

Document Scanner

Leopold Gaube and Florian Betz

May 5, 2022

Contents

1	Requirements	2
2.1	Document Localization	2
2.2	Perspective Transformation	2
2.3	Post-Processing	2
	List of Acronyms	3
	References	3

1 Requirements

There are already many document scanners on the market that can digitalize a paper or similar. The quality of the document scanning depends on the used algorithms. To guarantee a good scan, a perspective transformation should be used.

A comfortable library to use computer vision functions is Open Source Computer Vision Library (OpenCV). OpenCV is an open source computer vision and machine learning software library. [Ope22] You can use OpenCV with the programming languages C, C++, Python or Java.

2

Our algorithm for scanning a document in an image consists of three main parts, each with its own challenges.

First, we will have to localize the document in the image. Then we will use a perspective transform from the corner points of the document to the entire span of the image in order to obtain a top-down view without any distortions. Finally, we will use a binary filter to distinguish text from background.

2.1 Document Localization

Some smartphone scanner apps require the user to mark the four corner points of the document manually. However, we want to automate this process by detecting the corner points using Computer Vision.

Detecting the corner points accurately and consistently is probably the hardest and most crucial step of the entire algorithm.

2.2 Perspective Transformation

top-down view

2.3 Post-Processing

binary image

simple thresholding vs adaptive Thresholding

List of Acronyms

OpenCV Open Source Computer Vision Library

References

[Ope22] OpenCV. Opencv - about. <https://opencv.org/about/>, 21.04.2022.