Journal of Digital Contents Society Vol. 15 No. 6 Dec. 2014(pp. 737-743)
http://dx.doi.org/10.9728/dcs.2014.15.6.737

개선된 네이버 임베딩에 의한 초해상도 기법

엄경배*

요 약

단일 영상 초해상도 기법에는 보간 기반 방법과 표본 기반 방법 등이 있다. 보간 기반 방법들은 간결성에 강점을 가지고 있으나, 이들 방법들은 선지식을 이용할 수 없기 때문에 톱니 모양의 윤곽선을 가진 고해상도 영상을 생성하는 경향이 있다. 표본 기반 초해상도 기법에서서는 최근방 기반 알고리즘들이 널리 이용되어 지고 있다. 그들 중, 네이버 임베딩은 지역적 선형 임베딩이라는 매니폴드 학습 방법의 개념과 같다. 그러나, 네이버 임베딩은 국부 학습 데이터 집합의 크기가 너무 작은데에 따른 빈약한 일반화 능력으로 인하여, 시각적으로나 정량적인 척도에 의해 취약한 성능을 보인다. 본 논문에서는 이와 같은 문제점을 해결하기 위해 개선된 네이버 임베딩 알고리즘을 제안하였다. 저해상도 입력 영상이 주어지면 고해상도 버전의 화소 값들은 개선된 네이버 임베딩 알고리즘에 의해 구해진다. 실험 결과 제안된 방법이 바이큐빅 보간법이나 네이버 임베딩에 비해 정량적인 척도 및 시각적으로도 우수한 결과를 보였다.

키워드 : 초해상도, 개선된 네이버 임베딩, 서포트 벡터 회귀

Super Resolution Technique Through Improved Neighbor Embedding

Kyoung-Bae Eum*

Abstract

.For single image super resolution (SR), interpolation based and example based algorithms are extensively used. The interpolation algorithms have the strength of theoretical simplicity. However, those algorithms are tending to produce high resolution images with jagged edges, because they are not able to use more priori information. Example based algorithms have been studied in the past few years. For example based SR, the nearest neighbor based algorithms are extensively considered. Among them, neighbor embedding (NE) has been inspired by manifold learning method, particularly locally linear embedding. However, the sizes of local training sets are always too small. So, NE algorithm is weak in the performance of the visuality and quantitative measure by the poor generalization of nearest neighbor estimation. An improved NE algorithm with Support Vector Regression (SVR) was proposed to solve this problem. Given a low resolution image, the pixel values in its high resolution version are estimated by the improved NE. Comparing with bicubic and NE, the improvements of 1.25 dB and 2.33 dB are achieved in PSNR. Experimental results show that proposed method is quantitatively and visually more effective than prior works using bicubic interpolation and NE.

Keywords : Super Resolution, Improved NE, Support Vector Regression

※ Corresponding Author : Kyoung-Bae Eum
Received : December 08, 2014
Revised : December 30, 2014
Accepted : December 31, 2014
* Kunsan University, Dept. of Computer and Information Eng.
Tel: +82-63-469-4555 , Fax: +82-63-469-4560
email: kbeum@kunsan.ac.kr

1. Introduction

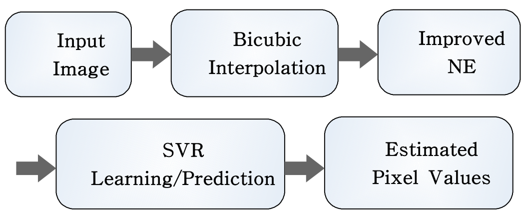
Super resolution (SR) techniques estimate an image at higher resolution from its low resolution observations. It has found useful in many applications, such as video surveillance and automatic target recognition [1]. The SR problem often implies multi-frame SR, where a high resolution image is obtained by combining the non-redundant informations found in multiple low resolution frames. Naturally, image registration and image reconstruction are the common two major steps in multi-frame SR algorithms. In this paper, I am concerned with single frame SR, which is different from the multi-frame case. There is only one observed low resolution image. The problem is highly ill posed. To explain why it is ill posed, I can view that a low resolution image is obtained by smoothing on the higher resolution image followed by a down sampling. The smoothing step is to prevent image aliasing. The goal of single frame SR is to estimate the high resolution image based on the observed low resolution image. Super resolution is to reverse the anti-aliasing and down sampling process. It is an inverse problem. Considering the anti-aliasing, an observed pixel in low resolution image can be viewed as a weighted sum of pixels from high resolution image pixels. The number of unknowns is more than the number of constraints. In this regard, like many other inverse problems such as blind image deconvolution, SR is an ill posed problem. There is no solution unless additional constraints are introduced to the problem. A commonly used constraint is the smoothness of an image [2].

For single image SR, interpolation based and example based algorithms are extensively considered. For interpolation based algorithms, interpolation function which reflects some explicit image priori information is used to obtain a much higher density of pixels in the resolving process. The interpolation algorithms have the strength of theoretical simplicity. However, those algorithms are limited because they are not able to use more priori information. Therefore, tending to produce high resolution images with jagged edges [3]. In this paper, I focus on the single image SR with example based algorithms. In general, the nearest neighbor based algorithms are extensively considered in previous example based algorithms [4,5,6]. For nearest neighbor based algorithms, a local training set is generated for each of test patches according to the similarlity between training and test patches. Given the local training set, a high resolution patch corresponding to the low resolution test patch is estimated by different learning algorithms. For instance, in [4], Freeman et al. firstly present an example based SR algorithm by employing Markov network to model the relationship between high resolution training patches and low resolution test patches. Another example based algorithm is introduced in [6] where neighbor embedding (NE) algorithm is adopted. They assume that a low resolution image and its corresponding high resolution image have similar local geometry, and generate a high resolution image patch by the nearest neighbors in the training set. However, these nearest neighbor based algorithms are weak in the performance of the visuality and the statistical measure such as peak signal to noise ratio (PSNR) because of the poor generalization of the nearest neighbor estimation [5]. The sizes of local training sets are always too small to improve the performance of NE. Bevilacqua [7] et al. have presented a new algorithm for single image SR based on external dictionary and nonnegative embedding. Vadukkoot [8] et al. have presented an advanced NE based algorithm for image SR reconstruction by combining the sparse neighbor search and subset selection based HoG clustering. However, these methods and NE are based on locally linear embedding(LLE). The LLE algorithm has a good performance provided that sufficient data points are sampled from the manifold. Otherwise, the local geometry of each patch could not be preserved. In this paper, an improved NE algorithm was proposed with Support Vector Regression (SVR) having an excellent generalization ability to solve this problem. Given a low resolution image, the pixel values are estimated in its high resolution version by using an improved NE. Experimental results verify that proposed method is quantitatively and visually more effective when comparing with a bicubic interpolation and conventional NE.

2. Proposed method

2.1 Patch representation

In the training stage, the high and low resolution image pairs are collected. For a low resolution image, a bicubic interpolation is used to synthesize its high-resolution version. All 5x5 patches from this synthesized image are extracted. Proposed method is used to get the neighbor representation of each patch in this high resolution version image. Obtained weight informations of each patch are used for SVR training. The pixel values in this synthesized image are updated by the learned SVR. (Figure 1) shows the flow chart of proposed method.



(Figure 1) Flow chart of proposed method

Patch representation algorithm in the proposed method can be summarized as follows: For each patch y_t^q in high-resolution version image Y_t:

(a) Find the set N_q of K nearest neighbors in high resolution training image Y_s.

(b) Compute the reconstruction weights of the neighbors that minimize the error of reconstructing y_t^q by equation (1).

Based on the K nearest neighbors identified, step (b) seeks to find the best reconstruction weights for each patch y_t^q in Y_t.

Optimality is achieved by minimizing the local reconstruction error for y_t^q

$$\varepsilon^q = \left\| y_t^q - \sum_{y_s^p \in N_q} w_{qp} y_s^p \right\|^2 \qquad (1)$$

which is the squared distance between y_t^q and its reconstruction, subject to the constraints $\sum_{y_s^p \in N_q} w_{qp} = 1$ and $w_{qp} = 0$ for any $y_s^p \notin N_q$.

Apparently, minimizing ε^q subject to the constraints is a constrained least squares problem. Let us define a local Gram matrix G_q for y_t^q as

$$G_q = (y_t^q 1^T - Y)^T (y_t^q 1^T - Y) \qquad (2)$$

where 1 is a column vector of ones and Y is a $D \times K$ matrix with its columns being the neighbors of y_t^q. Moreover, I group the weights of the neighbors to form a K dimensional weight vector w_q by reordering the subscript p of each weight w_{qp}. The constrained least squares problem has the following closed-form solution.

$$w_q = \frac{G_q^{-1} 1}{1^T G_q^{-1} 1} \qquad (3)$$

After repeating steps (a) and (b) for all N_t patches in Y_t, the reconstruction weights obtained form a weight matrix $W = [w_{qp}]_{N_t \times N_s}$. N_s is the number of all training patches.

2.2 SVR Learning and Prediction

The training data is made up of input / output pairs $(X_1;y_1),\ldots,(X_l;y_l)$, where X_i is a row vector corresponding to each patch from a weight matrix and y_i is the associated center pixel value of each patch in the high resolution image. Traditional linear regression estimates a linear function $W^T X + b$ that minimizes the mean square error:

$$\min_{w,b} \sum_{i=1}^{l} (y_i - (W^T X_i + b))^2 \qquad (4)$$

To address nonlinearly distributed input data, a mapping function $\phi(x)$ is introduced in the SVM[9] to map the data into a higher dimensional space in which the data can be linearly separated. In the high-dimensional space, overfitting occurs. To limit overfitting, a soft margin and a regularization term are incorporated into the objective function.

Support vector regression [10] has the following modified object function.

$$\min_{W,b,\xi,\xi^*} \frac{1}{2} W^T W + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \qquad (5)$$

$$\text{subject to } y_i - (W^T \phi(X_i) + b) \le \epsilon + \xi_i,$$
$$(W^T \phi(X_i) + b) - y_i \le \epsilon + \xi_i^*,$$
$$\xi_i, \xi_i^* \ge 0, i = 1, \ldots, l$$

ξ_i is the upper training error subject to the ϵ-insensitive tube $|y - (W^T \phi(X) + b)| \le \epsilon$ and ϵ is a threshold. C is the cost of error. The cost function ignores any training data that is within the threshold ϵ to the model prediction. This soft margin method increase the robustness of SVR. In the above object function, $\frac{1}{2} W^T W$ is a regularization term to smooth the function $W^T \phi(X_i) + b$ in order to limit overfitting. The parameters of the regression quality are the cost of error C, the width of the tube ϵ, and the mapping function ϕ. Similar to support vector classification, w is a high dimensional vector because ϕ maps data to a higher dimensional space, thus, the dual problem is solved instead [2]. The derivation of the dual is the same as in support vector classification. The primal dual relation shows that

$$w = \sum_{i=1}^{l} (-\alpha_i + \alpha_i^*) \phi(X_i) \qquad (6)$$

so the approximate function is

$$\sum_{i=1}^{l} (-\alpha_i + \alpha_i^*) K(X_i, X) + b \qquad (7)$$

$K(X_i, X)$ is the kernel function. Commonly used kernel functions are linear, polynomial, Gaussian, sigmoid etc.

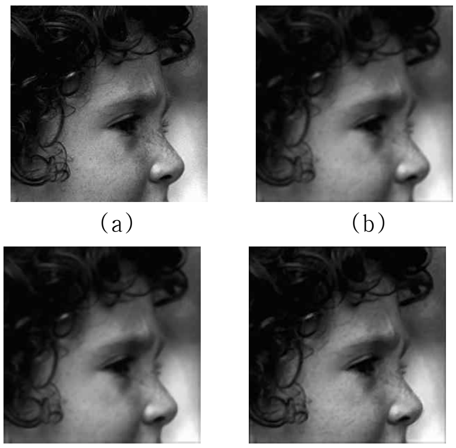
In this experiments, Gaussian kernels are used in SVR, and their parameters are selected via cross validation. The mean value of each patch is subtracted from its pixel values before calculating the neighbor coefficient ; this mean value is also subtracted from the corresponding pixel value in the high-resolution image. In testing, the mean value of each patch will be added to the predicted output pixel value. After the SVR model for patches is learned, they are used to predict the high resolution image of a given low-resolution test input. As the progress shown in (Figure 1) the high resolution version of the test input using a bicubic interpolation is synthesized. The patch representation algorithm is used to calculate the corresponding neighbor coefficient vector for each image patch. Finally, the pixel values are updated in the synthesized image using the previously learned SVR and the final SR image is obtained.

3. Experimental Results

Images from the USC-SIPI database and the medical image database of Univ. of Wisconsin are used in this experiments. I start with the original images as high resolution versions, and degrade them in a manner that is similar to the degradation I plan to undo in the images to be super-resolved. In the training set, input is formed by taking the neighbor coefficient for each patch in interpolated image and then converting it into a vector. The output corresponding to this input vector is the central pixel value in high resolution image. By sliding this neighborhood over all positions on the interpolated image, the training set for SVR is obtained. To super-resolve an image, the same size neighborhood window is used to create the input vector, and the pixel value in the interpolated image is updated by the trained model obtained in the training step. LibSVM[11], an implementation of SVR, is used in this experiments. To demonstrate the generalization ability of SVR, the training set is intentionally limited to very few(i.e. 1-2) images in the experiments. The test images are selected from several different catalogs. Natural images are gotten in the USC image database. The medical images such as brain and abdomen are obtained from the medical image database of Univ. of Wisconsin.

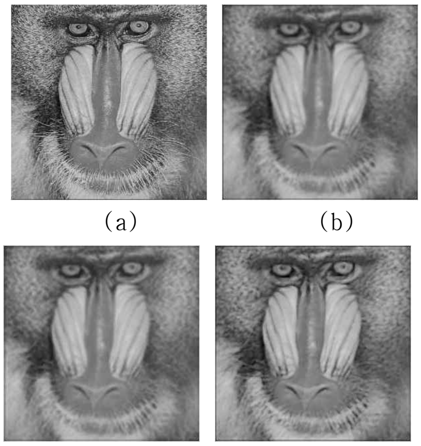
	Girl	Lenna	Man	Boat	Baboon
Bicubic	30.09	25.49	22.72	21.97	20.01
NE	30.41	25.87	22.90	21.96	19.99
Our method	30.89	30.12	23.84	22.65	20.32

<Table 1> PSNR values of test images



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

(Figure 2) Experimental results of Girl image
(a) Original image (b) Bicubic (c) NE
(d) Proposed method



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

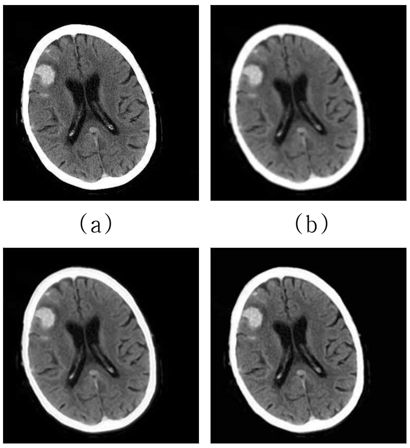
(Figure 3) Experimental results of Baboon image
(a) Original image (b) Bicubic (c) NE
(d) Proposed method

In this experiments, 2X magnifications are conducted on several related methods, such as bicubic interpolation and NE algorithm [6]. Objectively, PSNR is utilized to assess the performance of different algorithms, which are shown in the <Table 1> and <Table 2>. Also, the experimental results are visually evaluated. Given the high resolution training image, the corresponding low resolution image is obtained through blurring and down sampling.

Partition training high resolution image into patches of size 5*5 with overlapping 4 pixels. Few images for SVR training are used. The learned SVR models are used to super-resolve all test images. The SVR model is trained by LibSVM [11]. To validate the effectiveness of proposed method, the proposed method is compared with other algorithms including bicubic interpolation and NE.

	Brain1	Brain2	Abdomen1	Abdomen2
Bicubic	26.27	28.48	24.74	27.42
NE	28.79	31.47	26.66	28.83
Our method	29.79	32.98	27.42	30.20

<Table 2> PSNR values of medical images



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

(Figure 4) Experimental results of Brain1 image
(a) Original image (b) Bicubic (c) NE (d) Proposed method

To make a fair comparison, the same training images are used. All experimental results are shown in <Table 1>, <Table 2>, (Figure 2), (Figure 3) and (Figure 4). <Table 1> and <Table 2> show the PSNR results for the test images. As shown, proposed method outperforms the other methods with respect to PSNR. (Figure 2) shows the results of applying different methods. While the bicubic and NE smooth the texture on the face and hair region, proposed method shows better textures in these regions. (Figure 3) shows the results of Baboon image. There are more details in the result of proposed method than those of bicubic and NE, such as the beard in the lower part, which looks more sharp. (Figure 4) shows the results of brain1 image. Bicubic and NE show more smoothed results than proposed method. White line of brain center in the result of proposed method looks more clear than those of bicubic and NE. It is clear that proposed method is better than other algorithms in terms of the visuality and PSNR.

4. Conclusion

The nearest neighbor based algorithms are weak by the poor generalization of the nearest neighbor estimation, because the sizes of local training sets are always too small. In this paper, an improved NE algorithm based on support vector regression is proposed to solve this problem. Given a low resolution image, a neighbor representation is used to extract image patches, and the associated SVR model is learned to refine the pixel values in its high resolution version. The learned model by SVR is useful and robust for image super resolution. The experiments on different types of images are conducted and the results are promising. By comparing proposed method to the previous works, proposed approach produced very attractive SR images with better PSNR results than those with bicubic and NE. Future research will be directed at extensions of proposed approach to multi-scale SR problems. Following the above idea, it is interesting to study the algorithms for super resolution in the field of semi-supervised regression. In the future, I will pay more attention to the super resolution based on semi-supervised regression.

References

[1] R. S. Wagner, D. E. Waagen, and M. L. Cassabaum, "Image super resolution for improved automatic targ et recognition," in Proc. SPIE, vol. 5426, 2004.

[2] D. Li and Steven Simske, "Example based single fra me image super resolution by support vector regress

ion," Patten Recognition Research, vol. 5, no. 1, 2010.

[3] J. van Ouwerkerk, "Image super resolution survey," Image and Vision Computing, vol. 24, no. 10, 2006.

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

[5] Yi Tang and Xuelong Li, "Local semi-supervised regression for single image super resolution," IEEE Int. Workshop on MMSP, 2011.

[6] H. Chang, D. Yeung, and Y. Xiong, "Super resolution through neighbor embedding," in Proc. IEEE CVPR, vol. 1, 2004.

[7] M. Bevilacqua, A. Roumy, and C. Guillemot, "Super-resolution using Neighbor Embedding of Back-projection residuals," in 18th International Conference on Digital Signal Processing, 2013.

[8] Jili G Vadukkoot and Tinu Winson K, " A new image resolution technique using neighbor embedding," Int. Journal of Advances in Computer Science and Technology, vol. 3, no. 2, 2014.

[9] N. Cristianini and J. Shawe-Taylor, "An introduction to Support Vector Machines," New York:Cambridge Univ. Press, 2000.

[10] V. Vapnik, "The nature of statistical learning theory," Springer Verlag, New York, 1995.

[11] C. Chang and C. Lin, LIBSVM : a library for support vector machines, 2001.

엄 경 배

1986년 : 전북대학교 대학원 (공학석사)
1990년 : 전북대학교 대학원 (공학박사-전자공학)

1995년~1996년: 캐나다 토론토대학교 박사후 연구원
1999년, 2000년: 일본 동경대학교 방문 연구원
2011년~2012년: 미국 미주리대학교 교환교수
1989년~현재: 군산대학교 컴퓨터정보공학과 교수
관심분야 : 영상처리, 의료영상처리, 컴퓨터비젼