Text Recognition Using OpenCV Libraries For Devices Based on MeeGo

Iliya Mironov, Diana Ilina
Nizhny Novgorod State University
Nizhny Novgorod, Russia
frimen3000@mail.ru, ilina.diana@gmail.com

Abstract

OpenCV is an open source effective image processing library. This library expends image and video processing capabilities of MeeGo platform and can be used by various applications that process embedded cameras images and videos as well as any other images/videos. The library is cross-platform, and runs on Mac OS, Windows and Linux. The target of this project is analyzing the possibility of using OpenCV library with MeeGo devices and creating user application that use OpenCV. The tasks of our project are porting OpenCV to MeeGo, developing an example application that uses OpenCV, measuring resource consumption.

The aim of this project is to create a full application under MeeGo able to recognize text from imagesobtained from a camera mobile device and translate the resulting text into any other language.

To recognize the text we are plan to use the algorithms of computer vision library OpenCV. This library provides all the necessary algorithms to cover all phases of OCR. During of the work will be investigated performance of recognition algorithms on various mobile devices.

Index Terms: OpenCV library, MeeGo, cross-platform, OCR.

I. INTRODUCTION

The target of this project is to create applications for mobile devices under MeeGo able to recognize text from images of natural background, obtained from a camera mobile device and translate the resulting text into any other language. This article will be considered algorithms for detecting and text recognizers that exist today. Also during the development of algorithms may be used OpenCV library, which includes some of the necessary image processing functions

II. MAIN PART

The problem of recognizing text in images with the natural background is an important task for computer vision. Most of the algorithms recognize texts are designed for recognizers scanned text. In the case of text in natural images, the recognition problem becomes more difficult. But for all the algorithms recognizing text there is a sequence of steps that must be done. Next, consider the steps recognizers and analyzed each of them.

A. Steps recognition text

Typically, an algorithm for recognizing text is divided into several stages of image processing. When you implement the recognition algorithm, not all stages are required.

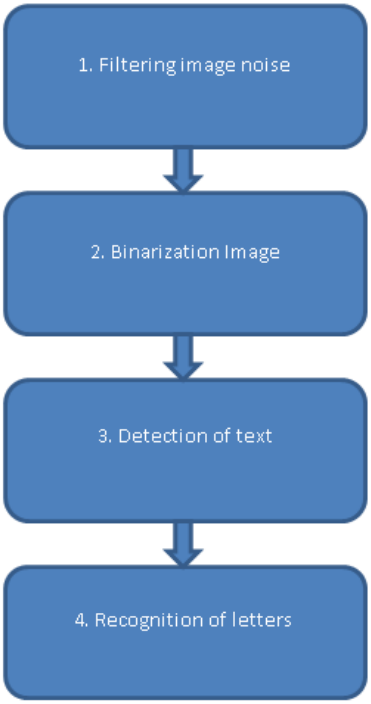

Fig. 1. Steps text recognition

The first two steps are not mandatory steps in recognizing the text. The process of filtering the image arises from the imperfections of the equipment lifts the image. If image filtering usually use Gaussian or median filtering. Binarization images simplifies the process of further image processing. Detection and recognition of text characters are the main stages. These stages can be used different algorithms depending on the task at hand. This article will describe the algorithms for detecting and recognizing text in a natural background.

B. Detection of text on the natural background

The text detecting algorithm though natural background is a difficult task of the machine vision. In this paper(note) will be presented a text detection algorithm based on the researches IDAR 2003[1], IDAR 2005[2] and also on the propose [3].

The main feature of the text that allocate it from another image elements is its nearly equal letters widths. This fact has been used in the work [3] for the text detection. The goal of that paper is to compute widths for that regions which a single considered pixel relate to. To compute stroke lets use the proposed in the paper [3] algorithm.

By an input image it is necessary to compute edges. For this we prefer to use Cany edge detector. Lets compute a gradient by initial image for each pixel related to the objects edges. So we shall take directions distribution that must be followed along objects width directions. Further consider a gradient direction for each pixel. From the considered pixel lets go by a ray along the pixel gradient direction. If the ray reach an opposite edge and both pixels gradients directions are different no more than $\pm\pi 6$, we

shall suppose that for all pixels under the ray width of the object which has a pixels group attached to is equal to the distance between the considered pixel, which the ray out of, and the nearest pixel by the opposite side, as depicted in Figure 2.

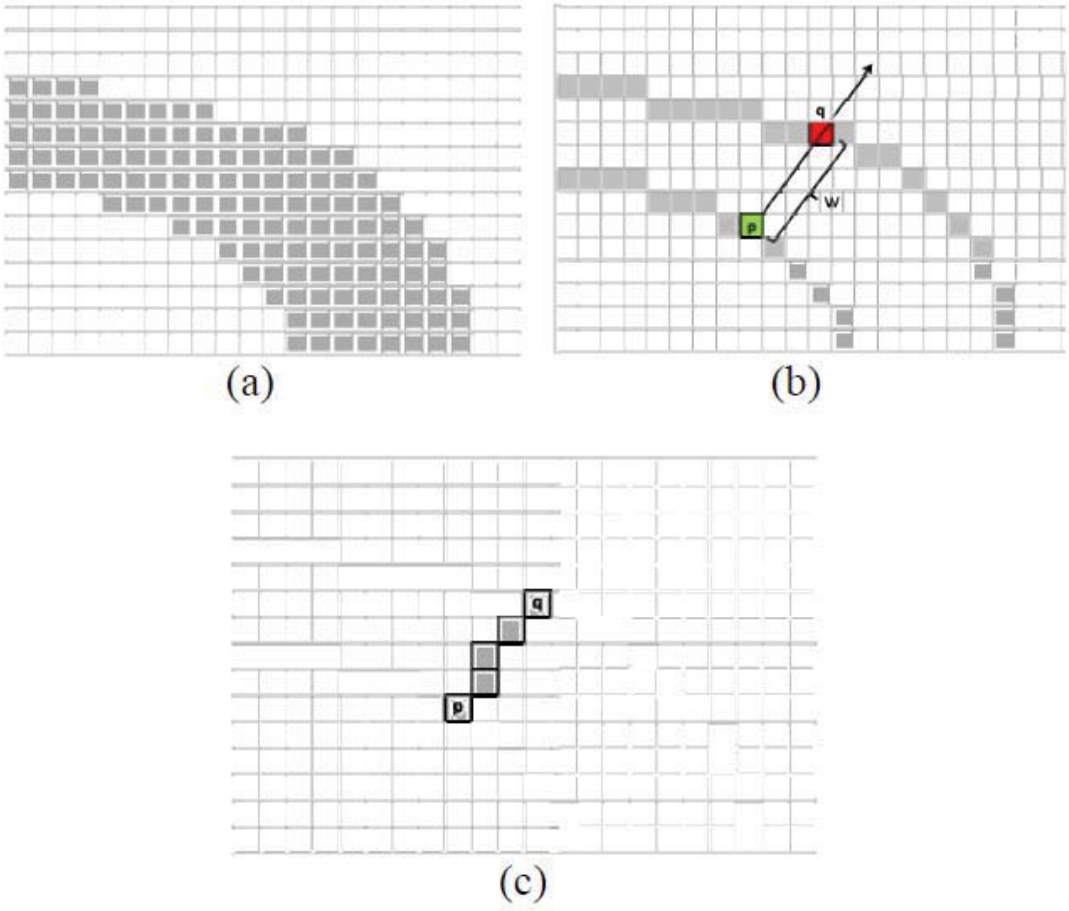

Fig. 2

(a)A typical stroke. The pixels of the stroke in this example are darker than the background pixels. (b) *p* is a pixel on the boundary of the stroke. Searching in the direction of the gradient at *p*, leads to finding *q*, the corresponding pixel on the other side of the stroke. (c) Each pixel along the ray is assigned by the minimum of its current value and the found width of the stroke.

Because objects on an image can take complex forms, so as strokes crosses, angles by the objects and closed objects contours, which can lead to wrong object's stroke width detection. Thus for the pixels which has covered by more than one ray it is necessary to compute average mean for the stroke width and set the considered pixel's stroke width equal to that average mean.

The next stage of the text detection is grouping pixels to letters. Two neighboring pixels can be grouped if theirs stroke widths are similar. However an object can have changeable stroke width. Thus to more accurate pixels grouping there are possible to take neighboring pixels widths ratio which must do not overdraw some threshold. From the grouped object we need to detect letters. For this there are computing stroke width dyspersia for each connected component and throwing aside that components for which stroke width overdraw the preassigned shreshold. Also we could limit letters sizes for reliability improving.

To futher text detection reliability improving it is necessary to select letters groups. Assumed that a text has a number of similar parameters, such as neighboring letters widths, width and height of letters. Therefore assume letters neighboring if their parameters ratio do not outnumber some threshold. At a result we would take letters sequence, which we can allocate on image.

C. Recognition of letters

Convolution neuron networks may be efficiently use to recognize patterns: faces and speech. They combine the following three architectures providing particular stability by scale, shift, rotation and other distorsion invariances:

- local receptor fields (allow take into account input data two-dimensional topology);
- paid total weights (provide general signs detecting reducing tuning parameters count) ;
- spatial hierarchical subsampling (allow build signs hierarchies).

REFERENCES

[1] “ICDAR 2003 robust reading competitions”, Proceedings of Seventh International Conference on Document Analysis and Recognition, 2003, pp. 682-68
[2] “*ICDAR 2005 text locating competition results*”, Eighth
International Conference on Document Analysis and Recognition, 2005. Proceedings. pp. 80-84(1)
[3] “Detecting Text in Natural Scenes with Stroke Width Transform”, Boris Epshtein, Eyal Ofek, Yonatan Wexler, IEEE 2010