

# Chapter 8

## Image Super-Resolution Based on MCA and Dictionary Learning

Kun Zhang, Hongpeng Yin and Yi Chai

**Abstract** Image super-resolution focuses on achieving the high-resolution version of single or multiple low-resolution images. In this paper, a novel super-resolution approach based on morphological component analysis (MCA) and dictionary learning is proposed in this paper. The approach can recover each hierarchical structure well for the reconstructed image. It is integrated mainly by the dictionary learning step and high-resolution image reconstruction step. In the first step, the high-resolution and low-resolution dictionary pairs are trained based on MCA and sparse representation. In the second step, the high-resolution image is reconstructed by the fusion between the high-resolution cartoon part and texture part. The cartoon is acquired by MCA from the interpolated source image. The texture is recovered by the dictionary pairs. Experiments show that the desired super-resolution results can be achieved by the approach based on MCA and dictionary learning.

**Keywords** Image super-resolution · Dictionary learning · Morphological component analysis · Cartoon · Texture

### 8.1 Introduction

Image super-resolution (SR) can produce the high-resolution (HR) enhancement version of single or multiple observed low-resolution (LR) images without changing the low-cost image acquisition sensors [1]. Therefore, image SR attracts

---

K. Zhang · H. Yin (✉) · Y. Chai  
College of Automation, Chongqing University, Chongqing 400044, China  
e-mail: yinhongpeng@gmail.com

H. Yin  
Key Laboratory of Dependable Service Computing in Cyber Physical Society,  
Ministry of Education, Chongqing 400030, China

Y. Chai  
Key Laboratory of Power Transmission Equipment and System Security,  
Chongqing 400044, China

much attention and always plays an important role in various applications, such as the remote sensing, the satellite image application, the medical imaging diagnosis system, the digital television system. As a result, various SR methods have been proposed, mainly including the interpolation-based methods [2], the regularized reconstruction-based methods [1–3] and the learning-based methods [4–12].

The learning-based SR methods have been comprehensively researched since the example-based approach in [4] was proposed. In these methods, the prior relationship between the HR and LR training set can usually be learned in terms of the image structure and content. The relationship is very beneficial to recovering high-resolution version of the LR test image. Consequently, various learning-based SR approaches are put forward [5–12]. Referring to manifold learning strategy, Chang et al. [5] utilize the principle of local linear embedding (LLE) in SR. Their approach is helpful in decreasing the scale of training set, but the fitting problem makes the approach imperfect in image sharpness [7]. In [6], the contourlet transform is introduced into the single-image SR to capture the smoothness contours by directional decompositions. The dictionary-based learning approach via sparse representation is further introduced into image SR by Yang et al. [7, 8]. After the work of Yang, some improved approaches based on sparse representation are also proposed. In [9], a texture constrained sparse representation is proposed. However, the approach is sensitive to noise. In [10], the dictionary is trained by the difference between the HR images and LR images. The approach performs well for image denoising and SR reconstruction. In [11], a hierarchical clustering algorithm is applied to optimize the parameter for the dictionary learning way in [7].

However, the aforementioned SR approaches focus to the omitted high-frequency texture component and should pay more attention to the other component. Inspired by the morphological component analysis (MCA) [12], a novel image SR approach based on MCA and dictionary learning is applied. In our approach, MCA is utilized to decompose the interpolated LR test image into the cartoon and texture component. The texture component is discarded. The cartoon component is reserved as the HR cartoon part of the expected HR image. The expected HR image's texture is estimated by means of the HR\LR dictionary pairs established from training image set. The reconstructed image result can be obtained by combining the HR cartoon and the HR texture. Thus, each component of the reconstructed HR image can be both focused on well. Experimental results show the desired super-resolution image quality can be achieved by the approach based on MCA and dictionary learning.

In the outline of this paper, Sect. 8.2 describes our approach based on MCA and dictionary learning in detail, Sect. 8.3 conducts some experiments and analyzes the results, and Sect. 8.4 shows the conclusion for this paper briefly.

## 8.2 Our Approach for Image Super-Resolution

In the proposed approach, two steps are necessary: the dictionary learning step and the SR reconstruction step. The former can provide a pair of dictionaries for the later to obtain a high-resolution enhancement version of the LR test image. In each step, the MCA theory is utilized to decompose images.

### 8.2.1 Image Decomposition Using MCA

In the view of MCA, an original image  $X$  is the overlap of  $M$  different morphology layers  $\{X_i\}_{i=1}^M$ , as shown in Eq. (8.1).

$$X = \{X_i\}_{i=1}^M = X_1 + X_2 + \dots + X_M \quad (8.1)$$

$$X_t = T_t \alpha_t, \quad X_s = T_s \alpha_s \quad (8.2)$$

In the SR application, two morphology layers  $\{X_t, X_s\}$  should be separated from  $X$ .  $X_t$  denotes the texture and  $X_s$  is for the needed cartoon. Correspondingly,  $T_t$  is the dictionary for  $X_t$  and  $T_s$  is for  $X_s$ . Then, the layers should be optimally represented by sparse representation  $\{\alpha_t, \alpha_s\}$  in Eq. (8.2), where,  $X_t, X_s \in R^N, T_t, T_s \in R^{N \times L} (L \gg N)$ . To realize sparse representation, the noise-constraint solution via the basis pursuit (BP) is often used in Eq. (8.3) [12]. Here,  $\varepsilon$  stands for the noise level as the prior information in the decomposed image.

$$\{\alpha_t^{opt}, \alpha_s^{opt}\} = \text{Arg} \min_{\{\alpha_t, \alpha_s\}} \|\alpha_t\|_1 + \min \|\alpha_s\|_1 \quad s.t. \|X - T_t \alpha_t - T_s \alpha_s\| \leq \varepsilon \quad (8.3)$$

$$X_s = X - T_t \alpha_t^{opt} \quad (8.4)$$

In the dictionary learning step for SR, the approach in Eq. (8.4) is adopted to further leave all the noise in the cartoon component after MCA image decomposition by Eq. (8.3). The solution is robust to generate the texture part with little noise from the HR training image in the super-resolution.

$$\{\alpha_t^{opt}, \alpha_s^{opt}\} = \text{Arg} \min_{\{\alpha_t, \alpha_s\}} \|\alpha_t\|_1 + \min \|\alpha_s\|_1 + \lambda \|X - T_t \alpha_t - T_s \alpha_s\|_2^2 + \gamma TV\{T_s \alpha_s\} \quad (8.5)$$

However, Eq. (8.3) is not suitable to extract smooth cartoon in SR reconstruction step. So Eq. (8.5) is adopted in this step [12]. In this solution, the noise-constraint condition is replaced by the unconstrained penalized item  $\lambda \|X - T_t \alpha_t - T_s \alpha_s\|_2^2$ . Besides, a total variation (TV) penalty is adopted to obtain the cartoon structure with pronounced edge. Here,  $TV\{T_s \alpha_s\}$  is the  $l_1$ -norm of the gradient. Figure 8.1



Fig. 8.1 Image decomposition by MCA. Left original image, Middle cartoon, Right texture

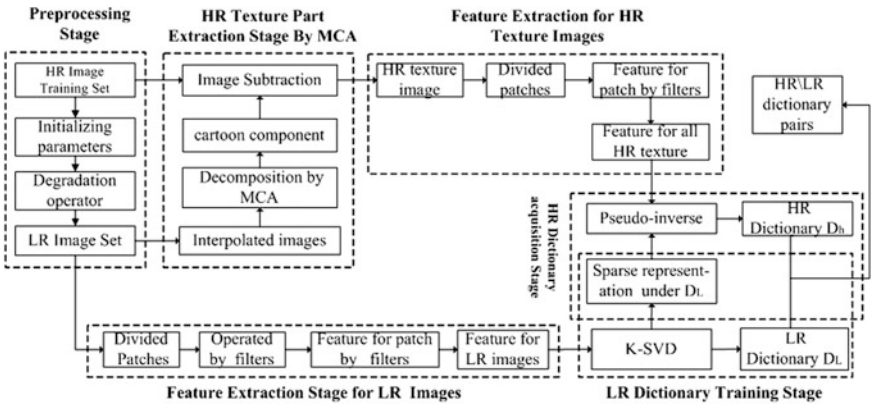


Fig. 8.2 The flowchart for dictionary learning step

shows image decomposition result by MCA with the TV penalty and the original image can be found in <http://pan.baidu.com/s/1c0Ix6Zu>. The cartoon component is completed and piecewise smooth. Thus, it can help the SR reconstruction step recover the cartoon part of the expected HR image efficiently.

### 8.2.2 The Training of the HR/LR Dictionary Pairs

In the dictionary training, MCA and dictionary learning method via sparse representation are combined to describe the feature of HR/LR images. The flowchart for this step is shown in Fig. 8.2.

In the feature extraction for LR dictionary, each LR image is interpolated and divided into patches with  $N \times N$  pixels and Eq. (8.6) are used as the derivative feature of image patches [5, 7]. Thus, the feature vector obtained for each LR patch

by concatenating the four vectors has a length of  $4N^2$ . To reduce the complexity, the principal components analysis (PCA) is employed to reduce feature space. The reduced feature  $y_i$  can reflect the corresponding low-resolution patch. Eventually, LR set can be described as shown in Eq. (8.7).

K-SVD [13] is used for training LR dictionary. The algorithm aims to iteratively improve the initial dictionary and achieve optimal sparse representations  $A$  for the feature set  $Y_l$  in Eq. (8.8) [13]. Here,  $\alpha_i$  is the  $i$ th column vector of sparse matrix  $A$  and denotes the sparse code for  $y_i$ ,  $T_0$  is the target sparsity constraint. Eventually, the optimal LR dictionary and the homologous sparse representations can be obtained by the algorithm.

In the HR texture extraction, Bicubic interpolation [2] are employed to magnify LR images into the HR image's size. The MCA with the noise-constraint in Eq. (8.3) are applied to separate cartoon component from the interpolated image. Thus, the HR texture is extracted from the HR image by image subtraction.

$$\begin{aligned} f_1 &= [-1, 0, 1], & f_2 &= f_1^T \\ f_3 &= [1, 0, -2, 0, 1], & f_4 &= f_3^T \end{aligned} \quad (8.6)$$

$$Y_l = \{y_i\}_{i=1}^n = \{y_1, y_2, \dots, y_n\} \quad (8.7)$$

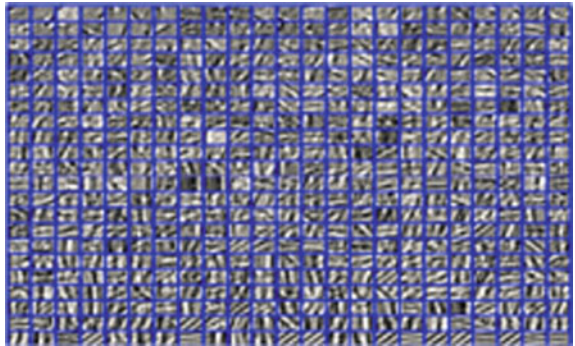
$$\min_{D_L, A} \{\|Y_l - D_L A\|_F^2\}, \quad s.t. \forall i \quad \|\alpha_i\|_0 \leq T_0 \quad (8.8)$$

$$X_h = \{x_i\}_{i=1}^n = \{x_1, x_2, \dots, x_n\} \quad (8.9)$$

In the HR feature extraction, each texture image is divided into patches and the feature is extracted by the same method in the feature extraction for LR images. Eventually, the HR texture image set can be described in Eq. (8.9).

The HR dictionary is acquired by assuming that the patches can be represented by the same sparse matrix  $A$  under the HR/LR dictionary pairs. The HR dictionary

**Fig. 8.3** The visual version of some HR dictionary's atoms from our experiment



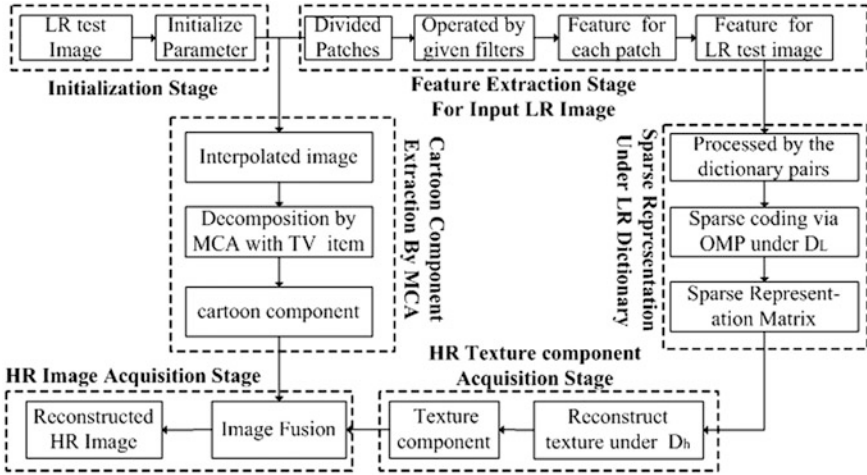


Fig. 8.4 The flowchart for SR reconstruction step

in Eq. (8.10) can be considered to solve the pseudo-inverse problem in Eq. (8.11). Figure 8.3 shows some HR dictionary's atoms obtained by the approach.

$$D_h = \arg \min \|X_h - D_h A\|_F^2 \quad (8.10)$$

$$D_h = X_h A^+ = X_h A^T (A A^T)^{-1} \quad (8.11)$$

### 8.2.3 The Reconstruction for Super-Resolution Image

In this section, an approach based on MCA and sparse representation is presented to obtain HR image using inputted LR image. The flowchart for the SR image reconstruction step is shown in Fig. 8.4.

In the step, the magnified LR image is decomposed into the cartoon and texture component by the MCA method with TV penalty item in Eq. (8.5). The decomposed cartoon component is retained as the HR cartoon part of the expected HR image. The same methods from the dictionary learning step are applied to extract the feature vector  $z_i$  of each LR patch. The input LR image can be described as shown in Eq. (8.12).

$$Z_{test} = \{z_i\}_{i=1}^{n_1} = \{z_1, z_2, \dots, z_{n_1}\} \quad (8.12)$$

In the sparse representation, the OMP algorithm is applied in sparse coding for the LR feature by iteratively approximating the solution of the problem in

Eq. (8.13) [13]. Thus, the sparse representation matrix for the input LR image can be obtained under the LR dictionary.

$$\gamma_i = \arg \min_{\gamma_i} \|z_i - D_L \gamma_i\|_2^2 \quad s.t. \quad \|\gamma_i\|_0 \leq T_0 \quad (8.13)$$

To estimate HR texture, each texture patch  $p_i$  can be calculated using the relevant sparse code and the HR dictionary in Eq. (8.14). Thus, the texture of the expected HR can be obtained using all the texture image patches.

$$p_i = D_h \gamma_i \quad (8.14)$$

Finally, the cartoon part from MCA image decomposition and the texture part from the sparse representation under the dictionary pairs are fused into the expected HR image.

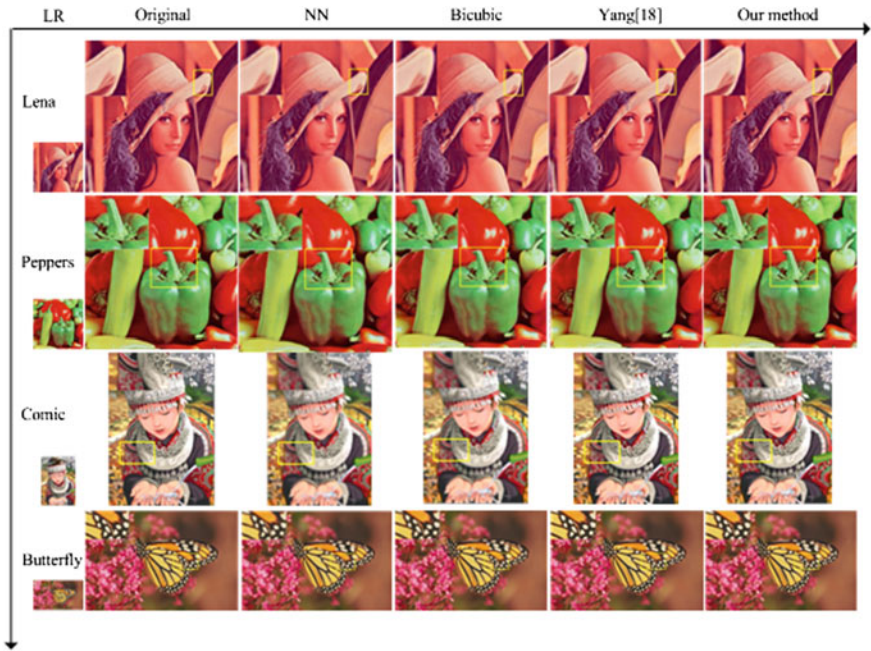
### 8.3 Experiment and Result Analysis

60 training images ( $t1.bmp \sim t60.bmp$  in the folder ‘CVPR08-SR/Data/Training’) are download from <http://www.ifp.illinois.edu/~jyang29/ScSR.htm> in the experiment. Each training image is divided into patches as the same way in [7] to train the HR/LR dictionary pairs. Four LR test images (*Lena*, *Peppers*, *Comic* and *Butterfly*) are used for SR reconstruction, which are the down-sampled images from <http://pan.baidu.com/s/1c0Ix6Zu>. For the color images, only the luminance channel is processed by various SR methods and the others are processed using Bicubic interpolation [2].

#### 8.3.1 The Quality Evaluation for SR Reconstruction

In this section, some experiments are implemented for image super-resolution by different SR methods, including NN [2], Bicubic interpolation [2], Yang’s in [8] and our approach. For each method, the magnification is always set as 3 and three evaluation indicators are employed to evaluate the results: PSNR, SSIM [14] and RMSE. The smaller RMSE and the bigger PSNR or SSIM mean the better image quality. The reconstructed HR images are shown in Fig. 8.5.

As shown in Fig. 8.5, the edges are usually jagged and some speckle rags stay in the SR results from NN and Bicubic interpolation. Yang’s result is a little smooth and the local texture isn’t enough prominent. In the result by our approach, the quality of the edges and texture is improved, such as the hat edges in *Lena*, the complex fold structures in the drawn rectangular region or the gangly pepper in *Peppers*. Besides, our approach shows the least freckles and jagged blocks among



**Fig. 8.5** The SR results using different methods. From *left to right* in each row: LR image, original HR image, the reconstructed results by NN, Bicubic, Yang’s and our approach

**Table 8.1** The evaluation values for the reconstruction results in different approaches

LR Image	Evaluation	NN	Bicubic	Yang [8]	Our
Lena (170*170 pixel)	PSNR	29.4889	31.6688	32.6294	32.9133
	SSIM	0.9430	0.9528	0.9563	0.9656
	EMSE	8.5525	6.6543	5.9576	5.7660
Peppers (170*170 pixel)	PSNR	29.8185	32.3616	33.3092	33.9546
	SSIM	0.9589	0.9681	0.9638	0.9770
	EMSE	8.2341	6.1441	5.5091	5.1146
Comic (83*120 pixel)	PSNR	22.0139	23.1121	23.9000	23.9522
	SSIM	0.6452	0.6987	0.7555	0.7603
	RMSE	20.2230	17.8211	16.2757	16.1781
Butterfly (256*170 pixel)	PSNR	27.0356	29.4222	30.7031	31.0197
	SSIM	0.9599	0.9700	0.9762	0.9803
	RMSE	11.3438	8.6185	7.4368	7.1706



**Table 8.2** PSNR for reconstructed images with additive Gaussian noise

Standard deviation	NN	Bicubic	Yang [8]	Our
$\sigma = 1$	22.0019	23.1025	23.8752	23.9276
$\sigma = 2$	21.9696	23.0708	23.8201	23.8538
$\sigma = 3$	21.9185	23.0290	23.6907	23.7302
$\sigma = 4$	21.8447	22.9644	23.5461	23.5848
$\sigma = 5$	21.7559	22.8870	23.3651	23.3751
$\sigma = 6$	21.6438	22.7923	23.1353	23.1587
$\sigma = 7$	21.5235	22.6853	22.8707	22.8872

these methods, such as the texture fold, the curved boundaries of the garment and the detailed decoration around the neck in *Comic*, and the marking of the butterfly's wings in *Butterfly*.

To avoid the visual error, PSNR, RMSE and SSIM are compared between result and original HR image in Table 8.1. PSNR in our approach are bigger and RMSE are smaller obviously. It certifies that the recovered HR images by our approach are more approximated to the original HR image. Besides, the bigger SSIM means that our results are more similar to the original HR image in terms of image structures [14].

### 8.3.2 The Reconstruction for Images with Additive Noises

In this section, several experiments are conducted to demonstrate that the proposed approach is more robust to the image with additive noises. The *Comic* image is selected in this section, because more complex textures and hierarchical structures are included in this image.

In these experiments, the LR *Comic* image with different additive noises are magnified into HR image with 249\*360 pixels from 83\*120 pixels. In each experiment, various Gaussian white noises with constant mean 0 and different standard deviations  $\sigma$  are added into the inputted original test image. The PSNR are shown in Table 8.2. In Table 8.2, the PSNR from our approach is biggest obviously. It proves that the proposed method is more robust to noise than these conventional methods.

## 8.4 Conclusions

In this paper, a novel approach is proposed for SR image based on MCA and dictionary learning. The MCA theory is used to decompose images well both in the dictionary learning step and HR image reconstruction step. The LR dictionary is

trained by the KSVD from the extracted feature of LR image patches. The HR dictionary is calculated directly via the sparse representation. It's helpful to reduce the complexity of the dictionary learning step. In the reconstruction step, the cartoon part can be obtained well by MCA with TV penalty item. The dictionary pairs make a contribution to recovering the texture part for the expected HR image. A series of experiments and results verify the validity and robustness of our approach.

**Acknowledgement** We'd like to thank all the researchers in the references for their related work and meaningful comments on image super-resolution. Further more, the work in this paper is funded by Chongqing University Postgraduates' Innovation Project, Project Number: CYS15026.

## References

1. Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. *IEEE Sig Process Mag* 20(3):21–36
2. Hadhoud MM, Abd EI-samie FE, EI-Khamy SE (2004) New trends in high resolution image processing. In: *The Workshop on Photonics and Its Application*, pp 2–23
3. Li YR, Dai DQ, Shen LX (2010) Multiframe super-resolution reconstruction using sparse directional regularization. *IEEE Trans Circ Syst Video Technol* 20(7):945–956
4. Freeman WT, Jones TR, Pasztor EC (2002) Example-based super-resolution. *Comput Graph Appl* 22(2):56–65
5. Chang H, Yeung DY, Xiong YM (2004) Super-resolution through neighbor embedding. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1:1275–1282
6. Jiji CV, Chaudhuri S (2006) Single-frame image super-resolution through contourlet learning. *EURASIP J Appl Sig Process*
7. Yang JC, Wright J, Huang T, Ma Y (2008) Image super-resolution as sparse representation of raw image patches. In: *Computer Vision and Pattern Recognition, CVPR 2008*
8. Yang JC, Wright J, Huang TS, Ma Y (2010) Image super-resolution via sparse representation. *IEEE Trans Image Process* 19(11):2861–2873
9. Yin HT, Li ST, Hu JW (2011) Single image super-resolution via texture constrained sparse representation. In: *Proceedings of International Conference on Image Processing, ICIP*, pp 1161–1164
10. Zheng ZH, Wang B, Sun K (2011) Single remote sensing image super-resolution and denoising via sparse representation. In: *2011 International Workshop on Multi-Platform/ Multi- Sensor Remote Sensing and Mapping*
11. Hu WG, Hu TB, Wu T, Zhang B, Liu QX (2011) Sea-surface image super-resolution based on sparse representation. In: *Proceedings of 2011 International Conference on Image Analysis and Signal Processing*, p 102–107
12. Elad M, Starck JL, Querre P, Donoho DL (2005) Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Appl Comput Harmonic Anal* 19(3):340–358
13. Aharon M, Elad M, Bruckstein A (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process* 54(11):4311–4322
14. Zhou W, Alan CB, Hamid RS, Eero PS (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612