.   .

Single Image Super-Resolution Based on the Feature Sign Method

LI Xiao-feng, ZENG Lei, XU Jin, and MA Shi-qi

(School of Communication and Information Engineering, University of Electronic Science and Technology of China   Chengdu   611731)

Abstract   Recently, the super-resolution methods based on sparse representation has became a research hotpot in signal processing. How to calculate the sparse coefficients fast and accurately is the key of sparse representation algorithm. In this paper, we propose a feature sign method to compute the sparse coefficients in the search step. Inspired by the compressed sensing theory, two dictionaries are jointly learnt to conduct super-resolution in this method. The feature sign algorithm changes the non-convex problem to a convex one by guessing the sign of the sparse coefficient at each iteration. It improves the accuracy of the obtained sparse coefficients and speeds the algorithm. Simulation results show that the proposed scheme outperforms the interpolation methods and classic sparse representation algorithms in both subjective inspects and quantitative evaluations.

Key words   feature sign method;   image reconstruction;   image resolution;   sparse representation

(611731)

;   ;   ;

TP301.6   A   doi:10.3969/j.issn.1001-0548.2015.01.003

Image super-resolution (SR) is a technique aiming at the estimation of a high-resolution (HR) image from one or several low-resolution (LR) observation images, which is widely used in remote sensing, biometrics identification, medical imaging, video surveillance, etc. There are three categories of super resolution methods. Interpolation methods are simple but short of high-frequency component, while reconstructed methods degrade rapidly when the desired magnification factor is large.

Recently, the third category called learning based methods are developed, which become the most active research area. In this method, a HR image can be predicted by learning the co-occurrence relationship between a set of LR example patches and corresponding HR patches. Standard methods are example based method[1], neighbor embedding[2], etc. Motivated by compressive sensing theories[3], Ref.[4] proposed an algorithm which used sparse representation of the input LR image to reconstruct the corresponding HR image with two jointly trained dictionaries. The method used the inherent data

message, mapped the data in some dictionaries, and reflected the data information by some sparse coefficients, which turned out to reconstruct the HR image greatly.

There are many classic sparse representation methods. Greedy methods, like orthogonal matching pursuit (OMP[5]), attack the problem heuristically by fitting the sparse models using greedy stepwise least squares. But they are often computationally slow. The least angel regression (LARS[6]) algorithm adds a new element to the active set of nonzero coefficients by taking a step along a direction having equal angles with the vectors in the active set at each iteration. But when the iteration number is large, the method is not efficient. Preconditioned conjugate gradients (PCG[7]) uses the internal Newton's method to minimize the logarithmic barrier function and receives good results. But it is time-consuming.

In this paper, the feature sign method for single image SR via sparse representation is proposed which can overcome above drawbacks. By determining the sign of the sparse coefficients, the non-differentiable problem is changed to an unconstrained quadratic optimization problem (QP) which can be solved easily. It is competent to capture the sparse coefficient of data in an over-completed dictionary efficiently and rapidly. Simulation results demonstrate that it obtains better reconstruction performance qualitatively and quantitatively over other classic SR methods.

1 Single Image SR Using Sparse Representation

The super-resolution image reconstruction problem is known to be an ill-posed inverse problem. Various approaches have been proposed to regularize it, while they all share two parts of constraints priors:

$$U(x) = \|y - Hx\|_2^2 + \lambda E(x) \qquad (1)$$

where H is reconstruction error constraint which represents data-fidelity, and λ is sparsity prior constraint which represents the desired sparse properties of the HR images. For the latter one, most explicit regularization techniques enforce ℓ_0-, ℓ_1- or ℓ_2-minimization properties of the HR image X, thus resulting in many different sparse algorithms.

In this paper, the ℓ_1-minimization is taken to calculate the sparse coefficient, which has been proved by Ref.[8] as long as the sparse representation coefficients of the ith patch α_i are sufficiently sparse, and the NP-hard problem can be efficiently solved by minimizing the ℓ_1-norm:

$$\min \|\alpha_i\|_1 \quad \text{s.t.} \begin{cases} F\|Z_l\alpha_i - y_i\|_2^2 & \varepsilon_1 \\ F\|x_i - Z_h\alpha_i\|_2^2 & \varepsilon_2 \\ \|RZ_h\alpha_i - v\| & \varepsilon_3 \end{cases} \qquad (2)$$

where F is an identity matrix or a linear feature extraction operator. It can be some kind of high-pass filter, for example, the first-order and second-order derivatives. Matrix R extracts the region of overlap between current target patch and previous reconstructed HR image, and v contains the values of the previous reconstructed HR image on the overlap. Formula (2) ensures that the obtained coefficient α_i can best extract sparse information of Y in Z_l, and so can be used to reconstruct X with Z_h.

The SR algorithm is a patch-wise based sparse recovery with the joint learning dictionaries. The input LR image Y is interpolated to the size of desired HR image at first, and then divided into a set of overlapping patches of size $p \times p$. For each LR image patch y_i, the feature y is extracted in the training phase, and sparse representation α_i is computed with respect to low-resolution dictionary Z_l. α_i is then used to predict the underlying HR image patch x_i (feature x) with respect to high-resolution Z_h. The predicted HR patches are tiled together to reconstruct the HR image X.

In this work, the pair of over-complete dictionary $\{Z_h, Z_l\}$ are jointly trained[9] by using the K-SVD[10] method:

$$\min_{Z_h, Z_l, \{\alpha_i\}_{i=1}^N} \sum_{i=1}^N (\|x_i - Z_h\alpha_i\|_2^2 + \|y_i - Z_l\alpha_i\|_2^2) + \lambda\|\alpha_i\|_1$$
$$\text{s.t.} \quad \|Z_x(:,k)\|_2 \quad 1, \|Z_y(:,k)\|_2 \quad 1 \qquad (3)$$

2 Feature Sign Method

How to solve formula (2) fast and efficiently is the fundamental of sparse representation and quality determination of reconstructed HR image. It is a typical ℓ_1-regularized least squares problem which

can be solved by many sparse algorithms.

To keep it simple, it can be rewritten in a vector form as:

$$\min_{\alpha} \sum_{i=1}^{m} \|y_i - \boldsymbol{Z}\alpha_i\|_2^2 + \lambda \sum_{i=1}^{m} \|\alpha_i\|_1 \qquad (4)$$

Let $\alpha_i^{(j)}$ be the jth coefficient of α_i, then the optimization problem is equal to:

$$\min_{\alpha_i} p(\alpha_i) = \|y_i - \boldsymbol{Z}\alpha_i\|_2^2 + \lambda \sum_{j=1}^{k} \left|\alpha_i^{(j)}\right| \qquad (5)$$

If the sign of $\alpha_i^{(j)}$ at each iteration is known, then the non-differentiable problem can be changed to an unconstrained quadratic optimization problem (QP) which can be solved easily.

The sign can be settled by 4 steps:

1) Define $p(\alpha_i) = \|y_i - \boldsymbol{Z}\alpha_i\|_2^2 + \lambda \,|\, \alpha_i^{(j)} |$ and set $f(\alpha_i) = \|y_i - \boldsymbol{Z}\alpha_i\|_2^2$, $g(\alpha_i) = \lambda \,|\, \alpha_i^{(j)} |$. Define $\nabla_i^{(j)} \,|\, \alpha_i \,|$ as the sub-differentiable value of the jth coefficient of α_i.

2) If $|\alpha_i^{(j)}| > 0$, $\nabla_i^{(j)} \,|\, \alpha_i \,|$ is given by the sign of $\alpha_i^{(j)}$. If $\alpha_i^{(j)} = 0$, then the sub-differentiable value $\nabla_i^{(j)} \,|\, \alpha_i \,|$ is given from the set of $[-1,1]$.

3) So the optimality conditions for achieving the optimal value of $p(\alpha_i)$ are translated to

$$\begin{cases} \nabla_i^{(j)} f(\alpha_i) + \lambda \text{sign}(\alpha_i^{(j)}) = 0 & \text{if } \left|\alpha_i^{(j)}\right| > 0 \\ \left|\nabla_i^{(j)} f(\alpha_i)\right| \quad \lambda & \text{if } \alpha_i^{(j)} = 0 \end{cases} \qquad (6)$$

4) Consider how to select the optimal sub-gradient $\nabla_i^{(j)} g(\alpha_i)$ when the optimality conditions are violated. Consider the case where $|\nabla_i^{(j)} f(\alpha_i)| > \lambda$ and $\alpha_i^{(j)} = 0$. Suppose $\nabla_i^{(j)} f(\alpha_i) > \lambda$, which means $\nabla_i^{(j)} p(\alpha_i) > 0$. In order to decrease $p(\alpha_i)$, $\alpha_i^{(j)}$ must be decreased. Since $\alpha_i^{(j)}$ is now at the null point, the very first decrease to $\alpha_i^{(j)}$ is changing it to a negative number. So set $\text{sign}(\alpha_i^{(j)}) = -1$. Similarly, if $|\nabla_i^{(j)} f(\alpha_i)| < -\lambda$, then set $\text{sign}(\alpha_i^{(j)}) = 1$.

3 Compute Sparse Coefficients

The algorithmic procedure of learning coefficients using feature sign is described in the following 5 steps.

An active set $\boldsymbol{H} \square \{j \mid \alpha_i^{(j)} = 0, |\nabla_i^{(j)} f(\alpha_i)| > \lambda\}$ is maintained for potentially nonzero coefficients and their corresponding signs $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_k]$, the image

point $y = [y_1, y_2, \cdots, y_n]$, the dictionary is Z, and sparse parameter is λ.

1) Initialize: $\alpha_i = \boldsymbol{0}, \boldsymbol{\beta} = \boldsymbol{0}, \boldsymbol{H} = \phi$, $\beta_j \in \{-1, 0, 1\}$ denotes $\text{sign}(\alpha_i^{(j)})$;

2) Activate: From zero coefficient of α_i, select $j = \arg\max_j |\nabla_i^{(j)} f(\alpha_i)|$. Activate $\alpha_i^{(j)}$ (add j to the active set $\boldsymbol{H}$) only if it locally improves the objective (6);

3) Feature sign: Let $\hat{\boldsymbol{Z}}$ be a sub-matrix of $\boldsymbol{Z}$ that contains only the columns corresponding to the active set of α_i and f_i, and $\hat{\boldsymbol{\beta}}$ be $\boldsymbol{\beta}$ corresponding to the active set. Compute the solution to the resulting unconstrained QP problem:

$$\min p(\hat{\alpha}_i) = \| y_i - \hat{\boldsymbol{Z}}\hat{\alpha}_i \|_2^2 + \lambda \hat{\boldsymbol{\beta}}^{\text{T}} \hat{\alpha}_i \qquad (7)$$

Let $(\partial p(\hat{\alpha}_i)/\partial \hat{\alpha}_i) = 0$, the optimal value of α_i under the current active set is

$$\hat{\alpha}_i^{\text{new}} = (\hat{\boldsymbol{Z}}^{\text{T}} \hat{\boldsymbol{Z}})^{-1} (\hat{\boldsymbol{Z}}^{\text{T}} y_i - \lambda \hat{\boldsymbol{\beta}}) \qquad (8)$$

Perform a discrete line search on the closed line segment from $\hat{\alpha}_i$ to $\hat{\alpha}_i^{\text{new}}$: Check the objective value at $\hat{\alpha}_i^{\text{new}}$ and all points where any coefficient changes sign, and update $\hat{\alpha}_i$ (and the entries in α_i) to the point with the lowest objective value. Remove zero coefficients of $\hat{\alpha}_i$ from the active set and update $\boldsymbol{\beta} = \text{sign}(\alpha_i)$;

4) Check the optimality conditions: Check the optimality conditions based on formula (6);

5) Output: The optimal coefficient matrix is $\alpha^* = [\alpha_1^*, \alpha_2^*, \cdots, \alpha_m^*]$.

4 Experiment Results

In the following section, simulation results are given to illustrate the performance of our scheme. The simulation and comparison are carried out by Matlab implementations. All experiments are executed on a 2.33 GHz Intel Core 2 Quad CPU Q8200 processor with 2 GB memory in Windows XP OS.

This paper samples 20 000 HR and LR patch pairs from the training images to learn the over-complete dictionary randomly. λ is 0.15 and the dictionary size is 1 024 in all experiments, which is proved to the best suitable number to balance between computation complexity and image quality. The input LR images are magnified by a factor of 3. The input patches are 5×5 pixels with an overlap of 3 pixels. For color

images, the SR algorithm is only applied on the *Y* (intensity) channel, and the Cb and Cr chromatic channels are only interpolated by Bicubic. The three channels are then combined to form our SR images. The results of various methods are evaluated both visually and qualitatively in root-mean-square error (RMSE) and SSIM (structural SIMilarity[11]).

4.1   Experiment results on image SR

In this part, some experiments are conducted to compare the proposed method and other sparse representation methods. Contrast algorithms are classic interpolation method bicubic, and sparse representation methods OMP[5], LARS[6], PCG [7].

a. inputting low image

b. bicubic   c. OMP

d. PCG   e. LARS

f. feature sign   g. original pictur

Fig.1   Reconstructed HR images (scaling factor 3) of Lena and the magnified brim of the hat by different methods.

Bicubic reconstructed images are quite fuzzy for lacking of high frequencies while OMP, LARS, and PCG images are short of image details. The edge blur and sawtooth effect are much more obvious in bicubic, OMP, LARS, and PCG, while feature sign pictures are more distinct and clearer. Specifically, as Fig.1 shows,

the line of the hat in Lena picture is more fluent and distinct than others in the reconstructed picture. And from the white part of the orange magnified in Fig.2, the proposed method has the clearest picture. Comparing the water ripple in yacht picture, bicubic picture is quite fuzzy, LARS, OMP, PCG pictures have blocking artifact, and the feature sign picture is more discerning.
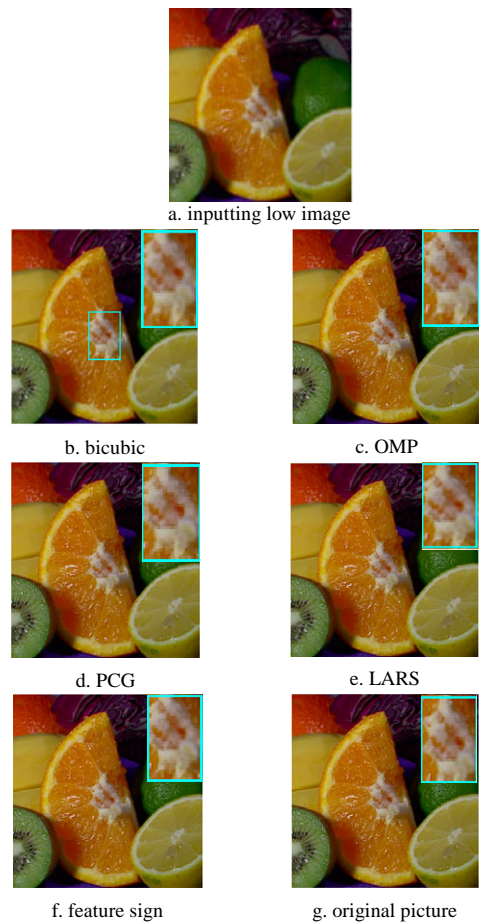

a. inputting low image

b. bicubic   c. OMP

d. PCG   e. LARS

f. feature sign   g. original picture

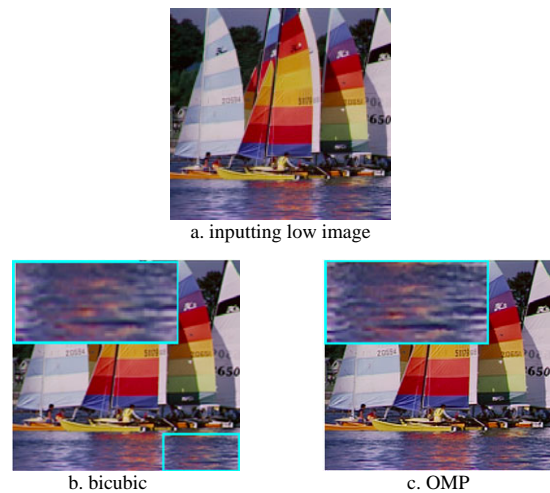Fig.2   Reconstructed HR images (scaling factor 3) of Fruit and the magnified core part by different methods.

a. inputting low image

b. bicubic   c. OMP

d. PCG   e. LARS
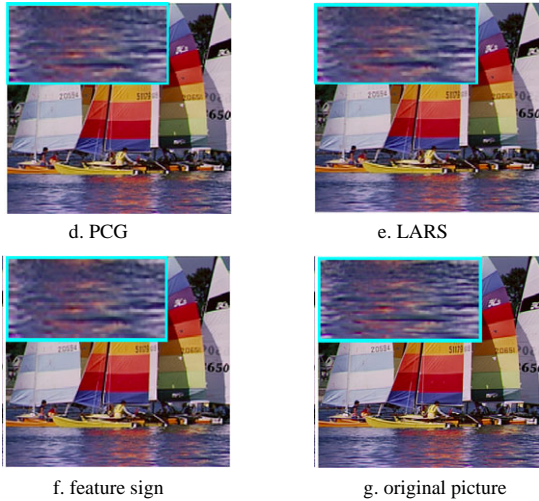
f. feature sign   g. original picture

Fig.3　Reconstructed HR images (scaling factor 3) of Yacht and the magnified water ripple by different methods.

Table 1 compares the RMSEs of the reconstructed images generated by different methods with different input images. The results show that the proposed algorithm achieves the lowest RMSE against bicubic, LARS, OMP, and PCG. Illustrated by the case of Lena, the RMSE reductions of feature sign over bicubic, LARS, OMP, and PCG are 1.472 1, 0.494 6, 0.480 3, and 0.391 9 respectively.

Table 1　RMSEs of different methods

methods	Lena	fruit	yacht
bicubic	6.593 9	5.183 0	10.254 0
OMP	5.616 4	4.546 2	9.316 3
LARS	5.602 1	4.554 8	9.278 3
PCG	5.513 7	4.505 9	8.960 2
feature sign	5.121 8	4.504 8	8.907 5

The reconstruction time of the algorithms is shown in table 2. It can be concluded that the feature sign method yields the best performance with almost the least time. The time of proposed method is nearly one half of LARS and one percent of PCG.

Table 2　Time of each method   s

methods	Lena	fruit	yacht
OMP	188.71	743.86	175.88
LARS	413.32	1 754.62	365.85
PCG	18 177.53	71 401.25	17 037.49
feature sign	206.33	824.56	192.82

As table 3 presents, the feature sign method can yield the biggest SSIM number, which indicates that the feature sign method can best restore the image.

Furthermore, using the feature sign method can improve the efficiency of sparse representation in super resolution.

Table 3　SSIM of the reconstructed HR images

methods	Lena	fruit	yacht
bicubic	0.963 0	0.945 0	0.933 2
OMP	0.971 8	0.952 8	0.943 6
LARS	0.972 2	0.952 9	0.943 9
PCG	0.972 5	0.953 0	0.944 2
feature sign	0.972 6	0.953 2	0.944 6

4.2　Experiment results on a 100-image SR

To test the efficiency of the proposed method more comprehensively, 100 SR experiments are performed by using different algorithms. The 100 test images are downloaded from the international open source image library. The average of RMSE and SSIM are computed in Table 4.

Table 4　General RMSE and time of 100 pictures

average	RMSE	time	SSIM
bicubic	6.658 1	/	0.934 3
OMP	6.092 7	21.791	0.943 6
LARS	5.951 8	47.139	0.955 4
PCG	5.767 4	1 830.051	0.952 6
feature sign	5.754 9	38.663	0.968 1

From table 4, the feature sign method yields the lowest RMSE with much less time, also the SSIM value is the biggest, It demonstrates the good quality of the proposed method used in image SR.

5　Conclusion

This paper proposes an efficient sparse representation method called feature sign for single image super-resolution. This method guesses the sign of sparse coefficients, then changes the complicated ℓ_1-norm question to a QP question. Simulation results demonstrate the advantage of the proposed scheme over existing schemes. Output images from the bicubic have edge blur, OMP pictures have badly jagged artifacts, those from LARS have some blocky effect, and PCG is time wasting, while feature sign reconstructions are distinct and have better visual performance in details. Reconstructed RMSE and SSIM all illustrate the good quality of the proposed method over other methods.

References

[1] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[C]//Proceedings of the Seventh IEEE International Conference on Computer Vision. Los Alamitos, CA, USA: IEEE Comput Soc. 1999.

[2] CHANG H, YEUNG D Y, XIONG Y. Super-resolution through neighbor embedding[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA: IEEE Comput Soc 2004: 275-282.

[3] CANDES, EMMANUEL J. Compressive sampling[C]//25th International Congress of Mathematicians, ICM 2006. United States: Asociacion International Congress of Mathematicians, 2006: 1433-1452.

[4] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing. 2010, 19(11): 2861-2873.

[5] DAVIS G, MALLAT S, AVELLANEDA M. Adaptive greedy approximation[J]. Constructive Approximation, 1997,13(1): 57-98.

[6] EFRON B, HASTIE T, JOHNSTONE I, et al. Least angle regression[J]. Annals of Statistics, 2004, 32(2): 407-499.

[7] KOH K, KIM S J, BOYD S. An interior-point method for large scale ℓ_1-regularized logistic regression[J]. Journal of Machine learning research, 2007,1(4): 1519-1555.

[8] DONOHO D L. For most large underdetermined systems of linear equations, the minimal ℓ_1-norm solution is also the sparsest solution[J]. Communication on Pure and Applied Mathematics, 2006: 59(7), 907-934.

[9] ZENG L, LI X F, XU J. An improved joint dictionary training method for single image super resolution[J]. COMPEL-The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 2013, 32(2): 721-727.

[10] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing over-complete dictionaries for sparse representation[J]. IEEE Transactions on Image Processing. 2006, 54(11): 4311-4322.

[11] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.