

Text detection

Zhixing CAO, Yuesong SHEN

20 février 2016

1 INTRODUCTION

The development of smartphones and the growing demands in content-based image understanding has made the text detection a crucial topic in machine-human interaction. It has been shown that the performance of image retrieval depends critically on the performance of text detection and recognition. For example, two book covers with different titles but identical background prove to be considered virtually indistinguishable without detecting and recognizing the text.

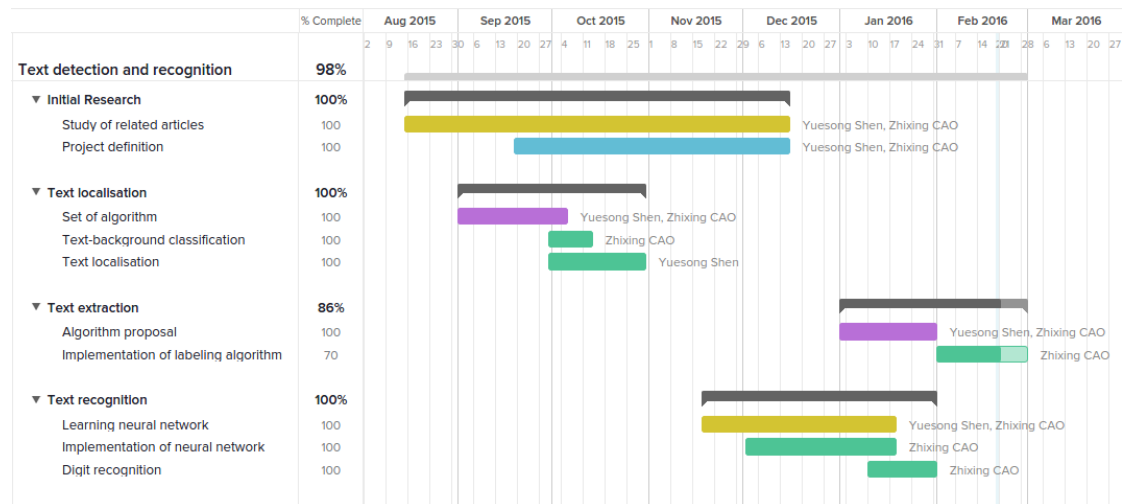
1.1 GOAL OF PROJECT

Early approach of text detection and recognition techniques such as OCR can be traced in early 1900s. Recent years, with the progress in the field of machine learning pattern recognition and text localisation techniques make new breakthroughs. With the study of some of those techniques, our project concern algorithms that can decode the text in images.

The approach of our project can be divided into two parts — text detection and text recognition. The detection part detect potential texts in images and output pure text candidates and the recognition part translate the image of text candidates to digital texts that can be read by machines.

1.2 WORK BREAKDOWN

The whole project is carried out by Zhixing CAO and Yuesong SHEN. Several works have been carried out independently during the project. Here is the overview of work breakdown.



2 STATE OF THE ART

2.1 TEXT DETECTION

As an essential prerequisite for text recognition, text within images has to be robustly located. This is challenge task due to the variety of the text form, such as variations in languages, font and style, geometric and photometric distortions, partial occlusion, and lightening conditions. Text detection problem have been studied a lot in recent researches and numerous methods are reported in the literature.

All the methods used in recent research can be classified into two catagories : method based on texture and method based on connected componet.

Texture-based method views texts as a special texture that is distinguish from its background. Features are extracted over some special regions and a classifier is used to identify texts areas. Zhong and al have segmented caption text regions from background and used the intensity variation information. Ye and al have proposed a method using multiscale wavelet feaitres and a coarse-to-fine algorithm to locate text lines under different backgrounds.

Different with the texture-based method, the connected component based approach ex-trats regions from the image and selects text candidates using some geometric rules. In IC-DAR 2005 text locating competition, the best result applies have used an adaptive binariza-tion method to find connected components and forms text lines based on geometric proper-ties. Recently Chen and al extract letter candidates by employing edge-enhanced Maximally Stable Extremal Regions and using geometric and stroke width information to exclude non-text objects. They have achieved an accuracy score of 95%.

2.2 TEXT RECOGNITION

Converting text data from image and deciphering into digits is an important problem. Early physical photocell-based OCR implemented matrix matching by comparing an image to a stored glyph on a pixel-by-pixel based. Those algorithms involve mostly extensive proces-

sing on the image such as thinning, smoothing contour analysis etc. because the majority of previous works uses geometrial and topological features. In recent research community the dominant approach to this problem is based on machine learning techniques — a general inductive process builds automatically a classifier by learning.

In the textbook *Pattern Recognition and Machine Learning*, Bishop reflects recent developments in the field of pattern recognition and machine learning and shows potential usages of machine learning method in patter recognition. The early effort has been made around 1986s by Burr, Mehr&Richfield to implement a neural networks in character recognition. Recently, with the development of paralel computation and the use of GPU, efficiant OCR system based on neural network is realized.

2.3 OUR APPROACH

In this report, we show you our approach of the text detection problem by combined texture-based method and connected component based method and our text recognition algorithm using neural network.

We adapte at first the method proposed by Liu and al to locate texts. This algorithm detects text candidate by applied classifier on contour pictures of the original image. Experimental results demonstrate that this approach is robust for font-size, font-color, background and languages, which can be used efficiently.

After locating the text, a connected component based method is used to extract word candidates. We find all components by using connected-component labeling algorithm. Then we use some geometric constraints and heuristic rules to merge them and extract letters and words candidates.

Our approach for text recognition is based on using neural network to classify characters. With a training set of different kind of characters, the neural network is constructed in order to match an input character to a learned one.

3 TEXT DETECTION

3.1 DETECTON OF CONTOUR

3.2 K-MEANS CLASSIFICATION

3.3 TEXT AREA IDENTIFICATION

3.4 CONNECTED COMPONENTS DETECTION

3.5 DETECTION OF WORDS

4 TEXT RECOGNITION

4.1 NEURAL NETWORK

4.2 PATTERN RECOGNITION

5 CONCLUSION