

Understanding Collections of Images

COS 521 Final Project Report

Steven Englehardt, Maciej Halber, Elena Sizikova

January 10, 2014

Contents

1	Introduction	2
1.1	Abstract	2
1.2	Background Work	2
2	Methods	4
2.1	Data	4
2.2	Implementation	4
2.3	Color Analysis	4
2.4	Fast Fourier Transform	4
2.5	Singular Value Decomposition	4
3	Analysis	5
4	Suggestions for Further Work	6

Chapter 1

Introduction

1.1 Abstract

This report explores a variety of image properties that make it possible to understand vast collections of images better. In particular, we look at how image color, saturation, sharpness, and detail can be extracted and compared between images using methods such as Fast Fourier Transform (FFT) and SVD (Singular Value Decomposition). We seek to understand how the theoretical underpinnings of these two algorithms affect the way the images are created in the first place. Ultimately, we provide a way of decomposing the image into mathematical notation (a descriptor) that differentiates well between a collection of images.

1.2 Background Work

There are many possible situations in which we would need to understand and compare image structure. For example, one might like to search for a location in which a photograph was taken, by looking at all the other available images, and finding the image closest to the search image. Alternatively, one may want to cluster images based on their content, and see what categories the image collection can be decomposed to. Both of these would be easy problems to solve, if the images were annotated with words: textual search is

a well-solved problem. However, when the images are not labelled (this is known as unsupervised learning), and the image collection is extremely large, it is impractical to label the images by hand. For such problems, it is important to analyze image content automatically.

Existing methods of image search by analyzing content of the image include Google Goggles and Google Image Search, both are based on similar technology [1], which checks for distinctive points, analyzes lines and textures and finally creates a mathematical model of the image. While the exact implementation is not available, Google Image Search does not provide clustering capabilities of analyzing existing input datasets. A more closer work is that of Oliva and Torralba [2] which create a GIST descriptor (cite Freedmans work!!!)

Chapter 2

Methods

2.1 Data

2.2 Implementation

2.3 Color Analysis

2.4 Fast Fourier Transform

2.5 Singular Value Decomposition

Chapter 3

Analysis

Chapter 4

Suggestions for Further Work

Bibliography

- [1] Johanna Wright, *Search by Text, Voice, or Image*. Inside Search: Official Google Search Blog, 2011.
- [2] Aude Oliva, Antonio Torralba, *Modeling the Shape of the Scene: a Holistic Representation of the Spatial Envelope*. International Journal of Computer Vision, Vol. 42(3): 145-175, 2001.
- [3] Leslie Lamport, *L^AT_EX: A Document Preparation System*. Addison Wesley, Massachusetts, 2nd Edition, 1994.