COLOR IMAGE DENOISING USING CLUSTERING

Dharshini.S, Sivaranjani.S

Abstract— Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Image denoising refers to the recovery of a digital image that has been contaminated by additive white Gaussian noise. In Existing a patch-based Wiener filter that exploits patch redundancy for image denoising. It uses both geometrically and photo metrically similar patches to estimate the different filter parameters. And these parameters can be accurately estimated directly from the input noisy image. So the grayscale denoising method can be applied to denoising the color images through such transformations. However such color–space conversions corrupt the noise characteristics. In this proposed system denoising in the RGB color space is performed using K-Means clustering technique. A locally optimal Wiener-filter-based method and have extended it to take advantage of patch redundancy to improve the denoising performance.

Index Terms— denoising, K-Means clustering, Wiener filter, image processing

.

1 INTRODUCTION

In RECENT years, images and videos have become integral parts of our lives. Applications now range from the casual documentation of events and visual communication to the more serious surveillance and medical fields. This has led to an ever-increasing demand for accurate and visually pleasing images. However, images captured by modern cameras are invariably corrupted by noise. With increasing pixel resolution but more or less the same aperture size, noise suppression has become more relevant. While advances in optics and hardware try to mitigate such undesirable effects, software-based denoising approaches are more popular as they are usually device independent and widely applicable. In the last decade, many such methods have been proposed, leading to considerable improvement in denoising performance. The problem from an estimation theory perspective to quantify the fundamental limits of denoising. The insights gained from that study are applied to develop a theoretically sound denoising method. The challenge of any image denoising algorithm is to suppress noise while producing sharp images without loss of finer details.

The first modern adaptive method to successfully address these contradictory goals can be attributed to Tomasi where the authors proposed a generalization of the SUSAN filter, which itself was an extension of the Yaroslavky filter. The proposed denoising by weighted aver- aging pixels similar in intensity within a local neighborhood. Under strong noise, identifying such similar pixels can be challenging. In, Takeda *et al.* proposed a signal-dependent steering kernel regression (SKR) framework for denoising. This method proved to be much more robust under strong noise. A patch-based generalization of the bilateral filter was proposed in and where the concept of locality was extended to the entire image. Although the results there were encouraging, the true potential for this nonlocal means (NLM) method was only realize. Another patch redundancy-based framework, i.e., BM3D, adopts a hybrid approach of grouping similar patches and performing collaborative filtering in some transform [e.g., discrete cosine transform (DCT)] domain. It ranks among the best performing methods that define the current state of the art. A significantly different approach to denoising was introduced in K-SVD. Building on the notion of image patches being sparse representable, Elad *et al.* proposed a greedy approach for dictionary learning tuned for denoising. In proposed a hybrid approach (K-LLD) that bridged such dictionary-based approaches with the regression-based frameworks. The motivation there was that similar patches shared similar *subdictionaries*, and such sub- dictionaries could be used for better image modeling. A similar observation was exploited in the form of a nonlocal sparse model (NLSM) to improve performance of the K-SVD framework.The dictionary-based methods provide implicit modeling for natural images. More explicit models have also been used for denoising.

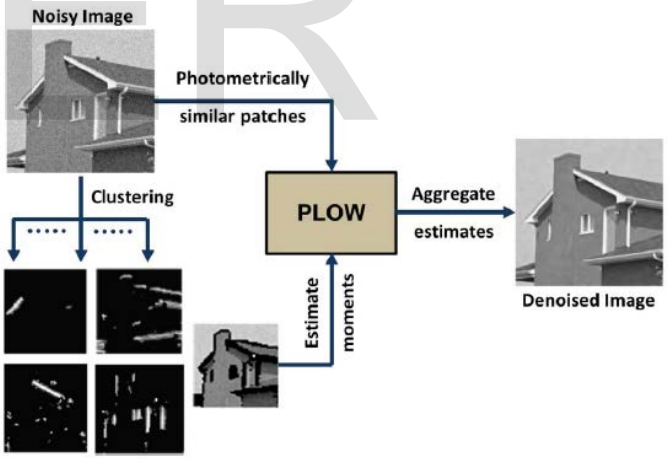

Fig 1: Outline of Proposed System

While most of the aforementioned approaches work in the spatial domain, a vast section of image denoising literature is devoted to transform domain methods. The main motivation in such methods is that, in the transform (e.g., DCT, wavelets, etc.) domain, it is possible to separate image and noise components, and denoising can be performed through the shrinkage of the transform coefficients. In Chang *et al.* showed that, using a spatially adaptive threshold parameter, along with the overcomplete wavelet basis, denoising performance can be considerably improved. Another wavelet-domain method in was considered the model for natural mages in the form of homogeneous Gaussian MRFs was used to improve the performance considerably. In Luisier *et al.* proposed a denoising method aimed at reducing the estimated mean-squared error

(MSE) through wavelet thresholding.

In propose system, a new denoising filter motivated by our statistical analysis of the performance bounds for patch- based methods. The contributions of our paper are as follows: We design a patch-based statistically motivated redundancy exploiting the Wiener filter, where the parameters of the method are learned from *both* geometrically and photo metrically similar patches. As will be clear from our discussions in the next section, our method is formulated to approach the performance bounds for patch-based denoising. As a side note, we also show that the NLM filter is an approximation of the optimal filter (in the MSE sense) obtained if one ignores the geometric structure of image patches. Although extensively used for denoising, the Wiener filter is usually used in conjunction with some transform basis. For example, the collaborative Wiener filter used in BM3D works in the DCT domain where an estimate of the ground truth (signal-to-noise ratio) is obtained through an initial filtering of the image. Our spatial domain method is motivated by our analysis of the image denoising bounds. The framework, graphically illustrated in Fig. 1, develop a locally optimal Wiener filter where the parameters are learned from both geometrically and photometrically similar patches.

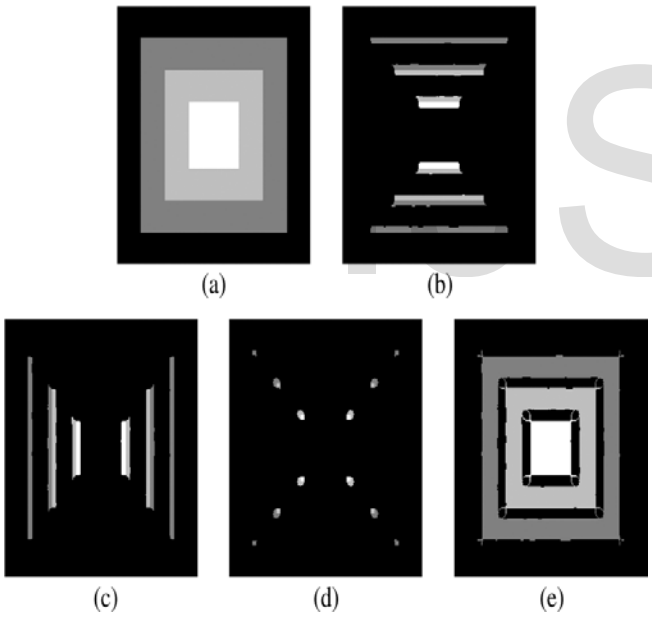

Fig2: Clustering of a simple image based on geometric similarity.

For this, the noisy image is first segmented into regions of similar geometric structure, as shown in Fig. 2. The mean and the covariance of the patches within each cluster are then estimated. Next, for each patch, we identify photometrically similar patches and compute weights based on their similarity to the reference patch. These parameters are then used to perform denoising patchwise. To reduce artifacts, image patches are selected to have some degree of overlap (shared pixels) with their neighbors. A final aggregation step is then used to optimally fuse the multiple estimates for pixels lying on the patch overlaps to form the denoised image.

2. PLOW FILTER

A. Motivation

In the performance bounds for the problem of image denoising. This was done from an estimation theory point of view, where we seek to estimate the pixel intensity z_i, at each location i from its noisy observation, i.e.,

$$y_i = z_i + \eta_i, \quad i = 1, \ldots, M. \tag{1}$$

Here η_i, is assumed to be independent and identically distributed (i.i.d.), and M is the total number of pixels in the image. In our paper, we specifically considered patch-based methods, where the observation model can be posed as

$$\mathbf{y}_i = \mathbf{z}_i + \boldsymbol{\eta}_i \tag{2}$$

$$E\left[\|\mathbf{z}_i - \hat{\mathbf{z}}_i\|^2\right] \geqslant \mathrm{Tr}\left[\left(\mathbf{J}_i + \mathbf{C}_{\mathbf{z}}^{-1}\right)^{-1}\right] \tag{3}$$

This covariance matrix captures the complexity of the patches and is estimated from all the geometrically similar patches present in the given image. Fig. 2 illustrates what we mean by geometric similarity, where it can be seen that each cluster groups together patches containing flat regions, edges in the horizontal or vertical directions, and corners of the simulated box image. Note that such grouping is done irrespective of the actual patch intensities. This is justified for intensity-independent noise when denoising performance is dictated by the complexity of patches, rather than their actual intensities. The values can be then directly estimated from the noisy image as the number of patches (including) that satisfy the aforementioned criterion.

Note that the condition for photometric similarity, as defined here, is stricter than that for geometric similarity. As such, photometric similarity can be expected to imply geometrically similar as well. The bounds expression in (3) thus takes into account the complexity of the image patches present in the image, as well as the redundancy level and the noise variance corrupting the image. In the bound was shown to characterize the performance of the optimal affine-biased denoising method. In particular, for the WGN, the right-hand side of is the performance achieved by the optimal linear minimum mean-squared-error (LMMSE) estimator, with and being the parameters of the estimator. The Wiener filter is, in fact, the LMMSE estimator that achieves this lower bound. Thus, a patch-based Wiener filter, where the parameters are accurately estimated, can lead to near-optimal denoising. This forms the basis of our approach. We outline the theory behind the proposed approach next.

B. Derivation and Analysis

Irrespective of the noise characteristics, the expression in leads to the lowest MSE theoretically achievable by any patch-based denoising method. This expression was derived in assum-

ing that the underlying unknown image patches are (independent) realizations of a random variable.

Algorithm 1 PLOW denoising

Input: Noisy image: $\mathbf{Y}$

Output: Denoised image: $\widehat{\mathbf{Z}}$

1: Set parameters: patch size $n = 11 \times 11$, number of clusters $K = 15$;
2: Estimate noise standard deviation $\hat{\sigma}$ [see (21)];
3: Set parameter: $h^2 = 1.75\hat{\sigma}^2 n$;
4: $\mathbf{Y}^0 \Leftarrow$ Prefilter image to obtain pilot estimate;
5: $\{\mathbf{y}_i, \mathbf{y}_i^0\} \Leftarrow$ extract overlapping patches of size n from $\mathbf{Y}$ & $\mathbf{Y}^0$;
6: $\mathbf{L} \Leftarrow$ compute LARK features for each $\mathbf{y}_i^0$;
7: $\mathbf{\Omega}_k \Leftarrow$ geometric clustering with K-means $(\mathbf{L}, K)$;
8: **foreach** *Cluster* $\mathbf{\Omega}_k$ **do**
9: Estimate mean patch $\widehat{\bar{\mathbf{z}}}$ from $\mathbf{y}_i^0 \in \mathbf{\Omega}_k$ [see (19)];
10: Estimate cluster covariance $\widehat{\mathbf{C}}_{\mathbf{z}}$ from $\mathbf{y}_i^0 \in \mathbf{\Omega}_k$ [see (20)];
11: **foreach** *Patch* $\mathbf{y}_i^0 \in \mathbf{\Omega}_k$ **do**
12: $\mathbf{y}_j^0 \Leftarrow$ identify photometrically similar patches [see (6)];
13: $w_{ij} \Leftarrow$ compute weights for all $\mathbf{y}_j^0$ [see (23)];
14: $\widehat{\mathbf{z}}_i \Leftarrow$ estimate denoised patch using $\mathbf{y}_j$ [see (17)];
15: $\mathbf{C}_{e_i} \Leftarrow$ calculate estimate error covariance [see (16)];
16: **end**
17: **end**
18: $\widehat{\mathbf{Z}} \Leftarrow$ aggregate multiple estimates from all $\{\widehat{\mathbf{z}}_i\}$ and $\{\mathbf{C}_{e_i}\}$ [see (27)].

3. PARAMETER ESTIMATION FOR DENOISING

The proposed denoising framework, graphically outlined in Fig. 1, requires us to infer various parameters from the observed noisy image. The procedure is algorithmically represented in Algorithm 1. We first identify geometrically similar patches within the noisy image. Once such patches are identified, we can use these patches to estimate the moments of the cluster, taking care to account for noise (steps 9 and 10 of Algorithm 1). Next, we identify the photometrically similar patches and calculate weights that control the amount of influence that any given patch exerts on denoising patches sim- ilar to it. These parameters are then used in to denoise each patch. Since we use overlapping patches, multiple estimates are obtained for pixels lying in the overlapping regions. These multiple estimates are then optimally aggregated to obtain the final denoised image. Since we use overlapping patches, multiple estimates are obtained for pixels lying in the overlapping regions. These multiple estimates are then optimally aggregated to obtain the final denoised image.

A. Geometric Clustering

The proposed filter was derived assuming geometrically similar patches to be sampled from some unknown pdf. So far, we have assumed such clustering to be available to us from an "oracle." To perform practical clustering, we need to identify features that capture the underlying geometric structure of each patch from its noisy observations. Such features need to be robust to the presence of noise, as well as to differences in contrast and intensity among patches exhibiting similar structural characteristics. An example of such variations among geometrically similar patches is shown in Fig. 2. Possible choices of features include contrast-adjusted image patches or principal components in conjunction with predetermined clustering guides. For where the image patches can be considerably noisy, we make use of the locally adaptive regression kernels (LARKs) introduced for denoising in and subsequently adapted as features for geometric clustering and object detection. We refer the interested reader to, where the design of the kernels is covered in detail. The number of clusters chosen affects the denoising result. In general, too few clusters can lead to structurally dissimilar patches being clustered together resulting in the incorrect estimation of the moments. On the other hand, too many clusters lead to too few patches within each cluster, making the moment estimation process less robust. Fortunately, the denoised output is not too sensitive to the choice of the number of clusters. In our experiments, we found that using a fixed value of yields good results for any given image, with the MSE fairly close to that obtained by tuning the number of clusters for that particular image.

B. Estimating Cluster Moments

Once the image is segmented into structurally similar regions, we estimate the moments, namely, mean and covariance, from the noisy member patches of each cluster. The simplest of such estimators, i.e., the sample covariance, is the maximum likelihood estimate. Although other estimators, for example, bootstrapping and shrinkage-based methods, exist, we found no discernible improvement in the denoising performance when more complex estimators were used. When the number of samples (patches in a cluster) are few compared with the dimensionality (number of pixels in each patch), the sample covariance can be unstable. For such cases, robust estimators proposed in may also be used. To estimate each patch independently without explicitly taking into account information from *estimates* of other overlapping patches, the estimation framework is in line with the assumption of the independence of underlying (noise-free) patches. However, in estimating the covariance matrix (in PLOW and also for the bounds), we do not enforce independence on the patches, and the covariance matrix estimated from overlapping patches is not necessarily diagonal. Therefore, both in our bounds and

our current paper, the correlation among the underlying noise-free patches are implicitly taken into account.

C. Calculating Weights for Similar Patches

To identify patches within the noisy image that are *photometrically similar* to a given reference patch. Once the similar patches are identified for a given reference patch, we perform denoising with the more similar patches exerting greater influence in the denoising process. This is ensured by the analytically derived weight which determines the con- tributing factor for patch in denoising the reference patch.

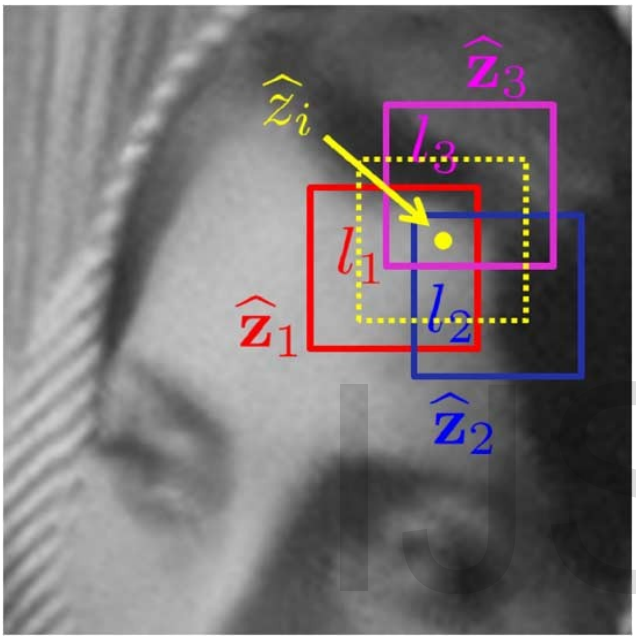

Fig3:Pixel estimation

D. Aggregating Multiple Pixel Estimates

Until now, we have estimated all the parameters needed to perform the filtering. The filter is run on a per-patch basis (although parameters are estimated from multiple patches), yielding denoised estimates for each patch of the noisy input. To avoid block artifacts at the patch boundaries, the patches are chosen to overlap each other. As a result, we obtain multiple estimates for the pixels lying in the overlapping regions. The simplest method of aggregating such multiple estimates is to average them. However, such naive averaging will lead to an over smoothened image. Alternatively, in keeping with earlier formulation, we can combine the multiple estimates in an LMMSE scheme that takes into account the relative confidence in each estimate.

4. CONCLUSION

The simplest method of aggregating such multiple estimates is to average them. However, such naive averaging will lead to an over smoothened image. Alternatively, in keeping with earlier formulation, we can combine the multiple estimates in an LMMSE scheme that takes into account the relative confidence in each estimate. While mainly developed for greyscale images, with trivial modification, our method achieves near state-of-the-art performance in denoising color images as well. Since the method works by learning the moments in geometrically similar patches, the interchannel color dependences can be implicitly captured in this framework. In a more practical setting where signal-dependent noise is observed, the clustering step needs to take into account color (or intensity) information as well. The noise in each cluster can be then assumed to be homogeneous, and the proposed filter can be independently applied in each cluster.

5.RESULTS

Here, we evaluate the proposed denoising method through experiments on various images at different noise levels. Since our method is motivated by our bounds formulation [1], we first compare the ideal denoising performance of our method (using "oracle" parameters) with the MSE predicted by the bounds. Later, we estimate the parameters directly from the noisy images, as outlined in Section III, and compare those results with various popular denoising methods. We also apply our method, with a minor modification, to color images. Finally, we address the practical case of denoising real noisy images where the noise characteristics are unknown and not necessarily Gaussian, or uncorrelated. In each case, we will show that our results are comparable, in terms of the MSE [peak signal-to-noise ratio (PSNR[6])], SSIM [45], and the recently introduced no-reference quality metric Q [46] (wherever applicable), with those obtained by state-of-the-art denoising methods and are, in many cases, vi- sually superior.

Since our method was specifically designed with the aim of achieving the theoretical limits of the performance, we first compare our results to the predicted performance bounds [1]. For this first experiment, we compute the "oracle" denoising parameters from the noise-free images. These "oracle" clusters are then used to estimate moments and $\mathbf{C}_\mathbf{z}$ from the latent image. We also use the ground-truth image to identify the photometrically similar patches and compute weights for each noise-free reference patch. The final denoising using the "oracle" param- eters is, of course, applied to the noisy image. A simple speedup for our method can be achieved by denoising only every third patch, bringing the average execution time down to approximately 17 s. Although this results in a minor drop of 0.2 dB in the PSNR, the visual differences are almost imperceptible.

MSE	PSNR	SSIM
BEAR		
31.1042	33.2026	0.8729
FIREMAN		
16.9641	35.8355	0.9622

TABLE:1 Denoising Performance

NOISY IMAGES

DENOISED IMAGES

REFERENCES

[1]. Priyam Chatterjee, *Student Member, IEEE*, and Peyman Milanfar, *Fellow, IEEE* "Patch-Based Near-Optimal Image Denoising", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 4, APRIL 2012.

[2] P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 895–911, Apr. 2010.

[3] P. Chatterjee and P. Milanfar, "Practical bounds on image denoising: From estimation to information," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1221–1233, May 2011.

[4] G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, Mar. 1994.

[5] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comput. Vis.*,Washington, DC, Jan. 1998, pp. 839–846.

[6] S. M. Smith and J. M. Brady, "SUSAN—A new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, May 1997.

[7] L. P. Yaroslavsky, *Digital Picture Processing*. Secaucus, NJ: Springer-Verlag, 1985.

[8] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[9] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising methods, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[10] S. P. Awate and R. T. Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 364–376, Mar. 2006.

[11] C. Kervrann and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image denoising and representation," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 45–69, Aug. 2008.

[12] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[13] 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.
[14] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[15] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1438–1451, Jul. 2009.