**Gaze Duration Error Estimation from Fixation Coordinate Prediction Models**

**Introduction**

Eye tracking applications are an important application of computer vision that permit the development of a wide variety of new forms of software interaction, as well as a tool for empirical investigations into the way people consume media. An eye tracking application will be trained on data sets containing combinations of facial images and fixation coordinates collected through careful application design. The resulting fixation coordinate predication model can then be applied sequentially to an interactive session by a user to determine how long their eyes were fixated on a given object: the gaze duration.

Fixation coordinate models will typically have an error profile that is dependent on both the nature of the media being presented, the variety of faces and lighting conditions in the training data. In addition, it has been observed that the error profile varies depending on the location of the true fixation point.

When fixation models are used sequentially for the purpose of estimating gaze duration (as is done in advertising media studies), then there is the potential for the error to either cancel out (reducing error in gaze duration estimation) or to compound (increasing error in gaze duration). Which of these outcomes occurs will depend on the scenario being investigated. Influencing factors include the ratio of true positives to true negatives in the data, and the error profile of the machine learning fixation coordinate model. Heretofore, gaze duration models have assumed the former situation, if indeed the possibility of error in gaze duration is even considered.

**[Need to check that this holds in the literature]**

In this work we develop a methodology for estimating session specific error profiles that relies on an empirical error profile of the underlying fixation model. The method allows us to profile whether an individual session has a compounding or cancelling error profile, and put empirical confidence bounds on the gaze duration estimates.

**Methodology**

We produce an error profile report using a hold out set of data from the fixation model training process. This error profile can be stratified into a grid across the screen dimensions. For each of these grid locations we calculate two error probabilities, the false negative probability (when true fixation is on the grid and the model produces coordinates outside the grid. And the false positive probability, when the true fixation is outside the grid location and the model produces a prediction inside the grid square.

We apply this grid of error probabilities in a Monte Carlo simulation that requires the following inputs:

1. The grid coordinates in which the item of concern is located.
2. The duration of true fixation on that item
3. The duration of the entire session of measurement.

The Monte Carlo process will then repeatedly simulate sessions of the given length and gaze duration. Each simulation will use the error profiles to produce a potential model output. Collectively these estimates model outputs constitute a probability distribution over outputs for the given scenario.

Study One

At this first stage we produce results analyzing what this distribution of measurements looks like for a given true duration. In particular we look at the screen locations for which we see either compounding or cancelling behavior in the aggregate formation of the gaze duration measurement.

Study Two

In real world applications we will have the true session length but rely on the model for the estimate of gaze duration. Meaning that the probability distribution we want is the distribution over true gaze durations for a given measurement, rather than the inverse which is what we have produced so far.

In order to invert the probability distribution, we start with the models’ predicted duration output as the estimate of true duration, we then repeatedly simulate in both increasing and decreasing directions the Monte Carlo process until the probability of the observed measurement falls below a given threshold. Crucially each of these simulations contains and equal number of samples and we make the additional assumption that the prior likelihood of all simulated true durations is equal.

These constraints allow us to use the collected samples to produce a probability distribution over true durations for a given observation. Each observation in the set of simulations was generated from an assumed true duration. We allocate that true duration a probability of occurrence proportional to the total number of observations of the predicted gaze duration.

<https://www.researchgate.net/publication/10919124_Cleaning_up_systematic_error_in_eye-tracking_data_by_using_required_fixation_locations>

<https://cognitiveresearchjournal.springeropen.com/articles/10.1186/s41235-019-0159-2>

<https://dl.acm.org/doi/abs/10.1145/968363.968391?casa_token=N_rmI3EBD1cAAAAA%3Ad2ZM6Yz9StJoJ_tUSmrGggZiQgrR53ZBro7GEueRM3mtPgIj4BLh_1rt6Xas5Op2IIrTzQ-gATju-Ow>

<https://www.researchgate.net/publication/254007815_Eye_tracker_data_quality_What_it_is_and_how_to_measure_it>

https://www.sciencedirect.com/book/9780124081383/eye-tracking-in-user-experience-design