Statistical Analysis and Selection of Visual Fixations

Umesh Rajashekar[1] umesh@ece.utexas.edu
Ian van der Linde[1,2] ianvdl@ece.utexas.edu
Alan C. Bovik[1] bovik@ece.utexas.edu
Lawrence K. Cormack[1] cormack@psy.utexas.edu

[1]Center for Perceptual Systems, The University of Texas at Austin, USA; [2]Anglia Ruskin University, UK

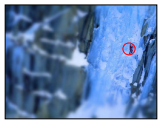
VSS 2006

1. Introduction

•Foveation
- •Non-uniform spatial resolution perception of the human eye
- •High resolution at point of gaze
- •Decreasing resolution towards periphery

Uniform resolution image

•Advantage of Foveation
- •Provides a large field of view without accompanying data glut

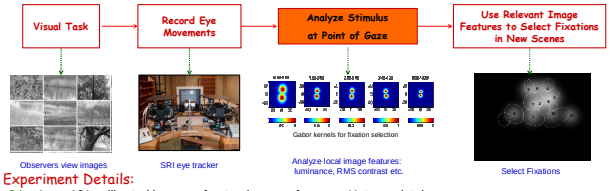


Multi-resolution retinal image when looking at ◯

•**Objective:** An important consequence of foveation is the need for eye movements to scan a scene. While eye movements are, no doubt, partially controlled by top-down (cognitive) mechanisms, their rapidity and sheer volume also suggests an influence of bottom-up (pre-cognitive) factors. Our goal was to evaluate the impact of low-level image features on eye movements, and to use this information to select fixations in scenes.

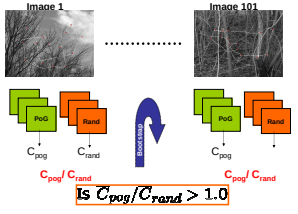
2. Methodology

Theme: *Determine which image features differ significantly between human and randomly selected fixations.*



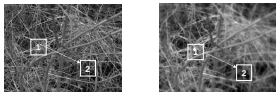
Experiment Details:
- •Stimulus : 101 calibrated images of natural scenes from van Hateren database
- •Task : 29 observers (24 naive) were instructed to scan the image 'efficiently' in 5s
- •Analysis : Statistics of local image features around point-of-gaze (***PoG***) were analyzed. Features included Luminance, Contrast, Bandpass profiles of luminance and contrast.

2a. Overview of Analysis



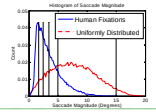
Do the statistics of image features around human **P**oint-**O**f-**G**aze (C_{pog}) differ significantly from those at **Rand**om fixations (C_{rand})?

2b. Eccentricity-based Analysis



Full-resolution analysis: All patches are analyzed at the resolution of the image.

Foveated analysis: Patch at fixation *n* is foveated based on its eccentricity from fixation *n-1* and analyzed.



Eccentricity-based Analysis: A patch at fixation *n* is classified into one of 5 saccade bins based on its eccentricity from fixation *n-1*. The figure shows the distribution of saccade magnitudes and boundaries of the 5 saccade bins.

3. Results

3a. Statistics of Luminance Ratios

■ Raised cosine weighted (w_i) luminance of patch *p*

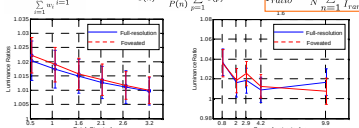
$$I(p) = \frac{1}{\sum_{i=1}^{M} w_i} \sum_{i=1}^{M} I_i w_i$$

■ Mean luminance of *P(n)* patches in image *n*

$$\bar{I}(n) = \frac{1}{P(n)} \sum_{p=1}^{P(n)} I(p)$$

■ Luminance ratio for *N* images

$$\bar{I}_{ratio} = \frac{1}{N} \sum_{n=1}^{N} \frac{\bar{I}_{pog}(n)}{\bar{I}_{rand}(n)}$$



Luminance ratios as a function of the patch size used to compute luminance

Luminance ratios as a function of saccade magnitude. Patch size = 1.6°

- • Luminance ratios are statistically greater than 1.0
- • Humans fixations, on average, land on brighter regions
- • Yet luminance is not a powerful attractor (max. ratio 1.04)

3b. Statistics of Contrast Ratios

■Weighted RMS contrast of a patch, *p*

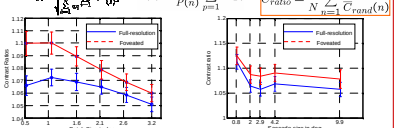
$$C(p) = \sqrt{\sum_{i=1}^{M} w_i \frac{(I_i - \bar{I}(p))^2}{\bar{I}(p)^2}}$$

■Average contrast for image *n*

$$\bar{C}(n) = \frac{1}{P(n)} \sum_{p=1}^{P(n)} C(p)$$

■ Contrast Ratio

$$\bar{C}_{ratio} = \frac{1}{N} \sum_{n=1}^{N} \frac{\bar{C}_{pog}(n)}{\bar{C}_{rand}(n)}$$

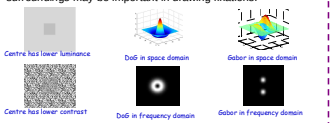


Contrast ratios as a function of the patch size used to compute contrast

Contrast ratios as a function of saccade magnitude. Patch size = 1.6°

•Contrast ratios are statistically greater than 1.0 and substantially higher than luminance ratios (max ratio. 1.11)
- • Foveated ratios are statistically greater than full-resolution ratios

Motivation: Variations of image features with respect to their surroundings may be important in drawing fixations.



Fixations are likely fall on the central patch despite lower luminance or contrast

Difference-of-Gaussians capture center-surround profiles

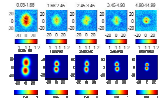
Gabors can be used to capture variations besides center-surround profiles

Problem: Determine the optimal size of the center-surround kernel

Solution: Analyze ratio of spatial frequency distributions of patches at human and random image fixations and model spatial frequencies that are significantly different.

$$F(i) = \frac{\frac{1}{P(i)} \sum_{p=1}^{p=P(i)} \text{abs}(FFT(p))_{PoG}}{\frac{1}{R(i)} \sum_{r=1}^{r=R(i)} \text{abs}(FFT(r))_{Rand}}$$

$$F_{ratio} = \frac{1}{N} \sum_{i=1}^{N} F(i)$$



Model significant spatial frequencies using Gabors

3c. Luminance Bandpass Ratios

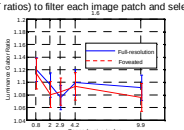
Goal : Compute the bandpass statistics of patch luminance

Approach : Use Gabors (best fits to FFT ratios) to filter each image patch and select maximum filtered output.

For patch *t(p)* at eccentricity *e* from previous fixation in image *i*, we compute:

$$G(i,e) = \frac{\frac{1}{P(i)} \sum_{p=1}^{p=P(i)} \max|Gab(e) * I(p)|_{PoG}}{\frac{1}{R(i)} \sum_{r=1}^{r=R(i)} \max|Gab(e) * I(r)|_{Rand}}$$

$$G_{ratio}(e) = \frac{1}{N} \sum_{i=1}^{N} G(i,e)$$



Luminance bandpass ratios as a function of saccade magnitude. Patch size = 1.6°

Bandpass-luminance profiles produce better separation between patches at human and random fixations than luminance ratios (max. ratio. 1.12).

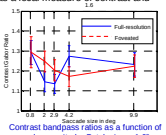
3d. Contrast Bandpass Ratios

Goal : Compute the bandpass properties of patch contrast

Approach: We use local image gradients as a local measure of contrast and analyze them as explained in 3c.

Original Image Patches

Gradient of Image Patches



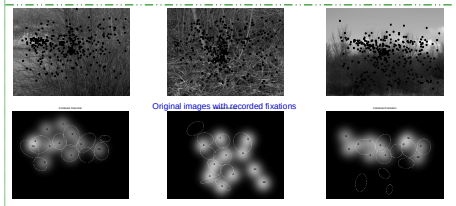
Contrast bandpass ratios as a function of saccade magnitude. Patch size = 1.6°

•Bandpass-contrast provides the maximum separation amongst the features examined (max. ratio of 1.29).
- •Thus, observers are very likely to fixate on regions with center-surround *contrast* profiles.

Laboratory for Image & Video Engineering: http://live.ece.utexas.edu

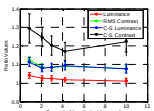
Center for Perceptual Systems: http://www.cps.utexas.edu

4. Fixation Selection Algorithm

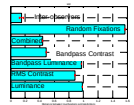


Comparing algorithmically selected fixations (bright regions) with clusters of human fixations (ellipses). Fixation selection maps were combined using linear weighting of the four features.

5. Conclusions



Ratios of image features (PoG / Rand) as a function of saccade magnitude. Patch size = 1.6°



Dissimilarity (Kullback-Leibler distance) between recorded fixations and selections

1. Image patches around human fixations have, on average, **higher values** of local patch luminance, contrast (RMS), bandpass outputs of patch luminance and contrast than patches selected randomly.
2. **Bandpass contrast** showed the greatest difference between human and random fixations, followed by bandpass luminance, contrast, and luminance.
3. A **foveated analysis of local contrast** resulted in even greater differences between human and random fixations than previously reported results that did not incorporate foveation.
4. Selecting image regions as likely candidates for fixation using these image features correlates well with fixations recorded from observers.

This research was supported by a grant from the National Science Foundation (ITR-0427372) and by ECS-0225451.