Efficient Exploration of Text Regions in Natural Scene Images Using Adaptive Image Sampling

Ismet Zeki Yalniz(✉), Douglas Gray, and R. Manmatha

A9.com, Palo Alto, USA
{izy,douggray,manmatha}@a9.com

Abstract. An adaptive image sampling framework is proposed for identifying text regions in natural scene images. A small fraction of the pixels actually correspond to text regions. It is desirable to eliminate non-text regions at the early stages of text detection. First, the image is sampled row-by-row at a specific rate and each row is tested for containing text using an 1D adaptation of the Maximally Stable Extremal Regions (MSER) algorithm. The surrounding rows of the image are recursively sampled at finer rates to fully contain the text. The adaptive sampling process is performed on the vertical dimension as well for the identified regions. The final output is a binary mask which can be used for text detection and/or recognition purposes. The experiments on the ICDAR'03 dataset show that the proposed approach is up to 7x faster than the MSER baseline on a single CPU core with comparable text localization scores. The approach is inherently parallelizable for further speed improvements.

Keywords: Adaptive image sampling · Scene text detection · 1D maximally stable extremal regions (1D MSER) · Mobile applications

1 Introduction

Recent advances in digital imaging technology enable users to take high quality digital pictures and videos in their natural environments. One can use these images and videos for automating various tasks such as product search, automatic navigation, license plate detection and recognition, surveillance and helping elderly or disabled people to recognize their environment. The existence of text in scene images provides valuable information about the content of the image. The research question is how to effectively detect and recognize text in scene images and perform it in real time using mobile devices.

Commercial Optical Character Recognition (OCR) systems are reasonably accurate (i.e., over 95 %) for recognizing text in document images [23]. However, text detection and recognition accuracies are generally much lower for natural scene images. In ICDAR'15, most methods performed below 40 % with the exception of "AJOU" [10] and "Stradvision-1". Both of these methods were based on variants of the MSER algorithm followed by different grouping approaches [8].

G. Hua and H. Jégou (Eds.): ECCV 2016 Workshops, Part I, LNCS 9913, pp. 427–439, 2016.
DOI: 10.1007/978-3-319-46604-0_31

(a) Low color contrast (b) Illumination effects (c) Unconventional text layout

Fig. 1. Example scene images from the ICDAR'03 dataset ([12]).

Image blur, low resolution, low contrast, unconventional text layout, non-uniform background, lighting and perspective changes are among the factors which makes the problem challenging as seen in Fig. 1.

The most common approach for recognizing text in scene images is to localize each word and/or character in the input image and then classify each one of them independently [3,22]. In these approaches, the performance of the overall text recognition framework heavily depends on the success of the text localization module. It is desirable to detect all the bounding boxes reliably with a high recall prior to recognition. There are also methods where the text detection and recognition modules are integrated [15,21]. This is achieved by creating several hypotheses about the location and the content of the text and refining the results with the help of a character/word classifier. Integrated text localization and recognition schemes are typically much slower because of the character/word classification overhead and the overall success of the framework depends on the classifier's accuracy. More recently, deep neural networks have also been successfully used for end-to-end recognition of text in natural images [7]. Deep neural network approaches are not elaborated further, since the primary focus of this paper is to provide real-time performance for effective text-detection on mobile devices with minimal use of computational resources.

In the case of text localization task, the aim is not to recognize the text but reliably find the bounding boxes for each word and/or character in the input image. Sliding window based approaches have been widely used for this purpose [2,3,11]. The main problem with the sliding window approaches is the computational overload which is not desirable especially for mobile applications. The localization accuracy also depends on the sliding interval which defines the total number of windows to be tested. In total, there are $O(n^2)$ number of candidate windows for an input image of size n pixels. More recently pixel grouping/merging approaches have been shown to provide the best text localization results with lower computational load [5,14,15,17,25].

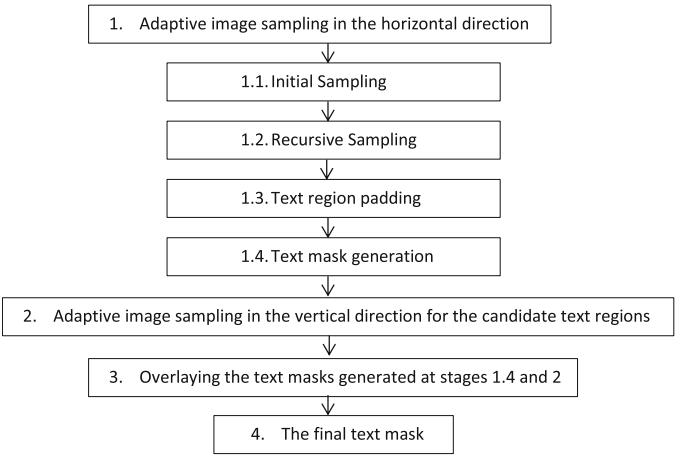


Fig. 2. The stages of the proposed image sampling framework for efficient text detection.

Text localization algorithms effectively use several features such as pixel intensity and local gradient information to group pixels and form candidate character glyphs for text detection purposes. Maximally Stable Extremal Regions (MSERs) [13] is currently the state-of-the-art approach to find text regions in scene images [14,15,24]. In a nut-shell, MSERs correspond to image blobs (i.e., connected components) in gray level images with the restriction that the size of the connected component is stable across several intensity thresholds. In the case of text localization, character glyphs are likely to be detected by the MSER approach, if there is sufficient contrast between the text and its background. The winning approaches in the ICDAR 2011, 2013 and 2015 competitions use MSERs to localize text in scene images [8,9,18]. The primary focus of our paper is, therefore, to improve upon the speed of the original MSER approach without losing its effectiveness. Text detection and recognition speed has prime importance especially for real-time mobile applications, where computational resources and battery power are always at premium [20].

In this paper, an adaptive image sampling framework is proposed for finding text regions efficiently in natural scene images. The observation is that most of the pixels in an input image do not belong to any text region and they can be eliminated at the early stages for efficient text detection. The proposed framework makes two assumptions for simplification purposes: (i) the text is composed of at least three characters, (ii) the text is aligned horizontally without significant skew. The requirement of at least three characters is a common assumption which is made by most text detection frameworks to eliminate false positives. The horizontal alignment of text is another general assumption which is also inherent in the ICDAR competition datasets.

The first stage of the proposed framework is to convert the input image to gray scale and sample a number of rows of pixels. The sampling frequency

is determined according to the height of the smallest possible character in the input image. Each row is independently tested for containing any text or not. If the row is determined to contain any text, the rows nearby are also sampled in a recursive manner until the character glyphs are fully contained. The same image sampling process is applied to the vertical dimension as well for the candidate regions which are determined to contain text in the previous stage. The final output is a binary image where the candidate regions of texts are marked. The entire process is repeated for detecting text in the other image polarity as well. Figure 2 shows a flow chart of these stages.

The text region proposal results on the most widely used and cited ICDAR'03 dataset show that the effectiveness of the proposed approach is comparable to the most efficient linear time implementation of the original 2D MSER approach [16] while providing up to 7x speed improvement on a single CPU core. An unoptimized single threaded implementation of the proposed framework provides 20 frames per second even on a relatively old cell phone with an Apple A4 single core processor (released in 2010). Therefore the proposed framework can be used almost on any smart phone being used in the market as of today in 2016. Indeed, the proposed text region proposal algorithm is inherently parallelizable. Each row or columns of pixels can be independently evaluated for containing text without any need for a global priority queue or union-find data structure. These data structures are actually used by the most efficient implementations of the original MSER algorithm [13,16]. Parallelization of the original MSER algorithm is therefore not trivial and the speed gains might be bounded because of the global data structures being maintained. To our best knowledge, there is no publicly available GPU/parallel implementation of the original MSER algorithm. Our evaluations are constrained to the single threaded execution model in the rest of the paper.

The contributions of the paper may be summarized as follows:

(a) An adaptive image sampling methodology to filter out non-text regions without the need for evaluating each pixel in the input image.
(b) 1D adaptation of the original MSER algorithm for efficient image sampling and text detection.
(c) A text region detection framework which
  (i) is comparable to the original MSER algorithm in terms of effectiveness,
  (ii) provides up to 7x speed improvements on a single CPU core,
  (iii) is inherently parallelizable for further speed improvements.

The rest of the paper is organized as follows: the proposed adaptive image sampling framework is elaborated first in Sect. 2. The experiments are discussed in Sect. 3. Conclusions are drawn in Sect. 4.

2 The Adaptive Image Sampling Framework for Efficient Text Detection

The proposed adaptive image sampling framework operates on gray scale images to find text regions. The first stage is to determine the location of the character glyphs on the vertical axis by sampling rows of pixels in the input image.

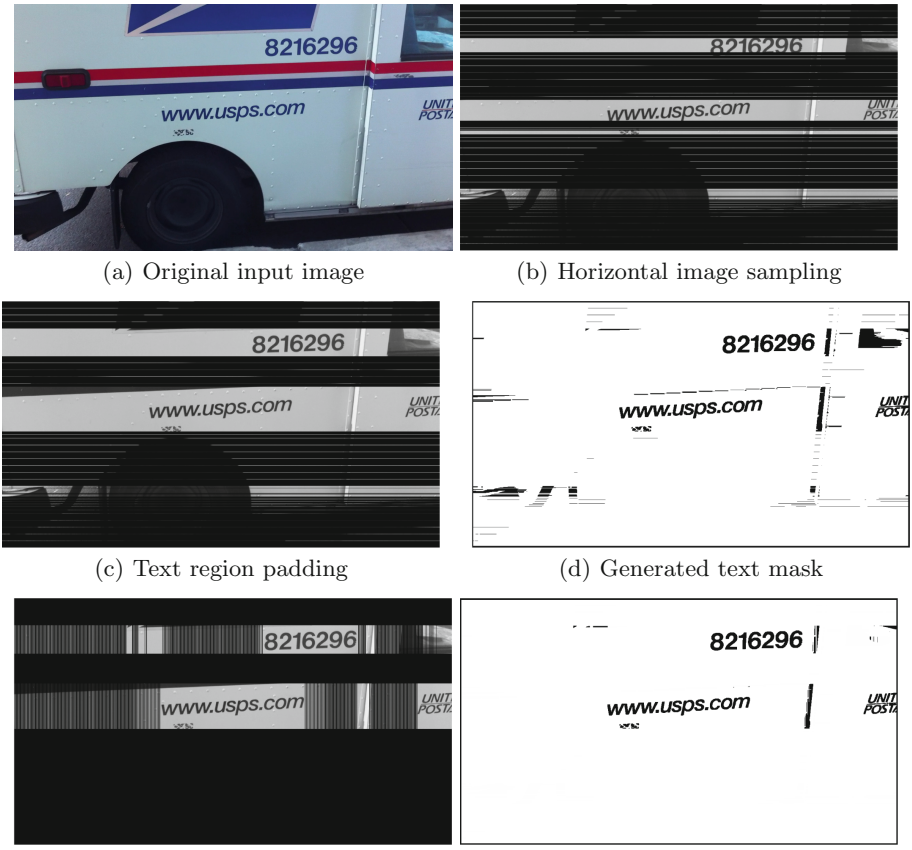
Each sampled row is tested for containing any text by analyzing the position and length of connected components in the row image across different intensity thresholds. In this particular case, the connected components are extracted by generalizing the original MSER approach for 1D images. The surrounding rows of pixels which are determined to contain text are recursively sampled as well to fully recover the text. At this point, all the connected components of each sampled row are encoded one by one on a 2D image mask output. The same sampling process is applied in the vertical direction as well for the regions which are determined to contain text in the horizontal sampling stage. Those connected components are overlaid on the 2D image output using the AND logical operator. The resulting 2D image contains candidate character glyphs. The same process is applied to the other image polarity as well to detect text written in different polarities (i.e., dark text against lighter background or vice versa). Figure 3 illustrates the stages of the proposed approach. The details are given in the following subsections.

2.1 Initial Sampling

The aim of the initial sampling stage is to find the approximate location of the text and ensure that all the character glyphs satisfying certain size constraints can be fully reconstructed. This is achieved simply by sampling the input image uniformly at a certain rate. According to the "Nyquist-Shannon" sampling theorem [19], the sampling frequency must be at least twice the highest frequency object in the image in order to detect and/or reconstruct the original signal. In the case of text detection, the highest frequency objects are the smallest characters of interest in the input image. The sampling interval in the horizontal and vertical direction are, therefore, set to 10 and 2 pixels respectively for any input image of size 640×480. In this way, any character of height 20 and width 4 and above can be reliably detected.

2.2 Detecting Candidate Text Regions

The task is to classify the line represented by a row of image pixels whether it passes through any text region or not in the input image. The problem is not trivial because rows of pixels lack two dimensional contextual information. Our observation is that the change in the intensity values along the line gives clues about the existence of the text. More specifically, the intensity values increase rapidly and stay steady for a period and then fall back again to the intensity value of the text background. These "extremal regions" are detected by adapting the original MSER algorithm for 1D images. The original formal definition of 1D MSER regions [13] holds for the 1D images as well. The difference is that, 1D extremal regions are characterized by their length, start and end coordinates. The area of a given 1D extremal region is defined to be its length in the 1D image. Notice that MSER regions can be extracted from 1D images quite efficiently because of increased spatial locality and simplicity of 1D MSER regions. In this work, the 1D MSER regions are extracted separately for each image polarity (i.e.,



(a) Original input image (b) Horizontal image sampling

(c) Text region padding (d) Generated text mask

(e) Vertical image sampling (f) Final text mask for the given image polarity

Fig. 3. The proposed adaptive image sampling scheme is illustrated for one image polarity of the input image (dark text against lighter background). Pixels which are not visited in stages (b), (c) and (e) are colored black.

negative and positive regions). For further speed improvements, it is possible to compute 1D MSER regions for both polarities at the same time with negligible computational overhead. Notice that the pixel gaps between extremal regions correspond to extremal regions in the other polarity for 1D images.

Having at least one MSER region is not a sufficient criterion for a row of image pixels to contain text. The rows are likely to have extremal regions even though the image does not contain any text. Therefore MSER regions with length less than two pixels and more than one sixth of the width of image are eliminated first. Next, a chaining mechanism is devised to further eliminate non-text regions. The key observation is that, if there exists any row which passes over a text region with at least three characters, then there must be at least

three 1D MSER regions which are relatively close to each other. Two 1D MSER regions are assumed to be close by if the distance between them is smaller than the width of the preceding region multiplied by a constant (trained empirically to be 8). This analysis is done for every row sampled in the initial sampling stage and the rows which are determined to contain text are forwarded to the recursive sampling stage.

2.3 Recursive Sampling Stage

The output of the initial sampling stage is a number of rows which are likely to contain text. Here the aim is to efficiently sample rows of those regions at finer levels and reconstruct the character glyphs. It is achieved as follows: the two adjacent image regions bounded by the previous and next sampled rows are added to a stack for further sampling. This is done for all rows in the initial sampling stage if they are identified to contain text. The stack now contains a list of intervals which designate the start and end coordinates of the image regions on the vertical axis. Each interval in the stack is popped and the row in the center of the interval is tested for containing any text. If the result is positive, then the two sub regions divided by the center row are added to the stack as well. This is done recursively for all sampling intervals in the stack until the sampling interval contains a single row. At the end of the iterations, all the image regions which contain text are sampled horizontally and the regions which are very unlikely are not sampled. The output of this stage is a list of labels for each row indicating whether it is sampled or not. This is forwarded to the text region padding stage.

2.4 Text Region Padding

Additional sampling might be necessary to reconstruct the character glyphs. For example, "pad" is a word where there is only one connected component along the rows passing over the upper part of letter "d" and the lower part of letter "p". Since there are less than three regions along those rows, certain parts of the letters are classified as non-text during horizontal adaptive sampling stage. As a result, those characters are partially detected. This is not desirable because the aim is to fully reconstruct the character glyph. The solution is to pad around the regions where the rows are densely sampled. For the case of the word "pad", only the middle zone is classified as text in the recursive sampling stage because there are at least three 1D MSER regions in chain along each row in the respective region. Given the height h and start position s of each dense sampled region, all the rows positioned between $(s - h/2)$ and $(s + 3h/2)$ are also sampled and 1D MSER regions are calculated if they were not sampled before.

2.5 Text Mask Generation

1D MSER regions extracted from the sampled rows are used to generate a 2D binary mask which has the same size as the input image. This is simply done by

printing all the 1D MSER regions of all the sampled rows on the respective row of the output image. If there are overlapping 1D MSER regions, they are printed on top of each other. This is actually equivalent to printing only the largest one among others. The output image contains extremal regions which are maximally stable along the horizontal direction.

2.6 Adaptive Image Sampling Along the Vertical Direction

The character glyphs exhibit similar characteristics when the pixels are sampled in the vertical dimension as well. In other words, the pixel intensity values increase and stay steady for a period if the column of pixels passes over a character glyph. In order to eliminate false positive text regions identified in the horizontal sampling stage, the adaptive image sampling is performed on the vertical direction as well for the candidate text regions. The sampling interval is determined to be half of the minimum width of a character which is set to 2 pixels. The column is classified to contain text if there exists at least one 1D MSER region without chaining the connected components. At the end, a mask image is generated using the extracted 1D MSER regions. Text masks obtained from the horizontal and vertical image sampling stages are combined simply using the logical AND operator over corresponding pixel values.

3 Experiments

The experiments are performed on the ICDAR'03 Robust Text Reading dataset. The dataset includes 251 scene images containing 1097 words and 5411 characters in total. The effectiveness of the proposed adaptive image sampling framework is compared to the fastest open source MSER implementation [1] which uses a linear time algorithm introduced by [16]. The evaluation metrics and the experimental results are discussed in the following subsections.

3.1 Evaluation Metrics

The aim is to evaluate the effectiveness of the proposed adaptive image sampling scheme for extracting the character glyphs from scene images. From the image sampling perspective, a high recall is desirable for recovering the character glyphs since the overall success of the text detection framework depends primarily on the success of the sampling stage. The evaluation measure therefore accounts for the overall amount of the character glyphs extracted from the test images and their localization accuracy. Given the ground truth bounding box t of a single character in the test image, the localization accuracy is determined by:

$$f(t,r) = \frac{|t \cap r|}{|t \cup r|} \tag{1}$$

where r is the estimated bounding box and the $|.|$ operator corresponds to the total number of pixels in the set respectively. If the two boxes fully overlap, the

localization score is equal to 1.0. If there is no overlap between the two boxes, then the score is zero. Given a set of estimated bounding boxes R, the best matching box r_t is defined as

$$r_t = arg \max_{r \in R} f(t, r) \tag{2}$$

for a ground truth character t in the test image. Precision and recall scores are computed for each character in the scene image as:

$$precision = \frac{|t \cap r_t|}{|r_t|}, recall = \frac{|t \cap r_t|}{|t|} \tag{3}$$

where the average precision and recall scores over all the characters in the ground truth set are reported. It should be noted that this is a slightly modified version of the evaluation scheme used in ICDAR competitions focusing on the success of the glyph extraction task prior to any classification.

3.2 Evaluation

The experiments are performed for various settings of the proposed adaptive image sampling framework. In the first experiment, the vertical sampling is skipped and the character glyphs generated from the horizontal adaptive sampling stage are used for evaluation. The second experiment is designed to investigate the success of the adaptive sampling approach over dense sampling. Dense sampling is run in two modes. In the first case, all the rows in the input image are sampled and all the extracted 1D MSER regions are encoded in to the final mask output if they satisfy the length constraints discussed in Sect. 2.2. In the other case, the task masks are generated along the horizontal and vertical direction independently in the same way and the output masks are merged using the AND logical operator.

The baseline is the original 2D MSER approach [13]. The MSER parameters are independently trained for each system using the training set. The input images are not downsampled at any point. The sampling parameters are automatically adjusted for each image based on the image height. Notice that the maximum variation parameter for the 1D MSER is calculated over the length of 1D connected components whereas the original 2D MSER algorithm uses the area of the 2D connected components. In the experiments, the other MSER parameter delta is varied from 6 to 20 with an increment of 1 for both 1D and 2D MSER. It should be noted that most applications operate in this particular interval for delta. There is no post processing or glyph classification stage involved in the experiments. The aim is to understand to what extent the proposed approaches are able to extract character glyphs from the input images. The F-measure is defined as:

$$\frac{2 \times precision \times recall}{precision + recall} \tag{4}$$

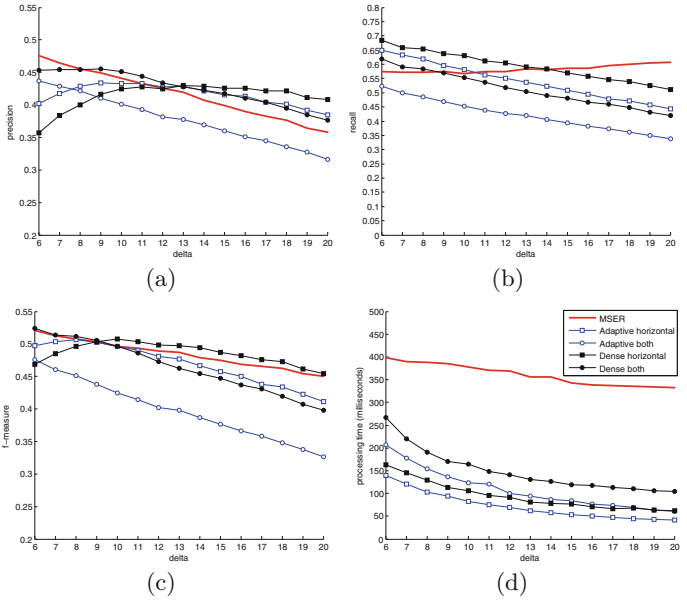
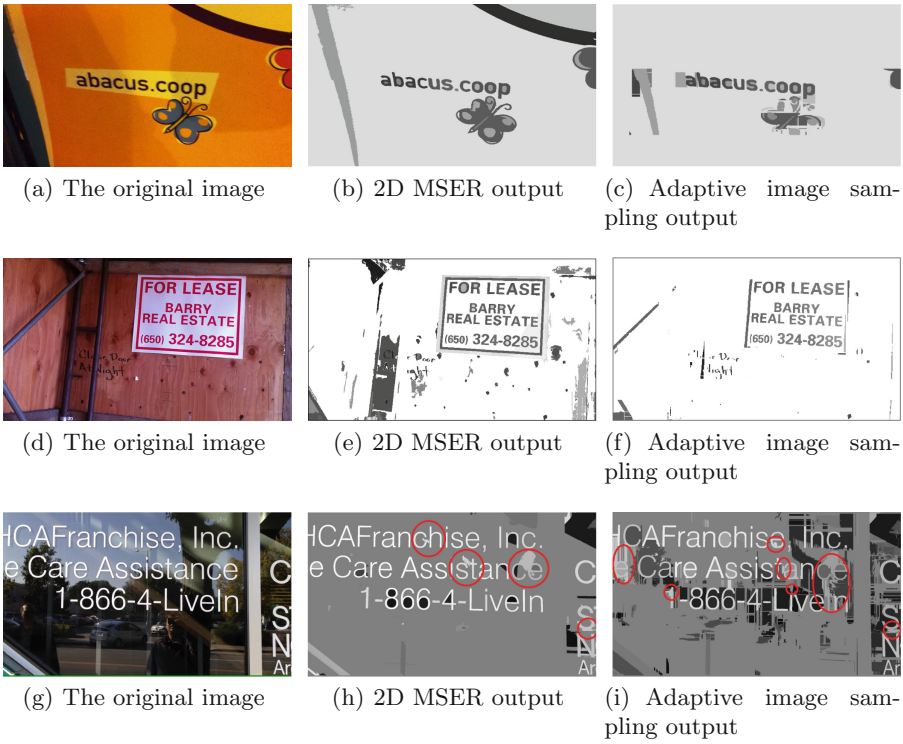


Fig. 4. Precision (a), recall (b), F-measure (c) and processing time plots (d) are given for different configurations of the adaptive image sampling framework in comparison with the original 2D MSER approach (baseline). Delta is the MSER parameter used for calculating the area stability of the region across different intensity levels. Depending on the application, it is typically set to be between 5 to 20 and it is set to 9 for the proposed application.

Figure 4 plots the localization scores, F-measure and processing time for all the configurations and the baseline for varying values of delta. It is clear that localization scores drop significantly if the vertical sampling stage is active for both adaptive and dense sampling tests. The dense sampling approach without vertical sampling provides the highest F-measure for values of delta above 8. The adaptive sampling approach with only horizontal sampling is the fastest method among others (up to 7x) providing F-measure scores comparable to the baseline. The speed gain becomes more drastic for increasing values of delta. The timing experiments are performed on a 32-bit operating system with an Intel i5 processor at 2.4 GHz. These approaches can be thought as different operating points in the space of precision, recall and processing time.

Figure 5 shows a number of examples for qualitative evaluation. The images are generated by coloring each connected component by averaging their pixel values in the original image. For the first two examples, all the character glyphs are extracted correctly in both images. The background of the text scene may look noisy due to sampling but this is not an issue for the task of text detection. In the last example, glyph extraction errors are shown with red circles for both the baseline 2D MSER output and the proposed approach. The errors are due



(a) The original image (b) 2D MSER output (c) Adaptive image sampling output

(d) The original image (e) 2D MSER output (f) Adaptive image sampling output

(g) The original image (h) 2D MSER output (i) Adaptive image sampling output

Fig. 5. The outputs of the 2D MSER and the proposed adaptive image sampling framework are visualized for three sample images. The glyph extraction errors are shown with red circles for the last example.

to connected glyphs and such errors are also expected under low contrast and image resolution for the MSER algorithm.

4 Conclusions

An efficient adaptive image sampling framework is presented for exploring text regions in natural scene images. The proposed approach exploits the fact that a small portion of pixels in the input images actually correspond to text regions and non-text regions should be avoided for efficient text detection. The sampling starts by sampling rows of pixels from the input image at a specific rate. Regions which may correspond to text regions are further sampled recursively until the text is fully contained. The same process is applied in the vertical dimension as well for the candidate text regions. The experiments on the ICDAR'03 dataset show that the effectiveness of the proposed approach is comparable to the MSER approach with a significant speed improvement. The proposed framework runs in real time on a mobile phone with an Apple A4 processor. It is expected to speed up the-state-of-the-art text detection and recognition systems with comparable

accuracy. The proposed approach is inherently parallelizable and further speed improvements are possible with an optimized parallel implementation.

Maximally Stable Extremal Regions have been used for several other computer vision tasks such as object detection and tracking [4,6]. Efficient implementations of MSERs is expected to improve such MSER-based object tracking and recognition approaches. Future work includes (i) optimizing the adaptive image sampling framework for further speed gains with a parallel GPU implementation, and, (ii) adapting it for other image recognition and object tracking tasks.

References

1. Bradski, G.: The OpenCV library. Dr. Dobb's J. Softw. Tools. **25**(11), 120–129 (2000)
2. Chen, X., Yuille, A.L.: Detecting and reading text in natural scenes. In: CVPR, pp. 366–373 (2004)
3. Coates, A., Carpenter, B., Case, C., Satheesh, S., Suresh, B., Wang, T., Wu, D.J., Ng, A.Y.: Text detection and character recognition in scene images with unsupervised feature learning. In: ICDAR, pp. 440–445 (2011)
4. Donoser, M., Bischof, H.: Efficient maximally stable extremal region (MSER) tracking. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2006, vol. 1, pp. 553–560. IEEE (2006)
5. Epshtein, B., Ofek, E., Wexler, Y.: Detecting text in natural scenes with stroke width transform. In: CVPR, pp. 2963–2970 (2010)
6. Gómez, L., Karatzas, D.: Mser-based real-time text detection and tracking. In: 2014 22nd International Conference on Pattern Recognition (ICPR), pp. 3110–3115. IEEE (2014)
7. Jaderberg, M., Simonyan, K., Vedaldi, A., Zisserman, A.: Deep structured output learning for unconstrained text recognition. In: ICLR (2015)
8. Karatzas, D., Gomez-Bigorda, L., Nicolaou, A., Ghosh, S., Bagdanov, A., Iwamura, M., Matas, J., Neumann, L., Chandrasekhar, V.R., Lu, S., et al.: ICDAR 2015 competition on robust reading, pp. 1156–1160 (2015)
9. Karatzas, D., Shafait, F., Uchida, S., Iwamura, M., i Bigorda, L.G., Mestre, S.R., Mas, J., Mota, D.F., Almazn, J., de las Heras, L.P.: ICDAR 2013 robust reading competition. In: ICDAR, pp. 1484–1493. IEEE Computer Society (2013)
10. Koo, H.I., Kim, D.H.: Scene text detection via connected component clustering and nontext filtering. IEEE Trans. Image Process. **22**(6), 2296–2305 (2013)
11. Lee, J.J., Lee, P.H., Lee, S.W., Yuille, A.L., Koch, C.: Adaboost for text detection in natural scene. In: ICDAR, pp. 429–434 (2011)
12. Lucas, S., Panaretos, A., Sosa, L., Tang, A., Wong, S., Young, R.: ICDAR 2003 robust reading competitions. In: ICDAR, pp. 682–687 (2003)
13. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. Image Vis. Comput. **22**(10), 761–767 (2004)
14. Neumann, L., Matas, J.: Text localization in real-world images using efficiently pruned exhaustive search. In: ICDAR, pp. 687–691 (2011)
15. Neumann, L., Matas, J.: Real-time scene text localization and recognition. In: CVPR, pp. 3538–3545 (2012)

16. Nistér, D., Stewénius, H.: Linear time maximally stable extremal regions. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5303, pp. 183–196. Springer, Heidelberg (2008). doi:10.1007/978-3-540-88688-4_14
17. Pan, Y.F., Hou, X., Liu, C.L.: Text localization in natural scene images based on conditional random field. In: ICDAR, pp. 6–10 (2009)
18. Shahab, A., Shafait, F., Dengel, A.: ICDAR 2011 robust reading competition challenge 2: reading text in scene images. In: ICDAR, pp. 1491–1496 (2011)
19. Shannon, C.E.: Communication in the presence of noise. Proc. Inst. Radio Eng. **37**(1), 10–21 (1949)
20. Takeda, K., Kise, K., Iwamura, M.: Real-time document image retrieval on a smartphone. In: 2012 10th IAPR International Workshop on Document Analysis Systems (DAS), pp. 225–229. IEEE (2012)
21. Wang, K., Babenko, B., Belongie, S.: End-to-end scene text recognition. In: ICCV, pp. 1457–1464 (2011)
22. Wu, V., Manmatha, R., Riseman, E.M.: Textfinder: an automatic system to detect and recognize text in images. IEEE Trans. Pattern Anal. Mach. Intell. **21**(11), 1224–1229 (1999)
23. Yalniz, I.Z., Manmatha, R.: A fast alignment scheme for automatic ocr evaluation of books. In: ICDAR, pp. 754–758 (2011)
24. Yin, X.C., Yin, X., Huang, K., Hao, H.: Robust text detection in natural scene images. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **36**(5), 970–983 (2014)
25. Zhang, J., Kasturi, R.: Character energy and link energy-based text extraction in scene images. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010. LNCS, vol. 6493, pp. 308–320. Springer, Heidelberg (2011). doi:10.1007/978-3-642-19309-5_24