Yorùbá Character Recognition: Image De-Noising For Effective Character Edge Detection

Yekeen S.A.
Department of Electrical and Electronics engineering
University of Ilorin
Ilorin, Kwara State, Nigeria.
E-mail: yeqeensirajudeen@gmail.com
Phone: +23408063785486)

ABSTRACT

In-spite of the numerous reported successes of applications of optical character recognition in languages based on Roman alphabets, the same success story is not true of Yoruba language. Though the twenty-five basic characters of Yoruba alphabets are drawn from Roman alphabets, the performance of optical character recognition of Yoruba document is abysmally low because Yoruba is a tonal language. Hence acceptable performance is only guaranteed by expanding the Yoruba's basic character set to include characters with diacritical tonal marks. Character image de-noising and effective character edge detection will go along way to improve the success rate in Yoruba character recognition.

Keywords: Edge detection, Filtering, Peak signal to noise ratio (PSNR) Similarity Structural index (SSIM), filter kernel.

iSTEAMS Proceedings Reference Format
Yekeen, S.A. (2019): Yorùbá Character Recognition: Image De-Noising For Effective Character Edge Detection. Proceedings of the 19th iSTEAMS Multidisciplinary Conference, The Federal Polytechnic, Offa, Kwara State, Nigeria.
7th – 9th August, 2019. Pp 57-64. www.isteams.net/offa2019 - DOI Affix - https://doi.org/ 10.22624/AIMS/iSTEAMS-2019/V19N2P6

1. BACKGROUND OF THE STUDY

Optical character recognition (OCR) is the process of converting digital image data of scanned character printed or handwritten into readable character streams by computer-based system. Conversion of handwritten document into computer editable version is important for easy accessibility and preservation of manuscripts related to our history. (Aradhana A.M. & Mitul M.P. 2014). Character recognition as an aspect of pattern recognition is a difficult task most especially the handwritten character due to the fact that each character differs with different writers, even differs with the same writer at different psychological condition. Character recognition majorly depends on the feature extraction of the characters, though other steps in character recognition includes data acquisition, pre-processing (smoothing/noise removal, edge detection de-skewing, morphological operation, segmentation). Feature extraction technique is generally classified into three group:

Statistical method, Structural method and Global transformation method. (G. Vamvakas , B Gatos,N Stamatopoulos, and S.J. Parentonis . 2008). Statistical method is based on statistical pixel distribution of an Image, statistical method approach includes portioning/zoning in regions, profile generation and projection as well as Distance and crossing approach. Structure and geometry method uses the physical structure of an image (character) which has the following metrics: vertical and horizontal lines, numbers of loops, numbers of strokes, major and minor axis, aspect ratio and cross points.

An OCR is widely used as a form of data entry from printed data records, e.g. in passport documents, invoices, bank statement so that it can be electronically stored, edited, searched. It also finds application in machine translation, text - to- speech and text mining. (Eric Bodden, Matte Clasen, Joachiu Kneisi. 2007). The predominantly existing optical character recognition is in Roman character, which has made conversation in English language (text and speech) across the globe easier, this also has helped in exchange of ideas and cultural values across the world. Yorùbá orthography comprises of Roman letters with diacritical marks (AbdulRahman O.I ,Odetunji A. O). Yorùbá character recognition will be more involving and challenging due to its diacritical marks which signifies the tone and meaning of Yorùbá words. Also, the under dots which differentiate similar character from each other for example OỌ oọ SṢsṣ EẸeẹ Yorùbá character like its Roman counterpart is divided into two sets of letters, the Consonants and the Vowels. It is the vowel letters that are given diacritical (tonal) marks, upper diacritical mark which indicates the tone and meaning of a word, out of these vowels two are given under dot. The remaining nineteen letters are Consonants letters, only one of them has under dot.

2. YORÙBÁ LANGUAGE ORTHOGRAPHY

Yorùbá language is a tonal language, the tone of the word defines its meaning. Words with the same orthography (spellings) will have different meaning depending on the tone with which they are called, for example Ìgbá (Locust tree), Ígbà (Climber), Igba(200), Igbà (Calabash) or Àjà (Dog), Ájá (Roof) etc. In these examples, the first four words are combination of same orthography likewise the last two, but with different diacritical tonal marks which actually give the meaning of each words. It is the combination of high (/), low (\) or mid ( ) tones that give right meaning of the combination of Yorùbá orthography.

The Yorùbá orthography is made up of twenty-five characters or alphabets, some of which are similar to Roman characters or letters. Out of these twenty-five characters, seven are vowels and the remaining eighteen are consonants. All the twenty-five letters in Yorùbá character set can assume upper and lower case. The twenty-five Yorùbá character set are as follows:

The upper case:
A B D E Ẹ F G GB I H J K L M N O Ọ P R S Ṣ T U W Y

The lower case:
a b d e ẹ f g gb i h j k l m n o ọ p r s ṣ t u w y

There are seven vowel letters in the Yorùbá character set, these letters are as follows:

Á É Ẹ Í Ó Ọ Ú — Vowel letters (Upper case) with high tone (acute)
À È Ẹ Ì Ò Ọ Ù — Vowel letters (Upper case) with high tone (grave)
á é ẹ í ó ọ ú — Vowel letters (Lower case) with high tone (acute)
à è ẹ ì ò ọ̀ ù — Vowel letters (Lower case) with high tone (grave)

3. EDGE DETECTION ALGORITHMS

Edge in a digital image may be defined as sudden change or discontinuity in pixel intensity of an image, its therefore, separate an object in an image from its background. Edge representation of a digital image drastically reduces the level of computational intensity in image processing and still retain the important features of the image for its recognition or classification (Mitra B. 2002). Edge detection algorithm are easily integrated into recognition algorithm for effective object classification or recognition used in computer vision application (Pratt W.K. 1978) (Rosenfeld and Kak A.C. 1982).

Over years there have been a number of methods of detecting edge from an image and subsequent feature extraction for pattern recognition. These edge detection methods include:

(1) Local or gradient-based edge detection method, this method includes Prewitt, Kirsch, Roberts, Sobel. The gradient edge detection methods detect the discontinuity in pixel intensity of an image by looking for the maximum and minimum in the first derivative of the image. (Yekeen S.A. & Ibiyemi T.S. 2018)

(2) Second Derivative or Zero Crossing techniques, examples of this operator is Marr-Hildreth. This is typically known as LoG (Laplacian of Gaussian) technique.

(3) Gaussian edge detection technique, examples of this detection technique includes Canny, Shen, Castan etc. This technique generally uses probability distribution function (PDF). Canny edge detector has best error rate so far. By variation method, Canny derives an optimal edge detection operator which turns out to be well approximated by the first derivatives of a Gaussian function (G. Vamvakas , B Gatos,N Stamatopoulos, and S.J. Parentonis (2008) (Olakanmi O. Oladayo 2015)..

(4) Colored edge detectors, are of three categories output Fusion methods, Multi-dimensional gradient methods and Vector methods (Jumoke F.A.,Stephen O.O., Elijah O.O, Odetunji O.O, 2015), Nandini N, Srikanta Murthy K., G. Hementha K. (2008)..

3.1 Noise in Digital Image

Noise in digital image is unwanted pixel values that do not reflect the true intensities of the objects in a digital image. Noise pixel intensities are either outrageously high or low, compared out of context of the image pixel intensities. Noise produces undesirable effects such as artefacts, unrealistic edges, unseen lines, corners, blurred objects and disturbs background scenes (Ajay K.B. and Brijdenra K.J. 2015). Noise therefore, pose difficulties in the process of image edge detection, feature extraction and other algorithm implementation in computer vision application and pattern recognition. The major source of digital image noise is during image acquisition and transmission in digital imaging system. Quantum noise arises from the discrete nature of electromagnetic radiation and its interaction with matter and also electronic noise which occur in detectors and amplifiers (Geoff Doughetry, 2009).

Other sources of digital image noise include:

(1) Variation in surface materials
(2) Illumination and contrast imbalance
(3) Error generated during analog digital conversion
(4) Ageing, may result to 1 and 2.
(5) For motion images electrical and electromechanical may be inclusive
(6) Atmospheric turbulence (Jumoke F.A.,Stephen O.O., Elijah O.O, Odetunji O.O, 2015).

There are several noise models in digital image. These models includes:
Gaussian (Normal noise), Rayleigh noise, Erlang (Gamma noise), Exponential noise, Impulse (Salt and pepper noise), Uniform noise, Periodic noise, Speckle noise e.t.c.

In this experimental study, Normal, Impulse and speckle noise will be considered being the most common digital image errors (noise) during acquisition and transmission.

In this experimental study Gaussian, speckle, impulse (salt and pepper) noise were used as case study.

3.2 Gaussian Noise

Gaussian noise is generated mostly in amplifiers and detectors, its therefore, called electronic noise (Geoff Doughetry, 2009), Gaussian noise is called normal or white noise, the probability density function (PDF) of Gaussian noise is given as

$$P(z) = \frac{1}{\sqrt{2\pi\sigma}} exp(-z-\mu)^2/2\sigma^2 \quad (1)$$

Where z - is the grayscale image pixel intensity

μ is the mean of the noise

σ is the variance or standard deviation of the noise pixel

3.3 Salt and Pepper Noise:

The salt-and-pepper noise are also described shot noise, impulse noise or spike noise. Its pixel value usually range from 0 – 255 (255 is salt and 0 pepper) impulse noise is usually caused by faulty memory locations, malfunctioning pixel elements in the camera sensors, or there can be timing errors in the process of digitization (Priyanka K. & Versha R. 2013).

Impulse noise also called Salt and Pepper noise, its PDF is given as

$$P(z) = \begin{cases} P_a\ ; & \text{for } Z = a \\ p_b; & \text{for } Z = b \\ 0\ ; & \text{otherwise} \end{cases} \quad (2)$$

3.4 Speckle Noise

Speckle is a multiplicative granular noise that exists in an image and degrades its quality. Speckle noise can be generated by multiplying random pixel values with different pixels of an image (Anisha V. 2018 https://medium.com/image-vision/noise-in-digital-image-processing-55357c9fab71). The model of speckle noise is given as:

$$p = i + \xi * i \quad (3)$$

Where, p is the speckle noise degraded image, i is the input image and ξ is the uniform noise image

4. EXPERIMENT AND RESULT ANALYSIS

In this experimental study, Yorùbá characters were generated using Matlab code. As it was stated earlier, Yorùbá characters are similar to Roman letters, this research therefore, concentrates on Yorùbá vowel characters with diacritical marks. Canny edge detection algorithm was preferred for its been the most powerful edge-detection method. The Canny method differs from the other edge-detection methods in that it uses pixel connected component analysis (CCA) and two different thresholds (to detect strong and weak edges). The weak edges are included in the output only if they are connected to strong edges. This method is therefore more likely to detect strong and true edges (Matlab documentation 2015).

Matlab code generated Yorùbá character images were degraded with Normal (Gaussian) noise, impulse (Salt and Pepper) noise and multiplicative (Speckle) noise. Several values of mean μ and variance (Standard deviation) σ of Gaussian noise were used. For impulse (Salt and Pepper) noise, several values of noise density were used. For speckle noise, several values of variance were used since the speckle is uniformly distributed noise, its therefore have zero mean.

The degraded character image was the filtered using linear and non-linear filters of 5X5 kernel. The filters used were Averaging filter, Weiner filter, Gaussian filter and non-linear median filter.

Digital Image Quality Metrics

There are several image quality measurement techniques used to evaluate and assess the quality of image for comparisons. These techniques include signal to noise ratio (SNR), peak signal to noise ratio (PSNR), mean squared error (MSE), universal image quality index (UIQI), structural similarity index method (SSIM), human vision system (HVS), featured similarity index method (FSIM) etc, the most efficient and commonly used metrics are SNR, PSNR and MSE. This work has opted for these three methods.

Signal to Noise Ratio (SNR)

SNR measures the sensitivity of imaging, its signifies the signal strength relative to the background noise (Amit K.S, Nomit S, Mayank D, Anand M, 2012)

The SNR of an images is given as

$$SNR = 10 * log_{10} \frac{\mu}{\sigma} \quad (4)$$

μ - The mean value of the input image pixels
σ - The variance of the image pixel and the background pixels
SNR - Signal to noise ratio

Peak Signal to Noise Ratio (PSNR)

PSNR is used to determine the degradation in the embedded image with respect to the host image (http://en.wikipedia.org/wiki/Signal_to_noise_ratio, June 2019). PSNR is calculated using the ratio of the maximum possible signal power to the power of the distorting noise which affects the quality of its representation. This ratio between the two images is computed in decibel form.

PSNR is given as:

$$PSNR = 10log_{10} \frac{L^2}{MSE} \quad (5)$$

L = is the peak signal value (image)
MSE = is the mean squared error

Structural Similarity Index Measure (SSIM)

The SSIM is one the most effective image quality metric used to measure the similarity between two images. SSIM has high correlation with human visual system (HVS) (Wang *et a*l 2004). SSIM measures the quality of an image by measuring the structural similarity that compares local patterns of pixel intensities that have been normalized for luminance and contrast. This quality metric is based on the principle that the human visual system is good for extracting information based on structure (Matlab Documentation 2015). SSIM as an objective image quality metric require the existence of a distortion-free image, called the reference image, that can be used for comparison with the image whose quality is to be measured. The similarity index is based on the multiplicative computation of three terms, the luminance, the contrast and the structural terms.

SSIM algorithm is given as

$$SSIM(p,q) = [l(p,q)]^{\alpha} * [c(p,q)]^{\beta} * [s(p,q)]^{\gamma} \quad (5)$$

Where

$$l(p,q) = \frac{2\mu_p\mu_q + C_1}{\mu_p^2 + \mu_q^2 + C_1} \quad (6)$$

$$c(p,q) = \frac{2\sigma_p\sigma_q + C_2}{\sigma_p^2 + \sigma_q^2 + C_2} \quad (7)$$

And

$$s(p,q) = \frac{\sigma_{pq} + C3}{\sigma_p\sigma_q + C_3} \quad (8)$$

Where $\mu_p, \mu_q, \sigma_p, \sigma_q$ and σ_{pq} are the local means, standard deviations and cross-covariance for the image p.

If α= β=γ=1, the similarity index simplifies to:

$$SSIM(P,q) = \frac{(2\mu_p\mu_q + C_1)(\sigma_{pq} + C_2)}{(\mu_p^2 + \mu_q^2 + C_1) + (\sigma_p^2 + \sigma_q^2 + C_2)} \quad (9)$$

The luminance comparison function $l(p,q)$ measures the correlation of the two images. Mean luminance ($\mu_{p,} and\ \mu_q$). This factor is maximally equal to 1 *iff* $\mu_p = \mu_{q\ ,}$. The contrast comparison function $c(p,q)$ measures the correlation of the contrast of the two images. The contrast is measured by the standard deviation $\sigma_p and\ \sigma_q$. This term is maximally equal to 1 *iff* $\sigma_p = \sigma_q$. And lastly the structural comparison function $s(p,q)$ measures the correlation coefficient between the two images $p\ and\ q$. The term σ_{pq} is the covariance between $p\ and\ q$. The values of the SSIM index are in (0,1). A value of 0 means no correlation between the two images, and 1 means that the two images are structurally equal. The constants C_1 C_2 and C_3 are used to avoid a null denominator.

The filtered images were subjected to following image quality metrics, SNR, PSNR and SSIM, the results obtained are shown in table1.

Table 1: Table 1: Image quality metrics showing SNR, PSNR and SSIM of Edge YORÙBÁ character image edge detection

Salt and Pepper Noise					
	Corrupted grayscale image	Averaging Filter	Wiener Filter	Median Filter	Gaussian Filter
PSNR	16.2439	26.2796	21.0013	38.9724	23.2216
SNR	15.9288	25.9646	20.6862	38.6574	19.3441
SSIM	0.6671	0.8375	0.9135	0.8927	0.8516
Speckle Noise					
PSNR	23.1081	25.6476	26.9041	31.1325	23.2034
SNR	22.7930	25.3325	26.5890	30.8174	22.3259
SSIM	0.6754	0.8374	0.9496	0.8130	0.9478
Gaussian Noise					
PSNR	23.0416	26.5566	27.4878	33.3119	36.4567
SNR	22.7266	26.2415	27.1727	32.9968	31.4532
SSIM	0.6148	0.8392	0.9496	0.7950	0.9380

Figure 1 shows the results obtained from the test image after being degraded with noises, filtered and edge detection carried out at each stages of the experiment.

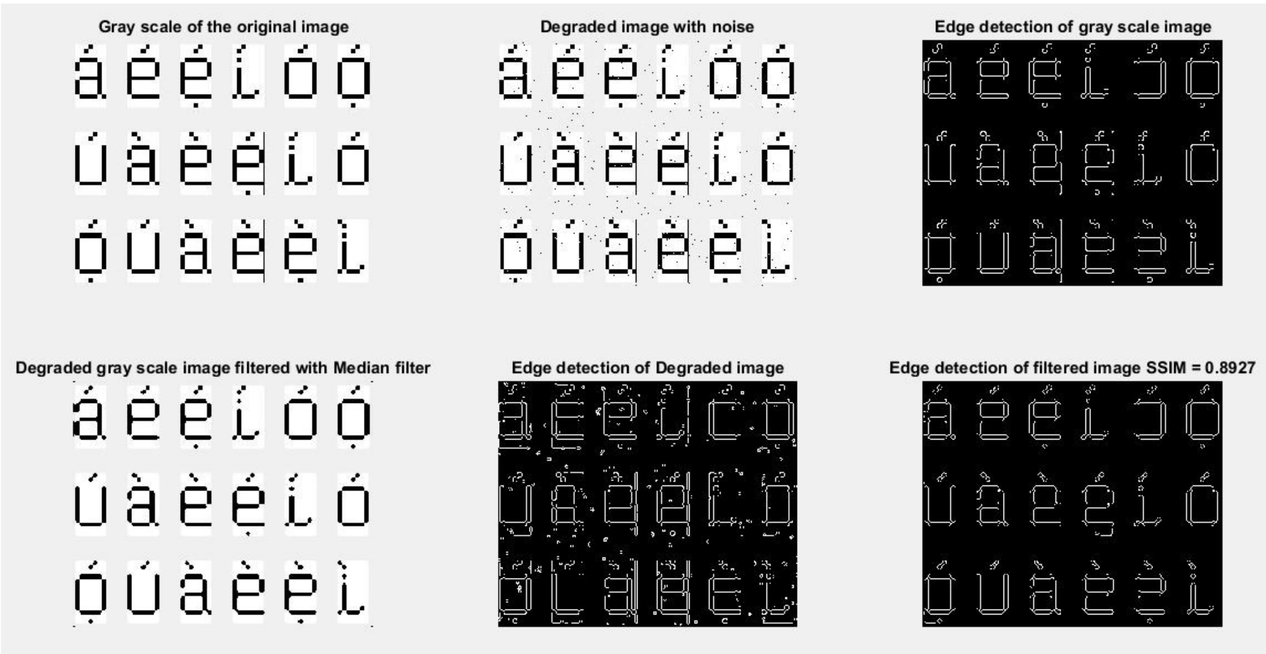

Figure 1: Preprocessing and edge detection of generated YORÙBÁ character image.

5. CONCLUSION

The image quality metrics (SNR, PSNR and SSIM) values obtained from the experiment, it shows that filtering improves quality of corrupted image. The level of image edge detection was also improved, these will eventually improve the YORÙBÁ character recognition as the featured extraction will be as close as possible to the expected values.
The result of the experiment also gave a clue of what filter have a better response or more sensitive to each type of noise. Therefore, if researchers are sure of the noise that degrades an image, a right filter could be chosen for the restoration process.

REFERENCES

1. AbdulRahman O.I ,Odetunji A. O. *Exploiting Handwritten Character Recognition System in African Language Documentation and Education*. OAU Ile-Ife (Unpublish).
2. Ajay K.B. Brijdenda K.J. (2015). *A review of noise model in digital image processing. Signal and image processing*: An international journal SPIJI vol6 No 2
3. Aradhana A.M, Mitul M.P. (2014). *Hand written Davanagari script recognition: A survey*. IOSR Journal of electrical and electronics engineering. IOSR – JEEE ISSN:2320 – 3331 vol9 issue 2 Ver II pg 80 – 87
4. Eric Bodden, Matte Clasen, Joachiu Kneisi. (2007). *Arithmetic coding revealed A guided tour from theory to praxis*. MC grill university school of computer science research group.

5. G. Vamvakas , B Gatos,N Stamatopoulos, and S.J. Parentonis (2008). *A complete optical character recognition methodology for historical document.* The eighth IAPR workshop on document analysis system IEEEcomputer society 978-0-76953337-7/08.
6. Geoff Dougherty, (2009) Digital Image Processing For Medical Applications; Cambridge University press, The Edinburgh building Cambridge CB2 8RU UK G.Dougherty, ISBN -13 978-0-511-53343-3
7. MATLAB Documentation 2015
8. Mitra Basu. (2002) Senior Member, IEEE, *Gaussian-Based Edge-Detection Methods-A Survey*. IEEE Transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 32, no. 3.
9. Nandini N, Srikanta Murthy K., G. Hementha K. (2008). *Estimation of skew angle in Binary Document Image using Hough Transform.* International Journal of Computer, Electrical, Automation, Control and Information engineering Vol 2 No6.
10. Olakanmi O. Oladayo (2015). *YORÙBÁ language and numeral offline interpreter using morpholigcal and template matching*. Telkominika Indonesia journal of electrical engineering, vol 13 No1 pp 166-173
11. Pratt W.K. *Digital Image processing*. New York Wiley inter science
12. Priyanka K, Versha R. (2013). *A brief study of various noise model and filtering techniques*. Journal of global research in computer science. Vol 4 No 4.
13. Ram Sarkar Malakar, Nibran Das, Subhadip Basu. (2011) *Word Extraction and character segmentation from text lines of unconstrained handwritten Bangla document images*. J.Intell system 20.page 227 – 260.
14. Rosenfeld A. and Kak A.C (1982) *Digital image processing*. 2nd edition New York academy.
15. Yekeen S.A. and Ibiyemi T.S.(2018) *Edge detection algorithm for YORÙBÁ character recognition*. Advances in Vision Computing: An International Journal (AVC) Vol.5, No.1/2/3/4.