

Automated Filtering of Eye Movements Using Dynamic AOI in Multiple Granularity Levels

Gavindya Jayawardena, Old Dominion University, USA
Sampath Jayarathna, Old Dominion University, USA

ABSTRACT

Eye-tracking experiments involve areas of interest (AOIs) for the analysis of eye gaze data. While there are tools to delineate AOIs to extract eye movement data, they may require users to manually draw boundaries of AOIs on eye tracking stimuli or use markers to define AOIs. This paper introduces two novel techniques to dynamically filter eye movement data from AOIs for the analysis of eye metrics from multiple levels of granularity. The authors incorporate pre-trained object detectors and object instance segmentation models for offline detection of dynamic AOIs in video streams. This research presents the implementation and evaluation of object detectors and object instance segmentation models to find the best model to be integrated in a real-time eye movement analysis pipeline. The authors filter gaze data that falls within the polygonal boundaries of detected dynamic AOIs and apply object detector to find bounding-boxes in a public dataset. The results indicate that the dynamic AOIs generated by object detectors capture 60% of eye movements & object instance segmentation models capture 30% of eye movements.

KEYWORDS

Area-of-Interest, Eye Tracking, Filtering

1. INTRODUCTION

Eye tracking can reveal objective and quantifiable information about the quality, predictability, and consistency of the underlying covert process of the human brain when carrying out cognitively demanding tasks (McCarley and Kramer 2008; Radach et al. 2003; Van der Stigchel et al. 2007). According to the eye-mind hypothesis (Just and Carpenter 1980), observers attend where their eyes are fixating. Thus, eye tracking measurements enable us to investigate cognitive behavior when visually exploring a stimulus. With the advancement of eye tracking technology, gaze tracking measurements have become reliable and accurate.

Eye gaze measurements include various metrics relevant to oculomotor control (Komogortsev et al. 2013) such as saccadic trajectories, fixations, and other relevant measures including velocity, duration, amplitude, pupil dilation (Krejtz et al. 2018). Studies have shown that the size of the pupil diameter correlates with the task complexity (Kosch et al. 2018) enabling the use of pupillary behavior as biomarkers of mental workload when completing a task. Several studies (Gehrer et al. 2018; Jayawardena et al. 2020) have incorporated eye tracking to obtain insights into underlying covert processes. As a standard practice in the community, upon successful completion of the study, performance of users is measured, traditional positional gaze metrics and advanced gaze metrics are

DOI: 10.4018/IJMDEM.2021010104

calculated, and statistical significance of computed numerous metrics are evaluated (Gehrer et al. 2018; Jayawardena et al. 2020).

Eye tracking experiments utilize area of interests (AOIs) to the aid the analysis process by extracting eye gaze metrics within a predefined AOIs. An AOI is a region of stimuli that is used to study the eye gaze metrics and link eye movement measures to the part of the area of the stimuli (Hessels et al. 2016). Studies in visual attention and eye movements (Noton and Stark 1971; Privitera and Stark 2000) have shown that humans only attend to a few AOIs in a given stimulus. Analysis of eye gaze metrics within AOIs can provide important cumulative clues to the underlying physiological functions supporting the allocation of visual attention resources. For instance, in the context of user interface interaction, the number of fixations within an AOI (e.g., a user interface component) indicates the efficiency of finding that component among others, whereas the maximum and average fixation duration within that AOI indicates the informativeness of that component (Goldberg and Kotval 1999). In addition, the fixation frequency and blink frequency within AOIs can indicate cognitive workload when interacting with the particular component of the user interface (Van Orden et al. 2001).

The analysis of eye movements in dynamic AOIs, such as in video sequences is not new (Marchant et al. 2009; Goldstein et al. 2007; Crossland et al. 2002; Timberlake et al. 2005). Studies with the primary focus on detecting fixation sequences within identified AOI have used clustering techniques to group gaze locations to determine the AOI (Nguyen et al. 2004), and various image processing algorithms (Privitera and Stark 2000) to automatically identify the AOI. For the analysis of cognitive workload and allocation of visual attention resources, many studies (Santella et al. 2006; Shanmuga Vadivel et al. 2015; Khosravan et al. 2016; Wang et al. 2019) have utilized saliency models of viewers in conjunction with visual information from video frames. In addition, computer vision techniques (Weibel et al. 2012) have been applied to generate AOI-mapped gaze coordinates by using a template of the desired object derived from a single frame of the eye tracking stimuli video. But it only works for pre-recorded eye tracking stimuli using the manual specification of the AOI template generated beforehand.

To overcome these challenges, we propose a computer vision-based deep neural network approach to identify the AOIs in video streams in real-time to filter gaze locations that fall into the identified AOIs for the analysis of both positional and advanced eye gaze metrics. We focus on two filtering techniques to dynamically generate AOI-mapped gaze locations on video streams:

- Pre-trained object detectors to identify bounding boxes of dynamic AOI in video sequences, and;
- Object instance segmentation models for offline detection of dynamic AOI via precise pixel-wise masks.

Since computer vision techniques can be adapted to detect a wide range of objects, we apply computer vision techniques to extract dynamic AOI-mapped eye movement data. We use transfer learning to remodel existing image classifiers for dynamic AOI detection. Upon detection of dynamic AOIs, we extract eye movement data which falls within the detected dynamic AOIs by checking if the gaze coordinate falls within any dynamic AOIs' polygonal boundaries. Our approach represents two levels of granularity in AOIs, (1) a lower level of granularity (i.e. bounding box), and (2) a higher level of granularity (i.e. pixel-wise mask). We evaluate our filtering methodology in terms of the percentage of the eye movements captured from each filtering technique proposed. We utilize our prior work on Real-Time Advanced Eye Movements Analysis Pipeline (RAEMAP) (Jayawardena 2020) designed to analyze traditional positional gaze measurements as well as advanced eye gaze measurements.

In contrast to the existing studies on analyzing eye movements within static AOIs (often pre-determined), which only works for pre-recorded eye tracking stimuli, our proposed methodology could extract eye movement data from dynamically detected AOIs in real time. In this study, we present pre-trained object detectors and object instance segmentation models for the detection of dynamic AOIs as the initial step towards extracting eye movement data from dynamically detected AOIs in

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/automated-filtering-of-eye-movements-using-dynamic-aoi-in-multiple-granularity-levels/271433

Related Content

Construct a Bipartite Signed Network in YouTube

Tianyuan Yu, Liang Bai, Jinlin Guo and Zheng Yang (2015). *International Journal of Multimedia Data Engineering and Management* (pp. 56-77).

www.irma-international.org/article/construct-a-bipartite-signed-network-in-youtube/135517

Emoticon Recommendation System to Richen Your Online Communication

Yuki Urabe, Rafal Rzepka and Kenji Araki (2014). *International Journal of Multimedia Data Engineering and Management* (pp. 14-33).

www.irma-international.org/article/emoticon-recommendation-system-to-richen-your-online-communication/109076

The Transforming Media Landscape

(2019). *Cross-Media Authentication and Verification: Emerging Research and Opportunities* (pp. 39-61).

www.irma-international.org/chapter/the-transforming-media-landscape/208000

Semantically Driven Multimedia Querying and Presentation

I. Cruz and Olga Sayenko (2008). *Multimedia Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 848-863).

www.irma-international.org/chapter/semantically-driven-multimedia-querying-presentation/27125

Video Face Tracking and Recognition with Skin Region Extraction and Deformable Template Matching

Simon Clippingdale and Mahito Fujii (2012). *International Journal of Multimedia Data Engineering and Management* (pp. 36-48).

www.irma-international.org/article/video-face-tracking-recognition-skin/64630