Eye movements on natural videos: Low-level feature correlations at successive fixations

Michael Dorr, Karl R. Gegenfurtner*, and Erhardt Barth
Institute for Neuro- and Bioinformatics, University of Lübeck, http://www.inb.uni-luebeck.de
*Abteilung für Allgemeine Psychologie, Justus-Liebig-Universität Gießen, http://www.allpsych.uni-giessen.de
{dorr|barth}@inb.uni-luebeck.de, Karl.R.Gegenfurtner@psychol.uni-giessen.de

Introduction

We investigated the relationship between low-level image features at successive fixations during free viewing of natural dynamic scenes. If there were a contribution of low-level features at the current centre of fixation to the selection of the next saccade target, we should find evidence in the differences of those features, going from one fixation to the next.

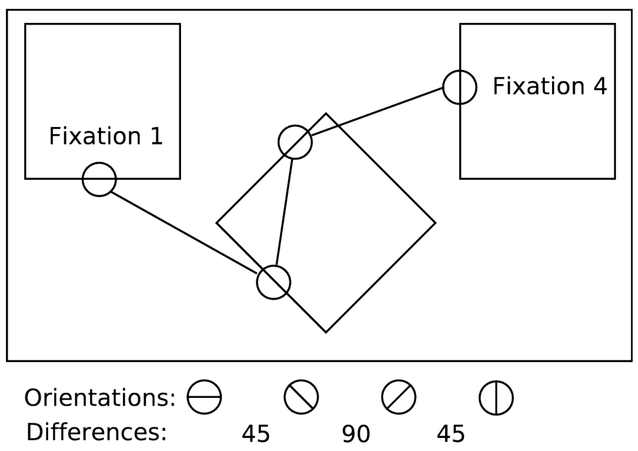


Fig. 1 Example sequence of low-level features along the scanpath (for a synthetic scene): the orientation at the first fixation (90°) differs from that at the second fixation (135°) by 45°, difference second to third (225°) is 90°, etc.

References
[1] M. Böhme, M. Dorr, C. Krause, T. Martinetz, E. Barth, Eye movement predictions on natural videos. *Neurocomputing*, 69(16-18):1996-2004, 2006.
[2] E. Barth, The minors of the structure tensor. In: Mustererkennung 2000, G. Sommer, N. Krüger, Ch. Perwass (eds.), Springer:221-228, 2000.

Methods

Data collection
We recorded eye movements from 54 subjects watching 18 high-resolution video clips of outdoor scenes (20 s each) using an EyeLink II eye tracker running at 250 Hz. Saccades and fixations were extracted by a velocity-based algorithm. Overall, 37000 fixations were detected.

Feature extraction
Colour, local orientation, and motion were computed on a Gaussian pyramid with 6 spatial and 3 temporal levels. Levels corresponded to a spatial resolution of 0.4, 0.8, 1.7, 3.4, 6.8, 13.6 cycles/degree and a temporal resolution of 7.5, 15, 30 fps, respectively. We also computed geometrical invariants that have been used to predict eye movements before [1].

Baseline scanpaths
As a baseline measure, we generated artificial scanpaths with varying degree of similarity to natural scanpaths.
Random: scanpaths were uniformly sampled.
Same lengths: fixation patches were drawn randomly, but their distances had the same distribution as saccadic amplitudes in natural scanpaths.
Scrambled: the fixation patches of a natural scanpath were brought in different order (so that saccadic amplitudes changed).
Synthetic: both the distribution of fixated patches and saccadic amplitudes were approximated.

Colour
The luminance (Y) and the colour opponency channels U (red/green) and V (blue/yellow) were extracted.

Local orientation
2D orientation was computed by eigenvalue analysis of the structure tensor *J*. Fixation patches with no dominant orientation (i.e. homogeneous areas) or with only very low contrast were excluded from further analysis.

Motion
Velocity was estimated based on the minors of the structure tensor, following the algorithm in [2].

Geometric invariants H, S, K
The geometric invariants *H*, *S*, and *K* were computed from the minors of the structure tensor. They indicate how many degrees of freedom are locally used by the image sequence.

$$J = \omega * \begin{pmatrix} f_x f_x & f_x f_y & f_x f_t \\ f_x f_y & f_y f_y & f_y f_t \\ f_x f_t & f_y f_t & f_t f_t \end{pmatrix} \qquad \begin{array}{lll} H & = 1/3\ \mathrm{trace}(J) & = \lambda_1 + \lambda_2 + \lambda_3 \\ S & = |M_{11}| + |M_{22}| + |M_{33}| & = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3 \\ K & = |J| & = \lambda_1\lambda_2\lambda_3 \end{array}$$

Results
Colour: random scanpaths show higher luminance differences than original scanpaths (fig. 2 bottom) because regions with high luminance (e.g. the sky) were fixated less often in original sps (fig. 2 top).
Orientation: differences along original scanpaths differ from those along random ones, but there is almost no difference to scanpaths with the same saccadic amplitudes (fig. 3 bottom).
Velocity/geometric invariants: random scanpaths show smaller differences (fig. 4/5 bottom) because original scanpaths have a bias towards regions with high velocity/structure (fig. 4/5 top).

Results

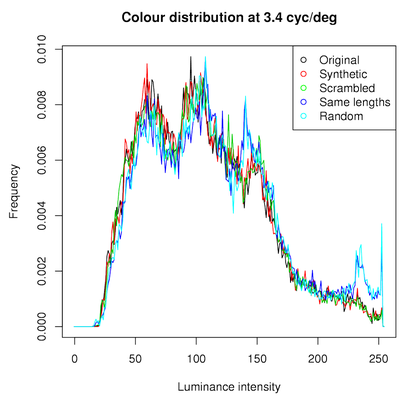


Fig. 2

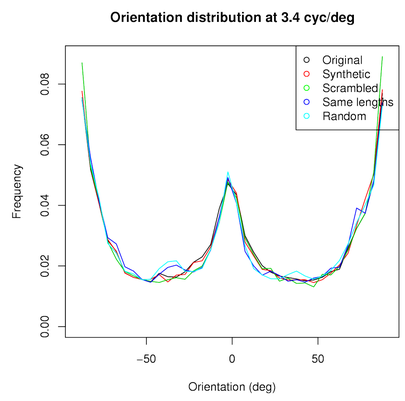


Fig. 3

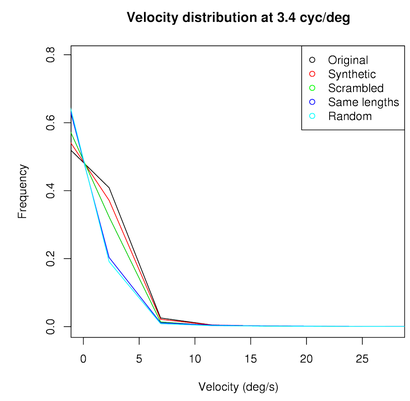


Fig. 4

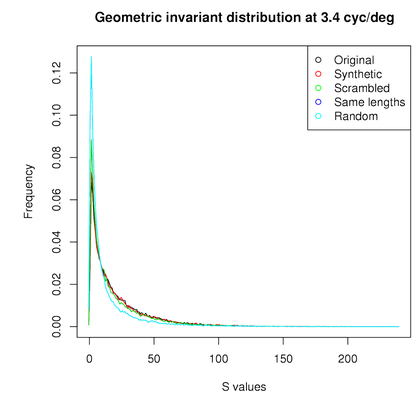
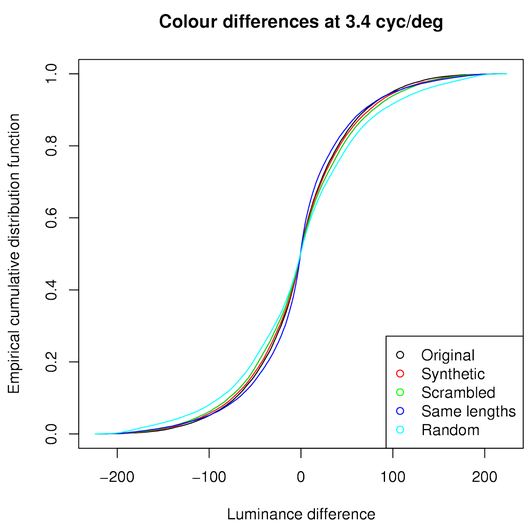
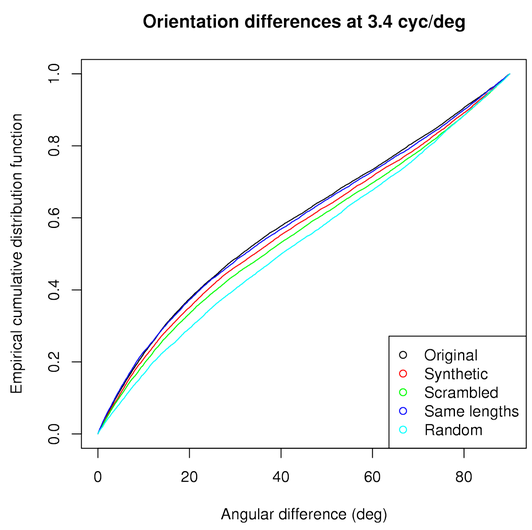
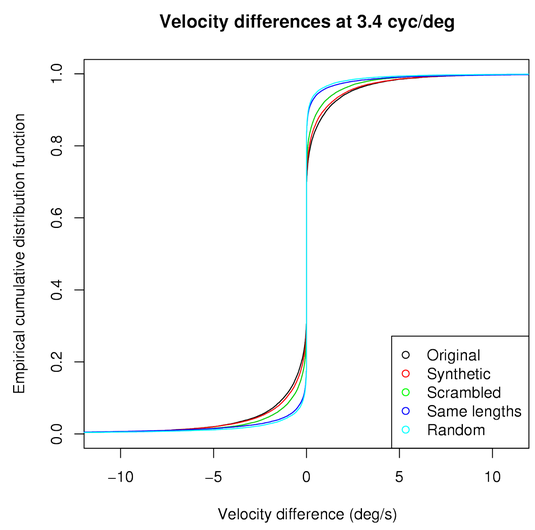
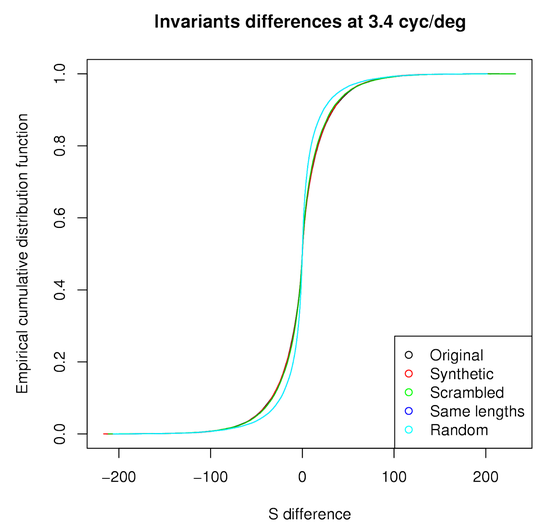


Fig. 5









Summary

- Local orientations are correlated along the scanpath, as are colour and motion
- These correlations can be explained by spatio-temporal correlations in natural scenes and a general bias in target selection
- We can conclude that low-level features at the current centre of fixation contribute little to the selection of the next saccade target

Our research has received funding from the European Commission within the GazeCom project (IST-C-033816) of the FP6.
All views herein are those of the authors alone; the European Community is not liable for any use made of the information.