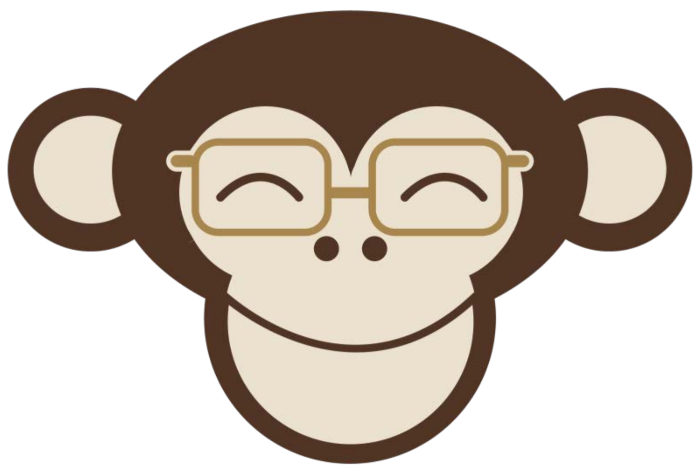
The PRIMATE Primer

Viewing behavior statistics from rhesus macaques during free viewing of natural scenes



By Seth Koenig

Buffalo Lab

Version 1.0-March 17, 2013

Image from http://skookummonkey.com/

Table of Contents

A. Introduction

A large number of studies in our lab as well as other utilize eye tracking systems to discern the location of visual attention in monkeys, other non-human primates, and humans. Despite the commonality and the extraordinarily large amount of eye tracking data, rarely ever is all this data combined in convenient lookup table. Even if we could collect this data from a large number of experiments and labs the variability in individual behavior may cloud our ability to properly analyze the data because behavior across different tasks, images, subjects/animals, and time can vary drastically. Furthermore, potentially one of the largest sources of outside variance in analysis of viewing behavior is the methods used to detect fixations and saccades. Variability in statistics as simple as fixation and saccade durations have been found to be unequivocal and often contradicting of each other across multiple methods of fixation and saccade detection.

Nonetheless, here I present the data the Buffalo lab has collected for monkeys who viewed natural scenes during a Scene Manipulation (SCM) task. Examples of the natural scenes used in SCM task can be found on the following page. Statistics present here were only calculated during the viewing of novel images and not during the viewing familiar or repeated images. Evidence from visual salience at fixation locations and Kullback-Leibler divergence indicate after approximately the first 5 fixations during viewing of familiar images monkeys look at less salient regions and at different regions of the image compared to during viewing of the then novel images. Therefore, there at exist at least some differences in the viewing behavior of familiar images compare to novel images; further investigation is necessary to determine if other statistics like fixation duration and saccade distances change between novel and familiar viewing behavior.

There are numerous factors influencing behavioral statistics, but the one factor that should not affect the calculated behavioral statistics is the method used to calculate such statistics. Currently, there does not exist a gold standard method for detecting fixations and saccades and thus virtually every experiment utilized different methods resulting in an additional outside source of variability affecting the calculated statistics. However, Elizabeth Buffalo and I have developed an amazing saccade and fixation detection algorithm utilizing k-means clustering in scan path state space which, to the best of our knowledge, is one of the most accurate algorithms with both a sensitivity and specificity likely exceeding 98%. The algorithm, called Cluster Fix, can typically detect small, short saccades unlikely to be detected by other algorithms due to being nearly indistinguishable from noise. To achieve both an increased sensitivity for smaller saccades and increased specificity for the transaction time between saccades and fixations, Cluster Fix globally and locally evaluates all saccades and fixations. For more details on Cluster Fix see König & Buffalo, 2013. A final caveat for calculating behavioral statistics is training may affect statistics; for example monkey trained to quickly find targets may make faster saccades with shorter fixations.

B. Example Images

Examples of images presented to the monkeys. These images range in variety and may include images of human or non-human faces, animals, urban scenes, outdoor or indoor scenes, as well as a variety of vantage points. All images were taken from Flickr.

C. Monkeys and Experimental Setup

All data was taken from 4 male rhesus macaques. At the beginning of experiments monkeys ranged in age from 7-8 years, and at the time of experiments monkeys had been trained to do visual tasks for 3-4 years with about 1 year of top-down search tasks in which monkeys had find a T among a set of L distractors before they would be rewarded (T's and L's task). Search tasks like the T's and L's may influence statistics like saccade rate since in theory monkeys would like to find the T as fast as possible so they can receive a reward. SCM was broken up into 2 parts: SCM and SCME. Presentation of all SCM sets occurred over period of less than 1 year from the presentation of SCM and SCME image sets. SCME sets were very similar to the SCM task except SCME image sets had images with qualitatively less prominent objects (i.e. simpler images). Both SCM and SCME sets are displayed in the previous section.

Scan paths were obtained at 200 Hz using an infrared camera (ISCAN) from the 4 monkeys while they freely viewed images. They sat head fixed in a dimly illuminate room 60 cm away from a 19" CRT monitor with a refresh rate of 120 Hz. Images of natural scenes were 600 pixels x 800 pixels large and subtended 25 x 33 degrees of visual angle (dva). Approximately 1 dva equated to a distance of 24 pixels. Experimental control software (CORTEX http://dally.nimh.nih.gov/) displayed images for 10 seconds each. Initial calibration of the infrared eye tracking system consisted of an initial 9 point calibration task. Drift was tracked throughout the experiment by presenting additional calibration trials between image viewing trials. I excluded eye tracking data more than 50 pixels (2 dva) outside of the image from any of the analysis.

All the data here is collected during SCM sets 6-9 and SCME sets 1-4. SCM sets 1-4 were excluded because the images were very boring containing images of the field station (monkey housing facility) lacking color and object variety or images of computer rendered home interiors. The monkeys did not particularly like viewing these boring images as indicated by their preference to not look at these images for the full 10 seconds. SCM image set 5 was excluded from analysis since there does not exist data from all 4 monkeys for this image set.

D. Known Sources of Variability

To determine what factors may influence viewing behavior statistics I conducted an ANOVA. By no means is this ANOVA complete. I do not further investigate the causal role of certain factors in influencing the statics except for saccade arc length because as you will see measures of saccade distance are the only factors in which the choice of SCM vs SCME was a significant source of variance. A correlation analysis revealed that saccade arc length was significantly correlated ($r = 0.681$, $p = 1.8$ e-5) with the amount of time spent looking outside of the image; the combination of the monkeys looking away as well as excluding viewing outside of the image from data analysis could have caused this correlation. The amount of time spent looking at a novel image is an indication of how interested monkeys are in those images.

In general, the only sources of variability were found to be individual variations among the monkeys and the images within a given image set. Therefore, I conclude there was a great deal of variability in how each monkey viewed an image, and each image led to different viewing dynamics most likely driven by bottom-up factors; simpler images containing less prominent objects (SCME) were not viewed differently than more complex images (SCM).

ANOVA in MATLAB does not always work well for a large number of factors if there is any dependence of one factor on another thus for example if the data are arranged in such a manner that image sets are dependent on SCM vs SCME, then MATLAB will throw you a NAN error for either the image set or SCM vs SCME because of a decrease in the degrees of freedom (lack of independence). Apparently if the data are arranged in a different manner (e.g. for salience at fixation locations) or the there are trends in the data that require both image sets and SCM vs. SCME to explain the variance then MATLAB does not throw you the NaN error. Therefore, some of the analyses I ran a 4 way ANOVA comparing the monkeys, image set, SCM vs SCME (or lesion), and fixation number as significant source of variance; however for some tests I had to run SCM vs. SCME and image set separately to determine if one, both, or none was a significant source of variance.

<u>All values listed are p-values.</u>

ANOVA All monkeys

Factor\ Parameter	Monkey Subject	Image Set	SCM vs SCME	Fixation Number
Salience at Fixation	9.3359e-036	5.0791e-050	0.6235	9.7357e-027
Salience during Fixations	1.4107e-008	0.1898	0.6894	0.0010
Salience contrast at fixations	0.0166	0.1237	0.4309	0.1466
Salience contrast during fixations	0.0012	0.0284	0.3399	0.0756
Image Intensity at fixations	2.3760e-010	5.5581e-034	0.5241	0.0184
Image Intensity during fixations	3.9077e-012	1.6239e-045	0.5732	4.6193e-004

Factor\ Parameter	Monkey Subject	SCM vs. SCME	Fixation Number		Image Set	SCM vs. SCME
Angle between fixations	1.5223e-004	0.1610	1.3003e-016		0.2776	0.1165
Distance between fixations	0	8.1127e-119	1.0642e-112		8.9772e-022	0.4416
Fixation Rate	2.1639e-051	0.0088	5.8801e-181		9.6237e-009	0.6079
Saccade distance (arc length)	3.9700e-063	4.7147e-066	2.1385e-004		0.6944	0.7361
Saccade amplitude	0	2.3065e-119	3.8031e-111		1.7446e-022	0.6025
Saccade angle leaving a fixation	9.9474e-010	0.4457	0.3958		0.9933	0.7113
Saccade angle entering a fixation	4.6681e-018	3.8712e-004	0.2348		0.3630	0.1397
Saccade Duration	2.3785e-011	0.0211	0.0128		0.8150	0.9229
Fixation duration	1.7109e-025	0.0191	4.5001e-101		1.0238e-004	0.6735

ANOVA TT-Lesion Effect

Parameter\Factor	Image Set	Pre vs Post Lesion & SCM vs SCME
Angle between fixations	0.1188	0.0232
Distance between fixations	9.1803e-024	0.4298
Fixation Duration	0.0063	0.9143
Fixation Rate	6.9962e-005 (3.1314e-006)	NaN (0.0024)
Saccade distance-arc length	3.5340e-012	0.7910
Saccade Amplitude	2.6974e-024	0.5340
Saccade angle leaving a fixation	0.9181	0.4819
Saccade angle entering a fixation	0.5673	0.8358
Saccade Duration	0.1458	0.8780

*() ran 1-way ANOVA since got NAN.

E. Behavioral Statistics

In this section we finally get to see the statistics you are looking for. Remember that the method used to calculate the statistics, the images that monkeys view, and the monkeys themselves are all significant sources of variability. I present both the “raw” statistics and “filtered” statistics in which I purposely exclude values from analysis. These will be clearly marked as RAW and FILTERED. Exclusion of certain ranges of values is solely arbitrary though I do my best to justify the exclusion of data based on percentiles, standard deviations, and when possible, realistic expectations of the distribution. For example—in a hypothetical situation—if across all 4 monkeys the saccade rate were to range from 0.01 Hz to 50 Hz I would say we don’t really care about fixations that occurred more than 600 ms apart as most fixations are less than 500 ms in duration and most saccades are less than 100 ms in duration, and saccade rate above 28.5 Hz are unrealistic since the minimum saccade duration is 10 ms and minimum fixation duration in 50 ms (1/35 ms = 28.5 Hz). Therefore the filtered range of data would be 1.6 Hz to 28.5 Hz.

As with a lot of behavioral data many distributions of data are not normal therefore standard tests that assume normality (e.g. t-test and z-test) are not applicable and typically I would use a ks-test. Furthermore, since normality is not the norm mean and standard deviation (STD) are not very useful so I will include mean, median, the 5^{th} and 95^{th} percentile, and inter-quartile range in all applicable statistics.

Please feel free to request statistics as well as make suggestions. Included with this document will be the original data. Unless otherwise specified the plots are of FILTERED data.

I. Fixation Duration

Definition: fixation duration is the length of time the monkey looks attentively at a region of the image during which the eye moves slowly over a small area.

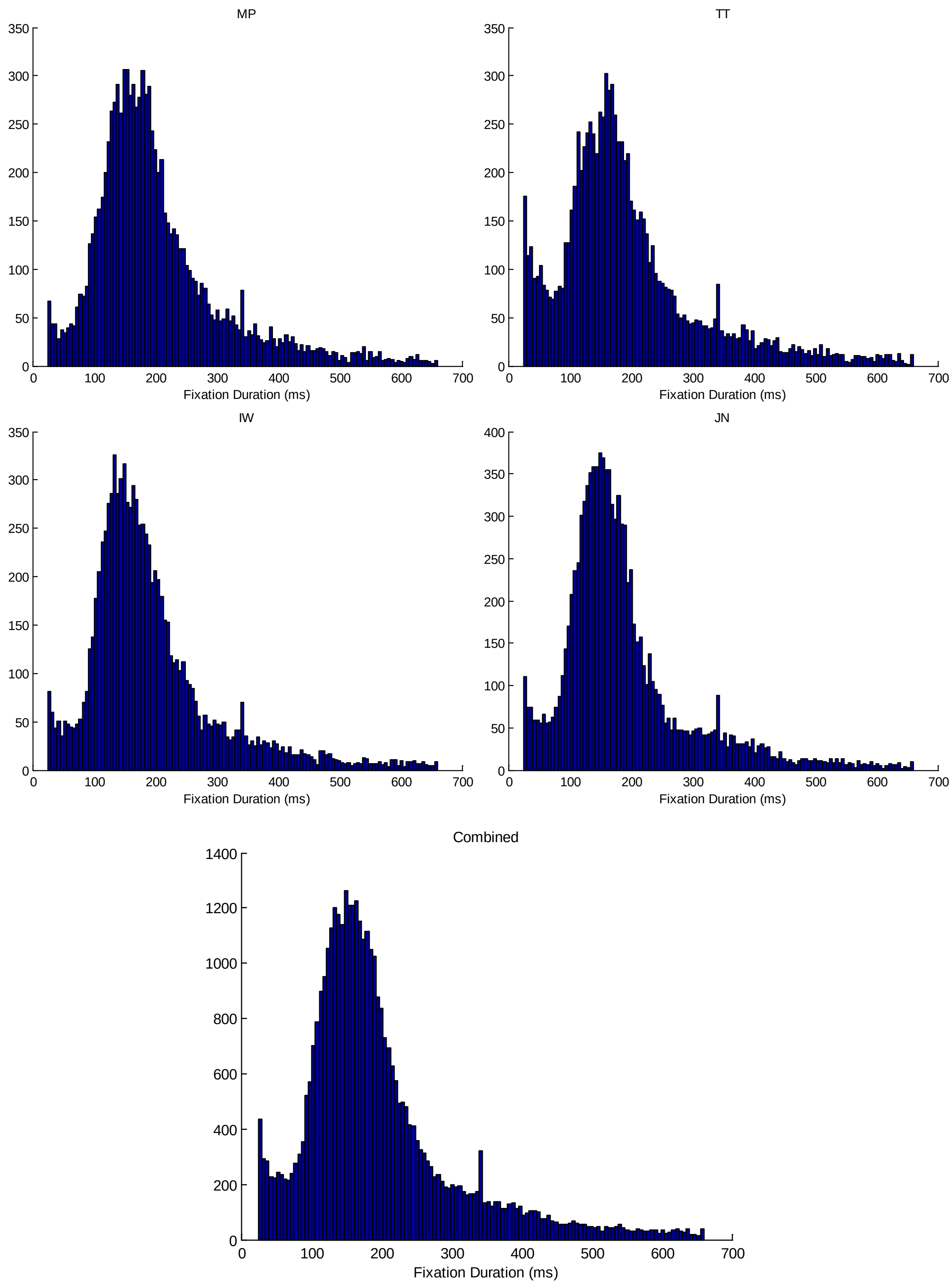
Units: time in ms

Filtered Range: 25 ms – 660 ms

Cutoff Justification: Fixations shorter than 25 ms in duration cannot be detected with Cluster Fix. Fixations much longer than 500 ms in duration are likely not fixations (i.e. noise or artifacts) or are multiple fixations with very small saccades (potentially microsaccades) in between that were not detected. A cutoff of 660 ms is the 97.5 percentile.

Raw	MP	TT	JN	IW	Combined
Mean	224.8	227.3	202.2	215.2	217.1
Median	180.0	175.0	165.0	175.0	175.0
STD	170.7	218.8	146.6	177.1	179.9
Minimum	25.0	25.0	25.0	25.0	25.0
Maximum	2465.0	6335.0	2955.0	4260.0	6335.0
5%	80.0	45.0	65.0	75.0	65.0
25%	140.0	125.0	130.0	130.0	130.0
75%	250.0	250.0	220.0	235.0	240.0
95%	520.0	575.0	450.0	495.0	510.0

Filtered	MP	TT	JN	IW	Combined
Mean	205.3	197.0	190.2	197.0	197.2
Median	180.0	170.0	165.0	170.0	170.0
STD	107.6	117.4	104.6	107.3	109.3
Minimum	25.0	25.0	25.0	25.0	25.0
Maximum	660.0	660.0	660.0	660.0	660.0
5%	80.0	45.0	65.0	75.0	60.0
25%	140.0	125.0	130.0	130.0	130.0
75%	240.0	235.0	215.0	230.0	230.0
95%	435.0	450.0	410.0	425.0	430.0


MP
TT
IW
JN
Combined
Fixation Duration (ms)
Fixation Duration (ms)
Fixation Duration (ms)
Fixation Duration (ms)
Fixation Duration (ms)

II. Saccade Duration

Definition: saccade duration is the length of time the monkey moves their eyes from one region of the image to another. Typically the eye is moving relatively quickly with high acceleration.

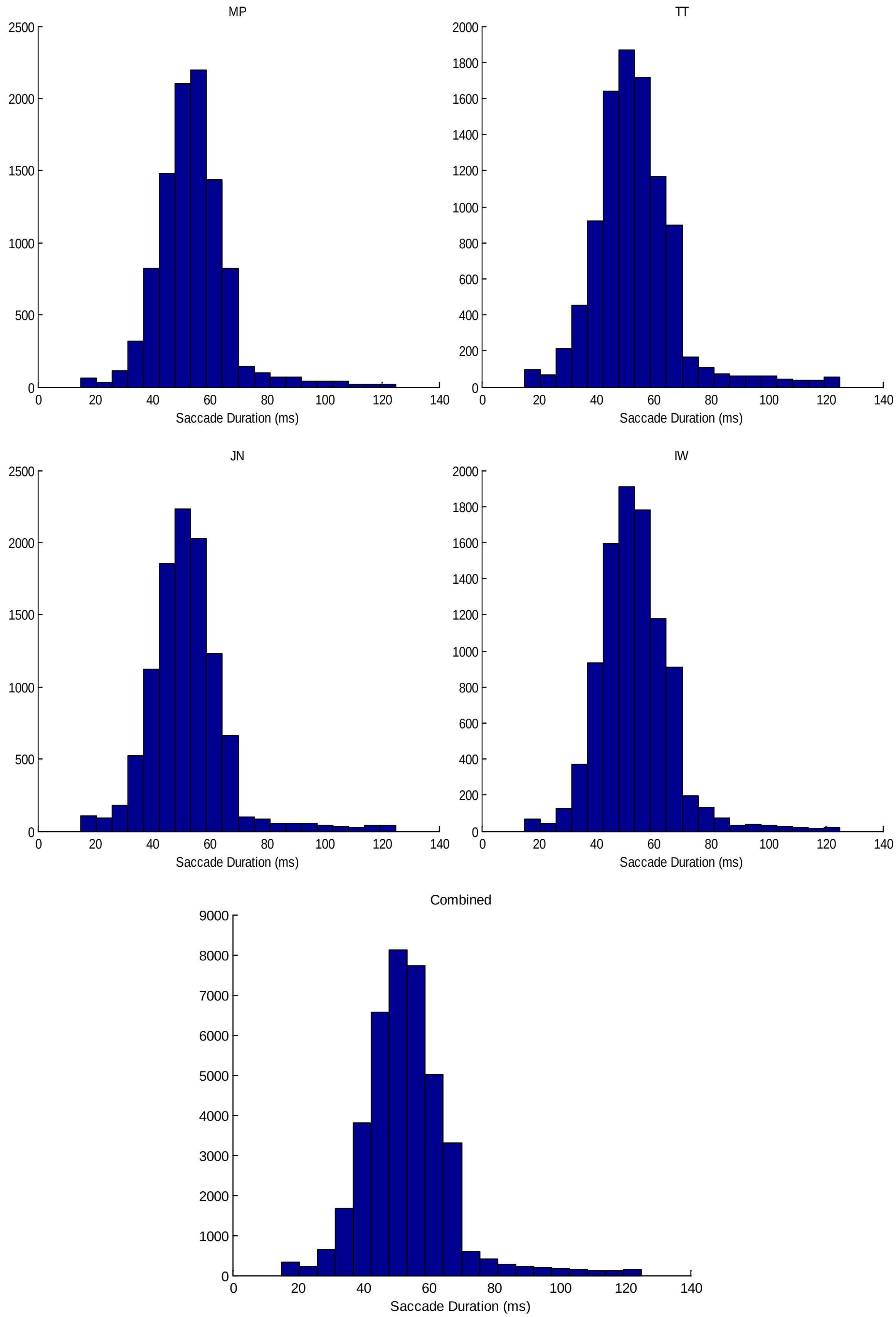
Units: time in ms

Filtered Range: 15 ms – 125 ms

Cutoff Justification: Saccades shorter than 10 ms in duration cannot be detected with Cluster Fix and Cluster Fix rounds to the nearest 5 ms; because of the rounding scheme, which I won't explain in detail, and the infrequent number of short saccades only quantized saccade durations of 15 ms were observed (I have double, double checked this code). Saccades longer than 125 ms in duration are likely due to the removal of data more than 2 dva (50 pixels) outside of the image or highly variable fixations with large velocities that look like saccades. A cutoff of 125 ms is the mean + 3 STDs.

Raw	MP	TT	JN	IW	Combined
Mean	54.7	56.4	54.6	54.2	54.9
Median	55.0	50.0	50.0	50.0	50.0
STD	19.6	25.9	25.3	18.6	22.7
Minimum	15.0	15.0	15.0	15.0	15.0
Maximum	690.0	435.0	470.0	355.0	690.0
5%	35.0	35.0	35.0	35.0	35.0
25%	45.0	45.0	45.0	45.0	45.0
75%	60.0	60.0	55.0	60.0	60.0
95%	75.0	95.0	85.0	75.0	80.0

Filtered	MP	TT	JN	IW	Combined
Mean	53.5	53.3	51.7	53.0	52.9
Median	55.0	50.0	50.0	50.0	50.0
STD	12.5	14.8	13.4	12.8	13.4
Minimum	15.0	15.0	15.0	15.0	15.0
Maximum	125.0	125.0	125.0	125.0	125.0
5%	35.0	35.0	35.0	35.0	35.0
25%	45.0	45.0	45.0	45.0	45.0
75%	60.0	60.0	55.0	60.0	60.0
95%	75.0	80.0	75.0	75.0	75.0


MP
TT
JN
IW
Combined
Saccade Duration (ms)
Saccade Duration (ms)
Saccade Duration (ms)
Saccade Duration (ms)
Saccade Duration (ms)

III. Fixation or Saccade Rate

Definition: fixation or saccade rate is the instantaneous rate at which fixations or saccades are made at though typically researchers simply call it the saccade rate. The instantaneous rate is calculated from the difference in consecutive mean fixation times.

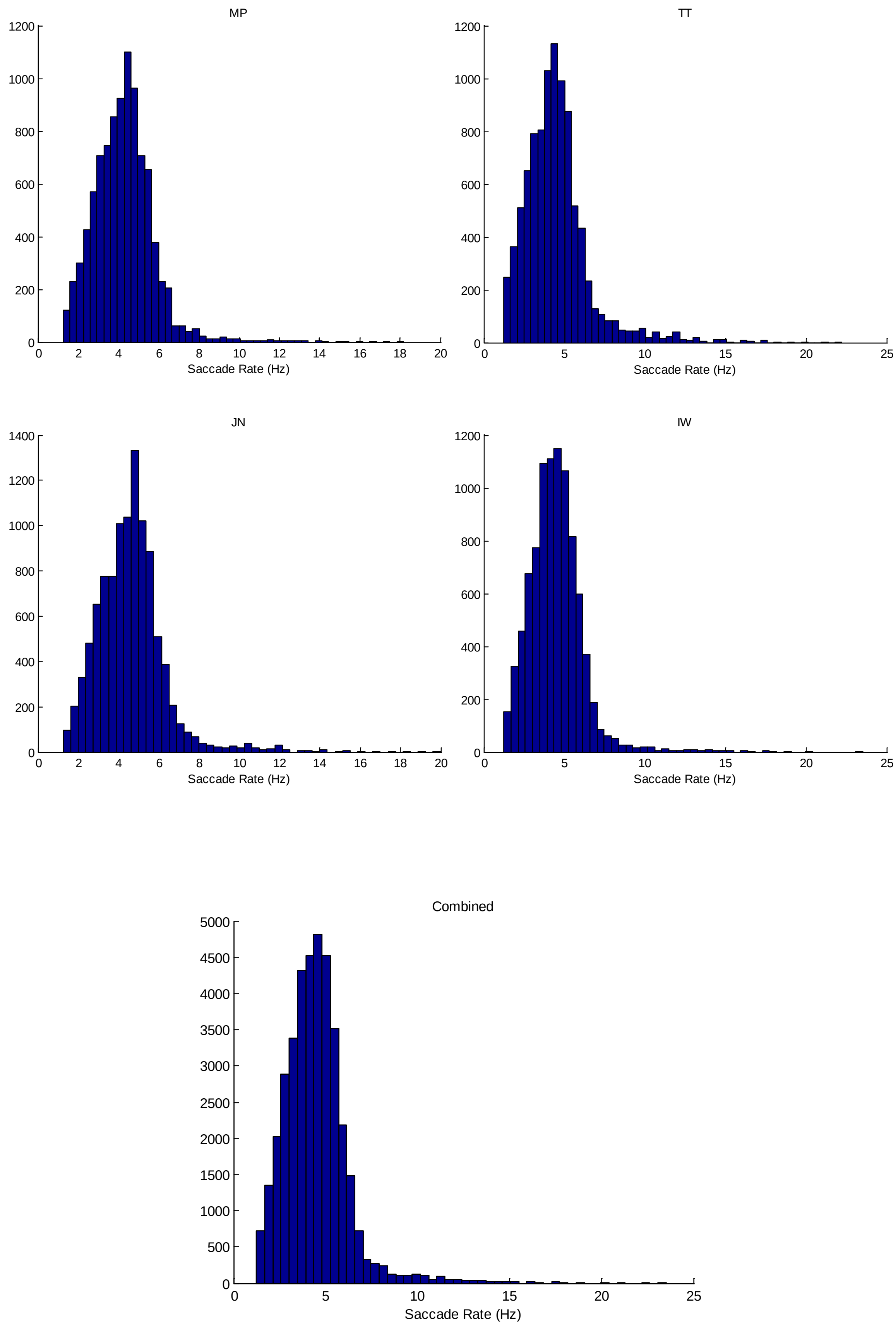
Units: rate/frequency in Hz

Filtered Range: 1.25 Hz – 25 Hz

Cutoff Justification: Fixations and saccades combined must be at least 35 ms in duration to be detected by Cluster Fix but we did not observe any fixation shorter than 25 ms or saccades shorter than 15 ms; therefore, we do not expect the rate to be higher than 25 Hz (i.e. the invers of 40 ms). Since fixations longer than 660 ms and saccades longer than 125 ms are likely artifacts of our processing of scan paths, we do not expect to see a combination of these two longer than785 ms so realistic rate should be higher than 1.25 Hz.

Raw	MP	TT	JN	IW	Combined
Mean	4.26	4.49	4.61	4.45	4.45
Median	4.21	4.30	4.55	4.35	4.35
STD	1.56	2.12	1.74	1.74	1.80
Minimum	0.65	0.27	0.61	0.39	0.27
Maximum	18.18	26.67	20.00	23.53	26.67
5%	1.99	1.81	2.25	2.06	2.01
25%	3.28	3.20	3.51	3.39	3.33
75%	5.06	5.26	5.41	5.26	5.26
95%	6.45	8.16	7.02	6.90	7.02

Filtered	MP	TT	JN	IW	Combined
Mean	4.29	4.53	4.62	4.47	4.48
Median	4.26	4.30	4.55	4.35	4.35
STD	1.54	2.08	1.73	1.72	1.78
Minimum	1.25	1.25	1.26	1.25	1.25
Maximum	18.18	22.22	20.00	23.53	23.53
5%	2.08	1.94	2.27	2.13	2.11
25%	3.31	3.23	3.51	3.42	3.36
75%	5.06	5.26	5.41	5.26	5.26
95%	6.56	8.16	7.02	6.90	7.10


MP
Saccade Rate (Hz)
TT
Saccade Rate (Hz)
JN
Saccade Rate (Hz)
IW
Saccade Rate (Hz)
Combined
Saccade Rate (Hz)

IV. Saccade Arc Length

Definition: saccade arc length is the arc length of the saccade. This is the distance over which the eye traverses during a saccade. This is not saccade amplitude (see next section).

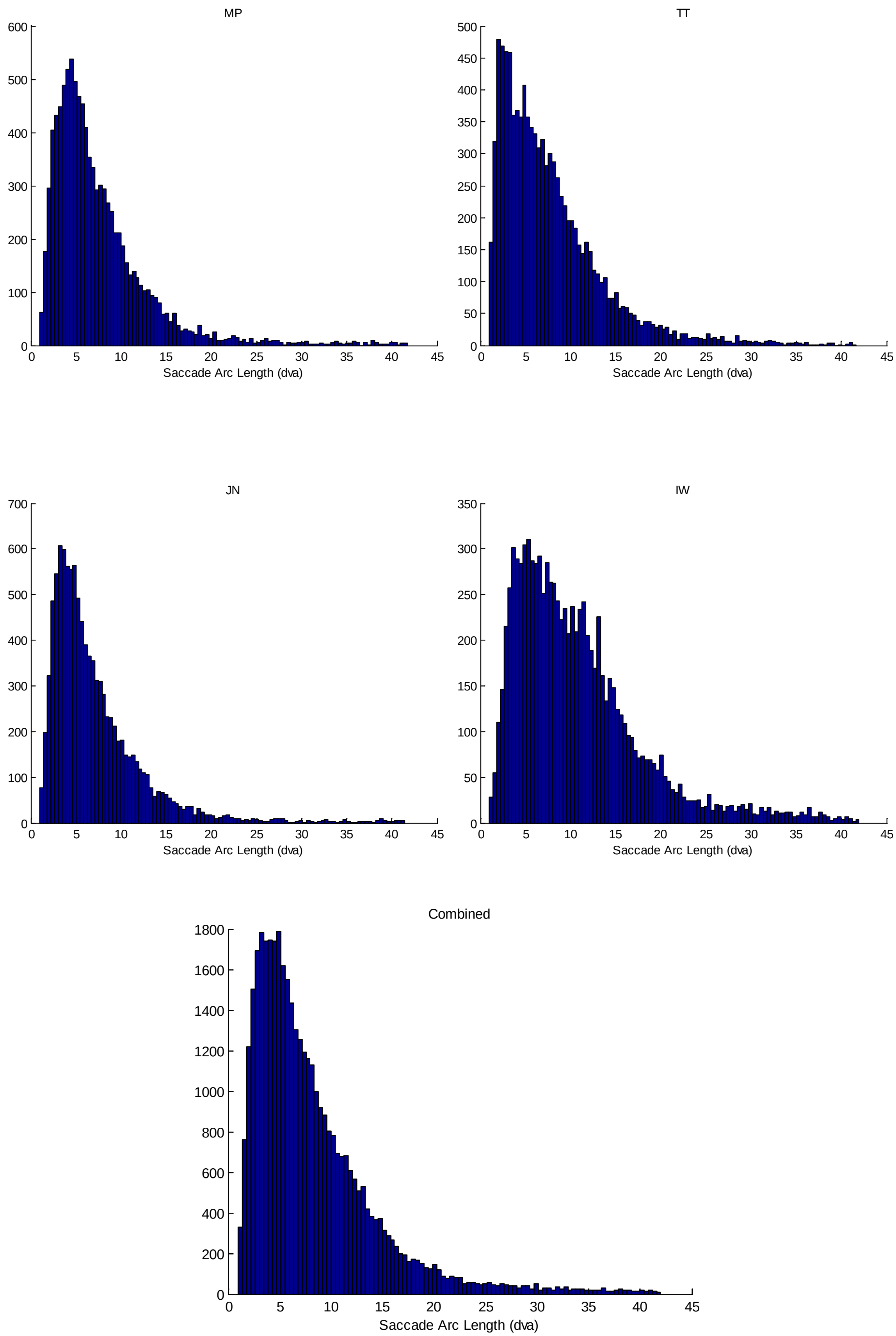
Units: distance in dva

Filtered Range: 1 – 42 dva

Cutoff Justification: Arc lengths less than 1 dva are more likely to be falsely detected saccades and saccades larger than 1000 pixels (42 dva) may be artifacts of the processing of scan paths or in some cases when saccades were not detected the arc length of several fixations.

Raw	MP	TT	JN	IW	Combined
Mean	8.37	7.96	10.40	11.58	9.57
Median	6.12	6.34	5.77	9.39	6.70
STD	20.49	6.67	25.89	9.30	17.80
Minimum	0.23	0.05	0.27	0.33	0.05
Maximum	1897.30	112.18	609.02	138.08	1897.30
5%	2.18	1.77	2.10	2.89	2.10
25%	4.01	3.53	3.70	5.66	4.09
75%	9.56	10.21	9.32	14.22	10.90
95%	19.70	19.58	21.94	29.19	22.72

Filtered	MP	TT	JN	IW	Combined
Mean	7.61	7.80	7.18	10.81	8.30
Median	6.08	6.36	5.66	9.25	6.63
STD	5.64	5.78	5.40	7.10	6.16
Minimum	1.03	1.01	1.02	1.02	1.01
Maximum	41.77	41.67	41.45	41.99	41.99
5%	2.21	1.83	2.14	2.90	2.15
25%	4.01	3.59	3.68	5.62	4.09
75%	9.41	10.19	8.95	13.86	10.64
95%	17.78	19.06	16.77	25.44	20.10


MP
TT
JN
IW
Combined
Saccade Arc Length (dva)
Saccade Arc Length (dva)
Saccade Arc Length (dva)
Saccade Arc Length (dva)
Saccade Arc Length (dva)

V. Saccade Amplitudes

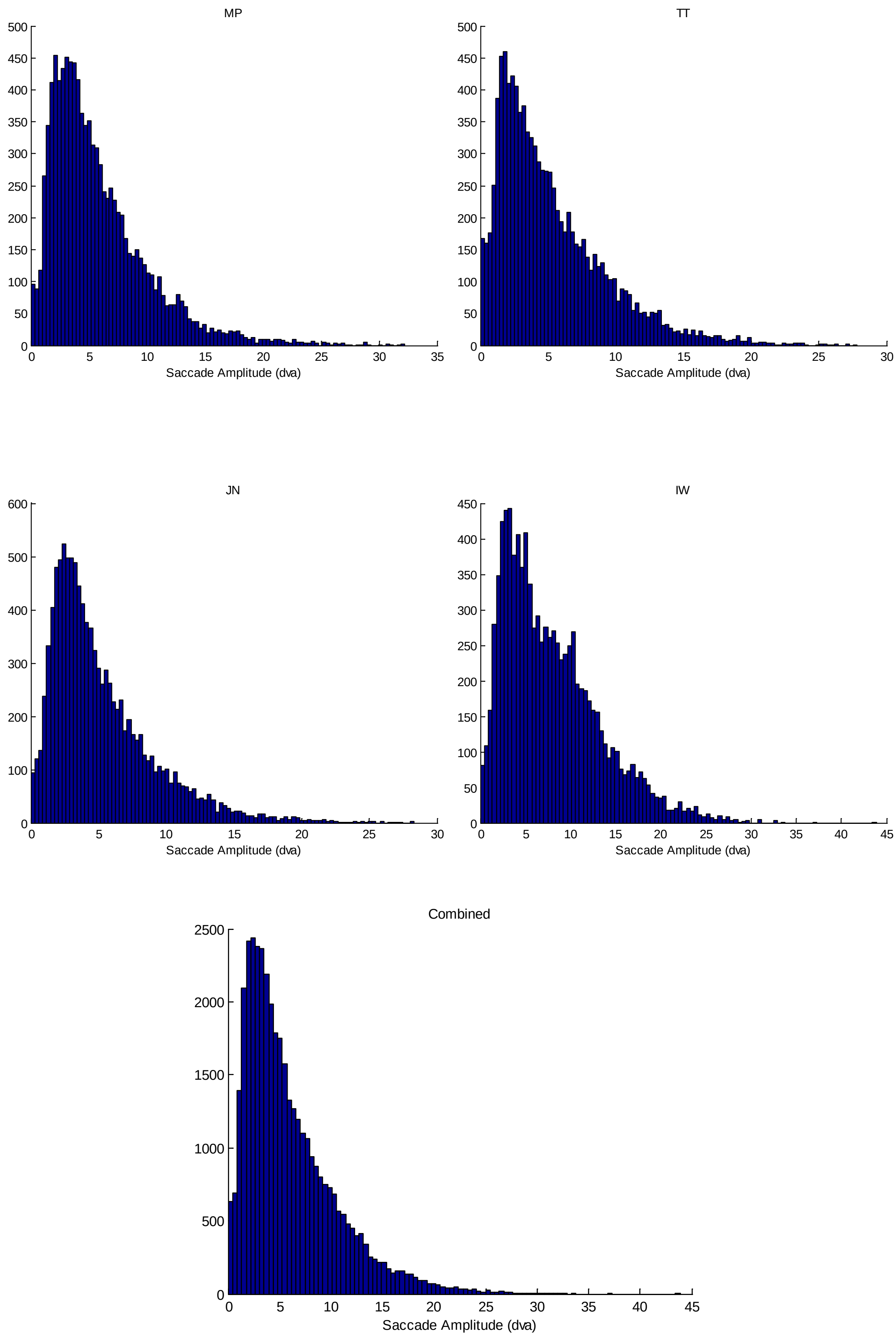
Definition: saccade amplitude is the end to end distance of a saccade. That is, the distance from the eye position of when the saccade begins to the eye position of when the saccade ends. As other algorithms may detect the onset of saccades differently, Cluster Fix may produce slightly different results. This data potentially includes microsaccades which are typically operationally defined as saccades with amplitudes less than 1 dva. In fact combined 3.9350% of all saccade amplitudes combined were less than or equal to 1 dva.

Units: distance in dva

Filtered Range: none

Cutoff Justification: Saccade amplitude statistics already have sufficient bounds from 0 dva to the maximum distance from any point on the image to another which is 1000 pixels (3,4,5 triangle ☺) or 42 dva though most saccades will not have this amplitude. The one apparent saccade by IW made that was greater than 42 dva may be due to a calibration issue or a purposeful saccade to the outside of the image that was within 2 dva of the image boarder and thus not excluded. Some extremely small saccade amplitudes may be microsaccades that could not be properly detected at 200 Hz and are extremely quantized.

Raw	MP	TT	JN	IW	Combined
Mean	5.87	5.26	5.26	7.87	6.03
Median	4.77	4.17	4.18	6.68	4.79
STD	4.29	4.03	3.85	5.35	4.52
Minimum	0.01	0.01	0.04	0.02	0.01
Maximum	32.21	27.83	28.34	43.86	43.86
5%	1.23	0.84	1.12	1.54	1.14
25%	2.83	2.26	2.51	3.65	2.72
75%	7.77	7.26	6.97	10.88	8.19
95%	14.03	13.17	12.92	18.24	15.05


MP
Saccade Amplitude (dva)
TT
Saccade Amplitude (dva)
JN
Saccade Amplitude (dva)
IW
Saccade Amplitude (dva)
Combined
Saccade Amplitude (dva)

VI. Distance between Fixations

Definition: the distance between fixations in the distance between the mean fixation locations of consecutive fixations. The distance between fixations is very similar to saccade amplitudes but not exactly the same since mean fixation locations are used and not the distance between the end of one fixation and the start of the next fixation.

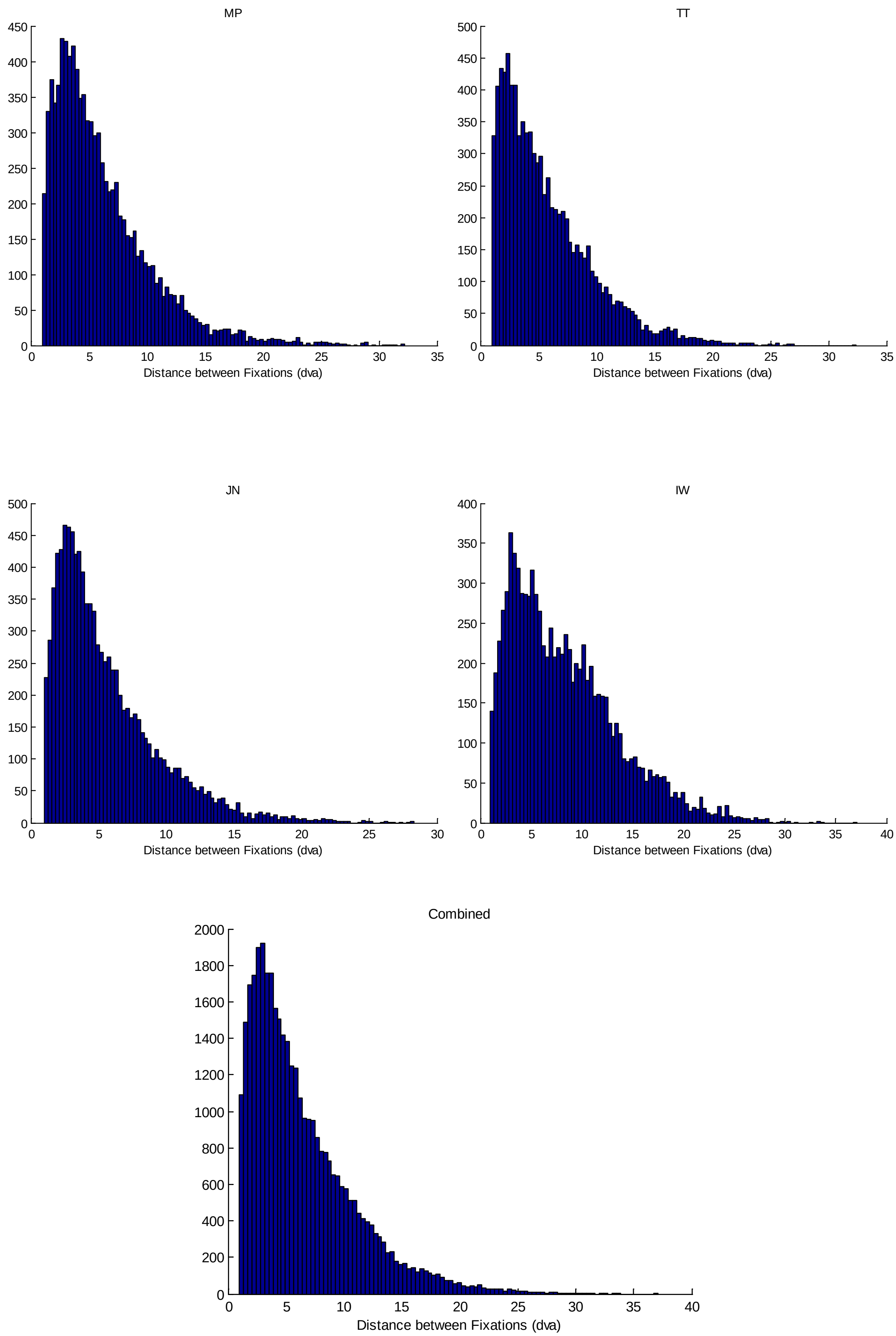
Units: distance in dva

Filtered Range: 1 – 42 dva

Cutoff Justification: Fixation distances less than 1 dva are probably the same fixation but with a microsaccade breaking the fixation into 2 fixations detected by Cluster Fix. Fixations greater than 42 dva (1000 pixels) are likely from fixations occurring outside of the image or a calibration issue.

Raw	MP	TT	JN	IW	Combined
Mean	5.92	5.31	5.36	8.02	6.12
Median	4.84	4.30	4.31	6.99	4.94
STD	4.32	4.10	3.92	5.38	4.58
Minimum	0.04	0.02	0.03	0.02	0.02
Maximum	32.19	32.36	28.27	43.32	43.32
5%	1.11	0.58	0.94	1.35	0.90
25%	2.86	2.26	2.57	3.78	2.77
75%	7.88	7.41	7.21	11.12	8.40
95%	14.17	13.21	13.06	18.31	15.17

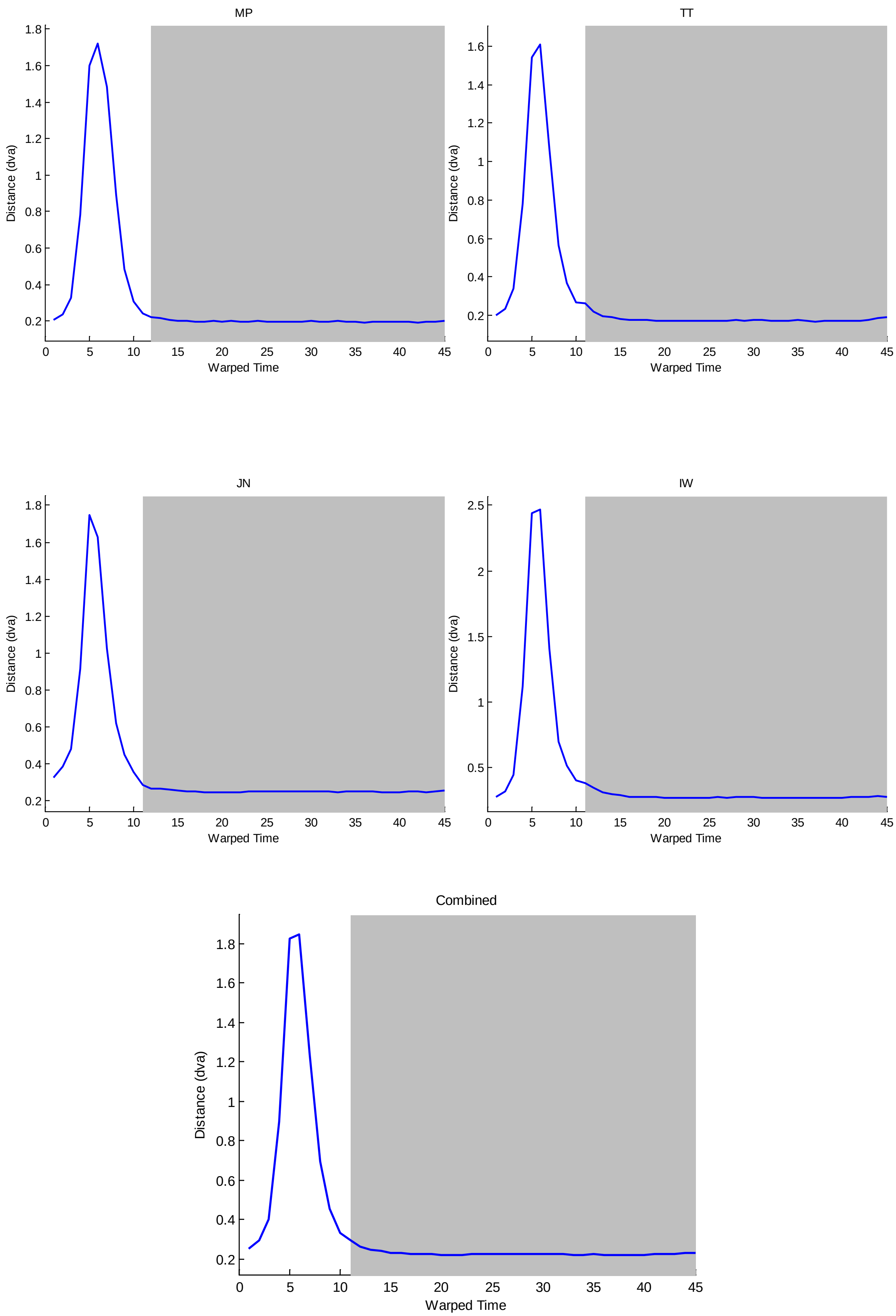
Filtered	MP	TT	JN	IW	Combined
Mean	6.17	5.78	5.63	8.30	6.45
Median	5.06	4.74	4.54	7.25	5.22
STD	4.26	3.99	3.85	5.26	4.49
Minimum	1.00	1.00	1.00	1.01	1.00
Maximum	32.19	32.36	28.27	37.08	37.08
5%	1.56	1.42	1.54	1.95	1.56
25%	3.09	2.72	2.82	4.10	3.09
75%	8.06	7.75	7.40	11.31	8.67
95%	14.32	13.52	13.23	18.41	15.40


MP
Distance between Fixations (dva)
TT
Distance between Fixations (dva)
JN
Distance between Fixations (dva)
IW
Distance between Fixations (dva)
Combined
Distance between Fixations (dva)

VII. Distance Profile

Definition: the distance profile is the distance over which the eye traverses from one time point to the next. Typically, the eye covers more distance during a saccade than during a fixation. Each profile is warped to the median length of saccades plus the median length of fixations, respectively. For the combined distance profile, all saccades and fixations were warped to the median length of all saccades and fixations across all monkeys, respectively. To be clear the distance profile is first the distance over which the eye covered from one time sample to the next (5 ms time resolution) during a saccade followed by the distance over which the eye covered from one time sample to the next during a fixation. Gray shading indicates fixation times.

Units: distance in dva

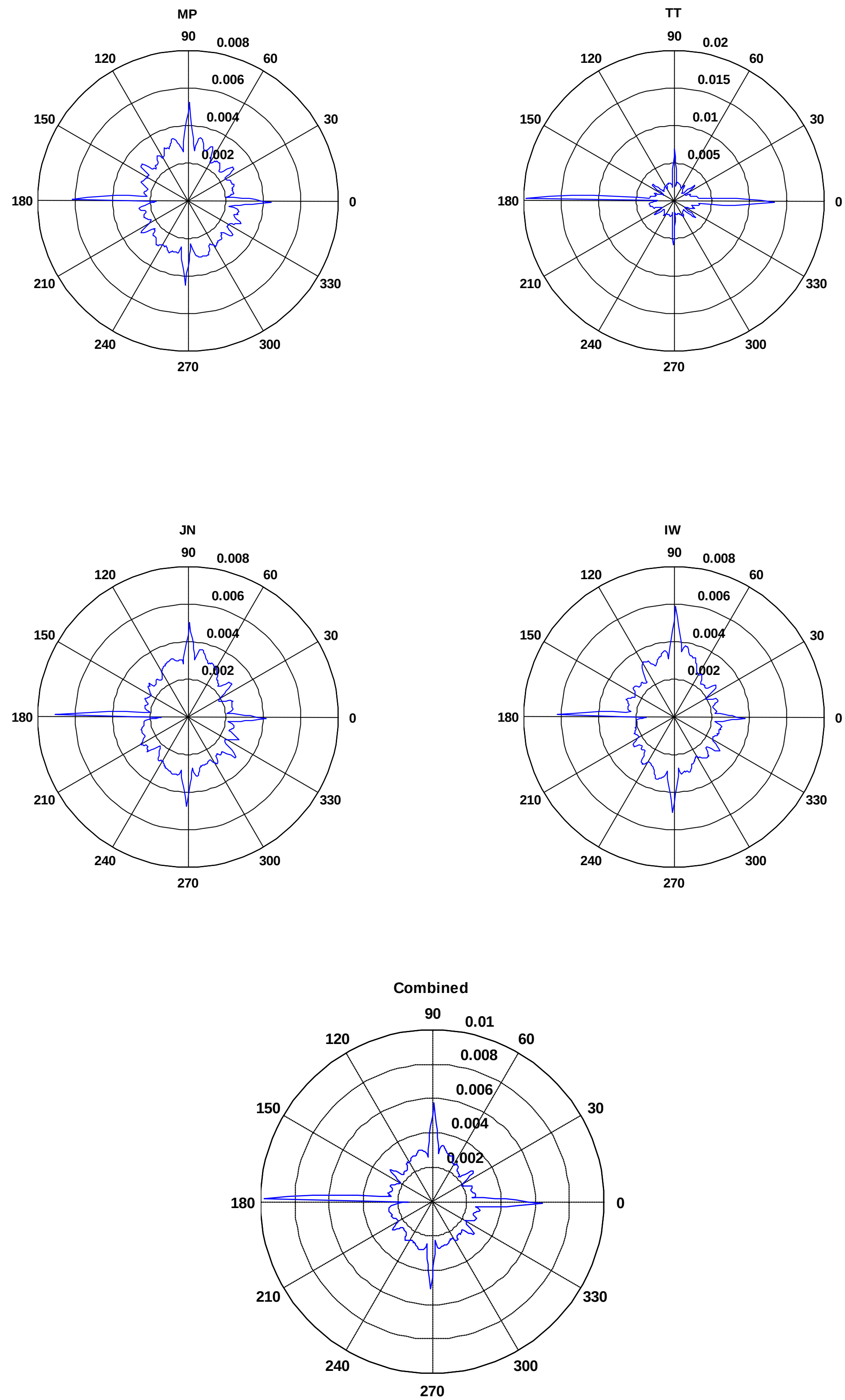

MP
TT
JN
IW
Combined
Distance (dva)
Warped Time

VIII. Saccade Angle Leaving a Fixation

Definition: the saccade angle leaving a fixation is the direction that the eyes move in as the monkey makes a saccade following a fixation.

Units: angle in degrees

The plots for all angles are filtered with a 6 degree moving average smoothing filter. The cardinal direction (i.e. N, E, S, W) have extreme biases which are smoothed out some.



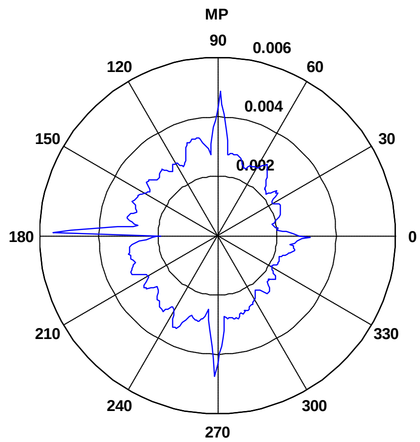
MP
90
0.008
60
120
0.006
150
0.004
30
0.002
180
0
210
330
240
300
270
TT
0.02
0.015
0.01
0.005
JN
IW
Combined
0.01

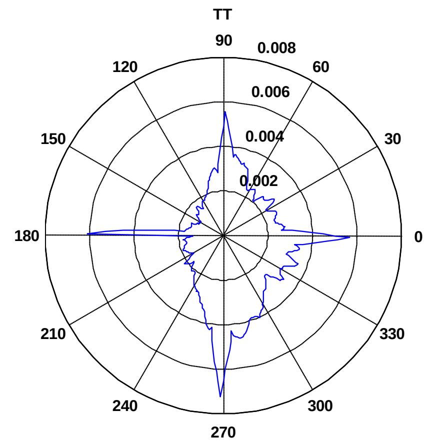
IX. Saccade Angle Entering a Fixation

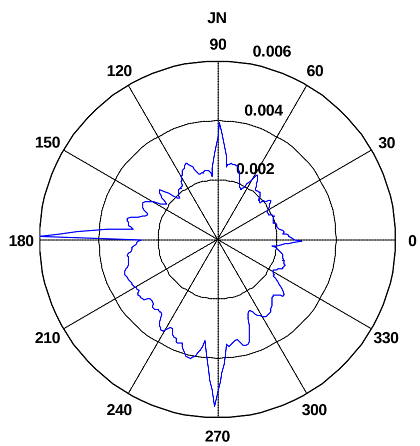
Definition: the saccade angle entering a fixation is the direction that the eyes move in as the monkey ends their saccade and enters a fixation.

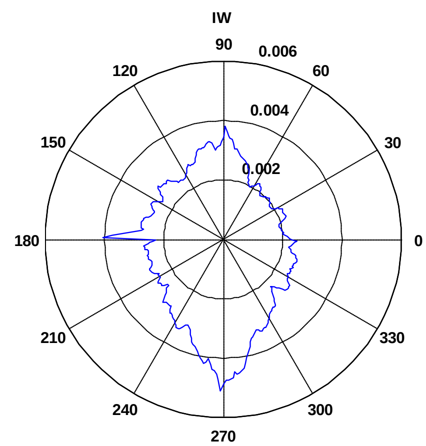
Units: angle in degrees

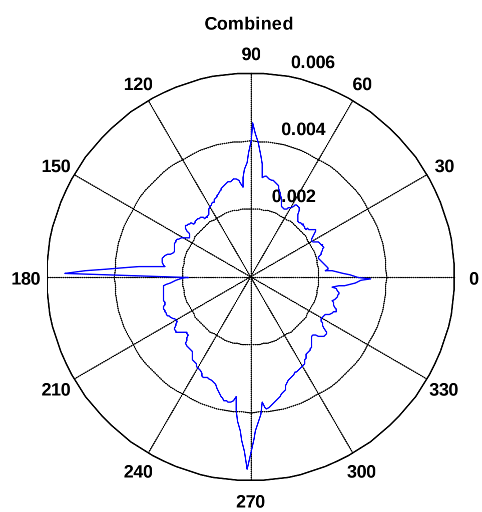
Again, the plots for all angles are filtered with a 6 degree moving average smoothing filter. The cardinal direction (i.e. N, E, S, W) have biases which are smoothed out some. The biases do not seem to be as strongly oriented towards the cardinal directions but more toward the N & S directions for the saccade angle entering a fixation compared to the saccade angle leaving a fixation.


MP
90
0.006
0.004
0.002
60
30
0
330
300
270
240
210
180
150
120


TT
90
0.008
0.006
0.004
0.002
60
30
0
330
300
270
240
210
180
150
120


JN
90
0.006
0.004
0.002
60
30
0
330
300
270
240
210
180
150
120


IW
90
0.006
0.004
0.002
60
30
0
330
300
270
240
210
180
150
120

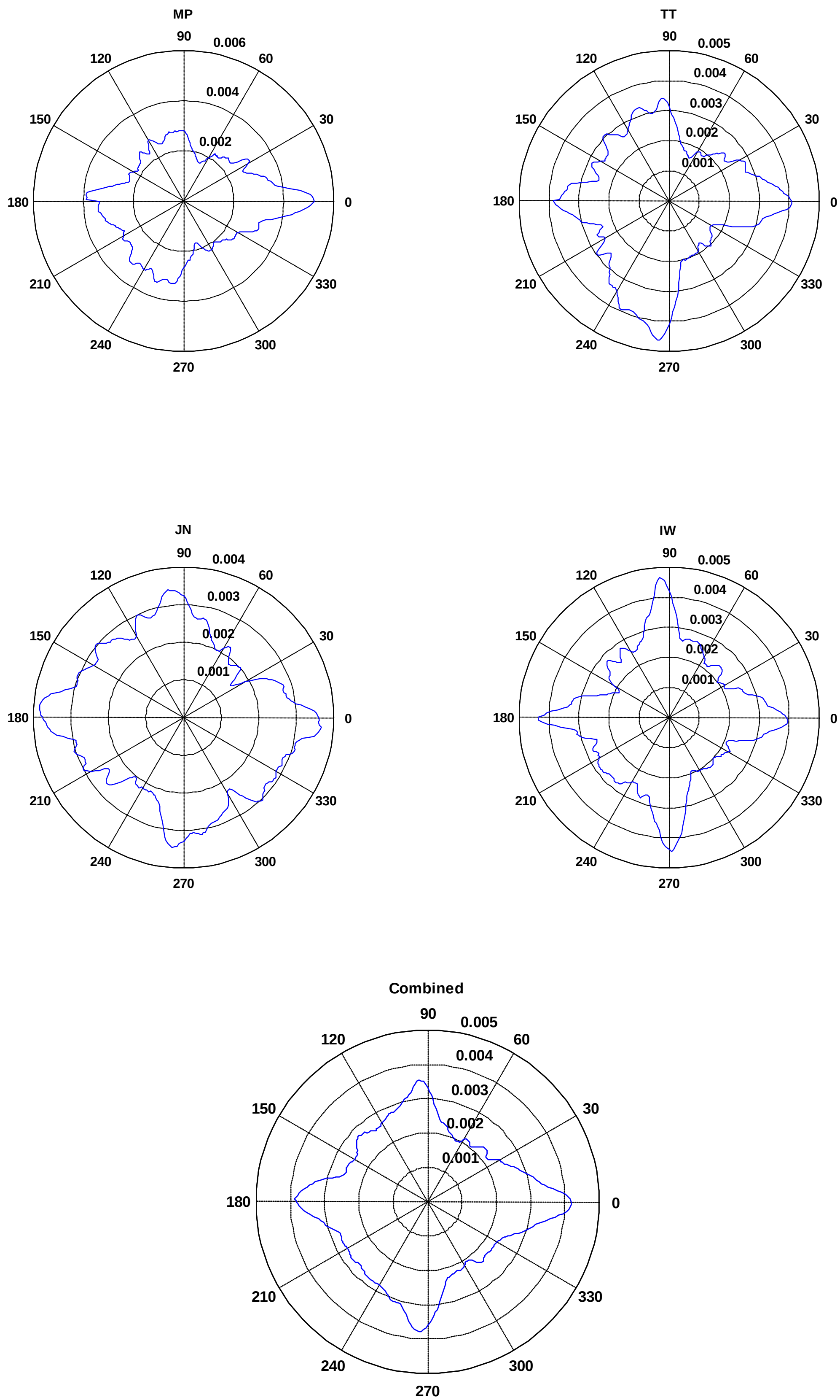

Combined
90
0.006
0.004
0.002
60
30
0
330
300
270
240
210
180
150
120

X. Angle between Fixations

Definition: the angle between fixations is the relative orientation of consecutive fixations to each other.

Units: angle in degrees

Again, the plots for all angles are filtered with a 6 degree moving average smoothing filter. While the saccade angles entering and leaving a fixation are extremely biased in particularly cardinal directions, the angle between fixations is relatively uniform implying that rapid eye movements may be constrained by the mechanics of the eye but fixation locations are not so much. For several of the monkeys—particularly IW—the angle between fixations may still follow a fairly biased distribution which likely causes the combined distribution to continue to be biased as well.


MP
90
0.006
120
60
0.004
150
30
0.002
180
0
210
330
240
300
270
TT
90
0.005
120
60
0.004
0.003
150
30
0.002
0.001
180
0
210
330
240
300
270
JN
90
0.004
120
60
0.003
0.002
150
30
0.001
180
0
210
330
240
300
270
IW
90
0.005
120
60
0.004
0.003
150
30
0.002
0.001
180
0
210
330
240
300
270
Combined
90
0.005
120
60
0.004
0.003
150
30
0.002
0.001
180
0
210
330
240
300
270

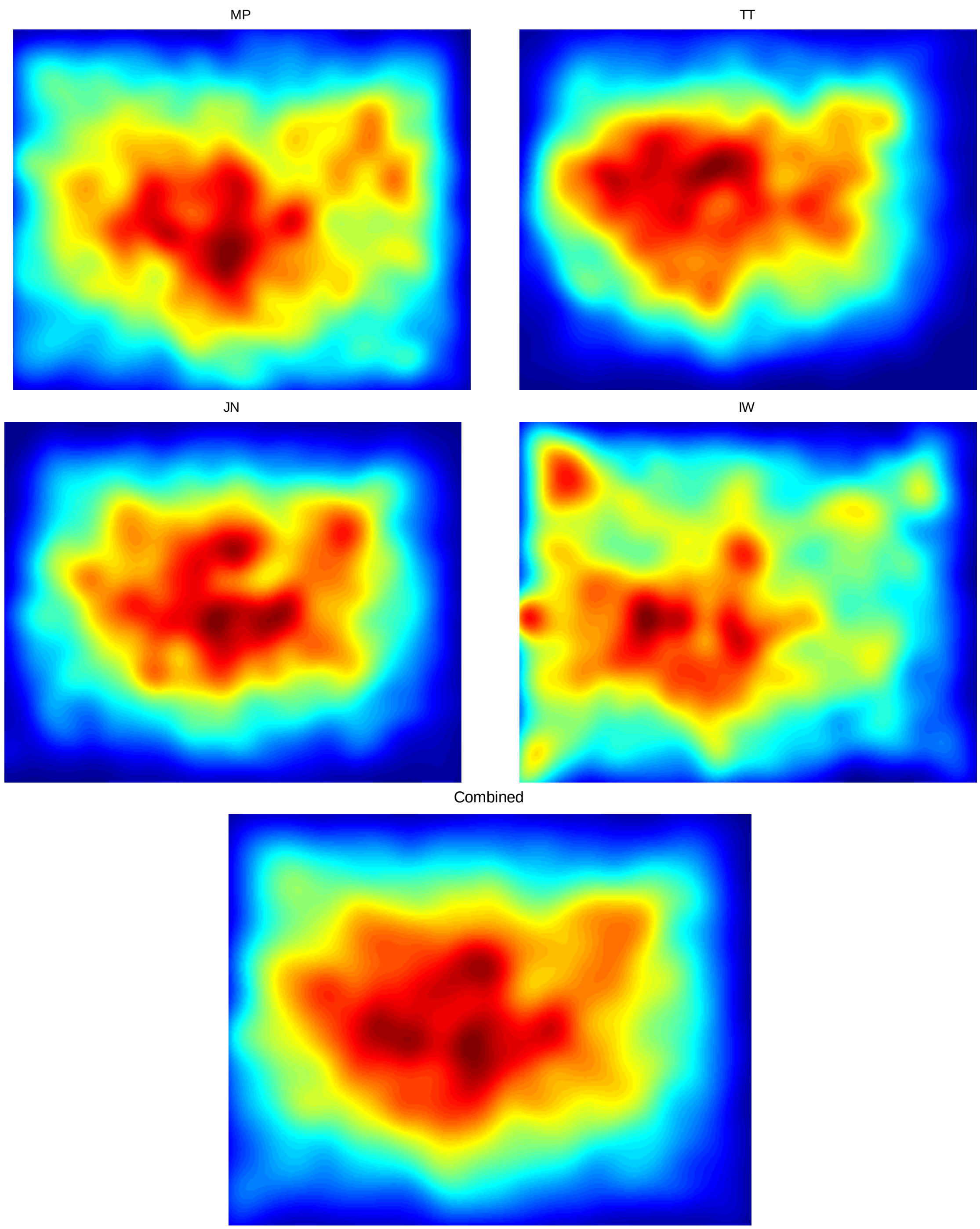
F. Fixation Location Statistics

In this section I describe some statistics pertaining to fixation locations including the distribution of fixation locations, salience at fixation locations, and image intensity at fixation locations. As with all other sections all these statistics will be defined including a simple explanation on how they were derived. Again, these statistics are currently only for the novel viewing of images and not the viewing of repeated or familiar images. For salience at least, the salience at fixation locations during the viewing of a repeated image appears to be significantly lower at every fixation except the first one. Therefore, I conclude these statistics in particularly are sensitive to memory. Conversely, I would suspect that the viewing behavior statistics in the previous section to be statistically similar (i.e. $p > 0.05$) between novel and familiar images because the previous statistics are more likely to be dependent on oculomotor system than memory; in other words I postulate viewing behavior statistics like those in the previous section are constrained by the statistics of the eye and motor control systems and are rather independent of memory unlike fixation locations

The fixation location statistics are considerably important when trying to fit a model of bottom-up viewing behavior. Bottom-up is defined as stimulus driven features such as visual salience. Therefore, bottom-up viewing behavior is viewing behavior that is driven by stimulus features, and in our case the stimuli are natural images. I have done a considerable amount of statistical analysis on how well salience and a BCRW predicts fixation location and fixation order including using KL-divergence and a classical ROC analysis. The BCRW, or biased correlated random walk, is model that incorporates viewing behavior statistics from the above section with a saliency map to predict viewing behavior. In general the results of the KL-diveragence and ROC analyses indicate that salience can be used to predict fixation locations and order, but the BCRW is a better predictor than salience alone. As a comparison image intensity (luminance) was used as a control, and image intensity faired only slightly better than chance as a predictor of fixation location and order. If you are interested in further details and analysis of fixation location statistics feel free to contact me!

I. Fixation Locations-Fixation 2D PDFs

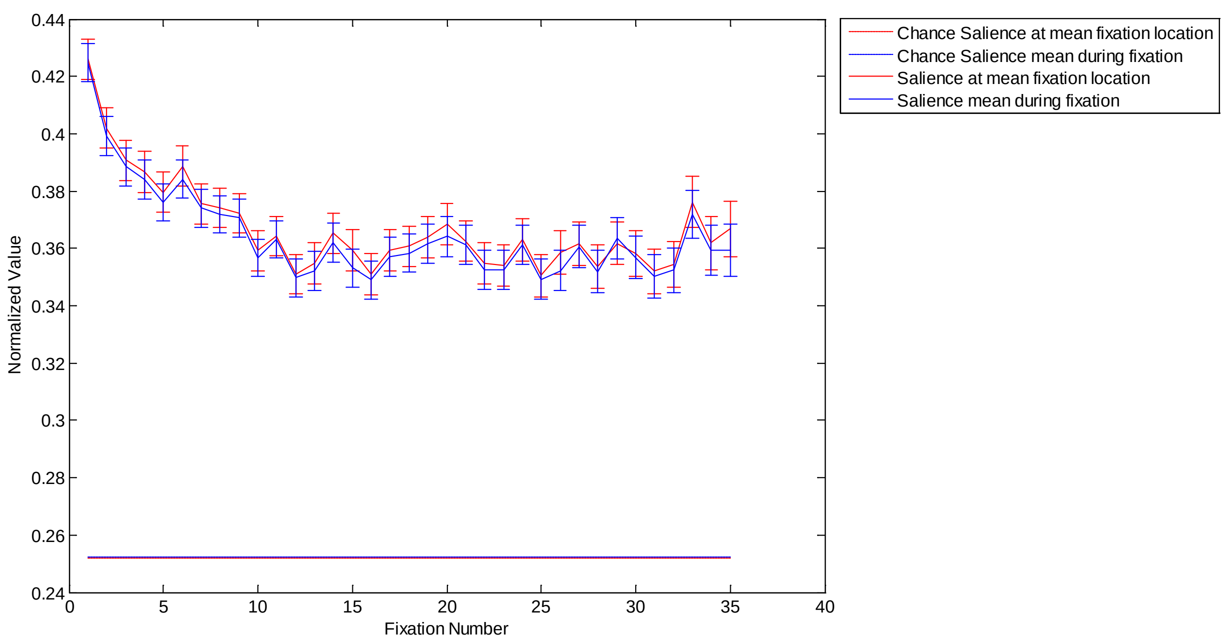
Fixation location was determined using Cluster Fix as the mean location during a fixation. To derive a 2D fixation location PDF each fixation was marked in a matrix the size of the images in SCM, and then the matrix was smoothed with a 1 dva Gaussian filter, and finally normalized to sum to 1. Each monkey has their own idiosyncratic bias, but in general fixations are strongly biased towards the center. Several studies have found that the central bias is independent of where the cross-hair is at the start of each trial. The combined PDF includes about 40,000 fixations with approximately 10,000 fixations per monkey.



II. Salience at Fixation Locations

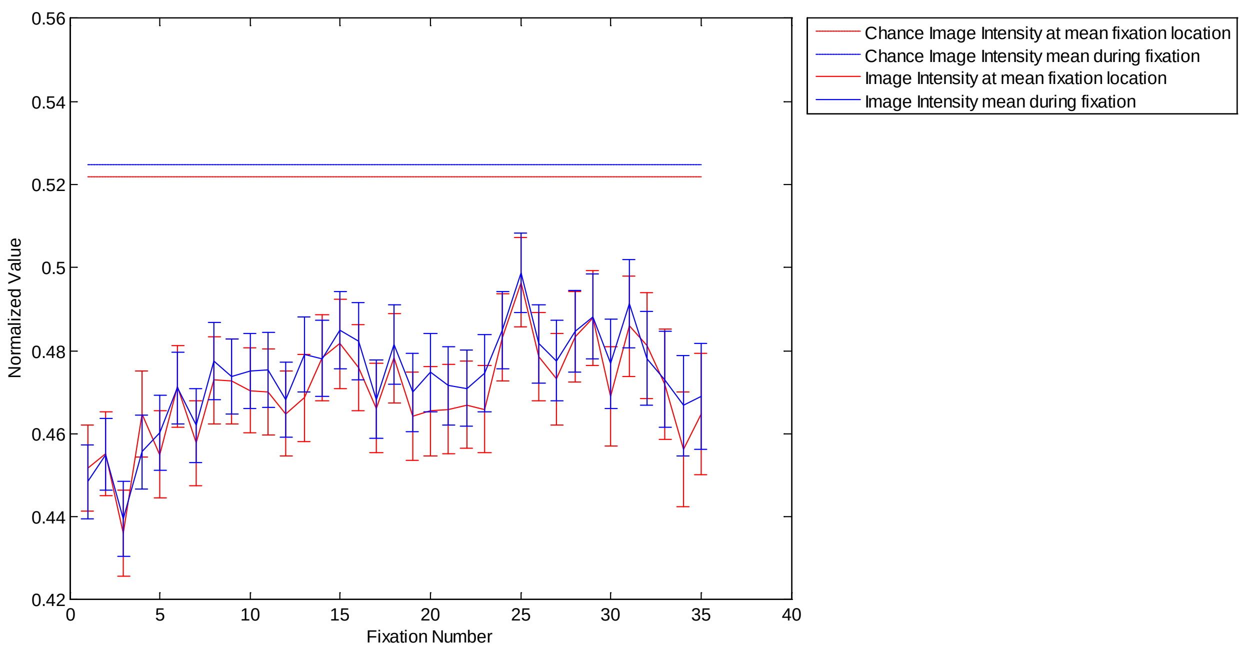
I calculated salience, or visual salience, using the Itti, Koch, and Niebur algorithm from their 1998 paper titled, “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis.” The algorithm I employ is extremely similar with only a few derivations mostly in the number of orientation contrast maps. In simple terms saliency maps are calculated from summing color, intensity (luminance), and orientation contrasts across multiple spatial scales. Itti, Koch, and Nieber modeled contrast filters after idealized receptive fields found in the retina, LGN, and V1 in hopes of creating a biologically plausible mechanism for explaining fixation location. Saliency maps often highlight faces and other prominent objects. My saliency maps are independently normalized with the lowest salient pixel in an image containing a salience of 0 and the most salient pixel contains a salience of 1. The distribution of salience values for each saliency map varies slightly but typically has heavy left tails meaning that most salience values are closer to 0 than 1.

For salience at fixation locations I calculated the salience at the mean fixation location (same as the above fixation location statistic) and the mean salience during a fixation. Since Cluster Fix determines the time points that constitute as a fixation, I can calculate the salience at each fixation time point and then average across all fixation time points for a single fixation to get the mean salience during a fixation. As a control for mean fixation locations, I calculated salience at random fixation locations to derive the chance salience. As a control for the mean salience during a fixation, I calculated the mean salience of random points was the same number of points as there were time points in a fixation. A 2-tailed z-test comparing the salience at the mean fixation location to chance levels showed that the salience at each mean fixation locations was significantly greater than chance ($p = 0$, yes I mean $p = 0$). A z-test comparing the mean salience during a fixation to chance levels showed that the mean salience during each fixation was also significantly greater than chance ($p = 0$). Error bars are mean +/- s.e.m. Data comes from approximately 40,000 fixations.



III. Image Intensity (Luminance) at Fixation Locations

Similar to calculating the salience at fixation locations, I also calculated image intensity (luminance) at mean fixations and the mean image intensity during a fixation. Image intensity was calculated as the gray scale pixel value ranging from 0 to 255 and then normalized to be from 0 to 1 by dividing pixel values by 255. It is more apparent here that values calculated at the mean fixation location are different than the values calculated from the mean during a fixation though not much different. Again, a 2-tailed z-test showed that image intensity values at each mean fixation locations were significantly less than chance ($p < 2e^{-51}$), and mean image intensity values during each fixation were significantly less than chance ($p < 3e^{-124}$). While the image intensity values at each mean fixation location were significantly less than chance, image intensity is not a great predictor of fixation location or order as indicated by both a ROC analysis and KL-divergence. Error bars are mean +/- s.e.m. Again, data comes from approximately 40,000 fixations.



G. Other Useful Data Relevant to Viewing Behavior

In this section I include several other figures and data that you may find useful for modeling, data analysis, and task design. The format of this section follows the format of all the previous sections.

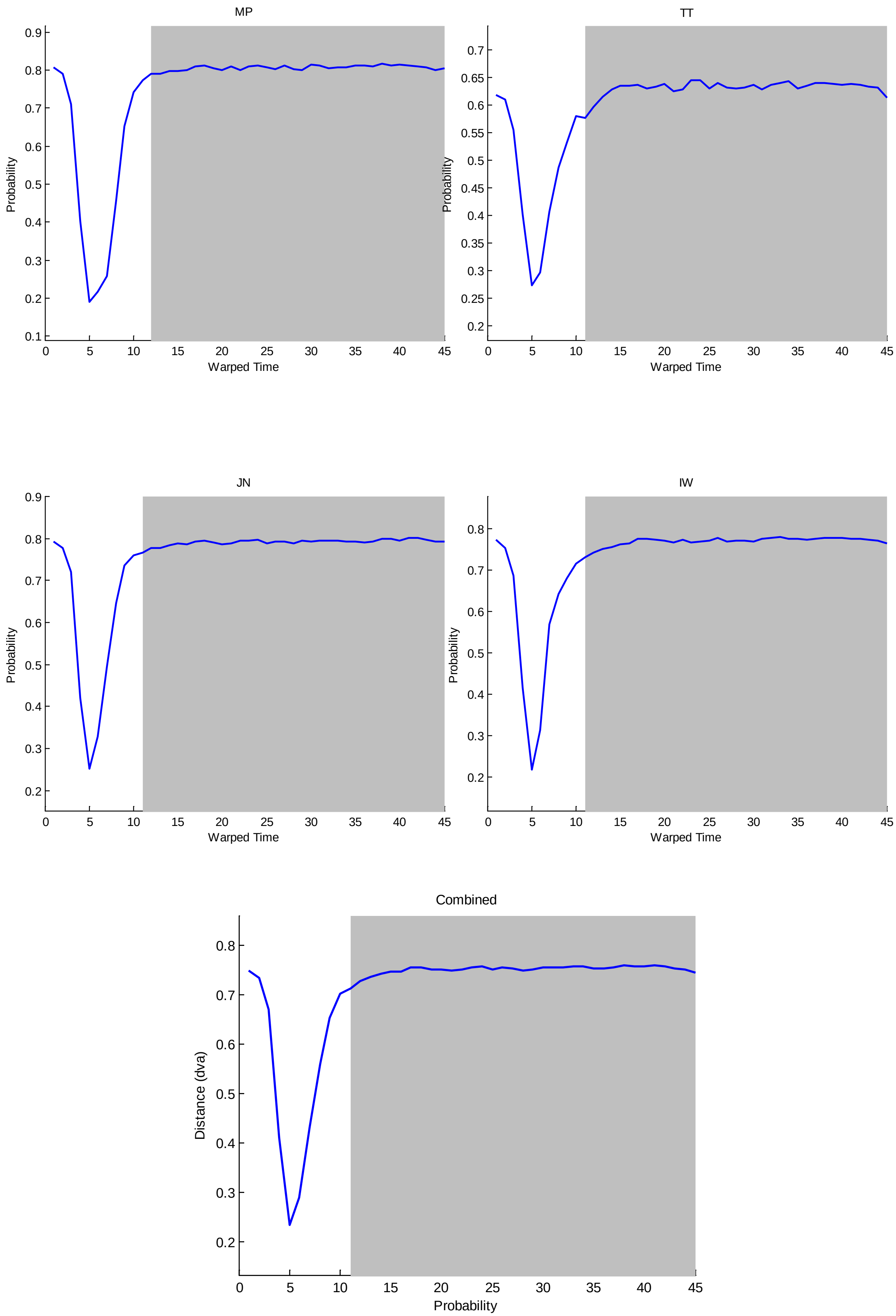
I. Persistence

Definition: persistence is the probability that the eye moves in the same direction over time. You can think of persistence as inertia or that eye movements tend to be correlated with each other over time. I have defined a direction threshold of 45 degrees. The selection of 45 degrees is based on the fact that the intermediate cardinal directions (i.e. NE, SE, SW, and NW) are separated by 45 degrees from the cardinal directions (N, E, S, and W). Furthermore, the approximate intermediate value of rotation of the scan path (see parameter profiles, next section) during a fixation and saccade is 45 degrees. Thus, mathematically persistence is defined as the probability that the direction of the eye changes by more than 45 degrees from one time point to the next.

Unfortunately, the way I defined persistence is counterintuitive because persistence implies constancy; however, with the calculation of persistence being the probability that the direction of movement changes by more than 45 degrees, persistence is more like the lack of constancy. Let me explain. During a saccade the eye moves in an arching fashion and the direction of movement is relatively smooth causing the persistence term be closer to 0. At the beginning and end of a saccade the eye moves less linearly and thus persistence values become higher. During a fixation, the center of the eye—thus the scan path—appears to fluctuate (tremor) around the mean fixation location in a chaotic fashion. This slow eye movement with an ever changing direction has low correlation and thus a high persistence value.

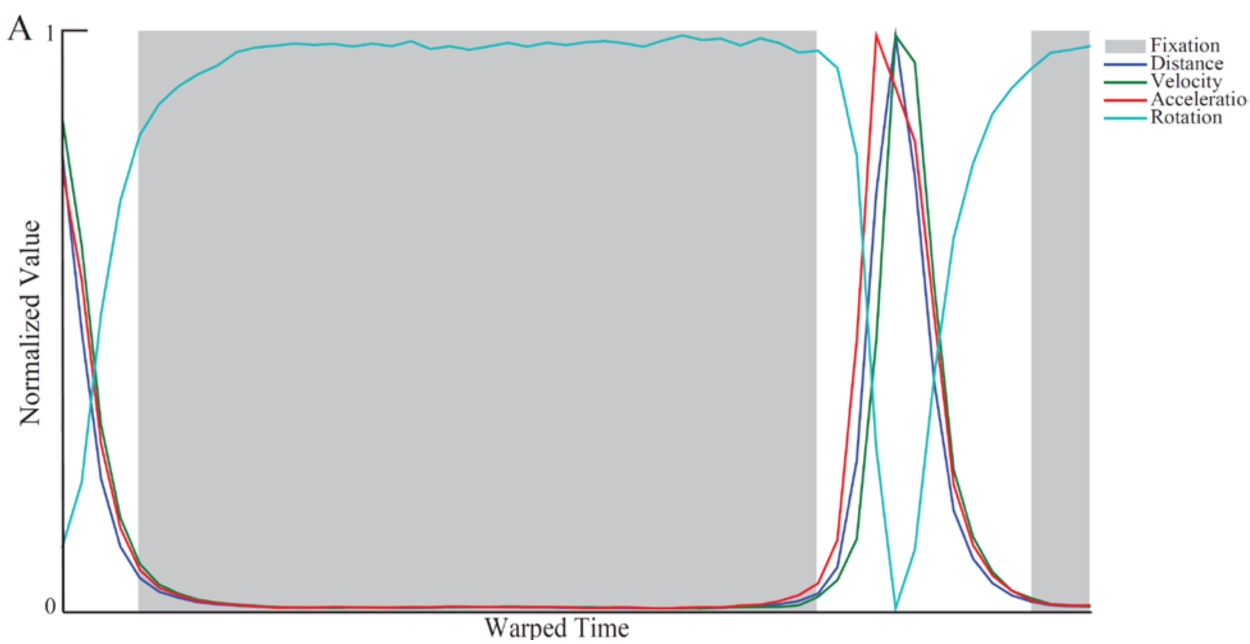
Like with the distance profile, I use time warping to analyze fixations and saccades of different time lengths. In the figures below, fixation times are shaded in gray.

Units: unitless/probability


MP
Probability
Warped Time
TT
Probability
Warped Time
JN
Probability
Warped Time
IW
Probability
Warped Time
Combined
Distance (dva)
Probability

II. Parameter Profiles

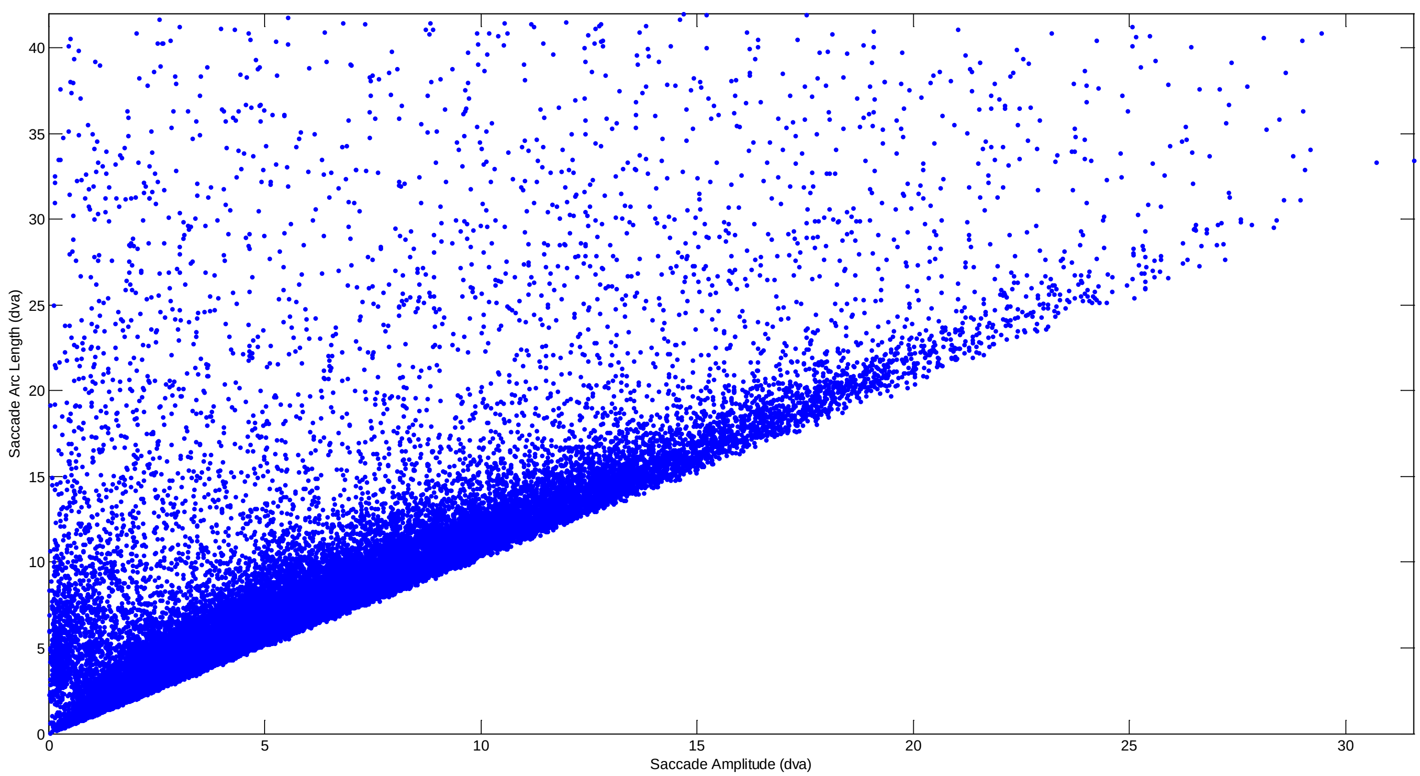
Parameter profiles are the extension of the distance profile but for 4 state space parameters—distance, velocity, acceleration, and rotation. Velocity and acceleration were computed as the first and second derivative of position, respectively. Distance was measured slightly differently than before as the distance traversed from one time point to two time points later (10 ms). Rotation, which is analogous to angular velocity, was calculated as the difference in the angle of eye movement from one time point to the next. Again, I use time warping to analyze fixations and saccades of different time lengths. These parameter profiles are integral in the Cluster Fix algorithm so here they are normalized to be from 0 to 1. I have placed the profiles here to demonstrate that these 4 parameters are qualitatively and quantitatively distinct between fixations and saccades. Any subsequent fixation and saccade detection algorithms may want to use these parameters as a starting point. I should note non-normalized parameter profiles are available. Parameter profiles were combined across approximately 40,000 fixations and saccades of all lengths from the raw data.



III. Correlation between Saccade Arc Length and Saccade Distance

I believe that saccade amplitude is a less sensitive measure than saccade arc length because saccade amplitudes do not account for the direction or necessarily the duration of a saccade. However, we would suspect that saccade amplitude and arc length would be significantly correlated with saccade arc lengths always being at least as long as saccade amplitudes. Here I will just show that saccade amplitude and saccade arc length are indeed correlated. On a side note, saccade arc length is logarithmically correlated with saccade duration.

For this analysis I use the same cutoffs as in the viewing behavior statistics section. I then compute the Pearson correlation coefficient, r, how significant this correlation is, p, and the slope-intercept that optimally describes the relationship between saccade amplitude and saccade arc length. As you can see, often amplitude and arc length are highly correlated (n = 39749, r = 0.786, p = 2), but there are a good number of data points which do not follow this trend (saccade arc length = 1.1086*saccade amplitude + 1.6895). There also appears to be a good number of saccades with short amplitudes (1 < dva) with longer (1-10 dva) saccade arc lengths that also do not follow this trend. On average saccade arc length is 1.3935 times larger than saccade amplitude.

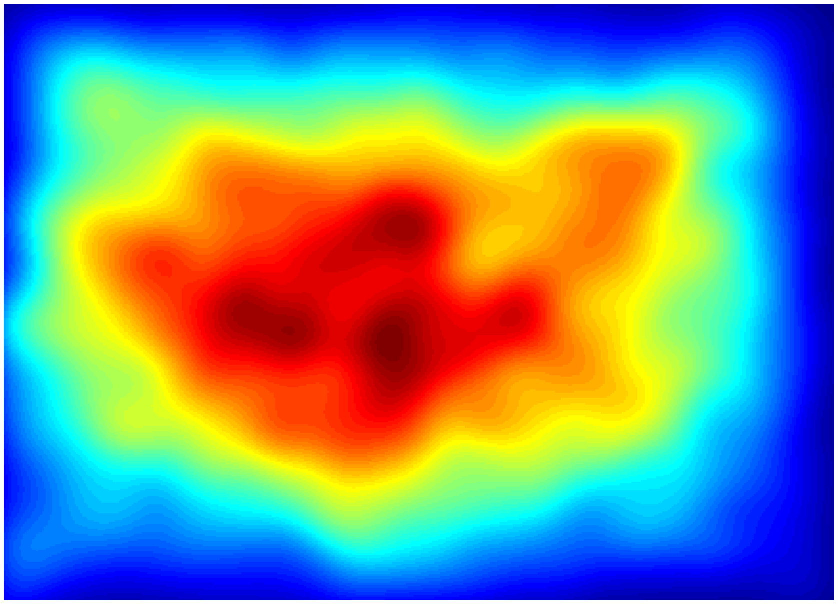


IV. Average Maps

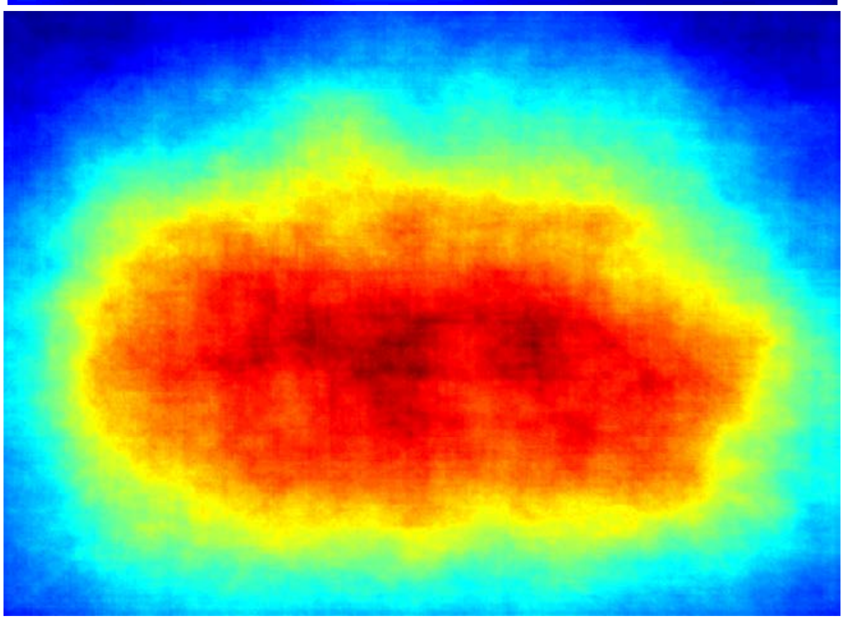
What is the arrangement of objects, fixations, etc. in SCM and other tasks that use natural scenes? Some of this curiosity comes from the observation that fixations have a strong central bias, and we often wonder why. As you will see below, the average saliency map has a strong central bias. Often the central bias in the saliency map is called the photographer's bias because the people who take the images of natural scenes often center objects in the scene or center the camera in order to optimize the perspective on the objects of interest. Averaging the image intensity [map] across the same image sets reveals nothing of real interest because we observe the sky is brighter than the ground. Combined with the results that fixations occur in highly salient regions, we conclude the central bias of fixation locations is highly related if not caused in great part by photographer's bias. Analysis of the BCRW salience model shows that oculomotor statistics (i.e. how the eye moves) also help create this central fixation bias.

Another interest of the average and individual saliency maps is the frequency of prominent objects in the image because Nathan, Mike, and Beth discovered grid cells in the macaque entorhinal cortex which have spatially periodic signals. Neither Nathan Killian nor I have found any indication of strong periodicity either in individual maps or average maps. Power, in the 2D fourier transform domain (power spectrum), across individual and average maps falls off with increases in spatial frequency as expected with natural images; most of the power is in lower frequencies.

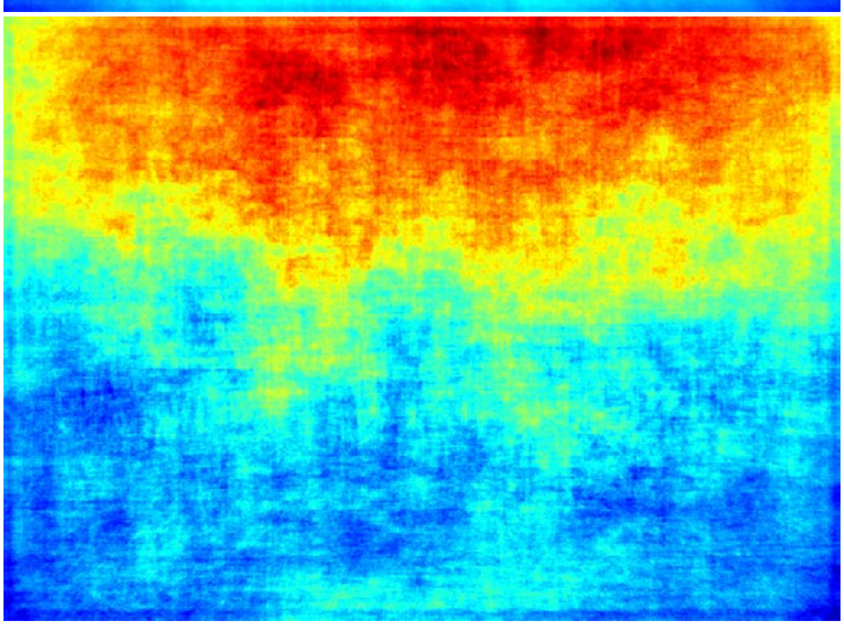
The figures on the next page are combined across all 4 monkeys over 288 images from both SCM and SCME tasks. The 2D fixation location PDF is an exact copy of the PDF from the fixation location statistics section.



Combined 2D Fixation Location PDf



Combined Saliency Map



Combined Image Intensity Map

Z. Description of Updates

As this is version 1.0 there are currently no updates.