International Journal of Electrical and Electronic Engineering & Telecommunications

ISSN 2319 – 2518 www.ijeetc.com
Vol. 3, No. 3, July 2014

Research Paper

COLOR HISTOGRAM BASED TEXT DETECTION IN IMAGES

Bhagyalaxmi P T[1]* and Bhaskar B[1]

*Corresponding Author: Bhagyalaxmi P T, ✉ bhagyajain2890@gmail.com

Texts in characteristic scene pictures convey rich semantic data, which might be utilized to help an extensive variety of provisions, for example, object distinguishment, picture recovery, and human machine cooperation. Text in many pictures taken commonly by a cell phone can have a critical relative or point of view twisting, making the current content location and the consequent OCR motors inclined to disappointments. Not at all like other content recognition approaches that for the most part depend on either content structure or color, we propose an edge work that joins together both by overseeing content situated shade diminishment process with adjusted SWT data. SWT pixels mapped to shade space vote for the color they compare to. Shades accepting high SWT vote in all probability fit in with content zones and are obstructed from being mean-moved away and than applying composition invariant low rank convert that can identify and correct messages in self-assertive introductions in the picture against complex foundations, so that the writings could be accurately perceived by basic OCR motors.

Keywords: Natural scene pictures, SWT, Color decrease, Texture invariant low rank transform, Arbitrary introduction

INTRODUCTION

Content Extraction from picture is concerned with concentrating the pertinent content information from an accumulation of pictures. These days the extent of the accessible computerized feature substance is expanding quickly. This prompts a pressing requirement for quick and compelling calculations for data recovery from interactive media content. Printed data in interactive media constitutes an exceptionally rich wellspring of abnormal amount semantics for recovery and indexing. Complex foundations, uneven light, and vicinity of just about boundless number of content text styles, sizes, and introductions posture incredible troubles even to state-of-the-craftsmanship content identification techniques. Not at all like record pictures, where content is typically superimposed on either clear or complex foundations and is in

[1] Department of Electronics &Communication Engineering, SJBIT, Bangalore, Karnataka, India.

this manner more unique (Chen and Wu, 2009; and Zagoris *et al.*, 2011), regular scene pictures bargain with scene content, which is as of now a piece of the caught scene and is regularly a great deal less different. All things considered, content discovery has turned into an extremely prominent examination territory because of its gigantic potential in numerous practical ranges, for example, sign interpretation, substance based web picture looking, and helping the outwardly debilitated. State-of-the-craftsmanship writing recognizes two significant content discovery approaches: composition based and area based. Surface based systems (Chen and Yuille, 2004; and Ye *et al.*, 2005) sweep pictures at diverse scales, review zone under the sliding window for content like gimmicks, and arrange it as content/non-message. They frequently need accuracy and are moderately moderate because of their scale-space approach. District based strategies (Epshtein *et al.*, 2010; and Yi and Tian, 2011), then again, work in a base up style by selecting pixels (or areas) with ordinary content properties and gathering them into joined segments that are further geometrically separated and assembled into content lines and/or words. Area based routines are not constrained to content size/ introduction and are (contrasted with composition based strategies) quicker. Other than the previously stated methodologies, mixture methodologies exist, which adventure preferences of both composition based and region based methodologies (Pan *et al.*, 2011). In the vision group, it is for the most part accepted that content location and distinguishment have been a broadly concentrated on issue and developed results should as of now exist. This couldn't have been a long way from reality. For one, most existing content recognition and distinguishment frameworks accept that the writings in the picture are taken from an about frontal perspective. This is frequently a long way from the case for web pictures or pictures taken by cellular telephones, The writings in the picture could be extremely contorted by a subjective planar (revolution, relative, or homography) convert from the picture plane. Practically all current content recognition routines perform inadequately on such pictures. Regardless of the possibility that such messages are by one means or another discovered the resulting distinguishment would additionally fall flat in light of the fact that the writings are not in their upright position which is frequently required by OCR motors for correct distinguishment. To purpose this issue, we propose a straight forward yet compelling plan that coordinates SWT and TILT systems and can naturally separate just about writings in a picture not withstanding their self-assertive beginning miss hapenings.

METHODOLOGY

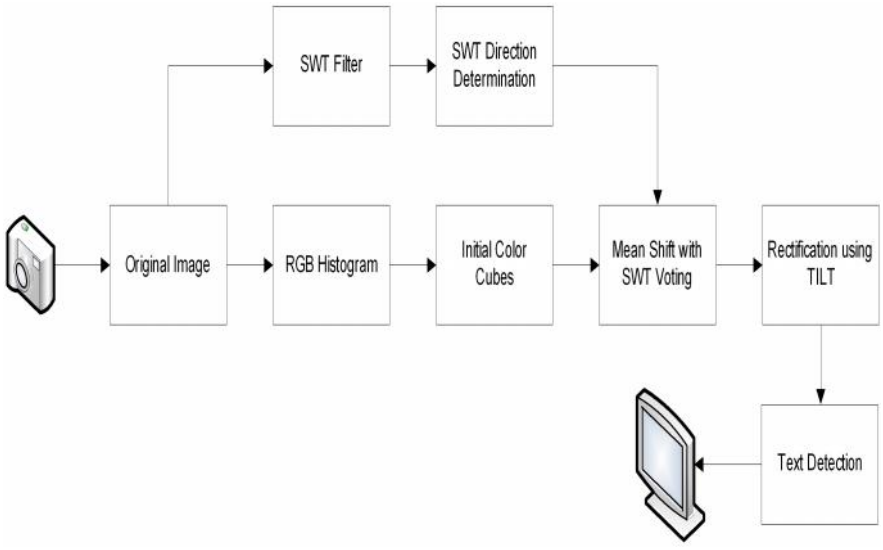
Text in common scene pictures is recognized from other picture structures and foundation by its trademark shape (character strokes are pretty much parallel) and shade consistency. (Epshtein *et al.*, 2010) is an area based content location system. It takes after the stroke width consistency supposition, which states that stroke widths stay consistent all through individual content characters. In the wake of getting an edge guide of a data picture, SWT strategy places sets of parallel edge pixels in the accompanying manner: for each one edge pixel p an inquiry beam in the edge inclination

course is genera ted, and the first edge pixel *q* along the hunt beam is found. In the event that *p* and *q* have about inverse angle bearings, an edge pair is structured and the separation in the middle of *p* and *q* (called stroke width) is figured. All pixels lying on the inquiry beam in the middle of *p* and *q* (counting *p* and *q*) are allocated relating stroke width. In the wake of doling out stroke widths to all picture pixels, the SWT strategy bunches pixels with comparable stroke widths into associated segments and channels out those that disregard geometrical properties of the content. At the point when the edge limit is sufficiently low, SWT ordinarily discovers all characters in the picture or in any event little partitions of each of them.

Text in scene pictures is recognized from other picture structures and foundation by its trademark shape (character strokes much parallel) and color consistency. Epshtein *et al.* (2010) is an area based content recognition technique.

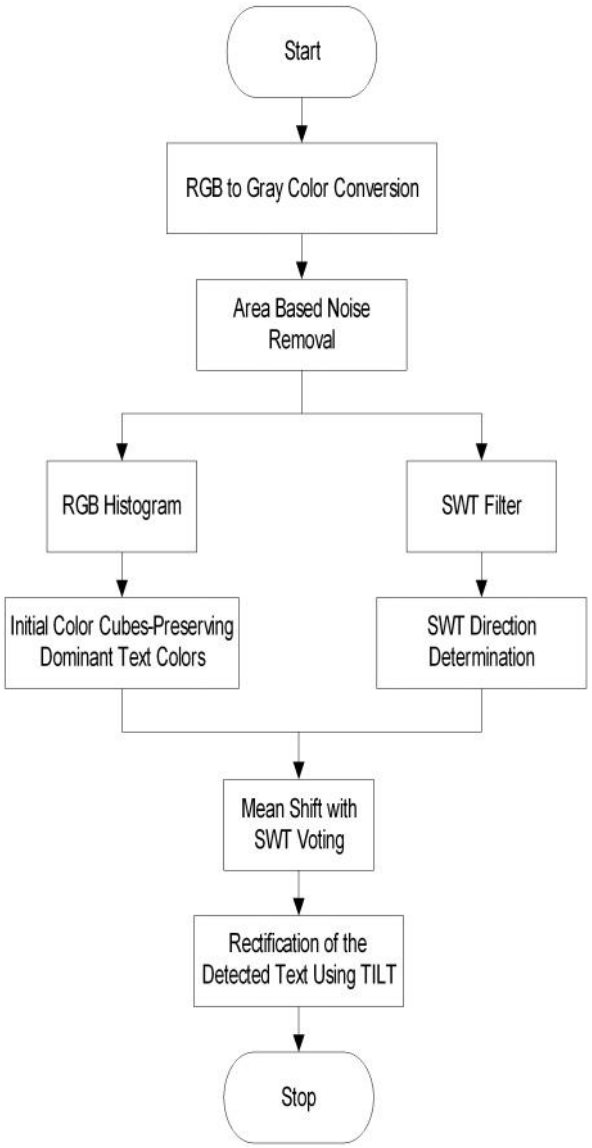
Text in regular scene pictures is recognized from other picture structures and foundation by its trademark shape (character strokes are pretty much parallel) and color consistency. [8] is an area based content recognition technique. It takes after the stroke width steadiness presumption, which states that stroke widths stay consistent all through individual content characters. In the wake of getting an edge guide of an info picture, SWT technique places sets of parallel edge pixels in the accompanying manner: for each one edge pixel p a pursuit beam in the edge inclination course is genera ted, and the first edge pixel q along the hunt beam is found. In the event that p and q have about inverse angle headings, an edge pair is framed and the separation in the middle of *p* and *q* (called stroke width) is After getting the SWT picture, all pixels with non-zero SWT qualities are mapped to the RGB shade space utilizing SWT lookup table Within request not to mean-movement genuine content shades, SWT voting is performed at every mean shift

Figure 1: Block Diagram of Proposed System



emphasis. On the off chance that the source and target blocks accept high and low amounts of SWT votes, individually, the shade is presumably leaving the safe content zone and is hindered from moving any further. In different cases, mean-moving is permitted. In place for SWT voting to work rightly, it is vital that SWT picture compares to genuine content characters. To manage both dim and light content situations, the first SWT execution runs the entire content discovery flowchart twice—in slope and counter-angle bearings—and consolidates the after effects of both headings. Consequently, we propose a SWT bearing determination strategy, which gives right SWT picture to SWT voting stage. We actualized the Nikolaou shade decrease system with changes. The introductory solid shapes are not chosen haphazardly as in Nikolaou and Papamarkos (2009) however in the accompanying way: RGB histogram receptacles are sorted in sliding request and starting 3d square focuses are constantly chosen from the highest point of the unvisited canisters rundown. Later stage we apply tilt calculation to get back just the content area from unique picture in frontal perspective point alongside shade.

Figure 2: Proposed System Flowchart



OUTLINE AND IMPLEMENTATION

Figure 2 shows the procedure stream outline of the entire framework which starts with the info picture. At first the information color picture is changed over to ash scale then the picture is at the same time given as info to shade decrease process and stroke width change process. The yield of both the procedure are consolidated and dispensing with of non content area is carried out focused around swt voting. Here at this stage color shapes shaped for a picture just the genuine nature (content colors) are held are matched with stoke width result and non message areas are mean moved focused around swt voting. The yield acquired at this stage is color picture holding just content district however the non caught content districts are still present which are redetected utilizing texture invariant low rank transform.

EXPERIMENTS

In this section, we present extensive experiments to verify the effectiveness of our system in extracting arbitrarily oriented texts in natural images, in comparison with many of the state-of-the-art systems and methods. In addition, we verify empirically how the results of our system could significantly impact on the performance of OCR engines. A commonly adopted way to evaluate algorithms is to use the f-measure, which is a combination of precision and recall. They are defined as follows:

Precession = $TP/(TP + FP)$

Recall = $TP/(TP + FN)$

fMeasure = (2 * Precission * Recall)/ (Precision + Recall)

where *TP* is true positive, *FP* is false positive and *FN* is false negative.

This dataset consists of text areas detected by SWT. SWT mainly detects texts that are nearly horizontal or with slight distortion.

This dataset consists of text areas detected by method (Chen and Wu, 2009). This method mainly detects horizontal texts and can also detect slight distorted text regions.

This dataset consists of text areas detected by Alex Chen's (Chen and Yuille, 2004) method.

Figure 3: Epshtein Dataset

Figure 4: Yi Dataset

Figure 5: Alex Chen Dataset

Figure 6: Unrectified Dataset

Figure 7: Rectified Dataset

This dataset consists of texts detected by our method in the original images. Most of them are not in upright position.

This dataset consists of the corresponding rectified text outputs from our method. Almost all are correctly rectified to their upright position.

Table:1: Performances of Different Text Detection Methods

Algorithm	Precision	Recall	f-measure
Our System	0.8333	0.9090	0.8695
Xin Zhang	0.7956	0.8299	0.8124
Epshtein	0.5501	0.6164	0.5814
Yi	0.3624	0.5686	0.4427
Alex Chen	0.3658	0.3254	0.3444

CONCLUSION

Our framework significantly beat existing content discovery frameworks on datasets that hold commonsense, uncontrolled pictures from the Internet or taken by cell telephones. Limits in the SWT voting condition focus affectability of the content/non-content move identification in the mean-movement stage. On the off chance that the edges are loose, considerably more content colors could be protected, however for this situation content characters are regularly part into a few content shades. Our future work will subsequently concentrate on uniting content shades that fit in with the same content character.

REFERENCES

1. Chen X and Yuille A L (2004), in Proceedings of the IEEE Conference on Computer Visionand Pattern Recognition, Detecting a Reading Text in Natural Scenes, pp. 366-373, IEEE, Washington.
2. Chen Y L and Wu B F (2009), "A Multi-Plane Approach for Text Segmentation of Complex Document Images", *Pattern Recognition*, Vol. 42, No. 7, pp. 1419-1444.
3. Epshtein B, Ofek E and Wexler Y (2010a), in IEEE Conference on Computer Vision and Pattern Recognition, Detecting Text in Natural Scenes with Stroke Width Transform, June 13-18, pp. 2963-2970, IEEE, San Francisco, CA, Washington.
4. Epshtein B, Ofek E and Wexler Y (2010b), "Detecting Text in Natural Scenes with Stroke Width Transform", in *CVPR*, pp. 2963-2970.
5. Nikolaou N and Papamarkos N (2009), "Color Reduction for Complex Document Images", *Int. J. Imaging Syst. Technol.*, Vol. 19, No. 1, pp. 14-26.
6. Pan Y F, Hou X and Liu C L (2011), "A Hybrid Approach to Detect and Localize Texts in Natural Scene", *IEEE Trans. Image Process*, Vol. 20, No. 3, pp. 800-813.
7. Ye Q, Huang Q, Gao W and Zhao D (2005), "Fast and Robust Text Detection in Images and Video Frames", *Image and Vision Comput*., Vol. 23, No. 6, pp. 565-576.
8. Yi C and Tian Y (2011), "Text String Detection from Natural Scenes by Structure-Based Partition and Grouping", *IEEE Trans. on Image Process*, Vol. 20, No. 9, pp. 2594-2605.
9. Zagoris K, Chatzichristofis S A and Papamarkos N (2011), "Text Localization Using Standard Deviation Analysis of Structure Elements and Support Vector Machines", *EURASIP J. Adv. Signal Process*, Vol. 47, pp. 1-12.