Summer July, 2016

Nonparametric Bayesian Dynamic Systems Analysis Applied to a Large Eyetracking Corpus

Joseph W Houpt
Mary E Frame, *Miami University - Oxford*

Nonparametric Bayesian Dynamic Systems Analysis Applied to a Large Eyetracking Corpus

Joseph W. Houpt Mary E. Frame

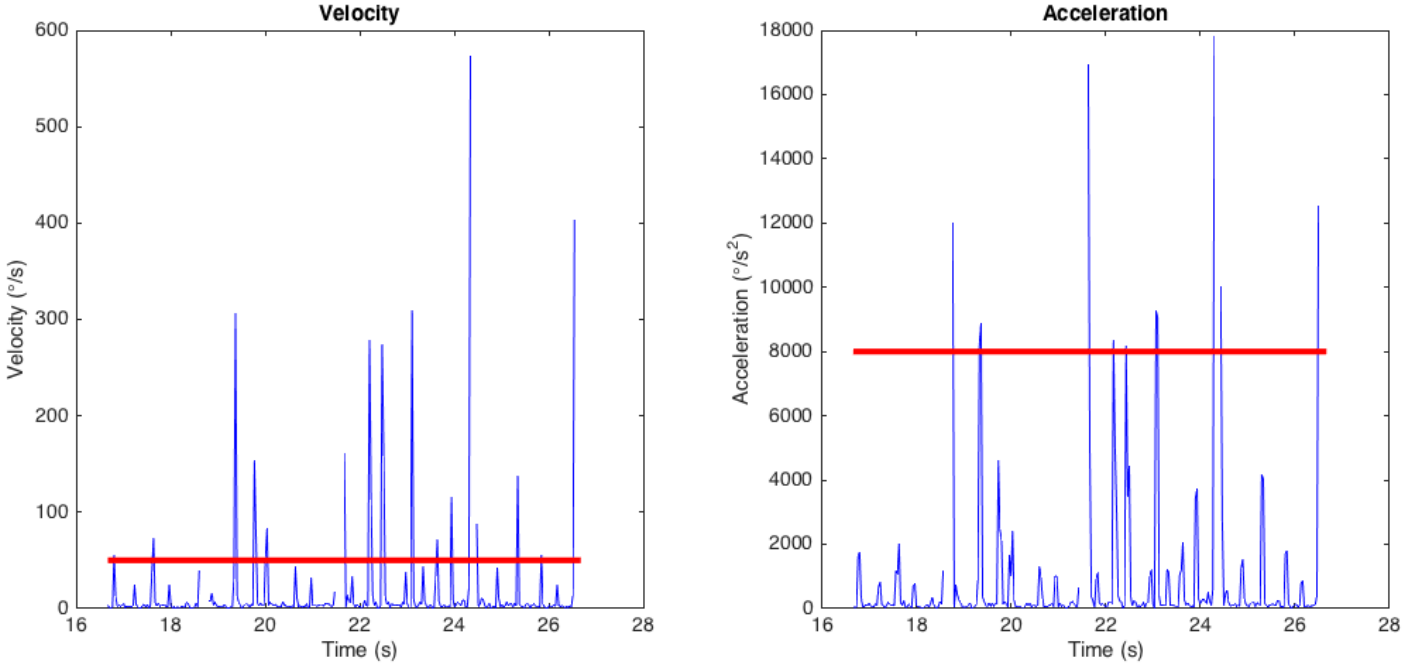
Wright State University Miami University

Society for Mathematical Psychology Annual Meeting

Outline

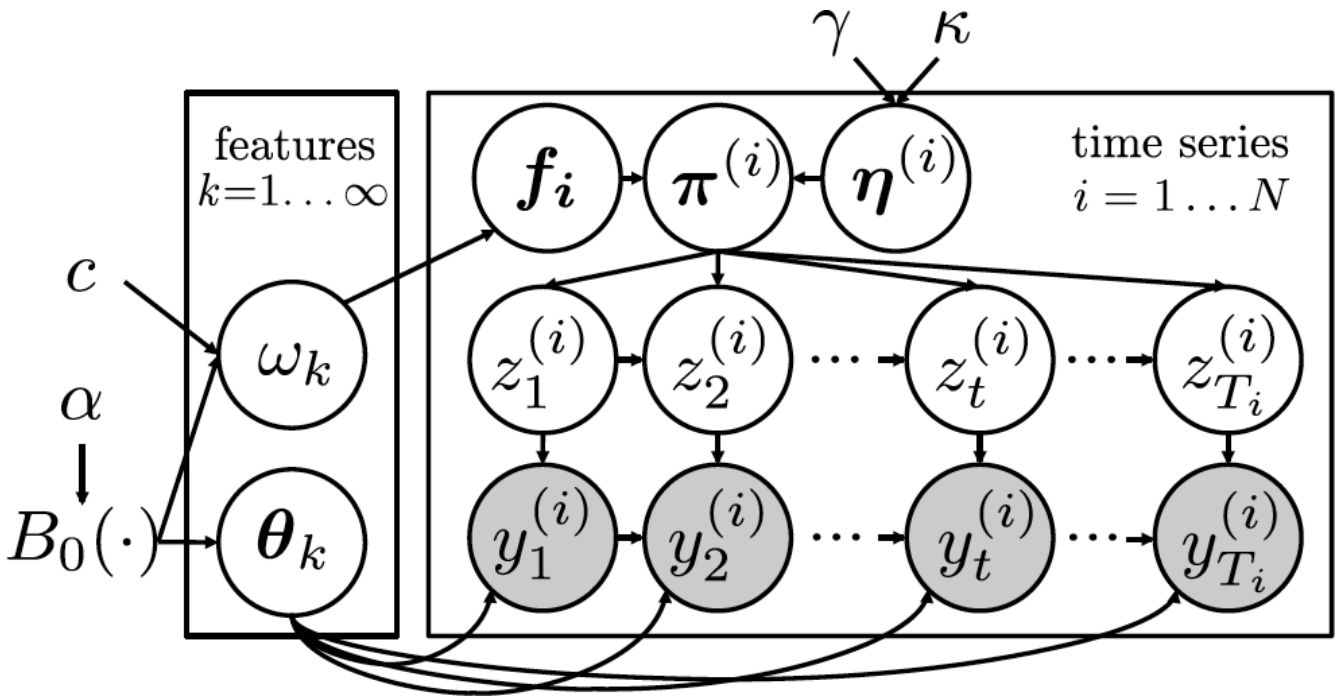

BBC HD
RBS

Existing Approaches



Outline

The BP-HMM



Fox, E. B., Hughes, M. C., Sudderth, E. B., and Jordan, M. I. (2014). Joint modeling of multiple time series via the beta process with application to motion capture segmentation. *The Annals of Applied Statistics, 8*(3), 1281-1313.

Vector Autoregressive Process

$$y_t \quad = \sum_{l \in (0,t-1)} a_l y_l + e_t \qquad e_t \sim N(0, \sigma^2)$$

Vector Autoregressive Process

$$
\begin{aligned}
y_t &= \textstyle\sum_{l \in (0,t-1)} a_l y_l + e_t & e_t &\sim N(0, \sigma^2) \\
\mathbf{y}_t &= \textstyle\sum_{l \in (0,t-1)} \mathbf{a}_l \mathbf{y}_l + \mathbf{e}_t & e_t &\sim N\left(0, \mathbf{\Sigma}\right)
\end{aligned}
$$

Vector Autoregressive Process

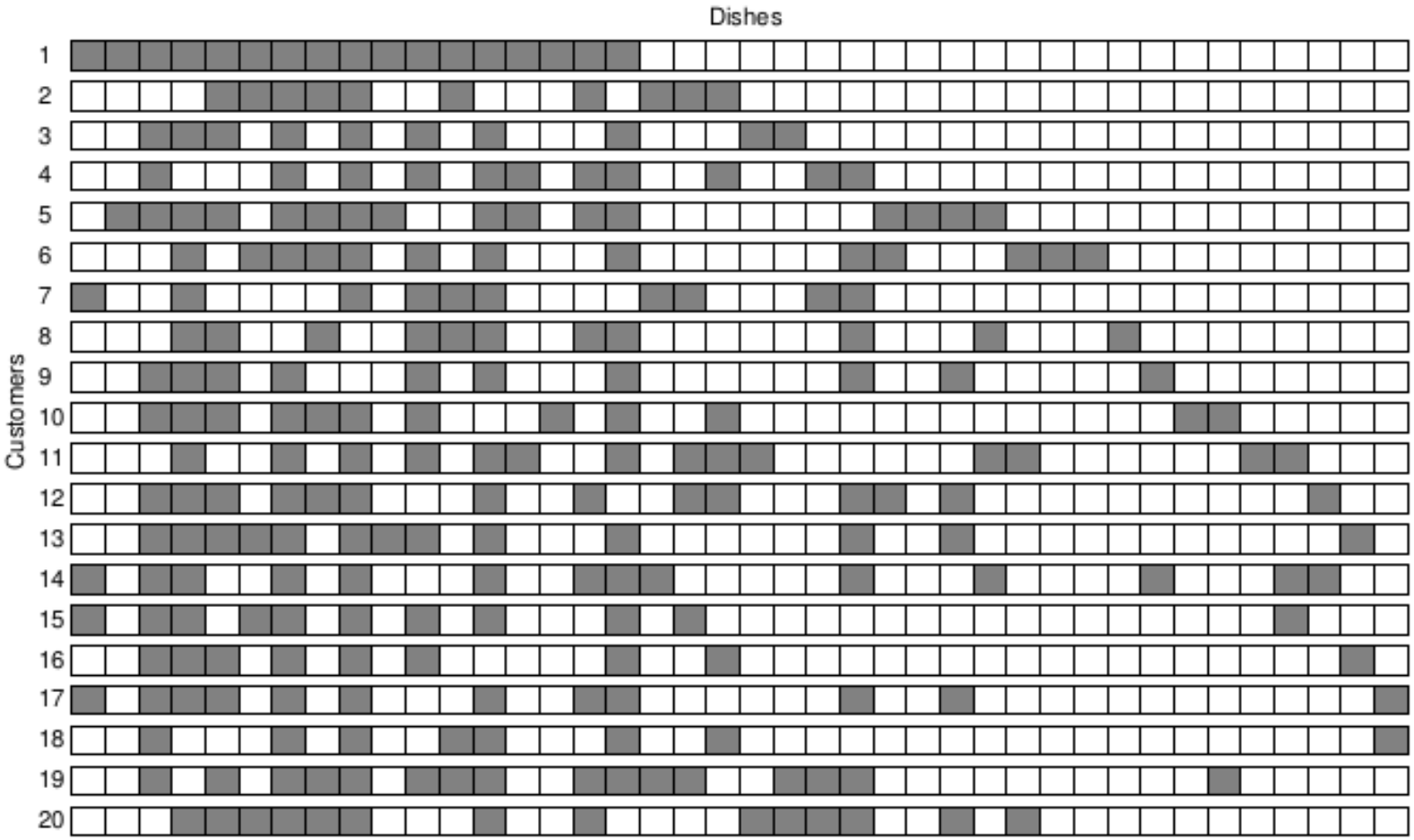
$$
\begin{aligned}
y_t &= \textstyle\sum_{l \in (0,t-1)} a_l y_l + e_t & e_t &\sim N(0, \sigma^2) \\
\mathbf{y}_t &= \textstyle\sum_{l \in (0,t-1)} \mathbf{a}_l \mathbf{y}_l + \mathbf{e}_t & e_t &\sim N\left(0, \mathbf{\Sigma}\right) \\
\mathbf{y}_t^{(i)} &= \textstyle\sum_{l \in (0,t-1)} \mathbf{a}_l \mathbf{y}_l^{(i)} + \mathbf{e}_t^{(i)} & e_t &\sim N\left(0, \mathbf{\Sigma}\right)
\end{aligned}
$$

Hidden Markov Model

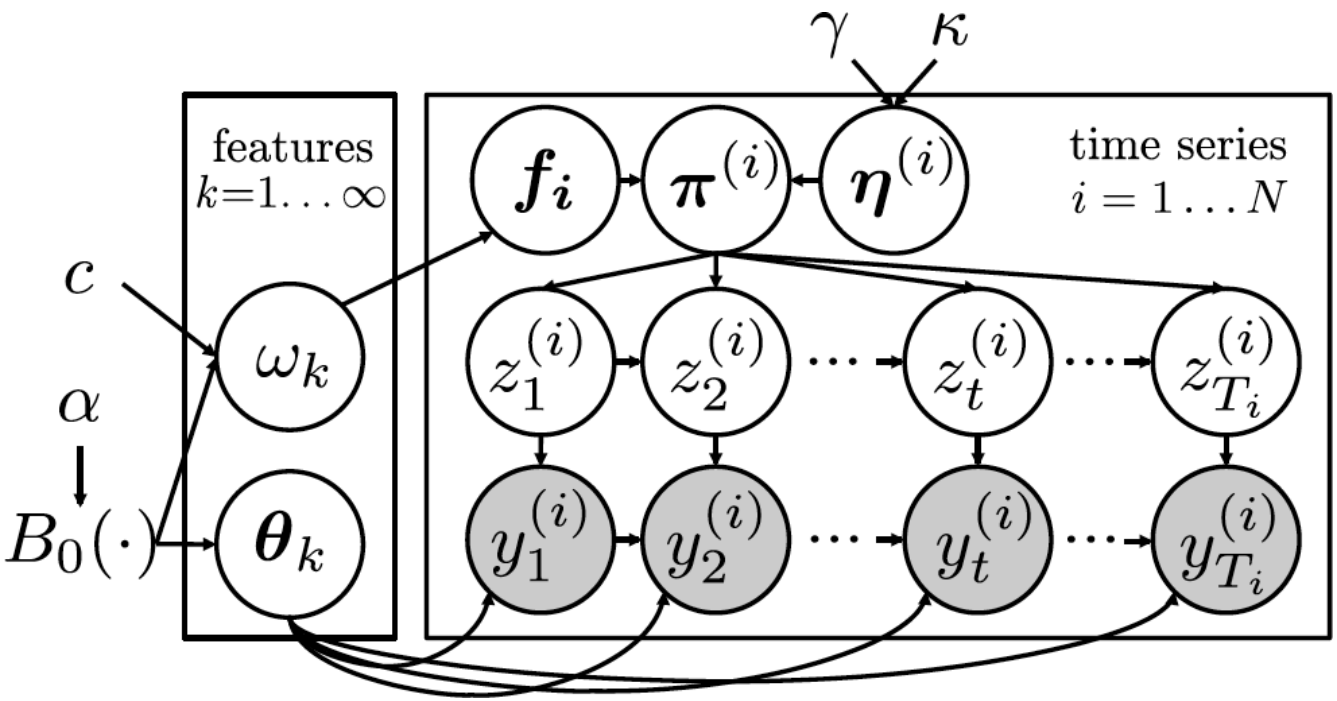
$$
\begin{array}{cccccccc}
 & z_1 & z_2 & z_3 & z_4 & z_5 & \cdots \\
z_1 & \pi_{11}^{(i)} & \pi_{12}^{(i)} & \pi_{13}^{(i)} & \pi_{14}^{(i)} & \pi_{15}^{(i)} & \cdots \\
z_2 & \pi_{21}^{(i)} & \pi_{22}^{(i)} & \pi_{23}^{(i)} & \pi_{24}^{(i)} & \pi_{25}^{(i)} & \cdots \\
z_3 & \pi_{31}^{(i)} & \pi_{32}^{(i)} & \pi_{33}^{(i)} & \pi_{34}^{(i)} & \pi_{35}^{(i)} & \cdots \\
z_4 & \pi_{41}^{(i)} & \pi_{42}^{(i)} & \pi_{43}^{(i)} & \pi_{44}^{(i)} & \pi_{45}^{(i)} & \cdots \\
z_5 & \pi_{51}^{(i)} & \pi_{52}^{(i)} & \pi_{53}^{(i)} & \pi_{54}^{(i)} & \pi_{55}^{(i)} & \cdots
\end{array}
$$

$$\pi_j \sim \text{Dirichlet}$$

Beta Process – Bernoulli Process



The BP-HMM


γ
κ
features
k=1...∞
time series
$i = 1 \ldots N$
$\boldsymbol{f_i}$
$\boldsymbol{\pi}^{(i)}$
$\boldsymbol{\eta}^{(i)}$
c
ω_k
α
$B_0(\cdot)$
$\boldsymbol{\theta}_k$
$z_1^{(i)}$
$z_2^{(i)}$
$z_t^{(i)}$
$z_{T_i}^{(i)}$
$y_1^{(i)}$
$y_2^{(i)}$
$y_t^{(i)}$
$y_{T_i}^{(i)}$

Outline

The DIEM Project

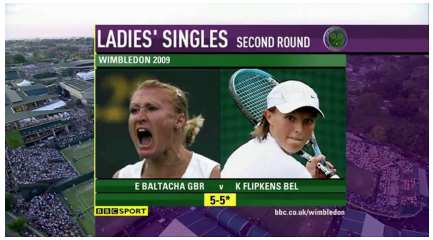
Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation, 3*(1), 5-24.

- Project led by John Henderson (UC Davis)
- Overview: https://thediemproject.wordpress.com/
- Examples: https://vimeo.com/visualcognition/
- Data: https://www.mediafire.com/?mpu3ot0m2o384

Technical Details

- 85 Video Clips
- Sports, TV Shows, Advertisements, Viral Videos, etc.
- Varying sizes and lengths
- $>$ 250 total participants
- Many viewers per clip (over 100 for some clips)

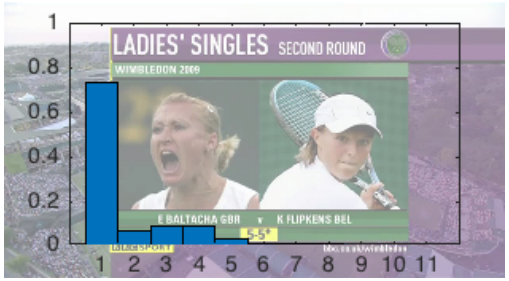
- Eyelink 2000, 1000Hz for each eye (down-sampled to 30Hz)
- 21" Viewsonic Monitor at 1280×960, 120Hz
- Observers at 90cm; using chin and headrest
  - Task was to rate enjoyment of the clip

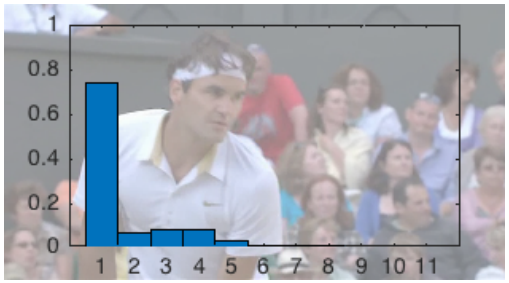


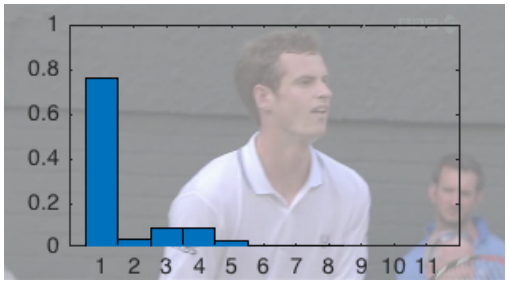
- 10 Video Clips
- 10 Observers Each

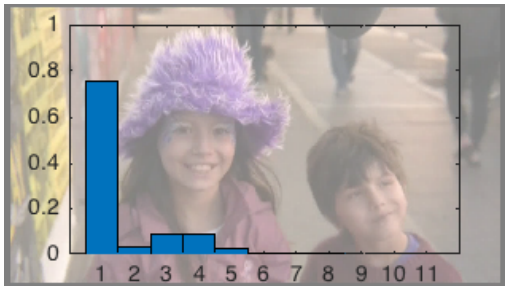
Outline

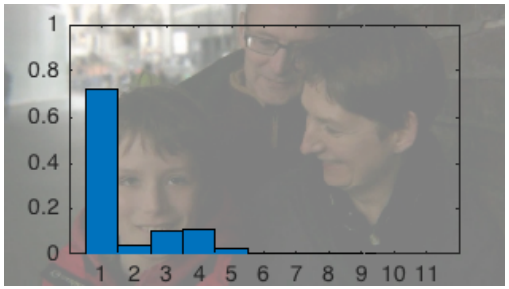
Distribution of Features

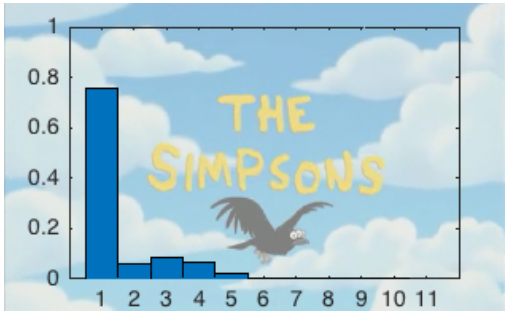


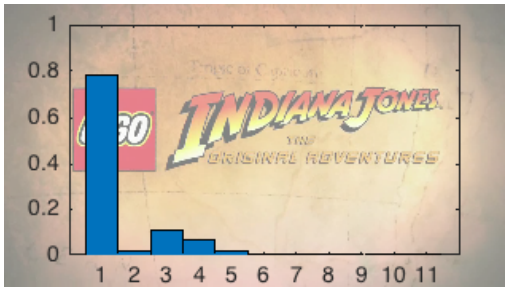


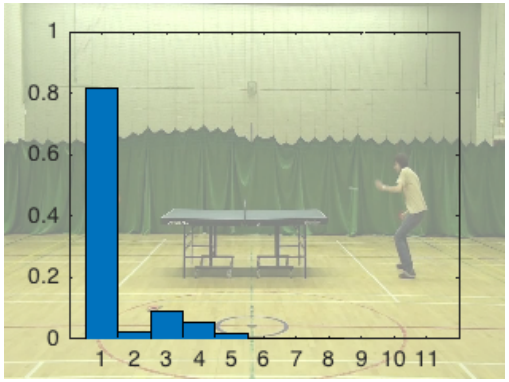




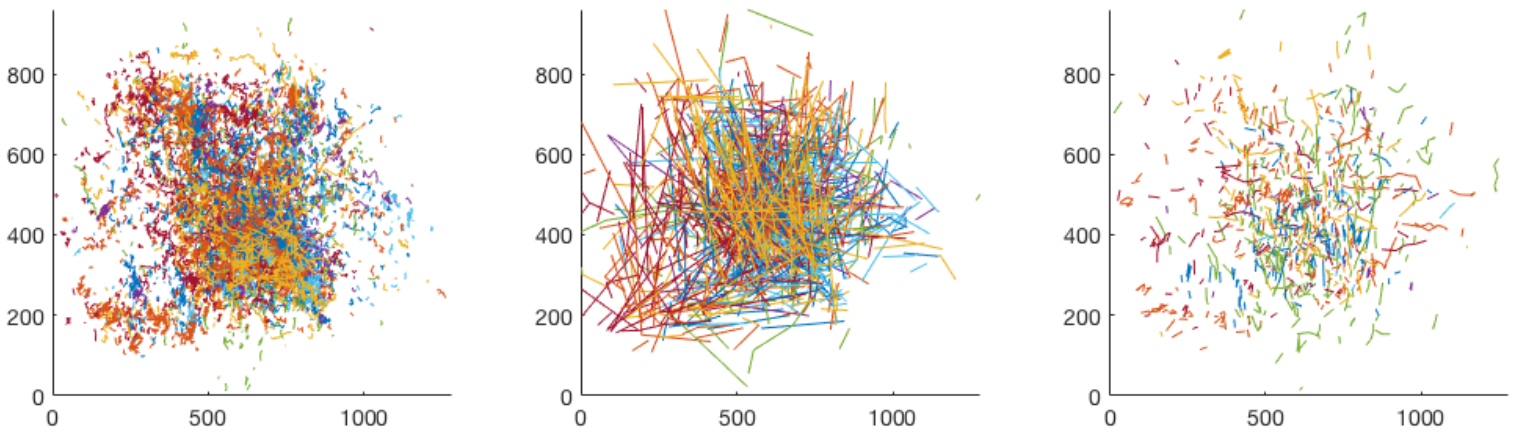




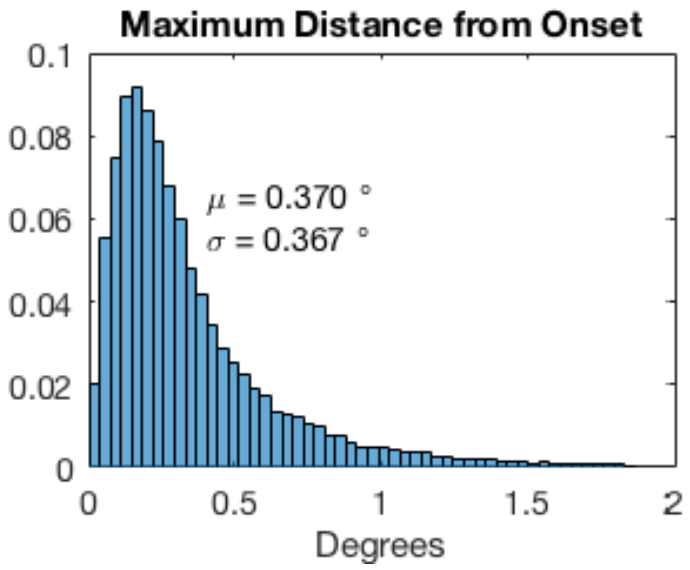


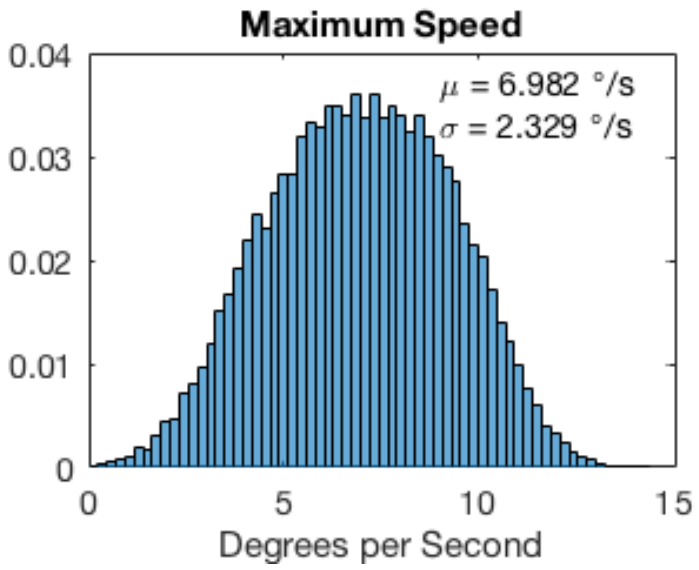


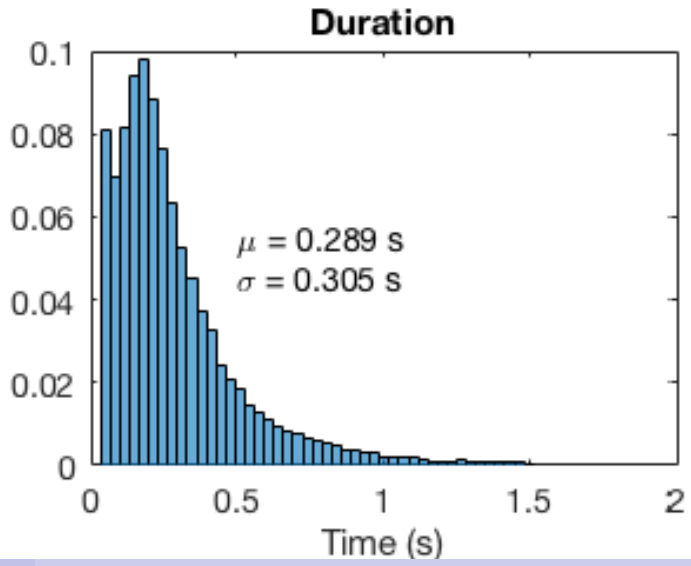
Features


800
600
400
200
0
0
500
1000
800
600
400
200
0
0
500
1000
800
600
400
200
0
0
500
1000

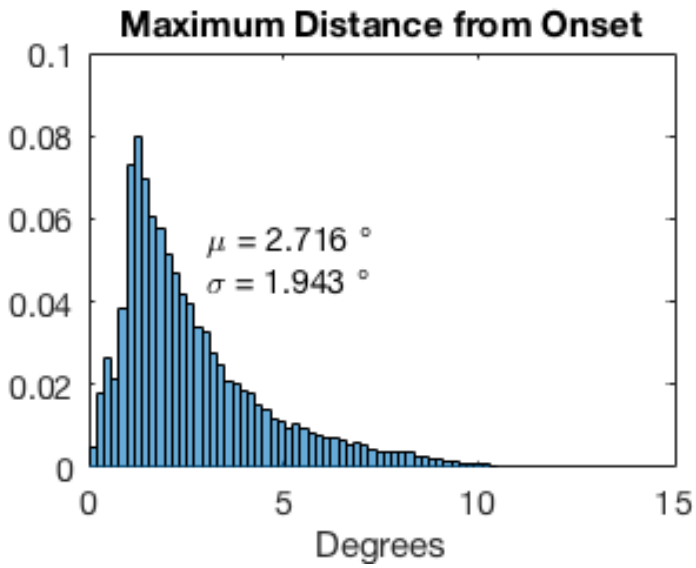
Fixations

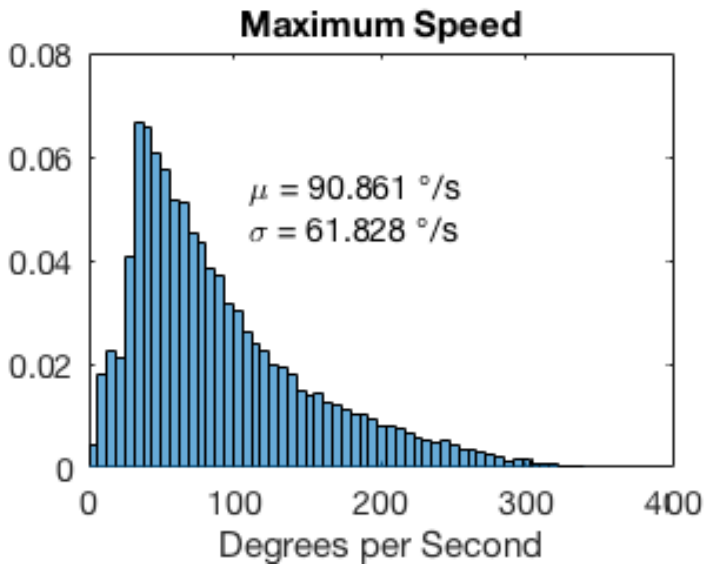

Maximum Distance from Onset
0.1
0.08
0.06
0.04
0.02
0
0
0.5
1
1.5
2
Degrees
μ = 0.370 °
σ = 0.367 °

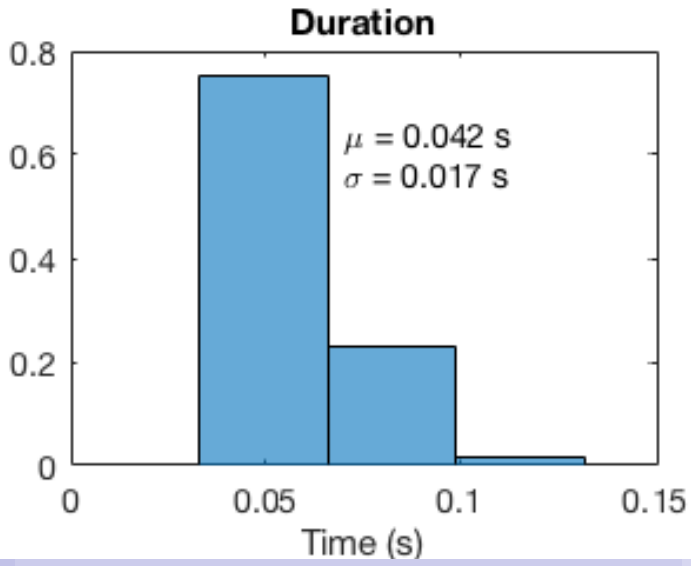

Maximum Speed
0.04
0.03
0.02
0.01
0
0
5
10
15
Degrees per Second
μ = 6.982 °/s
σ = 2.329 °/s


Duration
0.1
0.08
0.06
0.04
0.02
0
0
0.5
1
1.5
2
Time (s)
μ = 0.289 s
σ = 0.305 s

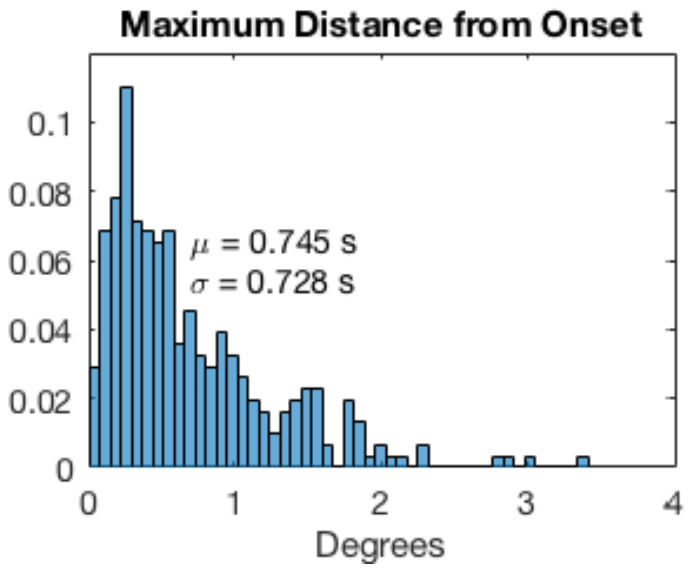
Saccades

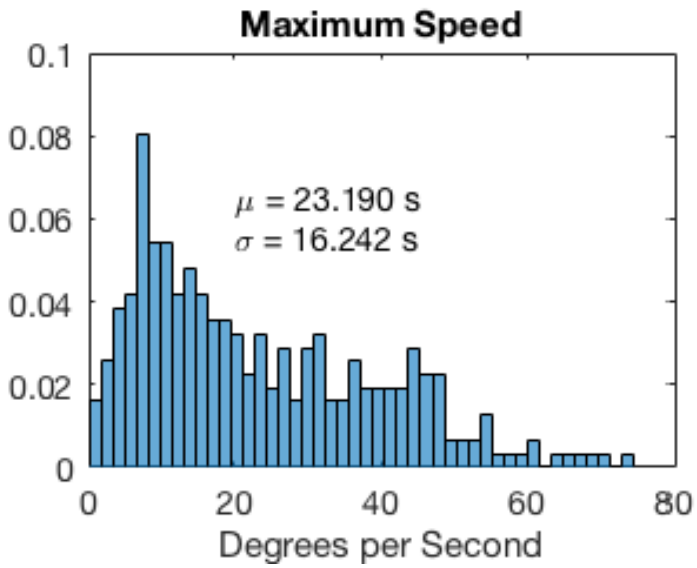

Maximum Distance from Onset
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
Degrees
μ = 2.716 °
σ = 1.943 °

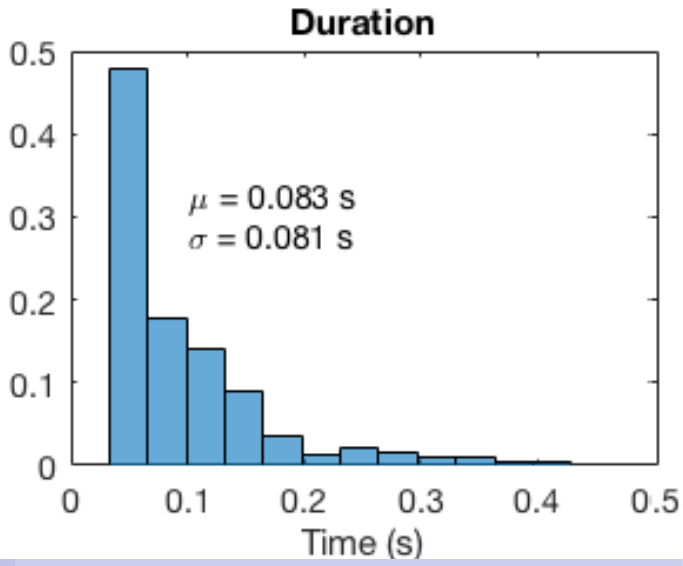

Maximum Speed
0.08
0.06
0.04
0.02
0
0
100
200
300
400
Degrees per Second
μ = 90.861 °/s
σ = 61.828 °/s


Duration
0.8
0.6
0.4
0.2
0
0
0.05
0.1
0.15
Time (s)
μ = 0.042 s
σ = 0.017 s

Glissades?


Maximum Distance from Onset
μ = 0.745 s
σ = 0.728 s
0.1
0.08
0.06
0.04
0.02
0
0
1
2
3
4
Degrees


Maximum Speed
μ = 23.190 s
σ = 16.242 s
0.1
0.08
0.06
0.04
0.02
0
0
20
40
60
80
Degrees per Second


Duration
μ = 0.083 s
σ = 0.081 s
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
Time (s)

Feature Transitions

Forward

		To		
		Fix	Scd	Gls
From	Fix		0.70	0.30
	Scd	0.55		0.45
	Gls	0.90	0.10	

Feature Transitions

Forward

		To		
		Fix	Scd	Gls
From	Fix		0.70	0.30
	Scd	0.55		0.45
	Gls	0.90	0.10	

Backward

		From		
		Fix	Scd	Gls
To	Fix		0.42	0.58
	Scd	0.91		0.09
	Gls	0.45	0.54	

Correspondence with Eyelink Parsing

	Fixation	Saccade	Glissade
Fixation	0.88	0.05	0.08
Saccade	0.44	0.45	0.11

Outline

- Data driven discrimination of eye-movement patterns
  - Across multiple videos and observers
  - A statistical model (i.e., we have a distribution over parse and parameters)
  - Described 100 time series with 3+ features
  - Based on limited data (30Hz rather than 1000Hz)
  - Features are similar to a priori categories: fixations, saccades, and glissades.

- Data driven discrimination of eye-movement patterns
  - Across multiple videos and observers
  - A statistical model (i.e., we have a distribution over parse and parameters)
  - Described 100 time series with 3+ features
  - Based on limited data (30Hz rather than 1000Hz)
  - Features are similar to a priori categories: fixations, saccades, and glissades.
- Future: Parse based on more information (particularly location)

- Data driven discrimination of eye-movement patterns
  - Across multiple videos and observers
  - A statistical model (i.e., we have a distribution over parse and parameters)
  - Described 100 time series with 3+ features
  - Based on limited data (30Hz rather than 1000Hz)
  - Features are similar to a priori categories: fixations, saccades, and glissades.
- Future: Parse based on more information (particularly location)

Thank you!

This work was supported by
AFOSR Grant FA9550-13-1-0087
The Analysis in Motion Initiative at Pacific Northwest National Laboratory