Webpage Saliency

Chengyao Shen[1,2] and Qi Zhao[2]

[1]NUS Graduate School of Integrated Science and Engineering, [2]Department of Electrical and Computer Engineering, National University of Singapore

Introduction

Motivation

Webpage is becoming a more and more important visual input. There are, however, few studies on saliency in webpages.

Contribution

1. An eye tracking dataset from 11 subjects on 149 webpages. It is the first dataset on webpage saliency.

2. A new computational model for webpage saliency. It is the first model for webpage saliency that is purely based on visual features.
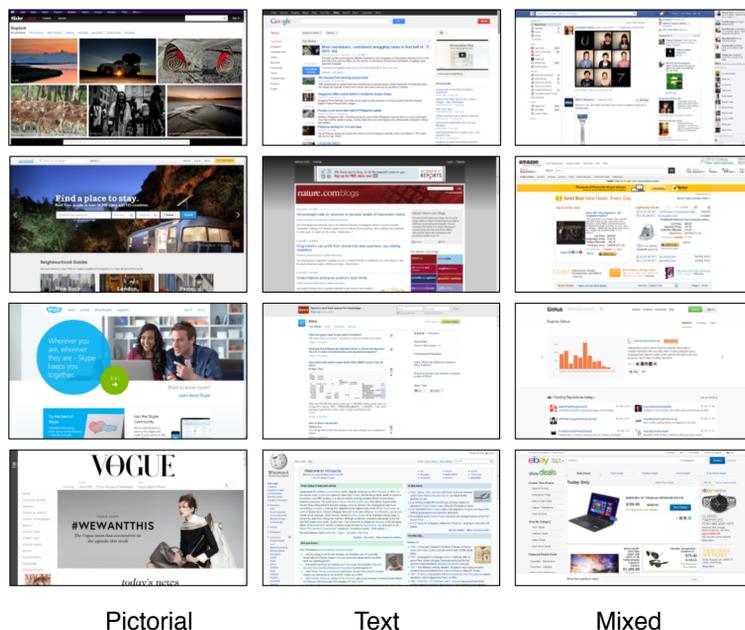
Dataset

FiWI Database

Fixations **i**n **W**ebpage **I**mages Database.

Stimuli: 149 screenshots of webpages rendered in Chrome browser in full screen mode (1360 x 768) in 3 categories

– **Pictorial**: One dominant picture or several large thumbnail pictures and with less text.

– **Text:** Informative text in high density.
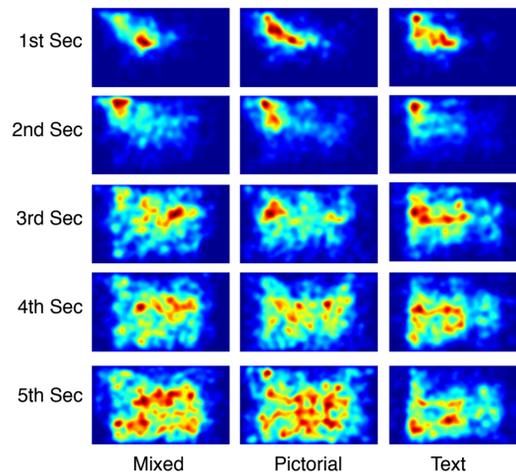
– **Mixed:** A mix of thumbnail pictures and text.

Pictorial — Text — Mixed

Subjects: 11 students from NUS (4 males and 7 females, 21-25)

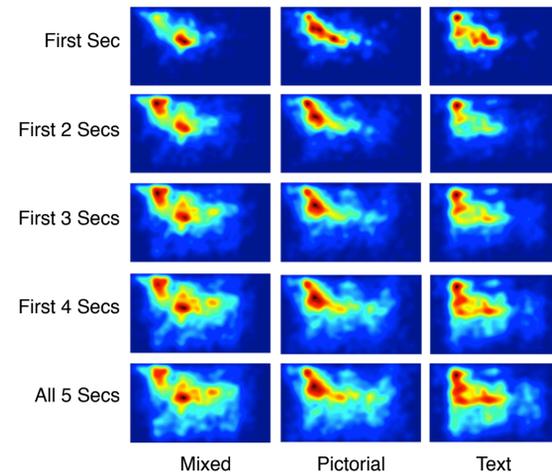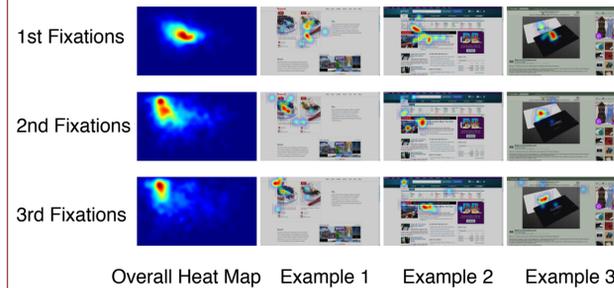**Procedure**: Free-viewing for 5 seconds.

Observations

Positional Bias and **Category Difference** in a second-by-second visualization

1st Sec / 2nd Sec / 3rd Sec / 4th Sec / 5th Sec

Mixed — Pictorial — Text

Positional Bias and **Category Difference** in accumulated fixation maps

First Sec / First 2 Secs / First 3 Secs / First 4 Secs / All 5 Secs

Mixed — Pictorial — Text

Positional Bias and **Object/Text Preference** in the first 3 fixations

1st Fixations / 2nd Fixations / 3rd Fixations

Overall Heat Map — Example 1 — Example 2 — Example 3

Object Representation from intensity and color maps in low spatial frequency.

Text Representation from four orientation feature maps in high spatial frequency

The Saliency Model

Feature Maps

– Multi-scale Feature Maps + Face Map + Positional Bias Map

Feature Integration

– Multiple Kernel Learning (MKL) for regression

Multi-scale Feature Maps:
Pyramidal center-surround and Gabor filters are applied on images in DKL color spaces in 6 scales. One more layer of threshold center-surround filter is used to inhibit false alarms caused by 'halo' effects and make the responses more concentrated on the objects.

Face Map:
Face detector and upper body detector are used.

Input Image — Intensity + Color (DKL Color Space) / Orientation — "Halo" effects are eliminated — Face Map / Positional Bias — Multiple Kernel Learning (MKL)
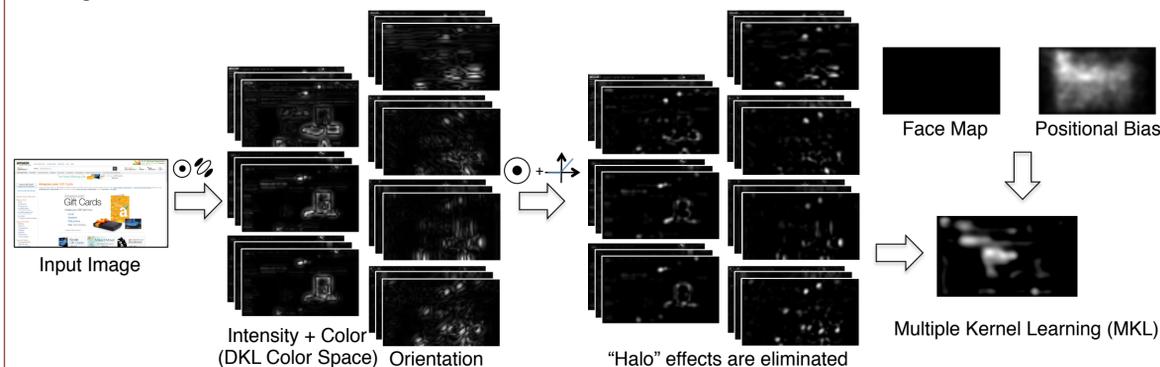
Results

Intra-Model Comparison

Feature Settings	CC	sAUC	NSS
Multiscale (no MKL)	0.2446	0.6616	0.8579
Multiscale	0.3815	0.7020	1.2000
Multiscale+Position	0.4433	0.6754	1.3895
Multiscale+Face	0.3977	**0.7206**	1.2475
Multiscale+Face+Position	**0.4491**	0.6824	**1.3982**

The performance of our model under different feature settings.

Feature Settings	CC	sAUC	NSS
Pictorial	0.4047	0.7121	1.2923
Text	0.3746	0.6961	1.1824
Mixed	0.3851	0.7049	1.1928

The performance of our model on three different categories in the FiWI dataset.

Inter-Model Comparison

CC / sAUC / NSS

GBVS — AIM — AWS — Signature — SUN — Judd — Our Model
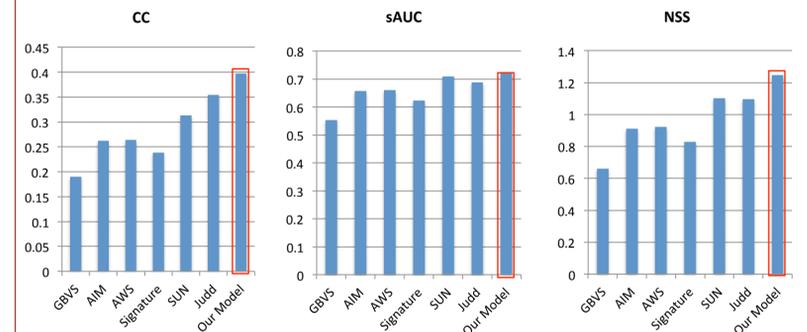
Qualitative Comparison

Stimuli — Human — Overlay — Ours — Ours+Position — GBVS — AIM — Signature — SUN — AWS — Judd

Resources