International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 3 Issue: 10
ISSN: 2454-4248
110 – 115

Recital Comparison of Bilingual Language Using Various Filters for Offline Handwritten Character

D. Shashi Kala
Research Scholar, Department of Computer Science
Mother Teresa Women's University, Kodaikanal
shashi.arul@yahoo.co.in

Dr. M. P. Indra Gandhi
Department of Computer Science,
Mother Teresa Women's University, Kodaikanal
mpyazhini@gmail.com

Abstract: Optical Character Recognition (OCR) of multilingual document containing Offline Handwritten Character (OHC) in regional languages of India, it is necessary to identify different script forms before running an individual OCR of the scripts. In this paper, novel approaches for offline character recognition are written in south Indian languages such as Tamil and Kannada. Preprocessing is one of the most important phases in OCR development. It directly affects the efficiency of any OCR. In this process an extracting of basic constituent symbols of the script. Different methodologies which are growing rapidly in the area of character recognition is South Indian Languages. In this paper, it is mainly focused on the existing methodology used in different stages of OCR to recognize offline handwritten character of bilingual regional languages of South India such as Tamil and Kannada are reviewed, summarized and documented.

Keywords: *OHC, Preprocessing, Salt and Pepper Noise, Gaussian Noise, Median Filter, Average Filter and Wiener Filter*

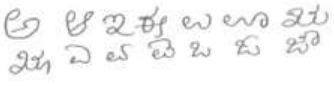
*****

I. INTRODUCTION

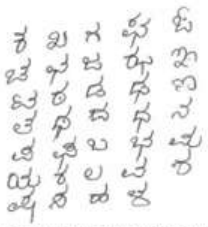
Offline Handwritten Character Recognition, generally condensed as OHCR, it is the method of converting handwritten text into machine processing form. OHCR is an active research area in the field of pattern recognition. Two types of Handwritten Character Recognition techniques are there, Online HCR, where in data are collected during the writing process with the help of a special pen and an electronic surface (tablet) and offline HCR where in data are scanned images of prewritten text, generally on a sheet of paper. Offline HCR is significantly different from Online HCR [1, 2]. In this proposed work, a method for offline handwritten charactersregional languages of South India is considered using Optical Character Recognition. A typical offline handwritten character recognition system involves the subsequent steps: Data Collection, Preprocessing, Segmentation, Feature Extraction, Classification and Post Processing.

II. LITERATURE SURVEY

This section gives a review of literature on the origin of handwritten script, Most of the Indian script are derived from the ancient Brahmi script, Since a majority of these scripts are mainly common in the Indian sub-continent they are also called Indic scripts. The South Indian bilingual script Tamil and Kannada largely share the same basic character set consisting of vowels and consonant. Kannada is the official language of state Karnataka. It is derived from the southern Bramhilipi. There are 15 vowels (swaras) and 34 consonants (vyanjanas) in Kannada alphabet. Tamil is a Dravidian language predominantly spoken by the people of Tamilnadu state. There are 12 vowels, 18 consonants and one special character (aayuthaezhuthu) are given below

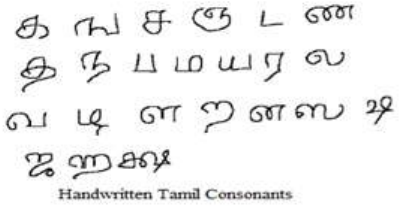


Handwritten Kannada Vowels



Handwritten Kannada Consonants

Handwritten Tamil Vowels



Handwritten Tamil Consonants

Data collected from a writer

 ISSN: 2454-4248
110 – 115

III. DATA COLLECTION

The emblematic database for South Indian character is both available freely or commercially, hence it have been collected as the sample of Tamil and Kannada handwritten character data from schools, colleges, and commercial sectors without considering ink or pen variations [6]. Three different writes (W1, W2, and W3) are asked to write the character in a paper and scanned the same is given in Figure. 1, the images are bilingual in nature and reflect on our experiment each character is segmented morphologically stored either as BMP, JPG or TIF format of pixel size65 x 65.

.
Figure 1: Tamil and Kannada Handwritten Characters

IV. PREPROCESSING

In the process of image analysis, preprocessing is the first phases during the course image are embellished, which is used for further processing called preprocessing. Preprocessing is the vital step in handwriting recognition system. In general, it is assumed that there is no skew and noise in the image, which make the image unfit for segmentation and feature extraction. Such features, leading to high classification error rates. In order to reduce these factors, the images must undergo a significant amount of preprocessing. Preprocessing is not the single step rather it contains sequences of steps [7]. The raw image is subjected to a number of preprocessing stages are explained.

4.1 Skew detection and correction: Skewness correction refers to the tilt in the bitmapped image of the scanned paper for character recognition system. It is usually caused if the paper is not fed straight into the scanner. Most of the character recognition algorithms are sensitive to the orientation (or skew) of the input document image, making it necessary to develop algorithms which can detect and correct the skew automatically

4.2 Skeletonization: Skeletonization is also called thinning. Skeletonization refers to the process of reducing the width of a line like object from many pixels wide to just single pixel. This process can remove irregularities in letters and in turn, makes the recognition algorithm simpler because they only have to operate on a character stroke, which is only one pixel wide. It also reduces the memory space required for storing the information about the input characters and no doubt, this process reduces the processing time too.

4.3 Noise Removal: The scanned image may contain some unwanted data which may lead to color charges, the shape or size of objects in the image and blurring of edges or dilution of details in the image it is referred as noise [6]. The noise may be inbuilt within the image or may be due to poorly photocopied of the scanning [3]. This unwanted information should be removed from the image. The major objective of noise removal is to remove any unwanted bit-patterns, which do not have any significance in the output. There are three common types of noise:

- ***Amplifier Noise (Gaussian Noise)***
- ***Salt-and-Pepper Noise (Spike Noise)***
- ***Speckle Noise (Multiplicative Noise)***

4.3.1 Amplifier Noise (Gaussian Noise): Gaussian Noise or Normal Noise is synonym of amplifier noise. It is a basic noise used in information theory. This type of noise is caused by random fluctuations in the signal. In color cameras where more amplification is used in the blue color channel than in the green or red channel, there can be more noise in the blue channel. Amplifier noise is a major part of the noise of an image sensor, that is, of the constant noise level in dark areas of the *image*. In Gaussian noise, each pixel in the image will be changed from its original value by a (usually) small amount. Meaning that each pixel in the noisy image is the sum of the true pixel value and a random, Gaussian distributed noise value. The noise is independent of intensity of pixel value at each point. Gaussian random distribution noise can be expressed by

$$p_G(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Where, z represents the grey level, μ the mean value and σ the standard deviation.

4.3.2 Salt and Pepper Noise (Spike Noise) Salt and pepper noise is sometimes called spike noise or random noise or independent noise. In salt and pepper noise (sparse light and dark disturbances), pixels in the image are very different in color or intensity unlike their surrounding pixels. Salt and pepper degradation can be caused by sharp and unexpected disturbance in the image signal. Generally this type of noise will only affect a small number of image pixels. After viewing the image it contains dark and white. Hence the term is said to be salt and pepper noise [5]. Salt & pepper distribution noise can be expressed by

 ISSN: 2454-4248
110 – 115

$$P(X)= \begin{cases} p1, & X=A \\ p2, & X=B \\ 0, & otherwise \end{cases}$$

Where, *p1*, *p2* are the Probabilities Density Function (PDF), *p(x)* is distribution salt and pepper noise in image and *A, B* are the arrays size image. Gaussian, salt & Pepper are called impulsive noise. In this paper it is mainly focused on*salt and pepper noise and Gaussian noise.*

V. FILTERING TECHNIQUES

Various techniques are used for removal of different types of noises based on the properties of the noise. Image filtering (noise removal) is not only used to improve image quality but also is used as a preprocessing stage in many application including image encoding, pattern recognition, image compression etc. Efficiency of noise removal algorithm depends on noise detection and noise replacement. With various techniques only the *Median filter, Average filter and Wiener filter* are considered here, Salt and Pepper noise and Gaussian noise are applied to the image using the above mentioned filters

5.1 Median Filter: The median filter is non-linear filter. It removes noise successfully with preserve sharp edges. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve noise. In this method it simply replaces each pixel value by the median of the intensity level in the neighborhood of the pixel [1]. It proves to be best in removing salt and pepper noise.

5.2 Average Filter:(mean filter): It performs smoothing of images (i.e. reducing variation of intensity between one pixel and the next). Average filter replaces each pixel by the average of pixel in a square window surrounding these pixels. Larger window can remove noise more effectively, but also blur image.

5.3 Wiener Filter: It provides linear estimation of a desired signal sequence from another related sequence [4]. Wiener filter provide solution for stationary signals in finding signal estimation problems. It provides successful results in removing noise from images .Wiener filter is based on statistical approach.

VI. PARAMETRIC DESCRIPTION

The performance parameters are most important criteria to justify the simulation results. Peak Signal Noise Ratio (PSNR), Means Square Error (MSE) and Root Means Square Error (RMSE) are considered parameters, the quality of denoised image is measured by[4]:

$$MSE = \frac{1}{M*N}\sum_{i=1}^{M} \quad \sum_{j=1}^{N}[x(i,j) - y(i,j)]^2$$

Where, x(i, j) is original image and y(i, j) is de-noised image.

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right)$$

Where R is maximum value of pixel present in an image, MSE between original and de-noised image with M*N size [4].

$$RMSE = sqrt(MSE)$$

VII. ALGORITHM

Step 1: Initialize an image and convert it into gray scale image

Step 2: Make it noisy by applying salt and pepper noise & Gaussian noise

Step 3: Noisy image is then converted into de-noisy image by applying different types of filters.

Step 4: Calculate PSNR and RMSE to check the performance of filter and de-noised image.

Step 5: De-noised image having high PSNR and low RMSE.

 ISSN: 2454-4248
110 – 115

VIII. Experimental Analysis

The *Experimental Analysis*ofthe different filtering techniques were implemented using Matlab– 12on PC equipped with 2.0 GHz and 4 GB of (RAM) are given below:

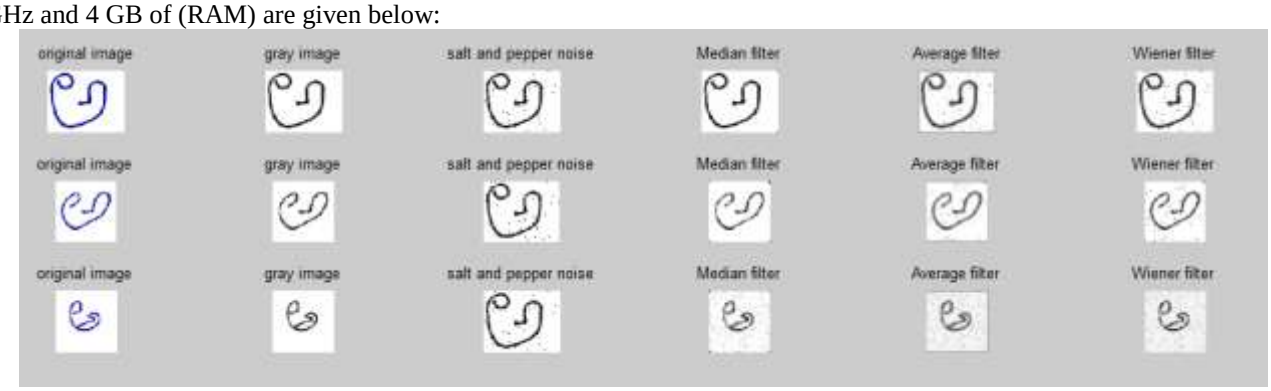


Figure 2: Salt and pepper noise de noised by filters for Kannada character

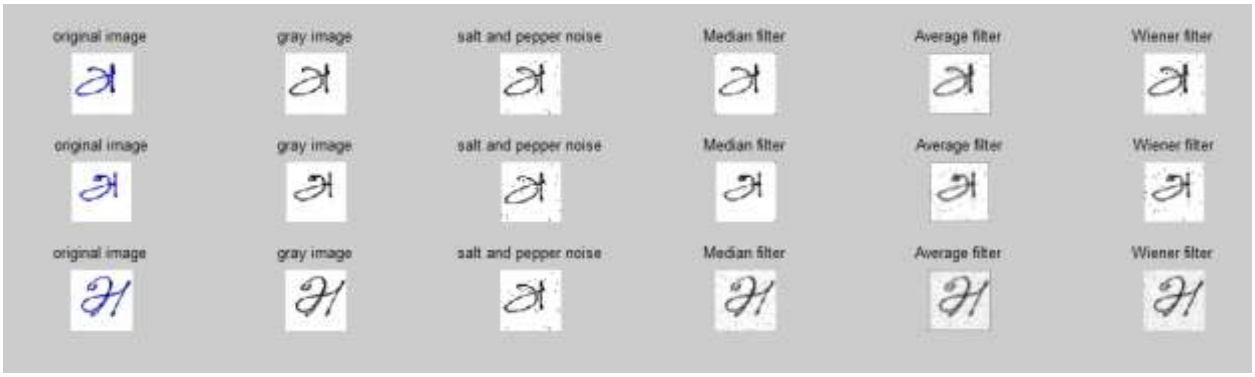


Figure 3: Salt and pepper noise de noised by filters for Tamil character

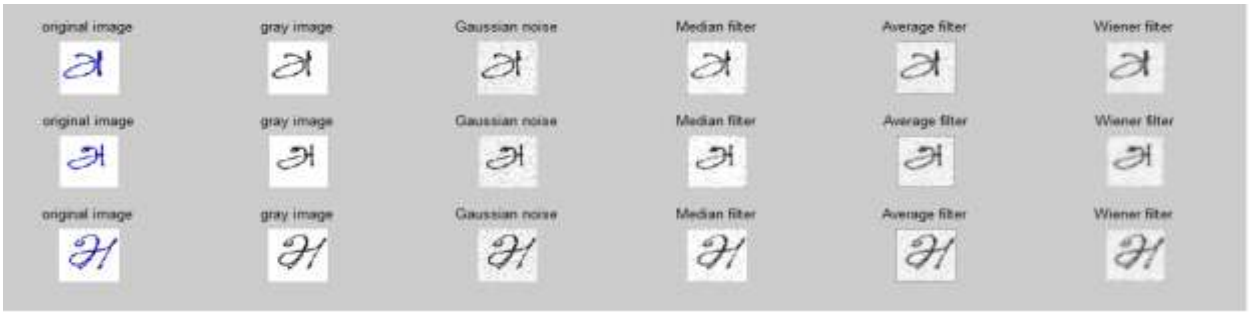


Figure 4:Gaussian noise de noised by filters for Tamil character

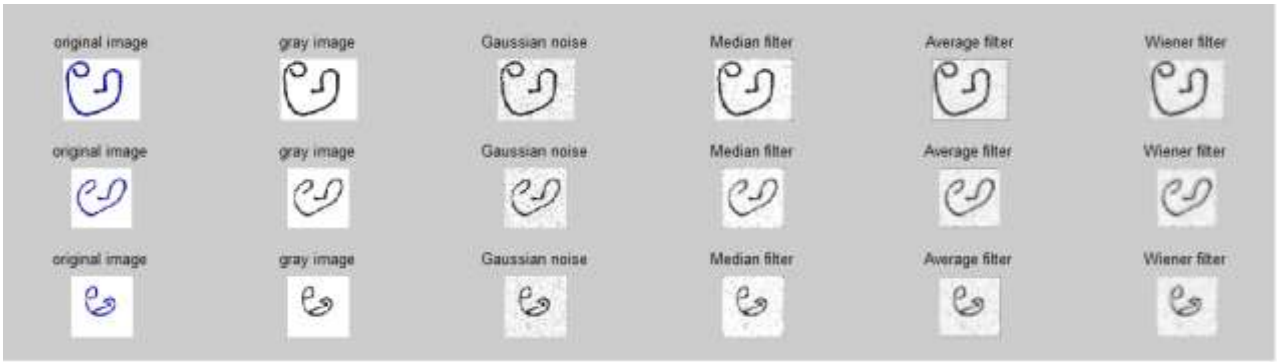


Figure 5: Gaussian noise de noised by filters for Kannada character

 ISSN: 2454-4248
110 – 115

The following given image is tested with salt and pepper noise model and de-noised with filters are shown for comparison in Table 1 and Table 2.

Table 1: PSNR and RSME Value for salt and pepper noise using different filter techniques in Kannada

Filter Techniques	W1		W2		W3	
	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE
Median Filter	42.689	1.8711	45.4698	1.3585	42.7793	1.8517
Average Filter	30.1987	7.8814	30.3066	7.7841	31.0267	7.1648
Wiener Filter	31.0928	7.1105	31.2264	7.0019	33.1318	5.6228

Table 2: PSNR and RSME Value for salt and pepper noise using different filter techniques in Tamil

Filter Techniques	W1		W2		W3	
	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE
Median Filter	41.8542	2.0598	41.6125	2.1179	40.2873	2.4670
Average Filter	30.746	7.4001	30.7016	7.43810	29.9979	8.0658
Wiener Filter	32.5170	6.03520	32.4708	6.0674	31.5264	6.7642

The following given image is tested with Gaussian noise model and de-noised with filters is shown for comparison in Table 3 and 4.

Table 3: PSNR and RSME Value for Gaussian noise using different filter techniques in Kannada

Filter Techniques	W1		W2		W3	
	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE
Median Filter	27.5972	10.6336	27.3475	10.943	27.5300	10.71620
Average Filter	24.7494	14.7595	24.5587	15.0870	24.4891	15.2085
Wiener Filter	24.7144	14.8190	24.5856	15.04050	24.4377	15.2988

Table 4: PSNR and RSME Value for Gaussian noise using different filter techniques in Tamil

Filter Techniques	W1		W2		W3	
	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE
Median Filter	29.3700	8.6704	29.0575	8.9881	29.5100	8.5318
Average Filter	26.0073	12.7695	25.8578	12.9913	26.0927	12.6445
Wiener Filter	26.0765	12.6683	25.9064	12.9187	25.9803	12.8092

Comparison of PSNR and RMSE for Salt and pepper noise de-noised by filters for Tamil and Kannada Languages using three different writers are shown inFigure 3, Figure 4 and Figure 5

 ISSN: 2454-4248
110 – 115

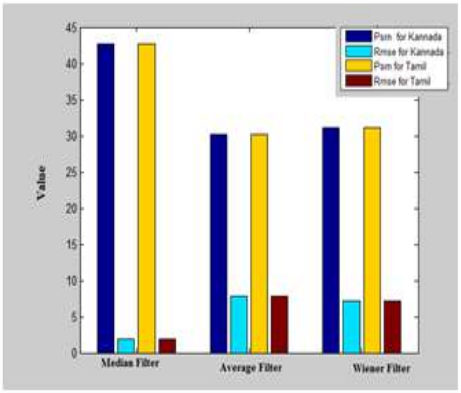


Figure 3 - writer 1

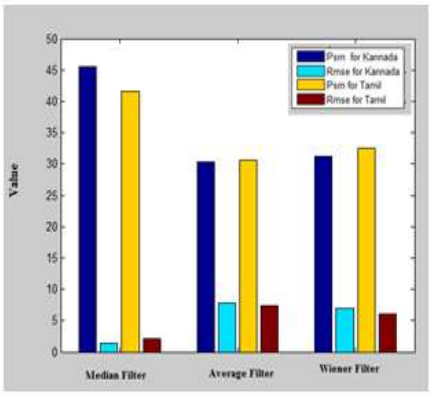


Figure 4 - writer 2

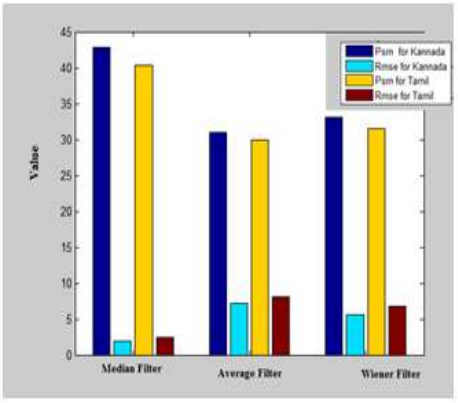


Figure 5 - writer 3

Comparison of PSNR and RMSE for Gaussian noise de-noised by filters for Tamil and Kannada languages using three different writers are shown in Figure 6, Figure 7 and Figure 8

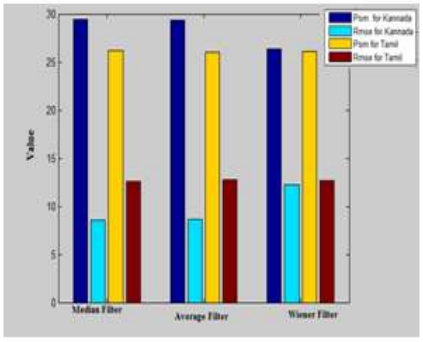


Figure 6 writer 1

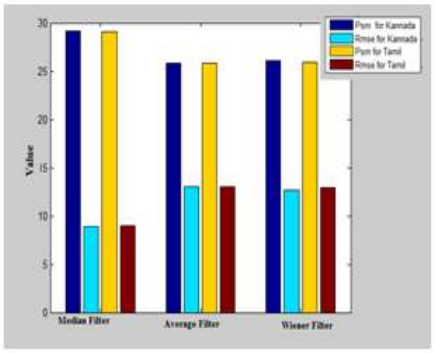


Figure 7 writer 2

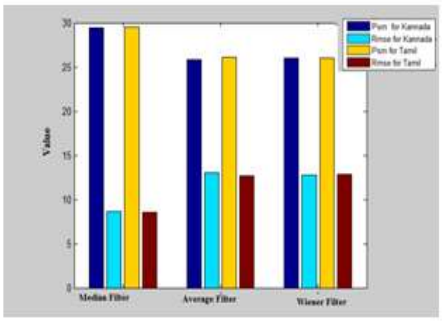


Figure 8 writer 3

IX. Comparative Study based as Result Analysis

The Table 1, 2, 3 and 4 showsthe PSNR and RSME value for salt and pepper noise and Gaussian Noise using different filtering techniques in Kannada and Tamil Languages for existing filter methods respectively. Comparison of our proposed method with the existing work is very difficult. Few works exist in thehandwritten character of Tamil and Kannada Languages which is experimented on different datasets of various complexities is expressed graphically.Figure 3, 4 and 5 shows the PSNRand RMSE value for *salt and pepper* noise using Median Filter, Average Filter and Wiener Filter for writer 1, 2 and 3 respectively. Figure 6, 7 and 8 shows the PSNRand RMSE value for *Gaussian*Noise using Median Filter, Average Filter and Wiener Filter for writer 1, 2 and 3 respectively.

X. Conclusion

In this paper three filtering techniques are applied on two different noise type of Gaussian Noise and Salt and Pepper Noise. From the experimental Analysis, the comparing of PSNR and RMSE value it is observed that the median filter is better in removing salt and pepper noise. The average and wiener filter are best to remove Gaussian Noise**.**

Reference

[1] Pal U, Chaudhuri BB. “Indian Script Character Rrecognition a Survey, Pattern Recognition”, 2004.

[2] Mamatha HosalliRamappa and Srikantamurthy Krishnamurthy,Skew Detection, Correction and Segmentation of Handwritten Kannada Document’.

[3] R. Singh, and M. Kaur, “OCR for Telugu Script using Back-propagation Based Classifier”, International Journal of Information Technology and Knowledge Management, 2010.

[4] V.R.Vijaykumar,P.T.Vanathi, P.Kanagasabapa thy and D.Ebenezer , “High Density Impulse Noise Removal using Robust Estimation Based Filter’’ IAENG International Journal of Computer Science, 35:3, IJCS,35,3,02 , 21 August 2008.

[5] R. Srinivas, Satarupa Panda, “Performance Analysis of Various Filters for Image Noise Removal in Different Noise Environment”, International Journal of Advanced Computer Research, ISSN (online): 2277-7970), Volume- 3 Number, 4 Issue, 13 December, 2013.

[6] Shi Zhong, Vladimir Cherkassky, “Image De-noising using Wavelet Thresholding and Model Selection”, IEEE, Proceedings, 2000.

[7] Q. Pan, “Two De-noising Methods by Wavelet Transform”, IEEE Transaction, Signal Processing, vol. 47, pp. 3401–3406, 1999.