Does luminance-contrast contribute to a saliency map for human overt attention?

Wolfgang Einhäuser & Peter König
Institute of Neuroinformatics (UNI/ETH Zürich)

558

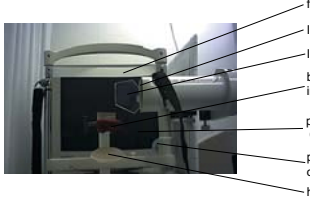
1 Abstract

In natural environments, humans select a subset of visual stimuli by directing their gaze to locations attended. In previous studies it has been found, that at fixation points luminance-contrast is higher than average. This lead to the hypothesis that luminance-contrast makes a major contribution to a saliency map of visual overt attention, consistent with a computation of stimulus saliency in early visual cortical areas. We re-evaluate this hypothesis by using natural and modified natural images to uncover the causal effects of luminance-contrast to human overt visual attention:

(1) We confirm that - viewing natural images - contrasts are elevated at fixation points. This, however, only holds for low spatial frequencies and in a limited temporal window after stimulus onset.
(2) However, despite this correlation between overt attention and luminance-contrast, moderate modifications of contrast in natural images do not measurably affect the selection of fixation points. Furthermore, strong local reductions of luminance-contrast do not repel but attract fixation.
(3) Neither contrast nor contrast modification is correlated to fixation duration.
(4) Even the moderate contrast modifications used fall into the physiologically relevant range, and subjects are well able to detect them in a forced choice paradigm.

In summary, **no causal contribution of luminance-contrast to a saliency map of human overt attention** is detectable. In conjunction with recent results on the relation of contrast sensitivity of neuronal activity to the level in the visual cortical hierarchy, the present study provides evidence, that saliency is computed not early but late during processing.

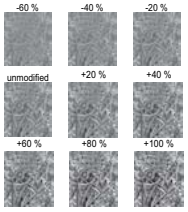
2 Setup



The eye-tracking setup consists of the actual oculometer (Dr. Bouis, Karlsruhe, Germany), the presentation screen and subject stabilization equipment. Stimuli are generated on a Macintosh G4/800. Eye-tracker output is recorded at **1 kHz** by a i80486DX4 PC. Presentation and recording are precisely synchronized by an optical signal.

3 Stimuli

One out of eight natural photographs used as stimuli for the experiments. Upon presentation the image spans 38.6 degrees of visual angle.



Enlargements of contrast modifications applied to stimuli. Five modifications of identical strength are applied to a stimulus and the influence of the modification on fixation probability is measured. Size of each square corresponds to 4.3 degrees of visual angle.

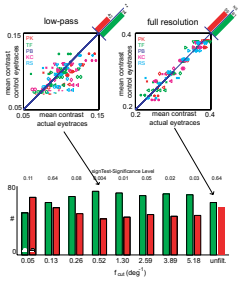
4 Results

The eye-tracking experiment consisted of 4 sessions. The first session was composed of three blocks, in each of which each of eight different natural images was presented once for 8s without any modification and in random order. In each of the sessions 2 to 4, each image was subjected to nine different peak contrast modification levels (see box 3). Using the same eight ' basis' images as in the first session, this yields a total of 72 stimuli per session.
For the unmodified natural images we analyze whether contrast correlates with fixation probability. For each unmodified image eye-traces obtained on the same image (' actual' condition - green traces) are compared to eye-traces obtained from different images but applied on the same image (' control' condition, red traces)

Eyetraces obtained from one subject on unmodified images. Green traces were obtained while viewing the shown image ('actual' condition), red traces on all other unmodified images ('control' condition)

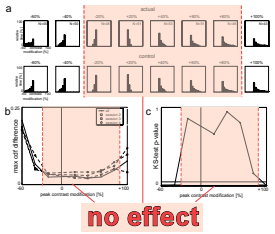
5 Results: Unmodified stimuli

Performing the analysis of box 4 for all images and all subjects yields a highly significant ($p<0.01$, sign-Test) difference between ' actual' and ' control' condition if contrast is defined on the low-pass (cut-off frequency 0.52cyc/deg) filtered image (top left panel) but no significant difference on the unfiltered images ($p>0.6$, sign-Test; top right panel). Analyzing this frequency dependence further, reveals that the difference between actual and control condition is maximal for cut-off frequencies around 0.5cyc/deg and remains significant over a wide range of cut-off frequencies. Nevertheless, very high as well as very low spatial frequency contrast has no significant relation to fixation behavior (bottom panel).



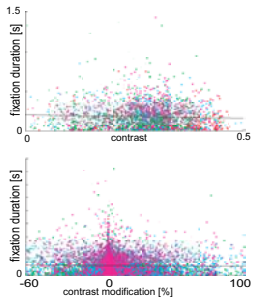
6 Results: Modified stimuli

As decisive step we address whether the observed correlation is a causal consequence of high contrast attracting attention. This is done by locally modifying luminance-contrast in the images. The distribution of contrast modification along the eye-traces are computed for each contrast modified stimulus and added up over subjects and basis images but separated according to peak contrast modification level (panel a, top row). For each peak modification level, this "actual" distribution is compared to the according "control" distribution, which represents the prediction if the modifications have no influence on fixation (panel a, bottom row).The similarity of actual and control distributions is assessed by a KS-test, using the cummulative density functions of both distributions. One finds, that only for unnaturally extreme modifications actual and control significantly differ (panel c). This is true for all sessions that include modified images (panel b), and thus cannot be attributed to some memory effect. Together with the fact that extreme contrast redcution attracts fixation rather than repelling it, this shows that high contrast by itself does not causally attract attention.



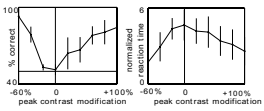
7 Fixation duration

The analysis done so far implicitly weighs fixation-probability with fixation duration. Thus we set out to control for correlation of the latter to contrast and to contrast modifications. Over all unmodified images and all subjects, we find no correlation between contrast and fixation duration ($|r|^2$ 0.051 for all cut off frequencies, top panel shows unfiltered images). This holds also true for individual subjects and individual images.
Neither do we find any correlation of contrast modification to fixation duration for the modified stimuli ($r = -0.016$; bottom panel). These controls show that neither contrast nor contrast modification is related to the duration of fixation and an implicit weighing with fixation duration thus cannot confound the analysis.



8 Detection Paradigm

In an additional 2-alternative forced choice paradigm, it is tested whether the modifications can be detected by subjects explicitly instructed to search for image abnormalities. Therefore the stimuli are split at the midline and while one half is subjected to the contrast modification as in the eye-tracking experiment, the other remains unmodified. Even at moderate contrast modification levels, which do not attract attention, subjects are able to identify the modified half. This shows that the contrast modifications are visible on the system level, even though they do not attract attention.



9 Discussion

In this study we show that - despite the observed correlation between contrast and fixation probability - high contrast by itself does not causally attract overt attention in natural scenes. As revealed by the detection paradigm, this is true, although the contrast information is available on the system level. In natural scenes luminance-contrast thus does not measurably contribute to the saliency map for human overt visual attention.
Neuronal responses in the early visual system are strongly sensitive to luminance-contrast as tuning curves to most low-level features scale with luminance-contrast. While proceeding through the visual hierarchy, on the other hand, responses become more and more invariant to contrast. Our finding that saliency does not causally depend on contrast thus yields the conclusion that - for natural scenes - the saliency map for overt visual attention does not originate in early cortical visual areas.

Ref:Einhäuser & König, Eur. J. Neurosci. 2003

contact: weinhaeu@ini.phys.ethz.ch http://www.ini.unizh.ch/~weinhaeu

This work was supported by Honda RI Europe (WE) and the Swiss National Science Foundation (PK, grant 31-61415.01).

Contact
Wolfgang Einhäuser
Institute of Neuroinformatics
University & ETH Zürich
Winterthurer Str. 190
CH-8057 Zürich
Switzerland
weinhaeu@ini.phys.ethz.ch
http://www.ini.unizh.ch/~weinhaeu