An Efficient Method of Text Localization in Complicated Background Scenes

He Xiao
Computer School, China West Normal University, Nan Chong, Sichuan, 637009, P.R China
Email: ho.xiao@aliyun.com

Yunbo Rao
School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, 610054, P.R China
Email: cloudrao@gmail.com

***Abstract*— We propose a novel method for text localization in low illumination images with complicated background scenes. Low illumination image, complex background and variations of text script make text localization problem challenging. The proposed method use density-based information and rectangle window in the residual edge image. The proposed method firstly is color space conversion from RGB to Ycbcr color space, secondly we extract Y component from Ycbcr color space and get image enhancement from Y component, then extract out the vertical edges of enhanced Y component, thirdly, long curve and random noise in complicated background are removed, finally, search the text localization and segment the text out from the original low illumination images. We evaluate our method using natural image data set and ICDAR data set. The experimental results show that the proposed method has more robust to interference characters and more accurate when compared with other methods.**

***Index Terms*—text localization, complicated background, low illumination, density-based information**

I. Introduction

Text localization are many application for such a technology, for example, recognizing book/CD cover, traffic light recognition, license plate recognition, image and video search engine, and web mining. This paper presents work towards automatic reading of text localization in low illumination images with complicated background scenes.

Various text localization methods are used for extracting text localization. Fig.1 shows some of the examples of different document and non-document images (scene images) with areas. Fig.1 (a) is an example of document images. Fig.1 (b-d) is non-document images with high illumination and simple background. For these image scenes, the state-of-the-art methods of text localization can resolve these problems well. So we don't study these cases of text localization in this paper. Various approaches are still some limitations and make text localization a challenging task. The main factors including: (1) font style and thickness, (2) low illumination images, (3) low quality images, (4) camera position introduce geometric distortions, (5) complicated background as well as foreground color and texture. In this paper, we only discuss the low illumination images of text localization with complicated background scenes case. Fig.2 shows a good example of the low illumination images for our experimental data.

Fig.1 (a) example of document images; (b), (c), and (d) non-document images, Note that, these cases with high illumination and simple background, we don't study these cases in this paper.

For the more accurate of text localization, the low illumination images should be enhancement as pre-processing in our method. Many image enhancement methods have been proposed. Such as histogram equalization, power lower transform, tone mapping function, Fourier, wavelet, discrete cosine transforms et al [1, 2]. However, these traditional techniques can't satisfy different illumination and complicated background case. At present, some existing software have already provided noise removal and contrast enhancement functions, it is likely that most of them introduce artifacts and could not produce desirable results for a broad variety of illumination images[2]. In this paper, we propose the tone mapping function to enhance illumination images as pre-processing of text localization.

doi:10.4304/jsw.9.6.1538-1544

Fig.2 Sample original images in our data sets, the problem we address in this paper is that of text localization.

The vertical edge images have many long background curves and short random noise edges. These complicated background and noise edges may interfere in the text localization. Morphological opening, closing, and connected component analysis are then performed on the binary masks to get rid of small and random noises, and to fill the holes [3]. However, this method is only resolve small and random noise, the long curves in background can not be removed by this method. Zhang et al.[4] propose three times scanning the edge image algorithm. The proposed scanning algorithm will record the edge lengths away from the top or bottom, and compute the actual edge lengths. However, this algorithm can’t remove a very long curve and complicated curve. In this paper, we also use similar method with work [4] to overcome these drawbacks and remove long curves in background and short random noise.

There has been a number of successful text localization works reported in [5-7]. Zheng et al. [5] describes a local-feature-based approach for text localization. The method use SIFT feature and geometric verification in complicated background scenes. Gatos et al [6] describes an efficient binarization and enhancement technique followed by a connected component analysis procedure for text detection in indoor/outdoor scene images. Chen et al [7] reports an approach of detection and recognition of sign from natural scenes. Laplacian of Gaussian (LOG) edge detector, color modeling, layout analysis and affine rectification are used to detect text. All these works have some limitations as follows: (1) the high computing work and can’t real-time implement. (2) For multiple texts, the algorithms can’t get all text localization. (3) These algorithms can’t consider low illumination images with complicated background scenes case. When these algorithms are implemented, low illumination images of text localization aren’t very satisfy results. In this paper, we consider the low illumination images with complicated background scenes case and propose a novel method for text localization, which is combine density-based information and rectangle window in the residual edge image.

Based on analysis above algorithms, in order to overcome these drawbacks of researcher’s mentions, we propose a novel approach for text localization in low illumination images with complicated background scenes. This paper introduces the following main results:

(1) Due to low illumination images can’t more accurate of text localization, we propose the tone mapping function to enhance illumination images as pre-processing of text localization.

(2) For remove a long curves in background and short random noise edges, we propose global scanning method to resolve previous method problems.

(3) For text localization, we use density-based information and rectangle window in the residual edge image to get text localization.

The remainder of the paper is organized as follows. Section II describes related work. Section III gives the proposed methodology. Experimental results are presented in Section IV. Finally, the conclusion is given in Section V.

II. Related Work

The research field of text localization receives a growing attention due to the proliferation of digital cameras and the great variety of potential applications. However, since there are problems of text localizations such as low illumination image, complicated background, and color similarity, the text localization is often difficult to be located accurately and efficiently. Many methods have been proposed in order to solve the problems. Such as: edge-based analysis, neural networks, color and fuzzy maps, vector quantization, texture analysis, et al.

Neural networks can be used as filters for analyzing small windows of an image and deciding whether each window contains the text. Park et al [8] propose to use neural networks to text localization. A post-processor combines the filtered images and locates the bounding boxes of text in the image is proposed. Text can be direct identified by scanning through the input image and looking for portions of the image that were not linked to other parts of the image. However, if the texts are linked to other part, the method can’t deal with these cases.

Edge-based methods the aim is to detect boundaries of an image, including edge statistics, edge features, and edge density. Due to sufficient information from edges in the text region, edge statistics yield promising results. This method is simplicity [9]. But using edge statistics information alone the rate of success is low especially in complex and low contrast images. Meanwhile, edge-based method uses the Half Transform [10], which has difficulty in extracting distorted or dirty image. It also has known as a complicated approach. Ming et al [11] propose using edge features and density of the text image, which can be used to successfully detect a number text location. Other researcher also developed a method to improve the edge image by eliminating the highest and lowest portions of the edge density to simplify the whole image. But some of the text region identity will be lost [12,13].

Color information of the low illumination image also plays an important role in text localization, where the unique color or color combination between the texts and complicated background are considered as the key feature to locate the text localization. Zhu et al. [14] use color features to locate text localization. However, this method is sensitive to the license plate color and brightness and needs much processing time. Meanwhile, fuzzy logic has been applied to the problem of locating texts. Zimic et al. [14] describe the text localization, and gave some membership functions for the fuzzy sets “bright” and “dark”, “bright and dark sequence” to get the horizontal and vertical text positions. This method is not robust enough to the different environments.

Vector quantization (VQ) image representation is a quadtree representation by the specific coding mechanism. Rodolfo et al.[16] devise a method based on VQ. It can give a system some hints about the contents of image regions, and such as information boosts location performance. However, it is a time consuming and complex method specially when applied to large images.

Texture analysis is another useful approach for text localization. Gabor filters have been one of the major tools for texture analysis [17, 18]. This approach takes the advantage of existing homogenous and frequent texture-like edges in text region. The process of this method is independent of rotation and scaling. It has the ability of studying images in an unlimited number of directions. However, this method is not satisfied real-time required.

Ezaki et al.[19] propose a novel based on connected components. The performance of the different methods depends on character size. Text regions are extracted from an image through edge extraction, enhancement and labeling. Since the texts in the images often have skewed and slant, the texts are recognized after the skew and slant correction. However, this method can’t deal with very dark images.

Currently, most researchers prefer a hybrid text localization method, where multiple features are involved in order to make the method more robust. In order to address these drawbacks of the methods, the method proposed in this paper is also a hybrid method. We propose a novel method, which is involved density-based information and rectangle window in the residual edge image. In next section, we discuss the proposed framework for text localization in low illumination images with complicated background scenes.

III. METHODOLOGY

By observing text localization in images, two main features are noticed. First, density-based information of text region are relatively strong and dominant. Second complicated background edges are usually either long curves or very short and with noise. These two important features and also low complexity for edge-based analysis motivate us to use edge information for the text localization.

The detailed procedures of the proposed method can be described as in Fig.3. The proposed method is composed of the following steps: (1) Conversion of RGB to Ycbcr and extract Y component, (2) image enhancement for low illumination images of Y component, (3) extract out the vertical edges of the low illumination image of Y component, (4) complicated background curve and noise removing, (5) search the text localization, and (6) segment the text out from the low illumination images. It should be noted that in Fig.3, the steps 1 in the proposed method also follows our previous work [3]. These steps are described in details in the following.

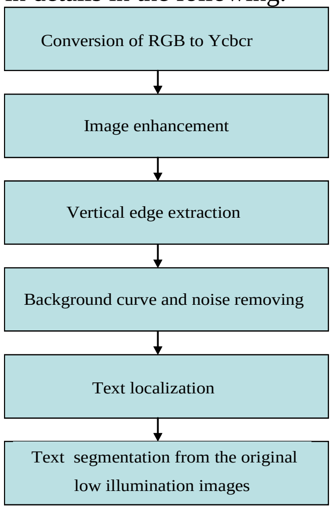

Fig. 3: A block diagram of the proposed method.

A. Conversion of RGB to Ycbcr

Our input low illumination with complicated background images is the RGB format which consists of three additive primaries: red, green, and blue. Since we perform enhancement method based on the illumination layer, a conversion is needed to convert input images from RGB to intensity+color components for achieving the intensity component. Due to Ycbcr is the most complete color model used conventionally to describe all the colors visible to the human eye, RGB color space is converted into Ycbcr color space where color information is not affected while modifying intensity values Y [2, 3].

The Ycbcr color components are given by the following equations [2, 3].

$$\begin{bmatrix} Y \\ cb \\ cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.00 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where *cb* and *cr* are color components respectively, Y is intensity component. Some examples of converted Ycbcr component are shown in Fig.4.

In this paper, we convert RGB color space to the Ycbcr color space for the Y+cb+cr decoupling. The Y component is taken as the intensity. The next processes are carried on the intensity component only.

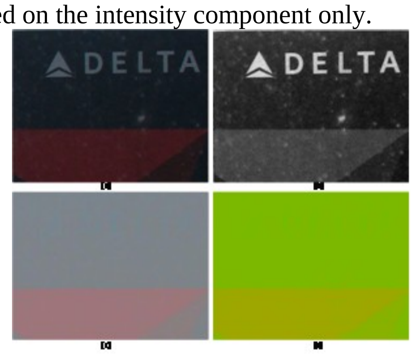

Fig.4 Conversion RGB to Ycbcr color space (a) original image, (b) Y component, (c) colored cb component, (d) colored cr component.

B. Image Enhancement

The goal of the traditional tone mapping of enhanced image is preserving image details and providing enough absolute brightness information in a low dynamic range image. Eric P. Bennett [20] proposes a tone mapping function to enhance underexposed, low dynamic range videos by adaptively and independently varying the exposure at each photoreceptor in a post-processing. The tone mapping function is given by:

$$m(x,\psi)=\frac{\log(\frac{x}{xMax}(\psi-1)+1)}{\log(\psi)} \qquad (2)$$

The white level of the input luminance is set by $xMax$ and ψ controls the attenuation profile. However, this method is the global enhancement image method. In our work, if we extract edge images directly form these text areas, a few vertical edge will appear in the text localization areas. Therefore, the low illumination images enhanced is important for text localization firstly. Meanwhile, the local (text) areas have low variances, which need to be enhanced in the low illumination image. In this paper propose a novel enhancement method of local areas for preserving image details and enhanced brightness information from Y components, which is inspired by [20,21]. The proposed method is given as follows:

$$L_{(x,y)}^{enh}=\frac{\log(\sigma_0)}{\log(\sigma_{(x,y)})}(L_{(x,y)}-\overline{L}_{(x,y)})+L_{(x,y)} \qquad (3)$$

where $L_{(x,y)}^{enh}$ is enhanced illumination of the input images. σ_0 is a standard deviation of the pixels in the slide windows, in this paper, the size of slide windows is 22×30 pixels. $\sigma_{(x,y)}$ is a standard deviation of the current area, which is scanned by slide windows. $\overline{L}_{(x,y)}$ is a pixels average of slide windows, $L_{(x,y)}$ is the illumination value of the current pixels. An experimental result is shown in Fig.5.

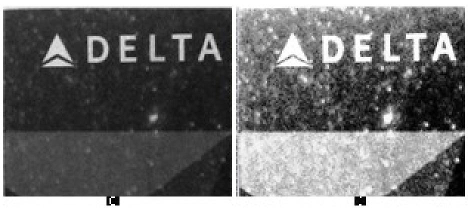

Fig. 5: (a) original low illumination image with complicated background scene. (b) enhanced low illumination image.

C. Vertical Edge Extraction

By observing inputted low illumination images, one of important features is that density of edge across scene are significant, while background edges are usually either long curves or very short. We firstly use Sobel to obtain the edge density of the input low illumination image. The selection of a proper threshold to extract strong edges from the gradient image and prevent to miss important edge information is relatively difficult. If a high threshold level may miss text edges, whereas a low level for the threshold results lots of weak edges in the clutter part of the scene. It is worth to note that the text region features significant density of vertical edges [2]. In this paper, the Sobel operator is shown as follow.

$$H=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ 1 & 0 & 1\end{bmatrix} \qquad (4)$$

In order to detect candidate regions for the text, we estimate edge density across the edge image by applying a Gaussian kernel on it. An experimental result is shown in Fig.6.

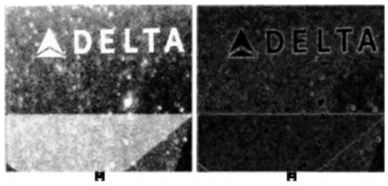

Fig.6 (a) low illumination images with complicated background scenes. (b) the vertical edges extracted from the low illumination image.

D.Background Curve and Noise Removing

In this step, we firstly use morphological opening, closing, and connected component to perform on the binary masks to get rid of small and random noises, and to fill the holes [3]. Due to complicated background, some long curves and big holes still have exiting on low illumination image. We use the similar method with work [4] to remove long curves and random noise. The proposed algorithm only requires us to scan the edge image for two times. The detail step of the proposed algorithm is shown as follows:

Step 1: input the binary images $I(x,y)$ and initialize the size of image m and n to zero matrixes.

Step 2: for each row i from left to right to scan, and record point of the curve position $P(i,j)$.

Step 3: for each column j from top to bottom to scan, and record point of the curve position $P(i,j)$.

Step 4: for each row i connect $P(i,j)$ to denote the background curve, for each column j also connect $P(i,j)$ to denote the background curve. If the curve length is close m or n, this curve is considered long curve, and should be remove. If the curve length is very short and closes one or two points, this curve is considered random noise, and also should be remove.

In this algorithm, we accumulate the edge lengths through observing the “concerned neighborhood pixels” of the current pixels $P(i,j)$. Fig.7 shows a experimental result, From Fig.7(b), most of the background curve and random noise have been eliminated, but the text localization edges are almost fully saved.

Fig.7 (a) the original binary image after vertical edge extraction. (b) the result of long curve and random noise removing from the original binary image.

E. Text Localization and Segementation

After the enhancement process and background curve or random noise removing, we get the new binary images $I(x, y)$. It take into account certain limits for the height and width of the connected component along with the appearance of neighboring connected components with almost the height in the horizontal direction.

We use rectangle window in the residual edge image to scan the connected area. In this paper, we assume that all connected areas (text localization region) by equation as follow:

$$R_{n_i \times n_j} = \{R_1, \cdots R_{m-1}, R_m\} \tag{5}$$

Where $n_i \times n_j$ is the size of connected area by rectangle window scanning. m denote the number of the connected area. Sum_m denote the pixel sum of the order m connected areas.

Given the density information ρ_{ave} for all of the connected regions, ρ_{ave} is defined to the total the connected divided by the total pixel of the connected area.

$$\rho_{ave} = \sum_{i=1}^{m} Sum_m / \sum_{i=1}^{m} R_{n_i \times n_j} \tag{6}$$

Given the density information ρ_m for every the connected area, ρ_m is defined to the connected divided by the pixel sum of the order m connected areas.

$$\rho_m = Sum_m / R_{n_i \times n_j} \tag{7}$$

If the density information inequality $\rho_m > \rho_{ave}$, this connected area is considered text localization. If the density information is satisfy inequality $T < \rho_m < \rho_{ave}$, connected area also is considered text localization. The threshold T is an experimental value in our paper.

For text segmentation, we use the work of [22]. Nomura et al. [22] propose a morphological thinning algorithm and the segmentation cost calculation automatically determine the baseline for segmenting the connected text.

IV EXPERIMENTAL RESULTS

In order to estimate the method efficient, various data sets of text are used in the proposed method. The first experiments have been conducted using more than 1000 images of texts with low illumination and complicated background. The images have been taken from natural scenes mainly with complex background and under different low illumination. Some input images and the result of text localization has been shown in Fig. 8.

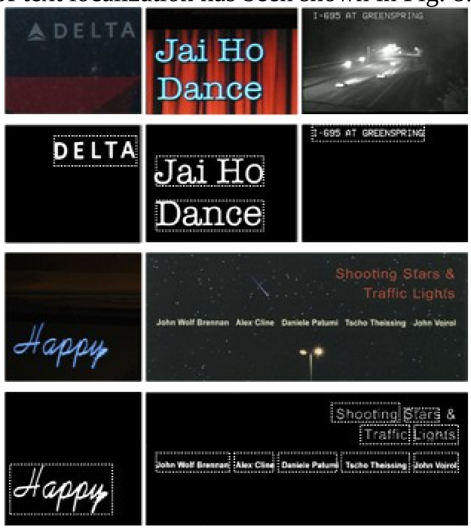

Fig.8 The first and third row is original low illumination image with complicated background scene, the second and fourth is experimental result of text localization.

TABLE 1
A COMPARISON OF SOME STATE-OF-THE-ART METHOD

Ref.	Localization rate	Time(sec)	Text localization method
[9]	90.08	38.2	edge-based analysis
[8]	87.91	36.5	neural networks
[13]	89.23	35.2	color and fuzzy maps
[15]	90.10	36.1	vector quantization
[16]	88.45	34.4	texture analysis
[18]	87.80	37.8	connected component
Our method	91.20	34.5	density-based

To make a tradeoff between accuracy and speed by altering the introduced threshold, we also give a comparison between the proposed method and some state-of-the-art well-reported methods using natural scenes images. The experimental results in the literature are given in Table 1.

In this experiment, 912 text localizations out of 1000 have been correctly extracted. The promising localization rate 91.2% has been achieved with low processing complexity so the proposed method is performed fast. Implemented time is an average time of 1000 images processing.

The third experiment is that the proposed method was tested using the public database of the ICDAR2005 Robust reading completion [23]. We focused on the data set for “Text localization” competition. The proposed method worked successfully even in case with shadows, non-uniform illumination, low illumination, and large signal-dependent noise. Experimental results are shown in Fig.9. The text areas are located in the final binary images while the non-text areas are eliminated.

Fig.9 The first and third row is original low illumination image with complicated background scene, the second and fourth is experimental result of text localization.

The fourth experiment is that we further perform the proposed method well by using precision rate and recall rate [8,24]. The precision and recall accuracy of text localization is computed as:

$$\text{Precision} \quad rate = \frac{\#of\ text \quad area \quad pixels}{\#of\ text \quad pixels \quad localization} \tag{8}$$

$$\text{Recall} \quad rate = \frac{\#of\ text \quad area \quad pixels}{\#of\ text \quad pixels \quad in \quad ground \quad truth} \tag{9}$$

Table 2 shows the precision and recall rates of state-of-the-art for 1000 image using our data set. From this table, we find our method is the better than other methods.

TABLE 2
AVERAGE PRECISION AND RECALL OBTAINED USING STATE-OF-THE-ART METHOD AND OUR METHOD ON SET OF 1000 IMAGES

	Ref[8]	Ref[9]	Ref[13]	Ref[15]	Ref[16]	Ref[18]	Our method
Precision rate	75.8	76.7	70.2	73.5	74.6	75.4	77.8
Recall rate	72.9	71.5	67.9	70.2	71.3	70.6	74.2

All the experiments were performed on PC running Windows XP. The PC is 2.8GHz Intel Xeon i5 3470 processor, Intel GMA HD 4000 graphics card, and 4GB of RAM. The average processing times for the five stages of the proposed method are listed in Table 3. A lot of the time is consumed on the first stage “image enhancement”. The total time of processing average 1000 image with size 640×480 is 34.5 ms, and it meets the requirement of real time processing.

TABLE 3
THE PROCESSING TIMES FOR THE FIVE STAGES IN THE PROPOSED METHOD (UNIT: MS)

color space conversion	image enhancement	vertical edges	long curve and noise removing,	text localization, and segmentation	Total time
5.6	10.4	3.1	7.9	9.5	34.5

V CONCLUSIONS

This paper strives toward a novel methodology that aids automatic localization, segmentation of visual text entities in low illumination with complicated background scene images. The proposed methodology is based on density-based information and rectangle window in the residual edge image. Experimental results show that by using the proposed method we achieve a good text localization rate for low illumination with complicated background scene images. It also shows more accurate detection results than the state-of-the-art methods of text localization.

In our future work, we plan to deal with the problem of text detection with dynamic video processing.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments. This work is partly supported by Scientific Research Fund of Sichuan Provincial Education Department (Grant No. 11ZB274), and National Science Foundation of China (Grant No. 61300092).

REFERENCES

[1] T. Arici, S. Dikbas, and Y. Altunbasak, A histogram modification framework and its application for image contrast enhancement, IEEE Transactions on Image Processing, vol.18, no.9, pp.1921-1935, 2009.
[2] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Person Prentice Hall, New Jersey, USA, 2008.
[3] Y.B Rao, Z.H Chen, M.-T Sun, Y.-F Hsu, and Z.Y Zhang, An effective nighttime video enhancement algorithm, Visual Communications and Image Processing (VCIP) , Nov 6-9, 2011, Taiwan
[4] D.N Zhang, Y.N Zhao, and J.X Wang, An efficient method of license plate location, Pattern Recognition Letters, vol. 26, pp.2431-2438, 2005.
[5] Q. Zheng, K. Chen, Y. Zhou, C.C Gu, and H.B Guan, Text localization and recognition in complex scenes using local features, Part III, LNCS 6494, pp. 121–132, 2011.
[6] B.Gatos, I. Pratikakis, K.Kepene, and S.J.Prtantonis, Text detection in indoor/outdoor scene images, In 1st Int. Workshop on Camera-based Document Analysis and Recognition, Seoul, Korea, Aug. 2005, pp. 127-132.
[7] X. Chen, J. Yang, J. Zhang, and A. Waibel, Automatic detection and recognition of signs from natural scenes, IEEE Transactions on Image Processing, vol.13, pp.87-99, 2004.
[8] S.H Park, K.I Kim, K. Jung, and H.J Kim, Locating car license plate using neural networks, Electron Letter, vol.35, no. 17, pp.1475-477, 1999.
[9] C.N.E Anagnostopoulos, I.E Anagnostopoulos, V Loumos, and E.Kayafas, A license plate recognition algorithm for intelligent transportation system applications, IEEE

Transaction on Intelligent Transportation Systems, vol.7, no.3, pp.377-391, 2006.
[10] T.D Duan, H. Du, T.L., Phuo, and T.V Hoang, Building an automatic vehicle license plate recognition system, In: Processing of Int. Conf. Comput. Sci. RIVF, pp. 59-63, 2005
[11] G.H Ming, A.L Harvey, P. Danelutti, Car number plate detection with edge image improvement, In: Fourth Internat. Symp on Signal Process and its Applications, vol. 2, pp. 597-600, 1996.
[12] H. Bai, C. Liu, A hybrid license plate extraction method based on edge statistics and morphology, In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 2, pp.831-834, 2004.
[13] W. Xu, Z. Liu, T. Wang, S.Y Liu, Sentiment recognition of online Chinese micro movie reviews using multiple probabilistic reasoning model, Journal of Computers, Vol. 8, No. 8, pp.1906-1911, 2013
[14] W.G Zhu, G.J Hou, and X Jia, A study of locating vehicle license plate based on color feature and mathematical morphology, In: 6th Internet. Conf. on Signal Processing, vol. 1, pp.748-751, 2002
[15] N. Zimic, J Ficzko, M. Mraz, and J. Virant,The fuzzy logic approach to car number plate locating problem. In: Processing Intelligent Information Systems, pp. 227-230, 1999
[16] Z. Rodolfo, R. Stefano, Vector quantization for license plate location and image coding, IEEE Trans. Industrial Electron, vol.47, no.1, pp.159-167, 2000.
[17] W.J Jia, H.F Zhang, and X.J He, Region-based license plate detection, Journal of Network and Computer Applications, vol.30, pp.1324-1333, 2007
[18] V. Abolghasemi, A. Ahmadyfard, A fast algorithm for license plate detection, In: Proceeding of the VISUAL, pp.468-477, 2007
[19] N. Ezaki, M.Bulacu, and L.Schomaker, Text detection from natural scene images: towards a system for visually impaired persons, In Proceedings of the International Conference on Pattern Recognition (ICPR), pp.683-686, 2004
[20] E. P. Bennett, and L. McMillan, Video enhancement using per-pixel virtual exposures, ACM Transactions on Graphics, vol. 24, no. 3, pp. 845-852, 2005.
[21] X.Y Liu, H. Fu, An effective clustering algorithm with ant colony, Journal of Computers, Vol.5, No.4, pp.598-605,2010
[22] S.G Nomura, K.J Yamanaka, O. Katai, H.S Kawakami, and T.Y Shiose, A novel adaptive morphological approach for degraded character image segmentation, Pattern Recognition, vol. 38, no.11, pp.1961-1975, 2005
[23] ABBYY. www.finereader.com 2005.
[24]W.Y Liu, T. Wang, Index-based online text classification for SMS spam filtering, Jouurnal of Computers, Vol.5, No. 6, pp.844-851, 2010

He Xiao received the B.S. and M.S. degrees from the China West Normal University and the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2005 and 2008, respectively. Since 2008, he has been an assistant professor at the Computer School of China West Norma University. His research interests include image and video enhancement, computer vision, and computer graphics. Contact him at ho.xiao@aliyun.com

Yunbo Rao received the B.S. and M.S. degrees from the Sichuan Normal University and the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2003 and 2006, respectively, both in School of Computer Science and Engineering (SCSE), and the PhD degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. He has been as a visiting scholar of Electrical Engineering of the University of Washington from Oct 2009 to Oct 2011, Seattle, USA. Since 2012, he has been an assistant professor at the School of Information and Software Engineering, University of Electronic Science and Technology of China. His research interests include video enhancement, computer vision, and crowd animation. Contact him at cloudrao@gmail.com.