

Image Super-Resolution Via Analysis Sparse Prior

Qiang Ning, Kan Chen, Li Yi, Chuchu Fan, Yao Lu, and Jiangtao Wen, *Fellow, IEEE*

Abstract—In this letter, we present a new algorithm for a single image super-resolution using the analysis sparse prior in the $l\alpha\beta$ color space. Experimental results show that our algorithm outperforms other existing state-of-the-art methods. In addition, due to the high scalability of our algorithm, key modules of the proposed algorithm can be integrated with other super resolution algorithms.

Index Terms—Analysis sparsity, image super-resolution, neighbor compatibility, variance prediction.

I. INTRODUCTION

THE GOAL of a single image super-resolution (SR) is to reconstruct a high-resolution (HR) image from a low-resolution (LR) image. Algorithms for image SR can be categorized into following families: Interpolation based [1], Reconstruction based [2], [3], Self-based [4], [5] and Learning based methods [6]–[8]. Interpolation based methods are the simplest and fastest. Reconstruction based methods apply constraints to the HR images based on priori such as categorization. Self-based methods utilizes redundancy across scales. Learning based methods make use of information from a library of images.

Based on the assumption that an image can be restored as a linear combination of atoms from an over-complete dictionary, sparse representation has been proposed for image SR [7], [8]. The sparsity model used in these papers is a generative model, *synthesis sparsity*. Recently, a different model, *analysis sparsity*, has been proposed in [9].

According to [9], the synthesis model is useful for finding where a signal can lie, whereas the analysis model can be used

to find where a signal cannot lie. When an over-complete dictionary with high coherence is used, it becomes difficult to determine where a signal may lie under the synthesis sparsity. On the contrary, finding where a signal can not be becomes easier. The algorithm proposed in this letter was motivated by this intuition.

In the remainder of the letter, we will first introduce our analysis sparsity based super-resolution model in Section II. Section III explains our algorithm in detail, while experiments and discussions are given in Sections IV and V.

II. PROBLEM FORMULATION

A. Analysis Reconstruction With Synthesis-Based Solvers

Under the traditional synthesis sparsity model, a signal $x \in \mathbb{R}^d$ is said to be k -sparse if it can be represented as a linear combination of atoms from $D \in \mathbb{R}^{d \times K}$.

$$x = D\gamma_S, \quad \text{s.t.} \quad \|\gamma_S\|_0 = k. \quad (1)$$

On the other hand, a signal x is said to be l -cosparse if it follows the analysis sparsity model and produces a sparse output when analyzed with an operator $\Omega \in \mathbb{R}^{K \times d}$.

$$\gamma_A = \Omega x, \quad \text{s.t.} \quad \|\gamma_A\|_0 = K - l. \quad (2)$$

Most recently, inherent connections between the two models has been revealed [9]. The equivalence of analysis reconstruction with an augmented synthesis problem was studied in [10]. As [10] puts, if $y = Mx$ has been observed, where $M \in \mathbb{R}^{m \times d}$ is the acquisition matrix, the solution of (3) is identical to the solution of an augmented synthesis problem (4):

$$\hat{x} = \arg \min_x \|\Omega x\|_0 \quad \text{with } y = Mx, \quad (3)$$

$$\hat{x} = D \times \arg \min_{\gamma} \|\gamma\|_0 \quad \text{with } \tilde{y} = \tilde{A}\gamma, \quad (4)$$

where $D = \Omega^\dagger$ is the pseudo-inverse of Ω , P_D is any projector on the nullspace of D , $\tilde{y} = [y, 0]'$, $\tilde{A} = [MD, P_D]'$.

Equations (3) and (4) show that the analysis sparsity model is a least-squares constrained synthesis sparsity model. Due to the robustness of synthesis solvers, when we refer to the analysis sparsity model in the next, we have the corresponding synthesis model in mind. We still use the synthesis dictionary training model proposed in [8] for our experiments.

B. Analysis-Based Super-Resolution

We use X and Y to denote the HR and LR images, and x and y the corresponding patches. Ω_I is the analysis operator for image I and $D_I = \Omega_I^\dagger$. S and U denote a down-sampling operator and a base upscaling method (for example, bicubic), respectively. H is the blurring filter. Δ is defined as the difference between an identity matrix I and USH .

Manuscript received November 19, 2012; accepted January 07, 2013. Date of publication January 23, 2013; date of current version March 07, 2013. This work was supported by the National Science Fund for Distinguished Young Scholars of China (Grant 61125102), the State Key Program of National Natural Science of China (Grant 61133008) and the National Significant Science and Technology Projects of China (Grant 2012ZX01039-001-003-2). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alexandros G. Dimakis.

Q. Ning, K. Chen, and L. Yi are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: qiangning1990@gmail.com; chenkan0007@gmail.com; ylyshh124@gmail.com).

C. Fan is with the Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: chuchufan1990@gmail.com).

Y. Lu was with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. He is now with the Department of Electrical and Computer Engineering, University of California-San Diego, La Jolla, CA 92093 USA (e-mail: luyao@ucsd.edu).

J. Wen is with the Department of Computer Science, Tsinghua University, Beijing 100084, China (e-mail: jtwen@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2013.2242198

Using the operators defined above, we have:

$$Y = SHX + n_0, \quad (5)$$

$$X = UY + \Delta X, \quad (6)$$

where n_0 represents gaussian noise. Note that we ignore the noise term in (6), and use back projection (BP) [11] at the end of our algorithm to denoise the result images.

The inverse problem then becomes a prediction of ΔX , instead of X by making use of the base method followed by deficiency compensation. The problem becomes less ill-posed after we do this replacement. To regularize the inverse process, we solve the following sparse representation problem,

$$\min_{\gamma} \|\gamma\|_0, \quad \text{s.t.} \quad \|\tilde{y} - \tilde{D}\gamma\|_2 < \varepsilon, \quad (7)$$

where $\tilde{y} = [Fy, 0]'$, $\tilde{D} = [FDy, P_{Dy}]'$ and the F operator is a feature extraction operator used to lay constraints on the closeness of the approximation of γ for y .

Similar to synthesis based SR proposed by [8], the HR and LR images are “connected” via a vector γ :

$$\widehat{\Delta x} = D_{\Delta X} \cdot \gamma, \quad (8)$$

and therefore we can recover X by putting all the reconstructed $\widehat{\Delta x}$ into ΔX and calculate (6).

In practice, we use the l_0 norm as opposed to the l_1 norm to convert (9) to a convex optimization problem, which is easier to solve and can still guarantee some sparsity.

$$\min_{\gamma} \|\gamma\|_1, \quad \text{s.t.} \quad \|\tilde{y} - \tilde{D}\gamma\|_2 < \varepsilon. \quad (9)$$

C. Iterative Refinements

From (6), we can see that our model can be generalized as an improvement to other existing methods. In our model, Δ can be regarded as a *deficiency learning* operator, as it represents information that cannot be recovered by U . Although there have been many studies on high frequency detail enhancement since, for example, Freeman [6], the deficiency learning perspective has not received sufficient attention until recently [12].

Since an upscaling method U can be refined using (6), if we denote this scheme by M_U , we can repeat the refinement using (6) iteratively until and if the iterations converge. According to (6), computational complexity of M is:

$$C(M_U) = C(U) + C(\Delta), \quad (10)$$

where $C(\Delta)$ represents the complexity of (7), and is constant for a given ε . Then,

$$C(M_U^n) = C(U) + nC(\Delta) \quad (11)$$

for the n -th iteration ($M_U^0 = U$). It is easy to see that the complexity of this iteration only grows linearly with respect to n .

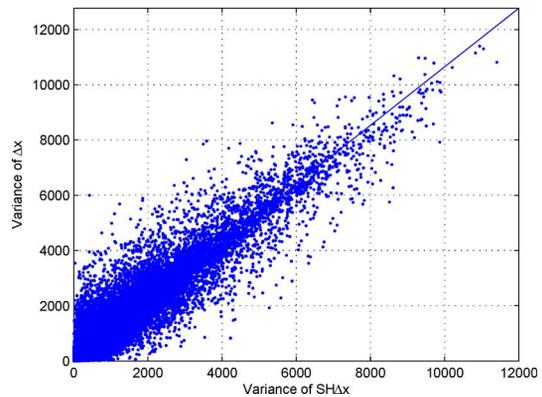


Fig. 1. The variance of Δx is linearly depended on $SH\Delta x$.

III. ALGORITHM DESCRIPTION

A. $l\alpha\beta$ Color Space

To up-scale color images, people usually conduct the SR only in one of the color channels, usually the luminance. The other two channels then undergo simple interpolation-based SR in order to reduce the computational complexity. When correlations exist between different color components, using different algorithms for different components may lead to noticeable degradation to color consistency after up-scaling.

In [13], Ruderman *et al.* developed the $l\alpha\beta$ color space, where the 3 principal components were produced using an orthogonal decorrelation. This allows for applying different operations in the 3 decorrelated channels without undesirable cross-channel incompatibilities. Our experiments in next section show that using the $l\alpha\beta$ space can lead to a significant and robust improvement in PSNR.

Usually, there are three steps when converting an image in the RGB space to the $l\alpha\beta$ space, namely RGB to XYZ, XYZ to LMS [14], and LMS to $l\alpha\beta$ [13]. In our algorithm, we use the following simplified conversion:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0.3475 & 0.8230 & 0.5559 \\ 0.2162 & 0.4316 & -0.6411 \\ 0.1304 & -0.1033 & -0.0268 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (12)$$

As the conversion matrix above is invertible, we can easily convert back to the RGB space.

B. Variance Prediction for Δx

In practice, (8) should turn into:

$$\widehat{\Delta x}_N = D_{\Delta X} \cdot \gamma, \quad (13)$$

The subscript N appears because the scale of Δx has been altered several times during the normalization both in the dictionary training step and reconstruction step. As a result, it is necessary to predict the scale of $\widehat{\Delta x}$ before obtaining the final result.

From (5) and (6) and ignoring noise,

$$SH\Delta X = Y - SHUY. \quad (14)$$



Fig. 2. Results of the “Zebra” image magnified by a factor of 2. Left to right: Glasner interpolation, SAI [1], ScSR [8], the Proposed Algorithm and GroundTruth.

Based on both intuition and experimental results (Fig. 1), we found that for a signal ω , $\text{Var}(SH\omega)$ has a strong linear correlation with $\text{Var}(\omega)$, so that Δx can be predicted as the following:

$$\widehat{\Delta x} = \beta \widehat{\Delta x}_N \frac{\sigma(Y - SHUY)}{\sigma(\widehat{\Delta x}_N)}, \quad (15)$$

where β is the slope, σ represents the standard deviation and $\sigma(\omega) = \sqrt{\text{Var}(\omega)}$.

C. Neighbor Compatibility

In [8], Yang *et al.* enhanced neighbor compatibility by incorporating information from previously reconstructed high-resolution patches to the patch that is currently being processed. This method suffers from several drawbacks. Firstly, the continuous *stitching* means that patches have to be processed one by one, making parallel computing impossible. Secondly, a previously reconstructed patch may introduce a strong enough component that “dwarfs” information obtained from other sources. In our algorithm, similar to [8], we also used four 1-D filters to extract first- and second-order derivatives:

$$f_1 = [-1, 0, 1], f_2 = f'_1, f_3 = [1, 0, -2, 0, 1], f_4 = f'_3. \quad (16)$$

The difference is that we also apply the above filters to lower scales, thereby incorporating multi-scale feature extraction and mitigating deficiencies introduced by single-scale features. In practice, however, too many scales may degrade too much of the importance of the original scale. So we choose two scales in experiments and also apply coefficient weighing the second scale features, so the final form of our feature would be following.

$$Fy = \begin{bmatrix} f_y \\ \mu f_{S_y} \end{bmatrix} \quad (17)$$

where f and F is the single- and multi-scale features extraction operator respectively, and μ is the weight of second-scale features.

Our algorithm can be explained using the pseudo code as following:

Require: Input color image Y , dictionary D_Y and $D_{\Delta X}$ using base method U

Ensure: High resolution reconstruction \hat{X}

- 1: Convert Y from RGB to $l\alpha\beta$ using (12)
- 2: Find the correct scale of ΔX using (14). Denote the result as P_s .
- 3: **for** each 4×4 patch y of Y , taken with overlap in each direction **do**
- 4: Solve the minimization problem in (9) and find the sparse representation γ

TABLE I
AVERAGE RESULTS ON A SET OF 80 RANDOM IMAGES¹

Measure(dB)	Bicubic	Step 1	Step 2
PSNR	39.47	31.52	40.90
Gain		-7.95	9.38
Total Gain		-7.95	1.43
Measure(dB)	Step 3	Step 4	Final
PSNR	41.46	41.71	43.19
PSNR Gain	0.56	0.25	1.48
Total Gain	1.99	2.24	3.72

¹ For fair comparison, PSNR/RMSE are measured in RGB space.

- 5: Extract patch p_s from P_s corresponding to the location of y and calculate its variance
- 6: Referring to (13), (15), recover the rescaled Δx
- 7: Add $\widehat{\Delta x}$ to $\widehat{\Delta X}$, averaging overlapped area
- 8: **end for**
- 9: Find \hat{X} using (6)
- 10: Denoising using BP

IV. EXPERIMENTS

In our study, we first investigated the benefit introduced by each step in our algorithm. In the experiments, the value of β in (15) was set to 1.

In Table I, the average gain provided by every step is presented. Step 1 is the technique of [8] (but without BP denoising), step 2 changes target from X to ΔX , step 3 converts the color space to $l\alpha\beta$ and finally step 4 replaces the synthesis model by the analysis model. Finally BP is applied to denoise the output from the previous steps.

In Table II we compare our results with Bicubic, SAI proposed by Zhang and Wu [1], self-based method Glasner by [4], and Yang *et al.* [8]. Examples are given in Figs. 2 and 3.

Furthermore, in Table III we show the results when the iterative refinement is applied to the Bicubic (B), SAI and M_B algorithms, without the BP denoising (therefore the results in this table were slightly different from Table II because BP was skipped).

V. DISCUSSION

Preliminary results reported in the previous section show that the proposed algorithm significantly outperforms other algorithms.

Due to the complexity involved in soloing the sparse representation, the current implementation of our algorithm is not yet real time. However, many optimizations of the algorithm exist, and, as discussed earlier, the proposed scheme is more parallel-processing friendly than many other state-of-the-art algorithms.

Some limitations exist for the proposed algorithm. Firstly, the performance of scale prediction degrades when BP is used



Fig. 3. Results of the “Lena” image magnified by a factor of 2. Left to right: Glasner interpolation, SAI [1], ScSR [8], the Proposed Algorithm and GroundTruth.

TABLE II
COMPARISONS AMONG CURRENT METHODS (DB)

Image	Bicubic	SAI	Glasner	Yang	Proposed
Brickhouse	36.04	33.84	38.73	37.86	39.16
Child	44.08	39.27	45.26	42.99	48.14
Lena	42.67	38.67	43.06	42.35	46.41
Mandrill	35.08	33.56	37.31	37.05	37.45
Statue	38.77	35.17	40.07	39.68	43.43
Sparrow	41.41	38.84	43.23	42.47	44.95
Zebra	35.23	30.40	35.85	36.66	40.10
Monarch	34.48	31.99	36.78	36.97	37.56
Gain		-3.25 ¹	1.57	1.03	3.68

¹ In our experiments using the SAI package, we found that although it gives visual pleasing results, it behaves poorly in respect of PSNR. Even this, it can still be improved by our scheme according to Table III.

TABLE III
A UNIVERSAL IMPROVEMENT METHOD (DB)

Image	Bicubic/ M_B	SAI/ M_{SAI}	M_B/M_{M_B}
Brickhouse	36.04/37.51	31.20/35.98	37.51/37.78
Child	44.08/46.41	36.26/43.20	46.41/46.70
Lena	42.67/45.02	35.12/42.36	45.02/45.38
Mandrill	35.08/36.29	31.08/34.50	36.29/36.52
Statue	38.77/41.18	31.65/38.41	41.18/41.88
Sparrow	41.41/43.50	35.50/41.31	43.50/43.92
Zebra	35.23/37.63	26.52/34.68	37.63/38.46
Monarch	34.48/36.12	30.32/34.42	36.12/36.53
Gain	1.98	5.90	0.45

within base up-scaling method U . According to (14) and (15), our prediction does not work if $Y - SHUY = 0$, which is what BP achieves. As a result, we remove the BP from U before iterative refinement. This limitation may be eliminated if a new prediction approach without calculating (15) is used.

In addition, theoretically, the proposed iterative refinement will work if the deficiency “image” of U could be sparsely represented. This is the case for linear up-scale operators of U , but may not be true universally.

Finally, it is theoretically unclear if the proposed iterative refinement converges and under what conditions.

VI. CONCLUSION

In this letter, we introduce a novel algorithm for color image super-resolution using iterative refinement and the analysis sparsity model (or enhanced synthesis sparsity model). Preliminary experiments with a relatively large number of images show significant improvement in SR performance as compared with other techniques.

REFERENCES

- [1] X. Zhang and X. Wu, “Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation,” *IEEE Trans. Image Process.*, vol. 17, pp. 887–896, 2008.
- [2] B. Morse and D. Schwartzwald, “Image magnification using level-set reconstruction,” in *IEEE Conf. CVPR*, 2001.
- [3] M. Ben-Ezra, Z. Lin, and B. Wilburn, “Penrose pixels superresolution in the detector layout domain,” in *IEEE Conf. Computer Vision*, 2007.
- [4] C.-Y. Yang, J.-B. Huang, and M.-H. Yang, “Exploiting self-similarities for single frame super-resolution,” in *ACCV*, Queenstown, New Zealand, Nov. 2010.
- [5] D. Glasner, S. Bagon, and M. Irani, “Super-resolution from a single image,” in *IEEE 12th Int. Conf. Computer Vision (ICCV)*, 2009.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor, Example-Based Super-Resolution Mitsubishi Electric Research Laboratories, 2001, Tech. Rep.
- [7] J. Mairal, M. Elad, and G. Sapiro, “Sparse representation for color image restoration,” *IEEE Trans. Image Process.*, vol. 17, no. 1, Apr. 2008.
- [8] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Trans. Image Process.*, vol. 19, pp. 2861–2873, 2010.
- [9] M. Elad, P. Milanfar, and R. Rubinstein, “Analysis versus synthesis in signal priors,” *Inv. Probl.*, vol. 23, pp. 947–968, 2007.
- [10] M. D. P. N. Cleju and M. G. Jafari, “Analysis-based sparse reconstruction with synthesis-based solvers,” in *IEEE ICASSP*, 2012.
- [11] M. Irani and S. Peleg, “Motion analysis for image enhancement: Resolution, occlusion and transparency,” *J. Vis. Commun. Image Represent.*, vol. 4, no. 4, pp. 324–335, 1993.
- [12] R. Zeyde, M. Elad, and M. Protter, “On single image scaleup using sparse-representations,” in *Curves and Surfaces*. Berlin/Heidelberg, Germany: Springer-Verlag, 2011, pp. 711–730.
- [13] D. L. Ruderman, T. W. Cronin, and C.-C. Chiao, “Statistics of cone responses to natural images: Implications for visual coding,” *J. Opt. Soc. Amer.*, vol. 15, no. 8, pp. 2036–2045, 1998.
- [14] N. Moroney, M. D. Fairchild, R. W. Hunt, C. Li, M. Luo, and T. Newman, “The CIECAM02 color appearance model,” in *IST/SID Tenth Color Imaging Conf.*, Nov. 2002.