國立臺灣大學電機資訊學院資訊工程學系

碩士論文

Department of Computer Science and Information Engineering
College of Electrical Engineering and Computer Science
National Taiwan University

Master Thesis

非監督式什麼的還沒想好

Unsupervised Clustering of Related Chinese Words

李嘉玲

Chia-Ling Lee

指導教授：許永眞 博士

Advisor: Jane Yung-jen Hsu, Ph.D.

中華民國 103 年 6 月

June, 2014

國立臺灣大學
資訊工程學系

碩士論文

非監督式什麼的還沒想好

李嘉玲 撰

103
6

國立臺灣大學碩士學位論文
口試委員會審定書

非監督式什麼的還沒想好
Unsupervised Clustering of Related Chinese Words

本論文係李嘉玲君 (R00922072) 在國立臺灣大學資訊工程學研究所完成之碩士學位論文，於民國 103 年 6 月 28 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：

_____

_____

_____

_____

_____

所主任 _____

Acknowledgments

I'm glad to thank . . .

致謝

感謝...

中文摘要

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爲有用，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，我們提出名爲「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，並結合這些資料成爲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爲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爲可能。

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.

Contents

口試委員會審定書 **i**

Acknowledgments **iii**

致謝 **v**

中文摘要 **vii**

Abstract **ix**

1 Introduction **1**
1.1 Motivation 1
1.2 Objectives 1

2 Background and Related Work **3**

3 Methodoloy **5**
3.1 Framework 5
3.2 Problem Definition 5
3.2.1 Terminology and Notation 5
3.2.2 Input and Output 5
3.3 Word-to-word Similarity 5
3.3.1 Bag-to-bag Similarity 5
3.3.2 Vector-based Similarity 5
3.4 Clustering 5
3.5 Ensembled Clustering 5
3.6 Walk-through Example 5

4 Experiment **7**
4.1 Dataset 7
4.2 Evaluation Metrics 7
4.3 Experimental Result 7
4.4 Human Experiment 7
4.5 Discussion 7

5 Conclusion **9**
5.1 Summary of Contributions 9
5.2 Future Work 9

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

1.2 Objectives

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.

Chapter 2

Background and Related Work

2.1 Preliminary Work

2.2 Related Work

Chapter 3

Methodoloy

3.1 Framework

3.2 Problem Definition

3.2.1 Terminology and Notation

3.2.2 Input and Output

3.3 Word-to-word Similarity

3.3.1 Bag-to-bag Similarity

3.3.2 Vector-based Similarity

3.4 Clustering

3.5 Ensembled Methods

3.6 Walk-through Example

Chapter 4

Experiment

4.1 Dataset

4.2 Evaluation Metrics

4.3 Experimental Result

4.4 Human Experiment

4.5 Discussion

Chapter 5

Conclusion

5.1 Summary of Contributions

5.2 Future Work