Class-specific Image Denoising Using Importance Sampling

Milad Niknejad
Instituto de Telecomunicacoes
Instituto Superior Tecnico
Universidade de Lisboa, Portugal
Email: milad3n@gmail.com

José M. Bioucas-Dias
Instituto de Telecomunicacoes
Instituto Superior Tecnico
Universidade de Lisboa, Portugal

Mário A. T. Figueiredo
Instituto de Telecomunicacoes
Instituto Superior Tecnico
Universidade de Lisboa, Portugal

Although image denoising has sometimes been considered as a solved problem [1], it is still a very active research topic. Although most methods being developed are for generic images, in some applications the image is known to belong to a certain class (*e.g.*, text, face, fingerprints) and this knowledge should be exploited. One way to do so is to use an external method, based on a dataset of clean images from that class, rather than a general purpose set of natural images [2], [3]. Specifically in patch-based methods, the rationale is that the statistical properties of the patches in a class-specific dataset are better adapted to the underlying clean image than if using a generic set of natural images.

Denoting a pair of clean and corresponding noisy patches as $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, the observation model is $\mathbf{y} = \mathbf{x} + \mathbf{v}$, where $\mathbf{v}$ is white Gaussian noise of known variance σ^2. If $P(\mathbf{x})$ is a prior for $\mathbf{x}$, its *minimum mean squared error* (MMSE) estimate is the posterior expectation

$$\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}|\mathbf{y}] = \int_{-\infty}^{+\infty} \mathbf{x}\, P(\mathbf{x}|\mathbf{y})\, d\mathbf{x} = \int_{-\infty}^{+\infty} \mathbf{x}\, \frac{P(\mathbf{y}|\mathbf{x})\, P(\mathbf{x})}{P(\mathbf{y})}\, d\mathbf{x}$$

where $P(\mathbf{y}|\mathbf{x}) \propto e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{x}\|_2^2}$. Much of the work on patch-based denoising can be seen as the search for good priors and/or ways to approximate the posterior expectation $\mathbb{E}[\mathbf{x}|\mathbf{y}]$.

External *non-local means* (NLM) methods estimate the underlying clean patch $\mathbf{x}$ as a weighted average of a set of clean patches in the external dataset $\mathbf{x}_j, j = 1, .., n$,

$$\widehat{\mathbf{x}} = \left(\sum_{j=1}^{n} w_j\right)^{-1} \sum_{j=1}^{n} w_j\, \mathbf{x}_j, \quad \text{with } w_j = e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{x}_j\|_2^2}. \qquad (1)$$

It turns out that (1) can be seen as a non-parametric Monte Carlo approximation of $\mathbb{E}[\mathbf{x}|\mathbf{y}]$ using self-normalizing *importance sampling* (IS, [4]), with $P(\mathbf{x})$ as a proposal density and assuming that $\mathbf{x}_j$, $j = 1, .., n$, are samples thereof (see a related observation in [5]).

At the other extreme of the parametric/non-parametric spectrum, the methods in [6], [7] approximate $P(\mathbf{x})$ as a single Gaussian fitted to a collection of similar patches in the noisy image, from which the MMSE estimate is obtained in closed form. Semi-parametric methods use Gaussian mixtures (GM) to approximate the prior, learned either from external data [8], or from the noisy image itself [9]. The MMSE estimate under a GM prior also has closed form [9]. Other non-Gaussian models (or mixtures of densities other than Gaussians) are more difficult to use, since in general the corresponding MMSE estimate under Gaussian noise does not have a closed form expression.

In this paper, we use IS to perform class-adapted image denoising, by using a more sophisticated and more adaptive prior than a Gaussian (or mixture thereof), for which the MMSE does not have a closed form expression. Our method consists of three main steps:

1) The external dataset of clean image patches is clustered by fitting a non-Gaussian mixture model. The rationale is that it has been shown that the statistics or image patches are better described by leptokurtic (*i.e.*, with heavier tails than a Gaussian) distributions [10]. Specifically, we adopt generalized Gaussian (GG) densities. In the parameter estimation step of the EM-type algorithm, we use the method in [11]. After running this clustering algorithm, the subset of clean patches assigned to the m-th cluster (GG component) is denoted as $\mathbf{X}_m^{GG}$.
2) Each noisy patch is assigned to one of the clusters obtained in the previous step. If the cluster densities of the clean patches were Gaussian (say, with mean $\boldsymbol{\mu}_m$ and covariance $\boldsymbol{\Sigma}_m$, for the m-th cluster), the corresponding likelihoods of noisy patches would also be Gaussian, with the same mean, and covariances $\boldsymbol{\Sigma}_m + \sigma^2\mathbf{I}$ [8], [9]. However, since we are using GG densities, the likelihoods do not have a simple expression, making it impractical to assign the noisy patches to the clusters via *maximum likelihood* (ML). To overcome this difficulty, we fit a Gaussian density to the subset of patches assigned to each GG cluster $\mathbf{X}_m^{GG}$, and classify each noisy patch into one of the clusters via ML using these Gaussian approximations.
3) Finally, after assigning each noisy patch to one of the clusters, the corresponding MSSE estimate is approximated via IS (see (1)), since it does not have a closed form expression under the GG cluster density. One approach would be to generate samples from the assigned GG distribution. Another approach, which we use in this work, is simply to sample clean patches from the assigned cluster $\mathbf{X}_m^{GG}$. In order to speed up the algorithm, for each noisy patch, only 800 patches from each distribution are selected in order to obtain the IS-based MMSE estimate. The IS viewpoint of (1) allows to use the improvements which have been recently proposed for this sampling method. In particular, one of the main shortcomings of using IS is degeneracy of the weights, due to the large dynamic range of the importance weights w_j. In order to alleviate this shortcoming, we use the method recently proposed in [12], which simply applies hard thresholding to the importance weights w_j before computing the sums in (1). The obtained patch are returned to their original position in the image and are averaged in the overlapped pixels.

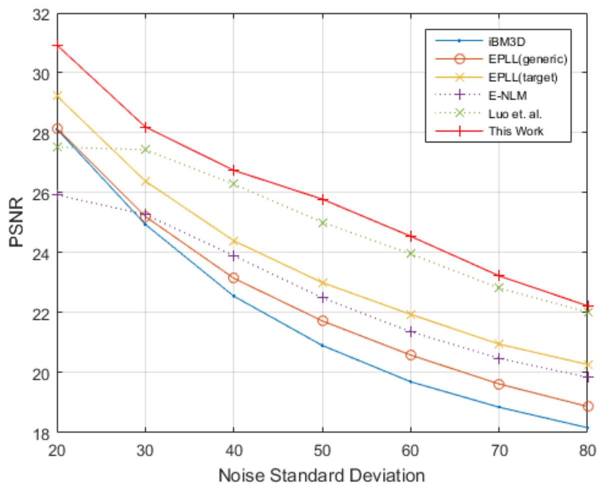
Notice that step 1 only needs to be applied once for a given dataset of class-specific images, while steps 2 and 3 are applied to every image to be denoised.

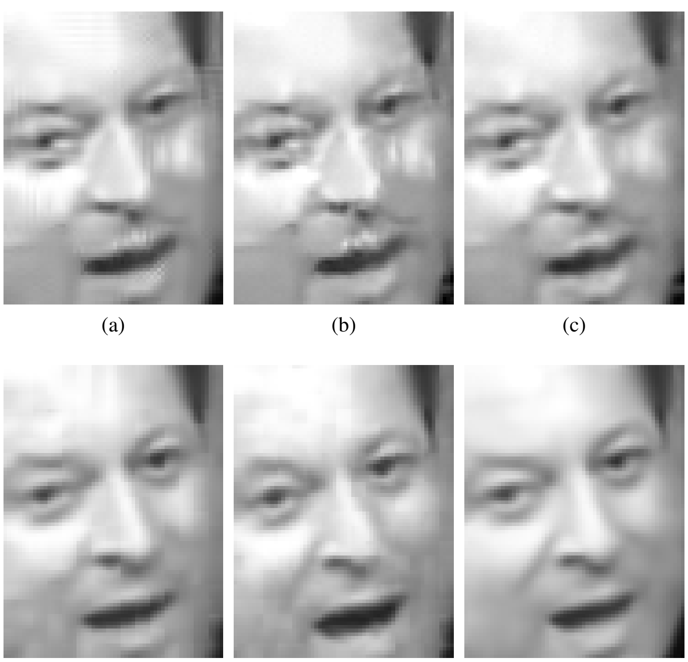
We report preliminary experimental results using the face image database in [13] and the text image database in [2]. For each dataset of images, 5 images were randomly selected as test images, and the other images were selected for training. In the prior learning step, the initialization is obtained by k-means algorithm with 20 clusters. For the multivariate GG density, the exponent was empirically set to 0.9, which was found to lead to good denoising results. In all the experiments the proposed method outperforms other general or class-specific methods for image denoising, as reported in Table I.

TABLE I: Denoising results for face image denoising using the Gore face database in [13]. The results are averaged over 5 test images.

	$\sigma = 20$	$\sigma = 30$	$\sigma = 40$	$\sigma = 50$
BM3D [14]	31.88	29.64	27.57	26.98
EPLL (generic) [8]	31.66	29.43	27.74	26.58
Class-specific EPLL	32.34	30.16	28.49	27.28
External Non-local means	31.81	30.08	28.75	27.48
Luo et. al. [2]	32.98	30.89	29.24	28.01
Ours (Guided multivariate Generalized Gaussian)	**33.09**	**30.99**	**29.48**	**28.08**

Fig. 1: Comparison of different external denoising methods for class-specific text images. For each method, the results arveraged for 5 text images.





(a) (b) (c)

(d) (e) (f)

Fig. 2: An example of denoising of a face image in the Gore database [13] (for $\sigma = 30$): (a) BM3D (PSNR=29.46); (b)EPLL (PSNR=28.97); (c) class-specific EPLL (PSNR=29.91); (d) external non-local means (PSNR=30.97); (e) Luo et. al. [2] (PSNR=32.20); (e) this work (PSNR=33.02).

References

[1] P. Chatterjee and P. Milanfar, "Is denoising dead?" *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 895–911, 2010.

[2] E. Luo, S. H. Chan, and T. Q. Nguyen, "Adaptive image denoising by targeted databases," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2167–2181, 2015.

[3] A. Teodoro, J. Bioucas-Dias, and M. Figueiredo, "Image restoration and reconstruction using variable splitting and class-adapted image priors," in *IEEE International Conference on Image Processing*, 2016.

[4] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer, 1999.

[5] A. Levin and B. Nadler, "Natural image denoising: Optimality and inherent bounds," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 2833–2840.

[6] M. Lebrun, A. Buades, and J. M. Morel, "A nonlocal Bayesian image denoising algorithm," *SIAM Journal on Imaging Sciences*, vol. 6, pp. 1665–1688, 2013.

[7] M. Niknejad, H. Rabbani, and M. Babaie-Zadeh, "Image restoration using Gaussian mixture models with spatially constrained patch clustering," *IEEE Transactions on Image Processing*, vol. 24, pp. 3624–3636, 2015.

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

[9] A. Teodoro, M. Almeida, and M. Figueiredo, "Single-frame image denoising and inpainting using Gaussian mixtures," in *4th International Conference on Pattern Recognition Applications and Methods*, 2015.

[10] H. E. Gerhard, L. Theis, and M. Bethge, "Modeling natural image statistics," *Biologically-inspired Computer VisionFundamentals and Applications. Wiley VCH*, 2015.

[11] F. Pascal, L. Bombrun, J.-Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized gaussian distributions," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5960–5971, 2013.

[12] E. Koblents and J. Míguez, "A population Monte Carlo scheme with transformed weights and its application to stochastic kinetic models," *Statistics and Computing*, vol. 25, no. 2, pp. 407–425, 2015.

[13] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

[14] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.