Racing Bib Number Localization on Complex Backgrounds

NOPPAKUN BOONSIM
Faculty of Applied science and Engineering
Khon Kaen University
112 Village No. 7 Nong Khai Campus, Khon Kaen University, Nong Khai 43000
THAILAND
boonsim@kku.ac.th

Abstract: - Racing bib number localization in marathon natural images is challenging because in those images, there are many texts appearing in the scene that will produce many false positive texts. This research presents a method to detect racing bib numbers on complex backgrounds based on edge detection technique. The algorithm initially extracts candidate texts by applied vertical edge detection and morphological operations. Then, face detection technique is applied for the verification step. The experiments were tested on over 400 marathon images. It was reported satisfactory by the performance that was improved from the original work.

Key-Words: - Racing bib number, Localization, Complex background

1 Introduction

Nowadays, wellness trends are more prominent around the world. Therefore, a number of marathon or distance running race events have been created and sparked interest from many people of all ages and genders. The runners or participants have a unique racing bib number (RBN) to identify themselves and this bib number appears on heterogeneous backgrounds and different materials. In running events, many images of runners are dramatically increased from organizers and photographers attending the events. It raises the problem of distinguish individual marathon images of a runner from all available marathon images. As organizers of the distance running race point of view, the security of the running events is important. Thus, videos and images of the running events should be kept and available to be retrieved. Photographers require this method to classify and identify participant's images in order to retrieve a particular image. The runners or participants also encounter difficulties in searching for themselves, their relatives and friends images. To solve these problems, image processing and computer vision techniques are employed in order to recognize the RBN which are printed on a paper or cardboard tag and pinned onto the runner's T-shirt before the race [1]. In addition, there are a few research studies applying image analysis in marathon events. It is an opportunity to experience the image analysis techniques for marathon images and events.

RBN recognition system contains three main processes: RBN detection, character segmentation and RBN recognition based on the state-of-the-art of text recognition [4]. First, RBN detection is a process to locate the RBN position in an image. After the RBN is detected, the character segmentation is employed to separate each character. Finally, each character is recognized and then shown the result. RBN detection is the first process in the pipeline and is considered a crucial process of the RBN system because the performance of RBN recognition depends on the accuracy of RBN localization.

In a review of text recognition works [4], many researches presented text recognition especially in natural settings. There are two works dedicated to recognize RBN. The first work was presented by Ami et al. [1] and they used the stroke width transform method [5] to extract characters in an image. The characters with similar stroke width are grouped for creating text region. After that, face detection is implemented and also locates the runner's torso. Finally, the given torso is used to verify the true RBN. The second work, Shivakumara et al. [2] presented a combining technique. The technique starts with torso detection. The considered output from torso detection are candidate texts. After that, there are two classifications. SVM is applied to classify the true RBN in both two classifications. In the first classification, text features are extracted by three feature detectors: Fourier, Polar and Zernike feature detectors. The second classification, the remaining texts are extracted with HOG features. Roy et al. [3] proposed a method called multi-modal to detect bib number in natural images.

This method utilized a color-based feature to locate text candidate regions and after that, extracted wavelet and color features of the given candidate regions. K-mean clustering technique was used to verify the potential candidate regions. Binarization technique [6] and Tesseract OCR [7] were applied to identify and recognize a bib number, respectively. There are some problems for a RBN detection process in marathon images. First, there are many other texts in the scene increasing the problem of locating the real RBN. Second, the captured images in natural scene may lose focus or be blurred with low contrast. Finally, uneven light intensities can occur on natural images. As a result, the images have unbalanced illumination throughout images.
This research aims to present an algorithm to detect RBN in complex background images. RBN candidate regions is extracted by the edge based method and the image processing techniques. Face detection also is applied in order to locate the runner's body and it can be used to verify the true RBN position. The remainder sections are organized as follows: the proposed method and detail are given in section 2. Section 3 illustrates the experimental data, tools, results and discussions. The conclusion of this research is shown in section 4.

2 Methodology

RBN detection on complex backgrounds is challenging because of the heterogeneous backgrounds, different materials and text-like-shape (signs, trees and buildings) that appeared in the images. Therefore, they seem to produce more false alarms than simple background image. It leads to reduce RBN detection accuracy. To overcome this problem, the research presents the edge-based method and image processing techniques to extract candidate texts. Face detection of runner is used to locate runner's torso and then is applied to verify the true RBN. The proposed algorithm for detecting RBN, RBN candidate extraction process and RBN verification process in shown in figure 1.

2.1 RBN Candidate extraction

The goal of candidate extraction process is to detect potential texts which could be contained in the real RBN. The most outstanding appearances of RBN are characters (bib number) with the similar size and shape and a number of character within the tag. There are a number of techniques to detect text in images such as color, edge, point, region, stroke texture feature and character appearance features [4].

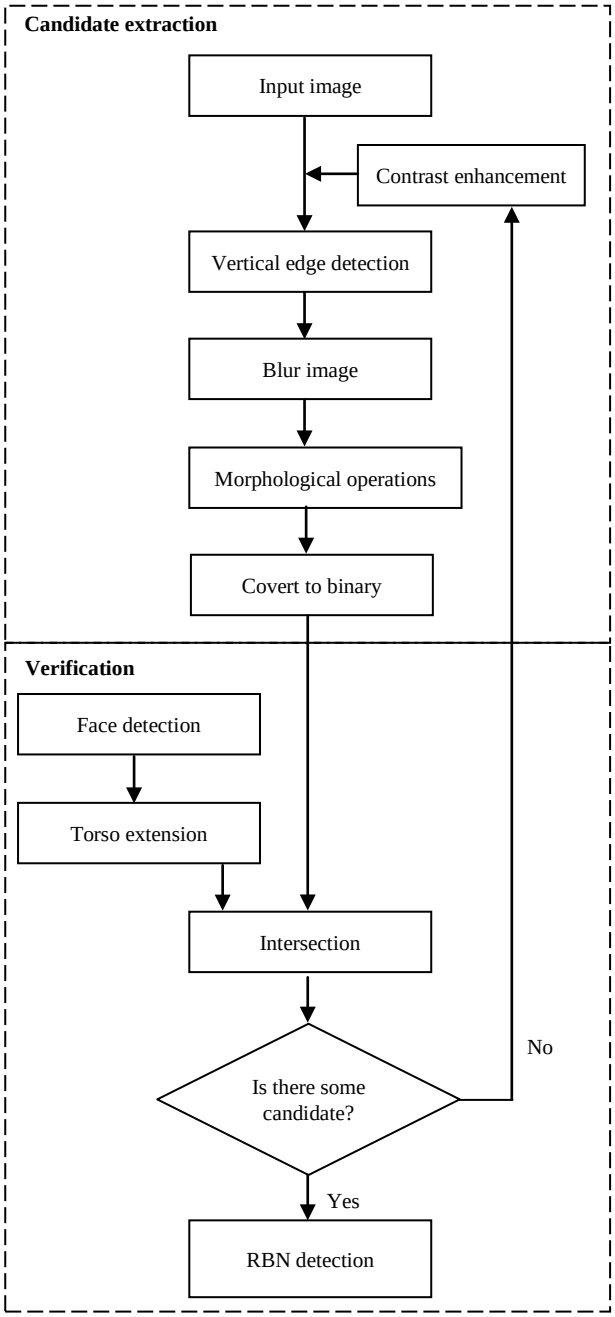
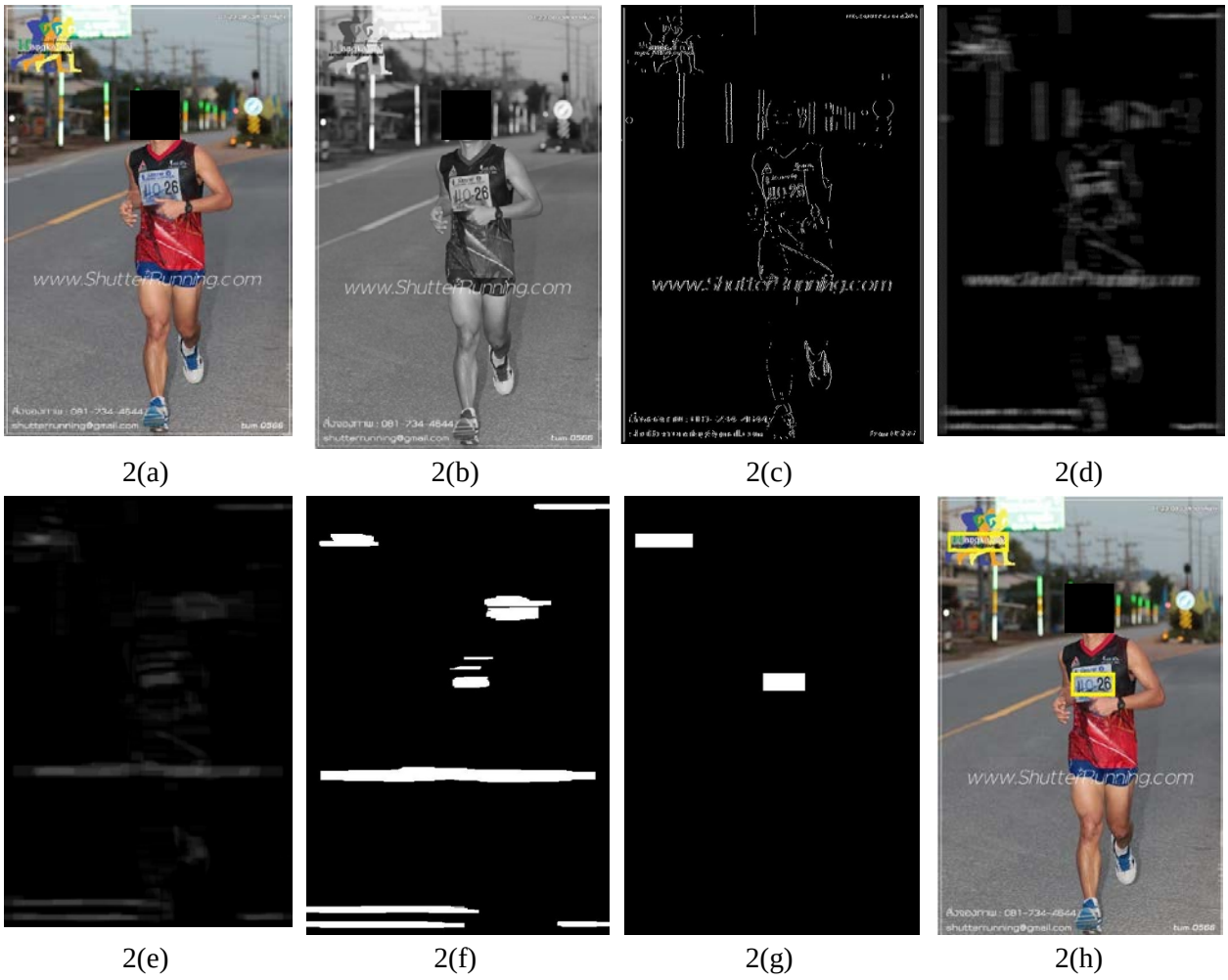


Fig.1 Algorithm for RBN localization

In this research, edge-based method is used to extract edges of an image and then mathematical morphological operations are applied to merge the edges for creating RBN region and faded small and large regions. After that, potential candidate regions are preserved by comparing to pre-defined RBN tag aspect ratio. As discussed, the RBN's character appearances have similar size and shape that can be used to extract the RBN tag.



2(a) 2(b) 2(c) 2(d)

2(e) 2(f) 2(g) 2(h)

Fig.2 Example images of candidate extraction step

Edge-based method is typically used to extract characters or texts in natural images [4] and the technique is reported rapidly computational time and high detection rate in license plate detection researches [8, 9].

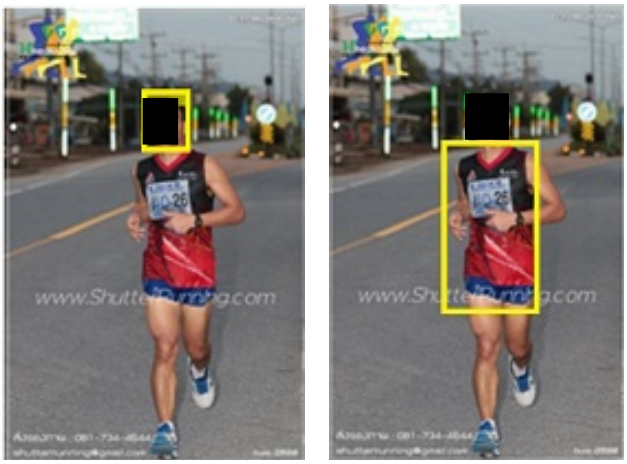
Edge-based technique contains numerous steps to extract RBN candidates and these steps are based on the work of Mendes et al. [10] which presented a method based on edge processing to detect license plates. The technique begins with converting RGB image, as shown in figure 2(a), to grey level image, given in figure 2(b). Then, Sobel's edge detection is applied to extract vertical edges of an image. Figure 2(c) shows vertical edges of image. Next, average filter method is employed to blur detected edges as shown in figure 2(d). The filter size is defined as 9 pixels height and 25 pixels width. After that, morphological operations are applied to manipulate the edges for creating RNB candidate region, to remove small and large regions. The morphological operations define thresholds as 9 pixels of minimum character height, 45 pixels maximum character height and 25 pixels maximum character distance, as shown in figure 3(e). Then, the image is converted to binary image by applying Otsu's method [11] for selecting automatic threshold, as given in figure 2(f). Finally, potential candidate regions are preserved by applying experimental RBN aspect ratio, between 1 and 6. Figure 2(g) shows final potential regions in binary image and figure 2(h) shows candidate regions in original image.

2.2 RBN Verification

With the given candidate, this step confirms all candidate to find the real RBN. As stated in previous section, RBN is normally pinned on the runner's T-shirt. Therefore, runner detection is used first and then the candidate region locates within the runner considering to the real RBN.

2.2.1 Runner detection

To detect runner, this research is designed to detect the runner's face first which is based on the algorithm of Viola et al. [12] and then can be extended to runner's body or torso. The algorithm uses HOG features of face training images and then trains and classifies by cascade of classifiers (weak classifiers). Figure 3(a) shows example image of face detection in the yellow bounding box. After face detection, face's position and size are obtained that can be used to locate runner's body. As the given face's width and height, it extends with 0.30 multiple face's width and 4.0 multiple face's height and locates under the runner's face, as shown in figure 3(b).



3(a) 3(b)

Fig.3 Example images of runner detection

2.2.2 RBN Verification

After runner detection, it can be used to verify the true RBN position. The verification process performs image intersection between the candidate regions image and runner's body image. The candidate regions image is obtained from the section 2.1 as can be seen in figure 4(a) and the runner's body image is given from the previous step (figure 4(b)). The result image of RBN position verification is presented in figure 4(c).

2.2.3 Contrast improvement

Marathon images are captured in natural scene which might be affected by uneven lights, lost focus, complex background and blur in RBN tag. Therefore, RBN is unable to detect by the proposed method. Image contrast enhancement could be employed in order to improve image contrast. In the research, local contrast enhancement strategy [13] is used to improve local contrast rather than global contrast because it experimentally deals with complex background images. The image is divided to 8 × 8 blocks and then applied contrast enhancement, histogram equalization, separately to each block. The algorithm is first applied to original image. If there is no the detected candidate, the contrast enhancement is employed on the image and the process will start from the beginning.

4(a) 4(b)

4(c)

Fig.4 Example images of RBN detection

3 Experimental results and discussion

3.1 Experimental results

In the evaluation process, Image data set was taken from the Benchmark for Enhancing the Standard of Thai language processing (BEST) competition 2016 [14]. The data set contained 400 marathon images and covered variety of age, gender and background in daytime. The original size of image dimensions are 800 × 533 pixels. Additionally, Precision (P), Recall (R) and F-measure (F) are used for evaluating the performance of the proposed technique.

Table 1 Performance of the proposed approach and existing approach on bib number detection

Works	***Images***	***Technique***	***Time(s)***	***P***	***R***	***F***
Ami et al.[1]	217	SWT + face	N/A	0.53	0.40	0.45
Shivakumara et al. [2]	212	Multi-modal	N/A	0.60	0.74	0.66
Roy et al. [3]	212	Multi-modal	N/A	0.76	0.55	0.64
Proposed work	400	Edge + face	~ 0.75	0.59	0.69	0.64

These metrics were employed in Ami et al. [1] where F-measure is defined by equation (1) to combine the precision and recall into a single measure.

$$F = \frac{P.R}{\propto.R+(1-\propto).P} \quad (1)$$

To compare to Ami et al. [1] as original work, therefore, this research also uses equation (1) to calculate F-measure value. The relative weight α is set to 0.5 [1]. The proposed technique was implemented on a computer with 2.5 GHz, Intel(R) Core i7-6500U CPU and 8GB RAM. The operating system installed on the computer was Windows 10 Pro 64-bit. MATLAB (version R2016b) is also used to implement the method. Image processing toolbox including edge detection technique, morphological operations, contrast enhancement connection component analysis and computer vision, face detection, are implemented. The experimental results of proposed work show the detected RBN 64% accuracy which is presented in the last row of table 1.

4 Discussion

A comparison of image data set, technique used, computational time and detection rate (precision, recall and f-measure) with other works, is shown in table 1. Ami et al. [1] presented stroke width transform to find text with similar stroke width within an image. Then, face detection is used to detect runner's face in order to locate runner's torso and then is used to find the RBN. The study was tested on 217 images and reported the detection accuracy, f-measure, at 45%. There was no computational time reported. To improve the accuracy, Shivakumara et al. [2] presented multi-modal technique including many steps as given in the introduction section. The detection accuracy was reported at 66%, implemented on 212 images.

Although, there are a number steps to detect the real RBN, but there is no time computation reported. Next, Roy et al. [3] also proposed a new multi-modal technique to detect bib number in natural images. They tested algorithm on data set in [2] and was reported detection accuracy at 64%. As we known, the color based method is very sensitive to the lighting conditions.

For the proposed algorithm, it uses the combined techniques: edge based method, image processing techniques, morphological operations and image contrast enhancement. Runner detection is employed to verify RBN position. The experiments were tested with 400 images and the computational time was reported 0.75 seconds (less than a second). The detection accuracy was reported at 64% which is a lot improved from [1], 45%, and a bit lower than [2], 66%.

The disadvantage of this method appeared for example, many texts are in image, the low contrast is in the RBN tag and the undetected face is affected by hat, mask and some face pose that causing cannot verify RBN. Figure 5(a) shows example of images that contain many texts on runner. Figure 5(b-d) indicates example of missed face detection.

5 Conclusions

This research presented a method to locate RBN position in marathon images which is the importance step of RBN recognition system. The method is based on edge detection technique, image processing and face detection. The algorithm was tested on 400 images and the performance was satisfied which is a lot improved from the original work [1] and a bit lower than the recent work [2]. In addition, the algorithm is fast computation.

The challenge with this method is when implemented on an extreme complex background, e.g. many texts and text-like-shape that will produce many false alarms. Future work will find methods or techniques to filter out non-RBN as much as possible. In addition, a new method to extract RBN appearances or features will improve the accuracy of the process. Last, the proposed face detection method has a problem with runner wearing hat, mask and various head pose. Future work will attempt to solve these problems that will improve the performance.

5(a) 5(b)

5(c) 5(d)

Fig.5 Example images of false alarm detection and undetected face

References:

[1] Ben-Ami, I., Basha, T. and Avidan, S. Racing Bib Numbers Recognition. *The British Machine Vision Conference (BMVC)*; 2012, pp.1–10.

[2] Shivakumara, P., Raghavendra, R., Qin, L., Raja, K.B., Lu, T. and Pal, U. A new multi-modal approach to bib number/text detection and recognition in Marathon images. *Pattern Recognition 2017*, Vol.61, pp. 479–491.

[3] Roy, S., Shivakumara, P., Mondal, P., Raghavendra, R., Pal, U. and Lu, T. A New Multi-modal Technique for Bib Number/Text Detection in Natural Images. *Proceedings of Pacific Rim Conference on Multimedia*; 2015, pp. 483-494.

[4] Ye, Q. and Doermann, D. Text detection and recognition in imagery. A survey. *IEEE transactions on pattern analysis and machine intelligence 2015*, Vol.37, No.7, pp.1480–1500.

[5] Epshtein, B., Ofek, E. and Wexler, Y. Detecting text in natural scenes with stroke width transform. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2010.

[6] Howe, N.R. Document binarization with automatic parameter tuning. *International Journal on Document Analysis and Recognition 2013*, Vol. 16, No. 3, pp.247-258.

[7] Smith, R. An overview of the Tesseract OCR engine. In Document Analysis and Recognition, 2007. ICDAR 2007. *Proceeding of the Ninth IEEE International Conference on Document Analysis and Recognition, 2007*, pp. 629-633.

[8] Deb, K., Chae, H.U. and Jo, K.H. Vehicle License Plate Detection Method Based on Sliding Concentric Windows and Histogram. *JCP 2009*, Vol.4, No.8, pp. 771–777.

[9] Du, S., Ibrahim, M., Shehata, M. and Badawy, W. Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology 2013*, Vol.23, No.2, pp. 311–325.

[10] Júnior, P.R.M., Neves, J.M., Tavares, A.I. and Menotti, D. Towards an automatic vehicle access control system: License plate location. *In Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference*, pp. 2916-2921.

[11] Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics 1979*, Vol.9, No.1, pp. 62–66.

[12] Viola, P. and Jones, M. Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2001.

[13] Zuiderveld, K. Contrast limited adaptive histogram equalization. *Graphics gems IV, Academic Press Professional 1994*, pp. 474-485.

[14] Benchmark for Enhancing the Standard for Thai language Processing 2016 [Internet]. 2017 [cited 2017 Jun 1]. Available from: https://thailang.nectec.or.th/best/