A novel approach for denoising of an image using modified gradient histogram preservation

G. Sankarbabu, Rajalakshmi. P, Surekha Poomathi. J
Department of Electronics and Communication Engineering

VELTECH. Chennai-62
sankarbabuvlsi@gmail.com, rajalakshmipadmanathan@gmail.com, poomathisuray412@gmail.com

Abstract - Image denoising is one of the basic problem in low level vision . It is one of the best method to evaluate various statistical image modeling methods. One of the most important problems in image denoising is how to preserve the fine scale texture structures while removing noise. To address this problem, we propose a texture enhanced image denoising (TEID) method by using the gradient distribution of the denoised image to be nearly close to the estimated gradient distribution of the original image .Some natural image priors, such as nonlocal self-similarity prior, sparsity prior, and gradient based prior, have been used mostly for noise removal. The denoising algorithms based on these priors tend to smooth the detailed image textures, degrading the image visual quality. A gradient histogram preservation (GHP) algorithm is developed to enhance the texture structures while removing noise.

Key words : Similarity prior, Sparsity prior, Gradient based prior, Gradient histogram preservation

1.Introduction

The main aim of image denoising is to estimate thebest clean image **x** from its noisy observation **y**. One commonly used observation model is **y = x + v** , where **v** is additive white Gaussian noise. Image denoising is a classical yet still active topic in image processing and low level vision, while it is an ideal test bed to evaluate various statistical image modeling methods. In general, we hope that the denoised image should look like a natural image, and therefore the statistical modeling of natural image priors is crucial to the success of image denoising.

With the rapid development of digital imaging, the resolution of imaging sensor is getting higher and higher. On one hand, more fine texture features of the object and scene will be captured; on the other hand, the captured high resolution image is more prone to noise because the smaller size of each pixel makes the exposure less sufficient. However, suppressing noise while preserving textures is difficult to achieve simultaneously, and this has been one of the most challenging problems in natural image denoising.

Unlike large scale edges, the fine scale textures have much higher randomness in local structure and they are hard to characterize by using a local model. Considering the fact image are homogeneous that texture regions in an and are usually composed of similar patterns, statistical descriptors such as histogram are more effective to represent them. Actually, in literature of texture representation and classification global histogram of some local features is dominantly used as the final feature descriptor for matching. Meanwhile, image gradients convey most of semantic information in an image and are crucial to the human perception of image visual quality. All these motivate us to use the histogram of image gradient to design new image denoising models.

With the above consideration, in this paper we propose a novel method for texture enhanced image denoising (TEID) Via gradient histogram preservation (GHP). From the given noisy image **y**, we will estimate the gradient histogram of original image **x**. Take this estimated histogram, denoted by $\mathbf{h}_r$, as a reference, we search for an estimate of **x** with GHP, i.e., the gradient histogram of the denoised image should be close to $\mathbf{h}_r$. , the proposed TEID method can well enhance the image texture regions, which are often over-smoothed by other denoising methods. The major contributions of this paper are as follows:

(1) A novel image denoising framework, i.e., TEID, is proposed, which preserves the gradient distribution of the original image. The existing image priors can be easily incorporated into the proposed framework to improve the quality of denoised image.

(2) A histogram specification operator is developed to ensure the gradient histogram of denoised image being close to the reference histogram, resulting in a simple yet effective GHP based TEID algorithm.

(3) A simple but theoretically solid algorithm is presented to estimate the gradient histogram from the given noisy image, making TEID practical to implement.

Many existing image denoising algorithms, including those sparsity and NSS priors based ones, tend to wipe out the image detailed textures while removing noise.

2. Denoising with gradient histogram preservation

In this section, we first present the image denoising model by gradient histogram preservation with sparse nonlocal regularization, and then present an effective histogram specification algorithm to solve the proposed model for texture enhanced image denoising.

Given a clean image **x**, the noisy observation **y** of **x** is usually modeled as

$$\mathbf{y} = \mathbf{x} + \mathbf{v}, \qquad (1)$$

where **v** is the additive white Gaussian noise (AWGN) with zero mean and standard deviation . The goal of image denoising is to estimate the desired image **x** from **y**. One popular approach to image denoising is the variational method, in which the denoised image is obtained by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ \tfrac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{x}\|^2 + \mu \cdot R(\mathbf{x}) \right\}$$

where $R(\mathbf{x})$ denotes some regularization term and μ is a positive constant. The specific form of $R(\mathbf{x})$ depends on the used image priors. One common problem of image denoising methods is that the image fine scale details such as texture structures will be over-smoothed. An over-smoothed image will have much weaker gradients than the original image. Intuitively, a good estimation of **x** without smoothing too much the textures should have a similar gradient distribution to that of **x**. With this motivation, we propose a gradient histogram preservation (GHP) model for texture enhanced image denoising (TEID).

Our intuitive idea is to integrate the gradient histogram prior with the other image priors to further improve the denoising performance. Suppose that we have an estimation of the gradient histogram of **x**, denoted by $\mathbf{h}_r$ (the estimation) method will be discussed in Section 4). In order to make the gradient histogram of denoised image ˆ**x** nearly the same as the reference histogram $\mathbf{h}_r$, we propose the following GHP based image denoising model:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x},F} \left\{ \tfrac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{x}\|^2 + \lambda R(\mathbf{x}) + \mu \|F(\nabla \mathbf{x}) - \nabla \mathbf{x}\|^2 \right\} \quad \text{s.t. } \mathbf{h}_F = \mathbf{h}_r , \qquad (2)$$

Where *F* denotes an odd function which is monotonically non-descending in (0, +∞), $\mathbf{h}_F$ denotes the histogram of the transformed gradient image |*F* (A**x**)|, and A denotes the gradient operator. By introducing the transform *F*, we can usethe alternating method for image denoising. Given *F*, we can fix A**x**0 = *F*(A**x**), and use the conventional denoising methods to update **x**. Given **x**, we can update *F* simply by the histogram operator introduced Thus, with the introduction of *F*, we can easily incorporate gradient histogram prior with any existing image priors *R*(**x**). The sparsity and NSS priors have shown promising performance in denoising, and thus we integrate them into the proposed GHP model. Specifically, we adopt the sparse nonlocal regularization term proposed in the centralized
sparse representation (CSR) model [7], resulting in the following denoising model:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x},F} \left\{ \begin{array}{l} \tfrac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{x}\|^2 + \lambda \sum_i \|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i\|_1 \\ +\mu \|F(\nabla \mathbf{x}) - \nabla \mathbf{x}\|^2 \end{array} \right\} , \quad \text{s.t. } \mathbf{x} = \mathbf{D} \circ \boldsymbol{\alpha},\ \mathbf{h}_F = \mathbf{h}_r \qquad (3)$$

where _ is the regularization parameter, **D** is the dictionary and _ is the coding coefficients of **x** over **D**.

Let's explain more about the model in Eq. (3). Let $\mathbf{x}_i = \mathbf{R}_i\mathbf{x}$ be a patch extracted at position *i*, *i* = 1, 2, . . . , *N*, where $\mathbf{R}_i$ is the patch extraction operator and *N* is the number of pixels in the image. Each $\mathbf{x}_i$ is coded over the dictionary **D**, and the coding coefficients is $\_i$. Let _ be the concatenation of all $\_i$, and then **x** can be reconstructed by

$$\mathbf{x} = \mathbf{D} \circ \boldsymbol{\alpha} \triangleq \left(\sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{D} \boldsymbol{\alpha}_i . \qquad (4)$$

In Eq. (3), β_i is the nonlocal means of $\_i$ in the sparse Coding domain. With the current estimate ˆ**x**, we use the blocking matching method as in [7] to find the non-local neighbors of Then β_i is computed as the weighted average

$$\boldsymbol{\beta}_i = \sum_q w_i^q \boldsymbol{\alpha}_i^q ,$$

where the weight is defined as

$$w_i^q = \tfrac{1}{W} \exp\left(-\tfrac{1}{h} \left\| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_i^q \right\|^2 \right), \qquad (6)$$

From the GHP model with sparse nonlocal regularization in Eq. (3), one can see that if the histogram regularization parameter μ is high, the function *F* (A**x**) will be close to A**x**. Since the histogram $\mathbf{h}_F$ of |*F* (A**x**)| is required to be the same as $\mathbf{h}_r$, the histogram of A**x** will be similar to $\mathbf{h}_r$, leading to the desired gradient histogram preserved image denoising. Next, we will see that there is an efficient iterative histogram specification algorithm to solve the model in Eq. (3).

3. Iterative histogram specification algorithm

Eq. (3) is minimized iteratively. As in [7], the local PCA bases are used as the dictionary **D**. Based on the current estimation of image **x**, we cluster its patches into *K* clusters, and for each cluster, a PCA dictionary is learned. Then for each given patch, we first check which cluster it belongs,
and then use the PCA dictionary of this cluster as the **D**. We propose an alternating minimization method to solve the problem in Eq. (3). Given the transform function *F*, we introduce a variable $\mathbf{g} = F(\nabla \mathbf{x})$, and update **x** (i.e., _) by solving the following sub-problem:

$$\min_{\mathbf{x}} \left\{ \tfrac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{x}\|^2 + \lambda \sum_i \|\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i\|_1 + \mu \|\mathbf{g} - \nabla \mathbf{x}\|^2 \right\} \quad \text{s.t. } \mathbf{x} = \mathbf{D} \circ \boldsymbol{\alpha} . \qquad (7)$$

To get the solution to the above sub-problem, we first use a gradient descent method to update $\mathbf{x}$:

$$\mathbf{x}^{(k+1/2)} = \mathbf{x}^{(k)} + \delta\left(\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{x}^{(k)}) + \mu\nabla^T\left(\mathbf{g} - \nabla\mathbf{x}^{(k)}\right)\right). \quad (8)$$

. Then, the coding coefficients α_i are updated by

$$\boldsymbol{\alpha}_i^{(k+1/2)} = \mathbf{D}^T\mathbf{R}_i\mathbf{x}^{(k+1/2)}. \quad (9)$$

By using Eq. (5) to obtain β_i, we further update α_i by

$$\boldsymbol{\alpha}_i^{(k+1)} = S_{\lambda/d}\left(\begin{array}{c}\frac{1}{d}\mathbf{D}^T\left(\mathbf{R}_i\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}_i^{(k+1/2)}\right)\\ +\boldsymbol{\alpha}_i^{(k+1/2)} - \boldsymbol{\beta}_i\end{array}\right) + \boldsymbol{\beta}_i, \quad (10)$$

Once the estimate of image $\mathbf{x}$ is given, we can update F by solving the following sub-problem:

$$\min_F \|F(\nabla\mathbf{x}) - \nabla\mathbf{x}\|^2 \text{ s.t. } \mathbf{h}_F = \mathbf{h}_r. \quad (11)$$

Finally, we summarize our proposed iterative histogram specification based GHP algorithm in **Algorithm 1**. It should be noted that, for any gradient based image denoising model, we can easily incorporate the proposed GHP in it by simply modifying the gradient term and adding an extra histogram specification operation

Algorithm 1: Iterative Histogram Specification for GHP

1. Initialize $k = 0, \mathbf{x}^{(k)} = \mathbf{y}$
2. Iterate on $k = 0, 1, ..., J$
3. Update $\mathbf{g}$:
   $\mathbf{g} = F(\nabla\mathbf{x})$
4. Update $\mathbf{x}$:
   $\mathbf{x}^{(k+1/2)} = \mathbf{x}^{(k)} + \delta\left(\begin{array}{c}\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{x}^{(k)})\\ +\mu\nabla^T(\mathbf{g} - \nabla\mathbf{x}^{(k)})\end{array}\right)$
5. Update the coding coefficients of each patch:
   $\boldsymbol{\alpha}_i^{(k+1/2)} = \mathbf{D}^T\mathbf{R}_i\mathbf{x}^{(k+1/2)}$
6. Update the nonlocal mean of coding vector $\boldsymbol{\alpha}_i$:
   $\boldsymbol{\beta}_i = \sum_q w_i^q\boldsymbol{\alpha}_i^q$
7. Update $\boldsymbol{\alpha}$:
   $\boldsymbol{\alpha}_i^{(k+1)} = S_{\lambda/d}\left(\begin{array}{c}\frac{1}{d}\mathbf{D}^T(\mathbf{R}_i\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}_i^{(k+1/2)})\\ +\boldsymbol{\alpha}_i^{(k+1/2)} - \boldsymbol{\beta}_i\end{array}\right) + \boldsymbol{\beta}_i$
8. Update $\mathbf{x}$
   $\mathbf{x}^{(k+1)} = \mathbf{D} \circ \boldsymbol{\alpha}^{(k+1)}$
9. Update F via histogram specification by Eq. (11)
10. $k \longleftarrow k+1$
11. $\mathbf{x} = \mathbf{x}^{(k)} + \delta\left(\mu\nabla^T(\mathbf{g} - \nabla\mathbf{x}^{(k)})\right)$

4. Reference gradient histogram algorithm

To apply the model in Eq. (3), we need to know the reference histogram $\mathbf{h}_r$, which is supposed to be the gradient histogram of original image $\mathbf{x}$. In this section, we propose a one dimensional deconvolution model to estimate the histogram $\mathbf{h}_r$. Assuming that all pixels in the gradient image $\nabla\mathbf{x}$ are independent and identically distributed (i.i.d.), we can view them as the samples of a scalar variable, denoted by x.

Then the normalized histogram of $\nabla\mathbf{x}$ can be regarded as a discrete approximation of the probability density function (PDF) of x. For the additive white Gaussian noise (AWGN) $\mathbf{v}$, we can readily model its elements as the samples of an i.i.d.

variable, denoted by v. Since $v \sim N$ _0, _2_ and let $g = \nabla v$, one can obtain that g is also i.i.d. Gaussian. with PDF [22]

$$p_g = \frac{1}{2\sqrt{\pi}\sigma}\exp\left(-\frac{g^2}{4\sigma^2}\right). \quad (12)$$

Since $\mathbf{y} = \mathbf{x} + \mathbf{v}$, we have A$\mathbf{y}$ = A$\mathbf{x}$ + A$\mathbf{v}$. It is ready to model A$\mathbf{y}$ as an i.i.d. variable, denoted by y, and we have $y = x + g$. Let p_x be the PDF of x, and p_y be the PDF of y. Since x and g are independent, the joint PDF $p(x, g)$ is,

$$p(x, g) = p_x \times p_g. \quad (13)$$

Then the PDF p_y is

$$p_y(y = t) = \int_a p_x(x = a) \times p_g(g = (t - a))\,da. \quad (14)$$

If we use the normalized histogram $\mathbf{h}_x$ and $\mathbf{h}_y$ to approximate p_x and p_y, we can rewrite Eq. (14) in the discrete domain as:

$$\mathbf{h}_y = \mathbf{h}_x \otimes \mathbf{h}_g, \quad (15)$$

where Ⓢdenotes the convolution operator. Note that $\mathbf{h}_g$ can be obtained by discretizing p_g, and $\mathbf{h}_y$ can be computed directly from the noisy observation $\mathbf{y}$. Obviously, the estimation of $\mathbf{h}_x$ can be generally modelled as a deconvolution problem:

$$\mathbf{h}_r = \arg\min_{\mathbf{h}_x}\left\{\left\|\mathbf{h}_y - \mathbf{h}_x \otimes \mathbf{h}_g\right\|^2 + c \cdot R(\mathbf{h}_x)\right\}, \quad (16)$$

5.Parameter setting

Our algorithm involves a few parameters to set, including the regularization parameters _ and μ in Eq. (7) to balance the effect of gradient preservation, constant _ in Eq.(8) and *d* in Eq. (10) to ensure convexity. For the parameter we use the same strategy as in [8] to adaptively update it according to the maximum a posterior (MAP) principle. Based on our experimental experience, we set the parameter μ to 5, and _ to 0.23 for noise level less than 30 while 0.26 for other noise levels. Based on the analysis in [6], to guarantee the convexity of surrogate function, *d* should be larger than the spectral norm of dictionary **D**. Since in our algorithm **D** is an orthonormal PCA matrix, any *d* greater than

1 will be fine, and we set it to 3 by experience. Christo Ananth et al. [1] proposed a system in which OWT extracts wavelet features which give a good separation of different patterns. Moreover the proposed algorithm uses morphological operators for effective segmentation. From the qualitative and quantitative results, it is concluded that our proposed method has improved segmentation quality and it is reliable, fast and can be used with reduced computational complexity than direct applications of Histogram Clustering. The main advantage of this method is the use of single parameter and also very faster. While comparing with five color spaces, segmentation scheme produces results noticeably better in RGB color space compared to all other color spaces. Note that these parameters are fixed to all images in our experiments.

6. Denoising results

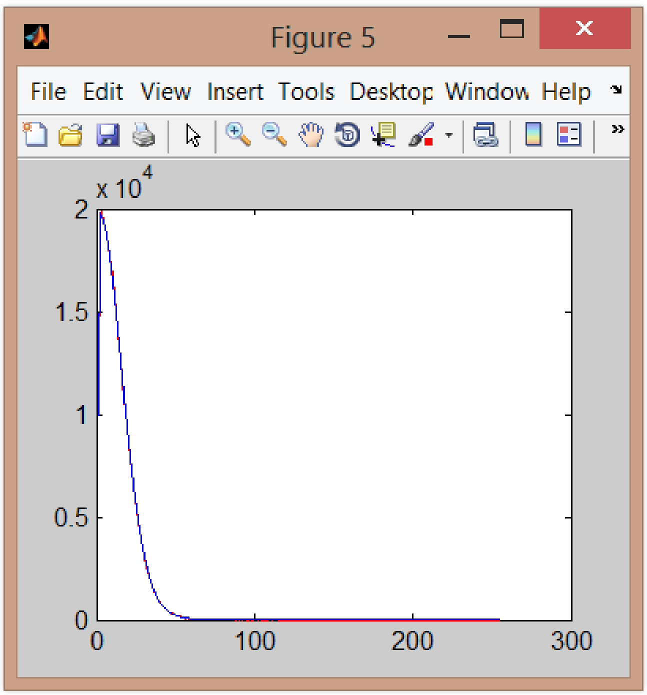


Fig 1. Graph of gradient

The above graph gives us the distribution of the gradients.

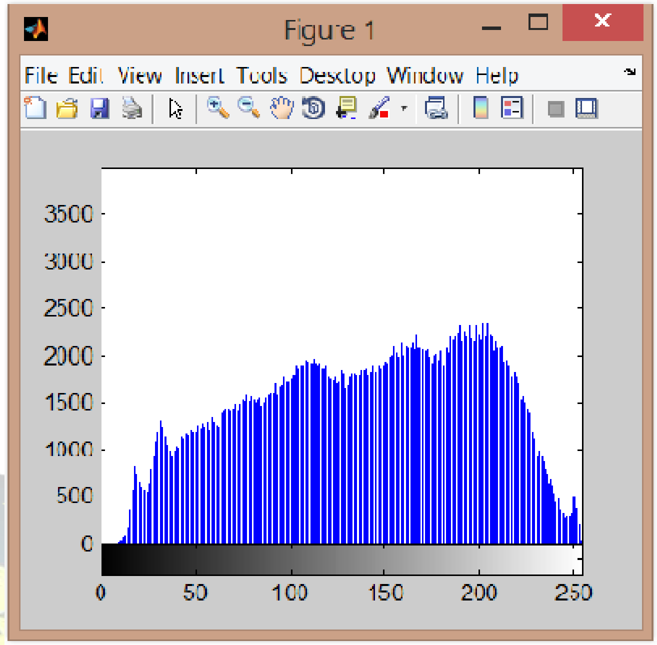


Fig 2. Graph of histogram

The above given figure1 give us the information regarding the distribution of the gradients.

Image gradients convey most of semantic information in an image and are crucial to the human perception of image visual quality

The figure 2 give us the information histogram distribution of the gradients

An over-smoothed image will have much weaker gradients than the original image.

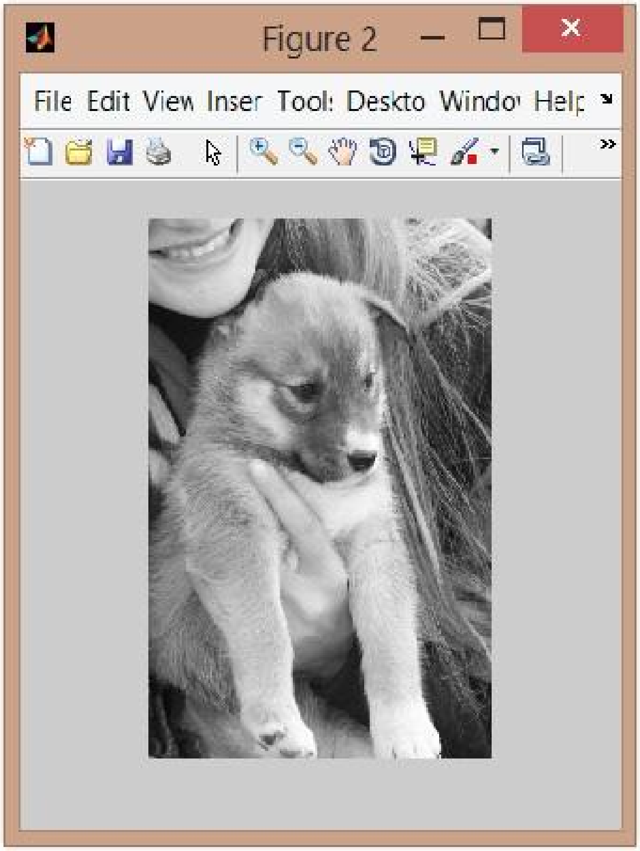
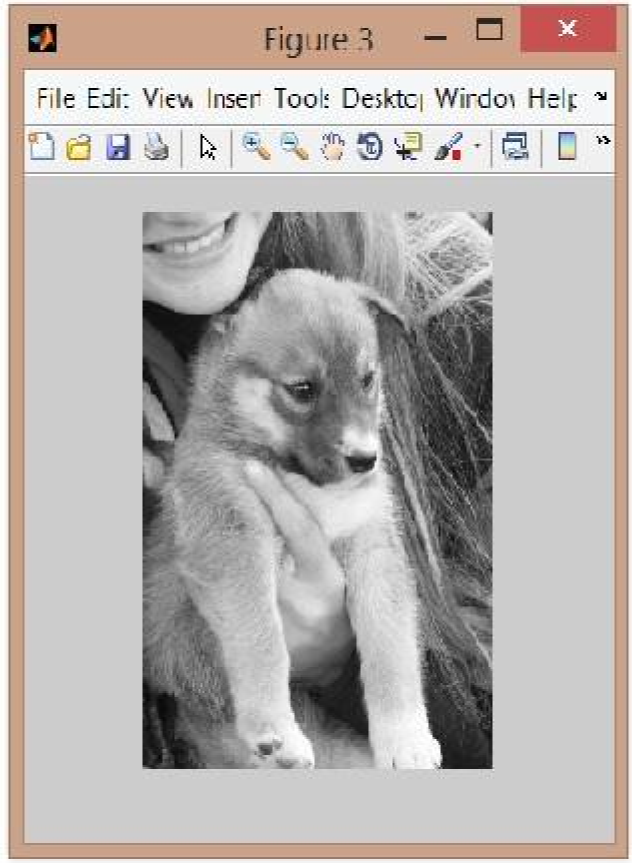


Fig 3. Denoised image

The above figure 3 gives us the denoised image which is obtained by the adding the Gaussian noise

The figure 4 gives us the clean image obtained by removing the noise from the denoised image.



Clean image

7. Conclusion

In this paper, we used a novel gradient histogram preserving (GHP) model for texture-enhanced image denoising(TEID). The GHP model can preserve the gradient distribution by pushing the gradient histogram of the denoised image toward the reference histogram, and thus is promising in enhancing the texture structure while re moving random noise. To implement the GHP model, we proposed an efficient iterative histogram specification algorithm.

REFERENCES

[1] Christo Ananth, A.S.Senthilkani, Praghash.K, Chakka Raja.M., Jerrin John, I.Annadurai, "Overlap Wavelet Transform for Image Segmentation", International Journal of Electronics Communication and Computer Technology (IJECCT), Volume 4, Issue 3 (May 2014), pp-656-658

[2] A. Buades, B. Coll, and J. Morel. A review of image de-noising methods, with a new one. Multiscale Model. Simul., 4(2):490–530, 2005.

[3] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: can plain neural networks compete with bm3d? Proc. CVPR, 2012.

[4] T. S. Cho, N. Joshi, C. L. Zitnick, S. B. Kang, R. Szeliski, and W. T. Freeman. A content-aware image prior. Proc. CVPR, 2010.

[5] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman. Image restoration by matching gradient distributions. *IEEE T-PAMI*, 34(4):683–694, 2012.

[6] I. Daubechies, M. Defriese, and C. DeMol. An iter-ative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math., 57(11):1413–1457, 2004.

[7] W. Dong, L. Zhang, and G. Shi. Centralized sparse representation for image restoration. *Proc. ICCV*, 2011.

[8] W. Dong, L. Zhang, G. Shi, and X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE T-IP*, 20(7):1838–1857, 2011.

[9] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE T-IP*, 15(12):3736–3745, 2006.

[10] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. Proc. ACM SIGGRAPH, 2006.

[11] B. Georgescu, I. Shimshoni, and P. Meer. Mean shift based clustering in high dimensions: A texture classification example. *Proc. ICCV*, 2003.

[12] R. C. Gonzalez and R. E. Woods. Digital image processing. Beijing: Publishing House of Electronics Industry, 2005.

[13] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar. Multiresolution histograms and their use for . recognition. *IEEE T-PAMI*, 26(7):831–847, 2004

[14] J. Jancsary, S. Nowozin, and C. Rother. Loss-specific training of non-parametric image restoration models: a state of the art. Proc. ECCV, 2012.

[15] N. Joshi, C. L. Zitnick, R. Szeliski, and D. Kriegman. Image deblurring and denoising using color priors. Proc. CVPR, 2009.

[16] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola. From Local kernel to nonlocal multiple-model image denoising. *IJCV*, 86(1):1–32, 2010.