Trans. Image Process., vol. 16, no. 2, Feb. 2007.
4. Freeman, W. T., Jones, T. R., Pasztor, E. C.: Example-based Super-resolution. IEEE Comput. Graphics and Applicat., vol. 22, no. 2, pp. 56--65, Mar./Apr. 2002.
5. Vapnik, V.: The Nature of Statistical Learning Theory, Springer Verlag, New York, 1995.
6. Li, D., Mersereau, R. M., Simske, S.J.: Blind Image Deconvolution through Support Vector Regression. IEEE Transactions on Neural Networks, Vol. 18, Issue. 3, pp. 931-935.
7. Li, D., R. M. Mersereau, R.M., Simske, S.J.: Image Denoising through Support Vector Regression. IEEE International Conference on Image Processing, pp.425--428, 2007.
8. Li, D., Simske, S.J.: Example Based Single-frame Image Super-resolution by Support Vector Regression. Journal of Pattern Recognition Research, Vol 5, No 1, pp. 104--118, 2010.
9. Xie, L.J., Li, D., Simske, S.J.: Feature Dimensionality Reduction for Example-based Image Super-resolution. Journal of Pattern Recognition Research, Vol 6, No 2, pp. 130--139, 2011.
10. University of Southern California - Signal & Image Processing Institute - The USC-SIPI Image Database, http://sipi.usc.edu/services/database/Database.html.
11. Chang, C. C., Lin, C. J.: LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.