ISSN:2231-0711

Leena Mary Francis et al | IJCSET(www.ijcset.net) | May 2016 | Vol 6, Issue 5, 144-148

Assemblage of Objects from Natural Scene

Leena Mary Francis[#1], K.C.Visalatchi[#2], Dr. N. Sreenath[#3]

[#] *Dept. of Computer Science and Engineering , Pondicherry Engineering College*
Pondicherry Central University, Pondicherry, India

***Abstract* - The first phase of any image recognition process includes image pre-processing process which enhances the image and makes it clear for further detection mechanism. Picture pre-handling can fundamentally build the dependability of an optical investigation. Some channel operations which improve or diminish certain image subtle elements empower a less demanding or speedier assessment. The AssemblagE of Objects from Natural Scene (AEON) proposes to pre-process the image and to collect the various objects from the natural scene. The image pre-processing phase includes two steps involved in it namely Image De-noising and Contrast Enhancement. Total variation approach is applied in order to remove the noise from the image and Otsu method is applied in order to enhance the clarity of the image. Then the processed image is given to the next phase in order to collect all the objects embedded in the natural scene, the Histogram of Oriented Gradients feature descriptor is used to collect the pool of objects. The dataset used for pre-processing and collection of objects is ICDAR 2015, on which performance measures such as precision, recall and f-measure are evaluated and gives good results.**

***Keywords* – contrast enhancement, de-noising, histogram of oriented gradients, image pre-processing**

I. INTRODUCTION

Detecting text in the natural scene images is a challenging research issue with applications in various domains such as accessibility computing, robotics etc. To enhance the process of detecting text in the images the preprocessing becomes an mandatory stage. The source shall be still images or videos. Most image preprocessing technqiues represents the source as a two dimensional medium. Pictures are additionally prepared as three-dimensional signs where the third-measurement being time or the z-pivot [1].

Object identification is an innovation identified with computer vision and picture handling that arrangements with examples of semantic objects of a specific class. Very much scrutinized spaces of item identification incorporate face location and person on foot discovery. Object recognition has applications in numerous ranges of PC vision, including picture recovery and video reconnaissance.

The processing of natural scene is considered to be most complicated process because of various factors such as very complex backgrounds, variations of light intensity and shadow, billions of colours and very difficult to perceive the image. Only after detecting various objects in the image any recognition process can be done over the image which can be used in various predictions such as face, text, things and even can be used in robot navigation. AEON (AssemblagE of Objects from Natural Scene) focus over image processing and to localize the objects in the image.

The remainder of this paper is strcutrued as follows: Section II briefly discusses about the various current literatures, Section III presents detail regarding the proposed system, Pre-processing is presented in Section IV, Object detection is stated out in Section V, Section VI discusses in detail the experimental setup for the system, conclusions and future directions is given in Section VII.

II. RELATED WORKS

Image processing of digital images is the utilization of various algorithms to perform processing of the image on digitized images. Than analog image processing, digital image processing have greater advantages, as it allows a greater range of techniques that can be applied to input data that can be solving the problems faced by noise and intensity of light varying across the image [2]. Since pictures represented using more than two metrics (maybe more) complex image preparing might be implemented as multidimensional frameworks.

Lien C Y et al states the random-values impulse noise is removed by low-cost VLSI architecture. The noise pixel is detected by decision-tree-based detector and gives out a potent design to find the edges. Moreover, a versatile innovation is used to improve the impacts of removal of noise. the experimental results exhibit that the proposed strategy can acquire better results as far as both quantitative assessment and visual quality than the past lower many-sided quality techniques [3]. Also, the execution can be equivalent to the higher many-sided quality techniques.

Taking into account a basic piecewise-smooth picture earlier, the author proposes a division based way to deal with consequently gauge and expel commotion from shading pictures. The NLF is acquired by evaluating the lower envelope of the standard deviations of picture change per portion. The chrominance of the shading commotion is altogether evacuated by anticipating the RGB pixel qualities to a line in shading space fitted to every fragment [4]. The commotion is expelled by detailing and unravelling a Gaussian contingent arbitrary field. Examinations were directed to test both the noise estimation and evacuation calculations. Our de-noising calculation gives good wavelet de-noising calculations on both engineered and genuine noise polluted pictures by creating shaper edges, delivering smoother level locales and saving unpretentious composition points of interest. These elements coordinate our unique criteria that the author proposed for a decent de- noising calculation.

Melange T et al in his paper have introduced a separate structure for removing noise. With a specific end goal to save the subtle elements however much as could be expected, the noise is evacuated regulated. The identification of uproarious shading components is in light of fluffy standards in which data from spatial and worldly neighbours and also from the other shading groups is utilized [5]. Identified uproarious parts are sifted in light of block matching where a noise versatile mean total contrast is utilized and where the pursuit locale contains pixels obstructs from both the past and current edge. The analyses demonstrated that the proposed strategy beats other best in class strategies both as far as target measures, for example, MAE, PSNR and NCD and visually.

Enhancement of contrast and preserving brightness are basic prerequisites for machine vision tasks. In any case, these are two contradicting goals when the picture is prepared by histogram equalization approaches. Current accessible strategies may not give comes about all the while fulfilling both necessities. Lin S C F et al states that by utilizing stretching colour data from a scene is re-established. Averaging against a uniform distribution empowers the yield picture to recover the data lost. Moreover, histogram re-mapping diminishes artefacts that frequently emerge from the adjustment system [6]. The procedure similarly utilizes a search procedure to discover ideal procedural arguments, such that the average difference of brightness between the input and output samples is reduced to the maximum possible extent. The adequacy of the proposed technique was tried with an arrangement of images caught in open environment and compared against existing methods. Performance was improved into better place.

M T Yildirim et al uses type 2 fuzzy logic image filtering operator to expel the noised from digitized images. The central prevalence of the proposed administrator over contending administrators is that it effectively expels impulse noise from digitized pictures while effectively protecting thin lines, edges, fine points of interest, and primary texture of the source image. The results conclude that the proposed operator can be utilized as an intense device for effective removal of noise from digial images without mutilating the valuable data inside the image [7].

An innovative complexity handling algorithm FC-CLAHE for handling anomalies in advanced mammogram is developed by Jenifer S et al [8]. The calculation alogorithm the choice of clasp cutoff in the conventional CLAHE method utilizing the soft deduction framework. From the examinations, it is clear that the variations of the picture has been enhanced by saving the general state of the source image. When histogram adjustment is connected to source pictures, an arrow scope of input intensity are mapped to wide range of output intensity value. At the point when difference constrained versatile histogram balance system is connected, the histogram is cut beyond a certain cutoff. FC-CLAHE system robotizes the determination of cliplimit. These accumulated extra entirety adds up to a significant value and when redistributed makes normal canister substantial in height.Now an extensive variety of input qualities is mapped to extensive variety of output values. The incline of the mapping capacity rises consistently over the whole scope of input qualities. At the expense of cut-out, a perfect mapping capacity is accomplished. It saves the contrast of the picture as well as enhances the picture complexity and entropy without decay of data in the original image.

Kim S E et al states that utilizing entropy scaling as a part of the wavelet space can improve the contrast of the image. The principle commitment of the research is that the proposed calculation is a first picture upgrade approach which utilized the local entropy scaling as a part of the wavelet domain. An inferred numerical works for entropy scaling in the wavelet change area to upgrade picture contrast, and built up a straightforward shading improving technique in HSI shading space. The proposed technique comprises of two stages. In the initial step, the differentiation of intensity component of the HSI shading space was improved. The high-recurrence coefficients were scaled by amplifying the entropy of the complexity characterized in the wavelet area [9]. In the second step, the immersion part of the HSI shading space was improved by straight scaling utilizing the changed intensity component. Simulated results demonstrate that the proposed technique produced great picture improvement execution.

Daniel E et al has shown improvement of scale quality with the enhanced genetic algorithm (EGA) [10]. The proposed strategy is tried for retinal information base pictures, in which both subjective and quantitative exhibitions are accomplished. In medicinal imaging procedures, the nature of pictures has more significance and needs large focus in this area.

Through the detailed literature survey, the necessity of pre-processing the image and localization is evidently shown. The basic requirement to do any processing with the image is that first the noise of the image should be removed and contrast is to be improved to enhance the image so that object localization can be done with much higher precision.

III. OVERVIEW OF THE PROPOSED SYSTEM

The related works gives the necessity of processing the image and the proposed work AEON does image pre-processing the object localization from the image.

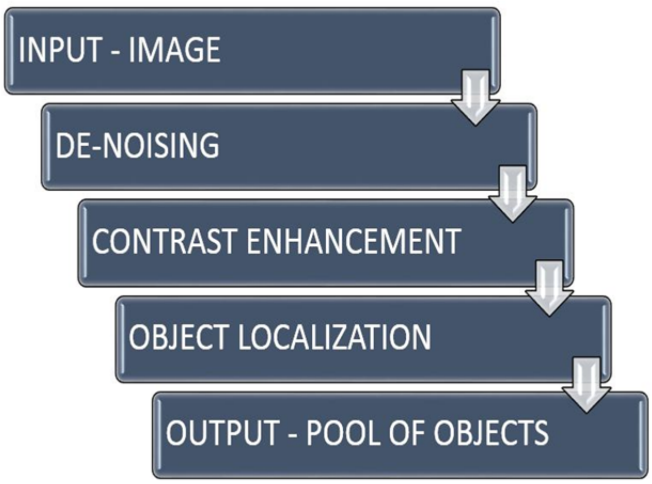

Fig. 1 Flow architecture of the AEON

The input to AEON is the natural scene image. The Figure 1 explains the progression of the process from one stage to another. After receiving the image first the image's noise is expelled using Total Variation Approach. Then the definition of the image is improved using Otsu method. After the pre-processing stage the image is ready for further processing. AEON aims at localization of objects from the image to form the pool of objects which needs to recognize further for any recognition problem which is not our area of focus. The object localization is done using Histogram of Oriented Gradients giving out pool of objects.

Fig. 2 Scene text images from ICDAR 2015

The Figure 2 gives the images from International Conference on Document Analysis and Recognition (ICDAR) which gives the various texts taken from natural environment. When objects are localized from these images, the pool of objects can be used in next phase of text recognition from natural scene which doesn't get primary importance in this paper, which mainly focusses on retrieval of objects from the natural scene.

IV. IMAGE PRE-PROCESSING

Considering specific objects of computerized noise diminishment innovation, we first need to comprehend "noise". Noise is any undesirable unsettling influence in a signal. It can square, twist, change or meddle with the signal. In signal handling it is any undesirable information that is created as an undesirable by-result of different exercises and is not used to transmit the sign. In video flags this appears to be fine static or "snow". Everyone comprehends that commotion is an undesirable unsettling influence in a sign which can block, mutilate or intersect with the signal. That is the reason noise removal framework has numerous advantages. Here AEON uses Total variation approach [11], otherwise called all out difference regularization is a procedure, frequently utilized as a part of image processing, that has applications in elimination of noise. It depends on the rule that signs with inordinate and conceivably spurious point of interest have high aggregate variety, that is, the indispensable of the outright slope of the sign is high. As per this rule, diminishing the aggregate variety of the sign subject to it being a nearby match to the first flag, expels undesirable point of interest whilst protecting vital subtle elements, for example, edges [12]. This noise evacuation procedure has focal points over straightforward strategies, for example, direct smoothing or middle sifting which lessen noise yet in the meantime smooth away edges to a more prominent or lesser degree. By complexity, absolute variety de-noising is amazingly powerful at the same time protecting edges whilst smoothing without end noise in level areas, even at low signal to noise proportions.

For a digital signal y_n, total variation V is defined as,

$$V(y) = \sum_{n} |y_{n+1} - y_n|$$

Fig. 3 Original Image before pre-processing

In computer vision Otsu strategy is utilized to naturally perform grouping based image thresholding [13]. The computation projects that the images containing two classes of pixels considering after the bi-modular histogram [14], it then figures the ideal edge isolating the two classes so that their consolidated spread (intra-class difference) is insignificant, or comparably, so that their between class fluctuation is maximal. Consequently, Otsu's technique is about a one-dimensional, discrete simple of Fisher's Discriminant Analysis [15]. The Otsu method intensively searches for the threshold that reduces the within-class varianc that is computed as a weighted sum of variances of the two classes:

$$\sigma_{\omega}^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

Weights $\omega_{0,1}$ are the probabilities of the two classes separated by a threshold t and $\sigma_{0,1}^2$ are variances of these two classes.

Fig. 4 Image is Pre-processed and objects are localized

V. Object Detection

Histogram of oriented gradients (HOG) [16] is a component descriptor used to distinguish objects. The technique works on the principle of counting the density of oriented gradients in the restricted segments of the image - window detection [17].

Usage of the HOG descriptor calculation is as per the following:

1. Partition the image into small section called cells, and for every cell process a histogram of direction of gradients or orientation of edges for the pixels inside the cell.
2. Discretize every cell into angular bins as per the inclination introduced.
3. Every cell's pixel contributes weighted inclination to its relating angular bin.
4. Gatherings of adjoining cells are considered as spatial districts called histograms. The gathering of cells into a block is the premise for gathering and standardization of histograms.
5. Standardized gathering of histograms speaks to the piece of block histogram. The arrangement of these block histograms speaks to the descriptor.

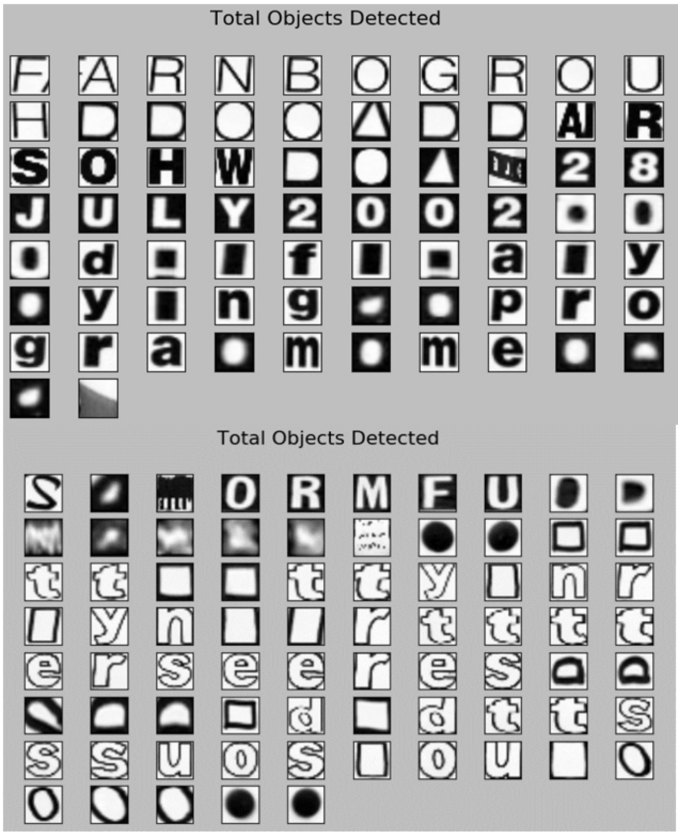

Fig. 5 Total objects detected from image

The final stage of AEON gives out the pool of objects for further research concern [18]. The figure 5 shows the pool of objects including some false positives also.

VI. Experimental Setup And Results

AEON has been implemented in Python, scientific programming language on Ubuntu environment. Precision is the percentage of selected items that are correct [19]. The precision is computed using the formula as shown below:

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall is the percentage of correct items that are selected. It's computed as the ratio between ground truth and match score [20]. The recall is computed using the formula as shown below:

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F-measure is the weighted harmonic mean of Precision and Recall [21].

$$F\ Measure = 2\frac{(Precision * Recall)}{Precision + Recall}$$

TABLE 1. PRF VALUES OF VARIOUS METHODS

Method	Precision	Recall	F- measure
Lien CY et al.	71.4	68.3	69.82
Melange T et al.	75.6	71.2	73.33
Daniel E et al.	69.5	63.2	66.20
Jenifer S et. Al	74.5	65.4	69.65
AEON	76.5	74.5	75.49

The various approaches precision, recall and f-measure values are tabulated in the Table 1. AEON shows better results of 76.5% of precision, 74.5% of recall and 75.49% of f-measure. The Surface chart clearly depicts the upward slope for AEON.

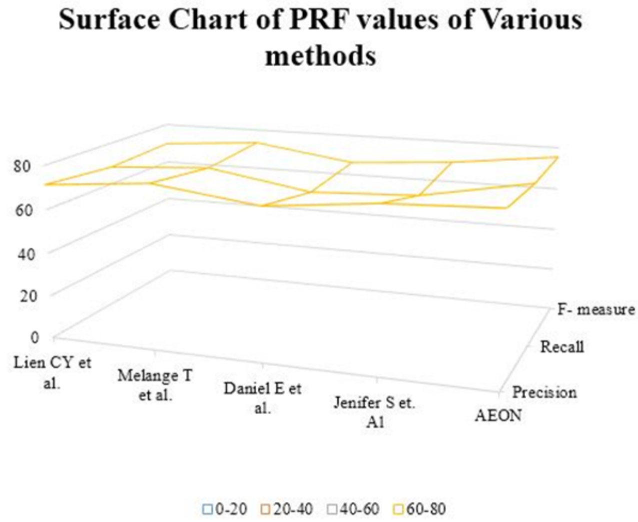

Fig. 6 Surface chart of PRF values

VII. CONCLUSION

This paper AEON proposes method to pre-process the image, and localize the objects. The noise was removed to give a clear image for processing using Total variation approach, and contrast was increased preserving the edges by Otsu method. After localizing the objects the various objects are extracted from the image and finally given as pool of objects as output. The evaluation of proposed system was evaluated over ICDAR2015 dataset and has given promising results. But few false positives were also deducted should slightly decrease the recall value. On the whole AEON performs well on natural scene images giving the pool of objects which can be further used for various recognition researches.

REFERENCES

[1] V. Bruni, I. Selesnick, L. Tarchi, and D. Vitulano, “An adaptive perception-based image preprocessing method,” in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, 2015, pp. 2331–2335.

[2] M. Ashour, S. Elhoshy, M. Ibrahim, T. El-Shabrawy, and H. Hammad, “An enhanced preprocessing dependent image theory algorithm for indoor coverage solutions,” in *2016 33rd National Radio Science Conference (NRSC)*, 2016, pp. 274–281.

[3] C.-Y. Lien, C.-C. Huang, P.-Y. Chen, and Y.-F. Lin, “An efficient denoising architecture for removal of impulse noise in images,” *Comput. IEEE Trans. On*, vol. 62, no. 4, pp. 631–643, 2013.

[4] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, “Automatic estimation and removal of noise from a single image,” *Pattern Anal. Mach. Intell. IEEE Trans. On*, vol. 30, no. 2, pp. 299–314, 2008.

[5] T. Mélange, M. Nachtegael, and E. E. Kerre, “Fuzzy random impulse noise removal from color image sequences,” *Image Process. IEEE Trans. On*, vol. 20, no. 4, pp. 959–970, 2011.

[6] S. C. F. Lin, C. Y. Wong, M. A. Rahman, G. Jiang, S. Liu, N. Kwok, H. Shi, Y.-H. Yu, and T. Wu, “Image enhancement using the averaging histogram equalization (AVHEQ) approach for contrast improvement and brightness preservation,” *Comput. Electr. Eng.*, vol. 46, pp. 356–370, 2015.

[7] M. T. Yildirim, A. Basturk, and M. E. Yuksel, “Impulse noise removal from digital images by a detail-preserving filter based on type-2 fuzzy logic,” *Fuzzy Syst. IEEE Trans. On*, vol. 16, no. 4, pp. 920–928, 2008.

[8] S. Jenifer, S. Parasuraman, and A. Kadirvelu, “Contrast enhancement and brightness preserving of digital mammograms using fuzzy clipped contrast-limited adaptive histogram equalization algorithm,” *Appl. Soft Comput.*, vol. 42, pp. 167–177, May 2016.

[9] S. E. Kim, J. J. Jeon, and I. K. Eom, “Image contrast enhancement using entropy scaling in wavelet domain,” *Signal Process.*, vol. 127, pp. 1–11, Oct. 2016.

[10] E. Daniel and J. Anitha, “Optimum green plane masking for the contrast enhancement of retinal images using enhanced genetic algorithm,” *Opt.-Int. J. Light Electron Opt.*, vol. 126, no. 18, pp. 1726–1730, 2015.

[11] “Total variation denoising,” *Wikipedia, the free encyclopedia*. 27-Jan-2016.

[12] “Total Variation Denoising (MM Algorithm).” [Online]. Available: http://eeweb.poly.edu/iselesni/lecture_notes/TVDmm/. [Accessed: 07-May-2016].

[13] “Global image threshold using Otsu’s method - MATLAB graythresh - MathWorks India.” [Online]. Available: http://in.mathworks.com/help/images/ref/graythresh.html. [Accessed: 07-May-2016].

[14] “Otsu’s method,” *Wikipedia, the free encyclopedia*. 29-Feb-2016.

[15] “Otsu Thresholding - The Lab Book Pages.” [Online]. Available: http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html. [Accessed: 07-May-2016].

[16] S. Tian, S. Lu, B. Su, and C. L. Tan, “Scene Text Recognition Using Co-occurrence of Histogram of Oriented Gradients,” in *2013 12th International Conference on Document Analysis and Recognition*, 2013, pp. 912–916.

[17] H. A. Rashwan, M. A. Mohamed, M. A. García, B. Mertsching, and D. Puig, “Illumination Robust Optical Flow Model Based on Histogram of Oriented Gradients,” in *Pattern Recognition*, J. Weickert, M. Hein, and B. Schiele, Eds. Springer Berlin Heidelberg, 2013, pp. 354–363.

[18] B. F. Wu, C. C. Kao, C. L. Jen, Y. F. Li, Y. H. Chen, and J. H. Juang, “A Relative-Discriminative-Histogram-of-Oriented-Gradients-Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 8, pp. 4228–4237, Aug. 2014.

[19] J. Wang, L. Shi, H. Wang, J. Meng, J. J.-Y. Wang, Q. Sun, and Y. Gu, “Optimizing top precision performance measure of content-based image retrieval by learning similarity function,” *ArXiv160406620 Cs*, Apr. 2016.

[20] “Precision and recall,” *Wikipedia, the free encyclopedia*. 13-Apr-2016.

[21] “F1 score,” *Wikipedia, the free encyclopedia*. 04-Feb-2016.