

Statistical Moments based Noise Classification using Feed Forward Back Propagation Neural Network

Shamik Tiwari
MITS Deemed University
Lakshmangarh, Sikar
Rajasthan, India

Ajay Kumar Singh
MITS Deemed University
Lakshmangarh, Sikar
Rajasthan, India

V.P. Shukla
MITS Deemed University
Lakshmangarh, Sikar
Rajasthan, India

ABSTRACT

A neural network classification based noise identification method is presented by isolating some representative noise samples, and extracting their statistical features for noise type identification. The isolation of representative noise samples is achieved using prevalent used image filters whereas noise identification is performed using statistical moments features based classification system. The results of the experiments using this method show better identification of noise than those suggested in the recent works.

General Terms

Image denoising, Pattern recognition.

Keywords

Noise models, moments, back propagation neural network.

1. INTRODUCTION

An image is often corrupted by noise in its acquisition or transmission. Noise is any undesired information that degrades the image and appears in images from a variety of sources. The goal of denoising is to remove the noise while retaining the important signal features to the extent possible, and can be done through filtering of different types. Denoising can be done through filtering, which can be of different types. When one has no prior knowledge of the nature of the noise, it is necessary to identify the nature of noise for filtering. The majority of filtering and contour detection algorithms [1-10] assume that the nature of the noise is known. In this work, we are interested in identifying the nature of the noise from the observed image so that we can apply the filtering or analysis algorithm.

A simple pattern classification based automatic noise identification method is proposed by Chen and Das [12]. These methods are based on the criterion of maximum likelihood for the selection of the most homogeneous masks as proposed by Lee[7][8] and Nagao & Matsuyamat [9], from which the values of the local standard deviations are calculated. Identification of the nature of noise and estimation of its statistical parameters by analysis of local histograms was suggested by Beaufrepaire et al. [13]. Their suggested method has shown various advantages as it not requiring a priori masks for the detection of homogeneous regions etc. Vozel et al. [14] suggested an unsupervised variational classification through a multi-thresholding method. Their work is effective for the determination of homogeneous regions in an image, and is a significant improvement in field of noise identification and estimation. Santhanam and Radhika [15] employed a novel approach to classify noises by using three

additive noises namely Salt & Pepper, Non Gaussian white & Gaussian white for comparative study of classification using Back Propagation Network and Multi Layer Perceptron. In this research, we extend the previous works with additive and multiplicative noises both. Our work shows better classification accuracy than the aforesaid methods.

2. NOISE MODELS

Basically, there are three standard noise models [19], which model the types of noise encountered in most images, they are additive noise, multiplicative noise and impulse noise.

An image function is given by $f(x, y)$ where (x, y) is spatial coordinate and f is intensity at point (x, y) . Let $f(x, y)$ be the original image, $g(x, y)$ be the noisy version and $\eta(x, y)$ be the noise function, which returns random values coming from an arbitrary distribution. Then the additive noise is given by the equation

$$g(x, y) = f(x, y) + \eta(x, y)$$

Additive noise is independent of the pixel values in the original image. Typically $\eta(x, y)$ is symmetric about zero. This has the effect of not altering the average brightness of the image. Additive noise is the good model for the thermal noise within photoelectric sensors.

Multiplicative noise or speckle noise is signal dependent form of noise whose magnitude is related to the value of the original pixel. The simple mathematical expression for a multiplicative noise model is given by

$$g(x, y) = f(x, y) + \eta(x, y).f(x, y)$$

or

$$g(x, y) = f(x, y).[1 + \eta(x, y)]$$

Third noise model is impulse noise, which has the property of either leaving a pixel unmodified with probability $(1-P)$ or replacing it altogether with a probability P . Restricting $\eta(x, y)$ to produce only the extreme intensities 0 or 1. The source of impulse noise is usually the result of an error in transmission or an atmospheric disturbance. Although in many cases noise is dependent upon physical properties of device, media, etc, people often consider noise as random variables, characterized by a probability density function (PDF). The following are some noise models and their PDFs which are commonly used in imaging processing. In this work, we consider four different

types of commonly occurring image noises, namely, uniform white, Gaussian white, salt-and-pepper and speckle noise.

Noise	PDF $p(z)$	Mean μ	Variance σ^2
Uniform Noise	$\begin{cases} \frac{1}{b-a} & \text{if } a \leq z \leq b \\ 0 & \text{otherwise} \end{cases}$	$\frac{1}{2}(a+b)$	$\frac{(b-a)^2}{12}$
Gaussian Noise	$\frac{1}{\sqrt{2\pi}b} e^{-z^2/2b^2}$ for $-\infty < z < \infty$	a	b^2
Impulse Noise	$\begin{cases} P_a & \text{for } z=a \\ P_b & \text{for } z=b \\ 0 & \text{otherwise} \end{cases}$	$aP_a + bP_b$	$(a-\mu)^2 P_a + (b-\mu)^2 P_b$
Speckle Noise	$p_z = \begin{cases} ae^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases}$	$1/a$	$1/a^2$

Table 1: PDF, Mean and Variance for selected noises

Among these four types, speckle noise is of multiplicative type, whereas the other three are additive in nature. Probability density function, mean and variance of the above said noises are shown in the table 1.

2.1 Uniform noise

Uniform noise is not often encountered in real-world imaging systems, but provides a useful comparison with Gaussian noise. The linear average is a comparatively poor estimator for the mean of a uniform distribution. This implies that nonlinear filters should be better at removing uniform noise than Gaussian noise. Figure 1 shows the histogram of uniform noise pattern.

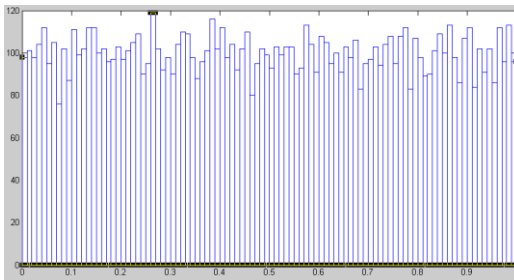


Fig. 1: Uniform noise pattern

2.2 Gaussian noise

The Gaussian distribution has an important property that in order to estimate the mean of a stationary Gaussian random variable, one cannot do any better than the linear average. This

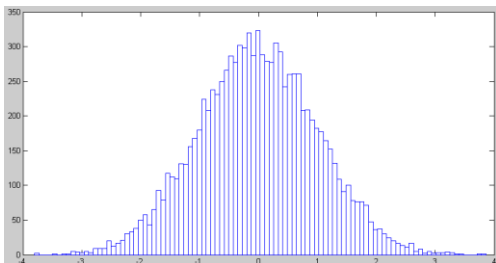


Fig 2: Gaussian noise pattern

makes Gaussian noise a worst-case scenario for nonlinear image restoration filters, in the sense that the improvement over linear filters is least for Gaussian noise. To improve on linear filtering results, nonlinear filters can exploit only the non-Gaussianity of the signal distribution. Figure 2 shows the histogram of Gaussian noise pattern with 10000 random numbers.

2.3 Impulse noise

Impulsive noise is sometimes called salt-and-pepper noise or spike noise. An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. This type of noise can be caused by dead pixels, analog-to-digital converter errors, bit errors in transmission etc. Figure 3 shows the histogram of impulse noise pattern with equal probabilities.

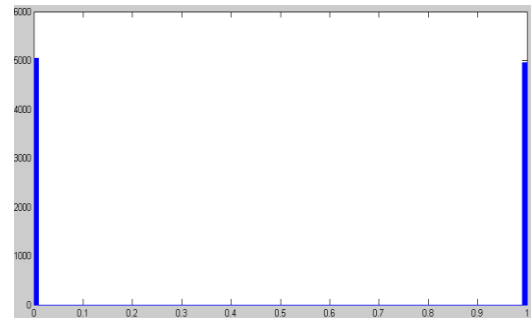


Fig 3: Salt & Pepper noise pattern

2.4 Speckle Noise

Another common form of noise is data dropout noise generally referred to as speckle noise. This noise is, in fact, caused by errors in data transmission. The corrupted pixels are either set to the maximum value, which is something like a snow in image or have single bits flipped over. Figure 4 shows the histogram of speckle noise pattern.

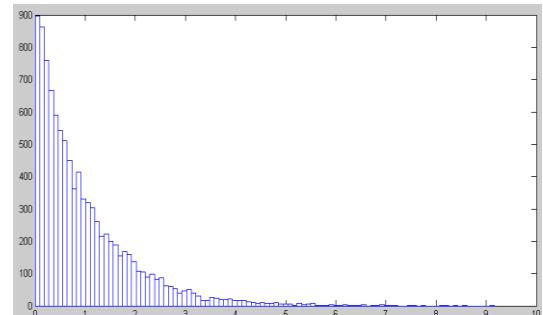


Fig 4: Speckle noise pattern

3. FEATURE SET

Statistical functions such as mean, median, standard deviation and moments are most common to characterize data set, which have been used as pattern features in many applications [16][17][18]. Statistical moments up to order 5 have been used as feature set in this work. The mean and the variance provide information on the location and variability (spread, dispersion)

of a set of numbers, and by doing so, provide some information on the appearance of the distribution (for example, as shown by the histogram) of the numbers. The mean and variance are the first two statistical moments, and the third and fourth moments also provide information on the shape of the distribution. Let z_i be a discrete random variable that denotes intensity levels in a image and let $p(z_i)$, $i = 0, 1, \dots, L-1$, be the corresponding normalized histogram, where L is the number of possible intensity values. A histogram component $p(z_j)$, is an estimate of the probability of occurrence of discrete intensity values z_j and the histogram may be viewed as an approximation of the probability density function. Thus the central moment function is defined by the equation:

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

where, n is the moment order .

Feature(1): Mean is the standardized first central moment of the probability distribution:

$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$

Feature(2): Standard deviation is the standardized second central moment of the probability distribution:

$$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2}$$

Feature(3): Skewness measures the asymmetry of a distribution. It is positive if there is a long tail to the right and negative for a tail on the left.

$$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$$

Feature(4): Kurtosis measures the sharpness of the distribution relative to a Gaussian distribution. Negative kurtosis implies that the peak is broader than a Gaussian. Positive kurtosis means that the distribution is sharper than a Gaussian.

$$\mu_4 = \sum_{i=0}^{L-1} (z_i - m)^4 p(z_i)$$

Feature(5): Central moment of fifth order is given by

$$\mu_5 = \sum_{i=0}^{L-1} (z_i - m)^5 p(z_i)$$

4. BACK PROPOGATION FEED FORWARD NEURAL NETWORK (BPN)

A successful pattern classification methodology [20] depends heavily on the particular choice of the features used by the classifier .The Back-Propagation is the best known and widely used learning algorithm in training multilayer feed forward neural networks. The feed forward neural net refer to the network consisting of a set of sensory units (source nodes) that

constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the network in a forward direction, from left to right and on a layer-by-layer basis. Back propagation is a multi-layer feed forward, supervised learning network based on gradient descent learning rule. This BPNN provides a computationally efficient method for changing the weights in feed forward network, with differentiable activation function units, to learn a training set of input-output data. Being a gradient descent method it minimizes the total squared error of the output computed by the net. The aim is to train the network to achieve a balance between the ability to respond correctly to the input patterns that are used for training and the ability to provide good response to the input that are similar. A typical back propagation network of input layer, one hidden layer and output layer is shown in figure5.

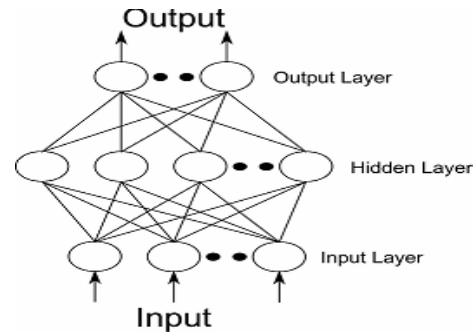


Fig 5: Feed Forward BPN

The steps in the BPN algorithm are:

Step 1: Initialize the weights.

Step 2: While stopping condition is false, execute step 3 to 10.

Step 3: For each training pair (x, t) perform steps 4 to 9.

Step 4: Each input unit X_i , $i=1, 2, \dots, n$ receives the input signal, x_i and broadcasts it to the next layer.

Step 5: For each hidden layer neuron denoted as Z_j , $j=1, 2, \dots, p$.

$$z_{inj} = v_{oj} + \sum_i x_i v_{ij}$$

$$z_j = f(z_{inj})$$

where v_{oj} is the bias on j^{th} hidden

Broadcast Z_j to the next layer. unit.

Step 6: For each output neuron Y_k , $k=1, 2, \dots, m$

$$y_{ink} = w_{ok} + \sum_j z_j w_{jk}$$

$$y_k = f(y_{ink})$$

Step 7: Compute δ_k for each output neuron, y_k

$$\delta_k = (t_k - y_k) f'(y_{ink})$$

$$\Delta w_{jk} = \alpha \delta_k z_j$$

$$\Delta w_{ok} = \alpha \delta_k \quad \text{since } z_0 = 1$$

where δ_k is the portion of error correction weight adjustment for w_{jk} i.e. due to an error at the output unit y_k , which is back propagated to the hidden unit that feed it into the unit y_k , and α is learning rate.

Step 8: For each hidden neuron

$$\delta_{inj} = \sum_{k=1}^m \delta_k w_{jk} \quad j = 1, 2, \dots, p$$

$$\delta_j = \delta_{inj} f'(z_{inj})$$

$$\Delta v_{ij} = \alpha \delta_j x_i$$

$$\Delta v_{oj} = \alpha \delta_j$$

where δ_j is the portion of error correction weight adjustment for v_{ij} i.e due to the back propagation of error to the hidden unit z_j

Step 9: Update weights.

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk}$$

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{ij}$$

Step 10: Test for stopping condition.

4. METHODOLOGY

The steps of the algorithm to classify noise are detailed in figure 6. These are seven major steps: image acquisition, preprocessing of images, noise inclusion, filtering, noise pattern extraction, feature extraction and classification. In the initial steps we introduced noises namely uniform, Gaussian, impulse and speckle to the preprocessed images. Then we filtered the noisy images using two commonly used filters, average filter and median filter. To get the noise patterns, filtered images have been subtracted from the noisy images in case of additive noises and divided in case of multiplicative noise. Once the noise patterns are acquired, the method demands the extraction of features. So, the statistical features as the moments up to fifth order have been calculated in the next step to prepare the training and testing database. Finally, the training and testing performed quite well with this feature database using feed forward back propagation neural network.

5. EXPERIMENTAL RESULTS

The general-purpose image database has been used to test the proposed technique (<http://wang.ist.psu.edu>). The database contains about 10,000 images of multiple categories. Then, the four different types of noises were introduced with intensity of 0.02, making the multiple databases of 1600 images (i.e., 400 images with uniform, 400 images with Gaussian noise, 400 images with impulse noise and the rest 400 with speckle noise).

All experiments were carried out using Matlab 7.0 simulations. Matlab function 'imnoise' is used to generate noises.

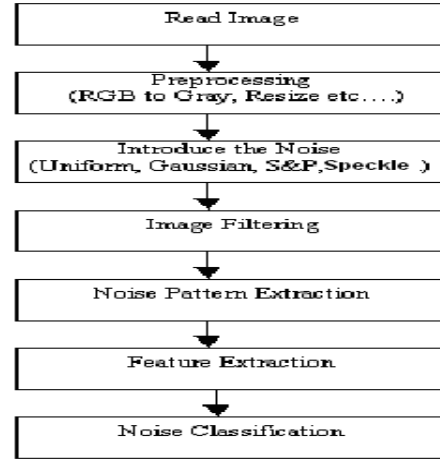


Fig 6. Block diagram of noise classification system

The classification is performed over a 5-10-4 feed forward neural network model that consist of one input layer with five neurons (selected for the five statistical moments), one hidden layer with ten neurons and one output layer with four neurons (for uniform, Gaussian, salt & pepper and speckle classes). Back propagation is pertained as network training principle where the training dataset is constructed by the extracted moment features of the noise patterns. The entire input features are normalized into the range of [0, 1], whereas the output class is assigned to one for the highest probability and zero for the lowest.

The various parameters for the neural classifier training for all the patterns are given in Table 2.

Parameters	
Learning Rate	0.01
Performance Goal	0.01
No. of Epochs taken to meet the performance goal	4000
Time taken to learn	32.28 Secs

Table 2: Parameters for neural classifier

The Performance graph of the neural classifier is shown in figure 7.

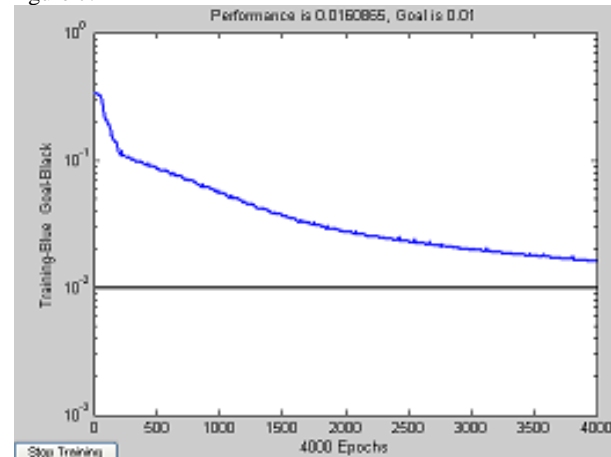


Fig 7: Performance graph neural classifier

The confusion matrices given in Table 3 give the percentage of images classified correctly by the BPN. Testing result shows 96 percent, 95 percent, 90 percent and 98 percent accuracy for uniform, Gaussian, impulse and speckle noises respectively. Result shows there are few chances to classify a noise in another class. For example, approximately 8 percent chances to recognize impulse noise as speckle noise, 4 percent chances to recognize uniform noise as impulse noise, 2 percent chances to recognize impulse noise as Gaussian noise.

	Uniform	Gaussian	Impulse	Speckle
Uniform	96.2	0	3.8	0
Gaussian	1.4	95	3	0.6
Impulse	0	2.3	90.19	7.5
Speckle	0.5	0	1.5	98

Table 3: Confusion matrix-performance analysis of BPN

6. CONCLUSION

In general the purpose of detection and classification of the noise has been achieved by using feed forward back propagation neural network. A maximum accuracy of 98 percent was found for speckle noise followed by accuracy of 96 percent, 95 percent and 90 percent for those of uniform, Gaussian and impulse noises respectively. This work shows that it can help in choosing the appropriate filter for image denoising.

7. REFERENCES

- [1] J.S LEE. March 1980, "Digital image enhancement and noise filtering by use of local statistics," IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2, vol 11: pages 165-168.
- [2] D.T. KUAN, , March 1985, "Adaptive noise smoothing filter for images with signal dependent noise," IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-7 vol 11: pages 165-177.
- [3] J. CANNY , 1986, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine intelligence, PAMI-7 vol 11: pages 679-698.
- [4] R. DERICHE, 1987, "Using Canny's criteria to derive an optimal edge detector recursively implemented," International Journal on Computer Vision, vol 1: pages 167-187.
- [5] D. KUNDUR and D. HATZINAKOS, May 1996, "Blind image deconvolution," IEEE Signal Processing Magazine, 13(3): pages 43-64.
- [6] M. SABRI - K. CHEHDI, June 1990 Lugano Suisse, "Likelihood decision rule for edge preserving smoothing of noisy pictures", IASTED International Conference on Signal Processing and Filtering, pp. 5-6.
- [7] J.S. LEE, 1983. "Digital image smoothing and the sigma filter", Computer Graphics and Image Processing, no 24, pp. 255-269.
- [8] J.S. LEE, 1981, "Speckle analysis and smoothing of synthetic aperture radar images", Computer Graphics and Image Processing, Vol no. 17, pp.24-32.
- [9] M. NAGAO, T. MATSUYAMAT, 1979, "Edge preserving smoothing", Computer Graphics and Image Processing no 9, pp. 394-407.
- [10] P.F. YAN and C.H. CHEN, 1986, "An algorithm for filtering multiplicative noise in wide range", Revue Traitement du Signal, vol. 3, no 2, pp. 91-96.
- [11] M. SABRI and K. CHEHDI , August 1990 Australia, "Automatic image noise identification" International Symposium on Signal Processing and its Application, pp. 612-615.
- [12] Chen, Y. and M. Das, 2007, An automated technique for image noise identification using a simple pattern classification approach. Proceedings of MWSCAS 07, IEEE Computer Society, USA., pp:819-822.
- [13] L. Beaulieu, K. Chehdi, and B. Vozel," Identification of the nature of noise and estimation of its statistical parameters by analysis of local histograms", Proc. 1997 IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 4,pp. 2805-2808.
- [14] Vozel, B., K. Chehdi, L. Klaine, V.V. Lukin and S.K. Abramov, 2006. "Noise identification and estimation of its statistical parameters by using unsupervised variational classification", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 14-19, IEEE Xplore Press, Toulouse, pp: 841-844.
- [15] T. Santhanam and S. Radhika, "A novel approach to classify noises in images using artificial neural network", Journal of Computer Science 6 (5): pp. 541-545, 2010.
- [16] Devendran V et. al., "Texture based Scene Categorization using Artificial Neural Networks and Support Vector Machines: A Comparative Study," ICGST-GVIP, Vol. 8, Issue IV, pp. 45-52, December 2008.
- [17] Said E. E. et al., , 2000, "Neural Network Face Recognition Using Statistical Feature Extraction," 17th National Radio Science Conference. Minufiya University, Egypt, C31, pp. 1-8.
- [18] Li L. et al., 2004, "Statistical modeling of complex backgrounds for foreground object detection," IEEE Trans. Image Proc., 13(11), pp. 1459-1472.
- [19] Gonzalez, R. and R. Woods, 2002. Digital Image Processing. 3rd Edn., Prentice Hall Publications, ISBN: 9788120336407, pp: 50-51.
- [20] Kumar, S., 2004. Neural Networks: A Classroom Approach. 1st Edn., Tata Mc-Graw Hill Publications, ISBN: 0-07-048292-6, pp: 169.