Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 13 (15): 3073-3078, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.3073.3078

Color Super-Resolution Reconstruction Based on A Novel Variation Model

Zhigang Xu and Honglei Zhu
School of Computer and Communication, Lanzhou University of Technology, No. 287, Langongping Road, Qilihe District, Lanzhou 730050, Gansu, China

Abstract: This study proposes a new color image super-resolution method in an adaptive and robust framework. In the framework, three channels of color images are reconstructed, respectively. Total variation regularization is used for edge preservation of single-channel component. In order to overcome the shortcoming of the total variation model which often leads to generate an undesirable staircase effect, a spatial adaptive function was designed based on locally structural features of images. This function was used to couple the low-order total-variation model with the beyond digital total-variation model. Then, an adaptive variational color super-resolution algorithm of image sequences was proposed to preserve image edge structures and remove the color staircase effect. Experimental results on real images were presented which demonstrate the effectiveness of the proposed method.

Key words: Super-resolution, total variation, locally structural feature, fuzzy entropy, adaptive regularization, color reconstruction

INTRODUCTION

Image Super-Resolution (SR) is a popular research area for producing High-Resolution (HR) images with better details. Such technology has wide applications in several domains of knowledge, such as surveillance, remote sensing, medicine, industrial automation, computer vision, image enhancement, high definition televisions, among others.

The goal of image super-resolution is to estimate a HR image from one or several Low Resolution (LR) images. A typical mathematical model that relates some observed LR images to an expected HR image can be described as:

$$y_k = DM_k B_k x + n_k \tag{1}$$

where, x is an expected HR image vector and the size of the vector is $LN_1 \times LN_2$; L denotes a down-sampling factor; y_k is a kth frame vector of LR images, its size is $N_1 \times N_2$; D is a $N_1N_2 \times LN_1LN_2$ down-sampling matrix; M_k denotes a motion matrix of the kth LR image; B_k is a blur matrix for representing the Point Spread Function (PSF). M_k and B_k are also $LN_1LN_2 \times LN_1LN_2$ matrixs; n_k represents an $N_1N_2 \times 1$ additive noise vector.

The reconstruction of the original HR image is a typical ill-posed inverse problem. There are mainly two categories of solution for this problem: reconstruction-based methods and learning-based methods. The reconstruction-based approaches estimate a HR image by enforcing some prior knowledge which is typically designed to reduce edge artifacts. Liu and Zhu (2007) presented an improved recursive steepest-descend iteration algorithm based on the regularizing least-mean-square reconstruction. Li *et al*. (2010) used two new regularization norms, termed as locally adaptive bilateral total variation and consistency of gradients, to improve the reconstructed images. Gevrekci *et al*. (2007) employed a new constraint set based on the spatio-intensity neighborhood and reconstructed a HR image based on Projection onto Convex Sets (POCS). Li *et al*. (2010) presented a variational approach to obtain HR images from multiframe low-resolution video stills. They adopted a two-phase iterative method for super-resolution reconstruction to represent a high-resolution image.

The learning-based approaches reconstructed high frequency details from a large training set of HR image patches that encode the relationship between HR and LR images. Yang *et al*. (2010) presented a method toward single image super-resolution based on sparse representations in terms of coupled dictionaries jointly trained from HR and LR image patch pairs. They sought a sparse representation for each patch of the LR input, and then used the sparse coefficients of this representation to estimate the HR image. Rezio *et al*. (2012) presented a learning-based method which

Corresponding Author: Honglei Zhu, School of Computer and Communication, Lanzhou University of Technology, Lanzhou, Gansu, China

generated HR images based on Locally Linear Embedding (LLE) and residual compensation. They used feature descriptors to represent HR residual images. Additionally, they applied the gradient profile prior for estimating the initial super-resolved image.

Nowadays, almost all super-resolution methods have still been designed to increase the resolution of monochromatic image. But there is little work addressing the problem of color SR. In this paper, we proposed a novel variational color image SR algorithm based on the new model of locally coupling digital total variation.

The remainder of this paper was organized as follows. In Section 2, we discussed the methodology of color super-resolution. Section 3 studies how to apply our method for generic color image super-resolution. We detail our formulation and solution to the image super-resolution problem based on locally coupling digital total-variation model. Various experimental results in Section IV demonstrate the efficacy of the new model as a prior for regularizing color image super-resolution.

COLOR SUPER-RESOLUTION

The currently applications of color images are wide. Relative to grayscale images, color images contain more information. But color SR faced some new problems and challenges. There are two reasons. First, a color image has three channels. There are strong correlations between the channels. Second, the human eye is more sensitive to color changes than gray scale change. The color artifacts in the reconstructed image are easily to perceive. According to the application conditions of color channel information, color SR can be categorized into two classes: separated reconstruction approaches and combined reconstruction approaches.

The separated reconstruction methods have two ideas. One applies monochromatic SR algorithms to each of the color components independently. The color HR image can be generated by combining three separate monochromatic images. The other is transferring the problem to a different color space where chrominance components are separated from luminance, and where Super-Resolution is applied to the luminance channel only. Since the human eye is more sensitive to the details in the luminance channel of an image than the details in the chrominance channels. Xiong *et al.* (2010) presented a practical solution which combines adaptive regularization and learning-based algorithm for color SR. This method performed the regularization on both the luminance and chrominance layers. At the same time, Learning-based pair matching was only performed on the luminance component. Sun *et al.* (2011) transformed the image from RGB color space to YUV color space. They only implemented image super-resolution on the Y layer. Firstly, they used the gradient profile prior and inferred the HR gradient field from the LR image by gradient field transformation. Secondly, the estimated HR gradient field was imposed a gradient domain constraint for the HR image. Finally, the HR image was inferred by coupling the gradient domain constraint and the image reconstruction constraint. Additionally, they used the bicubic interpolation to reconstruct the color components U and V.

The combined reconstruction methods use three color components to estimate the result of reconstruction. The methods strengthen the correlation between the channels. It is possible to obtain a better reconstructed image. However, applying super-resolution techniques to these components increases the computational burden. Sorrentino and Antoniou (2009) proposed two regularization terms and an intercolor dependencies penalty term operating in the YIQ color space, which penalize differences in edge locations and orientation across color channels. Li *et al.* (2010) described a SR method on color face image based on quaternion, combining with the ideas of Principal Component Analysis (PCA) and interpolation. They utilized the quaternion to deal with the three channels of R, G and B at the same time in order to increase the efficiency of the operation and use fully the correction of three channels. Maalouf and Larabi (2012) designed a grouplet-based structure tensor to combine geometric information of the different color channels of the image. They defined a functional on the multispectral geometry by the grouplet-based structure tensor. Then, this functional was minimized to insure the reconstruction of the SR image.

TOTAL-VARIATION-BASED COLOR SUPER-RESOLUTION RECONSTRUCTION

Super-resolution using the total variation model: Total variation is a powerful concept for robust estimation. The main advantage of TV model is that it recovers edges well and removes noise avoiding the ringing effect because the model does not penalize discontinuities in images. Arguably, the TV model realized a good balance between the ability to describe piecewise smooth images and the complexity of the resulting algorithms.

Since the advantages of the TV model, the TV-based regularizations have been widely utilized in denoising and SR in literature (Marquina and Osher, 2008; Xiao and Wei, 2011). Ng *et al.* (2007) proposed a SR algorithm to generate HR digital videos. They employed the TV regularization in the reconstruction model, termed as

Low-order Total Variation (LoTV). And they used fixed-point and preconditioning methods to solve the nonlinear TV-based SR reconstruction model. The LoTV model is defined as:

$$J_{TV}(f) = \sum_{\Omega} \sqrt{\left|\nabla f^1_{i,j}\right|^2 + \left|\nabla f^2_{i,j}\right|^2 + \varepsilon} \tag{2}$$

where, $\nabla f^1_{i,j} = f(i+1, j) - f(i, j)$ and $\nabla f^2_{i,j} = f(i, j+1) - f(i, j)$. ε is a small positive par a meter which ensures differentiability. Shao (2008) described the so-called Beyond Digital-TV (BDTV) model for SR. The BDTV model can preserve more details (e.g., edges) and inhibit the staircase effect. The model not only could ensure the flat areas of images, but also keep the clarity of important geometric structures. The BDTV model can be described as follows:

$$J_{BDTV}(f) = \sum_{\Omega} \sqrt{\left|\nabla f^1_{i,j}\right|^2 + \left|\nabla f^2_{i,j}\right|^2 + \left|\nabla f^3_{i,j}\right|^2 + \left|\nabla f^4_{i,j}\right|^2} \tag{3}$$

where $\nabla f^1_{i,j}$, $\nabla f^2_{i,j}$, $\nabla f^3_{i,j}$ and $\nabla f^4_{i,j}$ mean the differentials between the pixel f (i, j) and its four neighbors, respectively.

These total variation regularization models perform very well for some SR tasks. However, the approaches fail to consider the partial smoothness of an image. They are not locally adaptive, and thus they have limited adaptive capability in the process of SR reconstruction and cannot balance the suppression of noise against the preservation of image details. So the staircase effects maybe appear in homogeneous parts of the generated HR image. The smaller details, such as texture, are easily destroyed.

Color super-resolution based on the locally coupling digital total variation model: In response to these problems, we can use local structure features of LR images to distinguish edges and flat areas, because the local contrast of the pixels in flat areas is relatively small, while the contrast of edges is large. In addition, there are many small edges and details in natural images. Therefore, the regularization should be spatially adaptive to achieve anisotropic diffusion. This paper proposed a new TV model, termed as Locally Coupling Digital Total-variation model (LCDTV). The LCDTV is defined as:

$$J_{LCDTV}(f) = \sum_{\Omega} \sqrt{\left|\nabla f^1_{i,j}\right|^2 + \left|\nabla f^2_{i,j}\right|^2 + (1 - \Psi(i,j))\left(\left|\nabla f^3_{i,j}\right|^2 + \left|\nabla f^4_{i,j}\right|^2\right)} \tag{4}$$

where, $\Psi()$ is a locally adaptive measure function. The range of Ψ (i, j) is [0,1].

In this study, the locally adaptive measure function is designed to obtain from image local features, which responses to local smoothness. We applied a fuzzy entropy function $\Psi()$ to measure the local smoothness for the k×k patch centered at (i, j). $\Psi()$ looks like:

$$\Psi(i,j) = 1 - \frac{1}{\left(1 + \mathrm{abs}\left(\varphi_{i,j}\right)^m\right)^n} \tag{5}$$

where, $\varphi_{i,j}$ is a function of the fuzzy entropy-based neighborhood homogeneity (Jaffar *et al*., 2010). It can be described as follows:

$$\varphi_{i,j} = \frac{1}{k^2} \sum_{p=-k/2}^{k/2} \sum_{q=-k/2}^{k/2} H\left(u\left(f_{i+p,j+q}\right)\right) \tag{6}$$

where, u ($f_{i,j}$) is a fuzzy membership function of $f_{i,j}$, H(•) is the fuzzy entropy of u ($f_{i,j}$), H(•) can be written as:

$$H(u) = -u * \ln(u) - (1-u) * \ln(1-u) \tag{7}$$

And u ($f_{i,j}$) is defined as:

$$u_{s,t}\left(f_{i,j}\right) = \frac{1}{1 + \left|f_{i+s,j+t} - f_{i,j}\right| / W} \tag{8}$$

where, s and t mean an image patch which sizes is S×T and its center is (i, j), W is a constant of proportionality. It is obviously that Eq. 4 can reduce to the BDTV model when $\Psi()$ should be set with small values, i.e., $\Psi() \rightarrow 0$, while Eq. 4 can reduce to the LoTV model when $\Psi()$ is close to 1. The generalization of LDBTV is much more powerful because we exploited local features in the reconstruction. The edge preservation of this model approaches the LoTV model. And the performance of restoring texture information is close to the BDTV model. So our model provides efficient edge structure and detail preservation for natural images.

Based on the proposed LCDTV regularization term in Eq. 4, the Lagrangian energy function for HR color image reconstruction can be obtained:

$$\hat{f}_i = \arg\min_{f_i} \left[\sum_{k=1}^{p} \| g_{k,i} - DM_k B_k f_i \|^2 + \lambda J_{LCDTV}(f_i) \right] \tag{9}$$

$$I = R, G, B$$

where, $g_{k\,i}$ is the i component of kth LR image and k = 1, ..., p. λ is the regularization parameter.

The Euler-Lagrange equation for the energy functional can be given and the steepest descent algorithm is adopted in this study to find a solution of 9.

This leads to the following iteration:

$$\hat{f}_i^{n+1} = \hat{f}_i^n - \beta \left[\sum_{k=1}^{p} \left(\begin{array}{l} \left(B_k^T M_k^T D^T g_{k,i} - B_k^T M_k^T D^T D M_k B_k \hat{f}_i^n \right) \\ +\lambda L(\hat{f}_i^n)\hat{f}_i^n \end{array} \right) \right] \quad (10)$$

I = R, G, B

where, L (f) is an elliptic partial differential operator:

$$\nabla\left(\nabla / \sqrt{|\nabla f|^2}\right)$$

With an initial guess f^0, the high-resolution color image can be finally implemented by Eq. 10.

EXPERIMENT AND ANALYSIS

In this section, we tested our method on a variety of images. And we compare our approach with other state-of-the-art image super-resolution methods including the sparse SR method in (Yang *et al*., 2010) and Normalized Convolution (Pham *et al*., 2006). The Peak Signal to Noise Ratio (PSNR) and the Mean Structural Similarity (MSSIM) (Wang *et al*., 2004) were employed to measure results. In order to evaluate, the reconstruction results were transformed from RGB to YCbCr and only Y components of the results were calculated.

MSSIM is defined by:

$$\mathrm{MSSIM}(f,g) = \frac{1}{N}\sum_{i=1}^{N} \frac{(2\mu_{f_i}\mu_{g_i}+1)(2\sigma_{f_i}\sigma_{g_i}+9)(\sigma_{f_i g_i}+4.5)}{(\mu_{f_i}^2+\mu_{g_i}^2+1)(\sigma_{f_i}^2+\sigma_{g_i}^2+9)(\sigma_{f_i}\sigma_{g_i}+4.5)} \quad (11)$$

where, f and *g* are the original and the desired HR images, respectively; f_i and g_i are the image patches at the i-th local window; u_x is the mean intensity of the image patch f_i or g_i; σ_x is the standard deviation of the patch; σ_{xy} is the correlation coefficient of the patches f_i and g_i and N is the number of local windows in the image.

Figure 1 shows four HR natural images which were used as the experimental Data. The images were degraded by a symmetric Gaussian low-pass filter of size 5×5 with standard deviation 1. The degraded images were down-sampled horizontally and vertically with factor 2. And the Gaussian white noise with zero mean and 0.001 variance was introduced on the LR images. Additionally, these LR images were transformed by predefined shift vectors.

The experiment was conducted to compare the performance of the above-mentioned three methods. The parameters selection criterion for the methods was to choose parameters to produce most appealing results. The sparse SR method needs to train two dictionaries for the low- and high-resolution image patches. The size of LR patches is 3×3, with overlap of 2 between adjacent patches. And the HR patches is 6×6, with overlap of 4 between adjacent patches. The sparse regularization parameter was set $\lambda = 0.15$.

The visual results were presented in Fig. 2. As it is shown in Fig. 2, the reconstructed edges of the sparse SR method are relatively clear. But the sparse SR method using only the first frame of LR sequence to reconstruct, the effect of image noise suppression is not good. At the same time, the sparse SR method applies color space transformation from RGB to YCbCr and reconstructs the

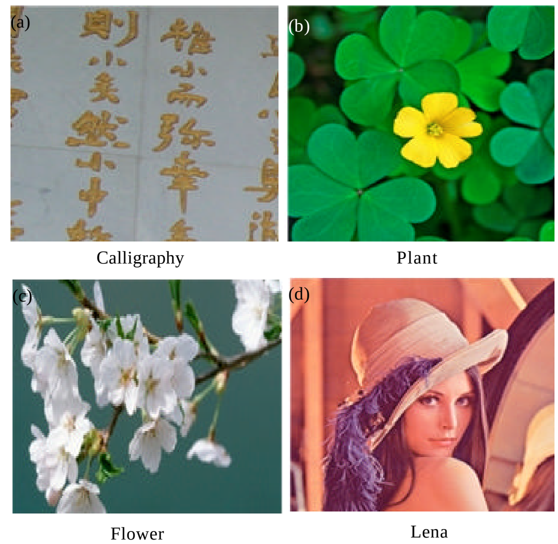

Fig. 1: Experimental images

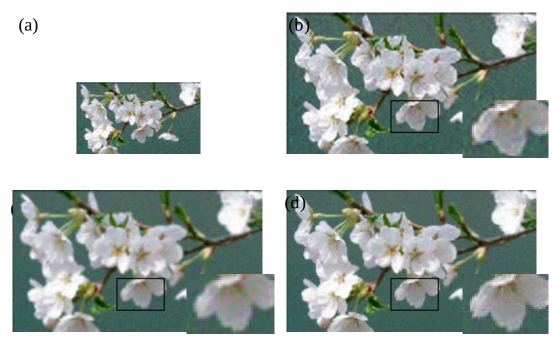

Fig. 2(a-d): Super-resolution on natural images, (a) one of the LR input images, (b) the result of sparse SR, (c) the result of NC and (d) the result of our method

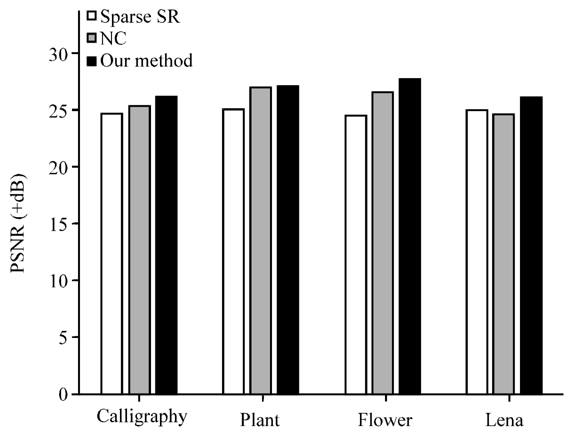

Fig. 3: PSNR comparison of three super-resolution methods

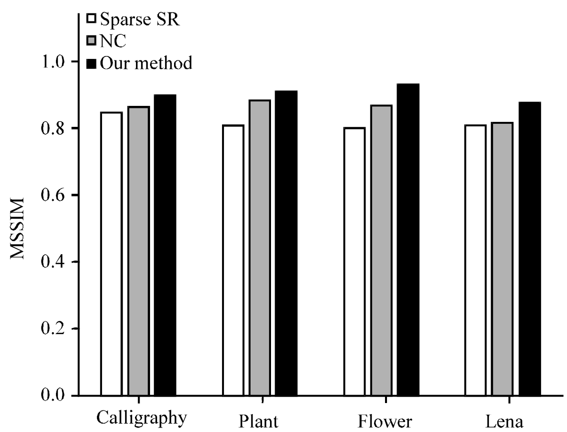

Fig. 4: MSSIM comparison of three super-resolution methods

luminance channel only. Color artifacts of the reconstructed image are obvious. The reconstructed image of NC method introduces undesired smoothing. Fig. 2d displays the result of our method. It can be seen that the preservation of edges is effective. And the obvious color artifacts do not been detected. The PSNR and the MSSIM comparisons of three methods are shown in Fig. 3 and 4, respectively. As we can see, our approach showed higher PSNR and MSSIM scores than the other methods.

CONCLUSION

The color super-resolution reconstruction is a more complex problem than the monochrome super-resolution reconstruction. Color super-resolution needs not only to consider the preservation and restoration of edge and texture details but also to consider the problem of matching color information. In this study, a new adaptive variational framework has been presented for robust color image super-resolution. A spatial adaptive function was designed based on locally structural features of images. The function coupled the low-order total-variation model with the beyond digital total-variation model. Subsequently, a unified framework based on the Euler-Lagrange equations was deduced to solve the color SR problem. Experimental results on real images have demonstrated the superior performance of the proposed algorithm.

ACKNOWLEDGMENTS

This research was supported in part by the Nature Science Foundation of Gansu Province, No: 1112RJZA015 and 1212RJYA017.

REFERENCES

Gevrekci, M., B.K. Gunturk and Y. Altunbasak, 2007. Pocs-based restoration of bayer-sampled image sequences. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, April 15-20, 2007, Honolulu, HI, pp: I-753-I-756.

Jaffar, M.A., A. Hussain and A.M. Mirza, 2010. Fuzzy entropy based optimization of clusters for the segmentation of lungs in ct scanned images. Knowl. Inform. Syst., 24: 91-111.

Li, T., X.H. Wang, G.Q. Song, J.K. Li and X.M. Yan, 2010. Color face image super-resolution reconstruction based on learning. Trans. Beijing Inst. Technol., 30: 193-196.

Li, X., Y. Hu, X. Gao, D. Tao and B. Ning, 2010. A multi-frame image super-resolution method. Signal Process., 90: 405-414.

Li, Y.R., D.Q. Dai and L. Shen, 2010. Multiframe super-resolution reconstruction using sparse directional regularization. IEEE Trans. Circuits Syst. Video Technol., 20: 945-956.

Liu, F. and X.C. Zhu, 2007. An improved superresolution reconstruction algorithm based on multiframe image. J. Nanjing Univ. Posts Telecommun. Nat. Sci., 27: 19-23.

Maalouf, A. and M.C. Larabi, 2012. Colour image super-resolution using geometric grouplets. IET Image Process., 6: 168-180.

Marquina, A. and S.J. Osher, 2008. Image super-resolution by tv-regularization and bregman iteration. J. Scientific Comput., 37: 367-382.

Ng, M.K., H. Shen, E.Y. Lam and L. Zhang, 2007. A total variation regularization based super-resolution reconstruction algorithm for digital video. EURASIP J. Adv. Signal Process., Vol. 2007., 10.1155/2007/74585

Pham, T.Q., L.J. van Vliet and K. Schutte, 2006. Robust fusion of irregularly sampled data using adaptive normalized convolution. EURASIP J. Applied Signal Process., 10.1155/ASP/2006/83268

Rezio, A.C.C., W.R. Schwartz and H. Pedrini, 2012. Image sequence super-resolution based on learning using feature descriptors. Proceedings of the International Conference on Computer Vision Theory and Application, February, 24-26 2012, Rome, Italy, pp: 135-144.

Shao, W., 2008. Image modeling based regularized multi-frame super-resolution reconstruction. Ph. D. Thesis, Dissertation, Nanjing University of Science and Technology, Nanjing.

Sorrentino, D.A. and A. Antoniou, 2009. Storage-efficient quasi-newton algorithms for image super-resolution. Proceedings of the 16th International Conference on Digital Signal Processing, July 5-7, 2009, Santorini-Hellas, pp: 1-6.

Sun, J., J. Sun, Z. Xu and H.Y. Shum, 2011. Gradient profile prior and its applications in image super-resolution and enhancement. IEEE Trans. Image Process., 20: 1529-1542.

Wang, Z., A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process., 13: 600-612.

Xiao, L. and Z.H. Wei, 2011. Research on non-local regularization model and algorithms for image super-resolution reconstruction. Chinese J. Comput., 34: 931-942.

Xiong, Z., X. Sun and F. Wu, 2010. Robust web image/video super-resolution. IEEE Trans. Image Process., 19: 2017-2028.

Yang, J., J. Wright, T.S. Huang and Y. Ma, 2010. Image super-resolution via sparse representation. IEEE Trans. Image Process., 19: 2861-2873.