Autoregressive Image Super-resolution And Image Recovery Using Nonlocal Prior Modeling

Ms. SHALY K NASAR

M.Tech Applied Electronics& Instrumentation
Department.of ECE
Younus college of Engineering& Technology,
Kollam, Kerala
shalyknasar@yahoo.com

Mr. NISHIL A

Asst. Professor,
Department of ECE
Younus college of Engineering& Technology,
Kollam, Kerala
rafi.nishil@gmail.com

Abstract—**Image interpolation and super-resolution are topics of great interest. Here developed a system to perform super resolution tasks by checking noise. A noisy image is processed under NSP (Nonlocal Spectral Prior). The non-noisy image will be superresolved by NARM (Nonlocal Autoregressive Modeling). NARM-based image Super-resolution method can effectively reconstruct the edge structures and suppress the jaggy/ringing artifacts, achieving the best image superresolution. A novel image prior model, namely non-local spectral prior (NSP) model, is then proposed to characterize the singular values of nonlocal similar patches. NSP builds a bridge between spectral analysis and image prior learning. We consequently apply the NSP model to typical image restoration tasks, including denoising, super-resolution and deblurring, and experimental results demonstrated the highly competitive performance of NSP in solving these low-level vision problems. The aim is to improve the resolution of a low resolution (LR) image to obtain a high-resolution (HR) one which is able to preserve the characteristics of natural images.**

Keywords- Image interpolation, nonlocal autoregressive model, nonlocal spectral prior, sparse representation, super-resolution.

I. INTRODUCTION

The aim of super-resolution is to generate higher resolution image from low resolution images. HR image offers high pixel density and there by more details about the original scene. Many image restoration problems such as denoising, super-resolution and de-blurring are inherently ill-posed inverse problems. Therefore, natural image prior models, which describe the `true' statistics of natural image, play an important role in image restoration. the NSP model, is proposed by learning the parameters of GGD from natural images. Gradient-based image prior modeling is based on the fact that natural images usually contain only a small part of edge/texture regions Therefore, multiple NSP models should be learnt and applied adaptively based on image content. In

this paper, we adopt vector quantization techniques to conduct multiple NSP model learning.. The image patches can be well-represented as sparse linear combination of elements from an appropriately choosen over complete dictionary. It takes advantage of redundancy of similar patches in natural images. By embedding NARM into sparse representation model, NARM can act as a kernel, so that it reflects image self-similarity and constraints image local structure by connecting a missing pixel to its nonlocal neighbors by using nonlocal redundancy techniques and assumed that nonlocal similar patches have similar coding coefficients [6]. The variable splitting and Augmented Lagrange Multiplier (ALM) techniques [17], are adopted to effectively solve NARM model. For robust image restoration, it is crucial to model the prior adaptively to image content. Gradient basis statistics are not robust to noise, it is difficult to evaluate robustly the parameter from noisy images. The NSP based super-resolution methods[16] in terms of PSNR, SSIM [8] and FSIM measures, offer better visual quality and suppress Gaussian noises.
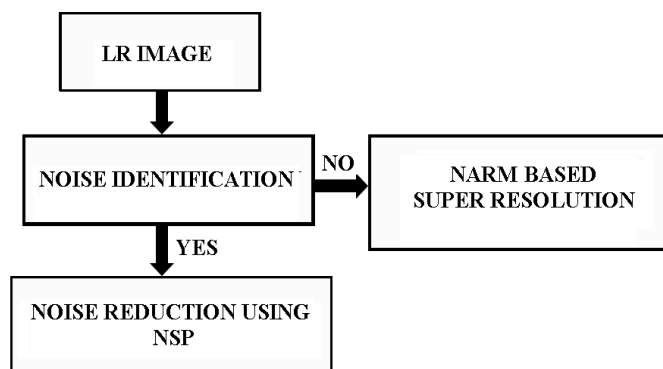

Fig. 1. Block Diagram of the system.

II. NONLOCAL AUTOREGRESSIVE MODELING

Nonlocal Autoregressive Modeling (NARM) refers pixels as a linear combination of its nonlocal neighbouring pixels. It

is a natural extension of traditional autoregressive modellingFor image interpolation, it is assumed that the low resolution (LR) image is directly down-sampled from the original high-resolution (HR) image. Thus, there is a great degree of freedom in recovering the missing pixels. With in the NARM all the pixels can be connected. Each patch has bunch of its nonlocal neighbors and weights assigned to it. To improve the SRM based image interpolation [6], we propose to improve the observation model y = Dx by incorporating the nonlocal self-similarity constraint. Since natural images have high local redundancy, many interpolation methods, including the classical bi-linear and bi-cubic interpolators and the edge guided interpolators, interpolate the missing HR pixel, denoted by xi as the weighted average of its local neighbors. Moreover employing a λ parameter is used for balancing the fidelity as well as the regularization term. The shape of the NARM matrix depends on the image content. Adopting ASDS strategy to learn local dictionaries [3].
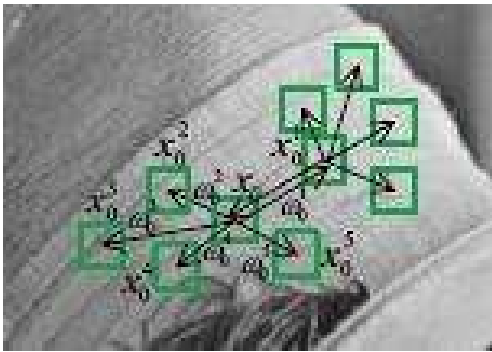

Fig. 2. Illustration of the NARM.

In the autoregressive model (ARM) is used to exploit the image local correlation for interpolation. Nonetheless, the image local redundancy may not be high enough for a faithful image reconstruction, especially in the region of edges, and thus artifacts such as ringings, jags and zippers often appear in the interpolated image. Apart from local redundancy, fortunately, natural images also have a rich amount of nonlocal redundancy The pixel x_i may have many nonlocal neighbors which are similar to it but spatially far from it. Clearly, those nonlocal neighbors (including the local ones) of x_i , denoted by x_i^j , can be used to approximate xi by weighted average:

III. NARM SUPERRESOLUTION

Superresolution algorithm can be adopted to the image,after intialising the image with a bicubic interpolator. Bicubic interpolation is one of the basic resampling techniques and the extension of linear or bilinear interpolation. It can be accomplished using Lagrange polynomials or cubic splines or cubic convolution algorithm. It uses 16 nearest pixel values

which are located in diagonal directions from a given pixel. It gives an exact result of intensity mapping of each pixel.

A. Algorithm

In this paper, we adopt the variable splitting technique [17] to solve the constrained minimization . Here the image is being classified in to k number of cluster patches. Learn a PCA for each cluster. The learned dictionary is being assigned to each cluster patches. Variable splitted patches can be solved using Augmented Lagrange Multiplier technique(ALM). Each patches can be grouped in to two equal regions, which can be constrained by ALM, that is by finding local maxima and local minima. Matrix completion and matrix recovery are also done at this stage. Finally conjugate gradient method is to compute the original HR image. The NARM for image super-resolution can be generally written as

$$\alpha^{\hat{}} = \arg\min_\alpha \{// y - DS\varphi_\alpha //_2^2 + \lambda.R(\alpha)\} s.t. y = D\varphi_\alpha$$
$$(1)$$

Running time of algorithm is shorter for smooth images than non smooth images because the smooth images are excluded from patch clustering.
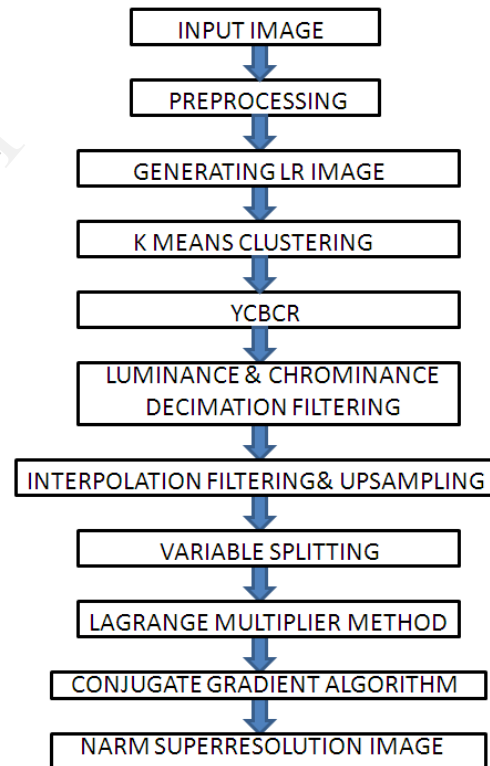

Fig. 3. Flow chart of the NARM super-resolution

B. Lagrange Multiplier Algorithm

The main purpose of this method is to remove blur solution on images caused by uniform linear motion. This method assumes that linear movement corresponds to an integral

number of pixels. Resolution of the restored image remains a very high level. The main contributions of this method are, increasing of ISNR, decreasing of MSE and reduces computational complexity compared to CG. Variable splitting technique divides the k clusters in to two set of constraints. Join the two variable with a new constraint The new linking constrained can be relaxed with a lagrange multiplier. Z is the Lagrangian multiplier, and μ is a positive scalar.

$$x^{(l+1)} = \arg \min L(x, Z^{(l)}, \mu^{(l)}) \qquad (2)$$

$$Z^{(l+1)} = Z^{(l)} + \mu^{(l)}(y - Dx^{(l+1)}) \qquad (3)$$

$$\mu^{(l+1)} = \tau.\mu^{(l)} \qquad (4)$$

Lagrange multipliers basically find local maxima and local minima of a function subject to equality constraint [17]. Employing local PCA dictionaries to span sparse domain for signal representations. Augmented Lagrange Multiplier method is employed for constrained optimization problems. The main motivation is to split the original problem in to several easier sub problems by introducing auxiliary variables.

C. Conjugate Gradient Algorithm

Conjugate gradient algorithm (CG)is a fundamental iterative method for effectively solving large scale systems with very little extra storage and has been applied to the super-resolution problems. The computational cost of Algorithm mainly comes from four sources, i.e., the clustering-based PCA (Principal Component Analysis) sub-dictionary learning in the NARM computing, the patch-based sparse coding and the conjugate gradient minimization. The patch clustering needs $O(u \cdot K \cdot q \cdot n)$ operations, where u is the number of iterations in K -means clustering, q is the total number of patches extracted for clustering, and n is the length of the patch vector. The computation of PCA sub-dictionaries needs $O(K \cdot (m2 \cdot n2 + n3))$ operations, where we assume that each cluster has m patch samples in average. Thus, the PCA sub-dictionary learning needs $O(T (u \cdot K \cdot q \cdot n + K \cdot (m2 \cdot n2 + n3)))$ operations in total, where T denotes the times of PCA sub-dictionary update in the whole algorithm implementation. This autoregressive modeling involves NL times of K nearest neighbor search and Conjugate gradient for minimization for solving Eq. (6), where NL is the number of LR pixels. Shape of S mainly depends on image content. Thus, this process needs $O(NL \cdot (s2 \cdot n + t1 \cdot p2))$ operations in total, where s is the width of searching window, t1 is the number of iterations in Conjugate gradient, and p is the number of nonlocal samples used for NARM modeling. The sparse coding by Eq. (1) needs $O(N \cdot (2 \cdot n2 + p \cdot n))$ operations, where N is the total number of patches extracted for sparse coding. Thus, the cost of Conjugate gradient minimization is $O(J \cdot \kappa \cdot (NL \cdot N \cdot (p+1)+N \cdot n))$,where κ denotes the iterations of the conjugate gradient (CG) algorithm and J is the total number of outer loop iterations in Algorithm. The final super-resolution result is insensitive to the initial regularization parameters λ and η,

$$X = [x_i^1, x_i^2 ......, x_i^J]^T \quad \omega_i = [\omega_i^1, \omega_i^2, ... \omega_i^J]^T \qquad (5)$$

$$\hat{\omega_I} = (X^T X + \gamma I)^{-1} X^T x_i$$

$$(6)$$

Gradient is an n*n 2D vectors. First computing gradient direction, then find the optimum in this direction. Choosing the residual value which would be the negative of the gradient. Then make it orthogonal to previous ones. Finally find the minimum along this direction. In CG method, only one matrix vector multiplication [15] is being done at each iteration. It is implemented as an iterative algorithm, applicable to sparse systems. Conjugate constraint is an orthonormal type constraint [14]. Mainly produces exact solution after finite number of iterations. Moreover it seeks only minima of nonlinear equations.

IV. EXPERIMENTAL RESULTS OF NARM SUPER-RESOLUTION

NARM method significantly outperforms image super-resolution method in terms of both quantitative metrics and subjective visual quality. It suppress jaggy artifacts and reconstruct edge structures. More spatial resolution indicates large number of pixels in the image, thus more details and good colour [23] transition can be achieved.
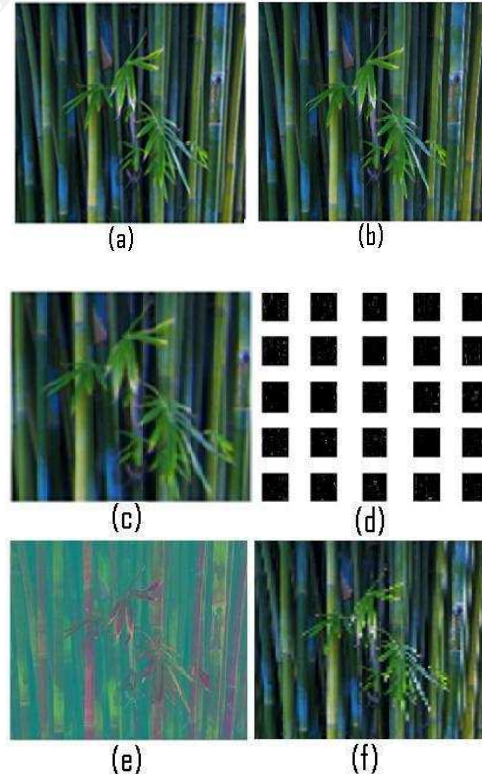

(a)　　(b)　　(c)　　(d)　　(e)　　(f)

Fig. 4. (a) Input image. (b) Preprocessed image. (c) Low resolution image. (d) K means clustering image. (e) Ycbcr image. (f) NARM super-resolution image.

V. NSP FOR IMAGE RECOVERY

Due to the variation of image content, learning a uniform NSP model for all nonlocal matrices is not accurate and robust. Therefore, multiple NSP models should be learnt and then applied adaptively based on image content [19]. In this paper, we adopt vector quantization techniques to conduct multiple NSP model learning. In the training stage, we extract patches from the sample images and partition the n training patches into clusters by a standard Gaussian-mixture clustering model which can be effectively solved by expectation-maximization algorithm. In each cluster, image patches share similar content, and we can assume that their nonlocal spectrums have similar GGD distribution. However, between different clusters, the nonlocal singular values have distinct distributions. Fig.5 has actually illustrated the content-awareness of NSP. After clustering, parameter estimation is conducted for each cluster. Since parameter estimation of GGD has no closed-form solution by methods of moments, numerical methods have to be used to estimate and by learning a support vector regression model from second and forth moments [13]. In this paper, we propose to estimate the GGD parameters by minimizing the KL-divergence. Since the KL-divergence of image patches is non-convex to λ and γ and the dimension of parameter space is only 2, we simply adopt a line searching strategy to learn the parameters of GGD in each cluster by minimizing the KL-divergence between empirical distribution and parametric GGD. We can compute k pairs of parameters λ and γ for the NSP model of each cluster.

$$[\hat{\gamma}, \hat{\lambda}] = \arg\min_{\gamma,\lambda} KL(PE(\sigma)/PD(\sigma))$$

(7)

$$[\hat{\gamma}, \hat{\lambda}] = \arg\min_{\gamma,\lambda} \int_{\sigma} PE(\sigma)\ln(PE(\sigma)/PD(\sigma))d\sigma$$
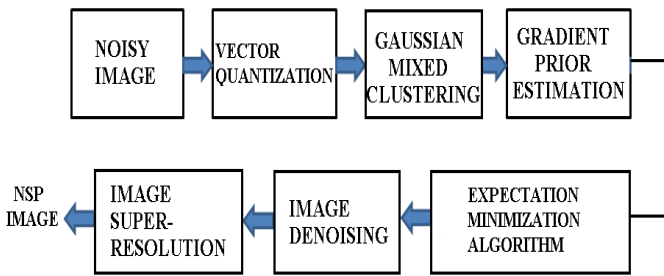
(8)

Fig. 5. Block Diagram of the NSP Modeling.

From Fig. 5, for different clusters, the parameters learnt are significantly different, which validates that NSP models differ with different image contents. For each cluster, due to the similarity of patch content, its learnt NSP model can better

model the nonlocal singular value distribution than using a globally trained model from natural images. In the image restoration stage, we only need to estimate the model parameters for each given patch approximately by clustering. For each nonlocal matrix to the given patch centered at location i, we estimate parameters (_i; i) by weighted sum via vector quantization. Obviously, the models of smooth patches and edge patches are very different.

Fig. 6 illustrates the patch based singular value maps of an image. First, for each local patch (size: 5*5) we collect 49 nonlocal similar patches to it, forming a 25*50 matrix. Then the singular values of the matrix formed by these patches are calculated by SVD. From Fig. 6, we can see that in some local regions the matrix of nonlocal similar patches cannot be considered as low-rank because even the first 10 largest singular values together occupy no more than 70% of the whole energy, which means that in those areas the low-rank approximation cannot well describe the fine structures of natural images. Therefore, it is necessary to explore the probabilistic distribution of singular values of matrices formed by nonlocal similar patches.
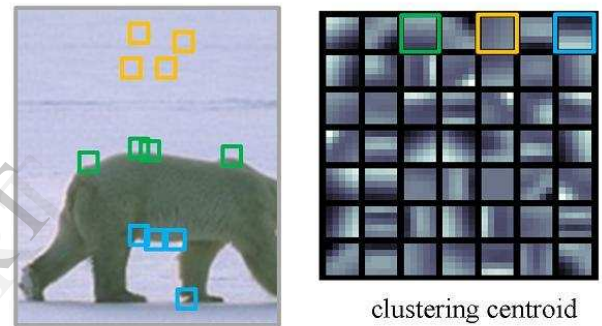

Fig. 6. An example training image, and clustering centroid,

The goal of image clustering is to find a mapping of the archive images in to classes or clusters such that the set of classes provide essentially the same information about the image archive as the image- set collection.

(a) $\lambda= 3.5$, $\gamma=0.59$ (b) $\lambda= 46.6$, $\gamma=1.48$ (c) $\lambda=3$, $\gamma=0.68$
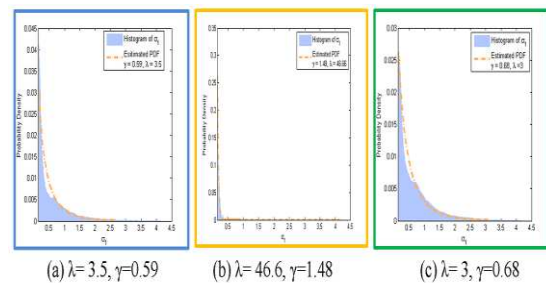
Fig. 7. Empirical distributions and learnt GGD prior models for three clusters labeled by blue, yellow and green, respectively.

The generated classes provide a concise summarization and visualization of image content. In fact, the nonlocal self-

similarity (NSS) has been successfully exploited in image restoration [19]. Despite the wide use of Nonlocal Self-Similarity (NSS), there lacks an in-depth analysis of the low-rank characteristics of nonlocal similar patches. NSS is highly content dependent, spatially variant, and the NSS induced nonlocal singular values are distributed with heavy-tails. In particular, we parameterize the heavy-taileddistribution of nonlocal singular values by generalized Gaussian distribution.

$$P(X_i) = (2\gamma_i \lambda_i^{1/\gamma_i})/(\Gamma(1/\gamma_i)\exp(-\lambda_i // \sigma(X_i) // x_i) \quad (9)$$

where Xi is the ith patch of image x and γi; λi are shape parameters. We choose the GGD for two reasons. First, it is exible to approximate various distributions of NSS induced singular values. Second, its parameters can be well estimated by some optimization approach. With GGD as the prior distribution of nonlocal singular values, proposed NSP can regularize image restoration by measuring its NSS. Empirically, we have found that NSP works effectively and stably across various natural images.

The image degradation process can be modeled as

$$y = (h \otimes k) + n \quad (10)$$

where x is the unknown clean image, h is the downsampling operator, k is the blurring kernel, n is additive Gaussian white noise and y is the degraded observation. Image restoration aims to recover x from the degraded image y, given kernel k and the distribution of random noise n . In case k is unknown, it will be a blind image restoration problem and we could estimate k before estimating x, or estimate them alternately.

A. Optimization for denoising and resolution.

Similar to the patch-based likelihood by using alternating optimization as:

Solve auxiliary variables {X i} by:

$$X_i^{\wedge} = \arg\min_{X_i}\{1/n^2 // P_i x - X_i //_2^2 + \sum_{i=1}^n \lambda_i // \sigma(X_i) // x_i\} \quad (11)$$

Reconstruct x by {X i} and perform gradient descent:
P i is the linear operator to extract nonlocal matrix at location i from image x . When $\gamma > 1$, it is a convex optimization problem which can be e effectively solved by gradient-based optimization techniques in matrix trace function. Motivated by those works, we propose the following iteratively reweighted singular vector thresholding algorithm. Considering the Equ.[11],then closed form of weighted nuclear problem be:

$$D\tau, \omega(Y) = Udiag(\{\max(\sigma\kappa - \tau\omega\kappa, 0)\})V \quad (12)$$

The coefficients of representation are used to generate the HR image. Edge priors are basically used to construct sharp images. In patch based image model,training database is being used. The quality of image depends on the initial set of clusters.

VI. EXPERIMENTAL OUTCOMES OF NSP

The performance of the proposed NSP model for various image restoration tasks, including denoising, deblurring and super-resolution and for prior lerarning [13]. For each cluster, the histogram of its singular values is computed and the parameters λ and γ of the GGD fitting models are estimated. NSP usually removes Gaussian noises.

A. Image Denoising

The denoising methods should not alter the original image u. Now, most denoising methods degrade or remove the fine details and texture of u [18]. Here applicable to remove mainly Gaussian noise [2]. In order to better understand this removal, we shall introduce and analyze the method noise. The method noise is defined as the difference between the original (always slightly noisy) image u and its denoised version. Let u be an image and Dh a denoising operator depending on a filtering parameter h. Then, define the method noise as the image difference (u-Dhu). Natural images also have enough redundancy to be re-stored by Nonlocal means. NL-means algorithm chooses a weighting configuration adapted to the local and non local geometry of the image. So the method noise should be very small when some kind of regularity for the image is assumed. If a denoising method performs well, the method noise must look like a noise even with non noisy images and should contain as little structure as possible. Since even good quality images have some noise, it makes sense to evaluate any denoising method in that way, without the traditional "add noise and then remove it" trick. We shall list formulas permitting to compute and analyze the method noise for several classical local smoothing filters, the Gaussian filtering, anisotropic filtering, total Variation minimization and neighborhood filtering.

The formal analysis of the method noise for the frequency domain filters fall out of the scope of this paper. NL-means algorithm under three well defined criteria: the method noise, the visual quality of the restored image and the mean square error, that is, the Euclidean difference between the restored and true images. The formulas are correlated by visual experiments.

The Generalized Gaussian distribution is continuous probability distributions, which add a shape parameter to the normal distributions. Nonlocal means does not update a pixel,s value with an average those of the pixels around it, instead, it updates it using a weighted average of the pixels, which depends on the distance between its intensity grey level vector and that of the target pixel [14]. Gaussian smoothing model is

also known to be Weierstrass transform. The Gaussian kernel is continuous.

The application of a denoising algorithm should not alter the non-noisy images. The algorithm especially used to solve Gaussian noises. In the image restoration stage, only need to estimate the model parameters for each given patch approximately by clustering. Wide applications from consumer electronics to bio-medical imaging.
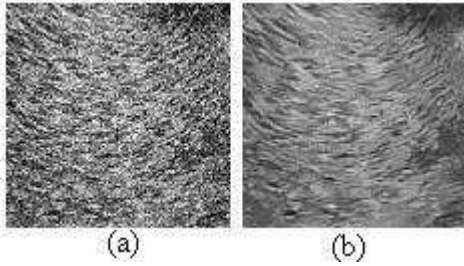

Fig. 8. Denoising results. (a).noisy image, (b).Nonlocal means method.

The denoising performance of the proposed NSP algorithm which contains 24 images of size 512 * 768. Gaussian white noise of 5 different standard deviations (10,15, 20, 25, 50) are added to the original images to simulate the noisy images. In this NSP based denoising algorithm with the trained NSP model, the balancing parameter λ is the only parameter to set. In our experiment, we set λ as 5 for all noise level.. Moreover, with the flexible mixture formulation our model outperforms the fixed low-rank regularization Visual quality is illustrated in Fig.9. Our NSP method preserves texture information well without generating much blurring effect, leading to very pleasing visual quality.
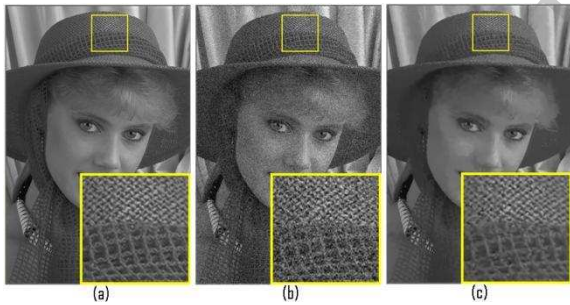

Fig. 9. Denoising results. (a).input image (b).noisy image, (c). proposed NSP method.

B. Image Deblurring

In this section, we first conduct experiments with simulated Gaussian blur kernel and uniform blur kernel, respectively, and then conduct experiments with real motion blur kernels. In the first experiment, we simulated the blur images by Gaussian and uniform kernels, respectively, and then added Gaussian noise to them (noise level: p2 and 2). We choose τ as 6 and gradient descent step length ζ as 0.6 in the

deblurring experiment. Note that the recently developed CSR has shown very strong deblurring capability.

This can see that our method outperforms all the competing algorithms in average. To validate more comprehensively the effectiveness of NSP, conduct de-blurring with real motion blur kernels. Edge priors are basically used to construct sharp images. The dataset adopted is a standard testing dataset for motion deblurring from priors. SVD is used to get the nuclear norm of all nonlocal matrices.
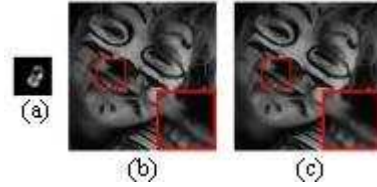

Fig. 10. Deblurring results. (a). blur kernel, (b). original image, (c). proposed NSP

In this dataset, there are 4 images in total with 8 real-world motion blur kernels. For each blurred image y, its corresponding original image x and blur kernel k are provided. Using two state-of-the-art methods for motion deblurring as our competing algorithms: Both the two competing approaches are based on gradient prior model. The algorithm results in very competitive performance although it is not specially designed for motion blur deconvolution.

C. Image Super-resolution in NSP

The image resolution describes the amount of information contained by images. Image super-resolution (SR) reconstruction is essentially an ill-posed problem, so it is important to design an effective prior. The desire for HR image come from the area of improvement in pictorial information. For this purpose, adopted a novel image SR method by learning non-local regularization priors from a given low-resolution image [11]. The goal of super-resolution is to impose different kinds of priors to narrow down the space of high resolution images. The non-local prior takes advantage of the redundancy of similar patches in natural images. For single image super-resolution, the low-resolution (LR) image is obtained by downsampling the blurred high-resolution (HR) image. In our experiments, 8 commonly simulated by first blurring the original HR image with a 7 *7 Gaussian kernel (standard deviation: 1.6) and then downsampling with a scaling factor of 3. In this experiment, two parameters of NSP need to be set: the balancing parameter τ and gradient descent step ξ choose $\tau = 6$ and $\xi = 1:0$ for all the images. In that method has very competitive performances with CSR and significantly outperforms all the other competing methods. The visual quality of several competing algorithms is compared in Fig.11.

NSP method can preserve sharp edges as well as complex texture regions. The content aware plays a crucial role in image restoration. Therefore, inorder for robust image restoration, it is crucial to model the prior adaptively to image content.

We compare our NSP based method with Sparsity based NARM method. So the proposed NSP suppress noises more than NARM method significantly outperforms better visual quality but the computational complexity is less (NSP). Based on this observation, a novel natural image prior model, that is nonlocal spectral prior (NSP) model, is proposed by learning the parameters of Generalized Gaussian Distribution (GGD) from natural images. NSP model can be technically seen as a conditional random fiield. Local smoothness is constrained merely by averaging on overlapped regions.
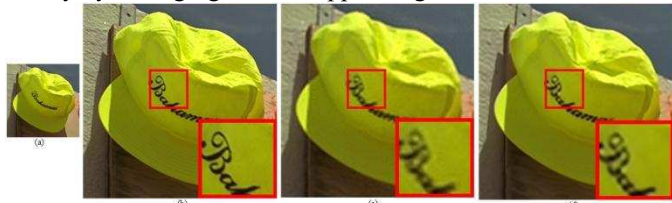

Fig. 11. Super-resolution results. (a). Low resolution image (LR) image, (b). High resolution (HR) image (c). sparsity NARM method, (d). proposed NSP method.

VII. CONCLUSION

We developed an effective system for image super-resolution by Nonlocal Autoregressive Modeling (NARM) and also a denoising and super-resolution algorithm NSP (Nonlocal Spectral Prior). By connecting a missing pixel with its nonlocal neighbors, the NARM can act as a new structural data fidelity term in SRM. NARM method significantly outperforms state-of-the-art image super-resolution methods in terms of both quantitative metrics and subjective visual quality. We showed that NARM can reduce much the coherence between the sampling matrix and the sparse representation dictionary. Furthermore, we exploited we proposed a novel image prior, namely nonlocal spectral prior (NSP), by analyzing the heavy-tailed distribution of singular values of matrices constructed by nonlocal similar patches. The NSP builds a bridge between spectral analysis and image prior learning. Our experiments on image denoising, deblurring and super-resolution demonstrated the effectiveness of the proposed NSP model. Wide applications from consumer electronics to bio-medical imaging. NSP model propose a fast parameter estimation approach which is robust to various kinds of image restoration.

REFERENCES

[1] Freeman, W., Jones, T., Pasztor, E.: Example-based superresolution.Computer Graphics and Applications, *IEEE 22 (*2002) 56-65.

[2] Tappen, M., Russell, B., Freeman, W.: Exploiting the sparse derivative prior for super-resolution and image demosaicing. In: In *IEEE Workshop on Statistical and Computational Theories of Vision.* (2003).

[3] Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE TIP 15* (2006) 3736-3745.

[6] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.,* vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[7] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.,* vol. 43, pp. 1531–1549, Apr. 2010.

[8] Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.,* vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[9] . Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration. In: *ICCV, IEEE* (2011) 479-486 .

[10] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in *image processing," Proc. IEEE,* vol. 98. no. 6, pp. 972–982, Jun. 2010.

[11] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in Proc. *IEEE Int. Conf. Comput. Vision, Sep.* 2009, pp. 349–356.

[12] W. Dong, G. Shi, L. Zhang, and X. Wu, "Super resolution with nonlocal regularized sparse representation," Proc. SPIE *Visual Commun. Image Process.,* vol. 7744, p. 77440H, Jul. 2010.

[13] Sun, J., Xu, Z., Shum, H., Xi'an, P.: Image super-resolution using gradient profile prior. In: CVPR, *IEEE* (2008) 1-8

[14] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.*, vol. 43, pp. 1531–1549, Apr. 2010.

[15] Cai, J., Candes, E., Shen, Z.: A singular value thresholding algorithm for matrix completion. *Siam Journal of Optimization* 20 (2010) 1956-1982.

[16] J. A. Tropp, "Algorithms for simultaneous sparse approximation," *Signal Process.,* vol. 86, no. 3, pp. 572–602, 2006.

[17] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," *Dept. Eelctr. Comput. Eng.,* UIUC, Urbana, Tech. Rep. UILU-ENG-09- 2215, Oct. 2009.

[18] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," Multisc. Model. Simul., vol. 4, no. 4, pp. 1091–1115, 2005.

[19] Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Non-local sparse models for image restoration. In: ICCV, *IEEE* (2009) 2272-2279.

[20] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Trans. Signal Process.,* vol. 58, no. 3, pp. 1553–1564, Mar. 2010.

[21] L. Zhang and X. Wu, "An edge guided image interpolation algorithm via directional filtering and data fusion," IEEE Trans. Image Process., vol. 15, no. 8, pp. 2226–2238, Aug. 2006.

[22] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process., vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[23] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.