MOBILE VISUAL SEARCH

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Computer Science

by

Khanh Le

June 2018

COMMITTEE MEMBERSHIP

TITLE:	Mobile Visual Search
AUTHOR:	Khanh Le
DATE SUBMITTED:	June 2018
COMMITTEE CHAIR:	Alexander Dekhtyar, Ph.D. Professor of Computer Science
COMMITTEE MEMBER:	Aaron Keen, Ph.D. Professor of Computer Science
COMMITTEE MEMBER:	Franz Kurfess, Ph.D. Professor of Computer Science

INTRODUCTION

Mobile Visual Search

Khanh Le

Visual search is an old and well-studied topic in computer vision and is integrated into many widely-used search engines such as Google, Bing, and DuckDuckGo. Instead of submitting a keyword or text to the system and getting relevant documents back, user will submit an image and the system will return images that have similar visual contents. However, mobile product visual search recently emerges to be an interesting and challenging topic again because of the growth of the smartphones and applications that go along with it such as online shopping. Ideally, a user can take a picture of a product using a mobile phone to identify the product and then retrieve visually similar products from the database.

Traditional approach to web-scale image retrieval relies on bag-of-visual-words model that is initially introduced in [1]. In this model, local features such as SIFT[2], SURF[3], or ORB[4] are extracted from the query images. Those features are then assigned to their closest visual words in a visual vocabulary. A visual vocabulary is constructed by extracting local features from all of the images in the database, then using k-means clustering to find the centers of each cluster. The centers are the visual words of the vocabulary. The query image is accordingly represented by a global histogram of visual words, and matched with database images by tf-idf weighting using inverted files [5].

Application of traditional method to mobile visual search is shown to be effective in identifying duplicate and rich-textured photos. However, its performance is generally poor in retrieving general product categories because of variations in illumination, scale, viewpoint, blur existence in the query image.Therefore, the results may be affected by redundant features extracted from the background of the query image.

In this paper, a object detection and localization scheme is integrated along with the bag-of-visual-words model to achieve better performance in image retrieval

ACKNOWLEDGMENTS

Thanks to:

- Andrew Guenther, for uploading this template

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

BIBLIOGRAPHY

[1] L. T. F. J. C. S. He, J. Mobile product search with bag of hash bits. 2011.

[2] D. Lowe. Distinctive image features from scale-invariant keypoints. 2004.

[3] S. H. Nister, D. Scalable recognition with a vocabulary tree. 2006.

[4] C. O. I. M. S. J. Z. A. Philbin, J. Object retrieval with large vocabularies and fast spatial matching. 2007.

[5] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos, 2003.