

# Entropy-based Adaptive Image Denoising

Long Ye, Haijun Gao, Qin Zhang

Key Laboratory of Media Audio & Video

Communication University of China

Beijing, China

{yelong, ghj, zhangqin}@cuc.edu.cn

**Abstract**—Image entropy, which not only describes the average amount of information about the image source, but also reflects the statistical characteristics of image data, can be described as the properties of image features and image processing basis. This paper is aimed to put forward an adaptive image denoising solution by setting the threshold and analyzing the impact of the original image data, which comes from the image noise under different entropy. For the blocks in image with low entropy value, we apply Haar wavelet method to denoise; for the medium entropy, we use field of expert (FoE) model as the prior information of the medium region denoising to achieve the image denoising of these blocks, and for the high entropy blocks, because little difference existed between ideal and noised image in human perception, no processing is implemented in this kind of blocks. The adaptability of our framework is embodied in the adjusting of the thresholds to classify different entropy blocks. Experiment results demonstrate the advantages of our framework measured by both PSNR and SSIM.

**Keywords**- Image Denoising, Entropy-based Adaptive Denoising, Haar Wavelet, Field of Expert

## I. INTRODUCTION

Image denoising, as a specific application of image restoration, has drawn many researchers' attention in the past decades. Given an ideal image  $x$  and a measured image  $y$  with the relationship as Eq.1,

$$y = x + v \quad (1)$$

where  $v$  denotes the additive zero-mean Gaussian noise, with standard deviation  $\sigma$ , the image denoising problem can be formulated as to reconstruct  $x$  by removing  $v$  from  $y$ . Therefore, one of the most important parts in image denoising is to find a transform domain that can represent the characteristics in  $x$  effectively. After the transformation,  $x$  is considered being separated with  $v$  by taking different energy distribution. One way to achieve the transforming is to build the dictionary with the existed bases directly, like DCT base [1], wavelet [2] and so on. Also, in order to better describe the details information, Curvelet[3], Contourlet[4] and Edgelet[5] are developed with the superiorities in representing the different kind of image contents separately. Without any prior information, this kind of direct transformations should be the best image denoising methods. Which is just the case of without any prior information, the denoising effect is limited.

To overcome this limitation, an image or image database prior is necessary, and the dictionary becomes the learned set of atoms. Olshausen B.A et al.[6] introduced a learned algorithm that attempts to find sparse linear codes for nature scenes, the resulting sparse image coding provided a more efficient representation because it processed a higher degree of statistical independence among its outputs. Michael.E et al. [7] denoised the image by training a dictionary on its content using K-SVD algorithm, and extended its deployment to arbitrary image sizes by defining a global image prior that forces sparsity over patches in every location in the image. Roth and Black [8] introduced the Fields of Experts (FoE) model which assumes a parametric, student T distribution for the potentials, but allows the filters to be arbitrary. If the measured image belongs to the category of training image database, the denoising efficiency is usually satisfactory. But if not, the learned dictionary may lead to a wrong guide; the content in the ideal image may be erased as it does not reflect the characteristics of the image database. Besides, the computational complexity in the dictionary training is also high.

By weighing the pros and cons of the two kinds of existed image denoising methods, this paper introduces an entropy-based adaptive image denoising framework. For the given noised image, we firstly divided it into non-overlapped  $8 \times 8$  blocks as the processing unit. And for each block, its entropy value is calculated as the guide to choose the suitable denoising way. The block with low entropy means that this block may not be noised seriously, and can be denoised easily, and then some direct transformations (Haar wavelets in this paper) are used to process in this kind of units. If with high entropy value, according to the comparison between the block from the ideal and measured image, we can see little difference subjectively. So in our framework, no process will be operated in the high entropy units. The denoising in units with medium entropy value is the key problem that greatly affects the denoising performance. We choose Fields of Experts (FoE) model which have been proven as the more effective image restoration method via prior information learning to denoise the blocks with medium entropy. The adaptability of our framework embodied in the adjusting of the thresholds to separate different entropy blocks.

The structure of our paper is as followed: In section 2, we formulate the relationship between entropy value and image denoising, and the entropy based adaptive image denoising algorithm is given; section 3 describes the framework in details, including Haar, FoE and the entropy adaptive scheme;

the experiments is showed and analyzed in section 4, the results demonstrate the advantages of our framework by measured both in PSNR and in SSIM.

## II. PROBLEM FORMULATION

In order to represent how entropy guiding the image denoising, an experiment is designed according to the statistical analysis of the 6 images with different entropy value showed as figure 1, where the image set is a sequence of 6 ivy wall images of taken at 6 viewing distances. The viewing distance of the  $i+1$ -st images is half that of the  $i$ -th image. The images display markedly different entropy value (reduce from 1 to 6) even though they are generated by the same scene. The size of the image is  $256 \times 256$ , means that every image can be divided into  $1024 \ 8 \times 8$  blocks(or Units).

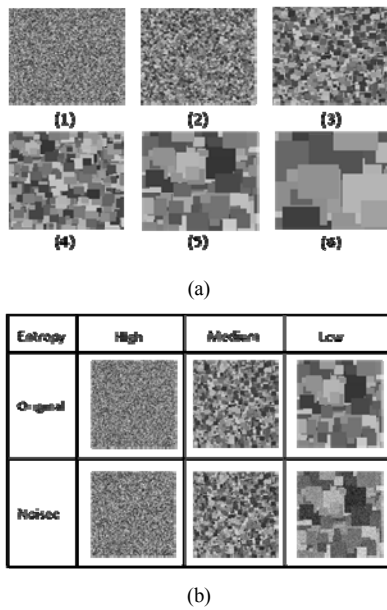


Figure 1: (a) 6 scales Ivy Wall image with decreasing entropy value from (1) to (6) and (b) Comparison between the original images and noised images with the high, medium and low entropy value separately.

After adding noise to the 6 scale ivy wall images, the comparison between the original images and noised images is showed as figure 2. To represent the noise influence to the images with different entropy values, we choose image (1), (3) and (5) as the high, medium and low entropy image separately. For high entropy image, because its structure is complex, the differences between original and noised image is not significant (not evaluated objectively but subjectively), no matter what method is to be used to denoise, the results are not particularly good. But for low entropy images, the structure is simple, in addition, it can still be able to identify the details of their images and other information after adding noise, it doesn't cause greater interference to its message, the use of relatively simple denoising method can be better for recovering the image. Therefore, for high and low entropy images, there respectively exist features: the former one has the feature of difficulties to distinguish the original image data and the noised, besides, it is not easy for human eye to

perceive; the latter one has the feature of being easy to distinguish the original image data and the noise, we apply Haar wavelet, which is not only efficient, but also thought to be the preferred method for denoising. Compared with high entropy image, medium entropy image is not particularly complex in structure; and compared with low entropy image, the noise causes greater interference to the message which the medium entropy image provides. So because the connection between the original image data and noise data is too tight to distinguish, we use expert field model, which we have proven to be a better method to achieve the medium entropy image denoising.

In summary, the entropy based adaptive image denoising algorithm can be described as below:

*Algorithm: Entropy-based Adaptive Image Denoising*

*Step 1: Divide the noised image into  $N$  non-overlapping  $8 \times 8$  blocks*

*Step 2: while  $i \leq N$*

*Calculate the entropy value of the  $i$ -th block as  $H_i$*

*Switch  $H_i$*

*Case  $H_i \leq \sigma_1$*

*Denoise the  $i$ -th block with Haar Wavelet;*

*Case  $\sigma_2 \leq H_i < \sigma_1$*

*Denoise the  $i$ -th block with FoE;*

*Case  $H_i \leq \sigma_2$*

*No denoise processing;*

*end*

*end while*

*Step 3: Adjust  $\sigma_1$  and  $\sigma_2$  according to the denoising result (Evaluated by PSNR or SSIM). If satisfy, end the program; if not, go to step 2.*

The adaptability of our framework embodied in the adjusting of the two thresholds  $\sigma_1$  and  $\sigma_2$ . As we mentioned before, the learned dictionary for medium entropy block is a double-edged sword. If the noised block can reflect more commonness contained in the learned dictionary, the denoising result will be satisfied; if not, the general transformation used in low entropy block could be better. As well, the setting of  $\sigma_2$  also should be considered in the same way, then the entropy-based adaptive image denoising have the formulation as:

$$[\sigma_1, \sigma_2]_{\text{opt}} = \arg \max_{[\sigma_1, \sigma_2]} P(E|\sigma_1, \sigma_2) \quad (2)$$

where  $E$  is the energy function Evaluated by PSNR or SSIM, and  $P(E|\sigma_1, \sigma_2)$  the likelihood corresponding to the energy function.

## III. ENTROPY-BASED ADAPTIVE IMAGE DENOISING

Form the description in section II, we can see that there are three main parts in the entropy-based adaptive image denoising framework: Haar wavelet-based and FoE based denoising, entropy based adaptability. This section will introduce the three parts in detail.

### A. Haar Wavelet based denoising

For the given noised color image, we divide it into RGB three matrixes, for each two-dimensional matrix, we use Haar wavelet decomposition and reconstruction respectively, and finally merge to get the final denoising images. 1D Haar wavelet function is defined in the form as:

$$\varphi(t) = \begin{cases} 1, & 0 \leq t < 0.5 \\ -1, & 0.5 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

When extending the Eq.3 to 2D bases, there are A, H, V and D four components respectively represent approaching, horizontal, vertical and diagonal component separately. The latter three components collectively refer the details or noise. We use Haar wavelet to decompose the image with the scale of 2. And, we get the approximate reconstructed image as the denoised result by computing  $\langle H1, I \rangle$ , where  $\langle \cdot \rangle$  denotes the inner production and  $I$  is the processing block. Similarly, we can also get the details of image by computing  $\langle H2, I \rangle$ ,  $\langle H3, I \rangle$  and  $\langle H4, I \rangle$ .

### B. FoE based image denoising

The dictionary trained by FoE can be applied to a Bayesian formulation with a prior term. Given a noised image  $y$ , our goal is to find the true image  $x$  that maximizes the posterior probability as:

$$P(x|y) \propto P(y|x)P(x) \quad (4)$$

Our experiments assume that the true image has been corrupted by additive, i.i.d. Gaussian noise with zero mean and known standard deviation  $\sigma$ . We thus write the likelihood as:

$$P(y|x) \propto \prod_j \exp\left(-\frac{1}{2\sigma^2} (y_j - x_j)^2\right) \quad (5)$$

where  $j$  ranges over the pixels in the image. In order to emphasize the practicality of the proposed model, we refrain from using expensive inference techniques. Instead we perform a gradient ascent on the logarithm of the posterior probability.

### C. Entropy-based Adaptability

After adding noise to the 6 images showed in figure 1, we firstly calculate the entropy value for each block. Then the entropy histogram can be received and showed in figure 2(a). As image (1), (3) and (5) represent the high, medium and low entropy image form separately, the initialized values of  $\sigma_1$  and  $\sigma_2$  can be set as the two lines in the figure 2(a). In detail, we set the threshold of the high-entropy and medium-entropy to 3.5, the medium-entropy and the low-entropy to 2.5. The blocks whose entropy value greater than 3.5 are set to blue, the entropy values between 2.5 and 3.5 are set to green, the ones which are less than 2.5 are set to red. Therefore, the entropy distribution is showed as figure 2(b). When the  $\sigma_1$  and  $\sigma_2$  values are initialized, the problem showed in Eq.2 becomes the optimization in the neighborhood of  $[\sigma_1 \ \sigma_2]_{\text{initialized}}$ .

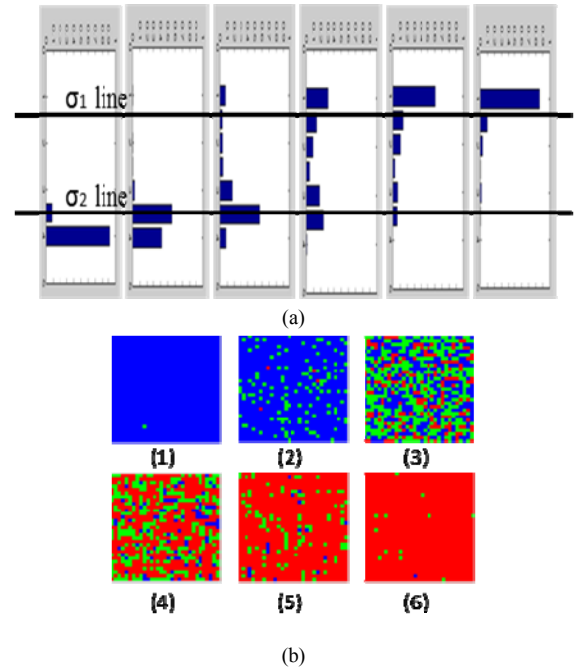


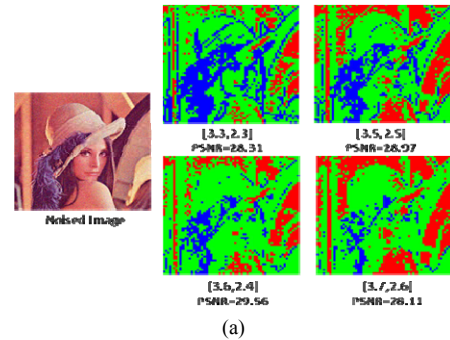
Figure 2: Threshold setting for the image blocks. Blue green and red present high entropy, medium and low entropy blocks separately.

## IV. EXPERIMENTS

### A. Optimization for the threshold $\sigma_1$ and $\sigma_2$

As described before, by analyzing the entropy values of the images and initializing the value  $\sigma_1$  and  $\sigma_2$ , the problem embodied in Eq.2 becomes the optimization in the neighborhood of  $[\sigma_1 \ \sigma_2]_{\text{initialized}}$ . We set the searching region as  $[\sigma_1 \ \sigma_2]_{\text{initialized}} \pm [0.5 \ 0.5]$ , and still set blocks whose entropy value greater than  $\sigma_1$ , between  $\sigma_2$  and  $\sigma_1$  and less than  $\sigma_2$  as blue, green, and red separately. Then we get the images in figure 3.

From figure 3, we see that different values of  $\sigma_1$  and  $\sigma_2$  can achieve different denoising performance (PSNR gain in this simulation). According to the optimization, we can find the optima in the neighborhood of  $[\sigma_1 \ \sigma_2]_{\text{initialized}}$  as  $[3.6 \ 2.4]$ , with the PSNR gain as 29.56. In this paper, we also compared our adaptive framework with the Haar wavelet-based and FoE-based image denoising methods.



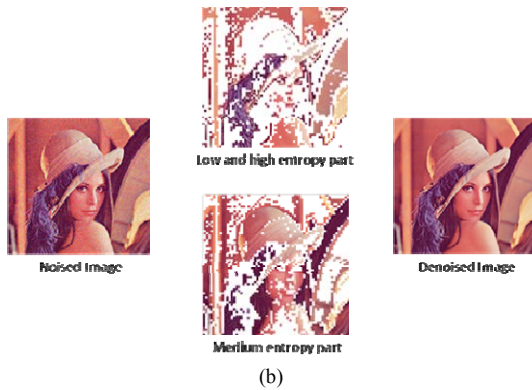


Figure 5: (a) Image denoising results comparison with different values of  $[\sigma_1 \sigma_2]$  and (b) Denoising results.

### B. Comparison with other methods

Basing on idea of the entropy-based adaptive framework, we divide the test ‘Lena’ image into two parts with the  $[\sigma_1 \sigma_2]=[3.6 \ 2.4]$ , which have been found as the optima, one contains the high and low entropy blocks, the medium entropy region is the other, showed as figure 6. Then we apply field of expert model to the medium entropy image to denoise and use Haar wavelet-based denoising method for part of low regions. Finally, we use pixel assignment method to get the final result.

In this paper, we compared our adaptive framework with the Haar wavelet-based and FoE-based image denoising methods. In addition to PSNR, we also computed a more perceptually-based similarity measure (SSIM) [9]. Table 1 shows the comparisons in the condition of adding the noise with different variances. The experiments result demonstrates the advantages of our method.

TABLE I: Denoising results comparison

$\sigma$	10	15	20	25
Haar(PSNR)	29.66	28.06	26.53	25.09
FoE(PSNR)	29.70	28.50	27.55	25.98
Our Method(PSNR)	29.74	28.86	27.82	26.71
Haar(SSIM)	0.9319	0.8726	0.8061	0.7477
FoE(SSIM)	0.9401	0.8733	0.8069	0.7489
Our Method(SSIM)	0.9406	0.8782	0.8079	0.7517

## V. CONCLUSION

In this paper, a novel image denoising framework is introduced based on the entropy. We applied our framework in the situation that different variances noise is added to the ideal image. No matter measured by PSNR or SSIM, our framework all showed its advantages.

In the future work, a new low entropy denoising method should be adapted in place of the Haar wavelet as the Haar may be too simple to denoise effectively. Besides, as the training step in FoE is with high computable complex, some off-line prior information should be introduced to make our framework more practical.

## ACKNOWLEDGMENT

This paper is supported by National Nature Science Foundation of China (60832004 and 61201236).

## REFERENCES

- [1] Moulin, P. and Liu, J. ‘Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors’, IEEE Transactions on Information Theory, Vol. 45, No. 3, pp. 909–919, April.1999.
- [2] Jansen, M. ‘Noise Reduction by Wavelet Thresholding’, SpringerVerlag, New York, 2001
- [3] Candès, E.J. and Donoho, D.L. ‘New tight frames of curvelets and the problem of approximating piecewise C2 images with piecewise C2 edges’, Comm. Pure Appl. Math., Vol. 57, pp. 219–266, February, 2004.
- [4] Do.M.N. and Vetterli, M. ‘Contourlets, Beyond Wavelets’, G. V. Welland ed., Academic Press, 2003.
- [5] Donoho, D.L. ‘Wedgelets: Nearly minimax estimation of edges’, Annals Of Statistics, Vol. 27, No. 3, pp. 859–897, June, 1998.
- [6] Olshausen, B.A. and Field. D.J. ‘Sparse coding with an overcomplete basis set: A strategy employed by V1?’, Vision Research, Vol. 37, pp. 311–325, 1997.
- [7] Aharon, M., Elad, M., and Bruckstein, A.M. ‘The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representation’, to appear in the IEEE Trans. On Signal Processing. 2005.
- [8] S. Roth and M. J. Black. ‘Fields of experts: A framework for learning image priors’ In IEEE Conf. on ComputerVision and Pattern Recognition, 2005.
- [9] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. ‘Image quality assessment: From error visibility to structural similarity’, IEEE Trans. Image Proc., 13(4):600–612, 2004.