

Image and Video Processing Coursework

CS3330: Image and Video Processing Coursework: Document Image Binarization

Damon Tran

Student Number: 160111190

Background

What is the aim of document image binarization?

Document image binarization is essentially a method that is used for taking an image of a text document and separating different objects within it generally foreground (text) from background, using various approaches via thresholding on a grayscale image, to separate the text from the rest of the picture to make it more readable [1].

The typical documents you would use this on, would be where the image is affected by various variables that make the document hard to read as displayed in figure 1 below [2]. You can see that the black text although visible, it is slightly difficult to see due to the dark patches and stains. However, through applying image binarization you can solve this issue.

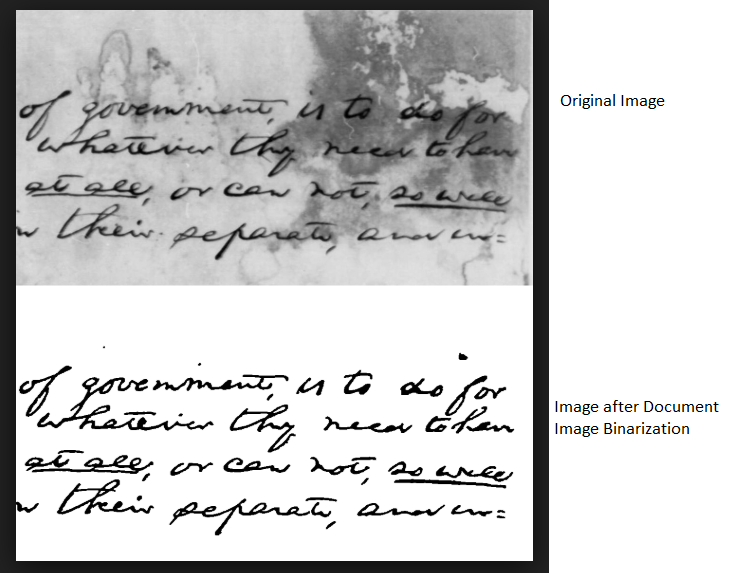


Figure 1

Why is binarizing a document image useful? (hint: binarization is a mid-level image processing technique, and the output of mid-level techniques is often fed into high-level techniques).

It is useful because mid-level processes take an image as an input, analyses and processes the image to obtain the relevant attributes which are then passed to a high-level processing technique to apply to said attributes [3]. Acting as a phase in the processing pipeline, setting the image up for other steps to achieve a specific outcome. Meaning that we can apply any high-level technique we wish to the outcome we’ve generated from the mid-level process, depending on our goal the flexibility of what we can do is quite varied.

For example, if we wanted to restore historical documents from many years ago despite the degradation of the paper and any issues that may obstruct us from reading it, If we wanted we could apply image binarization so that we could and then apply any high level technique depending on what we want to produce.

// In this case it is helpful because we are able to take the grayscale image of a text document and be able to retrieve the information, we need to binarize it in such a way that makes it able to read. Meaning that we can apply any high-level technique we wish to the outcome we’ve generated so depending on our goal the flexibility of what we can do is quite varied.

• What are the existing approaches to document image binarization? Briefly outline THREE different algorithms for document binarization found through your research and discuss the way in which the aims and performances of these algorithms differ in academic literature.

Existing approaches to document image binarization

(Note they differ because the approaches are only good depending on context, context mean that they are being applied to different images and thus different goals it is unfair to say that one algorithm is better than the other at something because they have different aims)

There are many approaches but the three I have researched and found are Edge Level Thresholding, Local Adaptive Binarization and Sauvola.

The Edge Level Thresholding (ELT) algorithm is based off the idea when you binarize and image with a threshold value it is impossible to find a single value that can do it efficiently for the whole image [4]. Instead, it uses multiple threshold values which are derived from each pixel and applied to various sub-regions to solve the problem of thresholding an image which is affected by illumination [4].

Local Adaptive Binarization (LAB) works with the concept of creating sub-images of the image to binarize and then identifying and classifying pixels in order to separate the image by text from background [5].

There are two sub-images, one small and large. The pixels in the large image are clustered together via Otsu’s threshold selection method [5]. Then regions are checked to see if it contains one or two classes. The smaller sub-image’s pixels are adjusted by classifying them in conjunction to the results of the larger image [5].

Penis

**References**

1. Karthika M, Ajay James , A Novel Approach for Document Image Binarization using Bit-Plane Slicing, Romania , Tirgu-Mures, 2014.
2. https://webee.technion.ac.il/labs/cgm/Computer-Graphics-Multimedia/Software/Software.html, 26th November 2018.
3. Dr. Harry Goldingay, Image and Video Processing CS3330 Introduction, Aston University.
4. J. R. Parker, C. Jennings, Thresholding Using an Illumination Model, ICDAR ’93, Japan, 1993.
5. Line Eikvil, Torfinn Taxt and Knut Moen, A Fast Adaptive Method for Binarization of Document Images, France, 1991.