JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 18, No. 8, pp. 1517-1522, Dec. 2017

TV-G 분해를 이용한 초해상도 알고리즘

엄 경 배* · **범 동 규**
군산대학교 컴퓨터정보공학과

Super Resolution Algorithm using TV-G Decomposition

Kyoung-Bae Eum* · **Dong-Kyu Beom**
Department of Computer and Information Eng., Kunsan National University, Kunsan 54150, Korea

[요 약]

단일 영상 기반 초해상도(SR) 기법 중 TV 기반 초해상도 기법은 에지 보존과 artifact가 없다는 점에서 성공적인 방법으로 평가되어 왔으나, 텍스쳐 성분에서는 개선을 보이지 못했다. 본 논문에서는 이와 같은 문제점을 개선하기 위해서 새로운 TV-G 분해 기반 초해상도 기법을 제안하였다. 제안된 초해상도 방법에서는 에지와 같은 구조적 성분의 해상도를 보다 더 개선하기 위해 SVR 기반 up-sampling 방법을 제안하였다. 또한, Neighbor Embedding(NE)을 개선하기 위해 완화된 제약조건을 이용한 Non-negative Embedding(NNE) 방법에 기반한 학습 방법을 이용하여 텍스쳐 성분의 해상도를 개선하였다. 실험을 통하여 본 논문에서 제안된 방법이 기존의 보간법, ScSR, 기존의 TV 및 NNE 기법들에 비해 정량적인 척도 및 시각적으로도 향상된 좋은 결과들을 보였다.

[Abstract]

Among single image SR techniques, the TV based SR approach seems most successful in terms of edge preservation and no artifacts. But, this approach achieves insufficient SR for texture component. In this paper, we proposed a new TV-G decomposition based SR method to solve this problem. We proposed the SVR based up-sampling to get better edge preservation in the structure component. The NNE used the relaxed constraint to improve the NE. We used the NNE based learning method to improve the resolution of the texture component. Through experimental results, we quantitatively and qualitatively confirm the improved results of the proposed SR method when comparing with conventional interpolation method, ScSR, TV and NNE.

색인어 : 초해상도, TV-G 분해, 구조 영상, 텍스쳐, NNE

Key word : Super Resolution, TV-G Decomposition, Structure Image, Texture, NNE

http://dx.doi.org/10.9728/dcs.2017.18.8.1517

Received 17 November 2017; **Revised** 13 December 2017
Accepted 25 December 2017

***Corresponding Author ; Kyoung-Bae Eum**

Tel: +82-063-469-4555
E-mail: kbeum@kunsan.ac.kr

Ⅰ. 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 medical imaging, satellite imaging, video surveillance and automatic target recognition [1]. In this paper, we are concerned with single frame SR. There is only one observed low resolution image. The low resolution image(LR) is obtained by smoothing on the high resolution(HR) image followed by a down sampling. The smoothing step is to prevent image aliasing. The goal of single frame SR is to estimate the HR image based on the observed LR image. SR is to reverse the anti-aliasing and down sampling process. It is an inverse problem. The number of unknowns is more than the number of constraints. In this regard, 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. In this paper, we focus on the single image SR with example based algorithms. We refer to Neighbor Embedding (NE) [3], as a single image example based approach, that performs for each LR output patch a K-NN (Nearest Neighbor) estimation. The NE assumes that a LR image and its corresponding HR image have similar local geometry, and generate a HR image patch by the nearest neighbors in the training set. However, the least squares(LS) solution of the NE is ‘too fitted’ on the LR data and thus generates undesired bad HR reconstructions. The performance for a certain critical point is fallen by using SUM1-LS as a NE method. Non-negative Embedding(NNE) avoid this problem by replacing the sum-to-one equality constraint by a more ‘relaxed’ inequality constraint. The generalization ability of NNE is improved by this relaxed constraint [4].

Among single image SR techniques, the total variation(TV) based SR approach seems most successful in terms of edge preservation and no artifacts . In [5], an effective TV based SR approach was introduced. It achieves sharp edge preservation with no artifacts. This method seems the effective solution for single image SR. However, it still does not achieve the real SR, because this method achieves good SR for edge components and insufficient SR for texture component.

In this paper, we proposed a new TV-G decomposition based SR method to solve this problem. The input LR image is decomposed into the texture component and the structure component by the TV-G decomposition. The selective SR methods suitable for each component are used in the proposed SR method. The quality of the initial structure image has a very critical impact on the final SR results, especially along the edges. So, we proposed the Support Vector Regression (SVR) based up-sampling to get better initial structure image and better edge preservation in the structure component. As a post-processing step, the enhancement filter is used to enhance strong edges in an intermediate HR image by the SVR based up-sampling . The LS solution of NE is ‘too fitted’ on the LR data and thus generates undesired bad HR reconstructions. The NNE method solves this problem by using a non-negativity constraint. So, the NNE get the increasing number of neighborhoods and produces better HR image than the NE. In the proposed method, the NNE based learning method is applied to improve the resolution of the texture component. Finally, the output HR image is obtained by adding the HR structure component and the HR texture component. The experimental results proved that the proposed SR method obtain sharp edges and fine details in visual perception. Comparing the proposed SR method with several other state-of-the-art image SR methods like Bicubic, ScSR[6], TV[5], and NNE[4] shows that this approach performs better in terms of both qualitatively and quantitatively.

Ⅱ. TV-G Decomposition based Super Resolution

2-1 The proposed SR method

An overview of the proposed SR method is shown in Figure 1. The input LR image is decomposed into the LR texture component and the LR structure component by the TV-G decomposition. The LR structure component is up-sampled by the proposed SVR based up-sampling to the HR structure component.

In the 1st step of proposed SVR based up-sampling, we get its HR version structure image by a bicubic interpolation from a LR structure image. The patches from this HR version structure image are extracted. Image patches can be represented as a NNE combination of elements from an appropriately chosen over-complete dictionary. The learned dictionary is a more compact representation of the patch, compared to the approach, which simply sample a large amount of image patchs, reducing the computational cost. Obtained coefficients of each patch are used for SVR training.

Finally, we update the pixel values in the HR version structure image using the learned SVR and obtain the intermediate HR structure image. As a post-processing step, we add an enhancement filter which consists of a high-pass filter(HPF) and a

TV filter. Then the edge component is sharpened by the HPF. The overshoot and undershoot is removed by the TV filter.

In the proposed SR method, the NNE based learning method as in the same manner of the conventional method is applied to improve the resolution of the LR texture component. The patch size of LR texture image is 3×3 and one pixel overlap is set between adjacent patches. The patch size of HR texture image is 5×5.

Finally, the output HR image is obtained by adding the HR structure component and the HR texture component. Fig. 1 shows the flow chart of the proposed SR method.

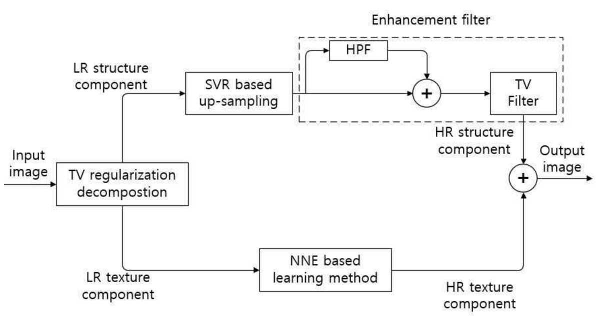

Fig. 1. The flow chart of the proposed SR method

2-2 TV-G Decomposition

Decomposing an image into meaningful components is an important and challenging inverse problem in image processing. An image is decomposed into its structural and textural parts [7]. The general concept of decomposition is that an image can be regarded as composed of a structural part, corresponding to the main large objects in the image, and a textural part, containing fine scale-details, usually with some periodicity and oscillatory nature [8].

Meyer suggests a new TV-G decomposition model [7]. An original image f is split into two components u and v, u containing the geometrical information and v the textural information. Our decomposition model is based on TV regularization approaches: a functional with two terms is minimized, a first one based on the total variation and a second one on a different norm adapted to the texture component [9]. Meyer proposes the following functional:

$$inf_{(u,v)\in BV\times G/f=u+v}\left(\int |Du|+\lambda \parallel v \parallel_G\right) \qquad (1)$$

where parameter λ is a Lagrange multiplier to control the fidelity scale of solution. The u is a structure and the v is a texture. The D is a gradient. The first item is used to retain the important edge feature of image. The Banach space G contains signals with large oscillations, and thus in particular textures and noise. A function belonging to G may have large oscillations and nevertheless have a small norm. In the second item, the norm on G is well-adapted to capture the oscillations of a function in an energy minimization method [7].

2-3 The reconstruction of the texture component

The input LR image is decomposed into the LR texture component and the LR structure component by the TV-G decomposition. In the proposed SR method, the NNE based learning method is applied to improve the resolution of the LR texture component.

NNE algorithm can be summarized as follows [4]:

We get the LR test patch x_t^i from the decomposed LR texture image X_t .

For each LR test patch x_t^i in LR texture image X_t :

(a) Dictionary is constructed by the pairs of LR and HR texture patchs in training data set.

(b) Compute the reconstruction weights of the neighbors that minimize the error of reconstructing x_t^i by equation (2).

$$w^i = \arg\min_w \left\| x_t^i - \sum_{x_s^p \in N_i} \omega x_s^p \right\|^2 \; s.t. \; w \geq 0 \qquad (2)$$

which is the squared distance between x_t^i and its reconstruction, subject to the non-negativity constraints $w \geq 0$. x_t^i is a LR test patch and x_s^p are the neighborhoods of this LR test patch in a dictionary.

(c) Compute the HR texture patch y_t^i by equation (3) using the appropriate HR dictionary patch of the neighbors and the reconstruction weights.

$$y_t^i = \sum w^i y_s^p \qquad (3)$$

y_t^i is a HR texture patch and y_s^p are the HR neighbor patchs of test patch in a dictionary. The LS solution of NE is 'too fitted' on the LR data and thus generates undesired bad HR reconstructions. The NNE method solves this problem by using a non-negativity constraint. This relaxed constraint improves the generalization ability of NNE. The NNE gets the increasing number of neighborhoods than the NE method. So, the NNE produces better HR texture image than the NE.

2-4 The reconstruction of the structure component

We proposed the SVR based up-sampling to improve the resolution of the input LR structure image decomposed by the TV-G method. In the 1st step, we get its HR version structure image by a bicubic interpolation from a LR structure image. The 5x5 patches from this HR version structure image are extracted. The proposed method is used to get the NNE neighbor representation of each patch in this HR version image. Obtained weight coefficients of each patch are used for SVR training. Finally, we update the pixel values in the HR version structure image using the learned SVR and obtain the intermediate HR structure image.

Support Vector Regression [10] is the extension of support vector machine. Using kernel tricks, the task of SVR is to use nonlinear functions to linearly estimate the output function in high-dimensional feature space. Similar to Support Vector Machine(SVM), the generalization ability makes the SVR very powerful in predicting unknown outputs.

In training, our SVR solves the following problem

$$\begin{aligned} \min_{w,b,\xi,\xi^*} \ & \frac{1}{2} w^T w + C \sum_{i=1}^{n} (\xi_i + \xi_i^*), \\ s.t. \ \ & y_i - (w^T \phi(\alpha_i) + b) \le \epsilon + \xi_i, \\ & (w^T \phi(\alpha_i) + b) - y_i \le \epsilon + \xi_i^*, \\ & \xi_i, \xi_i^* \ge 0, i = 1, \ldots, n. \end{aligned} \tag{4}$$

We note that y is the associated pixel value (at the same location as the center of the patch considered) in the original HR structure image, n is the number of training instances, $\phi(\alpha_i)$ is the neighbor representation of each patch in the transformed space, and w represents the nonlinear mapping function to be learned. C is the tradeoff between the generalization and the upper and lower training error ξ_i and , ξ_i^* subject to a threshold ϵ.

Gaussian kernel are used in our SVR, and their parameters are selected via cross validation. In this experiment, we subtract the mean value of each patch from its pixel values before calculating the neighbor coefficient α ; this mean value is also subtracted from the corresponding pixel value in the original HR structure image. In testing, the mean value of each patch will be added to the predicted output pixel value.

After the SVR model for each patch is learned, we use them to predict the HR version structure image of a given LR structure image. As a post-processing step, we add an enhancement filter which consists of a HPF and a TV filter. Then the edge component is sharpened by the HPF. The overshoot and undershoot is removed by the TV filter. Finally, we get the final HR structure image after this enhancing step.

Ⅲ. Experimental Results

Given the original HR image, we obtain the input LR image through blurring and down-sampling. The input LR image is decomposed into the texture component and the structure component by the TV-G decomposition. The LR structure component is up-sampled by the proposed SVR based up-sampling to improve the resolution. The SVR model is trained by LIBSVM [11]. In the proposed method, we update the pixel values in the HR version structure image using the learned SVRs and obtain the intermediate HR structure image. As a post-processing step, we add an enhancement filter to make the edge component sharp. The LR texture component is up-sampled by the NNE based learning method as in the same manner of the conventional method. Finally, the output HR image is obtained by adding the HR structure component and the HR texture component.

Objectively, PSNR(Peak Signal to Noise Ratio) and SSIM(Structural Similarity Index) values [12] are utilized to assess the performance of different algorithms, which are shown in the Table 1, Table 2, Table3, and Table 4. Also, the experimental results are visually evaluated.

The proposed SR method is evaluated on a variety of natural images from SET5, SET14, and URBAN100 data sets. To validate the effectiveness of the proposed SR, the proposed SR method is compared with other algorithms including Bicubic interpolation, ScSR, TV, and NNE. To make a fair comparison, the same training images are used.

Table 1. PSNR values of test images

	Barbara	Lenna	Girl	Boy	Woman
Bicubic	23.55	28.09	29.84	28.37	24.32
ScSR[6]	23.72	27.85	30.10	28.58	24.54
TV[5]	23.54	28.10	29.82	28.28	24.36
NNE[4]	23.70	28.41	30.10	28.58	24.74
Our Method	24.71	29.88	31.27	29.59	26.30

Table 2. SSIM values of test images

	Barbara	Lenna	Girl	Boy	Woman
Bicubic	0.708	0.865	0.723	0.850	0.843
ScSR[6]	0.723	0.870	0.731	0.858	0.853
TV[5]	0.706	0.864	0.721	0.850	0.842
NNE[4]	0.721	0.875	0.734	0.860	0.856
Our Method	0.761	0.900	0.762	0.881	0.876

Table 3. PSNR values of test images

	Urban	Parthenon	Boat	Zebra	Baboon
Bicubic	19.89	22.82	21.97	23.15	19.26
ScSR[6]	20.16	22.98	22.03	23.41	19.42
TV[5]	19.87	22.87	21.96	23.22	19.24
NNE[4]	20.16	23.14	22.14	23.75	19.41
Our Method	21.47	24.09	22.72	26.26	19.99

Table 4. SSIM values of test images

	Urban	Parthenon	Boat	Zebra	Baboon
Bicubic	0.353	0.628	0.596	0.705	0.477
ScSR[6]	0.383	0.641	0.611	0.729	0.500
TV[5]	0.351	0.629	0.595	0.705	0.475
NNE[4]	0.375	0.644	0.611	0.730	0.497
Our Method	0.451	0.688	0.657	0.774	0.570

As shown, the proposed SR method outperforms the other methods with respect to PSNR and SSIM.

Fig. 2. Experimental results of Barbara image
(a) Original image (b) Bicubic (c) ScSR
(d) TV (e) NNE (f) Our method

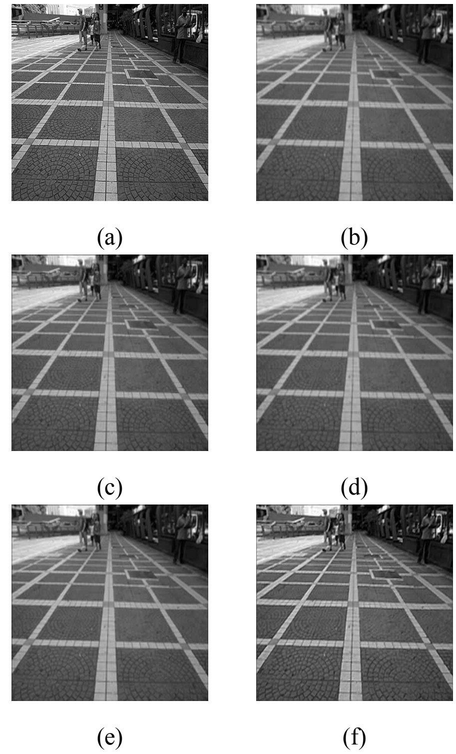

Fig. 3. Experimental results of Urban image
(a) Original image (b) Bicubic (c) ScSR
(d) TV (e) NNE (f) Our method

Fig. 2 shows the results of applying different methods. While the Bicubic, ScSR, TV, and NNE blurred the eye, eyebrow, lip and nose region, the result of the proposed SR looks more sharp in these regions. And, the proposed SR shows clear stripe pattern in a wrapping cloth when comparing with conventional SR methods. Fig. 3 shows the results of Urban image. Bicubic, ScSR, TV and NNE show more blurred and smoothed results than the proposed SR. But, the proposed SR shows clear texture pattern in the blocks of a sidewalk than those of Bicubic , ScSR, TV, and NNE. It is clear that the proposed SR is better than other algorithms in terms of the visuality, PSNR and SSIM.

Ⅳ. Conclusion

Among single image SR techniques, the TV based SR approach seems most successful in terms of edge preservation and no artifacts. In [5], an effective TV based SR approach was introduced. It achieves sharp edge preservation with no artifacts. This TV approach uses the interpolation and edge enhancement filter for the structure component and uses the interpolation for texture component. This method achieves good SR for edge components and insufficient SR for texture component.

In this paper, we proposed a new TV-G decomposition based SR method to solve this problem. The input LR image is decomposed into the texture component and the structure component. The selective SR methods suitable for each component are used in the proposed SR method. The quality of the initial structure image has a very critical impact on the final SR result. So, we proposed the SVR based up-sampling to get better initial structure image and better edge preservation in the structure component. As a post-processing step, the enhancement filter is used to enhance strong edges in an intermediate HR image by the SVR based up-sampling . The LS solution of NE is ‘too fitted’ on the LR data and thus generates undesired bad HR reconstructions. The NNE method solves this problem by using a non-negativity constraint. The generalization ability of NNE is improved by this relaxed constraint. The NNE get the increasing number of neighborhoods and produces better HR image than the NE method. So, we used the NNE based learning method to improve the resolution of the texture component. Finally, the output HR image is obtained by adding the HR structure component and the HR texture component.

We conducted the experiments on different types of images and the results are promising. The experimental results of the proposed SR method show sharp edges and fine details in visual perception. By comparing the proposed SR method to the previous works, the proposed SR approach produced very attractive SR images with better PSNR, SSIM results than those with Bicubic, ScSR, TV and NNE. In the future, we will pay more attention to the multi-scale SR problems.

References

[1] R. S. Wagner, D. E. Waagen, and M. L. Cassabaum, “Image super resolution for improved automatic target recognition,” in Proceeding of the SPIE, Vol. 5426, 2004.

[2] D. Li and Steven Simske, “Example based single frame image super resolution by support vector regression,” Patten Recognition Research, Vol. 5, No. 1, 2010.

[3] H. Chang, D. Yeung, and Y. Xiong, “Super resolution through neighbor embedding,” in Proceeding of the IEEE CVPR, Vol. 1, 2004.

[4] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi Morel, “Low Complexity Single Image Super Resolution based on Nonnegative Neighbor Embedding,” in Proceeding of the BMVC, 2012.

[5] Y. Sakuta, A. Tsutsui, T. Goto, M. Sakurai and R. Sakai, “Super Resolution utilizing Total Variation Regularization on CELL Processor,” in Proceeding of the International Conference on Consumer Electronics, 2012.

[6] Jianchao Yang, J. Wright, T.S. Huang, and Yi Ma., “Image Super Resolution Via Sparse Representation,” IEEE Trans. on Image Processing, Vol. 19, No. 11, 2010.

[7] J. F. Aujol, G. Gilboa, T. Chan, and S. Osher., Structure texture image decomposition – modeling, algorithms, and parameter selection, UCLA, CAM Report 05-10, 2005.

[8] D. Gabor., “Theory of communication,” Journal Inst. of Electrical Engineering, Vol. 93, No. 3, 1946.

[9] J. F. Aujol, G. Gilboa, T. Chan, and S. Osher., Structure Texture Decomposition by a TV-Garbor Model, UCLA, CAM Report 05-11, 2005.

[10] V. Vapnik, Statistical Learning Theory, Wiley Interscience, 1998.

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

[12] J. W. Kim and Y. S. Choi, “A Steganography Method Improving Image Quality and Minimizing Image Degradation,” Journal of DCS, Vol. 17, No. 5, 2016.

Kyoung-Bae Eum

1984 : Chonbuk National University (B.S.)

1986 : Chonbuk National University (M.S.)

1990 : Chonbuk National University (Ph.D.)

1995~1996 : Postdoctoral Researcher, Univ. of Toronto, Canada

1999, 2000 : Visiting Scholar, Univ. of Tokyo, Japan

2011~2012 : Visiting Scholar, Univ. of Missouri, USA

2016 : Visiting Scholar, Simon Fraser University, Canada

1989~current : Professor, Kunsan National University

※Research Interests : Image Processing, Computer Vision

Dong-Kyu Beom

2012~current : Undergraduate Student, Kunsan National University

※Research Interests : Image Processing, Computer Vision