

# Modelling collective viewing behaviour in natural video based on eye-tracking measurements

Marcus Nyström<sup>1</sup>, Kenneth Holmqvist<sup>2</sup>, Mirek Novak<sup>1</sup>

<sup>1</sup>Department of Information Technology, {marcus,mirek}@it.lth.se

<sup>2</sup>Cognitive Sciences {kenneth@lucs.lu.se}

Lund University, Sweden

This paper presents a method that finds common regions of multiple viewers' interest based on eye-movements collected from natural video. We intend in the future to use the knowledge of where these common regions are located to compress natural video in variable resolution, where regions with high gaze density are prioritized.

Recorded fixations and points of gaze can successfully be represented and analyzed by using Gaussian functions as devised by Velichkovsky et al. (1996) and Wooding (2002) under the names “attentional landscapes” and “fixation maps”, respectively. In essence, these representations are created by centring a 2-D Gaussian function at each point of gaze such that overlapping functions are superimposed. We refer to these Gaussian representations as “gaze density functions” (gdfs), and use them to model the likelihood that future viewers of the same video will direct their gaze towards a certain frame region.

There are problems inherent in using a gdf to mask visual displays such that visually important display regions remain in high resolution after masking. These problems include 1) Uniformly distributed gaze points may generate gdfs with several peaks in gaze density of the same magnitude. Consequently many high-resolution regions will appear after masking, which is confusing for future viewers. 2) Local maxima in the gdf represent visually important regions poorly.

To address these built-in problems that gdfs exhibit, we generalize gdfs into Quality Functions (QF). The gdf is first used to calculate how scattered the points of gaze are, which then controls the size of the high-resolution regions of the QF. Spread out points of gaze leave much uncertainty of where future viewers will direct their gaze and therefore leads to an increase in size of the high-resolution. The QF solves the second problem by an increasingly refined search for visually important frame regions, allowing regions with less gaze density than the global maxima to remain in high resolution after masking. Gaze points corresponding to large magnitudes in the gdf are prioritized in this search. The search is terminated when all gaze points needed to represent the main locus of viewers, as defined by a stop criterion, are found. The stop criterion takes into account the *collective* behaviour of future viewers, known limitations in eye-tracking equipment and calibration accuracy, and also previous research on individual variations in eye-movement behaviour.