IMAGE RESTORATION FROM MULTIPLE COPIES: A GMM BASED METHOD

Sandeep P and Tony Jacob

Department of Electronics and Electrical Engineering,
Indian Institute of Technology Guwahati, Guwahati-781039, India
{p.sandeep,tonyj}@iitg.ernet.in

ABSTRACT

Recovery of original images from degraded and noisy observations is considered an important task in image processing. Recently, a Piece-wise Linear Estimator (PLE) was proposed for image recovery by using Gaussian Mixture Model (GMM) as a prior for image patches. Despite having much lesser computational requirements, this method yields comparable or better results when compared with the widely used sparse representation techniques for image restoration. In many situations, we might have access to multiple degraded copies of the same image, and would like to exploit the correlation among them for better image recovery. In this work, we extend the GMM based method to the multiple observations scenario, where we estimate the original image by utilizing the collective information available from all degraded copies.

Index Terms— Image restoration, multiple observations, Gaussian mixture model, piece-wise linear estimator.

1. INTRODUCTION

Image restoration techniques attempt to solve the problem of recovering an original image $\mathbf{x}$, from a degraded/noisy observation $\mathbf{y}$. In many such problems, image degradation can be considered as the effect of a non-invertible linear operator acting on the original image $\mathbf{x}$ to produce $\mathbf{y}$. Common examples are random masking, subsampling, convolution (blurring) etc. We can write the observed image $\mathbf{y}$ as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where, $\mathbf{A}$ is the degradation operator and $\mathbf{n}$ is the zero mean additive white Gaussian noise. An image restoration technique then attempts to solve a linear inverse problem to recover $\mathbf{x}$ from $\mathbf{y}$.

Recently, sparse representations of natural images over a suitable overcomplete dictionary have been used extensively in different image restoration tasks[1, 2, 3, 4]. Different dictionary learning algorithms [5, 6] can be used to learn an overcomplete dictionary that is better adapted for the sparse representation of natural images. Such learnt dictionaries often give better results in image inverse problems, when compared to standard dictionaries. Given a dictionary, sparse representations are found using non-linear sparse estimation techniques such as l_1 minimization or Matching Pursuit (MP) algorithms. However, poor coherence properties and full degree of freedom in choosing dictionary atoms, can make these algorithms imprecise and unstable. Moreover, both the sparse coding and dictionary learning algorithms are computationally complex, creating difficulties in their practical implementations.

A different approach for solving general image restoration problems assumes Gaussian mixture model (GMM) as a prior for image patches[7, 8]. In [7], patches are estimated by maximizing the likelihood using a GMM prior, where the GMM parameters are learnt apriori from a large collection of natural image patches. In [8], a MAP-EM framework was developed to estimate the patches and to update the Gaussian model parameters in an iterative manner. GMM prior for image patches makes the signal estimation piece-wise linear (PLE), and can be implemented with a set of Wiener filters. Despite requiring considerably less number of computations, this method gives comparable or better results when compared with state of the art techniques. In this work, we study the problem of image restoration from multiple copies. The rest of this paper is arranged as follows. In section 2, we give a brief review of previous works in this area and a description of the GMM MAP-EM algorithm for image restoration from single observation. Section 3 explains the proposed algorithm which extends the original GMM algorithm into multiple observations. Experiment results are discussed in section 4, and section 5 concludes the work.

2. RELATION TO PREVIOUS WORK

In some situations, we may have multiple observations of the same image, where each of the observations is degraded independently. A common example is a denoising problem with multiple noisy observations of the same image. An algorithm for image denoising from multiple observations was proposed in [9] using a combination of averaging and wavelet coefficient thresholding. But, simple averaging techniques does not effectively utilize the correlation of signals among multiple observations. A technique which exploits the correlation among multiple signals was reported in the sparse representation literature [10, 11, 12], and uses simultaneous sparse representations of multiple measurement vectors. This method was applied for signal denoising, but was not used for image recovery from multiple images in general. In [13], a different technique using joint sparsity was proposed for image denoising from multiple observations contaminated with sparse noise. In this work, we extend the GMM based MAP-EM framework proposed in [8], to incorporate multiple observations. This new method utilizes the correlated information from all degraded observations to estimate the original image.

We now briefly describe the MAP-EM algorithm proposed in [8] for image restoration from a single degraded/noisy observation. In this algorithm, an image is split into patches of size $\sqrt{N} \times \sqrt{N}$ and the patches are lexicographically ordered as vectors $\mathbf{x}_i \in \mathbb{R}^N$, $i = 1, 2, ...I$, where I is the total number of patches. For each patch $\mathbf{x}_i$ of the original image, the corresponding degraded noisy patch $\mathbf{y}_i$ can be written using equation (1) as

$$\mathbf{y}_i = \mathbf{A}_i\mathbf{x}_i + \mathbf{n}_i, \quad i = 1, 2, ..I \tag{2}$$

where, $\mathbf{A}_i$ denotes the degradation matrix operating on i^{th} patch and $\mathbf{n}_i$ the zero mean additive white Gaussian noise vector added to i^{th} patch. GMM prior for image patches assumes that each patch $\mathbf{x}_i$ is drawn independently from a Gaussian model indexed by k_i from a mixture of K Gaussians. Gaussian parameters are denoted by $\{(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\}_{1\leq k\leq K}$. Given the degraded noisy observations $\{\mathbf{y}_i\}_{1\leq i\leq I}$, we have to find the estimates of the original image patches $\{\mathbf{x}_i\}_{1\leq i\leq I}$. A MAP-EM algorithm was developed to solve this problem in an iterative manner.

E-step: Assuming that the Gaussian parameters are known from the previous iteration, the E-step computes the MAP estimate $\tilde{\mathbf{x}}_i$ and the best Gaussian model $\tilde{k}_i$ for all patches, $i = 1, 2, ..., I$. Signal estimation and model selection is done by maximizing the log a-posteriori probability $\log p\left(\mathbf{x}_i|\mathbf{y}_i, \tilde{\boldsymbol{\Sigma}}_k\right)$, first over $\mathbf{x}_i$ and then over all Gaussian models. Since we can always perform a mean centering on every Gaussian cluster with respect to the cluster means, we assume that the Gaussian means $\{\boldsymbol{\mu}_k\}_{1\leq k\leq K}$ are all zero. For a fixed Gaussian model k, the MAP estimate is computed as

$$\tilde{\mathbf{x}}_i^k = \tilde{\boldsymbol{\Sigma}}_k\left(\mathbf{A}_i^T\mathbf{A}_i\tilde{\boldsymbol{\Sigma}}_k + \sigma^2\mathbf{I}_d\right)^{-1}\mathbf{A}_i^T\mathbf{y}_i \tag{3}$$

Once the estimates $\tilde{\mathbf{x}}_i^k$ are computed with all Gaussian models, the best model for i^{th} patch is selected as

$$\begin{aligned}\tilde{k}_i &= \arg\min_k\left(\left\|\mathbf{A}_i\tilde{\mathbf{x}}_i^k - \mathbf{y}_i\right\|_2^2 + \sigma^2\left(\tilde{\mathbf{x}}_i^k\right)^T\tilde{\boldsymbol{\Sigma}}_k^{-1}\tilde{\mathbf{x}}_i^k\right.\\ &\quad\left.+\sigma^2\log\left|\tilde{\boldsymbol{\Sigma}}_k\right|\right)\end{aligned} \tag{4}$$

After selecting the model $\tilde{k}_i$, estimate of the i^{th} patch is computed with the selected model as $\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i^{k_i}$.

M-step: Assuming the estimate $\tilde{\mathbf{x}}_i$ and the Gaussian model association $\tilde{k}_i$ are known for all patches $1 \leq i \leq I$, the M-step updates the parameters of all Gaussian models. Parameters of k^{th} Gaussian model is estimated by maximizing the likelihood of observation of all patches associated with this model. The ML estimate results in the empirical estimates of Gaussian parameters,

$$\tilde{\boldsymbol{\mu}}_k = \frac{1}{|\mathcal{C}_k|}\sum_{i\in\mathcal{C}_k}\tilde{\mathbf{x}}_i,\;\; \tilde{\boldsymbol{\Sigma}}_k = \frac{1}{|\mathcal{C}_k|}\sum_{i\in\mathcal{C}_k}\left(\tilde{\mathbf{x}}_i - \tilde{\boldsymbol{\mu}}_k\right)^T\left(\tilde{\mathbf{x}}_i - \tilde{\boldsymbol{\mu}}_k\right) \tag{5}$$

where, $\mathcal{C}_k$ denotes the set of all patches associated with k^{th} Gaussian model.

3. MAP-EM ALGORITHM FOR MULTIPLE IMAGES

In multiple observations, we have L degraded and noisy observations $\mathbf{y}_1, \mathbf{y}_2, ..\mathbf{y}_L$ of an image $\mathbf{x}$ satisfying,

$$\mathbf{y}_j = \mathbf{A}_j\mathbf{x} + \mathbf{n}_j,\;\; j = 1, 2, ...L \tag{6}$$

where $\mathbf{A}_j$ is the degradation matrix operating on j^{th} image and $\mathbf{n}_j$ is the additive zero mean Gaussian noise. Splitting all observed images into patches as in equation (2), we can write

$$\mathbf{y}_{ij} = \mathbf{A}_{ij}\mathbf{x}_i + \mathbf{n}_{ij},\;\; i = 1, 2, ..., I,\;\;\; j = 1, 2, ..., L \tag{7}$$

where $\mathbf{y}_{ij}$ denotes i^{th} patch of the j^{th} observation, $\mathbf{A}_{ij}$ is the corresponding degradation operator and $\mathbf{n}_{ij}$ is the additive white Gaussian noise vector added to i^{th} patch of the j^{th} image. Given all observations $\mathbf{y}_{i1}, \mathbf{y}_{i2}, ...\mathbf{y}_{iL}$ corresponding to a patch i, our problem is to compute the best estimate of the original image patch $\tilde{\mathbf{x}}_i$. We again use the MAP-EM framework as in the single observation case. In the E-step, for every patch, we compute the MAP estimate from all given observations with a fixed Gaussian model k. Then the best Gaussian model is selected in such a way that it describes all observations corresponding to a patch. In the M-step, we update the model parameters of a Gaussian by maximizing the likelihood the estimated patches associated with that Gaussian model.

E-step: Given all observations $\mathbf{y}_{i1}, \mathbf{y}_{i2}, ..., \mathbf{y}_{iL}$ of the i^{th} patch, assuming that the Gaussian parameters are known from the previous iteration, we compute the MAP estimate by maximizing the log a-posteriori probability $\log p\left(\mathbf{x}_i|\mathbf{y}_{i1}, \mathbf{y}_{i2}, ..., \mathbf{y}_{iL}, \tilde{\boldsymbol{\Sigma}}_k\right)$. MAP estimates $\tilde{\mathbf{x}}_i^k$ are computed independently with all Gaussians $k = 1, 2, ...K$ and the best Gaussian model $\tilde{k}_i$ is selected. Estimate of the i^{th} patch is then computed using the selected Gaussian model.

$$\begin{aligned}\left(\tilde{\mathbf{x}}_i, \tilde{k}_i\right) &= \arg\max_{\mathbf{u},k}\left[\log p\left(\mathbf{u}|\mathbf{y}_{i1}, \mathbf{y}_{i2}, ...\mathbf{y}_{iL}, \tilde{\boldsymbol{\Sigma}}_k\right)\right]\\ &= \arg\max_{\mathbf{u},k}\left[\log p\left(\mathbf{y}_{i1}, \mathbf{y}_{i2}, ...\mathbf{y}_{iL}|\mathbf{u}, \tilde{\boldsymbol{\Sigma}}_k\right)\right.\\ &\qquad\left.+\log p\left(\mathbf{u}|\tilde{\boldsymbol{\Sigma}}_k\right)\right]\\ &= \arg\max_{\mathbf{u},k}\left[\sum_{l=1}^{L}\log p\left(\mathbf{y}_{il}|\mathbf{u}, \tilde{\boldsymbol{\Sigma}}_k\right)\right.\\ &\qquad\left.+\log p\left(\mathbf{u}|\tilde{\boldsymbol{\Sigma}}_k\right)\right]\end{aligned} \tag{8}$$

Equation (8) follows from the independence of observations. Using the Gaussian pdf for image patches ($\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_k$), and noise vector ($\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \sigma^2\mathbf{I}_d$), we have,

$$\begin{aligned}\left(\tilde{\mathbf{x}}_i, \tilde{k}_i\right) = \arg\min_{\mathbf{u},k}&\left[\sum_{l=1}^{L}\left(\|\mathbf{y}_{il} - \mathbf{A}_{il}\mathbf{u}\|_2^2\right)\right.\\ &\left.+\sigma^2\mathbf{u}^T\tilde{\boldsymbol{\Sigma}}_k^{-1}\mathbf{u} + \sigma^2\log\left|\tilde{\boldsymbol{\Sigma}}_k\right|\right]\end{aligned} \tag{9}$$

Minimization is first done over $\mathbf{u}$ and then over k.

$$\tilde{\mathbf{x}}_i^k = \arg\min_{\mathbf{u}}\left(\sum_{l=1}^{L}\|\mathbf{A}_{il}\mathbf{u} - \mathbf{y}_{il}\|_2^2 + \sigma^2\mathbf{u}^T\tilde{\boldsymbol{\Sigma}}_k^{-1}\mathbf{u}\right) \tag{10}$$

Differentiating equation (10) and setting the result to zero, we have,

$$\begin{aligned}\tilde{\mathbf{x}}_i^k &= \left(\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{A}_{il} + \sigma^2\tilde{\boldsymbol{\Sigma}}_k^{-1}\right)^{-1}\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{y}_{il}\\ &= \tilde{\boldsymbol{\Sigma}}_k\left(\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{A}_{il}\tilde{\boldsymbol{\Sigma}}_k + \sigma^2\mathbf{I}_d\right)^{-1}\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{y}_{il}\\ &= \tilde{\boldsymbol{\Sigma}}_k\left(\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{A}_{il}\tilde{\boldsymbol{\Sigma}}_k + \sigma^2\mathbf{I}_d\right)^{-1}[\mathbf{M}_i]\,[\mathbf{v}_i]\end{aligned} \tag{11}$$

where, $[\mathbf{M}_i] = \left[\mathbf{A}_{i1}^T\mathbf{A}_{i2}^T...\mathbf{A}_{iL}^T\right]$ and $[\mathbf{v}_i] = \left[\mathbf{y}_{i1}^T\mathbf{y}_{i2}^T\mathbf{y}_{iL}^T\right]^T$. Equation (11) can be written in the form of a linear estimate using Wiener filter as

$$\tilde{\mathbf{x}}_i^k = \mathbf{W}_{k,i}\,[\mathbf{v}_i] \tag{12}$$

where,

$$\mathbf{W}_{k,i} = \tilde{\boldsymbol{\Sigma}}_k\left(\sum_{l=1}^{L}\mathbf{A}_{il}^T\mathbf{A}_{il}\tilde{\boldsymbol{\Sigma}}_k + \sigma^2\mathbf{I}_d\right)^{-1}[\mathbf{M}_i] \tag{13}$$

Once the estimates $\tilde{\mathbf{x}}_i^k$ are computed with all Gaussian models $k = 1, 2, ..., K$, the best Gaussian model for the i^{th} patch is selected using

$$\tilde{k}_i = \arg\min_k \left(\sum_{l=1}^{L} \left\| \mathbf{A}_{il}\tilde{\mathbf{x}}_i^k - \mathbf{y}_{il} \right\|_2^2 + \sigma^2 \left(\tilde{\mathbf{x}}_i^k \right)^T \tilde{\mathbf{\Sigma}}_k^{-1} \tilde{\mathbf{x}}_i^k + \sigma^2 \log \left| \tilde{\mathbf{\Sigma}}_k \right| \right) \quad (14)$$

Estimate of the i^{th} patch is then computed using the selected Gaussian model as,

$$\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i^{\tilde{k}_i} \quad (15)$$

As in [8], the E-step consists of a linear estimate of patches and a non-linear best model selection, together forming a Piecewise Linear Estimation (PLE). Implementation of E-step involves a set of Wiener filters which considerably reduces the computational complexity when compared with the sparse representation based methods.

M-step: In the M-step, we assume that the estimates $\tilde{\mathbf{x}}_i$ and selected Gaussian model $\tilde{k}_i$ are known for all patches $1 \leq i \leq I$. Then the parameters of k^{th} Gaussian are updated by maximizing the likelihood of all the estimated patches associated with the same Gaussian. Let $\mathcal{C}_k$ be the ensemble of estimated patches associated with k^{th} Gaussian. Then the ML estimates of model parameters can be calculated as

$$\left(\tilde{\boldsymbol{\mu}}_k, \tilde{\mathbf{\Sigma}}_k \right) = \arg\max_{\boldsymbol{\mu}_k, \mathbf{\Sigma}_k} \log p \left(\{\tilde{\mathbf{x}}_i\}_{i \in \mathcal{C}_k} | \boldsymbol{\mu}_k, \mathbf{\Sigma}_k \right) \quad (16)$$

With Gaussian model, the ML estimates result in the corresponding empirical estimates,

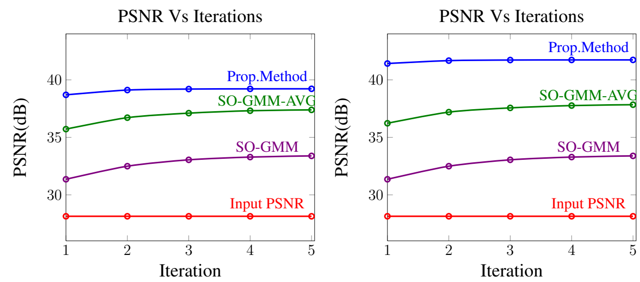
$$\tilde{\boldsymbol{\mu}}_k = \frac{1}{|\mathcal{C}_k|} \sum_{i \in \mathcal{C}_k} \tilde{\mathbf{x}}_i, \; \tilde{\mathbf{\Sigma}}_k = \frac{1}{|\mathcal{C}_k|} \sum_{i \in \mathcal{C}_k} \left(\tilde{\mathbf{x}}_i - \tilde{\boldsymbol{\mu}}_k \right)^T \left(\tilde{\mathbf{x}}_i - \tilde{\boldsymbol{\mu}}_k \right) \quad (17)$$

MAP-EM algorithm is initialized by the same method proposed in [8], where the covariance matrices are initialized to capture the directional regularities such as edges and contours contained in natural images.

4. EXPERIMENT RESULTS

This section describes the experiments conducted to evaluate the performance of the proposed algorithm in some image restoration tasks with multiple observations. We consider two common image restoration applications, namely, image denoising and image inpainting.

In image denoising, AWGN noise is added independently to multiple copies of a test image. The proposed algorithm is then used to recover the original image from these multiple noisy observations. We first compare our results with the results of the algorithm described in section 2, which uses only a single observation for image recovery. We denote this algorithm by Single Observation GMM (SO-GMM). As has been tried in literature with multiple observations, we next use an averaged version of SO-GMM to compare our results. In this method, SO-GMM algorithm is used to recover the original image independently from different copies and in every iteration, results from different observations are averaged to get the recovered image for that iteration. We denote this second method as SO-GMM-AVG. AWGN noise variance σ^2 is set to 100, and the experiments are repeated for two different number of multiple observations, $L = 10$ and $L = 20$. In all our experiments, we have used a 128×128 portion of gray level Lena image.

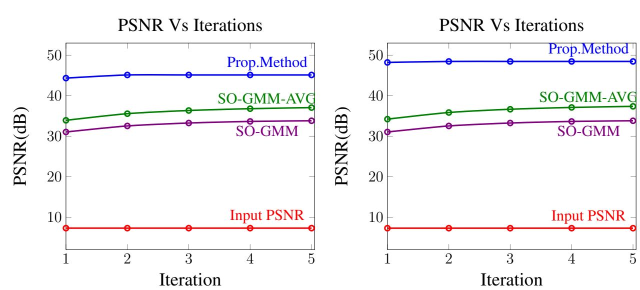


(a) Number of observations $L = 10$ (b) Number of observations $L = 20$

Fig. 1: PSNR Evolution in Image Denoising

Figures 1a and 1b show the evolution of PSNR in successive iterations of the MAP-EM algorithm for $L = 10$ and 20 respectively. From figure 1a, we can see that the proposed method achieves a PSNR of 39 dB from the second iteration onwards which is 6 dB more than the single observation SO-GMM. Due to noise averaging, SO-GMM-AVG gives a better PSNR of 37 dB, with 4 dB improvement compared to the SO-GMM method, but it is still 2 dB less than the proposed method. Figure 1b shows the case with $L = 20$. Since the number of observations are increased from 10 to 20, both the proposed method and SO-GMM-AVG give a significantly better PSNR compared to the SOGMM algorithm. In this case, the proposed method is 8.3 dB better than the SO-GMM method and 4 dB better than SO-GMM-AVG.

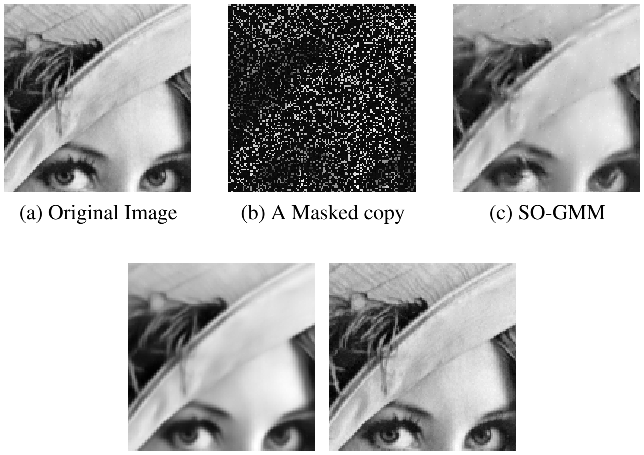
Next we see the performance of our algorithm, in an image inpainting problem with multiple observations. In this case, multiple copies of a test image are masked independently with different random masks. Then the proposed algorithm is used to recover the original image from these multiple observations. Results are compared in the same way as it was done in image denoising. SO-GMM method recovers the original image by using only a single observation among multiple copies, whereas the SO-GMM-AVG method independently performs inpainting on all observations, and the results are averaged in every iteration. Percentage of missing pixels is set to 50% for all images and the experiment is repeated for two different number of observations $L = 10, 20$. Figures 2a and 2b give the PSNR improvement in successive iterations of the inpainting algorithm. When the number of observations $L = 10$, it can be seen from figure 2a, that the proposed method is stable at a PSNR of approximately 45 dB from the first iteration itself. This is 8 dB



(a) Number of observations $L = 10$ (b) Number of observations $L = 20$

Fig. 2: PSNR Evolution in Image Inpainting

higher than the SO-GMM-AVG method and 11 dB higher than the SO-GMM method. We can see that the SO-GMM-AVG method improves only by 3.2 dB when compared to SO-GMM. When the number of observations goes from 10 to 20, as can be verified from figure 2b, the proposed method significantly improved the PSNR to reach 48.46 dB. We note that the performance of SO-GMM-AVG has not changed when the number of observations is increased.



(a) Original Image (b) A Masked copy (c) SO-GMM

(d) SO-GMM-AVG (e) Proposed Method

Fig. 3: Inpainting Results

To see how the proposed method improves the image recovery in the case of severely degraded images, consider the results of an inpainting problem shown in figure 3. Original image is degraded by masking about 80% of pixels. SO-GMM method recovers the image, but most of the fine details like thin edges are blurred. This can be clearly visible on Lena's hat at the top of the image. The recovered image using SO-GMM-AVG appears to be overly smooth when compared to the original image, and blurring of fine details is more compared to the SO-GMM method. But we can see that the proposed method recovers an image which is close to the original image and preserves most of the fine details.

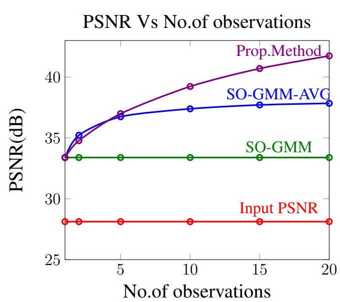


Fig. 4: PSNR vs Number of observations

Figure 4 shows the improvement in PSNR with the increase in number of observations, L. Since SO-GMM method uses only a single observation, its performance is same for all values of L and is shown for a comparison. When $L = 1$, naturally all algorithms have the same performance. As the number of observations increases from 1 to 5, both the proposed method and SO-GMM-AVG shows almost same performance. As L is increased further, performance of the proposed method becomes better, and improves significantly with each observation. In the case of SO-GMM-AVG, improvement in PSNR is almost constant beyond $L = 10$.

5. CONCLUSION

In this work, we have addressed the problem of utilizing multiple degraded observations of an image for better image restoration. Extending the work in [8], we have proposed an algorithm which utilizes the correlated information from all different observations to produce better reconstruction quality. Different experiments conducted to evaluate the performance demonstrate effectiveness of the algorithm in using correlation among multiple observations. As a future work, we consider to extend the proposed algorithm into more general situations as considered in [14]. This includes reconstruction from multiple images which are not perfectly registered due to camera motion and the effect of different exposure time for different images.

6. REFERENCES

[1] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736 –3745, Dec. 2006.

[2] Y. Lou, A. L. Bertozzi, and S. Soatto, "Direct sparse deblurring," *Journal of Mathematical Imaging and Vision*, vol. 39, no. 1, pp. 1–12, Jan. 2011.

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

[4] J. Yang, J. Wright, T.S. Huang, and Y. Ma, "Image superresolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861 –2873, Nov. 2010.

[5] K. Engan, S.O. Aase, and J.H. Husoy, "Method of optimal directions for frame design," in *Proceedings., IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1999, vol. 5, pp. 2443 –2446.

[6] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311 –4322, Nov. 2006.

[7] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *IEEE International Conference on Computer Vision (ICCV), 2011*, Nov. 2011, pp. 479 –486.

[8] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2481 –2499, May 2012.

[9] S.G. Chang, B. Yu, and M. Vetterli, "Wavelet thresholding for multiple noisy image copies," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1631 –1635, Sep 2000.

[10] S.F. Cotter, B.D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477 – 2488, July 2005.

[11] J.A. Tropp, A.C. Gilbert, and M.J. Strauss, "Algorithms for simultaneous sparse approximation: part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.

[12] J.A. Tropp, "Algorithms for simultaneous sparse approximation: part II: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, Mar. 2006.

[13] N. Yu, T. Qiu, and F. Ren, "Denoising for multiple image copies through joint sparse representation," *Journal of Mathematical Imaging and Vision*, pp. 1–9, 2012.

[14] B.K. Gunturk and M. Gevrekci, "High-resolution image reconstruction from multiple differently exposed images," *IEEE Signal Processing Letters*, vol. 13, no. 4, pp. 197–200, April.