Learning an Interest Operator from Eye Movements

Wolf Kienzle, Matthias O. Franz, Felix A. Wichmann,
and Bernhard Schölkopf

Abstract

In the computer vision community, so-called *interest operators* have become very popular in recent years. An interest operator is a function that, given an image, returns a set of *interesting* points that can then be used to e.g. categorize images or to detect objects within images. To our knowledge, all of the widely used interest operators are based on *heuristic* measures of local interestingness (e.g. contrast, second order structure, entropy, etc.). While this works well in numerous image processing applications, it seems natural to ask how (if at all) they relate to the visual system's notion of interestingness. This has been explored by e.g. Privitera and Stark [1], who found significant correlations between fixation points and locations selected by existing heuristic interest operators. Reinagel and Zador [2] took the opposite approach and reported that patches around fixation points tend to have higher contrast and that point-wise correlations decay with eccentricity.

In this work, we want to find regularities in the image structure around fixation points. Thus, we follow the spirit of [2], but instead of verifying *heuristic* hypotheses such as "higher contrast" or "correlations decay with eccentricity", our goal is to *infer* hypotheses *from the data*, using methods from statistical learning. To this end, we train a non-linear classifier on fixated vs. randomly selected image patches, i.e. we learn an interest operator from human eye movements.

At present, we are exploring various ways to deal with two major problems inherent in our approach. First, when a fixation point is recorded a some particular position, the size of the attracting image region around that point is unknown. Second, fixation positions are subject to "inaccurate" gaze positioning and measurement noise, which is substantial compared to pixel sizes. At the workshop, we will present first experimental results showing that statistical methods are capable of learning discriminative image features around fixation points.

References

[1] C. M. Privitera and L. W. Stark. Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(9):970 – 982, 2000.

[2] P. Reinagel and A. M. Zador. Natural scene statistics at the center of gaze. *Network*, 10:1–10, 1999.