

## **Image Enhancement by Hypothesis based Combination of multiple Filtering with multiple algorithms**

Najmal A, Kavitha N Nair

PG Scholar, Department of ECE, University College of Engineering, Muttom, Kerala  
Assistant Professor, Department of ECE, University College of Engineering Muttom, Kerala

---

**Abstract:** In image enhancement process involves removal of noise and distortions like artifacts ,to improve the visual perspective for a viewer. Due to unique characteristics the document image having different contents with different regions which distorted differentially by noises and several artifacts. In many applications, single filtering approach will not give a good result at each pixel locations. So we provide a new approach a Combined Filtering based on a hypothesis for image quality improvement . It utilizes an approach to select a set of filters to improve the quality of distorted region of image with different regions. The HCF extracted the featured vector to predict the performance of filtering for estimating pixel intensity in original image. Maximum likelihood estimates of the model parameters are calculated by unsupervised clustering Expectation Maximization(EM) algorithm and FCM clustering ,which used to weighting for the filter output. In this way, the HCF serves as a framework for combining the outputs of a number of different user selected filters, each best suited for a different region of an image. To improve contrast quality Bi histogram equalization techniques applied for various images such as remote sensing images and general images which are obtained at the output of combined filtering

The scheme consistently improves the quality of the decoded image over a variety of image content having different characteristics.

**Keywords:** Denoising, Artifacts,HCF,EM algorithm,FCM clustering,Maximum likelihood estimate,feature vectors,Ringing artifacts,Blocking artifacts,Bi histogram equalization

---

### **I. INTRODUCTION**

In many applications a single specific image enhancement filter will not produce the best quality at all locations in a complex image., will tend to blur edge detail. A document image typically consists of text, graphics, and natural images in a complex lay out. Each type of content is distorted differently by different types of noise and several JPEG artifacts due to its unique characteristics. For example text ,graphics region in natural image contain many sharp edges that lead to several types of artifacts. Typical images are also affected with different kinds of noise like Gaussian, random, speckle, salt and pepper noise[1][2]. A filter is designed to remove Gaussian noise from smooth regions in an image will tend to blur edge detail. Or a designed non linear filter which tend to preserve edges , may produce undesirable distortion in smooth regions in the image.

Denoising [3] is an important factor in image processing before it proceeding to the other image processing developments like edge detection, registration segmentation etc. The pixel characteristics in an image retained in the low frequency region , hence by traditional low pass filtering noise signal corrupted the pixel should be removed. But edge information is present at the high frequency components, so we need to maintain these information while denoising process.In the case of satellite images , which are degraded by noise during acquisition and transmission process.SAR uses microwave radiation so it can be illuminate the earth surface. The image details may be loss due to the intensity of impulsive noise has to be wide relatively high.



Fig. 1 a) High quality Image b) Artifacts affected image c)Gaussian noise affected image

A compression artifact is a noticeable distortion of media including caused by the application of lossy data compression. Lossy data compression involves discarding some of the media's data so that it becomes simplified enough to be stored within the desired disk space or be transmitted within the bandwidth. Artifacts or artefacts in an image[4][5 ] is a type of distortion cause by the application of lossy data or compression of image at pre processing in transmission of images. The presence of different artifacts can be visualising by loss of edge clarity or visible discontinuities at block boundaries of images. If the compressor could not reproduce enough

data in the compressed version to reproduce the original, the result is a diminishing of quality, or introduction of artifacts.

To overcome mentioned undesirable distortions and to achieve best visual perceptive result by using combination of suitable linear and non linear filters which being applied to regions of documented image for which is suited. In the thesis work, I have proposed a new approach for image quality enhancement. It involves a set of filters are selected to improve the quality of distorted images containing regions with different characteristics.

The Hypothesis Combined Filter (HCF) as a new approach for combining the outputs of distinct filters, each of which is chosen to improve a particular type of content in an image. In our formulation, we define M pixel classes where each pixel class is associated with one of the M image filters for the scheme. During the processing of an image, HCF performs a soft classification of each pixel into the M pixel classes through the use of a locally computed feature vector. After classification, the filter outputs are weighted by the resulting class probabilities and combined to form the final processed output.

At each pixel,a locally computed featured vector which provides prediction of relative performance of the filters.The major contributions of our research include the basic architecture of HCF and a novel probabilistic model which is used to define the pixel classes and to capture the dependence between the feature vector and the defined pixel classes.

Some researches have been proposed different methods for enhancement of image visual quality. Traditional linear filters [6] have the main advantage of simplicity and are backed by a rich theoretical foundation. In the realm of nonlinear filtering, median filters [7], weighted median filters [8], the order-statistic filters [8], and stack filters [9] are examples of the wide range of nonlinear filters that have been studied. More recently, many spatially adaptive methods of filtering have been developed to address the different aspects of image quality. Zhang *et al.* [10] proposed an adaptive version of the bilateral filter [11], [12] for image sharpening and denoising. The behaviour of the filter is locally adaptive based on the response of a Laplacian of Gaussian [13] operator.

Our scheme is taken as a spatially adaptive image enhancement technique. However, instead of applying a particular algorithm adaptively to the local content, our scheme is unique in that allow to apply more than one algorithm for modelling parameter estimation and also it provides a framework for combining any set of filters or whether they are linear or non-linear, iterative or non-iterative.

## II. COMBINED FILTERING METHOD BASED ON HYPOTHESIS

‘Hypothesis’ means we have develop a processed data from some un observed or missing data from the sample. Here HCF provides relative performance of the filters in estimating the corresponding pixel intensity in the original undistorted image. The prediction result then determines the preference of each filter used to obtain the final processed output.

### A. Structure Of Hypothesis Combined Filter (HCF)

In Fig. 2 shows the structure of combined filtering. We have selected a set of image filters each of which is effective in improving the image quality for certain type of content.The filtering process the input image  $x$  using M filters we have selected . Consider the input image is the distorted version of original image. For the n<sup>th</sup> pixel, the filter outputs, denoted by  $z_n = [z_{n,1}, z_{n,2}, \dots, z_{n,M}]^T$ , will each serve as a different intensity estimate of the corresponding pixel in the original image,  $X_n$ . Assume that a judiciously selected feature vector  $y_n$  carries information on how to combine the filter outputs to produce the best estimate of  $X_n$ . We also define M pixel classes corresponding to the M image filters. In the lower portion of Fig HCF computes the feature vector for the current pixel, followed by a soft classification of the current pixel into the M pixel classes. The soft classification computes the conditional probability of the pixel to belong to each one of the pixel classes, given the feature vector. The resulting probabilities are used to weight the filter outputs, which are then summed to form the final processed output. The original pixel intensity  $X_n$ , the feature vector  $y_n$ , and the filter outputs  $Z_n$  are realized from the distorted image as input.

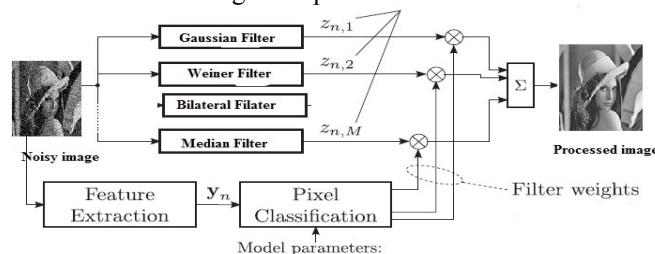


Fig.2 Structure of Hypothesis based Combined Filtering

### **B. Optimal Filtering Selection**

To apply the HCF to a particular application, we only need to specify the set of desired image filters, and a feature vector that is well suited for selecting among the filters. For example, Text regions contain many sharp edges and suffer significantly from ringing artifacts. Picture regions suffer from both blocking artifacts in smooth parts and ringing artifacts around sharp edges. The proposed employs different filters to eliminate these artifacts.

The first filter that select for the HCF is a Gaussian filter with kernel standard deviation  $\sigma = 1$ . choose  $\sigma$  large enough so that the Gaussian filter can effectively eliminate the coding artifacts in the smooth picture regions, even for images encoded at a high compression ratio. If it applied alone, however, the Gaussian filter will lead to significant blurring of image detail.

A bilateral filter [14]is being used as second Filter which operates as each pixel is replaced by the weighted average of its neighbourhood pixel so that it help in smoothening the images while preserving edges. Hence the bilateral filter can moderately remove the coding artifacts without blurring edges and image detail for non smooth picture regions .

To eliminate the ringing artifacts around the text, apply any other two filters in the HCF. A text region typically consists of pixels which concentrate around two intensities, corresponding to the background and the text foreground. Assuming that the local region is a two-intensity text region, the third filter estimates the local foreground gray-level, and the fourth filter estimates the local background gray level by clustering FCM approach.

In this approach requires some methodology for additional filters for combined filtering, among available set of filters ,to improve different content which is distorted differently due to several noises. So provide a Wiener Filter removes the additive noise and inverts the blurring simultaneously. By the process of recovering an image, degraded through some digital acquisition process, with the Wiener Filter. In other words, it minimizes the overall mean square error in the process of inverse filtering and noise smoothing.

### **C. Calculation Of Feature Vector**

For the picture regions, the feature vector also separates the smooth regions from the regions that contain textures and edges. For the text regions, the feature vector detects whether an individual pixel belongs to the background or to the text foreground.

1. For the first feature component, we compute the variance of the JPEG block associated with the current pixel. The block variance is used to detect the smooth regions in the image.
2. For the second feature component, we compute the local gradient magnitude at the current pixel using the Sobel operator S to detect major edges. The Sobel operator consists of a pair of  $3 \times 3$  kernels. Convolving the image with these two kernels produces a 2-dimensional local gradient estimate at each pixel. The second feature component is then computed as the norm of the gradient vector.
3. The third feature component is used to evaluate how well the JPEG block associated with the current pixel can be approximated by using only the local foreground gray-level and the local background gray-level. For the JPEG block  $s$ , suppose  $\{us_i : i = 0, \dots, 63\}$  are the graylevels of the 64 pixels in the block. We define the two-level normalized variance as

$$bs = \sum \frac{\min[u_s - cb(s)]^2, [us - cf(s)]^2}{cb(s) - cf(s)} \quad (1)$$

where where  $cb(s)$  and  $cf(s)$  are the estimates of the local foreground gray-level and the local background gray-level respectively,

4. The fourth feature component is the difference between the pixel's gray-level and the local foreground gray-level  $cf(s)$

$$Yn4=un-cf(s) \quad (2)$$

5. The fifth component is the difference between the pixel's graylevel and the local background gray-level  $cb(s)$ :

$$Yn5=un-cb(s)$$

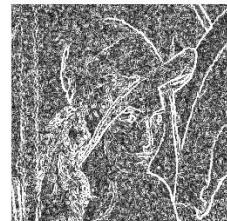


Fig.3 Sobel gradient: edge detection

#### **D. Probabilistic Model**

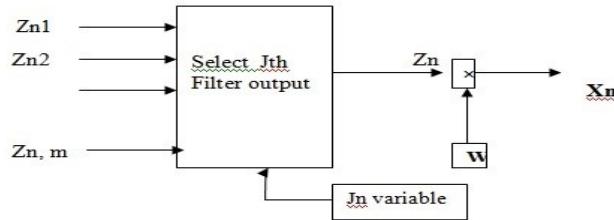


Fig.4 Probabilistic model for estimation of parameters

Our probabilistic model[16] is based on the original pixel intensity  $X_n$ , the feature vector  $Y_n$ , and the filter outputs  $Z_n$ . Based on this model, we derive the output of Combined filtering as the minimum mean square error estimate of  $X_n$ .

Let original pixel intensity  $X_n, Y_n, Z_n$  are the realization of random variables  $x_n, y_n$  and  $z_n$  respectively. When  $Z_n = z_n$ , the conditional distribution of  $X_n$  with M component is given by

$$F(x_n|j_n, z_n) = \sum_{i=0}^n \pi_j (X_n; Z_n, \sigma^2) \quad (4)$$

Where mean of the components of the filter outputs  $Z_n$  is  $\pi_j$  and  $\sigma^2$  are variance of parameters distribution. Also define an unobserved random variable  $J_n$  to estimate the output  $X_n$ . The approach of modeling  $f(z_n/x_n)$  requires a probabilistic model for estimating the processed value  $X_n$  with given  $Z_n$ . We assume that our training samples are of the form  $(x_n, y_n, z_n)$ , where  $x_n$  is the ground truth value of the original undistorted pixel,  $y_n$  is the feature vector extracted from the distorted image, and  $z_n$  are the corresponding filter outputs. Therefore, the absolute error made by each filter,  $(x_n - z_n, j_n)$ , can be computed [17]. Finally we obtain maximum likelihood estimates of these parameters through the expectation-maximization algorithm. The algorithm gives particular weighted values which provides particular preference to the filtered output. And the probabilistic model appropriately adjusts the weighting factors for the image filters

### **III. Expectation Maximization Algorithm**

For classification of pixel contents with the help of several parameters extracted from image data, using advanced clustering algorithms EM algorithm. The idea behind E M algorithm is that, even though do not know variables  $x_1, x_2$  knowledge of distribution is used to determine the estimate of  $p$ . This is done by first estimating the underlying data and use this data to update estimate of the parameter. This is repeated until it converge.

Expectation Maximization algorithm [18],[19] is an algorithm used in the estimation of the parameters of a probability distribution function. Perhaps the most frequently encountered estimation problem because direct access to the data necessary to estimate the parameters is impossible, or some of the data are missing

The EM algorithm consists of two major steps: an expectation step, followed by a maximization step

#### **1. Expectation step(E step)**

Compute the expected value of  $x$  data using the current estimate of parameter and the observed data. The expected value  $x_1$  given measurement  $y_1$  and based on current estimate of the parameter computed as,

$$x_1^{(k+1)} = E[x_1 | y_1, p^{(k)}] \quad (5)$$

#### **2. Maximization step(M step)**

Here the data from expectation step as if it were actually measured data to determine an Maximum likelihood estimate of the parameter. The estimated data sometimes called ‘imputed data’.

Here variables  $x_1^{(k+1)}, x_2^{(k+1)}$  are imputed and  $x_3$  available variable, the maximum likelihood estimate of the parameter is obtained by taking derivative of  $\log(x_1^{(k+1)}x_2^{(k+1)}, x_3/p)$  with respect to  $p$ , equal to zero, and solving  $p$ .

$$0 = \frac{d}{dt} [\log(x_1^{(k+1)}x_2^{(k+1)}, x_3/p)] \quad (6) \quad \text{Eqn.(3.8)}$$

#### **3. Maximum Likelihood Estimate(MLE)**

ML estimation[20] is a means of estimating the parameters of a distribution based upon observed data. Let  $\theta = \theta_1, \theta_2, \theta_3..$  denotes the set of parameters. Let  $x$  be the data observed from the distribution, parameterized by set  $f$  parameters. The key idea in ML estimation is to determine the parameter for  $\theta$  for which probability of observing the outcome  $x=x_1, x_2, x_3\dots$  is as high as possible.[21]

Let the function is  $L_x(\theta/x_1, x_2, x_3..) = f(x/\theta)$  is likelihood function is viewed as a function of parameter  $\theta$  with sample  $x$ . The ML estimates of the parameter is that value of parameters which can maximizes the likelihood function  $\theta_{ML} = \max L_x(\theta)$

#### IV. Fcm Clustering

Clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior. The Fuzzy Logic Toolbox is equipped with some tools that allow you to find clusters in input-output training data. A text region typically consists of pixels which concentrate around two intensities, corresponding to the background and the text foreground. Assuming that the local region is a two-intensity text region, the third filter estimates the local foreground gray-level, and the fourth filter estimates the local background gray-level. For estimating two intensities apply the *FCM* clustering algorithm to partition the 256 pixels of the window into two groups. Then, we use the two cluster medians,  $c_f(s)$  and  $c_b(s)$ , as the local foreground gray-level estimate and the local background graylevel estimate of the JPEG block.

For a text region that satisfies the two-intensity assumption, foreground pixels will then be clipped to the foreground gray-level estimate, and background pixels will be clipped to the background gray-level estimate. In our feature vector, described, a feature component is designed to detect whether the local region is a two-intensity text region. Another two feature components detect whether a pixel is a foreground pixel or a background pixel.

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. this algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center.

$$\sum_{i=0}^n U_{ij} (x_j - c)^2 \quad (7)$$

where  $m$  is any real number greater than 1,  $u_{ij}$  is the degree of membership of  $x_i$  in the cluster  $j$ ,  $x_i$  is the  $i$ th of d-dimensional measured data,  $c_j$  is the d-dimension center of the cluster.

#### V. Enhancement In Contrast

The problem is to enhance the contrast of an image in order to represent all the information in the input image. This can be done using several contrast enhancement techniques eg: several Histogram Equalization techniques. These techniques applied for various application such as remote sensing images and general images which are obtained at the output of combined filtering

To overcome these drawbacks several HE-based techniques are proposed and are more focused on the preservation of image brightness and the improvement of image contrast. So, Bi- Histogram equalization method with brightness preservation and contrast enhancement is applied at the combined filtered images in this approach

In this method, the separation intensity is presented by the input mean brightness value, which is the average intensity of all pixels that construct the input image. After this separation process, these two histograms are independently equalized. By doing this, the mean brightness of the resultant image will lie between the input mean and the middle gray level.

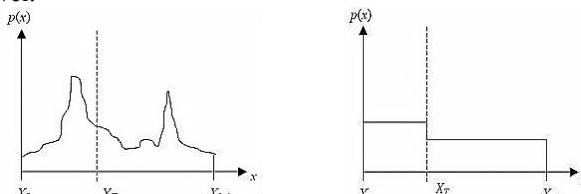


Fig.5 Bi- Histogram equalization (a)Input Histogram Divides into Two sub-histograms (b) Equalized Sub-histograms

#### VI. Quality Measures For Enhanced Image

To study the relative performance of cluster based segmentation methods the following quality measures are calculated.[22]

##### **A.PSNR: Peak Signal To Noise Ratio**

PSNR is the quality measurement between the original image and the reconstructed image which is calculated through the Mean Squared Error (MSE). PSNR is a good measure for comparing removal of distortion result of image. Typical value of image enhanced result is lies between 25 to 35 dB.

$$PSNR = 10 \log_{\frac{255*255}{MSE^2}} \quad (8)$$

#### **B.MSE: Mean Square Error**

The MSE represents the cumulative squared error between the distorted and the original image. In statistics, the mean squared error is one of the most ways to enumerate the variation between values disguised by an estimator and the true values of MSE between Two image is given by

$$\text{MSE} = \frac{1}{N} \sum_{i=0}^n (X_i - Y_i)^2 \quad (9)$$

Where  $i$   $x$  and  $i$   $y$  be the  $i$ th pixel in the reference  $x$  image and the distorted image  $y$ , respectively.

#### **C.SC:Structural content**

It is one of the correlation based measure. It measures the closeness between digital images .It measures the similarity between two images.Hence it is complementary to difference based measure.Structural content is given by:

$$SC = \frac{\sum C(i,j)^2}{\sum Cc(i,j)} \quad (10)$$

Where  $C$ :i,jth pixel of original image

$Cc$ : i,j th pixel of distorted image

#### **D.SSIM:Structural Similarity Index**

We define the structural information in an image as those attributes that represent the structure of objects in the scene, independent of the average luminance and contrast. Since luminance and contrast can vary across a scene, we use the *local* luminance and contrast for our definition.

$$\text{SSIM } S(x; y) = f(l(x; y); c(x; y); s(x; y)) \quad (11)$$

Where  $l(x,y)$  is the luminance comparison of images

$S(x,y)$  is the structural comparison

$C(x,y)$  is the contrast comparison

we consider one of the signals to have perfect quality, then the similarity measure can serve as a quantitative measurement of the quality of the second signal

#### **E. AD:Average Difference**

$$AD = [\sum X(i,j) - \sum X^*(i,j)] / MN \quad (12)$$

This measure shows the average difference between the pixel values, ideally it should be zero.

#### **F. IF:Image Fidelity**

IF parameter also measures the quality between original and distorted image

$$IF = 1 - [\sum X(i,j) - \sum X^*(i,j)] / X(i,j)^2 \quad (13)$$

Where  $X(i,j)$  be the processed image and  $X^*(i,j)$  be the distorted image

## **VII. Result And Discussion**

As a new approach for image quality enhancement, a Combined Filtering based on hypothesis with multiple algorithms has implemented. The performance of the system is evaluated by several types of image with different characteristics, are simulated using MATLAB version R2010a tool. And also the performance of new approach is tabulated based on different quality measures, which is compared with typical filtering methods .

Consider Fig.4.2 which shows the noise affected image Lena.jpg from database, and the processed image have less distortion from noise and artifacts. And Fig.4. 3shows cap.jpg enhancement of compressed image,suffer several ringing artifacts at sharp edges. Processed image should have eliminated from the artifacts around the sharp and smooth regions.Fig.4.5 vatican.jpg shows the distorted image of Vatican city obtained from satellite. By the proposed method the undesired noises and several distortion can be removed without affect information of land view in the satellite image

	PSNR	MSE	SC	SSIM	IF	AD
HCF	33.4	.008	.7964	.995	.951	.406
BF	31.3	.021	1.41	.995	.9291	.46
WF	30.16	.018	1.4	.990	.932	.43
GF	28.3	.023	1.46	.992	.920	.60

Table 1. Comparison of proposed method with common filtering methods based on quality measures



Fig.6 : Lena.jpg :Enhancement result for database image (a) original image (b)noise affected (c) processed image



Fig.7 : cap.jpg: Enhancement result for natural image (a)original image (b)distorted compressed image (c) artifacts removed from image



Fig.8 :Vatican city.jpg: Enhancement result for satellite image (a) satellite image (b) distorted data (c) Enhanced satellite image

The performance of method is evaluated based on several quality measures, and are listed in the table below. The different Quality parameter values for the finally processed image eliminated from distortion are presented graphically at the bottom.

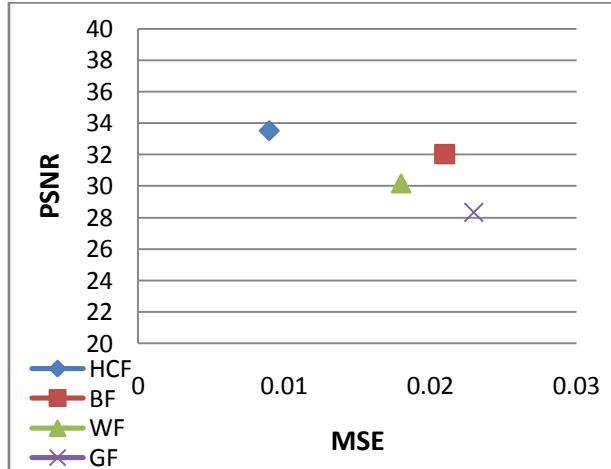


Fig 9:Graphical representation of PSNR vs MSR of combined filtering and comparison to individual filtering for satellite image.

### VIII. Conclusion

In our scheme, we incorporated different image filters for reducing the JPEG artifacts in the different types of image content that are common in document images, like text, graphics, and natural images. Hence the effectiveness of the HCF by it can apply as a post-processing step for JPEG decoder so as to reduce the artifacts due to the compression methods in the decoded document image In this scheme, implemented combined image filters for reducing the different distortions of image. The filtered high quality satellite images have low contrast and brightness values due to transmission, which have improved by Bi histogram equalization method And simulated with different type of images and are compared based on different quality measures.

Some future ideas can be develop with this new approach, By the application of this approach we can improve the visual appearance of video in MPEG FLV format for special application s as an active topic in computer vision.

## References

- [1] (R. Gonzalez and R. Woods, Digital Image Processing, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2007.)
- [2] A. K. Jain, *Fundamentals of Digital Image Processing*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 1989, ch. 9, pp. 347–357.
- [3] S. H. Kim and J. P. Allebach, “Optimal unsharp mask for image sharpening and noise removal,” *J. Electron. Imag.*, vol. 14, no. 2, pp. 23005–23018, May 2005.
- [4] G. K. Wallace, “The JPEG still picture compression standard,” *Commun. ACM*, vol. 34, no. 4, pp. 30–44, Apr. 1991.
- [5] T.-S. Wong, C. A. Bouman, I. Pollak, and Z. Fan, “A document imagemodel and estimation algorithm for optimized JPEG decompression,” *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2518–2535, Nov. 2009
- [6] A. C. Bovik and S. T. Acton, “Basic linear filtering with application to image enhancement,” in *Handbook of Image and Video Processing*. San Francisco, CA: Academic, 2000,
- [7] A. C. Bovik, T. S. Huang, and D. C. Munson, “A generalization of median filtering using linear combinations of order statistics,” *IEEE Trans. Acoust., Speech Signal Process.*, vol. 31
- [8] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, “Weighted median filters: A tutorial,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 43, no. 3
- [9] P. Wendt, E. Coyle, and N. C. Gallagher, “Stack filters,” *IEEE Trans. Acoust., Speech Signal Process.*, vol. 34, no. 4, pp. 898–911, Aug. 1988
- [10] B. Zhang and J. P. Allebach, “Adaptive bilateral filter for sharpness enhancement and noise removal,” *IEEE Trans. Image Process.*, vol. 17, no. 5, pp. 664–678, May 2000
- [11] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and colorimages,” in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.
- [12] 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.
- [13] P. A. Mlsna and J. J. Rodríguez, “Gradient and Laplacian-type edgedetection,” in *Handbook of Image aVideo Processing*. San Francisco, CA: Academic, 2000, pp. 415–431.
- [14] Medical Image Denoising Using Bilateral Filter Devanand Bhonsle1 I.J. Image, Graphics and Signal Processing, 2012, 6, 36-43 Published Online July 2012 MECS.
- [15] Anisotropic Diffusion in Image Processing Joachim Weickert Department of Computer Science University of Copenhagen Copenhagen, Denmark
- [16] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer-Verlag, 2009, ch. 6, pp. 214–215.
- [17] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, “An introduction to variational methods for graphical models,” *Mach. Learn.*, vol. 37, pp. 183–233, Nov. 1999
- [18] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *J. Royal Stat. Soc., Ser. B, Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.
- [19] C. Wu, “On the convergence properties of the EM algorithm,” *Ann. Stat.*, vol. 11, no. 1, pp. 95–103, Mar. 1983.
- [20] R. A. Boyles, “On the convergence of the EM algorithm,” *J. Royal Stat. Soc., Ser. B, Methodol.*, vol. 45, no. 1, pp. 47–50, 1983.
- [21] J. Rissanen, “A universal prior for integers and estimation by minimum description length,” *Ann. Stat.*, vol. 11, no. 2, pp. 416–431, Jun. 1983.
- [22] Image Quality Assessment: From Error Visibility to Structural Similarity Zhou Wang, Member, IEEE, Alan C. Bovik, Fellow, IEEE Hamid R. Sheikh, Student Member, IEEE, and Eero P. Simoncelli, Senior Member, IEEE