Impact of Gaze Uncertainty on AOIs in Information Visualisations

Yao Wang
University of Stuttgart
Stuttgart, Germany
yao.wang@vis.uni-stuttgart.de

Maurice Koch∗
University of Stuttgart
Stuttgart, Germany
maurice.koch@vis.uni-stuttgart.de

Mihai Bâce
University of Stuttgart
Stuttgart, Germany
mihai.bace@vis.uni-stuttgart.de

Daniel Weiskopf
University of Stuttgart
Stuttgart, Germany
daniel.weiskopf@visus.uni-stuttgart.de

Andreas Bulling
University of Stuttgart
Stuttgart, Germany
andreas.bulling@vis.uni-stuttgart.de

Figure 1: Areas of Interest (AOIs) are used extensively in eye tracking studies, however, it is currently unknown how gaze uncertainty impacts the resulting findings. To overcome this limitation, we study the uncertainty caused by the gaze estimation error of the eye tracker and amplified by the nearby AOIs in information visualisations. We propose two effective metrics, the Flipping Candidate Rate (FCR) and Hit Any AOI Rate (HAAR), to quantify the impact of uncertainty on the sample application domain of information visualisations.

ABSTRACT
Gaze-based analysis of different areas of interest (AOIs) is widely used in information visualisation research to understand how people explore visualisations or assess the quality of visualisations concerning key characteristics such as memorability. However, nearby AOIs in visualisations amplify the uncertainty caused by the gaze estimation error, which strongly influences the mapping between gaze samples or fixations and different AOIs. We contribute a novel investigation into gaze uncertainty and quantify its impact on AOI-based analysis on visualisations using two novel metrics: the Flipping Candidate Rate (FCR) and Hit Any AOI Rate (HAAR). Our analysis of 40 real-world visualisations, including human gaze and AOI annotations, shows that gaze uncertainty frequently and significantly impacts the analysis conducted in AOI-based studies. Moreover, we analysed four visualisation types and found that bar and scatter plots are usually designed in a way that causes more uncertainty than line and pie plots in gaze-based analysis.

CCS CONCEPTS
• Human-centered computing → Information visualization; Laboratory experiments.

KEYWORDS
gaze uncertainty, gaze estimation error, eye tracking study, information visualisations

ACM Reference Format:

1 INTRODUCTION
Gaze-based analysis of different areas of interest (AOIs) is widely used in information visualisation research, e.g. to understand how people explore visualisations [Polatsek et al. 2018] or assess the
quality of information visualisations with respect to memorability [Borkin et al. 2015] or graphic effectiveness [Goldberg and Helfman 2010]. On such visual stimuli, an AOI covers areas with semantic meaning such as the axes, the title, or different graphical elements. However, a key assumption in gaze-based AOI analysis is the correct assignment of gaze samples and fixations to the corresponding AOIs [Goldberg and Helfman 2010]. Despite recent advancements in eye tracking hardware and methods, there is still an inherent gaze estimation error [Ehinger et al. 2019] that causes uncertainty in the on-screen gaze locations. In addition, visualisation characteristics such as the distance between AOIs [Yun et al. 2013] and their sizes [Orquin et al. 2016] further magnify the impact of gaze estimation error on AOI analysis. Several studies have reported this impact on gaze-based applications, e.g. it hampers the conclusion of behavioural decision-making studies [Orquin et al. 2016] and the usability of gaze-based user interfaces [Barz et al. 2016]. To quantify this impact, Orquin and Holmqvist [2018] proposed the capture rate, a statistical metric to quantify the percentage of fixations assigned to a given AOI. The capture rate is dependent on the size of the AOI and the accuracy of the eye tracker. Still, it cannot provide insights into fixations that may belong to multiple AOIs, and is suitable for simpler user interfaces in perceptual studies. Inspired by previous work on gaze-based AOI analysis, this paper further assesses the impact of gaze uncertainty on the assignment of fixations to AOIs in information visualisations, especially for different visualisation types, such as line, bar, scatter, and pie plots. In perception studies [Orquin et al. 2016; Orquin and Holmqvist 2018], researchers can avoid the effect of gaze uncertainty by increasing the AOI sizes and distances between AOIs. However, AOIs in information visualisations are usually small and close to one another. For such stimuli, gaze uncertainty becomes crucial in analysing fixations that land at the borders of possibly multiple AOIs. Moreover, another consequence of gaze uncertainty caused by gaze estimation error could lead to fixations areas with no semantic meaning (e.g. white spaces) and prior research showed that such areas do not attract human attention [Matzen et al. 2017].

In this paper, to quantify gaze uncertainty and its impact on AOI-based analysis for information visualisations, we introduce two novel metrics: the Flipping Candidate Rate (FCR) and the Hit Any AOI Rate (HAAR). FCR quantifies the probability that fixations might flip between two or more AOIs. HAAR quantifies the percentage of fixations that land on at least one AOI, hence capturing the impact of gaze uncertainty when assigning fixations to AOIs. To further understand and assess the impact of gaze uncertainty when assigning fixations to AOIs, we artificially flipped assigned AOIs and compared the resulting scanpaths. We calculated the Sequence Score [Yang et al. 2020], which is a pairwise string similarity metric, between the original and altered scanpaths. A value lower than 1 for the sequence score means that the altered scanpaths are different from the original one due to different assignments of fixations to AOIs, which could lead to potentially very different conclusions based on the eye tracking study. The results of our analysis suggest that gaze uncertainty has a substantial influence on AOI-based evaluations, and bar and scatter plots are most commonly designed in a way that causes more uncertainty than line and pie plots in gaze-based analysis. The contribution of our work is two-fold: First, we analyse and demonstrate the impact of gaze uncertainty on the assignment of fixations to AOIs for 40 real-world information visualisations [Borkin et al. 2013]. Second, we propose two novel metrics, the FCR and HAAR, that quantify the impact of gaze uncertainty on AOIs.

2 RELATED WORK

Our work is grounded on 1) gaze-based area of interest (AOI) evaluations of information visualisations and 2) gaze estimation error, and it is positioned in the area of 3) uncertainty representation and visualisation. Below, we briefly survey the literature in these fields.

**Gaze-based Area of Interest Evaluation.** Gaze-based AOI evaluation is widely used for various kinds of visualisations, including web pages [Drusch et al. 2014], static visualisations [Borkin et al. 2015], and metro maps [Xie et al. 2021]. In general, it plays an important role in connecting eye tracking and visualisation research [Burch et al. 2017]. An overview of AOI-oriented data visualization is provided by Blascheck et al. [2017]. The impact of AOI sizes on decision-making studies was examined by Orquin et al. [2016], and capture rate was proposed to quantify the uncertainty about the amount of fixations to a given object [Orquin and Holmqvist 2018].

**Gaze Estimation Error.** The intrinsic gaze estimation error from eye-tracking devices is well studied [Ehinger et al. 2019; Orquin et al. 2016]. Zhang and Hornof [2014] focused on post-hoc error compensation, whereas Sattar et al. [2017] added “jitter” to gaze samples to simulate different eye trackers. For user interfaces, the implications of the accuracy of eye tracking for design was studied by Feit et al. [2017]. Error-aware gaze-based interfaces allow us to compensate the gaze estimation error [Barz et al. 2016, 2018]. Furthermore, there are filtering and visual-interactive cleansing methods to address data quality problems in eye tracking [Schulz et al. 2015].

**Uncertainty Representation and Visualisation.** Dealing with uncertainty for gaze and corresponding AOI assignments, we can relate our work to the general problem of representing and processing uncertainty. Skeels et al. [2010] provide a general discussion of different notions of uncertainty and their role in data visualisation. According to their terminology, gaze estimation error can be understood as uncertainty on level 1 (for measurement precision and similar); however, the derived AOI assignment and follow-up analysis can be seen as uncertainty on level 3 — the inference level. Uncertainty visualisation can then use uncertainty models to represent and analyse uncertainty. See, for example, the seminal work by Pang et al. [1997] and a most recent survey on the topic [Weiskopf 2022].

Previous literature either directly used fixations and saccades for further analysis, or mitigated the impact of gaze uncertainty by changing the stimuli. Given AOIs in information visualisations are close and small, a slight change of gaze position might substantially affect the mapping to AOIs. Thus, a fundamental investigation into the impact of gaze estimation error in information visualisations is necessary.
3 GAZE UNCERTAINTY ASSESSMENT

The impact of gaze uncertainty on AOI assignment is assessed from two perspectives: 1) the number of fixation candidates that might flip between AOs (FCR) and 2) the number of fixations that hit any AOI (HAAR). Both a high FCR and a low HAAR suggest a high level of gaze uncertainty. We conducted our assessment on VisRecall [Wang et al. 2021], a dataset consisting of hundreds of real-world visualizations from MASSVIS [Borkin et al. 2015, 2013], along with human gaze and AOI annotations. The AOs are annotated following the taxonomy by Borkin et al. [2015]. Gaze data from VisRecall were collected under a recallability task, which includes an encoding and a recalling phase. An EyeLink 1000 Plus eye tracker at 2 kHz was mounted on a 24.5’ monitor with a resolution of 1920 x 1080 px. Visualisations were scaled to fit around 21.1 x 14.8’ of visual angle in the centre. For our assessment, only the gaze data from the encoding phase are used. We randomly selected 10 visualisations from each visualisation type: bar, pie, line, and scatter plots.

3.1 AOI Flipping Candidates

As the first step, we determine the spatial uncertainty associated with fixations. To this end, we take the raw gaze samples within each fixation segment (i.e., the time span associated with a fixation), and apply kernel density estimation (KDE) to arrive at the gaze density distribution [Rayner 1998]. The extent of the estimation error for the input gaze points is controlled by the bandwidth $h$ of a Gaussian kernel $K_h$. We compute the overlaid gaze density by summing up the contributions from the gaze samples, resulting in the probability density for this fixation.

Our goal is the probability $p_i$ that describes the probability of assigning a fixation to the $i$th AOI. Therefore, in the second step, $p_i$ is obtained by integrating (i.e., summing over an area) densities over all points that are covered by the $i$th AOI, that is,

$$p_i = \int_{x \in \Omega} \mathbb{1}_{A_i}(x) \left( \frac{1}{n} \sum_{j=1}^{n} K_h(x - x_j) \right) d^2x$$

Here $x_1, \ldots, x_n$ are the gaze samples associated with the fixation, and $\mathbb{1}_{A_i}(x)$ indicates if the point $x$ in image space $\Omega$ is covered by the $i$th AOI (i.e., $\mathbb{1}_{A_i}(x)$ is the characteristic function for AOI $i$). Figure 2 illustrates the computation of AOI probabilities.

Figure 2: Overview of the basic steps to decide whether a fixation is considered a flipping candidate.

Fixations whose densities only overlay one AOI will have a probability distribution peaked at this AOI, whereas fixations that overlay two or more AOs to a similar extent will result in distributions that are closer to being uniform over the respective AOs. We consider the latter type of fixations as flipping candidates. Depending on the number of AOs that are overlaid by the fixation, we differentiate between different ranks of flipping candidates. For example, Figure 3 illustrates fixations of rank 2 and 3. Given a number of AOs, the flipping candidate score $s_k$ of rank $k$ is defined as follows:

$$s_k = \sum_{i=1}^{N} p_i - \left( \sum_{i=1}^{k} p_i - \frac{1}{k} \right)$$

Note that we assume the probabilities $p_i$ to be sorted in descending order, that is, $p_1 \geq p_{i+1}$. The first term of the score equation penalises fixations whose spread covers mostly white space and only few AOs. The second term captures the degree of uncertainty and is defined as the statistical distance between the probabilities $p_i$ and the discrete uniform distribution of length $k$.

We define a flipping candidate of rank $k$ if its score $s_k$ exceeds a predefined threshold $t$, and it does not receive higher scores on different ranks, that is, $k = \text{argmax}_k (s_j)$. In this work, we only consider rank $k \in \{2, 3, 4\}$ since it is unlikely that a fixation overlays more than four AOs. Examples of flipping candidates of rank 2 for scatter, bar, line, and pie plots can be found in Figure 4. Here, fixations that receive scores $s_2 > 0.5$ are marked in blue, whereas fixations that receive scores $0 \leq s_2 \leq 0.5$ are marked in orange. As an example, in the pie chart, the fixation located on the pie is marked in blue since its density overlaps with two adjacent data segments to a similar extent. In contrast, the fixation located between the legend and source text is marked in orange since its density mostly covers white space.

Figure 3: Examples of flipping candidates of rank 2 (left) and rank 3 (right).

3.1.1 Metrics. We propose the Flipping Candidate Rate (FCR) to quantify the probability that fixations might flip between two or more AOs. Given the above definition of a flipping candidate, we can count the number of flipping candidates occurring in a scanpath and use it as a measure of its uncertainty. Since we intend to compare scanpaths of different lengths, we normalize the number of flipping candidates $C$ by length $N$ to get the FCR = $\frac{C}{N}$ of a scanpath. Therefore, the FCR is a metric that describes how much the gaze estimation error affects the mapping to AOs.

We then go one step further and measure how much the uncertainty in the AOI mapping impacts later analysis, here for the example of scanpath analysis. To this end, we apply Sequence Score [Yang et al. 2020], a pairwise string similarity metric that is normalized between 0 and 1, where 1 suggests a perfect match. Each fixation in a scanpath is mapped to a character in a string based on its AOI label. We alter all flipping candidates in a scanpath to the second most probable AOI according to the FCR, and compute the Sequence Score between the original and flipped scanpaths.
3.1.2 Result. We use a 0.5° calibration error as the criterion to separate gaze data into two groups: low calibration error and high calibration error groups. We apply 0.05° KDE Gaussian bandwidth to the low calibration error group, and 0.25° to the high calibration error group. We observed a significant difference in line plots and pie plots across low and high calibration error groups (post hoc t-test: line: $t(165) = 2.43, p = 0.016$, pie: $t(161) = 2.51, p = 0.013$), but the difference cannot be confirmed for bar plots ($t(159) = 0.54, p = 0.588$) or scatter plots ($t(156) = 1.38, p = 0.172$).

3.2 Fixations Hit Any AOI

In most public gaze datasets of information visualisations, the raw gaze data are not accessible [Borkin et al. 2015; Polatsek et al. 2018]. It means that the calibration error is unknown, and the analysis in Section 3.1 is not applicable. However, we still can assess the impact of gaze uncertainty, that is, by enlarging the size of all AOIs uniformly. Given that the AOIs in visualisations can be close to one another, a drift and error in the gaze estimates make fixations land outside AOIs, or even hit different AOIs. In this analysis, we gradually enlarge the size of the AOI annotation to account for slight drifts in the gaze estimates.

3.2.1 Metrics. Since human attention is not naturally drawn by low saliency regions such as white spaces [Matzen et al. 2017], if there are too many fixations that land on empty or white spaces (in our case parts of a visualisation without any annotation), it is likely
indicative of large gaze estimation error. Hence, we propose Hit Any AOI Rate (HAAR), which is defined as the HAAR = \frac{\text{HIT}}{\text{HIT} + \text{OFF}}, where HIT represents the number of fixations that hit at least one AOI, and OFF represents the number of fixations that are not within any AOI.

3.2.2 Result. Fixations in this section are all adapted from the default fixation detection algorithm integrated in the EyeLink 1000 Plus eye tracker. Figure 8 demonstrates that pie plots have the highest HAAR among visualisation types without any enlargement. When the enlargement factor is larger than 0.2°, bar plots have the highest HAAR among visualisation types.

Figure 8: AOI enlargement factor by visual angle and the Hit Any AOI Rate (HAAR). When the AOI enlargement factor is larger than 0.2°, HAAR of bar plots exceeds pie plots to become the highest.

4 DISCUSSION

In this work, we demonstrated that gaze uncertainty can significantly impact subsequent AOI-based analysis such as scanpaths, including comparisons of scanpaths using similarity metrics. Using the two proposed metrics, the Flipping Candidate Rate (FCR) and Hit Any AOI Rate (HAAR), we identified several interesting findings that we discuss below.

Flipping candidate threshold and AOI enlargement factor. Since there is no AOI ground truth for gaze data, i.e., which AOI the participants really look at, it is difficult to identify whether fixations are assigned to the correct AOI. Hence, there is no deterministic rule to set the flipping candidate threshold and AOI enlargement factor. Figure 5 shows that the flipping candidate rate varies between 0.01 to 0.40 as the threshold changes, so a proper threshold becomes crucial for later analysis. Figure 8 illustrates the HAAR under different values of the AOI enlargement factor. A 1.0° AOI enlargement makes the HAAR increase from below 70% to above 80% for all kinds of visualisation types. This shows that AOI enlargement is an effective strategy to compensate for the influence of gaze estimation error, especially when only fixations are available [Borkin et al. 2015]. This is also a bio-plausible strategy since humans can perceive visual information in the foveal region [Jonas et al. 2015].

Guidelines on parameter selection. Here, we give a general guideline about how to choose a proper flipping candidate threshold and an AOI enlargement factor. When AOIs are so close to one another that the gaze estimation error cannot be ignored, a smaller flipping candidate threshold (e.g. 0.2) and a larger enlargement factor (e.g. 0.8°) are desired. A smaller flipping candidate threshold ensures that more ambiguous fixations can be detected, and a larger enlargement factor compensates more for the gaze estimation error to correct slight shifts or drifts of the gaze data. However, enlarging the size of an AOI too much will lead to overlapping AOIs, which then points to an increase in the flipping candidate rate.

Gaze uncertainty across visualisation types. Figure 5 shows that line and pie plots have fewer flipping candidates than bar and scatter plots. Moreover, in Figure 7, the sequence scores of line and pie plots were more severe than line and pie plots on MASSVIS dataset [Borkin et al. 2013]. While it is unclear whether such differences between visualisation types also hold on different datasets or visualisation designs, it opens up an opportunity for future research. Future work should investigate methods to design and optimise, e.g., different parameters for bar and scatter plots, to make them more robust against gaze uncertainty and limit such effects on any subsequent analysis. Figure 4 gives an impression of how different thresholds influence the flipping candidates in each visualisation type. The extent to which flipping candidates may threaten the validity of later analysis stages depends on the actual AOIs it overlays. A flipping candidate that overlays AOIs that share the same labels will not cause a difference in later analysis stages. As shown in Figure 6, the majority of flipping candidates occurring in scatter and pie plots can be considered insignificant, such as XX (Axis), SS (Source etc.), and DD (Data). Moreover, there are flipping candidates that overlay semantically different AOIs, such as TS (Title, Source etc.), SL (Source etc., Legend), and XD (Axis, Data). For example, in Figure 4, the bar plots shows two fixations that overlay an axis and a data item, while the red fixation in the pie plots is an example of SL.

5 CONCLUSION

This paper demonstrated and assessed the impact of gaze uncertainty on AOIs in information visualisations by introducing two effective metrics: Flipping Candidate Rate (FCR) and Hit Any AOI Rate (HAAR). Results showed that gaze uncertainty is a key factor that may impact the credibility of eye tracking studies and any resulting findings. While this paper addresses information visualisations, it also has a strong potential as an extension to other visualisation fields, such as volume visualisation or scientific visualisation in general.

ACKNOWLEDGMENTS

Y. Wang, M. Koch, and D. Weiskopf were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 251654672 - TRR 161. M. Bâce was funded by a Swiss National Science Foundation (SNSF) Early Postdoc. Mobility Fellowship (grant number 199991). A. Bulling was funded by the European Research Council (ERC; grant agreement 801708).

We would like to thank reviewers for their insightful comments, Constantin Ruhdorfer for the gaze data processing toolbox, Nuo
Chen for eye-tracking data collection, Dominique Thomas for paper editing support, as well as Guanhua Zhang and Sruthi Radhakrishnan for helpful comments on this paper.

REFERENCES


