Expert Systems with Applications 40 (2013) 6955–6963

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Performance improvement of character recognition in industrial applications using prior knowledge for more reliable segmentation

Martin Grafmüller [a,*], Jürgen Beyerer [a,b]

[a] Karlsruhe Institute of Technology (KIT), Vision and Fusion Laboratory, Adenauerring 4, 76131 Karlsruhe, Germany
[b] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), 76131 Karlsruhe, Germany

ARTICLE INFO

Keywords:
Character segmentation
Prior knowledge
Character recognition

ABSTRACT

In industrial applications optical character recognition with smart cameras becomes more and more popular. Since these applications mostly have challenging environments for the systems it is most important to have very reliable character segmentation and classification algorithms. The investigations of several algorithms have shown that character segmentation is one if not the main bottleneck of character recognition. Furthermore, the requirements of robust and fast algorithms related to skew angle estimation and line segmentation, as well as tilt angle estimation, and character segmentation are high. This is the reason for introducing such algorithms that are specifically adapted to industrial applications. Additionally, a method is proposed that is based on the Bayes theorem to take account of prior knowledge for line and character segmentation. The main focus of the investigations of the character recognition system is recognition performance and speed, since real-time constraints are very hard in industrial application. Both requirements are evaluated on an image series captured with a smart camera in an industrial application.

1. Introduction

Optical Character Recognition (OCR) using smart cameras becomes more and more common in addition to visual inspection tasks in industrial applications like pharmaceutical, food, electronic, or automotive industry. The aim is to read product or serial numbers to track goods over their entire life cycle automatically. Furthermore, it is possible to record in databases, which parts are built-in which products. Another application is to read packaging or expiration dates of drugs or foods to decide whether to deliver it where and when or to discard it. These are all examples where smart cameras substitute manual entries by a human. Thus, the main advantage of smart cameras is that they are reliable over time, since they do not get tired. The camera solution is much cheaper and usually significantly faster. The reading performance of humans considering only a short period of time is much higher though. The different industrial applications require hard constraints to the OCR system, e.g., it is very robust to changes of the material the characters are printed on, the font itself, and real-time demands. The experience has shown that in such kind of applications character segmentation is the main bottleneck of the system, since this part takes the longest processing time especially if also skew and tilt angle of the lines and characters must be estimated. Furthermore, the errors occurred during segmentation can not be corrected by the classifier. This means a reliable segmentation method is most important.

There are also constraints that can be used in a positive way to speed up the system and improve the reliability. Usually the text region, which contains all the characters to be read is selected by the user. This makes character detection in the whole image unnecessary. Moreover, the lines are generally aligned parallel, which does not mean that they are horizontally aligned to the image. The biggest advantage is that there is no big change from image to image except that the characters change normally. Thus, the number of lines and characters usually does not change in the image series like the line and character distances. This fact can be exploited in a way that the user can initialize the OCR system with the first picture captured. If so, the user can check the segmentation result and—if necessary—correct it manually. This information can be used as prior knowledge for segmentation of the characters for the following images, which may lead to significant improvements in segmentation and thus in classification. Furthermore, two different features and classifiers are evaluated with respect to accuracy and speed.

1.1. Related work

Optical Character Recognition is a broad field, which ranges from document analysis to license plate recognition. A paper that

* Corresponding author.
E-mail address: martin.grafmueller@vde-mitglied.de (M. Grafmüller).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.06.004

summarizes the last 40 years of character recognition was written by Fujisawa (2008). Some of the most important basics for character recognition can be found in Bunke and Wang (2000).

Similar to character recognition in industrial applications with smart cameras is license plate recognition, where a lot of articles have been published in recent years. A Chinese license plate recognition system was introduced in Shi, Zhao, and Shen (2005). The detection of the plates is based on color image processing. After detection of the license plates binarization of the image is performed using an adaptive threshold. Projection profiles are used for line segmentation. The threshold for line segmentation is linked to the height of the plates and thus depends on whether there are one or two lines in the image. Character segmentation is performed similarly. Segmentation errors are corrected by a post processing step, e.g., lines are horizontally aligned or fragments caused by the frame of the plate or rivets are removed. Furthermore, a correction of the alignment of every character is performed. For classification they use pattern matching with additional rules, which are applied to characters that are hardly discriminable. For further improvements of classification the system considers at which positions appear letters or numbers only. The system has a recognition rate of more than 90%. A license plate recognition system for Greek plates can be found in Anagnostopoulos, Anagnostopoulos, Loumos, and Kayafas (2006). The system is part of an intelligent transportation system. Moreover, this article also contains a detailed overview of other procedures used for license plate recognition. Some of them are compared with respect to recognition performance and processing time to the proposed system in Anagnostopoulos et al. (2006). Firstly, the system scales the license plates to fixed size. Afterwards, the authors apply a binarization method with a fixed threshold linked to the characters' size. The alignment of the characters is performed using connected component analysis. According to the size of the connected components fragments are rejected that do not fulfill certain rules. For classification they apply a probabilistic neural network, which leads to a recognition rate of more than 85% of correctly classified characters on five different datasets. A very detailed survey of license plate recognition was written by Anagnostopoulos, Anagnostopoulos, Psoroulas, Loumos, and Kayafas (2008). Wen et al. (2011) proposed a license plate recognition system that is capable for different characters, e.g., Latin, Arabic, Kanji or Chinese characters. For character segmentation an algorithm for local binarization is suggested, which is robust to changing illumination in the plates caused by shadows. Afterwards, noise in the image is removed by applying a median filter followed by a connected component analysis. The centers of the connected components are used for skew correction of the lines. A projection profile is used for slant correction and character segmentation. Geometrical and topological features are used for classification, which is performed with support vector machines. The authors report a overall recognition performance of more than 97% with a processing time of 126 ms for segmentation and recognition on a 1.8 GHz Intel Core processor with 1.5 GB main memory. The authors also report that segmentation fails if the resolution of the plate images is too low. Another system was introduced in Thome, Vacavant, Robinault, and Miguet (2011), where the authors emphasize that their system is fully independent of country specific features, which makes the system very universal and can be applied in most countries. The images are binarized by an adaptive thresholding that is necessary for connected component analysis. Based on some geometrical features of the components noise and other fragments are removed. Afterwards misalignment of the lines and characters are corrected. In the case that the width to height ratio of the components exceeds a given limit a projection profile is applied to segment the characters again. They report a recognition performance of more than 90%. The processing time on a 2.66 GHz Intel Pentium IV machine with 512 MB main memory is about 100 ms. A license plate recognition system running on video data was introduced in Sarfraz et al. (2011). Firstly, they apply a preprocessing step to remove noise and to amplify edges. The connected component analysis is used for character segmentation. Classification is done by a nearest neighbor approach. In order to improve the confidence they apply the recognition algorithm to every video frame and combine several results. Depending on the dataset, this leads to a recognition rate between 90% and 100%.

A further related research field covers character recognition in different kind of industrial applications or with mobile devices. One system for reading DOT numbers on tires was introduced by Koo and Oh (1996). The system is used to sort tires and to perform a statistical evaluation of stocking and market demand over months and years automatically. Before character segmentation they perform a preprocessing step to improve the image quality and the skew of the lines is corrected. Character segmentation is done using projection profiles. The segmented characters are classified by three neural networks, each of which having been trained with a subset of 36 classes. The recognition performance is about 95%. For reading flight and serial numbers of flight tickets automatically a system was proposed in Zhao and Wang (2003). Firstly, illumination differences are equalized in the images. Then the alignment of the flight tickets is performed, which is followed by an adaptive thresholding for binarization. Character segmentation is based on connected component analysis. A convolutional neural network is applied for classification of the ten classes. The system was successfully tested with an airline in Shanghai. It processed more than half a million flight tickets with a recognition rate of more than 93%. An alphanumeric character recognition framework for mobile phones was introduced by Laine and Nevalainen (2006). In the first step the image is binarized using a global threshold. Then skew of the lines is corrected, which needs about 50% of the overall precessing time. Character segmentation is performed on binarized images using a projection profile. They report a processing time of 40 s for classification of all characters in one image on a mobile phone Nokia 6630. Choi, Yun, Sim, and Kim (2010) proposed a character recognition system that reads slab identification numbers. The images are binarized with a global adaptive thresholding. Then the lines are segmented using a projection profile. Afterwards, connected component analysis is applied and components are rejected that do not fulfill geometrical properties of a character. According to the authors, the algorithm has more than ten parameters that must be adapted manually. The processing time is about 1 s per image on a computer with a 3 GHz Intel Core 2 Duo processor with 2 GB main memory. Systems for automatic meter-reading using mobile phones were also proposed like the system of Oliveira, dos Santos Cruz, and Bensebaa (2009), which can be used for reading an electric meter. Contrast is enhanced and illumination differences are equalized in a preprocessing step. Then an adaptive thresholding is applied. The nearest neighbor classifier is used for classification of the ten classes. A similar system for water meters was introduced by Nava-Ortiz, Gomez, and Diaz-Perez (2011). A local thresholding is applied for binarization, whereas the threshold is determined adaptively. For the best combination of features and classifiers tested, the authors report a recognition rate of approximately 93% with a processing time of 30 ms per character on a Nokia N80. Another system for reading codes on industrial containers was proposed by Bingxia, Na, Jia, and Quanming (2009). Classification is performed using pattern matching on geometrical features. This yields a recognition rate of 97%. An enhanced recognition system of container codes comes from Wu, Liu, Chen, Yang, and He (2011). They propose a thresholding method for binarization, which is robust to reflections. A projection profile is applied for character segmentation. Two support vector machines are used for classification, where one of them is trained