

## **Picture Search: Image to Information**

Jingjing Wang  
Zicheng Hu  
Victor Zhu

03/04/2016

### Introduction:

As early as the 1950s, there has been a lot of research on word recognition. However, much of the research is based on plain text pictures and extracting characters from them. In our project, we are going to extract information about objects by analyzing images of said object into a machine editable form.

We believe that this project will be able to direct use. By snapping a picture of an object, we can extract a bunch of information that can be used to learn more about it. Take for example, if you analyze a Lays bag of chips, theoretically our program would generate the phrases such as "Lay's", "Sour Cream & Onion", "Family Size!", and more! These words would be saved onto a text file, and the next step for the user or future software would be to search these terms on the internet for more information.

Some future implications for this project would be to possibly create software that will accept images of objects, and manage to pull up information on the object from the internet. Similar to how there are applications out there that can accept a picture of a bar code and pull up the exact product or link online, we would like to bring it a step up and make it an image of the actual object instead. In the real world, we often see things that we know nothing about, sometimes not even the name of it. Such an application would see great widespread use. Another future use would be restaurant menu breakdown, where we could have a software that analyzes a picture of a menu and gives the user information on all the options listed. With the large scale use of smartphones in this day, such an application would be simple, quick and easy to use on a smartphone.

### Goals:

In this project we will analyze objects in an image with the purpose of extracting the information printed on said objects. We will take the text off of objects, and print it in text format. Currently, we have found research on text and object recognition, and our job is to combine those two recognition methods and make them useful for our application.

### Difficulties:

- How to recognize texts along with the objects they rest on
- How to locate the regions that include texts
- How to distinguish texts from the object if they have similar color
- How to extract texts from object segment
- How to convert the image extraction into contents in the text file

### General structure:

In order to extract objects and text from a picture we need to combine object recognition and text recognition techniques. Because a picture may contain more than one object and for the purpose of matching words with responding object, we want to extract objects first.

To extract objects from an image, we are going to use the method already developed in MATLAB: Image Processing Toolbox. This toolbox provides comprehensive set of algorithms and functions for us to explore and convert into the version that we need. Some objects have words on them while some do not. We are going to categorize them into two groups for the conveniency of the next step: text extraction.

In the paper by Ivan Dervisevic, Dervisevic suggested optical character recognition depends on feature extraction and classification algorithms. His research is focused on single typewritten characters by viewing it as a data classification problem. However, most OCR (Optical Character Recognition) applications take advantage of the fact that text is presented on plain background that has sufficient high contrast to text. In our case, we are trying to extract text which is printed on the object, where the text may be the similar color as the object, like words on the black iPhone do. We are going to solve this problem by doing some data classification to put pixels of two colors into two categories. Then we are going to change the color of one group into its complementary color. This way we are able to extract text from a clear background.

After doing object/text recognition and data classification, we are going to finalize our project by outputting a file with object and information in a text file.

Distribution of individual efforts:

Victor will focus on object recognition research along with the coding part. Jingjing and Zicheng will focus on text recognition along with the coding part. All three of us will do on data gathering together. We will decide the amount of work at the beginning of every week.

## References

### Survey Papers:

Text Extraction from Grayscale Document Images Using Edge Information

<https://pdfs.semanticscholar.org/fdeb/b6719483e591df151d4665afe77d8a11a3d4.pdf>

Extraction of Text Regions in Natural Image

<http://scholarworks.rit.edu/cgi/viewcontent.cgi?article=1165&context=theses>

Machine Learning Methods for Optical Character Recognition

<http://perun.pmf.uns.ac.rs/radovanovic/dmsem/completed/2006/OCR.pdf>

### Methods References:

Automatically Detect and Recognize Text in Natural Images in MATLAB

<http://www.mathworks.com/help/vision/examples/automatically-detect-and-recognize-text-in-natural-images.html?requestedDomain=www.mathworks.com>

Feature extraction methods for off-line recognition of segmented (isolated) characters

<http://www.sciencedirect.com/science/article/pii/0031320395001182><http://www.sciencedirect.com/science/article/pii/0031320395001182>

Perform image processing, analysis, and algorithm development

<http://www.mathworks.com/products/image/>