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

A Benchmark for Region-of-Interest Detection in Images

Huang Tz-Huan

Advisor: Chuang Yung-Yu, Ph.D.

96 6
June, 2007

Acknowledgments

I'm glad to thank ...

...

(region of interest) (benchmark)
Photoshoot

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

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.

Chapter 2

Real-valued Optimization Algorithms

Overview of real-valued optimization

2.1 Covariance Matrix Adaptation Evolution Strategy

2.2 Standard Particle Swarm Optimization

2.3 Ant Colony Optimization for Continuous Domain

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 3

Multi-armed Bandit Algorithms

Describe the exploration vs. exploitation dilemma.

3.1 The Multi-armed Bandit Algorithm

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 4

Linear Transformation

Overview of real-valued optimization

4.1 Affine Transformation

4.1.1 Translation

4.1.2 Rotation

4.1.3 Scaling

4.1.4 Shear

4.1.5 Affine transformation matrix

Augmented Matrix

4.2 Projective Transformation

4.3 Optimization for Transformation Matrix

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 5

The New Bandit Technique

Overview of real-valued optimization

5.1 Framework of the New Bandit Algorithm

5.2 Initialization

5.3 Remain Evaluations Allocation

5.4 Recluster

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 6

Experiments

Overview of real-valued optimization

6.1 Test Problems

6.1.1 CEC2005 25 benchmark problems

6.2 Experiment Settings

CMA-ES initial mean and std, population settings SPSO2011 parameters (c, w) settings, and population settings

6.3 Results

6.4 Discussion

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 7

Conclusion

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.