Image Super-Resolution using a Hybrid Scheme with DCT Interpolation and Sparse Representation Method

Saulo R. S. Reis and Graça Bressan

Abstract — **Learning-based Super-Resolution methods have attracted much interest in recent years in many signal and image processing tasks. In this paper, we present an algorithm for single image super-resolution that use discrete cosine transform (DCT) interpolation and sparse learning-based super-resolution method. The input LR image is interpolated using both DCT interpolation and bicubic interpolation methods. The patches of bicubic interpolated image, undergoes a process sparse coding using OMP algorithm and training using k-SVD algorithm. The obtained sparse coefficients are multiplied with high-resolution dictionary generated in the training phase, resulting in the intermediate HR image. The final HR image is obtained by adding the DCT interpolated image and intermediate HR image. The experimental results demonstrate the effectiveness of the method proposed in terms of PSNR, SSIM and visual quality.**

Keywords — Super-Resolution, DCT domain, learning-based method, sparse representation.

I. INTRODUCTION

Super-Resolution (SR) is the technique of generate a high-resolution (HR) image from a sequence of observed low-resolution (LR) images. With the emergence of digital imaging devices and the need to improve of quality the captured images, super-resolution has become a very interesting research field in image processing and computer vision. Important applications that can be benefited by SR methods include remote sensing, medical imaging, surveillance and mobile TV. In the last years, numerous SR algorithms have been proposed in the literature. These algorithms can be divided into three main categories: interpolation-based methods [1-3], reconstruction-based methods [4-6] and learning-based methods [7-9].

The Interpolation-based methods are the most popular techniques used today to upscale a LR Image. These methods utilize a base function or kernel to estimate the pixels in HR grid. However, the HR images generated using interpolation-based cannot reconstruct high-frequency components and the final results are poor in terms of visual quality. The reconstruction-based methods assume that the high frequency components of image are present in the sub-pixels misalignments of LR observed images. These methods have been shown better results than interpolation-based methods, but with an increase in computational complexity. More recently, the learning-based methods aim to reconstruct a high-resolution image from the extraction of features contained in a training set composed of HR and LR image patches. Extensive research results have demonstrated their competitive performance.

In this paper, we propose a hybrid algorithm to single image super-resolution that combine the discrete cosine transformation (DCT) interpolation [10-12] and sparse representation learning-based SR method [7,8]. One important property of the DCT interpolation is better preserve the low-frequency components than others interpolations methods, such as bilinear e bicubic. The sparse representation learning-based method has been shown effective results in the reconstruction of high-frequency components using a reduced image data set. Hence, by exploiting these two important characteristics, we can achieve effective results in terms of Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) and visual quality.

The remainder of this paper is organized as follows. In section II, we briefly review the image acquisition model, DCT interpolation and sparse representation model used in the proposed method. In section III, we detailed the proposed method. Experimental results are presented in section IV and we conclude this paper in section V.

II. IMAGE SUPER-RESOLUTION MODEL

A. Image Acquisition Model

In this section, we present a traditional observation model for still images [4-9] in Super-Resolution Methods. This model includes the effects imposed during the acquisition process such as optical distortion, blurring and noise:

$$Y_L = SHX_H + \eta \qquad (1)$$

where $Y_L = [y_{L1}, y_{L2}, \cdots, y_{LN}]^T$ is a lexicographically ordered vector size $N = N_1 \times N_2$ pixels, which represents the observed LR image. X_H denotes the HR image of size $M = M_1 \times M_2$ pixels, represented by a lexicographically ordered vector as $X_H = [x_{H1}, x_{H2}, \cdots, x_{HM}]^T$ with $M > N$. The matrix H of size $M \times M$ represents the effects of the imaging system, such as optical distortion and blurring. The matrix S represents the downsampling operator, and η is zero-mean, white noise with variance σ_η^2.

The fundamental SR problem is the recovery of X_H from Y_L in eq. (1), without amplifying the effects of noise, blurring in the reconstructed HR image.

Saulo R. S. dos Reis and Graça Bressan, Departamento de Engenharia de Computação e Sistemas Digitais, Universidade de São Paulo, São Paulo-SP, Brasil, E-mails: saulo.reis@usp.br, g.bressan@larc.usp.br.

B. DCT interpolation

The discrete cosine transform (DCT) has been widely applied in areas such as compression, filtering and feature extraction. The main property of the DCT transform is to concentrate energy values in regions near to DC components, which does not cause much information loss, when eliminating high frequency components of image. The interpolation process using DCT transform, upscale the image, by adding coefficients of zero amplitude in high frequency components [10, 13]. As a simple example, suppose that the upscale factor equal to two and the LR image Y_L in eq. (1), the interpolation process is initially performed by applying the DCT transformed into (4×4) blocks. After adding coefficients of zero amplitude, the inverse transform DCT (8×8) is applied, resulting in the final interpolated image. The overall interpolation process is illustrated in Fig. 1. The experimental results have demonstrate the effectiveness of the DCT transform results in terms of visual quality than other interpolation methods, such as bilinear or bicubic. However, because of blocking effects, the results are still not satisfactory for upscale factors greater than 2.

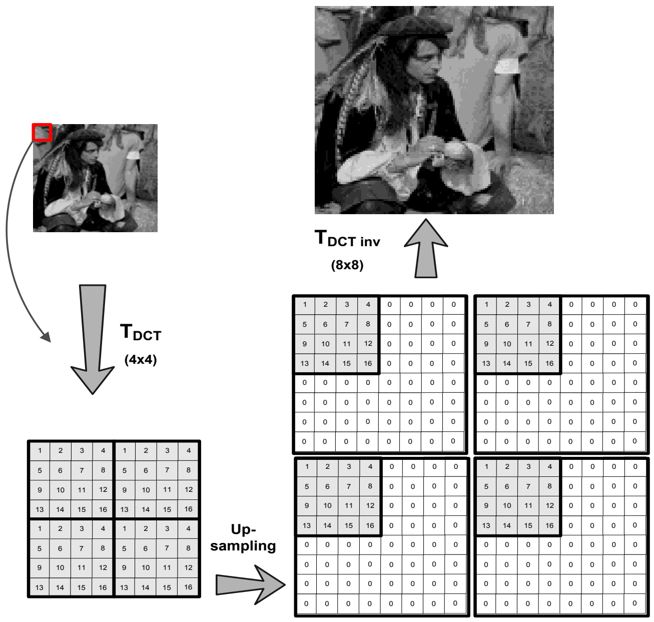

Fig. 1. Interpolation method in the discrete cosine transform (DCT) domain

C. Sparse Representation Model

The sparse representation is a process to represent a natural signal as a sparse combination of atoms with respect to a overcomplete dictionary. Given a signal $z \in \mathcal{R}^n$ and a dictionary $D \in \mathcal{R}^{n\times K}$ that contains K prototype signal-atoms for columns [14], the sparse representation can be defined as follow:

$$\min_\alpha \|\alpha\|_0 \qquad s.t.\ z = D\alpha \tag{2}$$

where α denotes the sparse vector of z and $\|.\|_0$ is the l^0 norm. Thus, for example, a LR patch $p_l^i = R_i \cdot Y_L$, can be represented as a sparse combination with respect to LR dictionary D_L:

$$D_L, \alpha = \arg min \|\alpha\|_0 \quad s.t. \|R_i \cdot Y_L - D_L\alpha\|_2^2 \leq \varepsilon \tag{3}$$

where R_i is a patch extraction operator in location i, and ε is related to noise. Likewise, other important characteristic of sparse representation model is the use of a reduced image data set when compared with traditional learning-based methods. Sparse model have been found in several applications, such as, image compression, feature extraction, regularization in inverse problems, and more

III. SR PROPOSED METHOD

In this section, we propose a single image SR method, by combining the interpolation DCT and sparse learning-based method. The key idea is to exploit of advantages of DCT interpolation in terms of visual quality and sparse representation in terms of reconstruction high-frequency components. The proposed algorithm can be summarized by the following steps:

1) The input LR image Y_L in eq. (1) is first interpolated using both DCT interpolation and bicubic interpolation resulting in the interpolated images Y_{BIC}^{LR} and Y_{DCT}^{LR}:

$$Y_{BIC}^{LR} = Q_{BIC} Y_L \tag{4}$$

$$Y_{DCT}^{LR} = Q_{DCT} Y_L \tag{5}$$

where Q_{BIC} and Q_{DCT} denote the bicubic and DCT interpolation operators respectively.

2) The interpolated image Y_{BIC}^{LR} undergoes a process of patch extraction and filtering using a set of high-pass filters. The Gradient and Laplacian high-pass filters used in [7, 8] are the same used in our algorithm:

$$p_l^i = R_i \cdot Y_{BIC}^{LR} \tag{6}$$

$$\tilde{p}_l^i = F_r * p_l^i \tag{7}$$

where p_l^i in eq. (6), represent the LR image patch. In eq. (7), $\tilde{p}_l^i$ is the LR image patch after filtering process, $F_r(r = 1, \cdots, R)$ represents a set of R high-pass filters and "*" denote convolution operation. In this case, is important to observe that $\tilde{p}_l^i > p_l^i$.

3) Considering the high dimensionality of LR patch $\tilde{p}_l^i$, a dimensionality reduction using Principal Component Analysis (PCA) algorithm is employed, in order to reduce computational complexity.

$$\hat{p}_l^i = B \cdot \tilde{p}_l^i \tag{8}$$

where B denote PCA operator.

4) Considering the resulting LR patches $\hat{p}_l^i$, the k-SVD iterative training algorithm [14] is applied, resulting in a LR dictionary D_L and a sparse vector α_i as follow:

$$D_L, \alpha_i = \arg min\|\alpha_i\|_0 \quad s.t. \ \left\|\hat{p}_l^i - D_L\alpha_i\right\|_2^2 \leq \varepsilon \quad (9)$$

5) To simplifies the computational cost, we use the same pseudo-inverse expression [7] to obtain a HR dictionary D_H. After this, we can obtain HR image patch p_h^i as follow:

$$p_h^i = D_H\alpha_i \quad (10)$$

6) Thus, the HR patches are aggregated using least-square solution, as follow :

$$I_{HR} = \left[\sum_i R_i^T R_i\right]^{-1} \sum_i R_i^T p_h^i \quad (11)$$

7) The final HR image is obtained by adding the DCT interpolated image Y_{DCT}^{LR} and image I_{HR} :

$$I_{HRF} = I_{HR} + Y_{DCT}^{LR} \quad (12)$$

The proposed method is also illustrated in fig. 2.

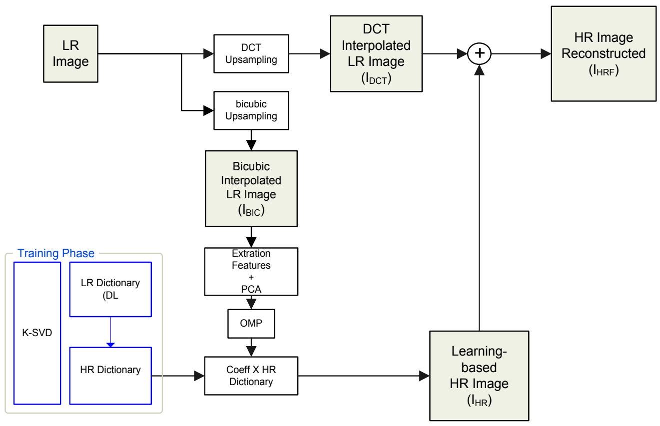

Fig. 2. SR proposed method using DCT interpolation and sparse representation

IV. EXPERIMENTS RESULTS

To demonstrate the performance of the proposed method, we conducted experiments with the classical images Lena, Bridge, Man, Pepper, Barbara, Coastguard and Monarch. We tested the propose method for upscale factor equal to two for all the experiments. The data set used for iterative training of D_L and D_H dictionaries, was composed by natural images with different characteristics and sizes, following the same data set used in [7, 8]. For the features extraction, we applied four 1-D filters Laplacian e Gradient used in [7, 8].

The experiments were conducted by utilizing 30 iterations of k-SVD algorithm for construction of D_L and D_H dictionaries. In Table I, we compare the proposed method with other methods, including bicubic interpolation and sparse-based SR proposed in [7].

TABLE I. SSIM RESULTS

Image	Size	Bicubic Interp.	DCT interp.	SMSR [7]	Our method
Lena	512×512	0,9903	0,9937	0,9965	0,9965
Bridge	512×512	0,9737	0,9835	0,9914	**0,9920**
Man	512×512	0,9801	0,9870	0,9930	**0,9933**
Barbara	720×576	0,9635	0,9789	0,9851	**0,9867**
Coastguard	352×288	0,7907	0,8053	0,8409	0,8397
Pepper	512×512	0,9926	0,9951	0,9971	0,9971
Monarch	768×517	0,9949	0,9968	0,9985	0,9983

In Table II, we also compare with the same methods, but in terms of PSNR.

TABLE II. PSNR RESULTS

Image	Size	Bicubic Interp.	DCT interp.	SMSR [7]	Our method
Lena	512×512	34,696	34,743	36,209	35,563
Bridge	512×512	26,523	26,629	27,436	27,173
Man	512×512	27,938	28,014	29,133	28,818
Barbara	720×576	27,938	28,002	28,627	28,387
Coastguard	352×288	29,138	29,301	30,384	30,063
Pepper	512×512	33,074	33,416	34,510	34,536
Monarch	768 ×517	32,970	33,016	35,670	34,708

The results shown that the proposed method reaches better results for images Bridge, Man and Barbara than other methods, in terms of SSIM. For others images, the results achieve the same level to the sparse-based method [7]. We can observe that in terms of PSNR, the results achieved the same level to the Sparse-based Method in [7]. In fig. 3 we compare the quality of the proposed method with the results produced by others method. We can observe an improvement of quality of final HR image.

IV. CONCLUSIONS

In this paper, we presented a single image super-resolution method combining the interpolation DCT and sparse learning-based method. The aim was to explore important characteristics of these two methods to achieve an enhancement in the final HR image. The experimental results demonstrate the effectiveness of the algorithm in terms of PSNR, SSIM and visual quality. For future investigations, the goal is to improve the visual quality of final HR images for upscale factors greater to two, and testing the algorithm proposed for SR videos and other test in noisier environments.

Fig. 3. Visual quality test for images “Man”, “Bridge” and “Pepper”. (a) Original Image, (b) Bicubic Interpolated, (c) Sparse Representation SR Method Ref. [7], (d) Our proposed Method

REFERENCES

[1] L. Zhang, and W. Xiaoling, “An Edge-Guided Image Interpolation Algorithm via Directional Filtering and Data Fusion,” *IEEE Transactions on Image Processing,* vol. 15, no. 8, pp. 2226-2238, 2006.

[2] X. Li, and M. Orchard, “New Edge-Directed Interpolation,” *IEEE Transactions on Image Processing,* vol. 10, no. 10, pp. 1521-1527, 2001.

[3] H. A. Aly, and E. Dubois, “Image Up-Sampling Using Total-Variation Regularization With a New Observation Model ” *IEEE Transactions on Image Processing,* vol. 14, no. 10, pp. 1647-1659, 2005.

[4] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical overview,” *IEEE Signal Processing Magazine,* vol. 20, no. 3, pp. 21-36, 2003.

[5] M. V. W. Zibetti, F. S. V. Bazan, and J. Mayer, “Estimation of the parameters in regularized simultaneous super-resolution,” *Pattern Recognition Letters,* vol. 32, no. 1, pp. 69-78, 2011.

[6] S. Baker, and T. Kanade, “Limits on super-resolution and how to break them,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, no. 9, pp. 1167-1183, Sep, 2002.

[7] R. Zeyde, M. Elad, and M. Protter, “On single image scale-up using sparse-representations,” in Proceedings of the 7th international conference on Curves and Surfaces, Avignon, France, 2012.

[8] J. Yang, J. Wright, T. S. Huang *et al.*, “Image Super-Resolution Via Sparse Representation,” *IEEE Transactions on Image Processing,* vol. 19, no. 11, Nov, 2010.

[9] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example-based super-resolution,” *IEEE Computer Graphics and Applications,* vol. 22, no. 2, pp. 56-65, Mar-Apr, 2002.

[10] E. M. Hung, Garcia, D.C., Queiroz, R.L.D., “Transform domain semi-super resolution,” in ICIP'11, 2011, pp. 1193-1196.

[11] W. Zhang, and W.-K. Cham, “Hallucinating Face in the DCT Domain,” *IEEE Transactions on Image Processing,* vol. 20, no. 10, Oct, 2011.

[12] D. Sun, and W.-K. Cham, “Postprocessing of low bit-rate block DCT coded images based on a fields of experts prior,” *IEEE Transactions on Image Processing,* vol. 16, no. 11, Nov, 2007.

[13] Z. Wu, H. Yu, and C. W. Chen, “A New Hybrid DCT-Wiener-Based Interpolation Scheme for Video Intra Frame Up-Sampling,” *IEEE Signal Processing Letters,* vol. 17, no. 10, Oct, 2010.

[14] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on Signal Processing,* vol. 54, no. 11, Nov, 2006.