Vision Research 228 (2025) 108546



Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres





The Polar Saccadic Flow model: Re-modeling the center bias from fixations to saccades

Rotem Mairon, Ohad Ben-Shahar *

Department of Computer Science, Ben-Gurion University of the Negev, Israel
School of Brain Sciences and Cognition, Ben-Gurion University of the Negev, Israel

ARTICLE INFO

Keywords:
Eye-movements
Saccades
Center bias
Polar representation
Mixture models

ABSTRACT

Research indicates that a significant component of human eye movement behavior constitutes a set of consistent biases independent of visual content, the most well-known of which is the central bias. While all prior art focuses on representing saccadic motion and biases in Cartesian retinotopic coordinates, here we propose the Polar Saccadic Flow model, a novel approach for modeling saccades' space-dependent biases in a polar representation. By breaking saccades into orientation and amplitude, the Polar Saccadic Flow model enables more accurate modeling of these components, leading also to a better understanding of the saccadic bias. Moreover, the polar representation also uncovers hitherto unknown patterns and biases in eye movement data, allowing for a more detailed and nuanced analysis of saccadic behavior. These findings have implications for the study of human visual perception, can help to develop more accurate eye movement models, and also may improve eye tracking technologies.

1. Introduction and background

The center bias is a prevalent tendency in human eye movement behavior. Regardless of various factors that influence eye movements, such as task demands, the visual stimuli itself, and individual differences, observers consistently exhibit a strong inclination to direct their gaze centrally. Understanding this persistent center bias is crucial as it can shed light on the underlying mechanisms driving visual attention and perception.

The center bias holds substantial implications and is key for disentangling content-related viewing preferences and inherent biases in gaze planning. Clearly, by effectively modeling the center bias, computational models may produce more realistic and more accurate predictions of human eye movement behavior. For instance, Kümmerer, Bethge, and Wallis (2022) employed an ablation strategy to decompose the center bias into components due to image content, initial central fixation, and other factors, providing a nuanced understanding of how different elements contribute to the center bias. Such approaches enhance our comprehension of the mechanisms driving human eye movements and improve our ability to predict these movements based on visual information.

Research into the center bias in eye movements has been facilitated greatly by the use of human *fixation maps*. These maps show how fixations are distributed across a visual stimulus, and were key in identifying the prevalence of central fixations in the first place. At the same time, fixation maps completely lack temporal information and thus they are unable, for example, to provide information on *when* the center bias manifests itself, e.g., is it more prevalent at stimulus onset, reflecting an initial capture of attention, or do central fixations appear perhaps later during a trial, indicating a more strategic use of attention? Rothkegel et al. (Rothkegel, Trukenbrod, Schütt, Wichmann, & Engbert, 2017) addressed the first part of this question by examining changes in fixation distances from the center when the initial saccade is delayed relative to the sudden onset of a stimulus. Their findings confirm that the center bias is most pronounced at the onset of viewing, illustrating the limitations of static fixation maps and emphasizing the need for more dynamic measures to fully understand the center bias in scene viewing.

Studies that do explore temporal aspects of human eye-movement behavior focus on where people fixate *next*, i.e., given the last fixation, or perhaps the sequence of past fixations, where would an observer saccade next. Such studies suggest that regularities may be found not only in where we look, but also in the way we shift our gaze from one location to another. It has been shown, for example, that when free-viewing images on a screen, observers favor nearby locations in the displayed image over distant ones (Gajewski, Pearson, Mack, Bartlett, & Henderson, 2005; Pelz & Canosa, 2001; Tatler, Baddeley, & Vincent, 2006) and that horizontal saccades are more dominant (Bair & O'Keefe,

* Corresponding author at: Department of Computer Science, Ben-Gurion University of the Negev, Israel.
E-mail address: ben-shahar@cs.bgu.ac.il (O. Ben-Shahar).

https://doi.org/10.1016/j.visres.2025.108546
Received 26 September 2023; Received in revised form 7 September 2024; Accepted 17 January 2025
Available online 31 January 2025
0042-6989/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1998; Lee, Badler, & Badler, 2002; Moeller, Kayser, Knecht, & König, 2004).

The existence of saccadic biases clearly indicates the need to include it in any attempt to model and predict entire scanpaths computationally. Yet, most existing models tend to account for these biases only implicitly or to overlook them entirely. Early models incorporate the transition between fixations based on their saliency measure (Itti & Koch, 2000; Itti, Koch, & Niebur, 1998; Walther & Koch, 2006), implying that scanpaths are formed only *after* the scene has been preprocessed. Others view scanpath formation as a dynamic process driven by information maximization (Lee & Yu, 1999; Wang et al., 2011), while yet another general approach is to learn from human scanpath data, thereby capturing saccadic regularities only indirectly (Jiang et al., 2016; Liu et al., 2013; Ngo & Manjunath, 2017). Indeed, only a limited number of models in the literature incorporate saccadic biases explicitly. For example, one such approach uses the joint probability distribution of saccade amplitudes and orientations to model the conditional probability of making a saccade to the next location for scanpath generation (Le Meur & Liu, 2015).

When it comes to the center bias, models that focused thus far on predicting the distribution of fixations often leverage this bias in order to enhance their predictive capabilities (Judd, Durand, & Torralba, 2012). This is typically achieved by weighting predictions with a static spatial distribution that reflects the tendency for observers to direct their gaze centrally, resulting in improved fixation predictions. The same approach has also been applied in models aimed at predicting entire scanpaths (Kümmerer et al., 2022). However, the specific dynamics of how the center bias in manifested in saccadic movements themselves during free viewing, beyond spatial weighting techniques, remains relatively unexplored.

In light of this discrepancy, the temporal sequence of fixations plays a crucial role in understanding the dynamic nature of visual attention. This becomes particularly relevant in the context of developing computational models that aim to predict complete human scanpaths, as these models require both spatial accuracy and temporal precision. However, while some saccade prediction models incorporate certain temporal aspects like fixation durations (Engbert, Trukenbrod, Barthelmé, & Wichmann, 2015; Schütt, Rothkegel, Trukenbrod, Reich, Wichmann, & Engbert, 2017; Sun, Chen, & Wu, 2019), they typically do not explicitly address any potential influence of the center bias throughout the sequence of saccades. Moreover, not unlike its critical contribution in fixation prediction, it is clearly essential to investigate the role of the center bias in predicting saccades too. Incorporating these two aspects (i.e., temporal regularities and the center bias) into predictive scanpath models is likely to improve their power significantly.

As mentioned above, the prior art largely lacks a comprehensive exploration of the dynamic manifestation of the center bias during free viewing. However, Clarke, Stainer, Tatler, and Hunt (2017) introduced the *Saccadic Flow* (SF) model, which quantifies the probability of saccade endpoints conditioned on their launch locations. Consequently, this allows capturing the inherent tendency of saccades to "flow" in specific directions or towards particular regions of the visual field, regardless of the visual stimulus, providing a key component in seeking models of the temporal dynamics of scanpath programming.

At the backbone of the SF model are 2D normal distributions of saccade endpoints. However, in this paper, we show that such a spatial representation in Cartesian coordinates may not fully capture the underlying dynamics of the center bias. Alternatively, we propose the *Polar Saccadic Flow* (PSF) model, a novel approach that employs a polar representation of saccades, separating them into orientation and amplitude components. This alternative representation allows for more precise computational modeling of each component, to better capture their distinct properties, and to provide deeper understanding of the dynamic nature of the center bias, and the intricate nature of eye movement behavior in general. We emphasize that while this approach involves separate modeling of orientation and amplitude, it does not assert or make any assumption about their independence (or lack thereof), but rather lays the ground for future studies to explore potential interactions or dependencies between these components more comprehensively.

2. Previous related work

As mentioned above, the center bias is commonly observed and discussed in the context of fixation maps. These maps convey the spatial organization of fixations across a visual stimulus, but they do not provide insights into *temporal* aspects of eye movements. Some studies do explore temporal aspects such as fixation duration (Nuthmann, 2017) and saccade latency (Sumner, 2011), while others peek into ordinal data such as the spatial organization of consecutive fixations. This type of data was found particularly constructive in identifying certain biases in saccadic events, such as preferences for specific saccade amplitudes (Bahill, Adler, & Stark, 1975) or dominance of horizontal saccades under free-viewing conditions (Bair & O'Keefe, 1998; Lee et al., 2002; Moeller et al., 2004).

The idea that consecutive fixations bear significance is not only intuitive but was also formalized by Clarke, Green, Chantler, and Hunt (2016) in their stochastic search model of saccades. Part of that model relies on the finding that the distribution of saccade endpoints very much depends on their starting position (see Fig. 1a using the much larger and diverse CAT2000 data (Borji & Itti, 2015)). This concept was re-implemented in a follow-up model dubbed the aforementioned Saccadic Flow (SF) (Clarke et al., 2017), a generalization of the center bias that incorporates the saccade origin too. More formally, instead of characterizing the prior probability $p(x_i, y_i)$ of fixating at (x_i, y_i), the SF model seeks the conditional probability $p(x_i, y_i | x_{i-1}, y_{i-1})$.

More specifically, to characterize how the endpoint distributions vary with saccade origin, the SF model employs a sliding window approach across the image space. All saccade landing sites that *originate* from an $n \times n$ window are used to fit a bivariate normal distribution. Multivariate polynomial regression is then used to fit fourth-order polynomials to each of the distribution parameters as they vary across the image (see Fig. 1b).

As long as saccades are represented as pairs of (start and end) points in the Cartesian plane, using a bivariate normal distribution to model how endpoint fixations are distributed in space may be natural. However, comparing the actual and modeled distributions in Fig. 1a, this choice appears sub-optimal, not only by the occasional discrepancies between the modeled means (i.e., the center of the Gaussian ellipses) and the peaks of the actual distributions, but also in their overall shape.

In the sections that follow, we explore a different representation of saccades that leads to a better modeling of the space-dependent bias in the actual data. Specifically, we propose the Polar Saccadic Flow model (PSF), based on a polar representation of saccades that divides them into their orientation and amplitude components. This facilitates the modeling of each component individually, leads to a more precise model as a whole, and provides new insights about eye-movement behavior and the mechanisms that govern it.

3. Dataset and preprocessing

To explore the effect of saccade representation in polar space, eye movement data is essential. In our analysis and modeling we used the publicly available CAT2000 eye-tracking dataset (Borji & Itti, 2015), with 2,000 natural indoor and outdoor scene images, each $1,920 \times 1,080$ pixels in size, from 20 categories (including "Action", "Art", "Indoor", "Object", "Pattern", "Social", and others). Each image was observed by 18 subjects while their eye movements were recorded using the Eyelink 1000 device. Each trial began with observers fixating on a central cue marker, after which they freely viewed an image. Saccades were identified based on velocity and acceleration thresholds, as recommended by the device manufacturer.

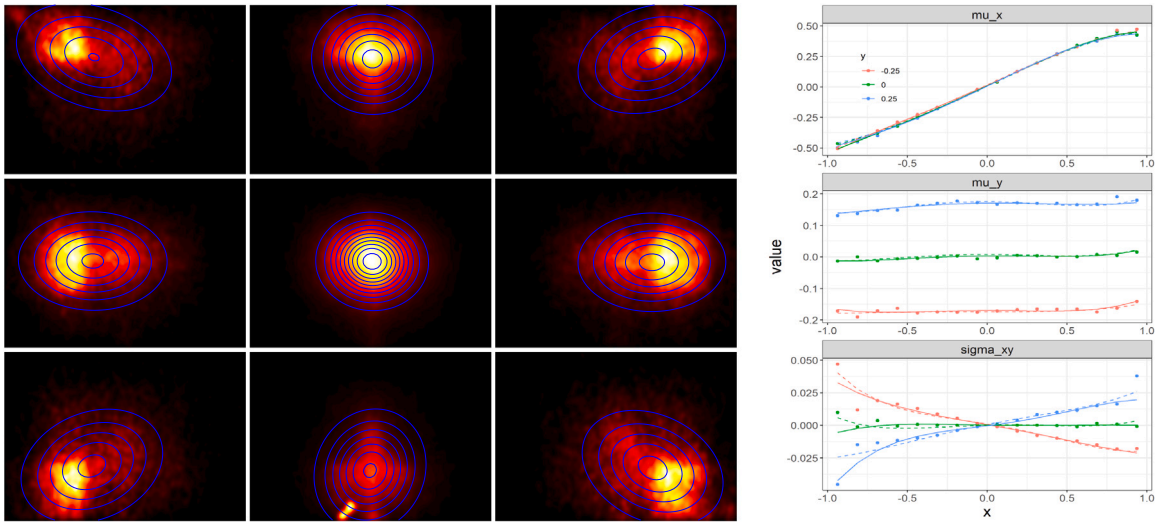


Fig. 1. Visualization of space-dependent biases in saccadic behavior and an approximation of the SF model using endpoint fixation data for the CAT2000 dataset (Borji & Itti, 2015). **(a)** Each heat map shows the spatial distribution of endpoints of those saccades that were launched from the corresponding region in image space after it was divided into a 3 × 3 grid. These heat maps are similar to those presented in Clarke et al. (2017), as both depict endpoint fixation distributions, but here we use the more comprehensive CAT2000 data. A bivariate normal distribution (blue ellipses) was fitted to each spatial distribution and superimposed on each heat map to approximate the behavior expected from the SF model, here without a global fitting procedure. **(b)** The three graphs show how the mean and covariance parameters of the distribution vary with the saccade launch site as it is moves across the x-axis for 3 different y-axis values in the normalized display ($y = -0.25$, 0, and 0.25). These graphs were generated using code provided by Clarke et al. (2017) and illustrate the global model's parameter changes, with both standard (dotted) and robust (solid) polynomial regression fits.

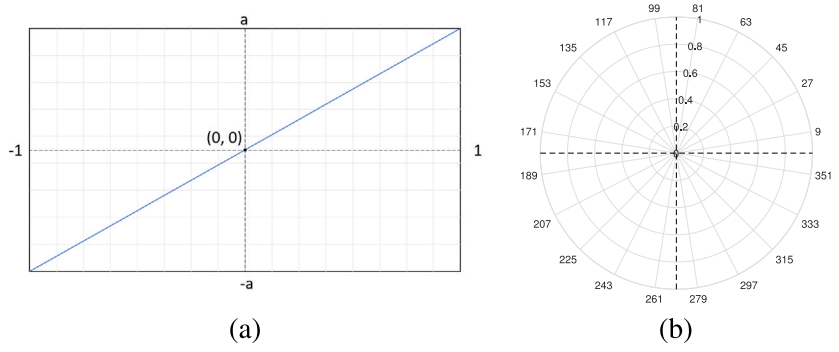


Fig. 2. Technical setup of the Polar Saccadic Flow model. **(a)** The normalized two-dimensional space of fixations. The parameter a denotes the aspect ratio of the image, which is 0.5625 for all images in the CAT2000 dataset (Borji & Itti, 2015). The blue diagonal line indicates the maximum possible saccade amplitude of $2\sqrt{1+a^2}$. **(b)** The quantization scheme of the polar representation used for our analysis of saccade orientations and amplitudes. Due to space constraints and for clarity, the figure only shows 5 representative amplitude bins out of the total 20 discussed in the text.

The purpose of the cue marker employed during the CAT2000 data collection was to ensure that all observers begin from the same point. However, in the context of the center bias, this may contaminate or confound the data or otherwise produce irrelevant information. To ensure data integrity and relevance to our type of analysis, we excluded the first saccade in each trial as well as saccades beyond stimulus boundaries. In addition, we adopted Clarke et al.'s (Clarke et al., 2017) approach to normalize fixation to a canonical domain (Fig. 2a). Technically, we normalized all fixations (x, y) to Cartesian region $(-1, 1)\times(-a, a)$, where a is the aspect ratio of the image (0.5625 for all images in the CAT2000 dataset (Borji & Itti, 2015)). We then remapped and represented all saccades in polar coordinates by their orientation and amplitude. Note that in normalized coordinates the maximum possible saccade amplitude is given by $2\sqrt{1+a^2}$, the length of the domain diagonal (Fig. 2a).

To allow constructive characterization of saccade orientation and amplitude as a function of saccade origin, the image space must be quantized into processable launch sites. Because all of the images in the dataset have a $9:16$ aspect ratio, we divided the image space into 9×16 grid of square, equiareal launch sites. For the images used in our analysis this translated to rectangular regions of 120×120 pixels in size (see Fig. 2a). We then collected all saccades originating from each site and represented them by their orientations and amplitudes, the latter being further normalized by the maximal possible saccade amplitude mentioned above (~2.3 in our case).

We followed by calculating a pair of corresponding histograms depicting how the orientation and amplitude were distributed for their site. Because the maximum normalized saccade amplitude in the data did not exceed 1.0 (and in reality, reached ~0.84), amplitude values were binned into 20 bins of equal width in the range $[0, 1]$. Similarly, to calculate orientation histograms, we divided the unit circle into 20 bins with equal angular width of 18 degrees in the range $[-\pi, \pi]$ (see Fig. 2b). The resultant sets of $9 \times 16 = 144$ pairs of histograms are presented in Figs. 4 and 6, and serve as the raw data for the statistical analysis and computational modeling discussed next. In addition, Fig. 10 is a Cartesian depiction of Fig. 4 and note that all 3 figures contain additional modeling information beyond the raw data, to be discussed in the remainder of this paper. Finally, Fig. 3 is a full-size depiction of one selected launch site from all 3 figures (in this case, launch site $(1, 12)$), brought as a reference and legend for the different

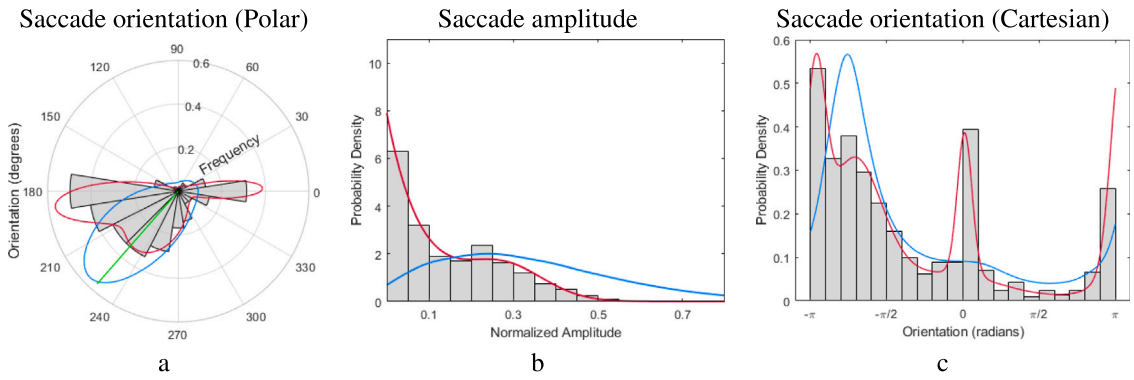


Fig. 3. Full-size depiction of one selected launch site (1, 12) from Figs. 4, 6, 10. In all figures the bars (either polar or Cartesian) depict the raw data, the blue curves represent SF model predictions, and the red curves are predictions by the proposed PSF model. **(a)** Polar representation of the orientation data and models, as in Fig. 4. The green line denotes the direction from the launch site to the center of the image. **(b)** Representation of the amplitude data and models as in Fig. 6. **(c)** Cartesian representation of the Orientation data and models in as in Fig. 10.

types of information incorporated in each graph in the full arrays. When inspecting Figs. 4, 6, and 10, please zoom in accordingly on individual panel using the digital version.

4. Saccade orientation

We begin by analyzing and modeling the orientation distributions. As we suspected, the orientation histograms just described (Fig. 4) reveal more informative spatial patterns that are not available in the classical fixation maps. To begin with, one can notice two main tendencies: in launch sites closer to the margins of the image space, orientation distributions can be generally characterized by a single, low-variance, and dominant mode that tends to point towards the center of the image space. On the other hand, in launch sites near the center of the image space, orientation distributions become multi-modal with a larger total variance.

This behavior is not unexpected. The viewer has little reason to saccade beyond the image boundaries (although clearly, this in itself does not necessarily imply a preference towards the center of the image space or a unimodal distribution). Additionally, as we move away from the boundaries, the eye has a greater range of potential targets to move to, thus the observed increase in orientation variance. However, considering the full 9-by-16 grid of launch sites reveals intriguing non-trivial patterns, and in particular certain symmetries in the orientation distributions, manifested along both the horizontal and vertical dimensions.

To parse the intricate symmetries present in the orientation distributions more easily, let us delineate and focus on the first (i.e., top right) quadrant of the image space, namely rows 1 to 5 and columns 9 to 16. In the top row of this quadrant, the rightmost distribution exhibits a distinct mode oriented predominantly leftward (and somewhat downward, as will be discussed later). As we move closer to column 9, we observe this leftward mode gradually dispersing, giving rise to additional rightward and downward modes. A similar trend is observed in the following rows in the same quadrant.

In essence, the distributions within each row, starting from column 16, primarily point leftward and exhibit two additional "forces" as we traverse left to more central launch sites: a downward force pulling towards the bottom of the quadrant and a repulsive force pushing away from its left boundary. These forces present two intriguing characteristics. First, their intensities appear to increase non-linearly, and second, they begin to manifest later in rows that are closer to the horizontal axis of the image space.

Upon examining the column-wise variations in distributions, another detail surfaces that further refines our understanding of the downward pulling force. In the rightmost column, the single leftward mode shifts gracefully towards a horizontal orientation as we descend the rows towards the row 5. This pattern suggests a consistent downward pull across the first quadrant, manifested as an increasingly horizontal component in the rightmost column's distributions and as a downward mode in distributions closer to the vertical center of the image space.

Taken together, these findings delineate a saccadic landscape in the first quadrant dominated by a leftward and downward bias, accentuated near the central axes of the image space. Intriguingly, also noticeable is a repulsion away from the vertical axis—a phenomenon that has no counterpart along the horizontal axis. This saccadic behavior in the first quadrant mirrors horizontally and vertically across the grid, revealing a saccadic flow with a pronounced center-oriented bias. Saccades are horizontal across the horizontal axis but are not exclusive to this orientation near the center. Similarly, near the vertical axis, saccades trend towards the vertical but also exhibit horizontal orientations. This nuanced balance between symmetric and asymmetric forces underscores the dynamic nature of saccadic behavior and the central bias in eye movements.

To better understand these patterns, we next seek to model the orientation behavior in each launch site in a formal manner. As certain launch sites exhibit multiple dominant orientations (cf. Fig. 4), we hypothesize a mixture model of four components, each distributed normally in the orientation domain and representing one of the cardinal directions. Importantly, to accommodate the circular nature of the domain, we suggest a Von-Mises Mixture Model (VMM), which is defined as follows.

$$p(\theta) = \sum_{i=1}^{4} \pi_i f_i\left(\theta_i|\mu_i,\kappa_i\right) \tag{1}$$

where each component is a Von-Mises distribution

$$f_i\left(\theta_i|\mu_i,\kappa_i\right) = \frac{e^{\kappa_i \cos(\theta_i-\mu_i)}}{2\pi I_0\left(\kappa_i\right)} \tag{2}$$

and $I_0(\kappa)$ is the modified Bessel function of order 0, i.e.,

$$I_0\left(\kappa_i\right) = \sum_{n=0}^{\infty} \frac{1}{(n!)^2}\left(\frac{\kappa_i}{2}\right)^{2n}. \tag{3}$$

The parameters $\pi_i > 0$ are the mixture weights and satisfy $\sum_{i=1}^{4} \pi_i = 1$. The parameter μ_i of each VMM distribution represents its "mean" direction, while the parameter κ_i is its "concentration", representing how spread out it is, akin to the variance of normal distribution in flat domains. Given the non-linear nature of the mixture model, fitting it to the measured distribution in each launch site can be done with Expectation-Maximization (EM). We initialize the EM algorithm using K-means clustering to obtain a rough estimate of the mean direction and concentration.

Following the above, and in addition to the raw data, Fig. 4 shows the 4-VMM (red curve) that best fits its corresponding data in each

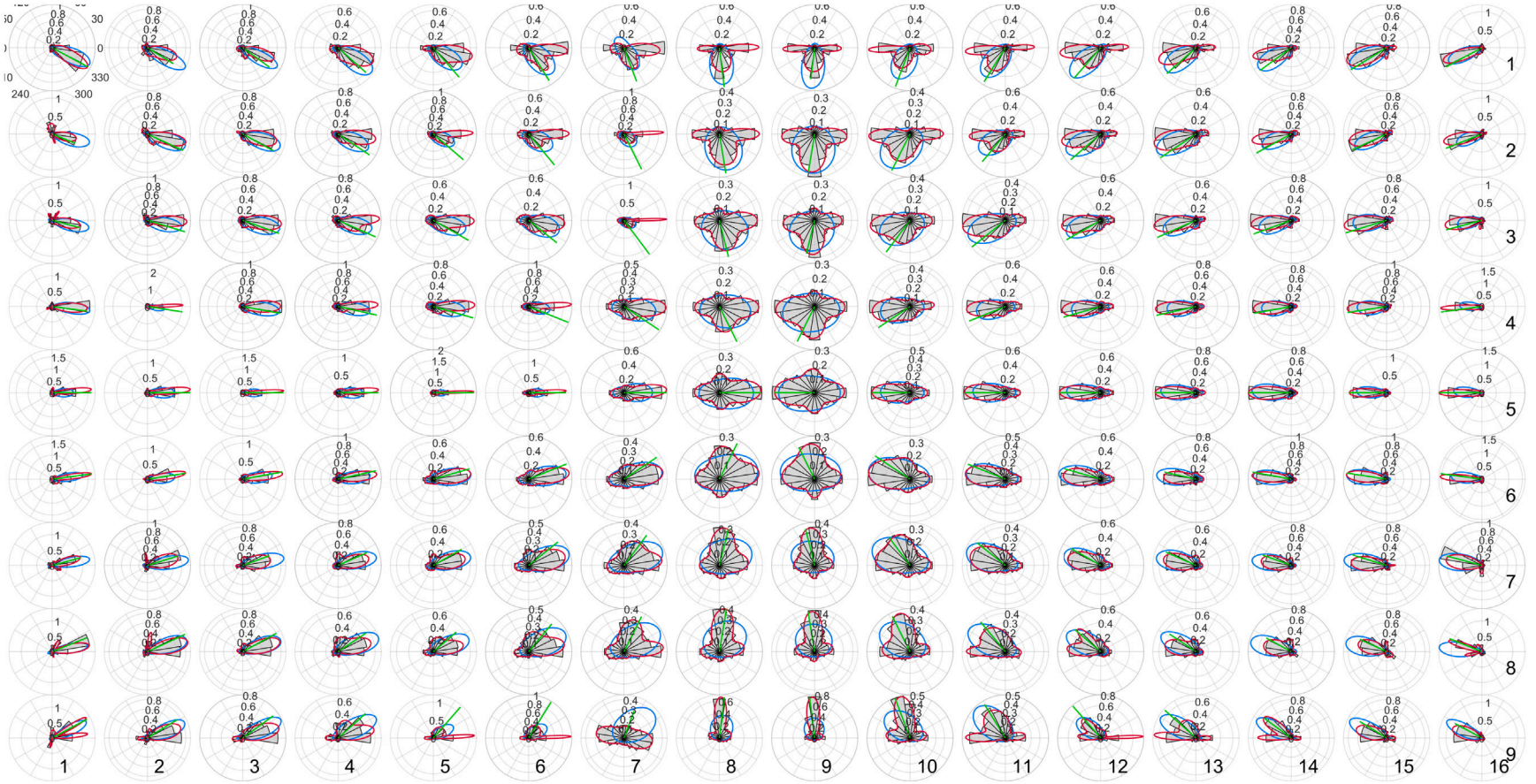


Fig. 4. Raw saccade orientation distributions and PSF modeling. Polar histograms depict saccade orientation distributions for individual launch sites. Raw data (cf. Sec. 3) is presented as polar bar plots, while the red polar curves represent the optimal four-component Von Mises Mixture model fitted to each launch site. Blue polar curves depict the corresponding SF model via marginalization of orientation and interpolation. Note the differences between the two models make vis-a-vis the raw data. For reference, the green line within each launch site marks the direction from the center of the launch site to the center of the image space. Please use the digital version to zoom in on the details of each launch site of interest.

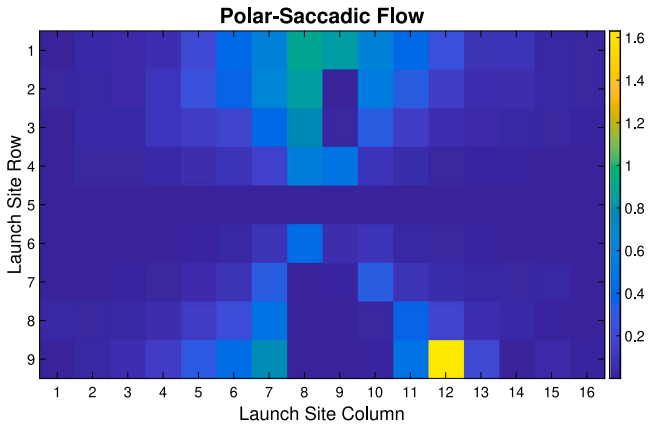


Fig. 5. Color-coded angular deviation (in radians) of the preferred saccade orientation from the direction of the image center (with respect to each launch site). The preferred orientation is modeled in each site as the mean of the most dominant component in the VMM distribution at that site. With the exception of one salient outlier, a systematic bow-like pattern clearly emerges.

launch site on the 9-by-16 grid (The figure also presents the prediction of the SF model for comparison, but this is discussed later in Section 6). Each panel also shows a green reference line of the direction towards the center of the image space. Notably, the dominant orientation does not always align with this center-pointing direction in all launch sites. For instance, while the dominant orientation in the top-left launch site $(1, 1)$ is almost perfectly aligned with the direction towards the image center, other launch sites, such as $(5, 2)$, show noticeable discrepancies.

This variability in alignment between the dominant orientation in each launch site and the respective center-pointing direction underscores the complexity of the saccadic flow patterns within the image space. It suggests that the center bias is not merely a uniform tendency to direct gaze towards the center from each launch site, but is rather the emergent outcome of a more intricate flow of saccades. This flow, as previously outlined, manifests in dominant tendencies that are not only directed towards the horizontal and vertical centers of the image space but also involve more subtle, spatially-dependent biases.

To quantify these discrepancies across different launch sites, we measured the angular difference between the mean of the most dominant component of the fitted distribution and the direction towards the center. The color-coded results are shown in Fig. 5, where a bow-like pattern clearly emerges. As can be observed, the discrepancy between the dominant bias component in saccade orientation and the image center is influenced by the proximity of launch sites to either the cardinal or the oblique axes of the image space. The preferred bias orientation generally aligns with the direction to the center when considering launch sites along the cardinal axes. However, as launch sites get closer to the oblique axes from the vertical center, the two angles begin to deviate.

The structured pattern in Fig. 5 and the fitted models in Fig. 4 thus reveal that the center bias is far more complex than a simple space-invariant tendency towards the center. Saccades tend towards the *horizontal* and *vertical* centers when launched from locations closer to the oblique axes, and more directly towards the center when launched from locations closer to the cardinal axes. The set of local computational models (red graphs in Fig. 4) quantifies these biases formally.

5. Saccade amplitude

Next to the analysis of orientation from the previous section, the raw saccade amplitude distributions (i.e., bar histograms from Fig. 6) can be analyzed separately. Again, a horizontally symmetric pattern emerges across the image space, where launch sites closer to the center of the image space exhibit saccade amplitudes that are significantly weaker and less variable. As we move away from the center, the amplitude becomes more variable, and larger saccades are present also.

A closer examination of the data, however, reveals a more nuanced aspect of these distributions, particularly in launch sites far from the center. There the amplitude distributions exhibit clear *bi-modal* behavior, indicating *two* distinct underlying processes. Specifically, these processes appear to generate a rapidly declining sub-distribution of short (small amplitude) eye movements and a symmetric unimodal component of longer saccades. This extends the commonly accepted property of saccades by which their frequency simply decreases with their amplitude (Castelhano, Mack, & Henderson, 2009; Wang et al., 2011). The data in Fig. 6 clearly suggest a more complex, dual-process behavior, and an amplitude–frequency relationship which is also space-dependent. Indeed, even without the details of a dual-process mechanism, the mean amplitude of saccades in each launch site varies spatially. In particular, the salient space-dependent concentric pattern that emerges in Fig. 7 clearly indicates a dependency on the distance from the center of the image.

The different observations about saccade amplitudes inspire a natural mixture model combining an exponential distribution for the short saccades and a normal distribution for the longer-range ones, which together may capture the complexity of human gaze during free viewing. Formally, we suggest an expo-normal model where the distribution of saccade amplitudes at a given image space location (i.e., launch site) is given by

$$f(x) = p \cdot \mathcal{N}(x|\mu, \sigma^2) + (1 - p) \cdot \mathcal{E}(x|\lambda) \tag{4}$$

where

$$\mathcal{N}(x|\mu, \sigma^2) \equiv \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5}$$

and

$$\mathcal{E}(x|\lambda) \equiv \lambda e^{-\lambda x} \tag{6}$$

and the parameter $0 \leq p \leq 1$ is the mixing coefficient.

Fig. 6 depicts in red the best-fitted expo-normal model at each launch site. It is easy to observe the excellent match to the raw data. With such a model in place, we are now able to scrutinize its components, as representing two distinct processes. It is particularly tempting to seek the properties of the normally distributed longer-saccades, not only because this process is missing from previous models, but because it too appears to be space-dependent. For example, Fig. 8 depicts the mean of that latter process in isolation. Clearly, a roughly concentric pattern emerges again, suggesting that the amplitude of longer saccades increases as we move away from the center.

6. Comparison with the saccadic flow model

To assess the contribution of a polar representation of saccades to modeling local saccadic behavior and its biases, we sought to compare the PSF model to its SF counterpart (Clarke et al., 2017). However, the latter employs a spatial bivariate normal distribution $f(x, y) = \mathcal{N}(\mu, \Sigma)$ to model explicitly the *location* of each saccadic endpoint relative to its launch site, with no direct modeling of the orientation and amplitude components. Formally, for each site, the endpoints of saccades launched from that site are fitted with a bivariate normal distribution, $f(x, y) = \mathcal{N}(\mu, \Sigma)$ in Cartesian coordinates.

To compare, we thus must converge to a common representation and derive univariate distributions of saccade orientation and amplitude from each SF bivariate distribution of fixation locations. To that end, we transformed each of these SF distributions from Cartesian to polar coordinates, where θ and r denote orientation and amplitude, respectively. Using this representation, we then computed the marginalized, uni-variate distributions of θ and r, as illustrated in Fig. 9. (For a detailed derivation of the marginalization, the reader

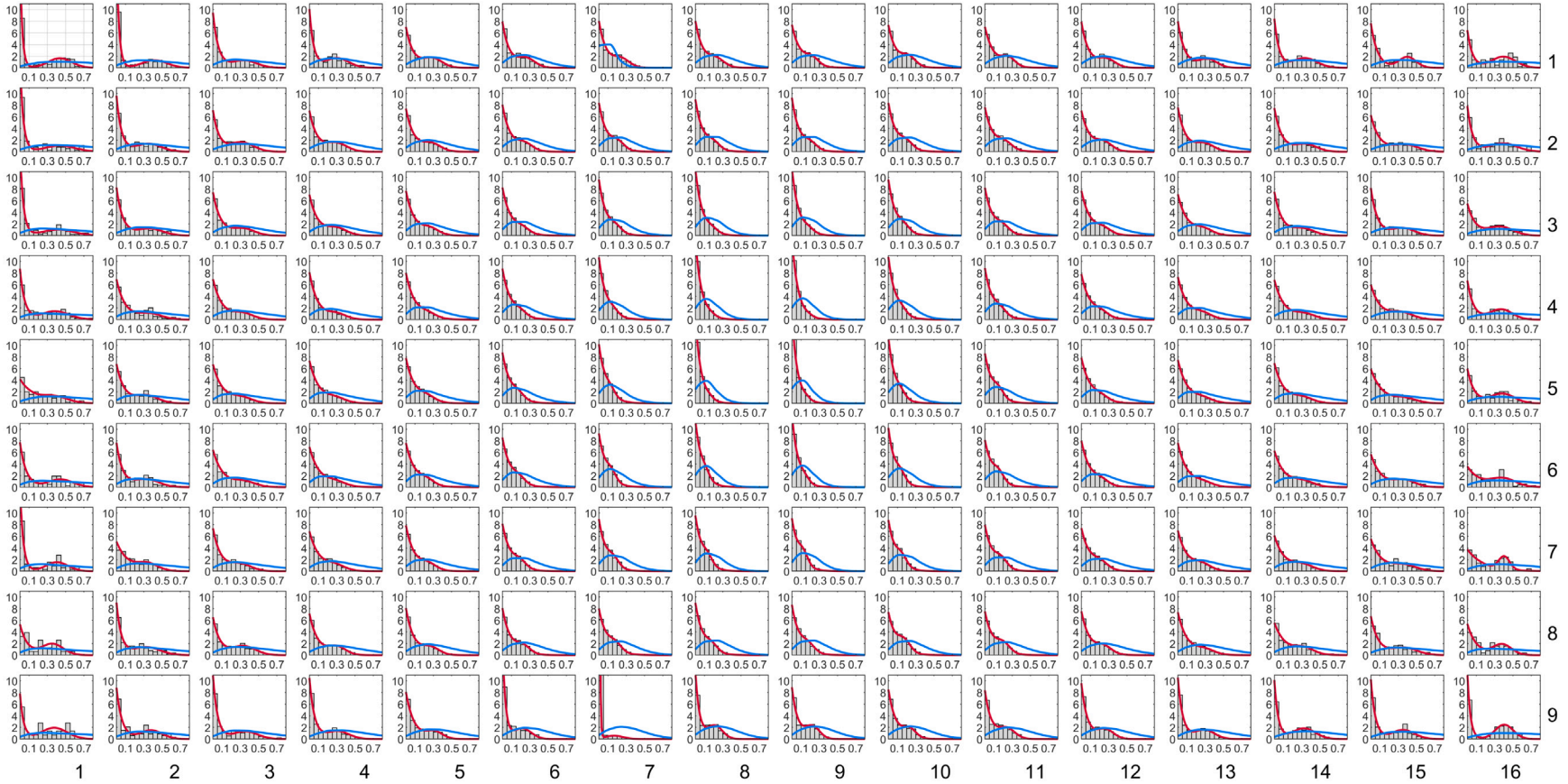


Fig. 6. Raw saccade amplitude distributions and modeling. Histograms depict the distributions of saccade amplitudes for individual launch sites. The raw data (cf. Sec. 3) is presented as bar plots. The red graphs represent the expo-normal distribution fitted to each launch site, while the corresponding SF model is depicted in blue via amplitude marginalization and interpolation. Please use the digital version to zoom in on each launch site of interest.

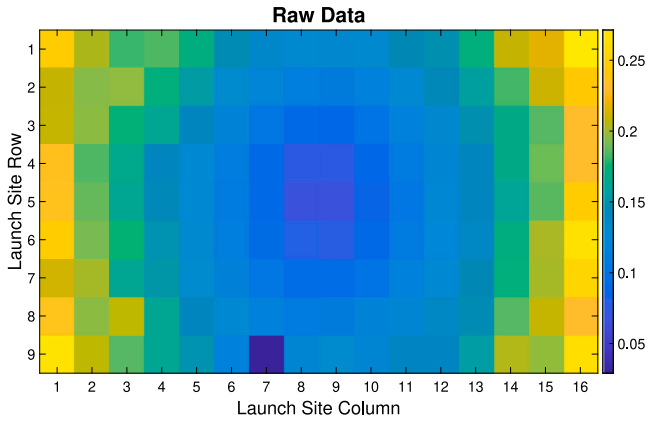


Fig. 7. Mean saccade amplitude across the image space. Note the concentric pattern, indicating dependency on the distance from the center.

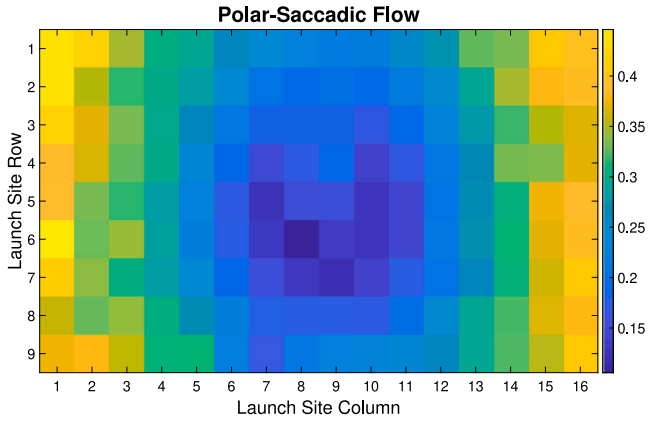


Fig. 8. Mean of the normal component of the expo-normal mixture model fitted to the raw saccade amplitude in each launch site. Note the general dependency on the distance from the center of the image.

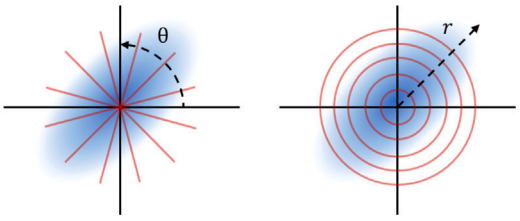


Fig. 9. Marginalizing a bivariate normal distribution. **Left**: Radial marginalization, which integrates the values of the bivariate distribution across radii to produce the probability in terms of orientation, θ. **Right**: Angular marginalization, in which the probability is expressed as a function of the distance from the origin, r, regardless of orientation.

is kindly referred to Cooper and Farid's technical paper, "A Toolbox for the Radial and Angular Marginalization of Bivariate Normal Distributions" (Cooper & Farid, 2020)). Importantly, this separation of orientation and amplitude is essentially a change of representation that does not affect their inherent dependence or independence status in the original data. Instead, it simply facilitates a more proper comparison between the PSF and SF models within a common framework.

6.1. Orientation

Fig. 4 presents (in blue) the prediction of the SF model in terms of saccade orientations next to the VMM component of the PSF model (in red) that was already discussed in Section 4. For completion, Fig. 10 shows the same two models (and the raw orientation data) after "flattening" the polar domain to a standard Cartesian coordinate frame. Such flattening may emphasize the differences between the predictive power of the competing models and is provided here as an additional presentation aid. Both models appear to produce distributions with a similar structure in peripheral launch sites. However, when the observed data exhibit multiple modes, the VMM model fits the data much better. For instance, at the launch site (2, 3) in Fig. 10, the VMM distribution (red graph) has two modes, which more closely matches the observed distribution than the single mode distribution derived from the SF model. The ability of the VMM to more accurately capture the multi-modality in observed distributions becomes more apparent as we examine launch sites closer to the center of the image space.

To measure the discrepancy between the competing models and the observed data at each launch site, we calculated the KL-divergence for the orientation component across the center points of the histogram bins. Fig. 11 shows the estimated KL-divergence value per launch site in a color-coded map. Both models exhibit divergence values that are larger at certain launch sites than others, and in particular, both seem to perform somewhat better *closer to* the horizontal center than away from it. However, as they are plotted at the same divergence scale (color map), it is easy to see that the KL-divergence values of the PSF's VMM model, and their variance across the visual field, are much smaller, a likely result of its better capacity to capture the multi-modality of observed saccadic distributions in different launch sites. Quantitatively, the average KL-divergence over the entire field of view was found to be $\sim 0.17 \pm 0.18$ (mean $\pm$ std) for our proposed VMM distributions, less than half of the $\sim 0.41 \pm 0.32$ (mean $\pm$ std) when using the distributions derived from the SF model. An independent two-sample t-test comparing these divergence values showed their difference to be strongly statistically significant ($t(286) = 8.01, p < 10^{-5}$).

6.2. Amplitude

Fig. 6 compares the expo-normal amplitude component of the PSF model with the corresponding marginal distribution derived from the SF model. The difficulty of the SF model to match the density around the low amplitude values is evident. The expo-normal distribution, on the other hand, appears to allow for a significantly better representation of the data, including capturing the bimodal behavior when it exists.

Fig. 12 shows the KL-divergence values per site between the observations and the two competing models. The expo-normal component of the PSF model significantly outperforms the amplitude component of the SF model. The average KL-divergence values (over the entire field of view) are $\sim 0.85 \pm 2.26$ (mean $\pm$ std) for the PSF model and $\sim 23.28 \pm 5.54$ (mean $\pm$ std) for the SF model. Statistical significance analysis using an independent two-sample t-test yields $t(286) = 44.96$, $p < 10^{-5}$, confirming highly significant differences.

6.3. Log-likelihood comparison

In addition to the previously described evaluations, we conducted a quantitative comparison using log-likelihood (llh) scores to directly measure the predictive advantage of both the PSF and SF models relative to a uniform distribution. For this analysis, we utilized the CAT2000 dataset and randomly partitioned it into 80% training and 20% testing subsets. Both models were trained (i.e., fitted) on the training subset, and log-likelihood calculations were performed on the testing subset. Consistent with the methodology of the original SF study (Clarke et al., 2017), we computed llh scores through a sampling scheme that involved 100 repetitions, where each constitutes 1,000 random samples from the test set. The comparative results are visually summarized in Fig. 13.

For the amplitude component, our PSF model demonstrates a significant improvement in predictive accuracy. The mean llh score for our PSF model in this test was 1187, nearly doubling the score of 625

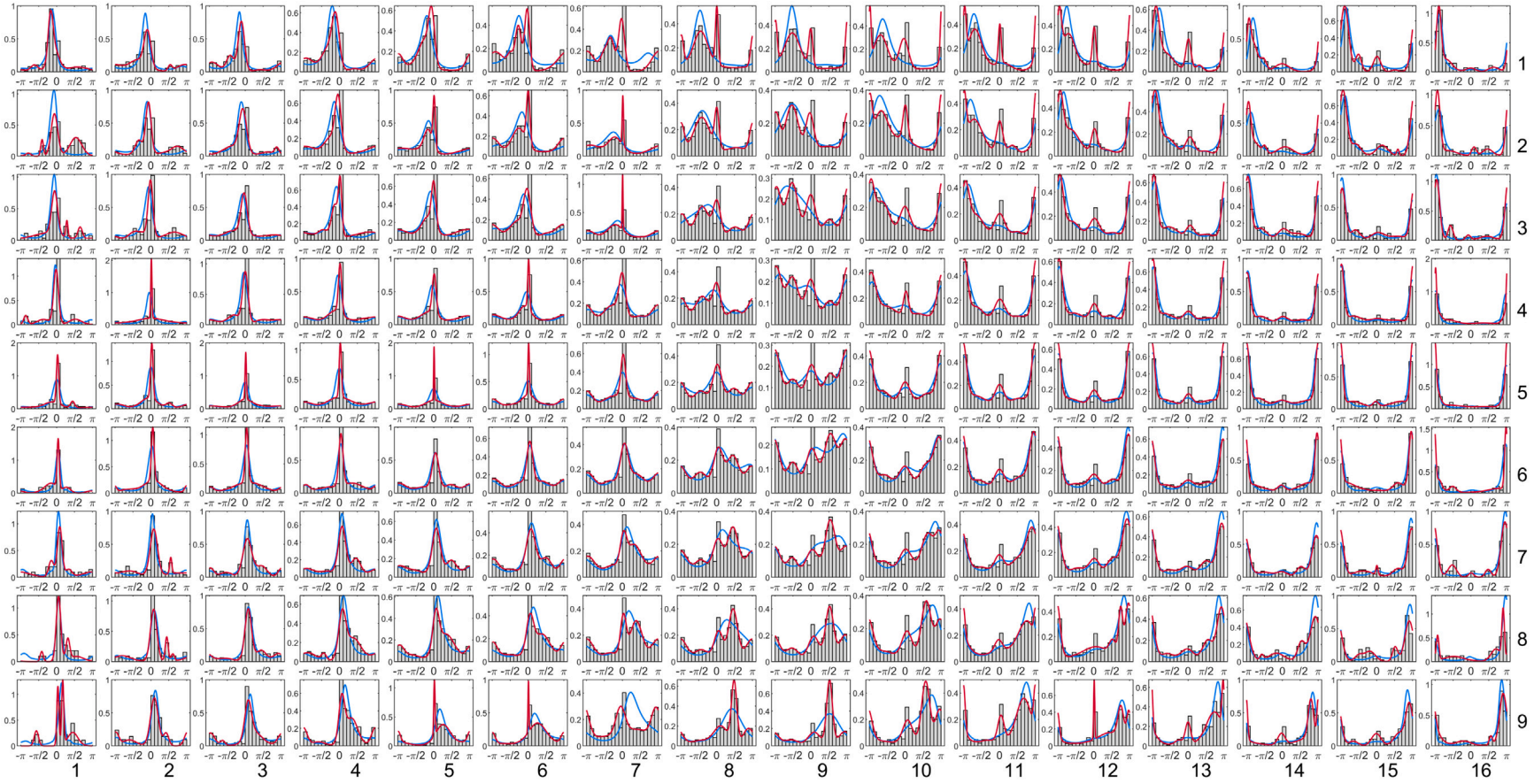


Fig. 10. Orientation of saccades across selected launch sites. Each site's observed orientation distribution is fitted with a 4-component VMM, denoted by a red line. The blue line represents the corresponding marginal distributions derived from the SF model.

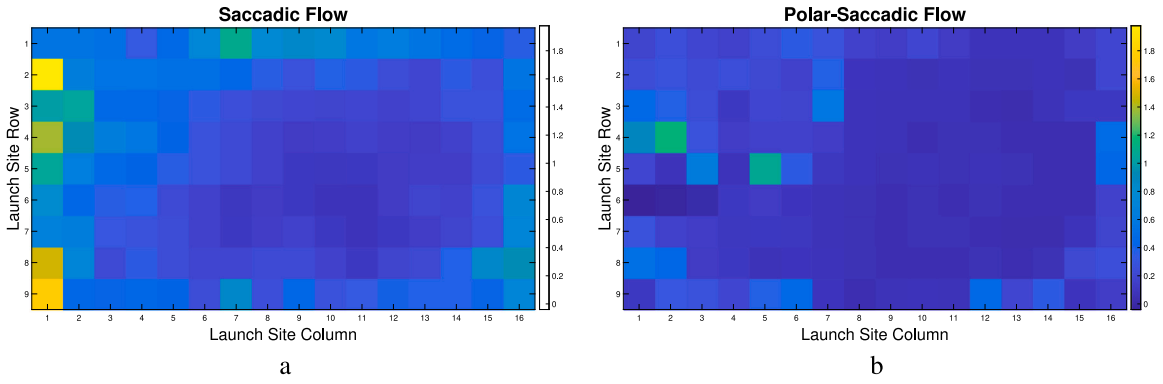


Fig. 11. Kullback–Leibler Divergence (KL-Divergence) between the empirical histogram and model histograms for the orientation component per launch site. (a) The KL-Divergence for the orientation component of the SF model. (b) The KL-Divergence for the VMM component of the proposed PSF model. In both cases, the divergence appears to decrease nearer to the horizontal center. However, the divergences in the VMM distributions are significantly smaller, as indicated by an independent two-sample t-test with $t(286) = 8.01$ and $p < 10^{-5}$). Averaged over the entire image space, the divergences are approximately $\sim 0.17 \pm 0.03$ for the VMM component and $\sim 0.41 \pm 0.10$ for the SF model.

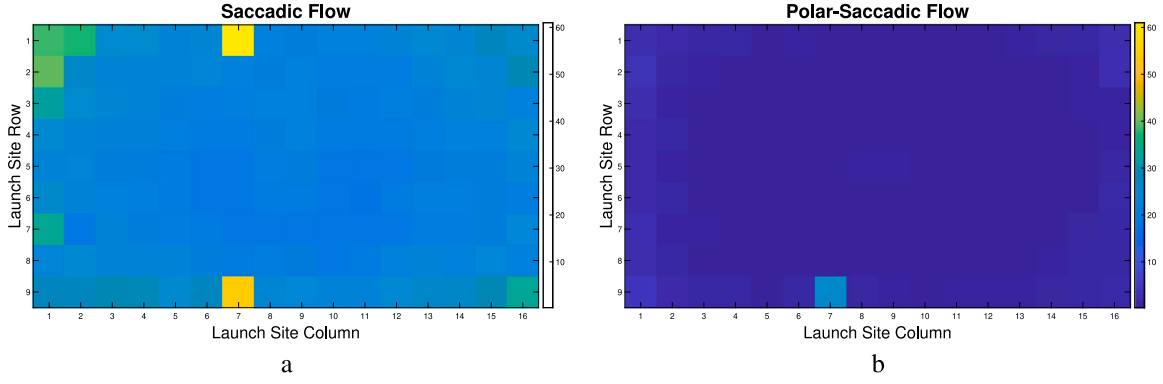


Fig. 12. Kullback–Leibler Divergence (KL-Divergence) between the empirical distribution and the competing amplitude models per launch site. (a) The divergence for the amplitude component of the SF model. (b) The divergence for the expo-normal component of the proposed PSF model. Averaged over the entire image space, the KL-divergence values were found to be $\sim 0.85 \pm 2.26$ for the expo-normal component and $\sim 23.28 \pm 5.54$ for the SF model. Statistical analysis using an independent two-sample t-test confirms the difference between the models to be highly significant with $t(286) = 44.96$, $p < 10^{-5}$.

achieved by the SF model (recall that both are relative to the predictive performance of uniform distribution). An independent two-sample t-test revealed this difference to be statistically significant ($t(198) = -190.39, p < 10^{-5}$). This substantial enhancement in llh score underscores the significant advantage of the PSF model in accurately capturing the nuances of saccadic amplitudes.

Similarly, the orientation component of our model showed improved performance. The mean llh score for the orientation predictions was 145, compared to 116 for the SF model. Again, an independent two-sample t-test revealed this difference to be statistically significant ($t(198) = -11.20, p < 10^{-5}$). This improvement in prediction performance highlights the effectiveness of the Von Mises Mixture (VMM) model in capturing the complex, multimodal nature of saccade orientations, and supports the advantage of the PSF model at predicting saccadic behaviors.

7. Summary and discussion

While the prior art of human fixation maps informed the community that viewers tend to focus more near the center than near the periphery, our present work attempts to scrutinize and formally model the space-dependent tendencies that underpin this phenomenon. Inspired by the SF model (Clarke et al., 2017), and analyzing saccades rather than fixations, we first proposed a critical new perceptive that breaks saccades to their orientation and magnitude components and showed that center-related biases are present in both components throughout the duration of viewing. Leveraging this representation we then derived a novel set of models that describe the tendencies and biases of saccadic programming much more accurately than the prior art.

The computational modeling proposed in our work is still local in nature, with a space-dependent description (via orientation and amplitude distributions) derived for each launch site. A global description clearly will provide added value and insights and is left for future work. Noteworthy, attempts for more global modeling were already made in the SF model, where the changes in the parameters of the local models across launch sites were fitted global polynomials along sections of space. While still not dense (as it relates to few discrete 1D cross sections), the extraction of such a description was made easier by the unimodal nature of the underlying 2D Gaussian model. The same task, however, becomes much more complicated (or even ill-defined) when the underlying local models are multimodal, as is the case with the VMM formulation. The richer local descriptors thus yield a more technically challenging investigation for a global representation, and we hope that concrete proposals can be deliberated in the community in subsequent research.

Following the prior art, we too leveraged a large and diverse dataset of eye movement sequences that emerge from a wide range of visual scenes. Indeed, using datasets such as CAT2000 facilitates the exploration and identification of intrinsic, image-independent biases in saccadic movements, the main goal of our present (as well as other's previous) work. Nevertheless, this does not indicate that specific contents do not have any effects on the pattern of saccades, and since it is divided into multiple categories, datasets such as CAT2000 allows to explore these questions too. While an exhaustive study of this issue is reserved for future work, here we conducted an initial exploratory analysis for content-dependent effects using selected categories from

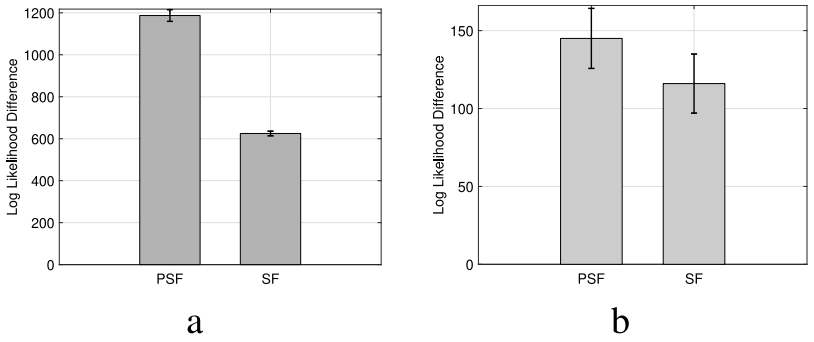


Fig. 13. Differential log-likelihood (llh) scores, calculated as the difference between the llh of each model and that of a uniform distribution. The graphs illustrate a marked improvement in llh scores for our PSF over the SF model, exhibiting far superior accuracy in capturing amplitude variations with the expo-normal distribution (a) and more precise prediction of saccade orientations using the Von Mises Mixture model (b). Error bars are $\pm 1 STD$.

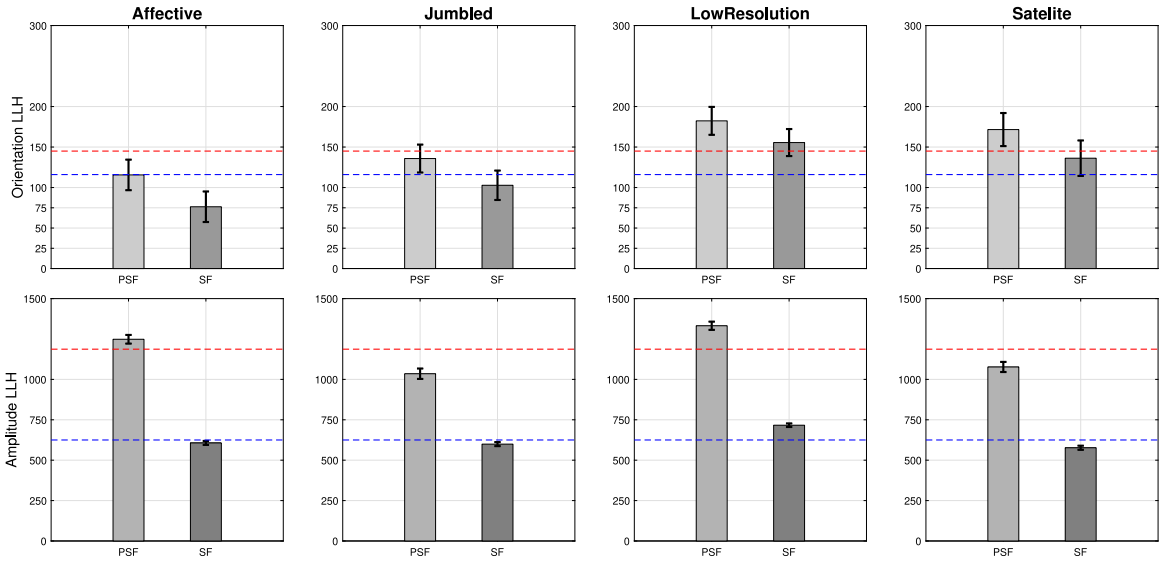


Fig. 14. Differential llh scores for orientation (top row) and amplitude (bottom row), relative to uniform distributions, across four CAT2000 categories. Each panel shows the scores for both the PSF and SF models, benchmarked against their respective baseline differential llh values (dashed lines — PSF in red, SF in blue). These results highlight the nuanced impacts of specific image content on the predictive capabilities of each model, emphasizing both inter-model and category-specific variations.

the CAT2000 dataset. Specifically, using the same llh-based analysis as described above, we examined the predictive power of our proposed model for specific classes, while hypothesizing that it would be different (either better or worse) than the pooled performance from Section 6.3.

For this pilot test, we selected the "Low Resolution", "Affective", "Jumbled", and "Satellite" categories for their diverse visual properties, and assessed the predictive power of both the PSF and SF models for each, compared to the baseline improvement of each as presented in Section 6.3. This initial analysis reveals non-negligible variations in predictive performance both within and between the models when tested on these different categories. For instance, while both models perform better (than their respective baselines) when tested on the "Low Resolution" class, they exhibit lower llh scores on the "Jumbled" class, and showed mixed results on the "Affective" and "Satellite" classes. The "Affective" class is particular intruiging since unlike the rest, it triggered different qualitative behavior of the two models compared to their pooled baselines. While the SF model performed worse in both orientation and amplitude, the PSF model deteriorated only in orientation but improved on amplitude. These findings highlight the nuanced impact that specific image contents can have on saccadic behavior, illustrating the complexity of content-dependent effects. It also emphasizes the difference between the models. The comparative results are summarized visually in Fig. 14 and underscore the need for dedicated future work on content-dependent effects in saccadic flow, potentially through class-based modeling approaches.

Despite the challenges ahead, in particular in understanding if and how saccade orientation and amplitude interact, better local modeling of saccadic biases, and the new polar representation advocated in this paper, already provide important benefits. First and foremost, with much smaller errors compared to the measured data, the PSF clearly allows a better modeling baseline, thereby facilitating the upcoming steps in the quest to model and predict where people tend to look next. Further theoretical research of our proposed models may be linked to mechanisms, both kinematic and perceptual, leading to a deeper understanding of saccadic programming in general. With better models of the inherent biases, saccade data can also be separated better into their intrinsic vs. stimulus-dependent ingredients, facilitating a better understanding of the latter too. Finally, these models, with their predictive power, may improve eye-tracking technology, feeding back better data to eye movement research.

CRediT authorship contribution statement

Rotem Mairon: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Ohad Ben-Shahar:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing.

Data availability

Data used in this research was obtained from publically available open repositories shared by others.

References

Bahill, A., Adler, D., & Stark, L. (1975). Most naturally occurring human saccades have magnitudes of 15 degrees or less.. *Investigative Ophthalmology and Visual Science*, *14*(6), 468–469.

Bair, W., & O'Keefe, L. P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Visual Neuroscience*, *15*(4), 779–786.

Borji, A., & Itti, L. (2015). Cat2000: A large scale fixation dataset for boosting saliency research. arXiv preprint arXiv:1505.03581.

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6–6.

Clarke, A. D., Green, P., Chantler, M. J., & Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision*, *16*(7), 4–4.

Clarke, A. D., Stainer, M. J., Tatler, B. W., & Hunt, A. R. (2017). The saccadic flow baseline: Accounting for image-independent biases in fixation behavior. *Journal of Vision*, *17*(11), 12–12.

Cooper, E. A., & Farid, H. (2020). A toolbox for the radial and angular marginalization of bivariate normal distributions. arXiv preprint arXiv:2005.09696.

Engbert, R., Trukenbrod, H. A., Barthelmé, S., & Wichmann, F. A. (2015). Spatial statistics and attentional dynamics in scene viewing. *Journal of Vision*, *15*(1), 14–14.

Gajewski, D. A., Pearson, A. M., Mack, M. L., Bartlett, F. N., & Henderson, J. M. (2005). Human gaze control in real world search. In *Attention and performance in computational vision: second international workshop, WAPCV 2004, prague, czech Republic, May 15, 2004, revised selected papers 2* (pp. 83–99). Springer.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259.

Jiang, M., Boix, X., Roig, G., Xu, J., Van Gool, L., & Zhao, Q. (2016). Learning to predict sequences of human visual fixations. *IEEE Transactions on Neural Networks and Learning Systems*, *27*(6), 1241–1252.

Judd, T., Durand, F., & Torralba, A. (2012). *A benchmark of computational models of saliency to predict human fixations*: *Tech. rep.*, MIT CSAIL Technical Reports.

Kümmerer, M., Bethge, M., & Wallis, T. S. (2022). DeepGaze III: Modeling free-viewing human scanpaths with deep learning. *Journal of Vision*, *22*(5), 7–7.

Le Meur, O., & Liu, Z. (2015). Saccadic model of eye movements for free-viewing condition. *Vision Research*, *116*, 152–164.

Lee, S. P., Badler, J. B., & Badler, N. I. (2002). Eyes alive. *ACM Transactions on Graphics*, *21*(3), 637–644.

Lee, T. S., & Yu, S. (1999). An information-theoretic framework for understanding saccadic eye movements. *Advances in Neural Information Processing Systems*, *12*.

Liu, H., Xu, D., Huang, Q., Li, W., Xu, M., & Lin, S. (2013). Semantically-based human scanpath estimation with HMMs. In *Proceedings of the IEEE international conference on computer vision* (pp. 3232–3239). IEEE.

Moeller, G. U., Kayser, C., Knecht, F., & König, P. (2004). Interactions between eye movement systems in cats and humans. *Experimental Brain Research*, *157*(2), 215–224.

Ngo, T., & Manjunath, B. (2017). Saccade gaze prediction using a recurrent neural network. In *2017 IEEE international conference on image processing* (pp. 3435–3439). IEEE.

Nuthmann, A. (2017). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, *24*(2), 370–392.

Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, *41*(25), 3587–3596.

Rothkegel, L. O., Trukenbrod, H. A., Schütt, H. H., Wichmann, F. A., & Engbert, R. (2017). Temporal evolution of the central fixation bias in scene viewing. *Journal of Vision*, *17*(13), 3–3.

Schütt, H. H., Rothkegel, L. O., Trukenbrod, H. A., Reich, S., Wichmann, F. A., & Engbert, R. (2017). Likelihood-based parameter estimation and comparison of dynamical cognitive models. *Psychological Review*, *124*(4), 505.

Sumner, P. (2011). Determinants of saccade latency. *The Oxford Handbook of Eye Movements*, 413–424.

Sun, W., Chen, Z., & Wu, F. (2019). Visual scanpath prediction using IOR-roi recurrent mixture density network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(6), 2101–2118.

Tatler, B. W., Baddeley, R. J., & Vincent, B. T. (2006). The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, *46*(12), 1857–1862.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407.

Wang, W., Chen, C., Wang, Y., Jiang, T., Fang, F., & Yao, Y. (2011). Simulating human saccadic scanpaths on natural images. In *2011 IEEE conference on computer vision and pattern recognition* (pp. 441–448). IEEE.