國立臺灣大學工學院機械工程學系

碩士論文

Department of Mechanical Engineering
College of Engineering
National Taiwan University

Master Thesis

系統最佳化實驗室碩士論文

System Optimization Lab

張祐晨

Yu-Chan Chang

指導教授：詹魁元博士

Advisor: Kuei-Yuan Chan, Ph.D.

中華民國 104 年 7 月

July, 2015

# 國立臺灣大學（碩）博士學位論文
# 口試委員會審定書

論文中文題目

論文英文題目

本論文係○○○君（○學號○）在國立臺灣大學○○學系、所完成之碩（博）士學位論文，於民國○○年○○月○○日承下列考試委員審查通過及口試及格，特此證明

口試委員：

（簽名）

（指導教授）

系主任、所長 （簽名）

（是否須簽章依各院系所規定）

# 誌謝

感謝…

# Acknowledgements

I’m glad to thank…

# 摘要

本論文提出了一影像中使用者感興趣區域 (region of interest) 偵測之資料集 (benchmark)。使用者感興趣區域偵測在許多應用中極為有用，過去雖然有許多使用者感興趣區域之自動偵測演算法被提出，然而由於缺乏公開資料集，這些方法往往只測試了各自的小量資料而難以互相比較。從其它領域可以發現，基於公開資料集的可重製實驗與該領域突飛猛進密切相關，因此本論文填補了此領域之不足，我們提出名為「Photoshoot」的遊戲來蒐集人們對於感興趣區域的標記，並以這些標記來建立資料集。透過這個遊戲，我們已蒐集大量使用者對於感興趣區域的標記，並結合這些資料成為使用者感興趣區域模型。我們利用這些模型來量化評估五個使用者感興趣區域偵測演算法，此資料集也可更進一步作為基於學習理論演算法的測試資料，因此使基於學習理論的偵測演算法成為可能。

關鍵字： 關鍵字

# Abstract

This thesis presents a benchmark for region of interest (ROI) detection. ROI detection has many useful applications and many algorithms have been proposed to automatically detect ROIs. Unfortunately, due to the lack of benchmarks, these methods were often tested on small data sets that are not available to others, making fair comparisons of these methods difficult. Examples from many fields have shown that repeatable experiments using published benchmarks are crucial to the fast advancement of the fields. To fill the gap, this thesis presents our design for a collaborative game, called Photoshoot, to collect human ROI annotations for constructing an ROI benchmark. With this game, we have gathered a large number of annotations and fused them into aggregated ROI models. We use these models to evaluate five ROI detection algorithms quantitatively. Furthermore, by using the benchmark as training data, learning-based ROI detection algorithms become viable.

**Keywords:** keyword

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Attention plays an important role in human vision. For example, when we look at an image, our eye movements comprise a succession of *fixations* (repetitive positioning of eyes to parts of the image) and *saccades* (rapid eye jump). Those parts of the image that cause eye fixations and capture primary attention are called *regions of interest* (ROIs). Studies in visual attention and eye movement have shown that humans generally only attend to a few ROIs. Detecting these visually attentive regions in images is challenging but useful in many multimedia applications, such as automatic thumbnail cropping, object recognition, content-based image retrieval, adaptive image compression and automatic browsing in small-screen devices.

Many algorithms have been proposed for automatic ROI detection in images. Unfortunately, these methods were often evaluated only on specific and small data sets that are not publicly available. The lack of published *benchmarks* makes experiments non-repeatable and quantitative evaluation difficult. However, as recommended by the latest ACM SIGMM retreat, repeatable experiments using published benchmarks are important for advancing the multimedia research field [**?**].

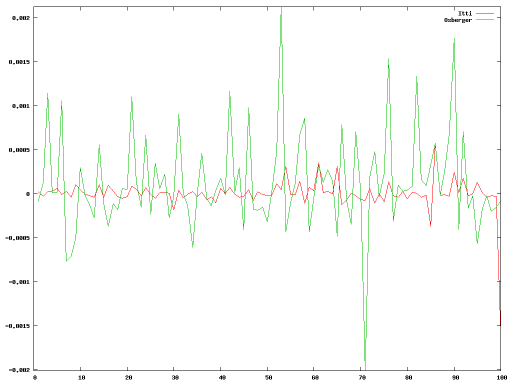


Figure 1.1: kl-distance

| | Itti's method | Fuzzy growing |
|---|---|---|
| Precision | 0.4475 | 0.4506 |
| Recall | 0.5515 | 0.5542 |

Table 1.1: Evaluation of FOA sets.