

Research on Character Recognition Technology based on Computer Vision

ZHAO Guotao^{1, a}, Ding Jie²

¹Hubei Engineering University, Xiaogan 432000, China

²College of Technology of Hubei Engineering University, Xiaogan 432000, China

^azgt234800@163.com

Keywords: Computer Vision; Character Recognition; Adaboost algorithm

Abstract. Computer vision is an important research direction for the man-machine interactive intelligent visual image processing. It is a key technology to the text information extraction and recognition of video or image. As a new research direction of video analysis and application field, Video OCR has the certain application and theory significance. In this paper, the technical route of Video OCR system has been deeply researched. Adaboost low resolution and scanning technology based on character recognition has been proposed. Some numerical simulations are made to test the validity and capability of the proposed technology.

Introduction

Computer vision is the research of how to make the machine to "see". In other words, it refers to with camera and computer to replace human eyes on target recognition, tracking and measuring machine vision, etc. As a scientific subject, theory and technology of computer vision research related, trying to build can be obtained from the images or multidimensional data information system of artificial intelligence [1-3]. Intelligent Video OCR is a computer interactive visual image processing, an important research direction, it is mainly involved in Video detection, tracking and text information extraction, recognition and retrieval technology. Video character recognition (Video OCR) is a kind of to Video or image of the character or character information extraction and recognition of Video information processing technology.

As digital Video is finding wider and wider application in all fields, the extraction of Video information, retrieve, query technology more and more important, Video OCR research has become a hot research topic in the field of Video information processing [1]. A complete Video OCR system is mainly composed of the text frame detection, Video text localization, text tracking, Video text enhancement and segmentation, character segmentation and recognition of several parts. Furthermore, it involves many difficulties, including the establishment of theoretical model and the determination of technical route, the current related research work in a new theory and technology innovation constantly [4-6].

In this paper the technical route of Video OCR system has been deeply researched. Adaboost low resolution and scanning technology based on character recognition has been proposed. Numerical simulation is performed to validate the performance of the proposed technology.

The rest of the paper is organized in the following sequence. In Section 2, the proposed technology is introduced in detail. In Section 3, numerical simulation results are made and given to demonstrate the validity of the method. Concluding remarks are summarized in the final section.

The overall model framework of Video OCR system

A complete Video OCR system is mainly composed of the following parts: the text frame detection, Video text localization, text tracking, Video text enhancement and segmentation, character segmentation and recognition, its structure diagram is shown in figure 1.

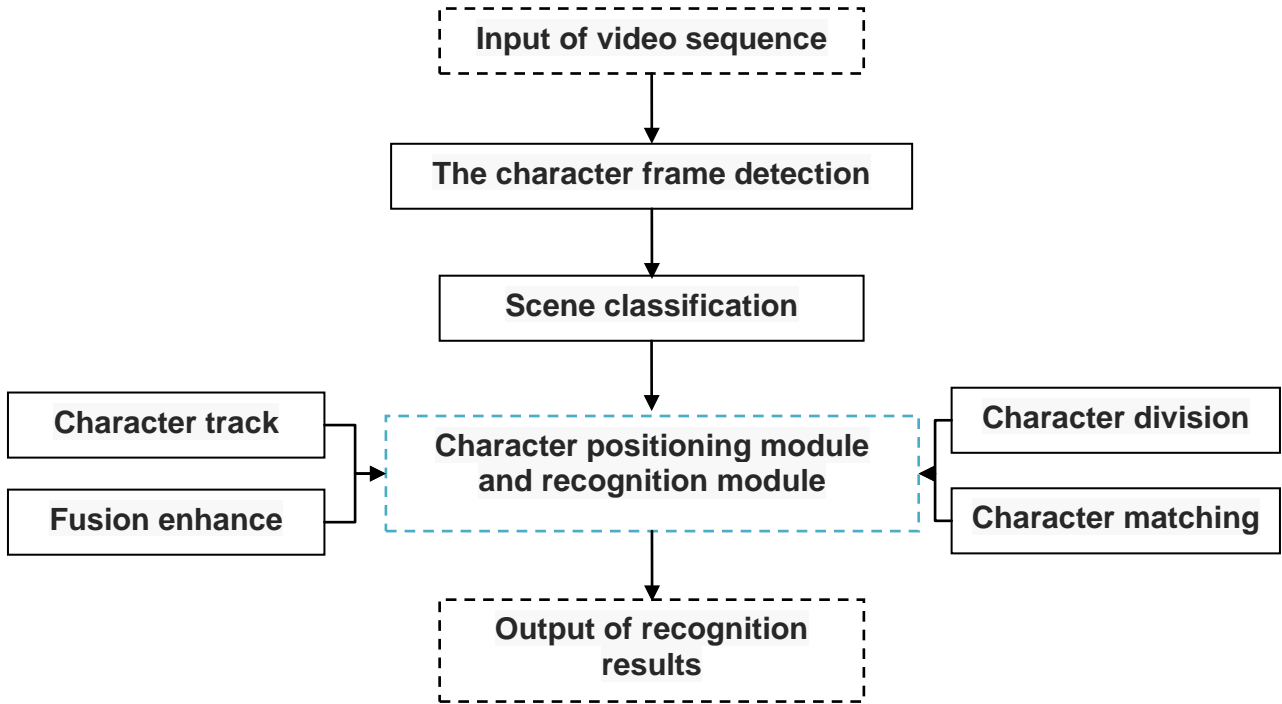


Fig. 1. The flowchart Video OCR system

Character tracking

Composite are studied in this paper the Wolf [7], Li [8] and Crandall [9] and others for video character tracking related research methods, combining with the video text has time redundancy is studied. Video and image character in its content of form is the biggest difference between character with temporal redundancy of video. Assumptions, is now able to locate certain text area to track, the time redundancy of video text has the following functions:

- 1) Effectively reduce and reduce the cost at a cost to the text location;
- 2) This time redundancy can also be used to remove the video frames contained in the text information, identifying the main basis is the non-text information does not generally like text in a number of consecutive video frames;
- 3) Based on multiple consecutive video frames contained in the same character information enhancement operation, can effectively improve the video text recognition rate.

Character classification

Co In order to get the capability of classification of text classifier, this paper USES the SVM learning machine, the extraction of the text characteristic to carry on the training study. SVM is proposed by Vapnik [10] based on statistical learning theory proposed by a machine learning method, this algorithm classified due to its outstanding performance and is widely used in the classification problem. Its main idea is to use the largest interval method to seek the optimal division of the plane to classification point between class interval is the largest. In the case of linear inseparable, then used a kernel function to the points of the nonlinear mapped to high-dimensional space and make the linearly separable. Different from traditional classification methods of empirical risk minimization, the minimum structure risk of SVM classification method USES is stronger classification ability. By learning and training data, the resulting character classifier can be described as [11]:

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i y_i K(x, x_i) + b\right) \quad (1)$$

where x denotes the input feature vector. α_i, b are the decision parameters obtained from the training. (x_i, y_i) and $K(x_i, y_i)$ represent support vector and Kernel function, respectively. If

$f(x) > 0$, then the area according to the input feature vector x including characters. If $f(x) \leq 0$, the area including no character.

The text precise positioning method based on SVM, can greatly eliminate the initial location of the noise interference, achieve better location performance. Part of the positioning rendering is shown in Figure 2.

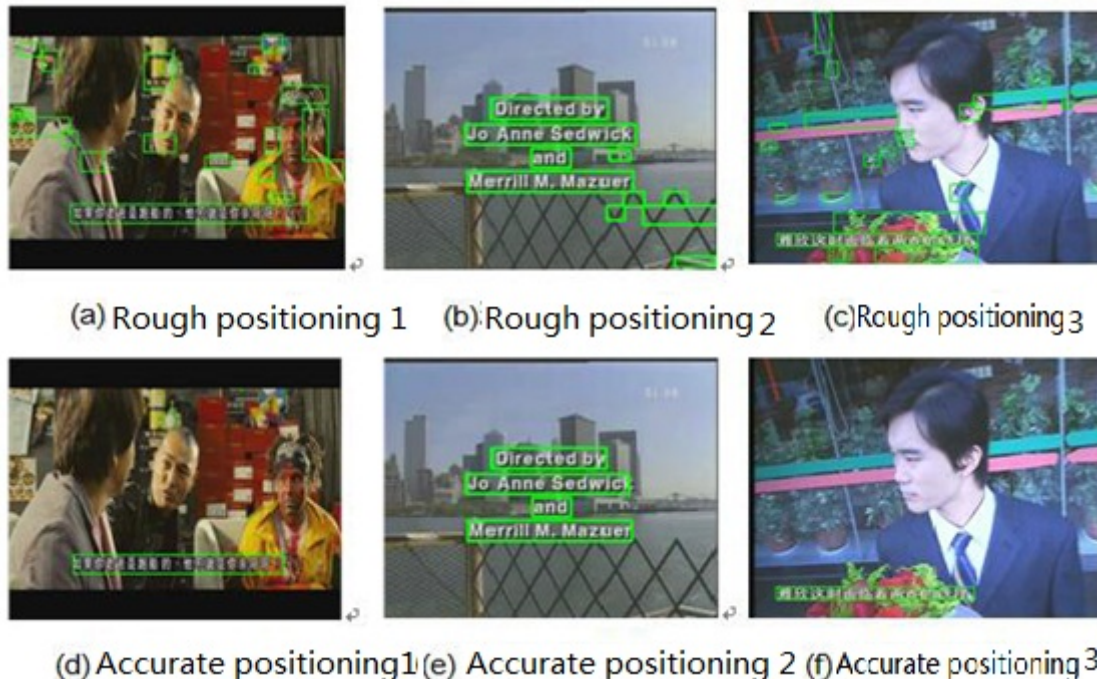


Fig. 2. The character rendering accurate positioning method based on SVM

The text precise positioning technology based on Adaboost

In order to better solve the problem of text location under complex background, this paper based on the text of the precise positioning strategy [11-15]. Adaboost algorithm is a strong classifier classification performance. As a kind of iterative algorithm, Adaboost algorithm is mainly through changing the distribution of the data to implement the algorithm process, it not only classification accuracy is high, and because the time complexity can be adjusted according to different classification task, its classification efficiency is very considerable, so the algorithm has been widely used in target detection, recognition and positioning.

The algorithm to determine the main basis of each sample is weighted includes two aspects:

- 1) In each round of training, the classification of each sample is correct;
- 2) After the previous round training samples of all the overall classification accuracy.

On the basis of the two basis, the algorithm is obtained by training the classifier to constantly fusion, will serve as the final decision classifier, the weight of each sample of the specific decision process in this paper are described in details. Among them, accounting for the weight of the sample, characterization of its being weak classification for the probability of training set. If a sample in a round of training without can be classified correctly, its corresponding weight would increase accordingly, then under construction a set of training when the selected probability also would increase accordingly, on the contrary, the weights to reduce.

With such a large number of features, using Adaboost algorithm training would be a good choice. This is because the essence of Adaboost method is by training a large number of weak classifier (e.g., consisting of a single feature classifier) were combined to form a strong classifier.

Through a large number of initial samples training such strong classifier will have a strong ability of classification, the classifier is more than qualified to the distinction between the text and background, for precise positioning of text have obvious effect.

Conclusion

Computer vision refers to with camera and computer to replace human eyes on target recognition, tracking and measuring machine vision. As a new research direction of video analysis and application field, Video OCR has the certain application and theory significance. In this paper, the technical route of Video OCR system has been deeply researched. Adaboost low resolution and scanning technology based on character recognition has been proposed. Some numerical experiments are made to test the validity and capability of the proposed technology. Several conclusions were drawn in this paper:

- 1) In text positioning module, the text precise positioning method based on SVM, can greatly eliminate the initial location of the noise interference;
- 2) Use trained classifier Adaboost algorithm, by updating the weight of candidate text filtering step by step. Through the experiments, the results show that this algorithm has stronger stability and robustness.

References

- [1] Lienhart R. Video Mining[M]. Springer US, 2003:155-183.
- [2] Datong Chen, Jean-Marc Odobez and H. Bourlard. Text Detection and Recognition in Images and Video frames. Pattern Recognition. 2004(37):595-608.
- [3] Na Y, Wen D. Advances in Multimedia Information Processing - PCM 2010[M]. Springer Berlin Heidelberg, 2010:392-403.
- [4] Wing Hang Cheung, Ka Fai Pang, Michael R. Lyu, Kam Wing Ng, Irwin King. Chinese Optical Character Recognition for Information Extraction from Video Images[C]. Proceedings of International Conference on Imaging Science,2001.
- [5] Datong Chen, Jean-Marc Odobez and H. Bourlard. Text Detection and Recognition in Images and Video frames. Pattern Recognition. 2004(37):595-608
- [6] Mao Wen-ge, Chung Fu-lai, Lam K, et al. Hybrid Chinese/English text detection in images and video frames. Proceedings of 16th International Conference on Pattern Recognition, Quebec, Canada, 2002:1015-1018.
- [7] Wolf C, Jolion J M, Chassaing F. Text Localization, Enhancement and Binarization in Multimedia Documents[C]. Proceedings of Pattern Recognition. 2002, 2:1037 -1040.
- [8] Li H, Doermann D, Kia O. Automatic Text Detection and Tracking in Digital Video[J]. IEEE Trans on Image Processing. vol.9, no.1 pp.147-156, 2000.
- [9] David Crandall. Extraction of Special Effects Caption Text Events from Digital Video[J]. International Journal on Document Analysis and Recognition. 2003(5): 138-157.
- [10] Kwang In Kim, Keechul Jung, and Jin Hyung Kim. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence. 2003.
- [11] Wang X, Huang L, Liu C. A video text location method based on background classification[J]. International Journal on Document Analysis and Recognition (IJDAR), 2010, 13(3):173-186.
- [12] Chong-Wah Ngo, Chi-Kwong Chan. Video text detection and segmentation for optical character recognition, Multimedia Systems, Vol.10, 2005.3.
- [13] Dongqing Zhang, Raj Kumar Rajendran, And Shih-Fu Chang. General And Domain-Specific Techniques For Detecting And Recognizing Superimposed Text In Video[C]. ICIP. 2002.
- [14] Minoru Mori. Video text recognition using feature compensation as category- dependent feature extraction[C]. In Proceedings of 7th International Conference on Document Analysis and Recognition. 2003(2):645.
- [15] Osamu Hori. A video text extraction method for character recognition[C]. In Proceedings of 5th International Conference Document Analysis and Recognition. Bangalore, India, 1999:25-28.