Saccadic Biases

Alasdair D. F. Clarke, Matthew J. Stainer

July 14, 2015

Abstract

More bias modelling! Cause who can be bothered running actual experiments?

Much effort has been made to attempting to explain eye guidance during natural scene viewing [Tatler, 2009 VIS RES special issue]. However, underlying fixation placement appears to be a set of consistent biases in eye movement behaviour (e.g. see Clarke and Tatler, 2014). We present a model that parametrically accounts for where saccades are directed dependent on where in the bounds of a scene an observer is currently fixating. We find... [some very interesting stuff]. Given that much of our understanding in eye guidance is derived from how people look at pictures on a computer screen, it is important that we use these biases to form a frame upon which to build more sophisticated models of eye guidance that can account for the allocation of gaze above oculomotor behaviours that are independent of the image.

1 Introduction

Improve on last year's [Clarke and Tatler, 2014] effort. More sophisticated biases. And some examples of how to use biases for improved data analysis.

The human fovea provides a small window of high acuity vision to the world, and as such the locations that we select to view in the world can tell us about how we seek the information necessary to complete the task we are currently undertaking. Current understanding of eye guidance would suggest that fixation locations are selected based on a combination of low-level factors (such as visual salience [**?**] or orientation information [**?**]) and high-level factors [**???**]. However, there are also strong observable biases in eye movement including directional and amplitudinal [XXis that a word?XX] biases in saccades [**???**], and a strong tendency to fixate near to the centre of images [**??**]. Importantly, these biases are independent of the viewed content. If we are to gain a complete understanding of the factors that govern eye movement, we must therefore build models of eye guidance on the framework of these underlying biases.

1.1 The central bias

There is a strong tendency for people to look close to the centre of pictures [**????**] and movies [**?**] presented on computer screens. There have been a number of suggestions for why this might be. One possibility for this effect is that the muscles of the eye show a preference for the 'straight ahead' position, re-centring in the orbit of the eye socket for most comfortable contraction of the ocular muscles (an *orbital reserve* [**?**]). As most scene viewing experimental set-ups stabilise the head to increase the accuracy of the eye tracking, and most scenes are presented in the centre of computer displays, such a re-centring mechanism would mean that the centre of images would indeed be preferentially selected. However, when scenes are scrambled into four quadrants, fixations are located near to the centre of each quadrant, rather than the display centre, suggesting that the central tendency is responsive to the viewed content [**?**] rather than the frames of the computer monitor.

Another possibility for the central fixation bias is that it represents a *photographer bias* as photographers tend to frame their shots to include the most important content in the centre of the scene. However, when **?** presented scenes where the image features were biased towards the edge of the scene, the central fixation bias persisted. The final possibility is that as a consequence of repeated exposure to photographer bias, the centre of scenes is simply where people are *trained* to look at images [**?**]. Such learning of spatial probabilities of targets can explain why, for example, people tend to look around the horizon when searching for people in natural scenes [**???**]. Expecting to find interesting content in the centre of scenes might be a consequence of this hypothesis typically being correct.

Clarke and Tatler [2014] revealed that the characteristics of the central bias is remarkably consistent across a series of eye movement databases.... [obviously you are in a better place than me to talk about this paper!]

1.2 Behavioural biases in saccades

Further to the observed bias towards the centre of images, it has been revealed that there are underlying biases in the characteristics of eye movement (in terms of the directions and amplitudes of saccades). It has been noted by several researchers that when viewing scenes, there is a higher proportion of eye movements in horizontal directions than vertical or oblique movements (Brandt, 1945; Crundall & Underwood, 1998; Gilchrist & Harvey, 2006; Foulsham, Kingstone & Underwood, 2008; Tatler & Vincent, 2008). There are a number of possibilities as to why this tendency exists (as discussed in Foulsham, Kingstone, & Underwood, 2008). Firstly, there may be a muscular or neural dominance making oculomotor movements in the horizontal directions more likely. Secondly, the characteristics of photographic images may mean that content tends to be arranged horizontally by the photographer. In such

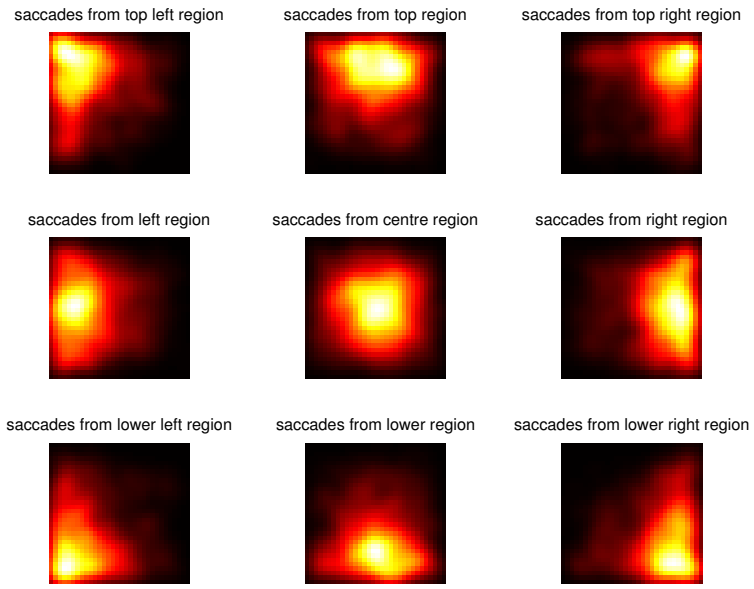

Figure 1: Empirical example of saccadic flow from Clarke et al. [2009]. Fixations come from a visual search experiment for a target in noise.

situations, horizontal saccades may be the most efficient way to inspect scenes. Thirdly, using horizontal saccades in scene viewing might be a learned strategy. Observers may learn the natural characteristics of scenes based on previous experience, and therefore demonstrate an increased likelihood of moving in the horizontal direction. A final alternative explanation is that this tendency is a consequence of the aspect ratio of visual displays, which normally allow for larger amplitude saccades in the horizontal than vertical directions (Wartburg et al., 2007).

Foulsham and colleagues have presented two interesting exceptions to the horizontal direction bias. Foulsham, Kingstone and Underwood (2008) found that when the orientation of an image is rotated, the distribution of saccade directions follows the orientation of the scene. A second exception comes from using circular apertures (Foulsham & Kingstone, 2010). When a scene is presented in a circular aperture, the tendency to make horizontal saccades disappears, being replaced by a tendency to make vertical saccades relative to the image orientation. However, when using fractal images (where images do not have an obvious orientation), observers tend make horizontal saccades, regardless of the angle that the image is presented.

1.3 The present study

The aim of the present study is to characterise the underlying biases of eye movement with which to understand fixation selection in natural scenes. We

	Observers	Images	Task	Display dura
Clarke, Coco, and Keller [2013]	24	100	object naming	5000 ms
Yun, Peng, Samaras, Zelinsky, and Berg [2013] - SUN	8	104	image description	5000 ms
Tatler, Baddeley, and Gilchrist [2005]	14	48	memory	variable
Einhäuser, Spain, and Perona [2008]	8	93	object naming	3000 ms
Tatler [2007] - free	22	120	free viewing	5000 ms
Judd, Ehinger, Durand, and Torralba [2009]	15	1003	free viewing	3000 ms
Yun et al. [2013] - PASCAL	3	1000	free viewing	3000 ms
Ehinger et al. [2009]	14	912	visual search	variable
Tatler [2007] - search	30	120	visual search	5000 ms
Asher, Tolhurst, Troscianko, and Gilchrist [2013]	25	120	visual search	variable

Table 1: Summary of the 10 datasets used throughout this study.

	Eye tracker	Viewing distance	Screen size	Image size	Viewing angle	Chin / head rest
Clarke et al. [2013]	EyeLink II	50 cm	21"	800 x 600	31 x 25°	no
Yun et al. [2013] - SUN	EyeLink 1000	?	?	?	?	?
Tatler et al. [2005]	EyeLink I	60 cm	17"	800 x 600	30 x 22°	no
Einhäuser et al. [2008]	EyeLink 1000	80 cm	20"	1024 x 768	29 x 22°	yes
Tatler [2007] - free	EyeLink II	60 cm	21"	1600 x 1200	40 x 30°	no
Judd et al. [2009]	?	2 feet	19"	1024 x 768*	?	yes
Yun et al. [2013] - PASCAL	EyeLink 1000	?	?	?	?	?
Ehinger et al. [2009]	ISCAN RK-464	75 cm	21"	800 x 600	23.5 x 17.7°	yes
Tatler [2007] - search	EyeLink II	60 cm	21"	1600 x 1200	40 x 30°	no
Asher et al. [2013]	EyeLink 1000	55 cm	?	1024 x 1280	37.6 x 30.5°	yes

Table 2: Details of the experimental setups in each of the 10 datasets analysed in the present study. We provide only information reported in the original articles. Question marks indicate information not reported in the original article. *For the Judd et al dataset images varied in pixel dimensions but the majority were at 1024 x 768.

extend the previous work by examining [XXstuffXX] and

2 Methods

2.1 Datasets

An overview of the datasets used is given in Tables 1 and 2. Hopefully we can add some more datasets to this. Note, we will note use the Ehinger, Hidalgo-Sotelo, Torralba, and Oliva [2009] dataset for training as previous results show that the central bias is due to contextual effects, this dataset behaves slightly differently from the others. However, we will include in in the testing stage.

2.2 Pre-processing

As with Clarke and Tatler [2014] we have normalised all fixations to the image frame, keeping the aspect ratio constant. ie, $(x, y) \in (-1. - 1) \times (-a, a)$ with typically $a = 0.75$. The initial fixations and saccades were not included in the analysis. Saccades with a start or end point falling outside of the image frame were also removed.

When fitting saccadic flow models, we *mirrored* the set of fixations, but adding in reflected copies of the data (reflected in the horizontal, vertical and both midlines). This has two advantages. (i) It is an easy way to make saccadic flow biases in the horizontal or vertical directions. This is similar to how the central bias was defined Clarke and Tatler [2014], but by a different mechanism (with the central bias, the model fitting procedure is much simpler and so we just enforced zero mean and 0s in the covariance matrix). (ii) It increases the amount of data available for fitting by a factor of four. This is important as (due to the central bias) there are relatively few saccades that originate from the corners of the images. By equating all corners, we can pool the data and obtain more stable estimates for the underlying distribution.

2.3 Training and Testing

As with Clarke and Tatler [2014] We will train models on all available data (saccades from all datasets grouped together) and report the resulting model as the baseline we recommended using. In order to check for over fitting, we will additionally evaluate the modelling procedure by training to one single dateset and evaluating over all the others. Our pooled dataset currently consists of over 600,000 saccades over seven datasets.

When evaluating over the full dataset we will also explore the effect of window-size on the results. The smaller the window, the more accurately we can characterise the behaviour for fixations close to the edge of the image. However, due to the central bias, there are far fewer fixations in these areas and so the distributions fits for these areas are increasingly noisy as the window is moved closer to the edge and with a smaller size.

3 Biases

We will model and discuss saccadic flow, coarse-to-fine, and left v right.

3.1 Saccadic Flow: Gaussian

Saccadic flow can be thought of as a generalisation of the central bias. Instead of computing the distribution of all saccadic endpoints in a dataset, we look at the distribution of saccade endpoints given the start points. So for a saccade

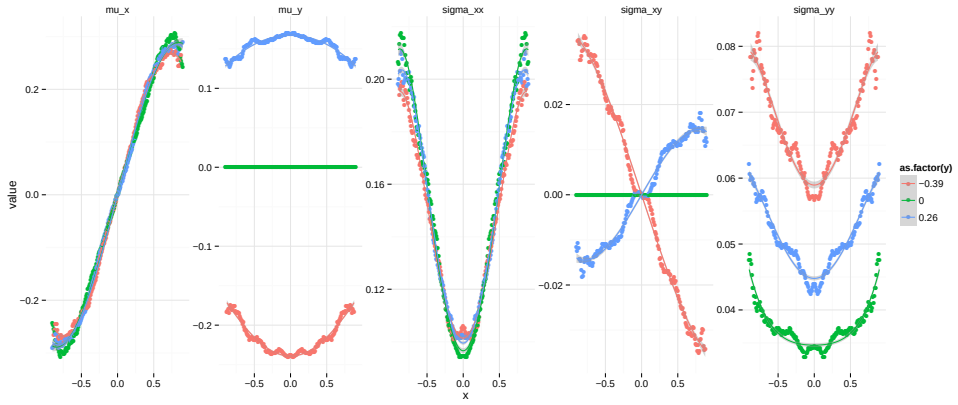

Figure 2: Flow:normal parameters over space. Note this is only a place holder for now.

from (x_0, y_0) to $(x_1, y1)$ we want to model $p(x_1, y_1|x_0, y0)$ This is illustrated in Figure 1.

3.1.1 Modelling

To characterise how the distribution of saccadic endpoints varies with the start point, we used a sliding window approach. All saccades that originated in a $n \times n$ window were taken and used to fit a multivariate Gaussian distribution. This window was then moved over the stimuli in steps of $s = 0.01$. Parameter sets estimated from windows containing less than 250 datapoints were removed. Multivariate polynomial regression was then used to fit 4-th order polynomials to each of the parameters. We experimented with varying the window size ($n \in \{0.05, 0.1, 0.2\}$). However, as this parameter was found to have a negligible result, we only report the results for $n = 0.1$.

3.1.2 Results

Figure 2 shows how the parameters for the multivariate Gaussian distribution vary over horizontal position for a selection of vertical positions. The regression coefficients given in Table 3 allow us to estimate the conditional probability of a saccade to $(x_1, y1)$ given the starting fixation (x_0, y_0).

How well does this model account for the fixations in our datasets? Figure 3 shows the deviance of the flow model expressed as a proportion of the deviance of the Clarke-Tatler central bias. For reference, we also show the results for re-fitting the central bias to each dataset. From this figure, we can see that the flow-normal model approximately halves the deviance.

As the flow:normal model is significantly more complex, requiring nine times as many parameters, it is important to test for robustness. Rather than carry-

parameter	equation
$\Omega_{x,x}$	$= 0.33 + 0.38x^2 - 0.29y^2 + 0.02x^4 + 0.22y^4$
$\Omega_{x,y}$	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$
$\Omega_{y,x}$	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$
$\Omega_{y,y}$	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$
α_{x^2}	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$
α_{y^2}	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$
ν	$= x + y + x^2 + y^2 + x^3 + y^3 + x^4 + y^4$

Table 3: Parameter model - clearly I still have to fill in all the coefficients!

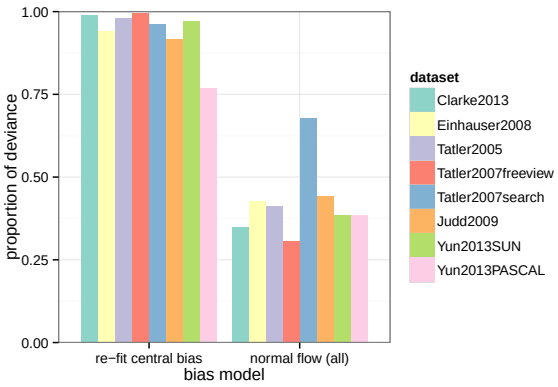

Figure 3: Flow:normal deviance results. We can see that re-fitting the central-bias to each specific dataset offers little improvement over using the Clarke-Tatler model, while the flow:normal model decreases the deviance by half.

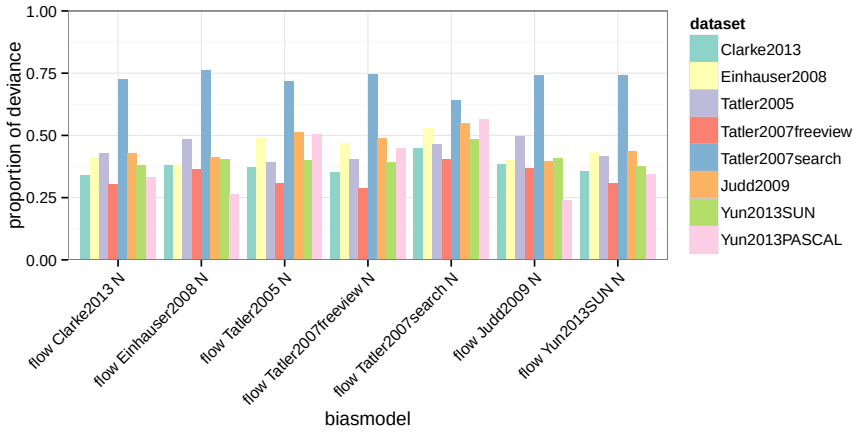

Figure 4: Flow:normal deviance results over datasets. In general, we can see that bias models trained on different datasets all explain around the same amount of variance in the datasets.

ing out k-fold cross validation, as we have a large amount of data, we will use a more stringent test, and evaluate how well the flow:normal model trained on just one dataset can explain the variance on the others. These resutls are shown in Figure 4.

3.1.3 Discussion

We put the Flow:normal model forward as a robust prior for image-content independent saccadic behaviour. This model can be thought of as as partner of the Clarke-Tater central bias, and we expect that in some cases, the simpler central bias will be more appropriate, while in others, the more complex flow model is a better choice. We have demonstrated that although this model requires more parameters, it generalises well from one dataset to another and is a far better baseline for modelling a scan-path than the central bias.

There are two main simplifications to our modelling work. First of all, we are using an unbounded distribution (ie, $(x, y) \in \mathbb{R}^2$) to model bounded data. While it is possible to deal with this issue, by either applying a transform $(-1, 1) \rightarrow \mathbb{R}$ (such as $z = log(\frac{x'}{1-x'})$, where $x' = \frac{x+1}{2}$), or fitting a truncated multivariate Gaussian, we decided that given the good performance of the model as is, it was not worth adding the additional complexities to our model at this time.

The second simplification is that we are treating the data as normal. From Figure 1 we can see that the data is clearly skewed, particularly in the corners. We will attempt to address these issues in the following section.

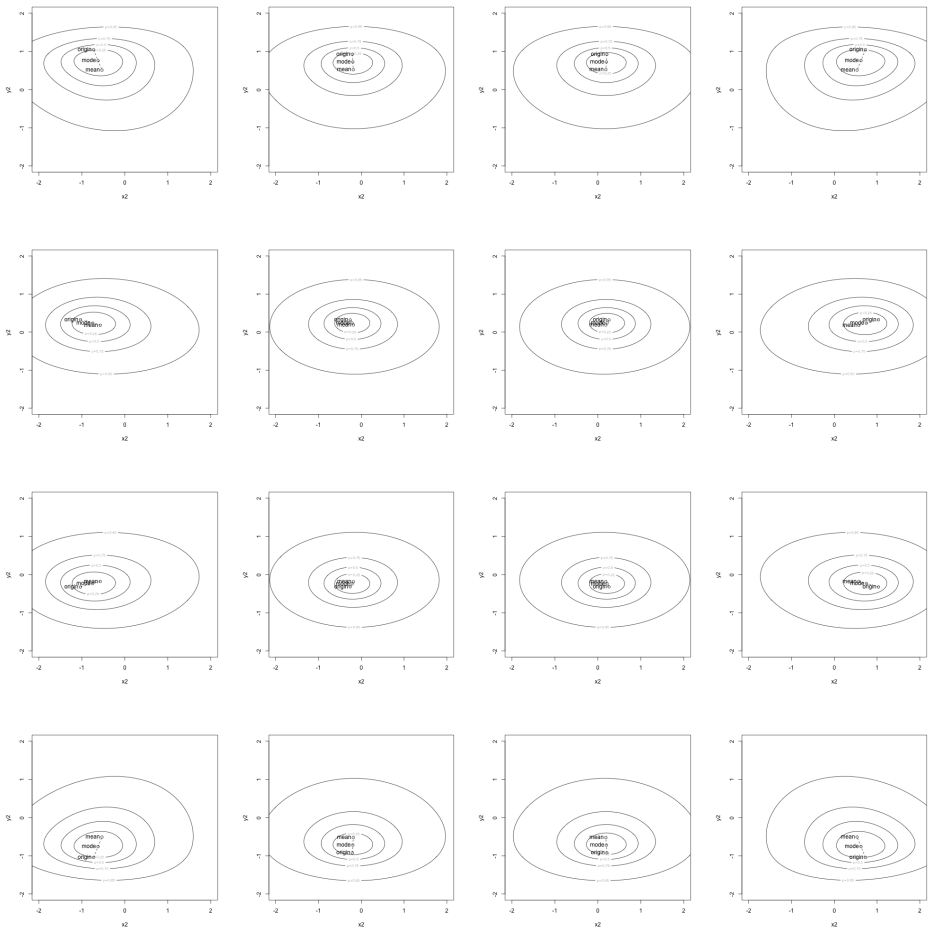

Figure 5: Multivariate skew-t distributions fitted to fixation location, by saccade start point.

3.2 Saccadic Flow: Skew Normal

We will model saccadic flow using multivariate skew-t distributions [Azzalini, 2015]. The multivariate skew-normal distribution [Azzalini and Dalla Valle, 1996] is given by:

$$\phi(z;\lambda) = 2\phi(z)\Phi(\lambda z) \tag{1}$$

for $z \in \mathbb{R}$. I think.

An example of the the distributions vary with saccadic start point is showing in Figure 5.

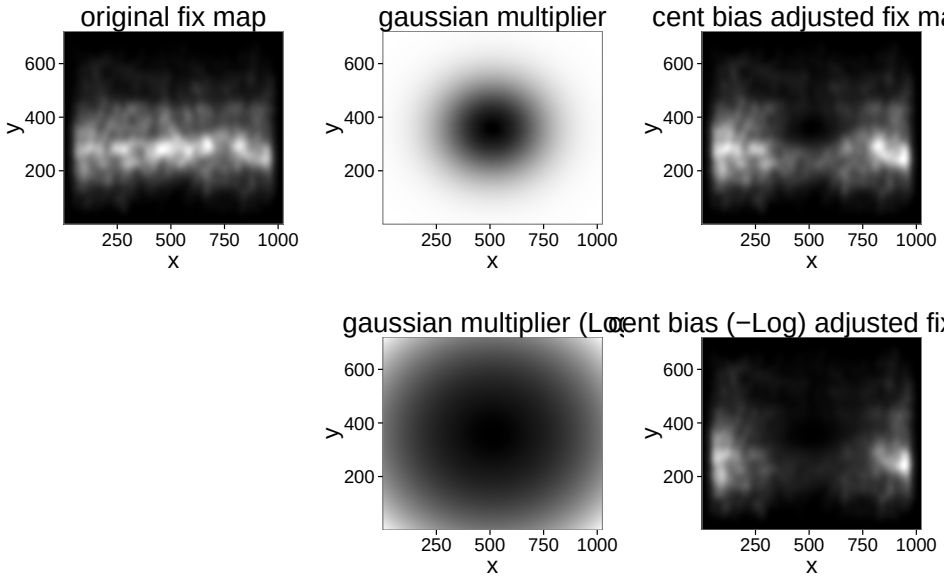

Figure 6: Traditional 'heat map' plots of fixations normalised by the central bias. This method allows us to characterise fixations that are less accountable for by image-independent central biases.

3.3 Coarse-to-fine

People make shorter saccades over time. Include $1/f$ dynamics?

3.4 Left v Right

Initially more fixations to the left half of the image [Nuthmann and Matthias, 2014].

4 Using Biases for Better Analysis

We will use the the central bias [Clarke and Tatler, 2014] and *saccadic flow* in some different contexts to see what biases can do for vision research. :p

4.1 Attentional Landscapes

HELLO THERE MY FRIEND

Or do we call them hotspot maps?

4.2 ROC Analysis

Example of using our models rather than shuffle approaches.

4.3 Flow and Coarse to fine

To what extent does saccadic flow account for coarse-to-fine dynamics

4.4 Inverse Yarbus

Do these biases allow us to improve inverse yarbus performance?

4.5 Salience

Does salience explain the less likely saccades?

5 Discussion

5.1 Scenes and natural viewing behaviour

That observers organise their viewing behaviour on computer screens around the reference frames provided by the bounds of scenes (see also **?**) causes problems for relating findings of eye guidance in scenes to eye guidance in natural behaviour, as the bounds of such reference frames are unclear in the real world. While it has been suggested that we tend to fixate near to the centre of our 'straight ahead' head position [FOULSHAM WALKING, CRISTINO AND BADDELEY?], there are no discrete edges as are typical in computer based scene viewing paradigms. If fixation locations are constrained by the bounds of the scene, this highlights the care we must take about the generalisations we make from findings in the lab to the real world (see [kingstonepaper 2010]).

Acknowledgements

Thanks to Adelchi Azzalini for advice on using the `sn` package for `R`. And mention grants.

References

Matthew F Asher, David J Tolhurst, Tom Troscianko, and Iain D Gilchrist. Regional effects of clutter on human target detection performance. *Journal of vision*, 13(5): 25:1–15, 2013.

A. Azzalini. *The R package `sn`: The skew-normal and skew-t distributions (version 1.2-2)*. Università di Padova, Italia, 2015. URL `http://azzalini.stat.unipd.it/SN`.

Adelchi Azzalini and Alessandra Dalla Valle. The multivariate skew-normal distribution. *Biometrika*, 83(4):715–726, 1996.

Alasdair D. F. Clarke, Moreno I. Coco, and Frank Keller. The impact of attentional, linguistic, and visual features during object naming. *Frontiers in psychology*, 4:927, 2013.

Alasdair DF Clarke and Benjamin W Tatler. Deriving an appropriate baseline for describing fixation behaviour. *Vision research*, 102:41–51, 2014.

Alasdair DF Clarke, Mike J Chantler, and Patrick R Green. Modeling visual search on a rough surface. *Journal of Vision*, 9(4), 2009.

Krista A Ehinger, Barbara Hidalgo-Sotelo, Antonio Torralba, and Aude Oliva. Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*, 17(6-7):945–978, 2009.

Wolfgang Einhäuser, Merrielle Spain, and Pietro Perona. Objects predict fixations better than early saliency. *Journal of Vision*, 8(14):18: 1–26, 2008.

Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba. Learning to predict where humans look. In *Computer Vision, 2009 IEEE 12th international conference on*, pages 2106–2113. IEEE, 2009.

Antje Nuthmann and Ellen Matthias. Time course of pseudoneglect in scene viewing. *Cortex*, 52:113–119, 2014.

Benjamin W Tatler. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14):4: 1–17, 2007.

Benjamin W Tatler, Roland J Baddeley, and Iain D Gilchrist. Visual correlates of fixation selection: Effects of scale and time. *Vision research*, 45(5):643–659, 2005.

Kiwon Yun, Yifan Peng, D. Samaras, G.J. Zelinsky, and T.L. Berg. Studying relationships between human gaze, description, and computer vision. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 739–746, 2013. doi: 10.1109/CVPR.2013.101.