Sequential Selection, Saliency and Scanpaths

Ramin Fahimi

Advisor: Prof Neil Bruce

University of Manitoba
Department of Computer Science

20 Nov 2018

Outline

Attention

- Humans and animals have a significant but limited amount of processing resources.
- Attention is the process of selectively concentrating on some portion of available information
- Critical to Information Reduction and filtering.
- Visual attention is a form of attention that allows humans and animals to control and allocate their limited visual processing resources
- Sequences of successive eye movements during scene exploration are described as visual scanpaths

Visual attention - Fixation and saccades

FIGURE 1 – Scanpaths and salient regions – the ordering reveals a compelling difference among observers, even when over location of fixations are consistent.

eye-movements

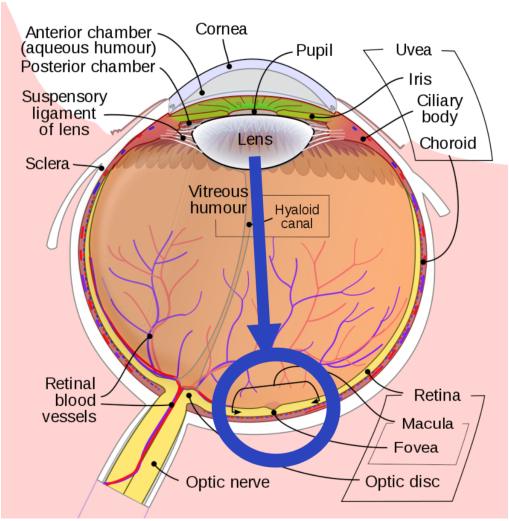


FIGURE 2 – Inside humans eyes - Fovea [1]

Saccade	voluntary fast & ballistic movements
Vestibular-Ocular Reflex	compensatory changes in eye position as head moves
Smooth Pursuit	voluntary tracking of moving stimuli
Nystagmus	Reset by a primitive saccade similar to reaching the orbit limit
Optokinetic Nystagmus	low-frequency rotations at constant velocity.
Vergence	Coordinated movements - converging or diverging
Torsion	Coordinated rotation, dependent on head tilt and eye elevation.

datasets

Name	Size	Observer	duration	Fixations	Type
OSIE [22]	700	15	3	✓	
CROWD [14]	500	16	5	✓	
PASCAL-S [17]	850	12	2	✓	
CAT2000 [4]	2000	18	5	✓	
FIGRIM [5]	2787	15	2	✓	
LOWRES [22]	168+25	15	3	✓	
SALICON [13]	10000	60	-	✓	

TABLE 2 – Available Datasets for training & evaluating models of gaze behaviour. Note that this list derives from careful consideration of all available datasets, their duration, and modifications in how information is extracted to engage in both spatial and temporal analysis.

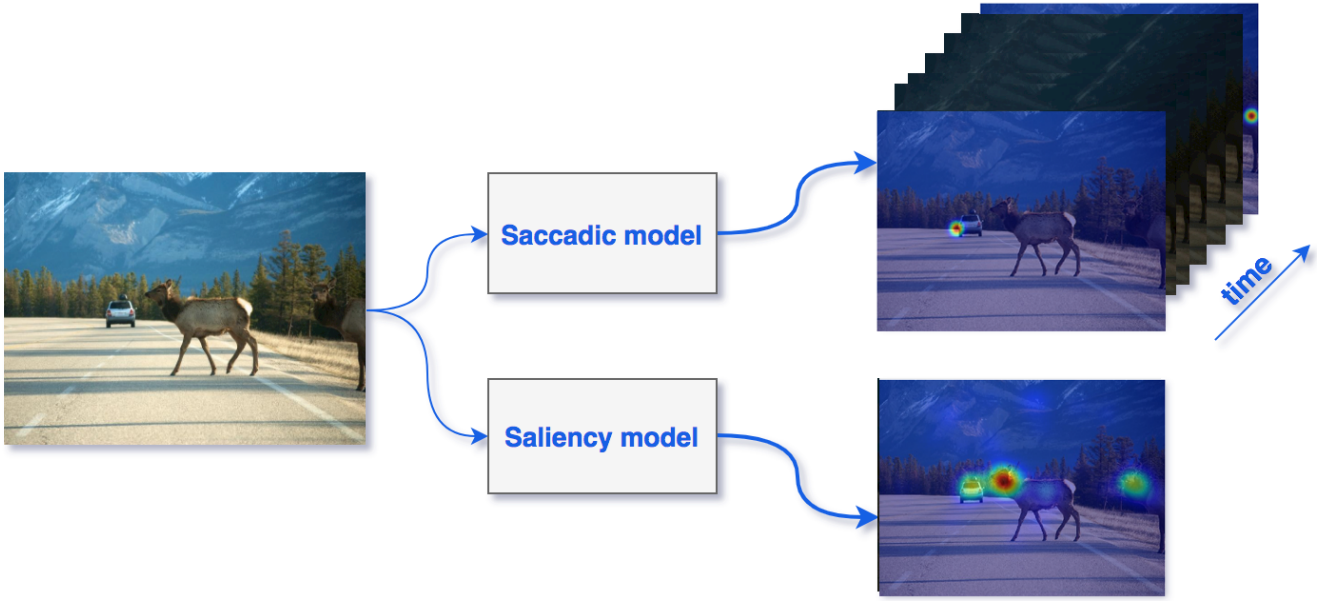
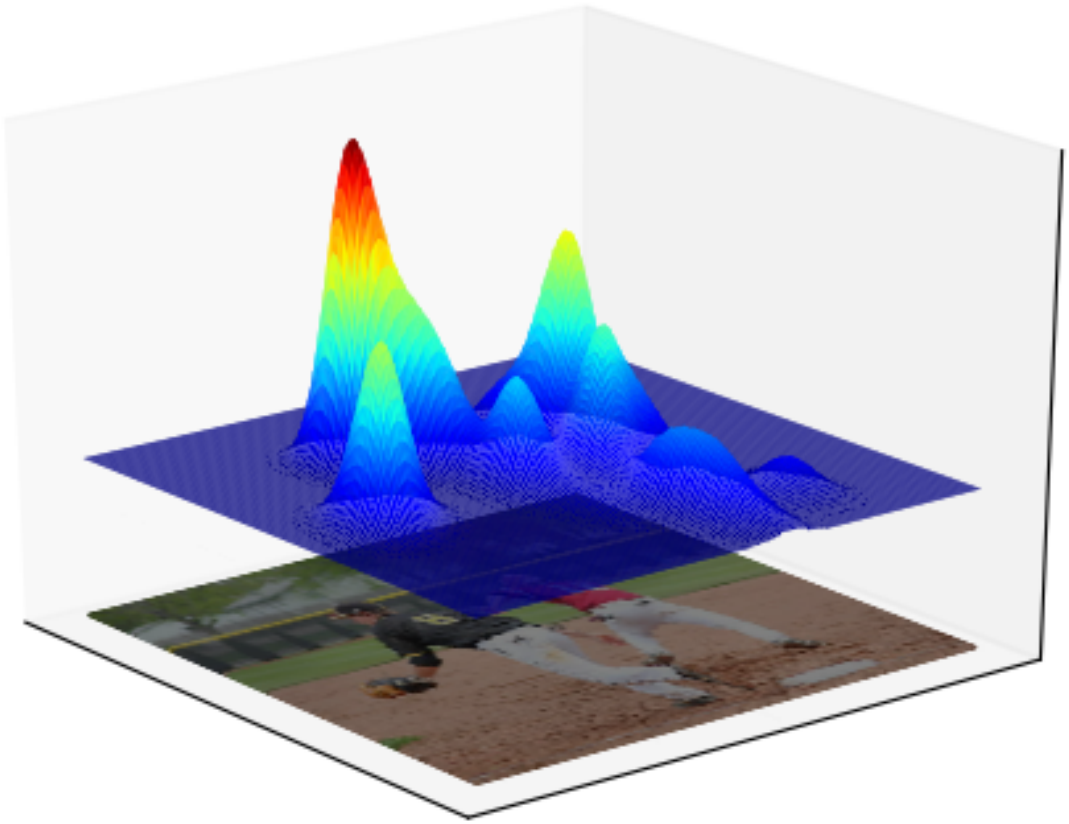


FIGURE 3 – Saliency (static) models vs. Saccadic models

Static



Static metrics

	AUC	sAUC	SIM	CC	KL	IG	NSS	EMD
Implementation								
Bounded	✓	✓	✓	✓				
Location-based, parameter-free	✓	✓				✓	✓	
Local computations, differentiable			✓	✓	✓	✓	✓	
Symmetric			✓	✓				✓
Behavior								
Invariant to monotonic transformations	✓	✓						
Invariant to linear transformations (contrast)	✓	✓		✓			✓	
Requires special treatment of center bias		✓				✓		
Most affected by false negatives			✓		✓	✓		
Scales with spatial distance								✓

TABLE 3 – A summary of static metrics for comparing gaze patterns (by Bylinskii et al. [6])

Sequential

- tries to predict **Position, order, shape, duration**
- Started by scanpath theory anad koch model.
- continued with many more (within - between)
- Applications involves all the existing + high focus on active camera control

Sequential Metrics

	Metric	abrv	Quantization	Target
1	Euclidean distance	EUC	Direct	Position
2	Mannan distance [18]	MAN	Direct	Position
3	Eyeanalysis [19]	EYE	Direct	Position
4	Levenshtein distance [20]	LEV	Grid	Position, Order
5	ScanMatch [7]	SMT	Grid/Temporal	Position,Order,Duration
6	Hausdorff distance [12]	HAU	Direct	Position
7	Frechet distance [9]	FRE	Direct	Position, Order
8	Dynamic time warp [3]	DTW	Direct	Position, Order
9	Time delay embedding [21]	TDE	Direct	Position
10	MultiMatch Shape [8]	MM_S	Direct	Shape
11	MultiMatch Direction(angualr)[8]	MM_A	Direct	Direction
12	MultiMatch Length [8]	MM_L	Direct	Length
13	MultiMatch Position [8]	MM_P	Direct	Position
14	MultiMatch Duration [8]	MM_D	Direct	Duration
15	Recurrence [2]	REC	Radius	Position
16	Determinism [2]	DET	Radius	Fixation Trajectories
17	Laminarity [2]	LAM	Radius	Fixation Persistence
18	Corm [2]	COR	Radius	Leading/Following

TABLE 4 – Common metrics for evaluation of Scanpaths.

Sequential metrics - commons

Given two scanpaths P and Q with lengths N and M :

- $P = (p_1, p_2, ..., p_n)$
- $Q = (q_1, q_2, ..., q_m)$
- $d_{i,j} = D_{EUC}(P_i, Q_j)$
- Stimuli resolution (W, H)

Euclidean, Mannan, Eyeanalysis

Euclidean distance

$$D_{EUC}(P,Q) = \sum_{i=1}^{\min(N,M)} \sqrt{(X_P - X_Q)^2 + (Y_P - Y_Q)^2}$$

Mannan distance

$$D_{MAN} = [1 - \frac{D}{D_r}] * 100$$

$$D^2 = \frac{N * \sum_{j=1}^{M} \min d_{i,j}^2 + M * \sum_{i=1}^{N} \min d_{i,j}^2}{2 * N * M * (W^2 + H^2)}$$

Eyeanalysis

$$D_{EYE} = \frac{(\sum_{i=1}^{M} \min d_{i,j}^2 + \sum_{j=1}^{N} \min d_{i,j}^2)}{max(N,M)}$$

Levenshtein

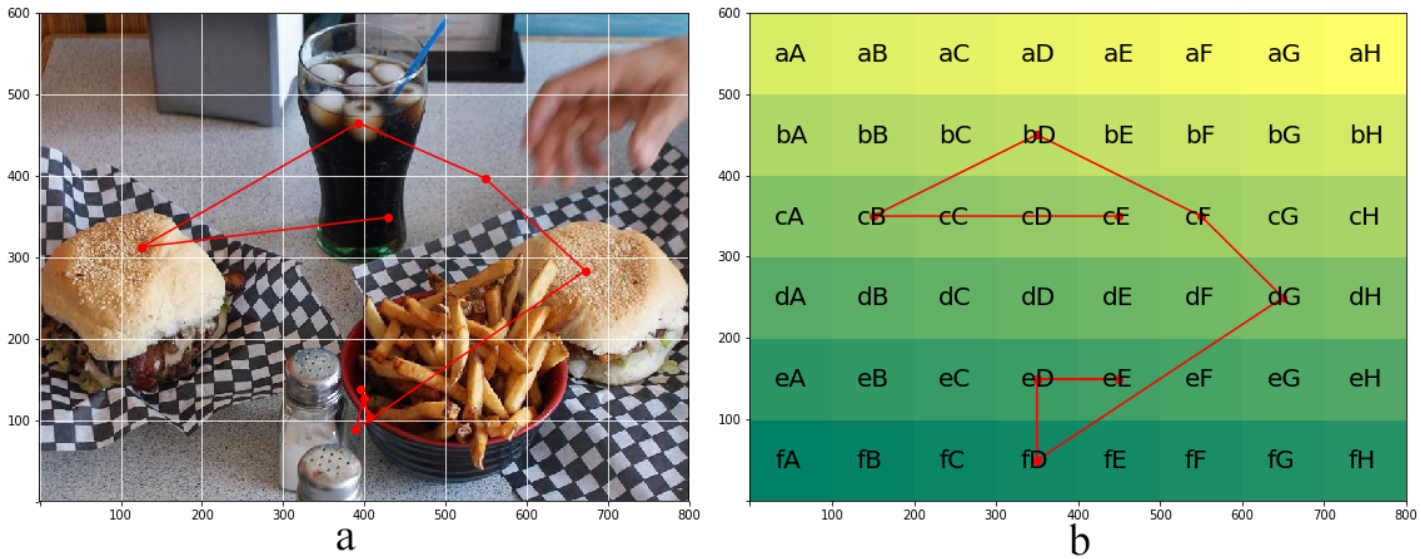
example : kitten vs sitting

kitten → sitten (substitution of "s" for "k")
sitten → sittin (substitution of "i" for "e")
sittin → sitting (insertion of "g" at the end)

Levenshtein

$$d_{A,B}(i,j) = \begin{cases} max(a,b) & if \quad min(i,j) = 0 \\ min\begin{cases} d_{A,B}(i-1,j)+1 \\ d_{A,B}(i,j-1)+1 \\ d_{A,B}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & otherwise \end{cases}$$

ScanMatch



X	Y	T
394.9	138.4	253
405.6	102.7	251
399.4	127.7	273
389.6	88.3	143
672	283.4	198
549	396.7	184
392.6	465	395
125.5	312.4	327
429.4	349.1	205

c

FIGURE 5 – a) Visualization of scanpath record in c. b) pre-processing stage for transforming scanpaths to strings. "eDeEeDfDdGcFbDcBc" without considering duration and" eDeDeEeEeDeDfDdGcFbDbDbDcBcBcBcEcE" when T=100ms.

String-based problems

- Close fixations could fall into different bins
- fixations in a specific bin are treated equally

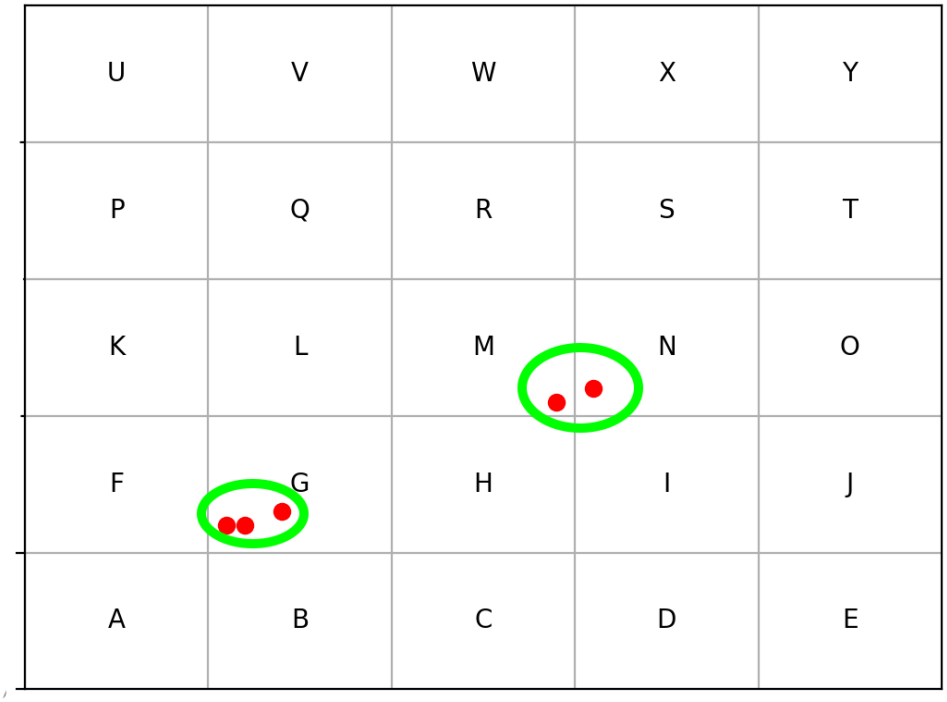


FIGURE 6

Dynamic time warping I

Algorithm 1 Dynamic Time Warp

procedure DTW(P, Q)

$D \leftarrow array[0..n, 0..m]$

$D[:, 0] \leftarrow infinity$

$D[0, :] \leftarrow infinity$

$D[0, 0] \leftarrow 0$

for $i \leftarrow 1 : N$ **do**

for $j \leftarrow 1 : M$ **do**

$D[i, j] \leftarrow minimum(D[i-1, j],\ D[i, j-1],\ D[i-1, j-1])$

$D[i, j] \leftarrow D[i, j] + d(P[i], Q[j])$

end for

end for

$return D[n, m]$

end procedure

Dynamic time warping II

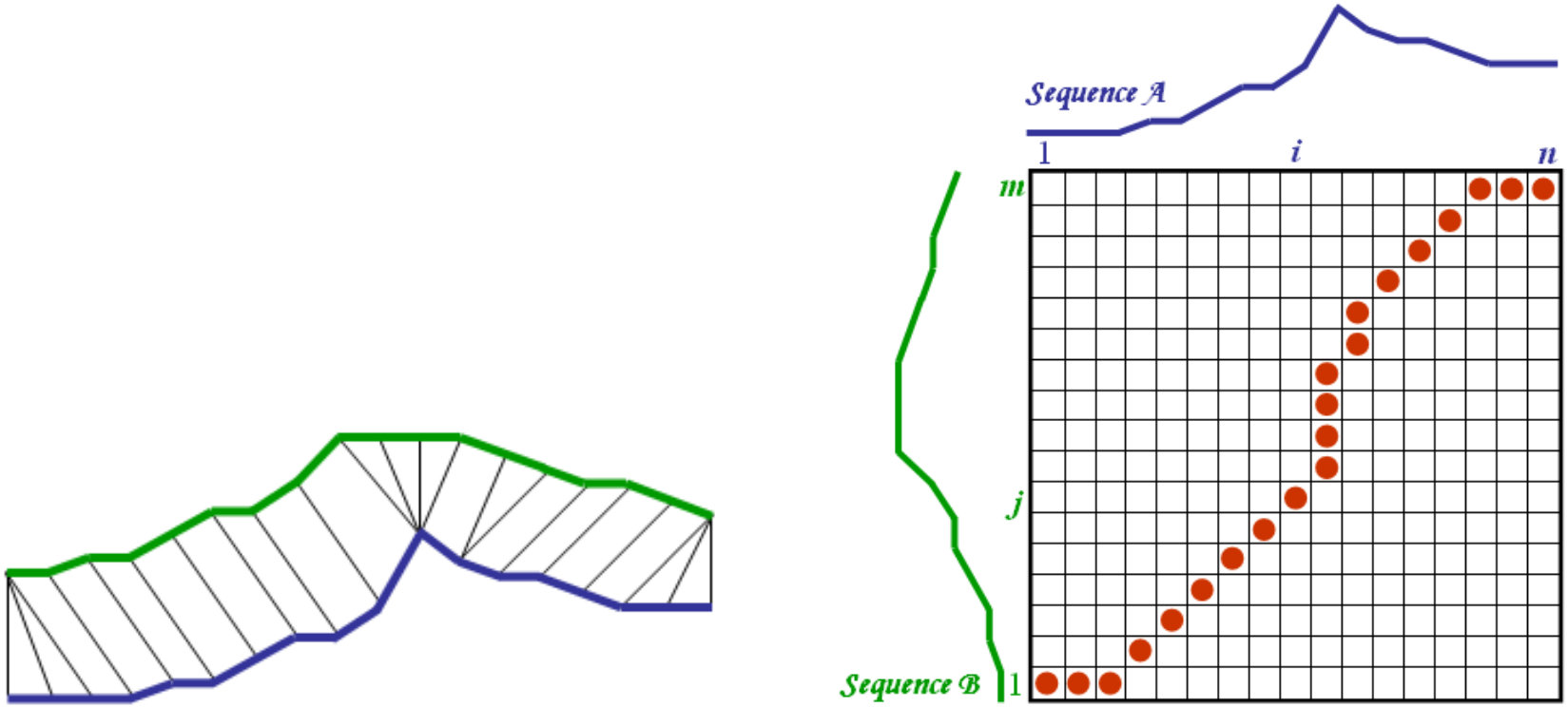


FIGURE 7 – [10]

Frechet distance

Frechet Distance

$$D_{FRE(A,B)} = \inf_{\alpha,\beta} \max_{t\in[0,1]} \left\{ d\Big(P(\alpha(t)),\, Q(\beta(t))\Big) \right\} \tag{1}$$

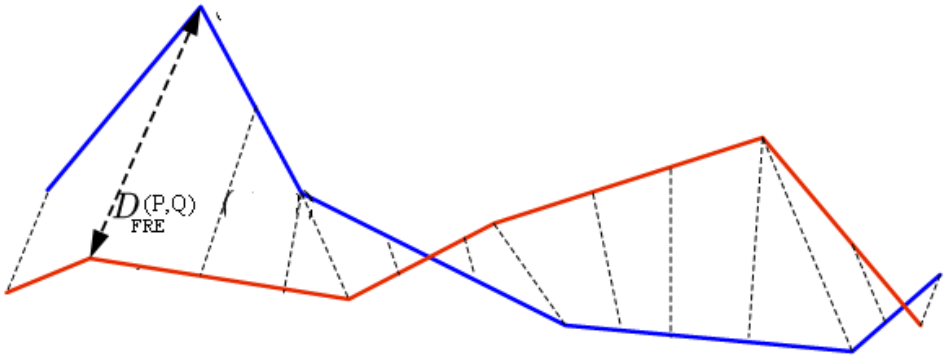


FIGURE 8 – [10]

Hausdorff distance & Time delay embedding (TDE)

Hausdorff Distance

$$D_{HAU}(P,Q) = max(h(P,Q), h(Q,P)) \tag{2}$$
$$h(A,B) = \max_{a \in P} \min_{b \in Q} ||a-b||$$

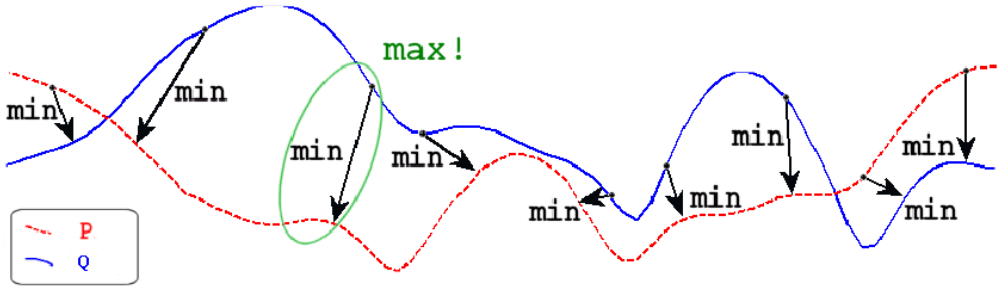


FIGURE 9 – [16]

MultiMatch

1. Simplification
   - Direction
   - Length
2. Alignment : Dijkstra algorithm
3. Calculate distance
   - Shape
   - Length
   - Direction
   - Position
   - Duration

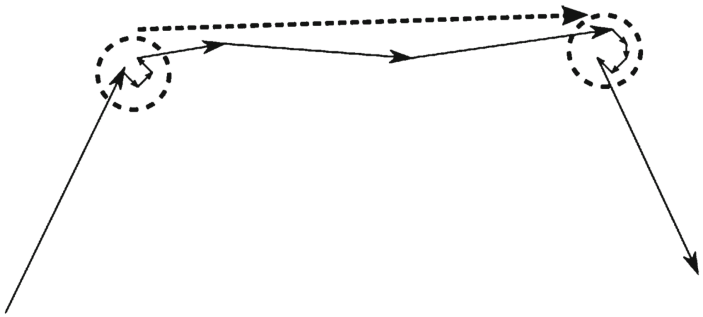
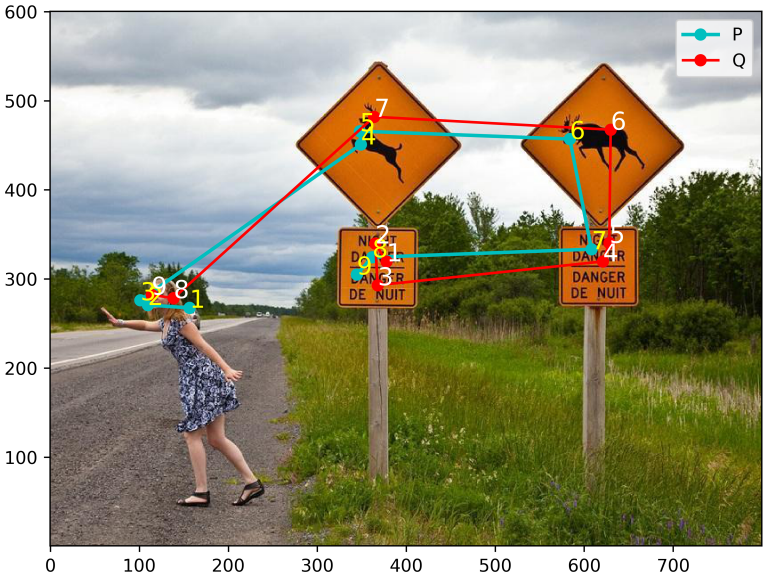


FIGURE 10 – Amplitude-based (dashed circles) and direction-based (dashed arrow) clustering [8, 11]

Recurrence



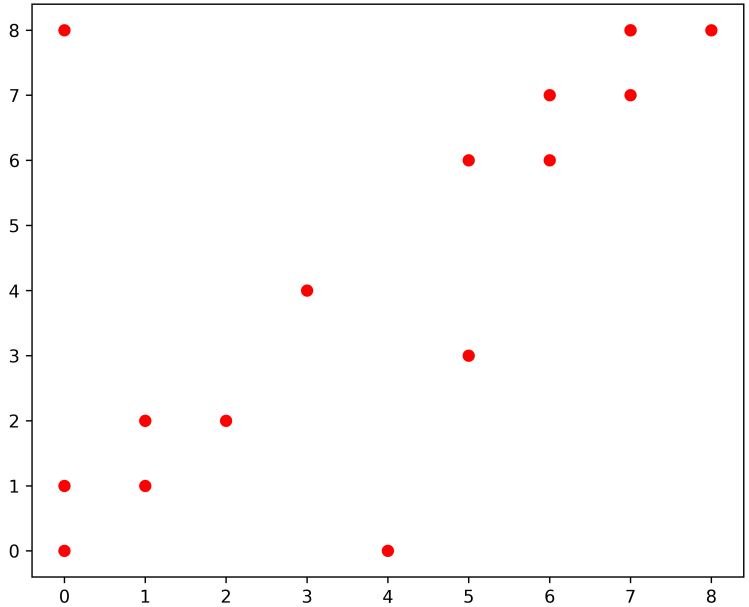


FIGURE 11 – Recurrence matrix based on euclidean distance with $\rho = 2 * 24$(visual angle)

$$r_{P,Q}(i,j) = \begin{cases} 1 & d(P[i], Q[j]) < \rho \\ 0 & otherwise \end{cases}$$

$$D_{REC}(P,Q) = 100 * \frac{C}{N^2}$$

Determinism, Laminarity, CORM

LAM

$$D_{LAM}(P,Q) = 100 * \frac{|H_L|+|V_L|}{2*C}$$

DET

$$D_{DET}(P,Q) = 100 * \frac{|D_L|}{R}$$

CORM

$$D_{COR}(P,Q) = 100 * \frac{\sum_{i=1}^{\min(M,N)} \sum_{j=1}^{\min(M,N)} ((j-i)r_{ij}}{(\min(M,N)-1)C}$$

Inter-observer Congruency (IOC)

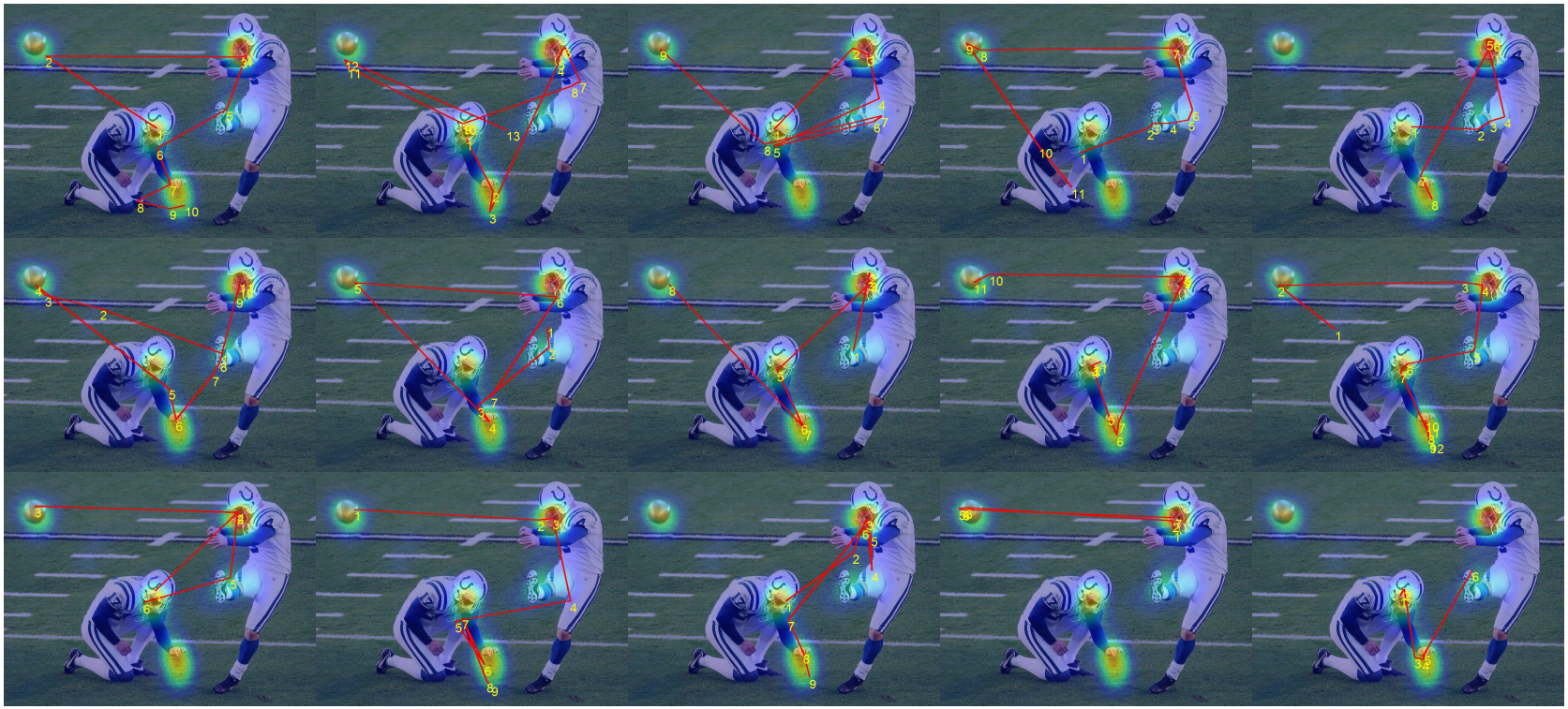


FIGURE 12 – Visualization of 15 observers viewing the same image revealing differences in viewing. Eye tracking data is subjective and the extent of agreement is measured as IOC.

What can a similarity metric do ?

A similarity metric should have the capability to weigh differences and similarities among models

1. Detecting differences between predefined set
2. Supporting scanpath theory
3. Diagnostic use
4. Data-driven clustering

What makes a metric suitable and what's missing ?

1. Should be capable of measuring distances between sequences of fixations.
2. Should have an interpretation that is intuitive
3. Should be able to effectively capture the order, position, duration of fixations and shape of scanpath[8].
4. Should provide a level of sensitivity that allows for reasonable separation of models that produce good vs. poor sequences
5. Should not consist only of coarse grained saccade or fixation statistics

Area under intersection

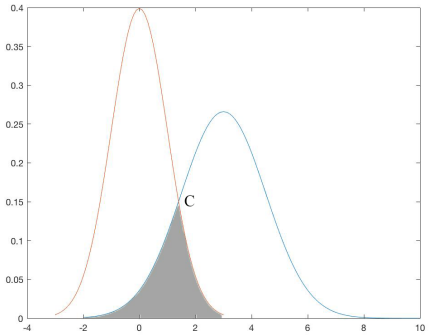


FIGURE 13 – Intersection

$$P(X_1 > c) + P(X_2 < c) = 1 - F_1(c) + F_2(c)$$

$$= 1 - \frac{1}{2}\text{erf}\left(\frac{c-\mu_1}{\sqrt{2}\sigma_1}\right) + \frac{1}{2}\text{erf}\left(\frac{c-\mu_2}{\sqrt{2}\sigma_2}\right)$$

where C can be calculated by :

$$c = \frac{\mu_2\sigma_1^2 - \sigma_2\left(\mu_1\sigma_2 + \sigma_1\sqrt{(\mu_1-\mu_2)^2 + 2\left(\sigma_1^2-\sigma_2^2\right)\log\left(\frac{\sigma_1}{\sigma_2}\right)}\right)}{\sigma_1^2-\sigma_2^2}$$

Noise

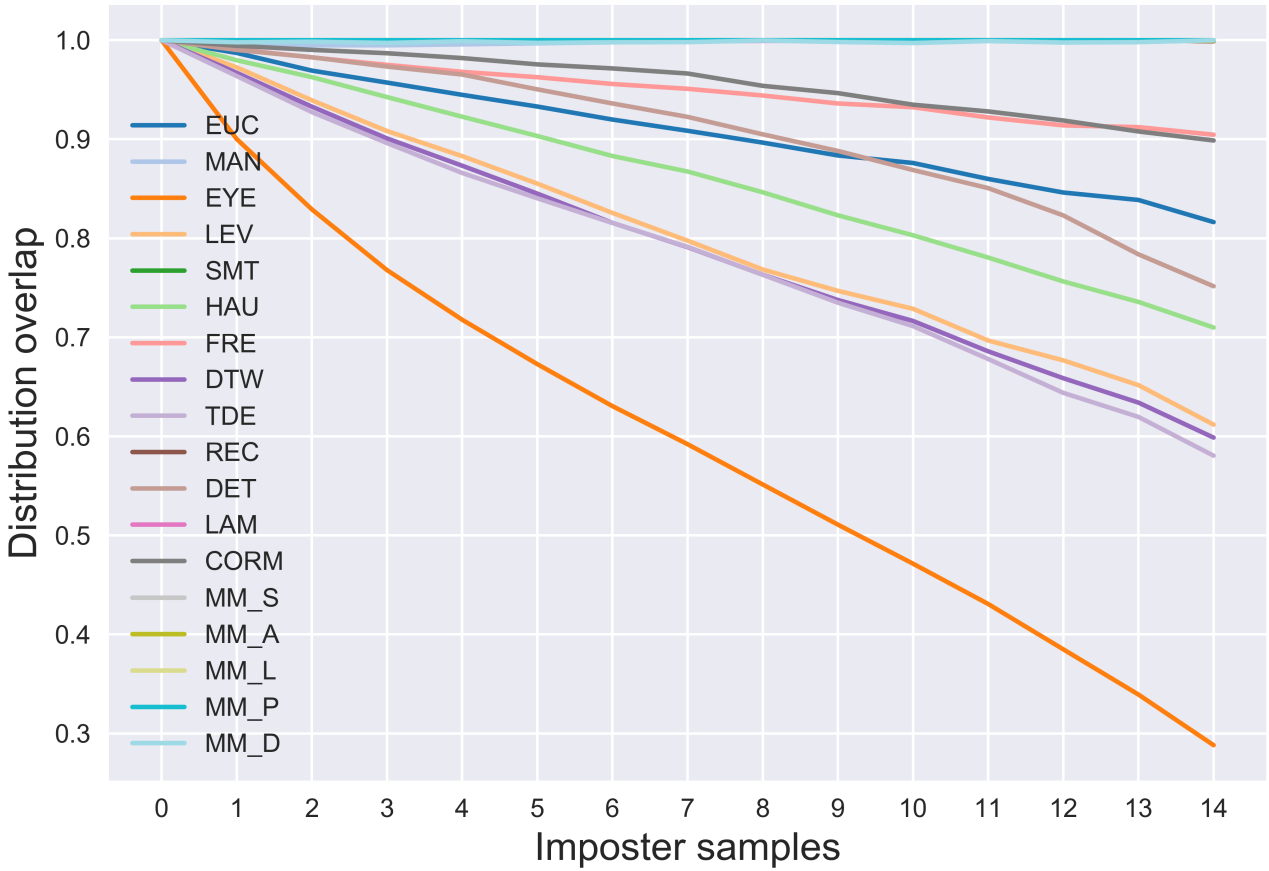


FIGURE 14 – Sensitivity to increasing number of imposter samples in measuring inter-observer distances.

Intersection

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
MEAN	0.82	1.00	0.29	0.61	1.00	0.71	0.90	0.60	0.58	1.00	0.75	1.00	0.90	1.00	1.00	1.00	1.00
SPP	0.67	1.00	0.15	0.49	1.00	0.46	0.71	0.37	0.46	1.00	0.46	1.00	1.00	1.00	0.98	1.00	1.00

TABLE 5 – Area of intersection for two distributions. One is based on distances of observers viewing the same image, and the other is between images. In practice, a strong metric should elicit a very different distribution.

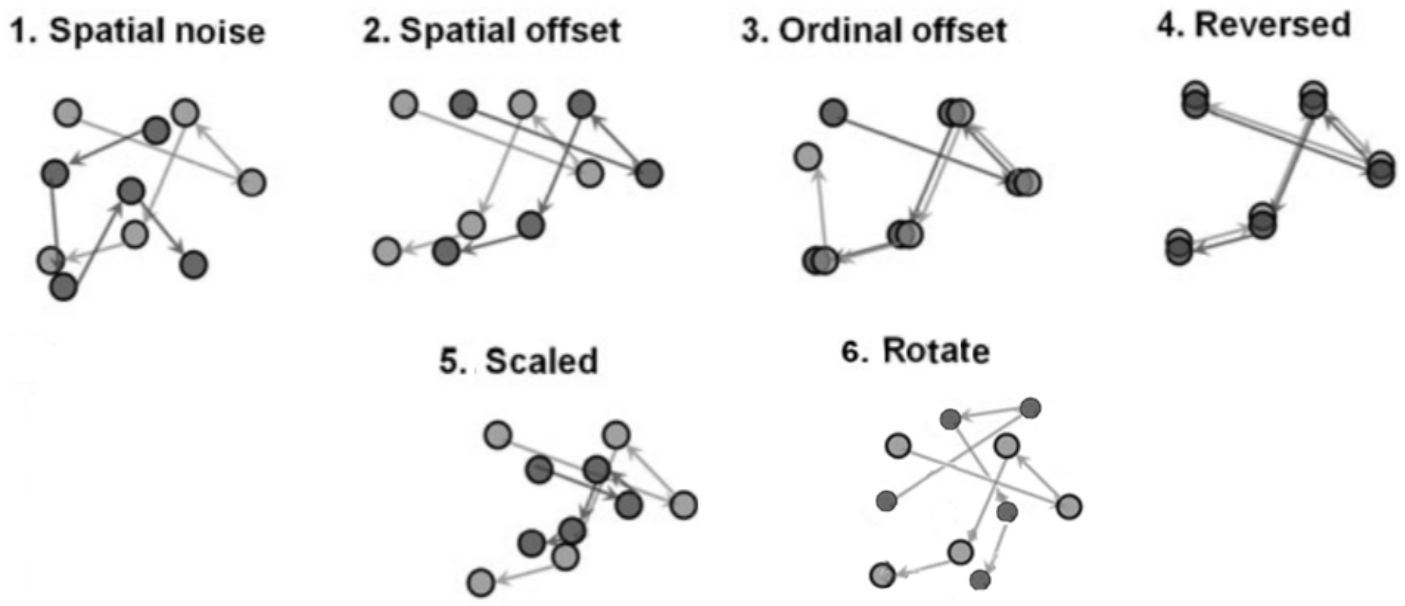


FIGURE 15 – A depiction of different conditions that may reveal sensitivity of sequential metrics. These may happen by virtue of small differences in viewing patterns, noise in data capture, or the overall stochastic nature of the process. (Derived from [8])

Spatial Noise

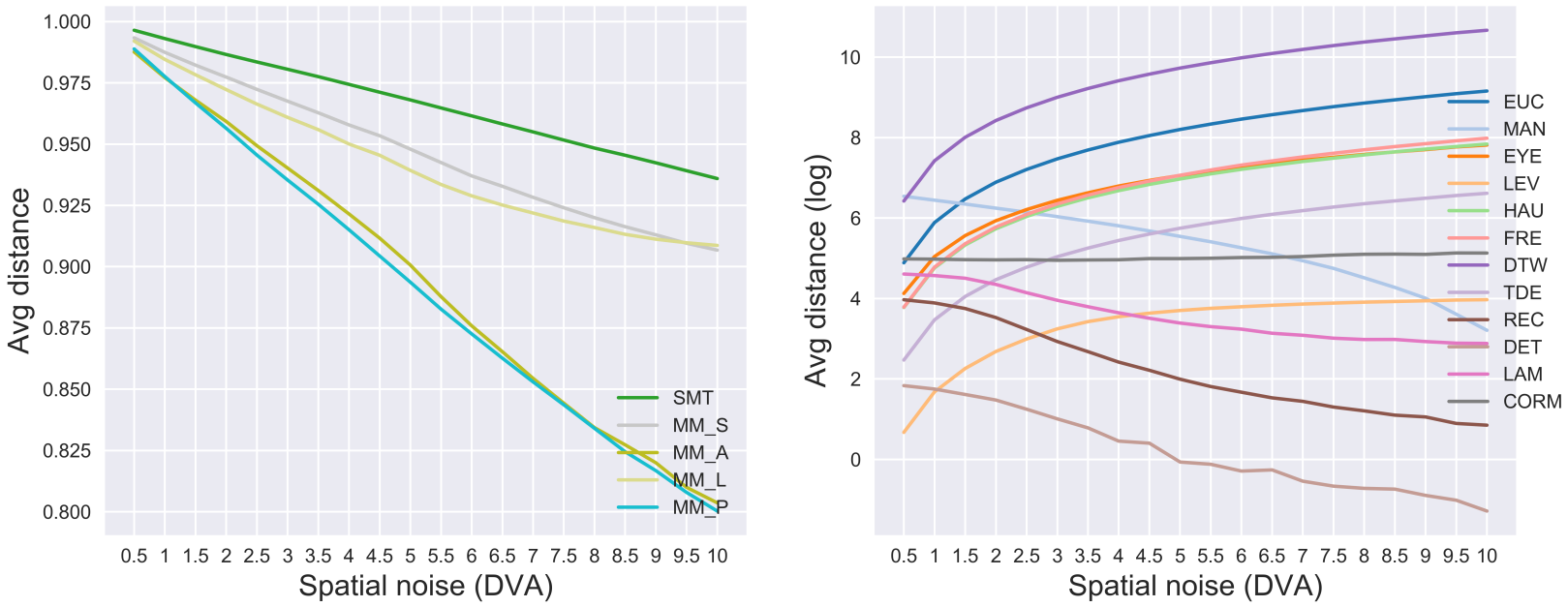


FIGURE 16 – Spatial noise : Each fixation has been moved according to a random sample from a Gaussian distribution with σ according to degree of visual angle

Spatial offset

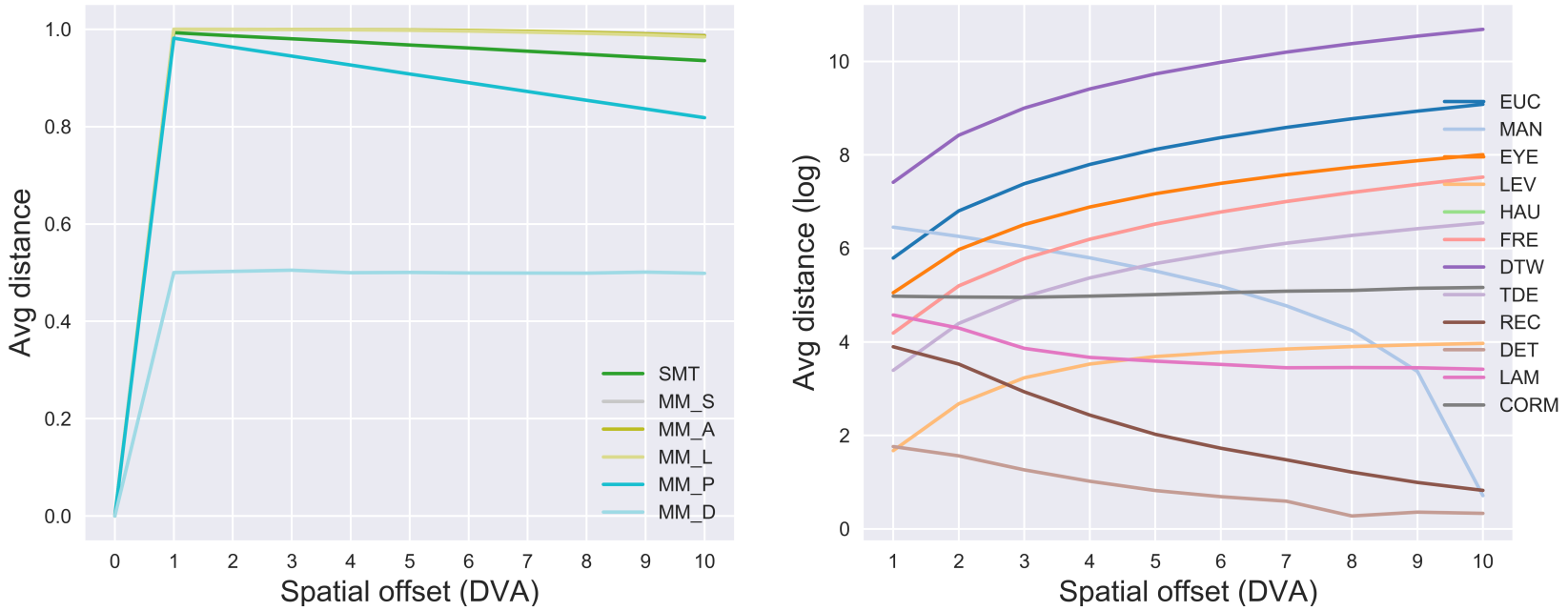


FIGURE 17 – Spatial offset : all of fixation have been moved according to a random sample from a Gaussian distribution with σ according to degree of visual angle

Ordinal offset

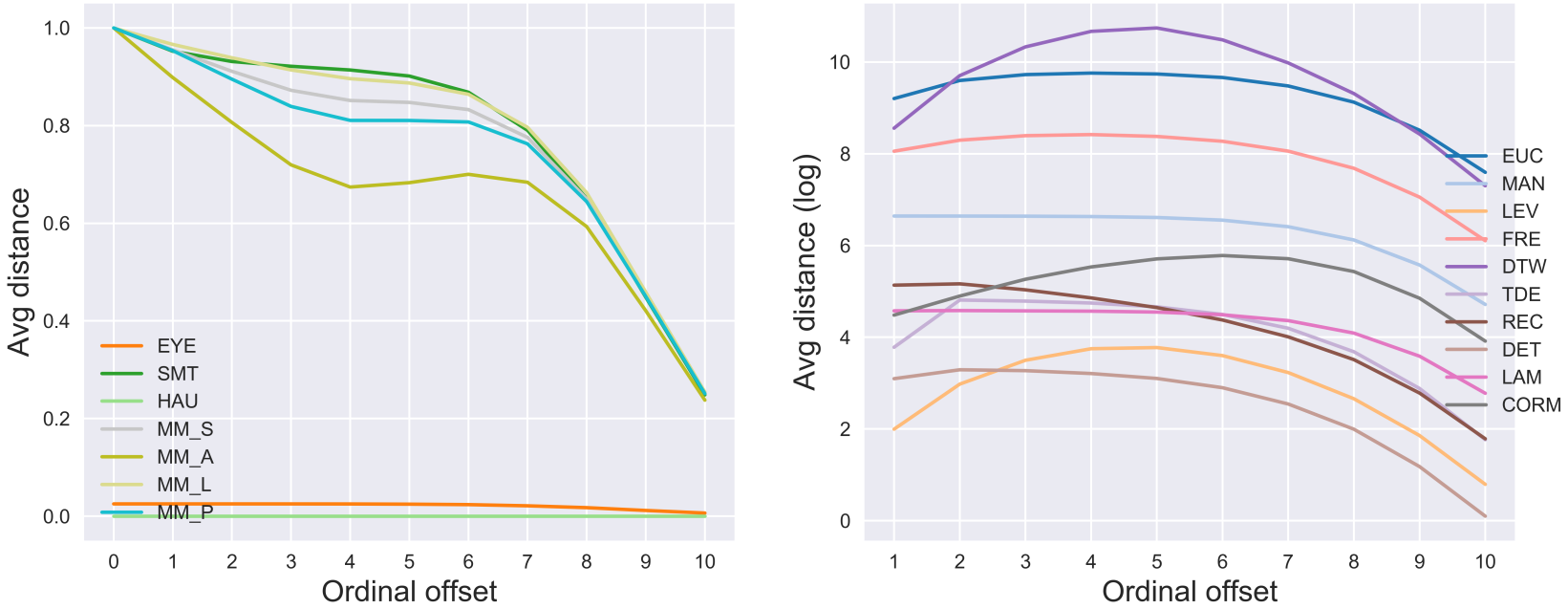


FIGURE 18 – Ordinal offset : average distance when scanpaths have temporally shifted in clockwise

Reverse ordinal offset

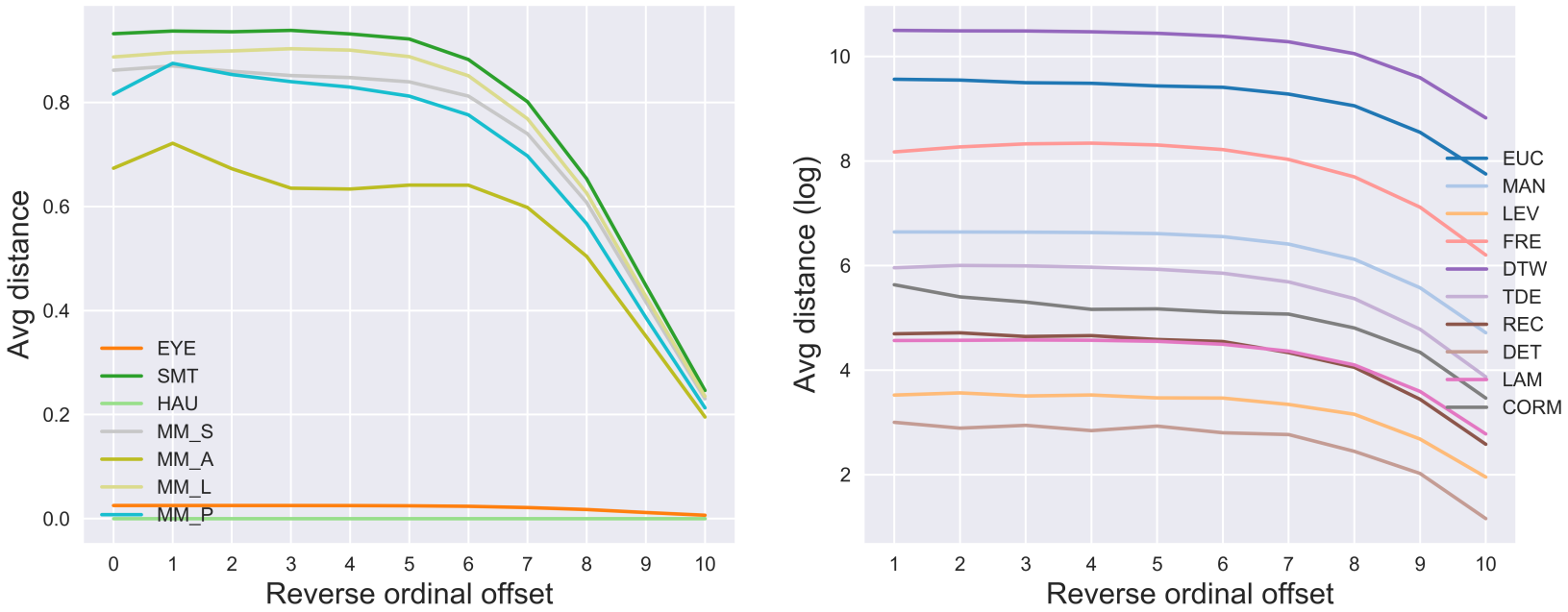


FIGURE 19 – Reverse Ordinal offset : average distance when scanpaths have temporally shifted in counter-clockwise

Rotation

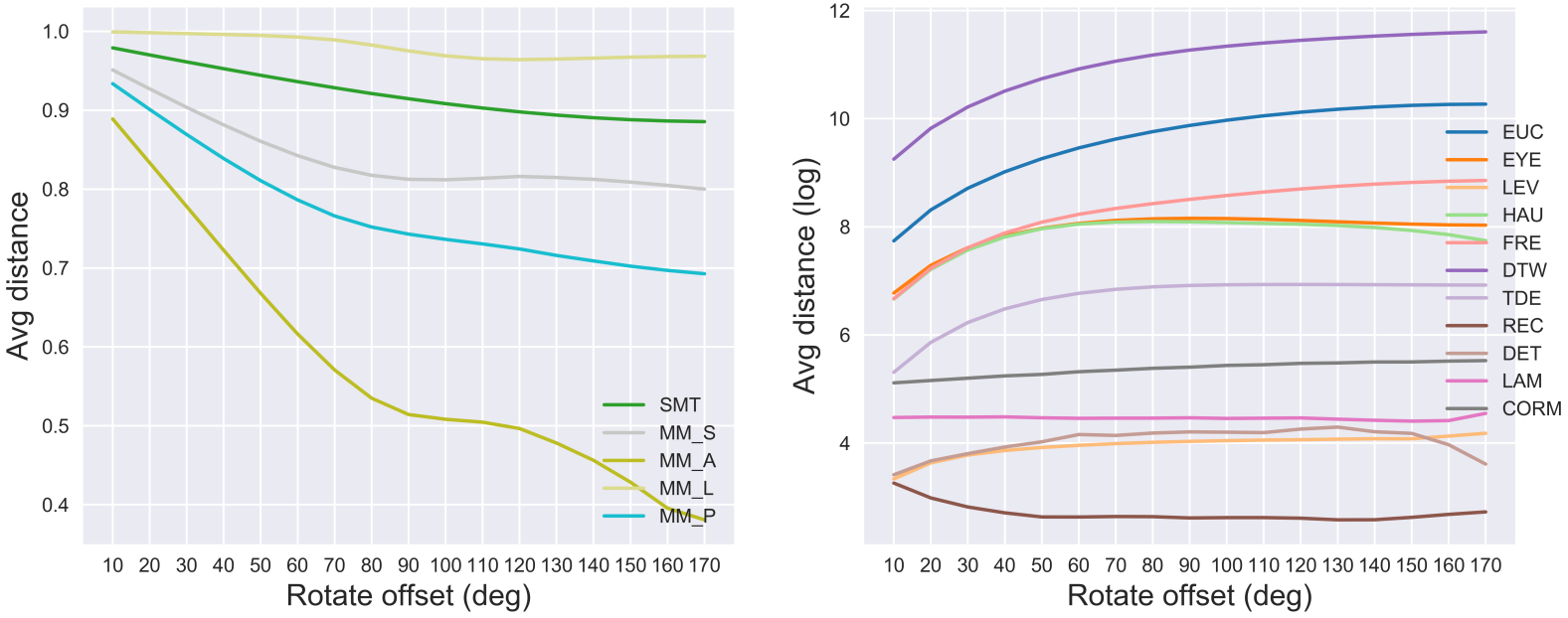


FIGURE 20

Similarities in ranking | I

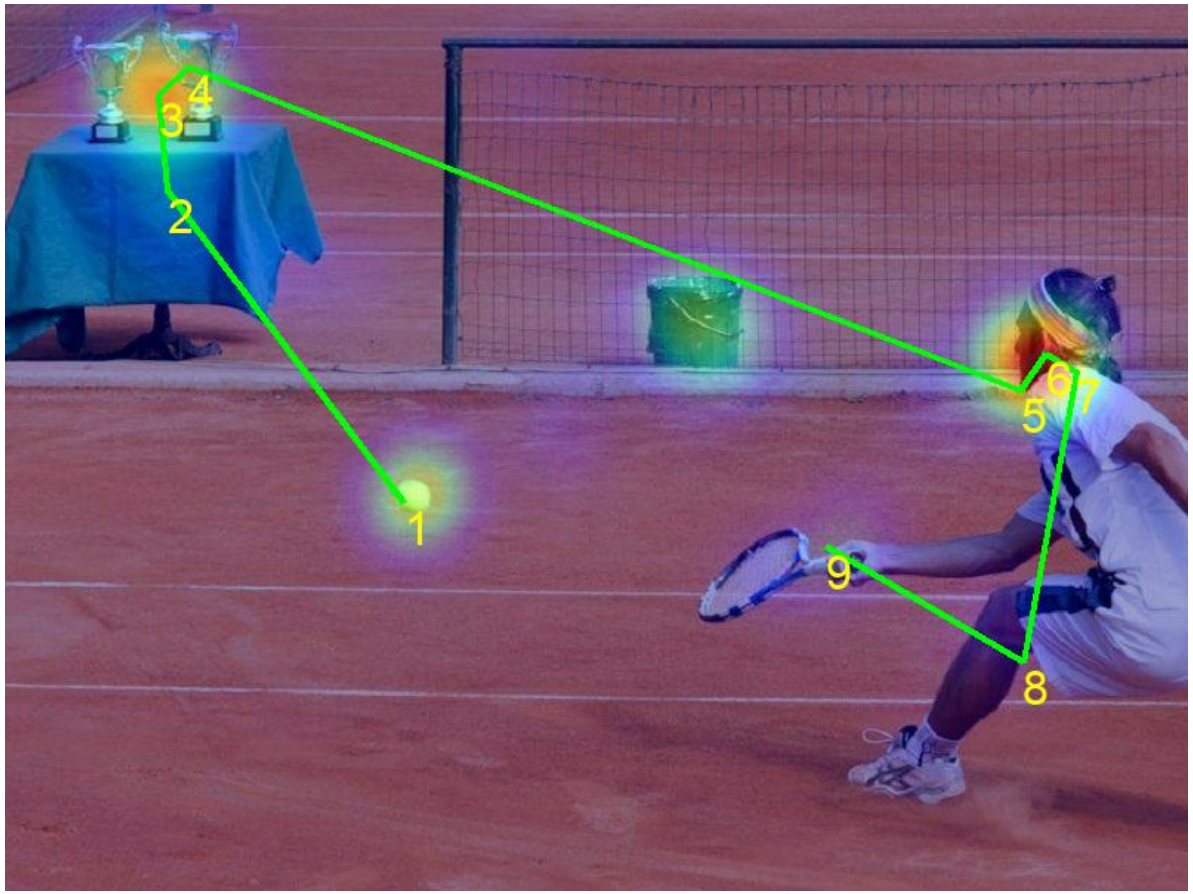


FIGURE 21 – Base scanpath

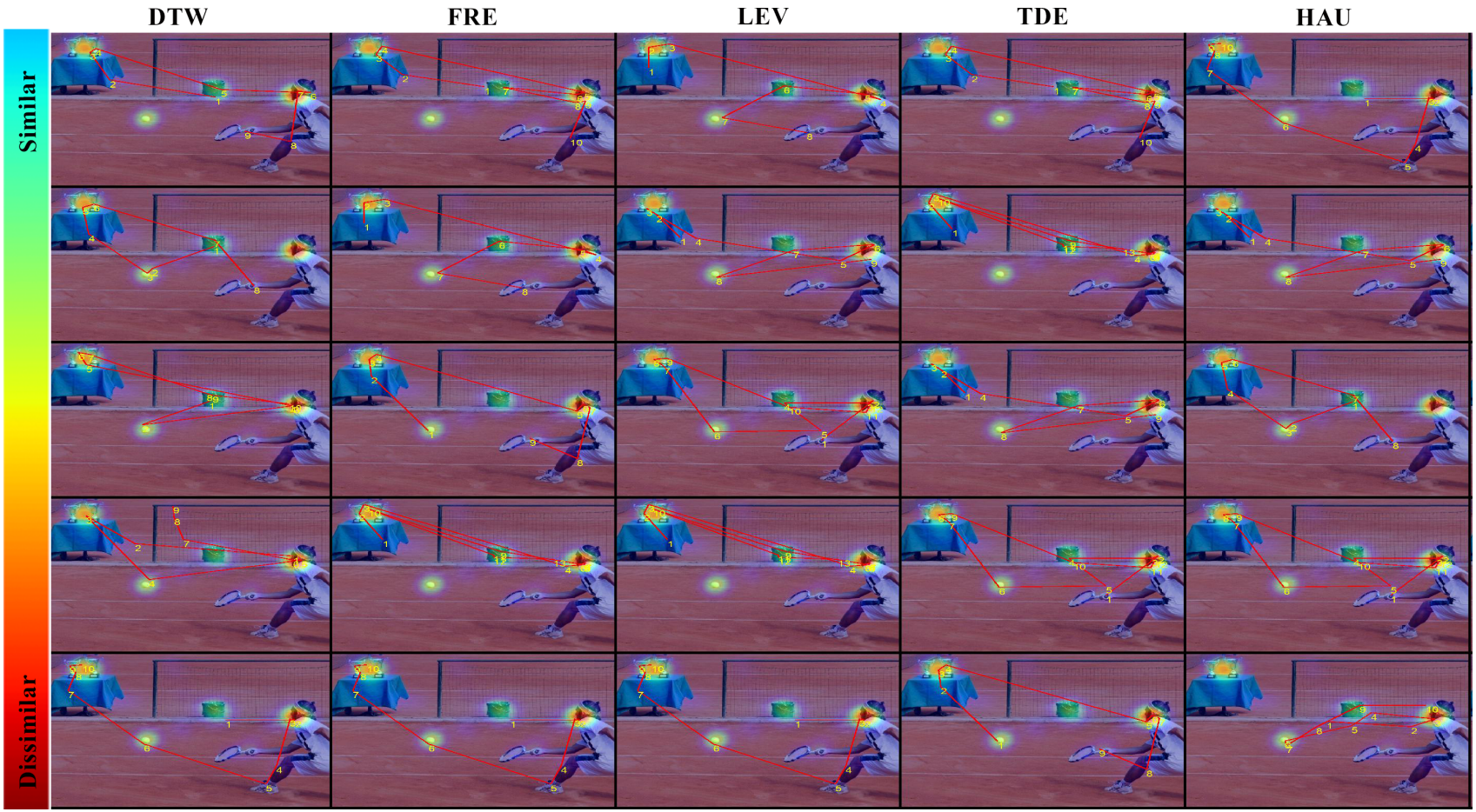


FIGURE 22 – Base scanpath

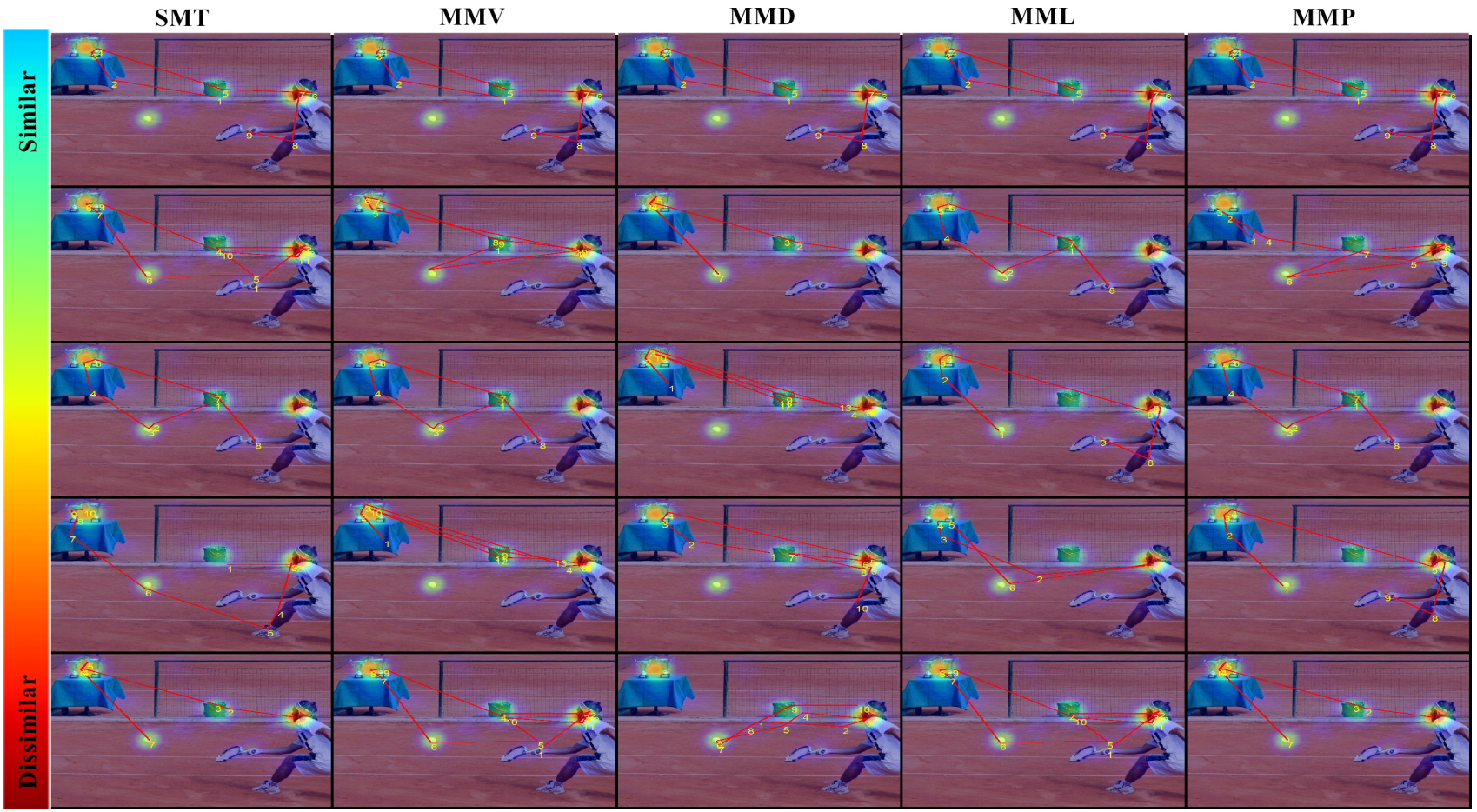
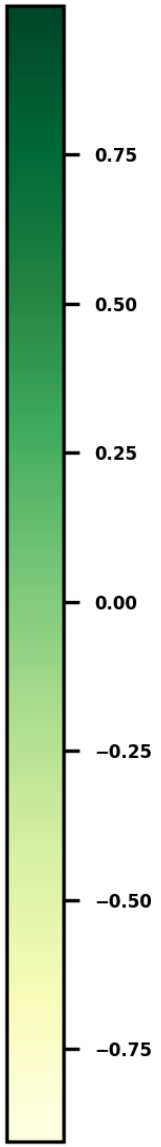


FIGURE 23 – Base scanpath

Analysis

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	COR	MM_S	MM_A	MM_L	MM_P
EUC	1.00	-0.29	0.28	0.48	-0.38	0.17	0.67	0.80	0.43	-0.45	-0.03	0.02	-0.01	-0.53	0.01	-0.44	-0.53
MAN	-0.29	1.00	-0.91	-0.35	0.46	-0.89	-0.38	-0.52	-0.46	0.35	0.08	0.08	0.12	0.27	0.22	0.23	0.41
EYE	0.28	-0.91	1.00	0.31	-0.58	0.77	0.36	0.50	0.59	-0.45	-0.16	-0.19	-0.06	-0.27	-0.23	-0.24	-0.44
LEV	0.48	-0.35	0.31	1.00	-0.57	0.18	0.25	0.70	0.16	-0.52	-0.12	-0.08	-0.07	-0.14	-0.06	-0.09	-0.34
SMT	-0.38	0.46	-0.58	-0.57	1.00	-0.33	-0.36	-0.63	-0.65	0.48	0.25	0.21	-0.21	0.33	0.23	0.21	0.51
HAU	0.17	-0.89	0.77	0.18	-0.33	1.00	0.39	0.36	0.38	-0.16	-0.02	-0.03	-0.09	-0.27	-0.23	-0.21	-0.33
FRE	0.67	-0.38	0.36	0.25	-0.36	0.39	1.00	0.69	0.44	-0.22	0.03	0.02	0.00	-0.60	-0.18	-0.46	-0.51
DTW	0.80	-0.52	0.50	0.70	-0.63	0.36	0.69	1.00	0.49	-0.55	-0.09	-0.02	-0.01	-0.50	-0.17	-0.36	-0.64
TDE	0.43	-0.46	0.59	0.16	-0.65	0.38	0.44	0.49	1.00	-0.43	-0.27	-0.19	0.10	-0.41	-0.16	-0.27	-0.49
REC	-0.45	0.35	-0.45	-0.52	0.48	-0.16	-0.22	-0.55	-0.43	1.00	0.59	0.17	0.13	0.13	-0.03	0.10	0.33
DET	-0.03	0.08	-0.16	-0.12	0.25	-0.02	0.03	-0.09	-0.27	0.59	1.00	0.27	-0.04	-0.05	-0.05	-0.03	0.05
LAM	0.02	0.08	-0.19	-0.08	0.21	-0.03	0.02	-0.02	-0.19	0.17	0.27	1.00	-0.04	-0.07	0.00	0.02	0.06
COR	-0.01	0.12	-0.06	-0.07	-0.21	-0.09	0.00	-0.01	0.10	0.13	-0.04	-0.04	1.00	-0.02	-0.02	-0.00	-0.04
MM_S	-0.53	0.27	-0.27	-0.14	0.33	-0.27	-0.60	-0.50	-0.41	0.13	-0.05	-0.07	-0.02	1.00	0.53	0.62	0.66
MM_A	0.01	0.22	-0.23	-0.06	0.23	-0.23	-0.18	-0.17	-0.16	-0.03	-0.05	0.00	-0.02	0.53	1.00	0.12	0.44
MM_L	-0.44	0.23	-0.24	-0.09	0.21	-0.21	-0.46	-0.36	-0.27	0.10	-0.03	0.02	-0.00	0.62	0.12	1.00	0.41
MM_P	-0.53	0.41	-0.44	-0.34	0.51	-0.33	-0.51	-0.64	-0.49	0.33	0.05	0.06	-0.04	0.66	0.44	0.41	1.00



The benchmark

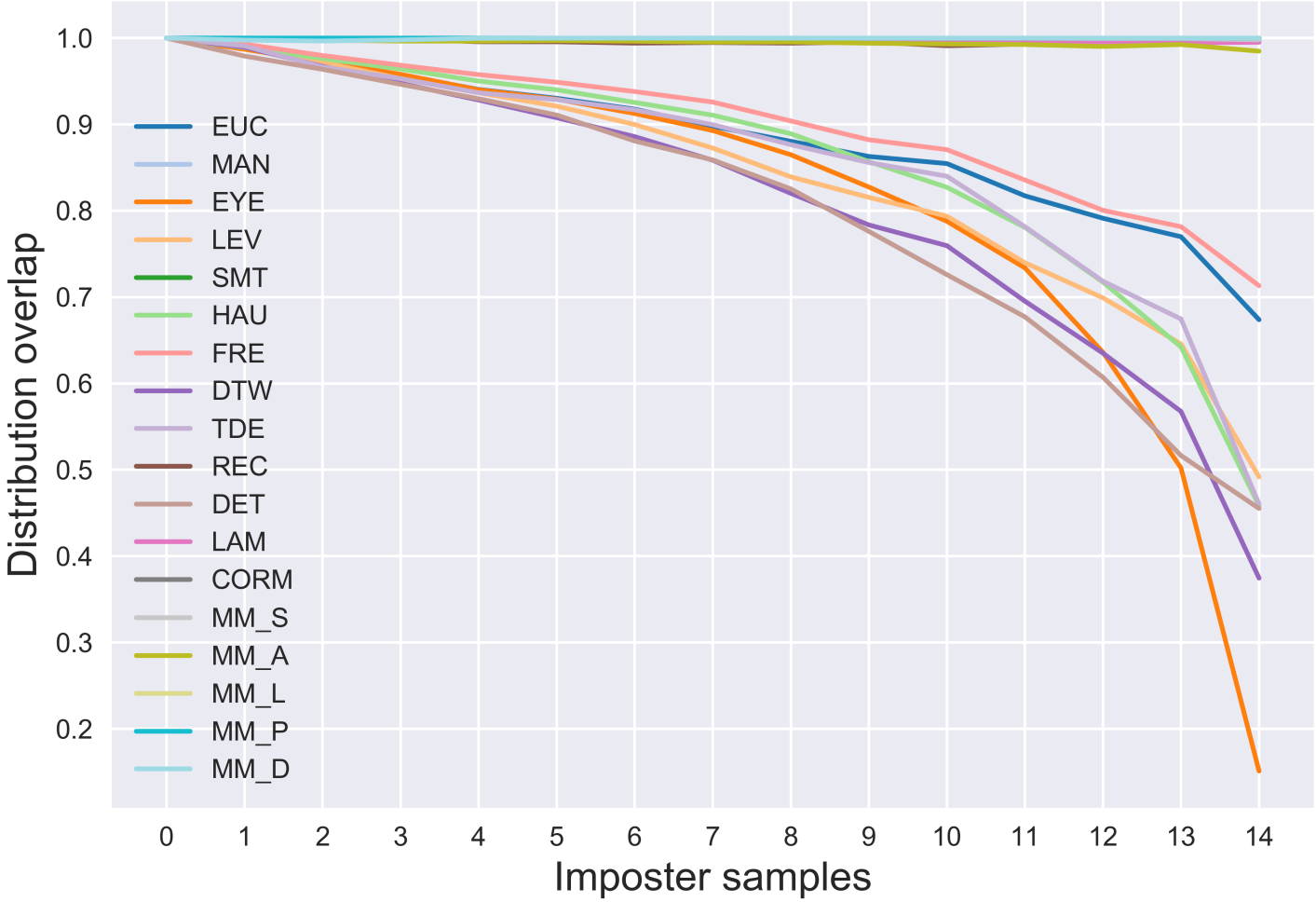
Sequential bechmark

Model	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
AIM	831.78	23.27	168.98	15.35	0.76	242.37	386.23	2032.22	113.64	4.33	0.66	8.27	40.86	0.84	0.67	0.86	0.77
AWS	828.08	23.48	164.34	15.11	0.76	244.82	383.03	2010.82	113.07	4.90	0.83	9.25	40.15	0.84	0.67	0.86	0.77
CAS	761.36	24.29	167.34	15.19	0.77	230.77	353.70	1910.34	102.37	4.63	0.93	7.94	41.36	0.85	0.66	0.88	0.79
CVS	898.08	1.83	235.55	16.07	0.76	291.37	413.44	2353.33	128.56	2.70	0.45	6.42	42.24	0.84	0.67	0.86	0.76
DVA	847.99	19.93	176.01	15.48	0.76	251.78	392.54	2088.77	116.76	4.13	0.70	8.35	39.27	0.84	0.67	0.86	0.76
GBVS	707.28	32.96	147.79	14.85	0.77	202.24	323.15	1730.56	94.50	5.64	1.13	8.41	41.84	0.86	0.67	0.89	0.81
IKN	777.06	25.81	164.95	15.20	0.77	228.45	358.76	1913.76	105.72	4.47	0.75	7.73	41.67	0.85	0.67	0.88	0.79
IMSIG	796.70	23.28	166.60	15.12	0.77	236.84	370.47	1959.85	105.25	5.01	0.94	8.72	40.49	0.85	0.67	0.87	0.78
QSS	909.77	9.17	200.28	15.77	0.76	287.47	424.04	2267.87	130.19	3.77	0.65	8.09	40.51	0.84	0.67	0.86	0.76
SSR	874.44	15.22	188.71	15.54	0.76	262.36	408.29	2159.36	120.03	4.16	0.76	7.98	40.24	0.84	0.67	0.87	0.77
SUN	850.72	16.39	185.09	15.59	0.76	256.88	393.81	2117.18	117.92	3.83	0.69	8.13	40.05	0.84	0.67	0.86	0.76
cG	599.08	18.56	187.09	15.11	0.78	220.50	286.53	1693.72	87.20	4.46	1.14	7.03	45.62	0.85	0.49	0.85	0.81
SAM-VGG	728.45	47.33	106.02	14.03	0.77	187.48	340.53	1600.90	94.71	6.10	1.06	10.60	39.63	0.86	0.70	0.88	0.82
OpenSalicon	730.94	41.46	120.34	14.19	0.77	200.00	348.58	1639.66	91.99	6.57	1.11	10.83	38.31	0.85	0.69	0.88	0.82
SALGAN	1112.63	14.12	213.76	15.66	0.76	424.14	555.35	2611.46	171.14	4.92	0.81	11.04	33.30	0.83	0.69	0.83	0.77
PathGAN	1218.31	80.62	443.05	17.52	0.73	414.35	552.11	3711.61	173.52	0.17	0.17	8.10	32.01	0.67	0.48	0.61	0.60
LeMeur	427.38	109.58	116.62	12.48	0.57	128.85	182.75	1246.36	56.40	0.01	0.00	0.00	12.18	0.66	0.12	0.63	0.50

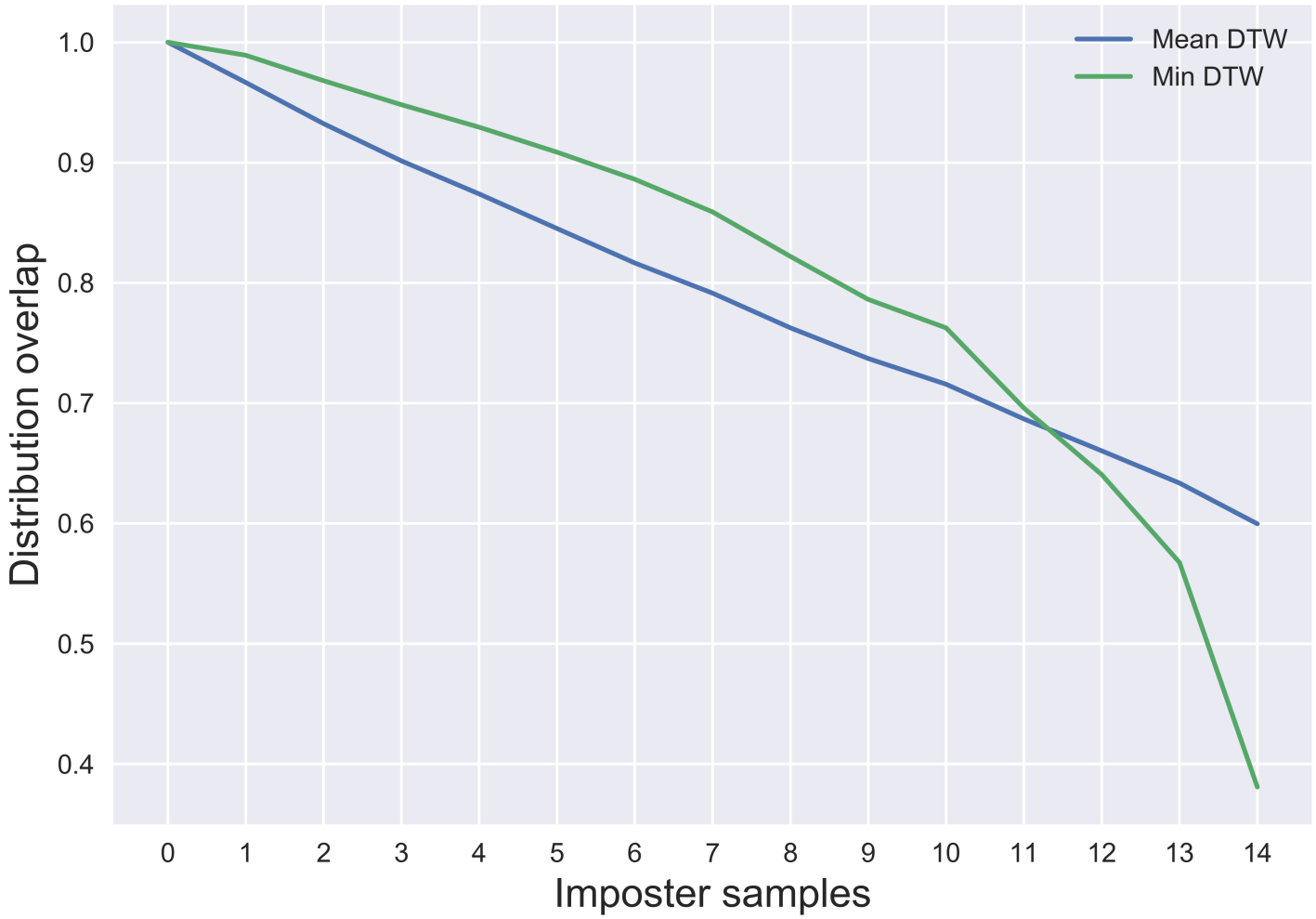
Scanpath Plausibility

- Consider a scene with k regions that tend to be gazed at
- When comparing, should we take average ?
- If N observers produce $\approx M$ different strategies for viewing an image, and a model is successful in reproducing one of these strategies, this still implies that $\frac{M-1}{M}$ of instances included in the metric are in some sense noise

SPP results


Distribution overlap
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
EUC
MAN
EYE
LEV
SMT
HAU
FRE
DTW
TDE
REC
DET
LAM
CORM
MM_S
MM_A
MM_L
MM_P
MM_D
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Imposter samples

SPP results ||


Mean DTW
Min DTW
Distribution overlap
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Imposter samples

Applications

Stimuli effect

Inter observer results per category for CAT2000 - std 11

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
Affective	1554.69	32.17	200.55	30.73	0.40	429.68	571.34	5638.05	113.33	4.00	2.75	13.08	34.18	0.90	0.70	0.91	0.86
BlackWhite	1468.51	31.68	210.74	30.06	0.39	431.53	558.50	5470.13	116.43	3.59	2.34	11.85	34.55	0.90	0.70	0.91	0.86
Fractal	1632.23	26.11	227.07	31.16	0.37	452.68	598.80	6069.34	127.71	3.16	2.10	11.34	34.82	0.89	0.70	0.90	0.84
Inverted	1703.50	25.84	225.79	31.94	0.37	442.19	595.13	6287.62	129.76	2.58	1.91	10.63	34.67	0.89	0.71	0.91	0.84
Line Drawing	1683.26	25.44	229.55	32.54	0.35	441.17	593.74	6444.01	133.39	2.60	1.75	10.56	34.89	0.89	0.69	0.91	0.84
Noisy	1208.06	33.65	220.39	27.44	0.36	440.41	536.13	4771.37	116.06	4.22	2.06	11.79	36.71	0.89	0.65	0.90	0.86
Outdoor	1513.83	30.19	212.57	30.72	0.38	438.72	581.14	5697.78	118.53	3.50	2.21	11.38	34.89	0.89	0.70	0.90	0.85
Pattern	1375.97	29.68	223.07	29.57	0.37	445.03	558.78	5400.19	120.59	3.66	2.03	11.42	35.92	0.89	0.67	0.91	0.85
Satellite	1547.59	26.45	231.41	30.17	0.36	454.31	588.80	5813.68	128.96	3.06	1.96	11.11	34.96	0.89	0.69	0.91	0.85
Social	1463.49	37.10	187.82	30.25	0.41	396.76	536.32	5327.27	107.14	4.09	2.73	13.00	34.43	0.90	0.71	0.92	0.87

Resolution I

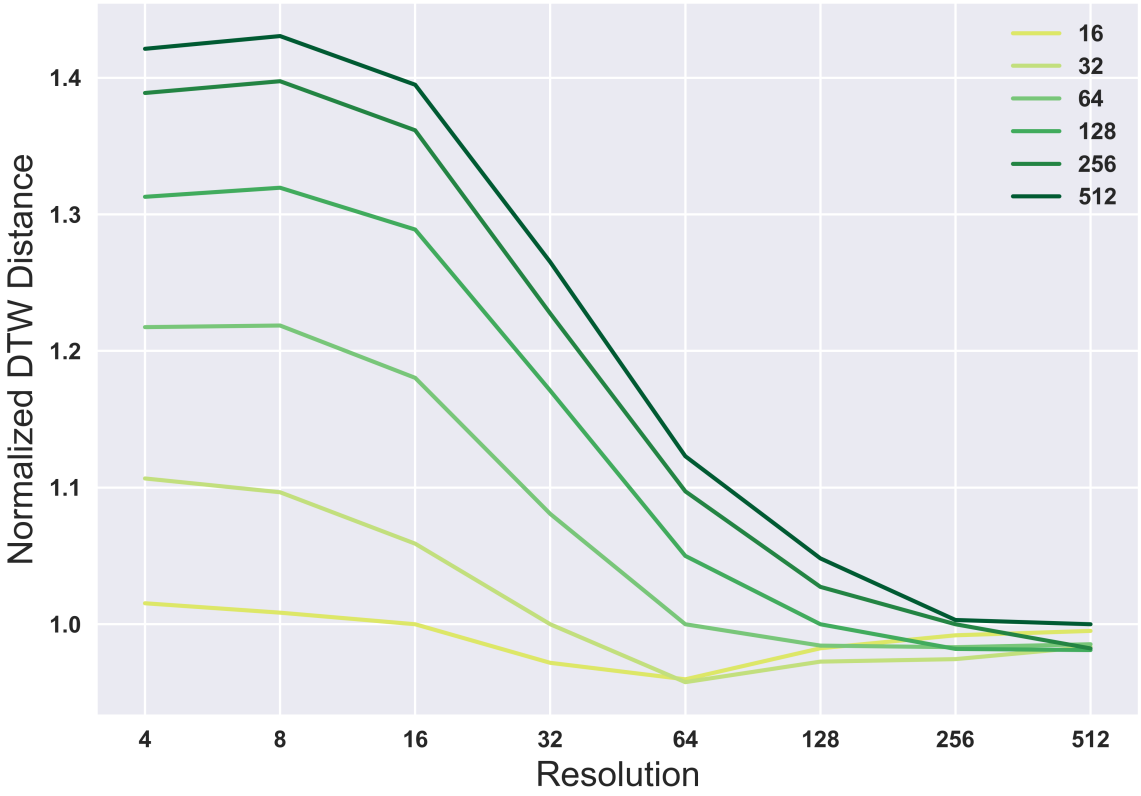


FIGURE 25 – Inter observer score for different resolutions of an image. Results have been normalized normalized by the inter-observer distance of the base resolution.

Resolution II

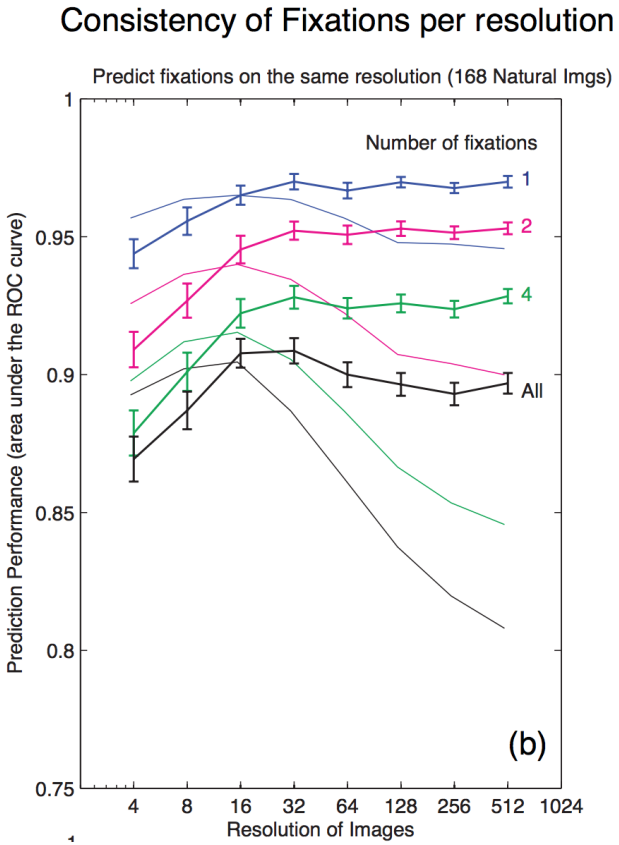


FIGURE 26 – [15]

Eye-tracking vs mouse-tracking

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
mouse-mouse	703.29	27.26	136.40	10.75	0.77	179.95	314.08	1446.57	89.85	4.56	1.27	0.87	34.60	0.71	0.62	0.73	0.66
mouse-mouse spp	432.03	20.81	83.52	7.89	0.72	94.09	188.37	906.49	58.87	0.06	0.00	0.00	12.35	0.62	0.41	0.61	0.54
Mouse-eye	824.88	8.72	213.97	15.97	0.76	256.13	370.25	2145.35	121.03	6.45	23.08	17.64	44.92	0.83	0.66	0.85	0.76
Mouse-eye SPP	591.79	49.19	143.45	13.26	0.74	169.53	247.31	1528.37	85.19	4.10	22.67	13.37	18.58	0.74	0.42	0.70	0.64

TABLE 7 – Mouse tracking vs Eye tracking data on OSIE dataset. This shows a remarkable difference in considering within and between class distances, and provides strong evidence for the view that mouse-tracking proxies can't help saccadic models as effectively as static models.

Saliency in time

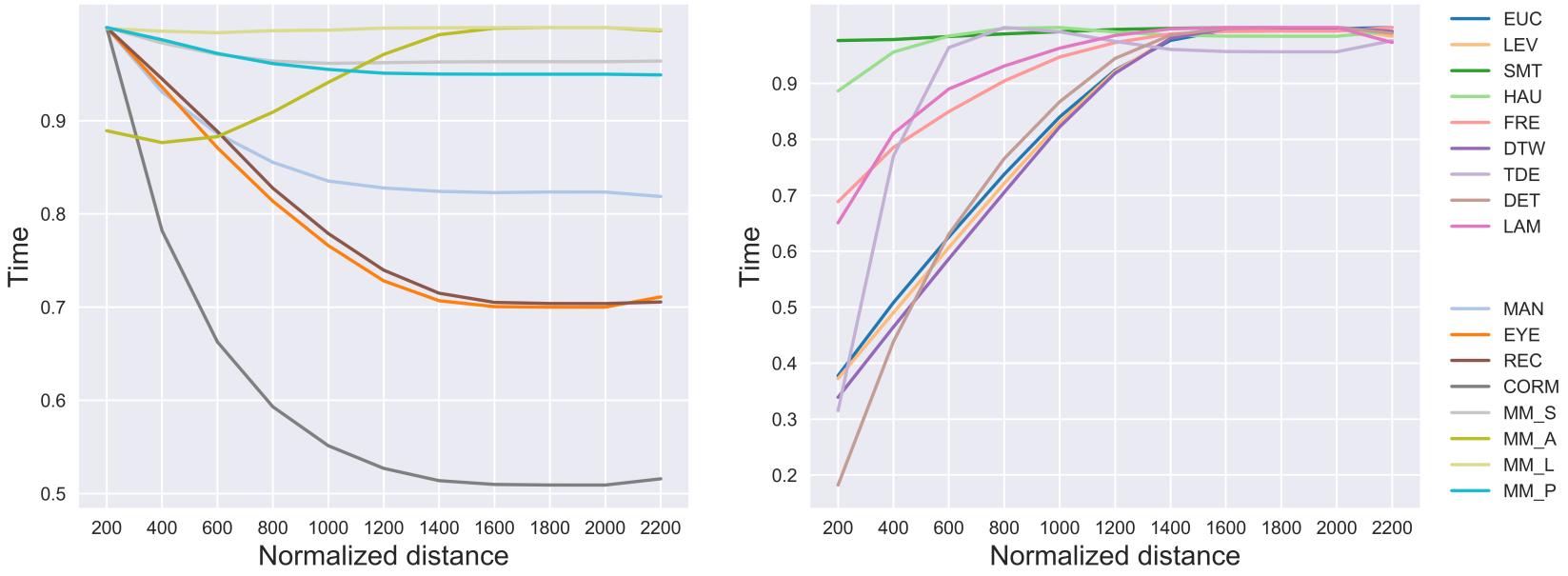


FIGURE 27 – Inter observer distance by quantizing according to duration of viewing and normalized by maximum IOC distance pr metric

Summary

- Review of existing dataset
- Comprehensive study of sequential metrics
- Careful and empirical study of metrics
- Making the first benchmark of saccadic models
- Review of possible applicable areas

Limitations & future work

- Top-down vs bottom up
- Memory
- Emotion
- Human vs Machine
- Revisiting the literature

Thank you

Thank you !

https ://github.com/rAm1n/msc-thesis

References I

American academy of ophthalmology.
https://www.aao.org/, Sept. 2018.

N. C. Anderson, W. F. Bischof, K. E. W. Laidlaw, E. F. Risko, and A. Kingstone.
Recurrence quantification analysis of eye movements.
Behavior Research Methods, 45(3) :842–856, Sep 2013.

D. J. Berndt.
Using dynamic time warping to find patterns in time series.
In *KDD workshop*, volume 10, pages 359–370. Seattle, WA, 1994.

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

Z. Bylinskii, P. Isola, C. Bainbridge, A. Torralba, and A. Oliva.
Intrinsic and extrinsic effects on image memorability.
Vision research, 116 :165–178, 2015.

Z. Bylinskii, T. Judd, A. Oliva, A. Torralba, and F. Durand.
What do different evaluation metrics tell us about saliency models ?
IEEE transactions on pattern analysis and machine intelligence, 2018.

References II

F. Cristino, S. Mathôt, J. Theeuwes, and I. D. Gilchrist.
Scanmatch : A novel method for comparing fixation sequences.
Behavior research methods, 42(3) :692–700, 2010.

R. Dewhurst, M. Nyström, H. Jarodzka, T. Foulsham, R. Johansson, and K. Holmqvist.
It depends on how you look at it : Scanpath comparison in multiple dimensions with multimatch, a vector-based approach.
Behavior research methods, 44(4) :1079–1100, 2012.

T. Eiter and H. Mannila.
Computing discrete fréchet distance.
Technical report, Citeseer, 1994.

N. Guo, M. Ma, W. Xiong, L. Chen, and N. Jing.
An efficient query algorithm for trajectory similarity based on fréchet distance threshold.
ISPRS International Journal of Geo-Information, 6(11) :326, 2017.

References III

K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer.
Eye tracking : A comprehensive guide to methods and measures.
OUP Oxford, 2011.

D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge.
Comparing images using the hausdorff distance.
IEEE Transactions on pattern analysis and machine intelligence, 15(9) :850–863, 1993.

M. Jiang, S. Huang, J. Duan, and Q. Zhao.
Salicon : Saliency in context.
In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 1072–1080. IEEE, 2015.

M. Jiang, J. Xu, and Q. Zhao.
Saliency in crowd.
In *European Conference on Computer Vision*, pages 17–32. Springer, 2014.

T. Judd, F. Durand, and A. Torralba.
Fixations on low resolution images.
Journal of Vision, 10(7) :142–142, 2010.

References IV

D. Kaspar.
Application of directional antennas in RF-based indoor localization systems.
PhD thesis, Master's thesis, Swiss Federal Institute of Technology Zurich (ETH), 2005.

Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille.
The secrets of salient object segmentation.
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 280–287, 2014.

S. K. Mannan, K. H. Ruddock, and D. S. Wooding.
The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images.
Spatial vision, 10(3) :165–188, 1996.

S. Mathôt, F. Cristino, I. D. Gilchrist, and J. Theeuwes.
A simple way to estimate similarity between pairs of eye movement sequences.
Journal of Eye Movement Research, 5(1), 2012.

C. M. Privitera and L. W. Stark.
Algorithms for defining visual regions-of-interest : Comparison with eye fixations.
IEEE Transactions on pattern analysis and machine intelligence, 22(9) :970–982, 2000.

References V

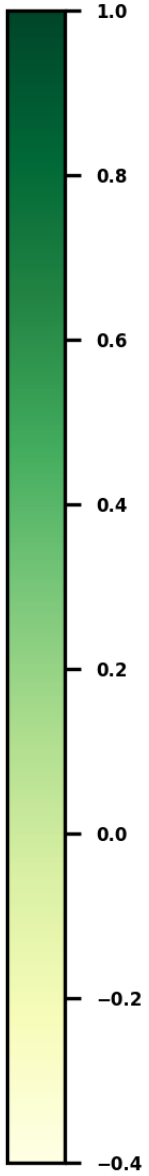
W. Wang, C. Chen, Y. Wang, T. Jiang, F. Fang, and Y. Yao.
Simulating human saccadic scanpaths on natural images.
In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 441–448. IEEE, 2011.

J. Xu, M. Jiang, S. Wang, M. S. Kankanhalli, and Q. Zhao.
Predicting human gaze beyond pixels.
Journal of vision, 14(1) :28–28, 2014.

Frame Title

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P	MM_D
0	703.85	44.50	101.73	13.97	0.45	196.94	326.33	1559.30	76.91	24.34	8.02	26.48	39.39	0.86	0.69	0.89	0.83	0.50
1	711.85	41.39	110.30	14.18	0.43	201.49	328.95	1609.35	79.37	23.08	7.71	25.97	39.35	0.86	0.68	0.89	0.83	0.50
2	722.77	38.41	118.70	14.47	0.42	205.32	331.15	1666.83	82.14	21.79	7.43	25.45	39.31	0.86	0.68	0.89	0.82	0.50
3	730.13	35.13	127.61	14.74	0.41	209.79	333.05	1722.08	84.53	20.64	7.15	24.83	39.32	0.86	0.67	0.88	0.82	0.50
4	737.48	32.13	136.09	14.96	0.40	214.14	334.86	1771.37	86.98	19.32	6.81	24.29	39.27	0.86	0.67	0.88	0.82	0.50
5	744.68	29.01	144.46	15.21	0.38	218.35	336.20	1822.85	89.06	18.20	6.57	23.84	39.30	0.85	0.67	0.88	0.81	0.50
6	752.79	25.80	152.97	15.47	0.37	222.74	337.91	1877.78	91.14	17.17	6.35	23.32	39.31	0.85	0.66	0.88	0.81	0.50
7	759.13	22.99	161.08	15.72	0.36	225.95	338.95	1923.53	93.17	16.17	6.10	22.65	39.45	0.85	0.66	0.88	0.80	0.50
8	765.75	19.82	169.88	15.98	0.35	230.27	340.36	1975.00	95.57	15.17	5.80	21.91	39.94	0.85	0.65	0.88	0.80	0.50
9	772.88	16.74	178.57	16.16	0.33	234.68	342.07	2021.15	97.93	13.99	5.44	21.31	40.07	0.85	0.65	0.88	0.79	0.50
10	776.90	13.88	186.98	16.32	0.32	238.56	342.81	2059.89	99.91	12.83	5.02	20.55	40.41	0.85	0.65	0.87	0.79	0.50
11	785.33	10.89	195.21	16.58	0.31	242.51	345.37	2113.29	102.55	11.46	4.58	19.88	40.46	0.84	0.64	0.87	0.79	0.50
12	792.32	7.83	204.33	16.74	0.30	246.62	347.13	2159.77	105.23	10.06	4.14	19.09	40.66	0.84	0.64	0.87	0.78	0.50
13	795.56	5.03	212.60	16.93	0.28	249.84	346.87	2199.74	107.08	9.16	3.79	18.40	41.07	0.84	0.63	0.87	0.78	0.50
14	806.46	2.21	220.63	17.21	0.27	253.53	348.14	2251.63	109.64	7.97	3.40	17.55	41.43	0.84	0.63	0.87	0.77	0.50

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	COR	MM_S	MM_A	MM_L	MM_P
EUC	1.00	-0.15	0.30	0.40	-0.20	0.23	0.45	0.53	0.47	-0.16	0.10	-0.10	-0.12	-0.15	0.15	-0.09	-0.28
MAN	-0.15	1.00	-0.34	-0.17	0.32	-0.36	-0.26	-0.26	-0.13	0.23	-0.07	0.05	0.20	0.12	0.04	0.19	0.23
EYE	0.30	-0.34	1.00	0.37	-0.35	0.78	0.45	0.57	0.46	-0.28	0.27	-0.21	-0.20	-0.04	0.07	-0.08	-0.19
LEV	0.40	-0.17	0.37	1.00	-0.27	0.30	0.34	0.65	0.24	-0.24	0.09	-0.10	-0.17	0.04	0.15	0.03	-0.09
SMT	-0.20	0.32	-0.35	-0.27	1.00	-0.30	-0.26	-0.30	-0.40	0.26	-0.23	0.26	0.10	0.15	0.08	0.12	0.29
HAU	0.23	-0.36	0.78	0.30	-0.30	1.00	0.45	0.48	0.32	-0.23	0.21	-0.11	-0.18	-0.05	0.03	-0.10	-0.18
FRE	0.45	-0.26	0.45	0.34	-0.26	0.45	1.00	0.75	0.42	-0.13	0.13	-0.07	-0.05	-0.13	0.12	-0.10	-0.25
DTW	0.53	-0.26	0.57	0.65	-0.30	0.48	0.75	1.00	0.47	-0.23	0.13	-0.13	-0.12	-0.03	0.18	-0.05	-0.20
TDE	0.47	-0.13	0.46	0.24	-0.40	0.32	0.42	0.47	1.00	-0.18	0.13	-0.26	-0.02	-0.07	0.13	-0.05	-0.20
REC	-0.16	0.23	-0.28	-0.24	0.26	-0.23	-0.13	-0.23	-0.18	1.00	0.24	0.18	0.62	-0.01	-0.05	0.01	0.09
DET	0.10	-0.07	0.27	0.09	-0.23	0.21	0.13	0.13	0.13	0.24	1.00	0.06	0.05	-0.07	-0.04	-0.03	-0.14
LAM	-0.10	0.05	-0.21	-0.10	0.26	-0.11	-0.07	-0.13	-0.26	0.18	0.06	1.00	0.01	-0.06	-0.05	-0.01	0.02
COR	-0.12	0.20	-0.20	-0.17	0.10	-0.18	-0.05	-0.12	-0.02	0.62	0.05	0.01	1.00	0.01	-0.03	0.02	0.08
MM_S	-0.15	0.12	-0.04	0.04	0.15	-0.05	-0.13	-0.03	-0.07	-0.01	-0.07	-0.06	0.01	1.00	0.50	0.52	0.53
MM_A	0.15	0.04	0.07	0.15	0.08	0.03	0.12	0.18	0.13	-0.05	-0.04	-0.05	-0.03	0.50	1.00	0.19	0.26
MM_L	-0.09	0.19	-0.08	0.03	0.12	-0.10	-0.10	-0.05	-0.05	0.01	-0.03	-0.01	0.02	0.52	0.19	1.00	0.36
MM_P	-0.28	0.23	-0.19	-0.09	0.29	-0.18	-0.25	-0.20	-0.20	0.09	-0.14	0.02	0.08	0.53	0.26	0.36	1.00



	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
AIM	593.92	22.19	116.54	11.61	0.58	158.59	273.56	1444.21	86.60	0.35	0.00	0.20	19.86	0.77	0.42	0.76	0.66
AWS	593.23	23.72	109.86	11.38	0.58	161.36	271.98	1407.45	85.77	0.57	0.00	0.43	18.85	0.76	0.41	0.75	0.66
CAS	532.63	21.02	112.93	11.50	0.58	152.50	242.07	1312.11	75.73	0.44	0.00	0.16	20.11	0.78	0.41	0.78	0.68
CVS	659.01	52.73	170.68	12.43	0.57	213.22	304.00	1735.46	100.06	0.16	0.00	0.06	22.62	0.77	0.43	0.75	0.65
DVA	615.52	27.60	121.59	11.84	0.58	166.80	280.03	1491.45	89.73	0.35	0.00	0.26	18.96	0.76	0.42	0.75	0.65
GBVS	487.70	7.35	100.39	11.13	0.58	130.77	215.79	1176.68	69.77	0.74	0.00	0.33	19.59	0.79	0.42	0.79	0.70
IKN	550.60	16.57	114.01	11.49	0.58	151.81	250.14	1347.33	80.05	0.43	0.00	0.21	19.38	0.77	0.42	0.77	0.67
IMSIG	562.50	23.95	110.11	11.38	0.58	152.12	258.72	1346.12	77.43	0.60	0.00	0.28	19.04	0.77	0.41	0.76	0.66
QSS	672.74	40.35	143.32	12.11	0.57	203.34	312.80	1675.07	103.47	0.27	0.00	0.13	19.73	0.76	0.42	0.75	0.65
SSR	639.80	32.53	132.01	11.83	0.58	176.81	295.52	1558.77	92.35	0.41	0.00	0.17	19.93	0.77	0.42	0.76	0.66
SUN	613.25	30.00	128.63	11.89	0.58	172.51	281.01	1506.95	90.78	0.28	0.00	0.22	19.68	0.77	0.43	0.75	0.65
cG	425.96	25.20	124.23	11.30	0.59	153.05	186.99	1147.19	60.49	0.39	0.00	0.10	22.07	0.77	0.21	0.70	0.73
SAM-VGG	502.56	6.08	66.11	10.20	0.58	96.95	227.88	1052.29	68.58	0.64	0.00	1.07	16.90	0.78	0.43	0.77	0.70
opensalicon	493.05	2.53	75.73	10.35	0.58	105.61	227.40	1064.80	66.15	0.81	0.00	0.72	16.81	0.77	0.41	0.76	0.68
salgan	815.41	69.93	158.61	11.96	0.59	337.08	436.59	2054.97	144.33	0.50	0.00	0.78	16.77	0.75	0.43	0.73	0.66
pathgan	878.71	165.25	330.67	13.88	0.55	327.80	408.63	2598.86	128.85	0.00	0.06	1.42	19.95	0.57	0.21	0.48	0.46

TABLE 9 – The benchmark - SPP results

Time	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P
600	216.10	52.81	166.30	4.18	0.42	188.70	217.59	420.02	26.95	24.41	0.97	13.87	80.40	0.89	0.57	0.89	0.88
800	290.85	49.15	155.68	5.50	0.42	203.50	248.20	575.89	65.88	23.06	2.33	17.28	62.87	0.87	0.56	0.88	0.86
1000	358.09	46.80	144.83	6.81	0.43	209.60	268.46	726.90	82.44	21.68	3.35	18.97	53.28	0.86	0.56	0.88	0.85
1200	422.14	45.17	135.30	8.10	0.43	212.52	285.75	874.13	85.51	20.20	4.07	19.85	47.69	0.86	0.58	0.88	0.84
1400	480.79	44.11	127.39	9.32	0.43	212.85	299.27	1019.08	84.85	19.01	4.61	20.53	44.33	0.85	0.60	0.88	0.84
1600	528.57	43.72	121.09	10.35	0.43	211.13	307.42	1137.91	83.39	18.05	5.02	21.01	42.37	0.86	0.62	0.89	0.83
1800	559.07	43.52	117.55	11.01	0.43	209.98	312.34	1214.99	82.16	17.45	5.24	21.26	41.31	0.86	0.63	0.89	0.83
2000	569.69	43.45	116.50	11.20	0.43	209.56	313.70	1236.92	81.83	17.21	5.30	21.31	40.99	0.86	0.64	0.89	0.83
2200	570.96	43.49	116.40	11.23	0.43	209.51	313.91	1239.15	81.79	17.18	5.31	21.31	40.94	0.86	0.64	0.89	0.83
2400	571.02	43.49	116.40	11.23	0.43	209.51	313.89	1239.25	81.79	17.18	5.31	21.31	40.94	0.86	0.64	0.89	0.83
2600	572.29	43.23	118.23	11.05	0.43	211.45	315.94	1230.42	83.50	17.22	5.26	20.75	41.47	0.86	0.63	0.88	0.83
2800	703.85	44.50	101.73	13.97	0.45	196.94	326.33	1559.30	76.91	24.34	8.02	26.48	39.39	0.86	0.69	0.89	0.83

TABLE 10 – Saliency in time - SPP results

	EUC	MAN	EYE	LEV	SMT	HAU	FRE	DTW	TDE	REC	DET	LAM	CORM	MM_S	MM_A	MM_L	MM_P	MM_D
Affective	685.96	32.24	107.11	9.30	0.14	190.35	201.39	2712.41	53.60	6.11	5.57	9.82	17.03	0.04	0.17	0.06	0.06	0.19
BlackWhite	659.40	29.71	100.18	9.12	0.14	180.38	192.49	2524.98	52.31	5.90	5.53	10.07	18.35	0.04	0.16	0.06	0.06	0.19
Fractal	761.61	31.14	100.95	9.13	0.13	183.53	197.63	2613.20	56.01	6.08	5.47	10.30	19.26	0.04	0.16	0.06	0.06	0.18
Inverted	686.92	30.29	96.57	8.58	0.13	172.23	187.76	2431.32	52.90	4.75	5.31	10.08	19.64	0.03	0.14	0.05	0.06	0.18
Line Drawing	708.23	29.50	93.04	8.78	0.12	168.59	177.14	2491.01	54.75	6.06	5.26	10.37	20.35	0.03	0.14	0.05	0.06	0.18
Noisy	622.31	30.75	100.73	9.89	0.15	178.69	197.67	2421.20	58.80	8.56	5.64	10.92	20.08	0.04	0.18	0.06	0.06	0.20
Outdoor	683.76	30.95	96.43	9.06	0.14	184.42	198.42	2465.37	52.76	6.48	5.49	10.13	18.84	0.04	0.17	0.06	0.06	0.19
Pattern	653.08	31.71	105.42	9.80	0.14	180.26	195.58	2639.35	56.26	7.19	5.43	10.55	19.83	0.04	0.16	0.06	0.06	0.19
Satellite	719.24	32.91	106.05	8.91	0.14	185.62	203.48	2628.73	58.27	5.88	5.42	10.58	19.64	0.04	0.16	0.06	0.06	0.19
Social	712.32	30.71	96.41	9.39	0.14	180.79	205.10	2691.31	49.06	6.68	5.71	9.85	17.33	0.04	0.18	0.05	0.06	0.19

TABLE 11 – Standard deviation of Inter observer distances per category for CAT2000