Motion Perception Under Uncertainty

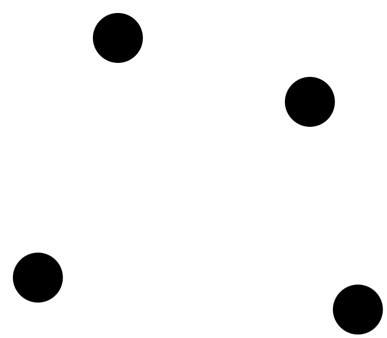
Hongjing Lu
Department of Psychology
University of Hong Kong

Outline

- Uncertainty in motion stimulus
  - Correspondence problem
- Qualitative fitting using ideal observer models
  - Based on signal detection theory
  - Based on Bayesian framework
- Quantitative fitting using Bayesian models with priors
  - Bayesian model with generic priors
  - Expectation-Maximization (EM) implementation

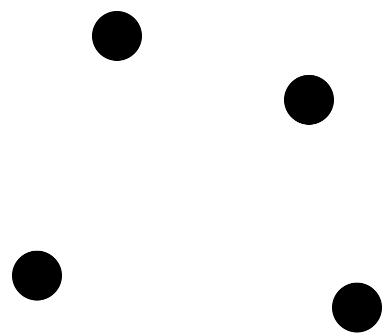
Perceive Motion from Two Frames

Frame 1



Perceive Motion from Two Frames

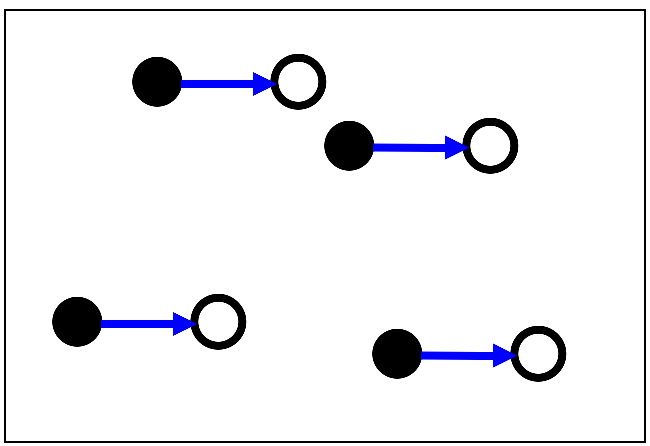
Frame 2



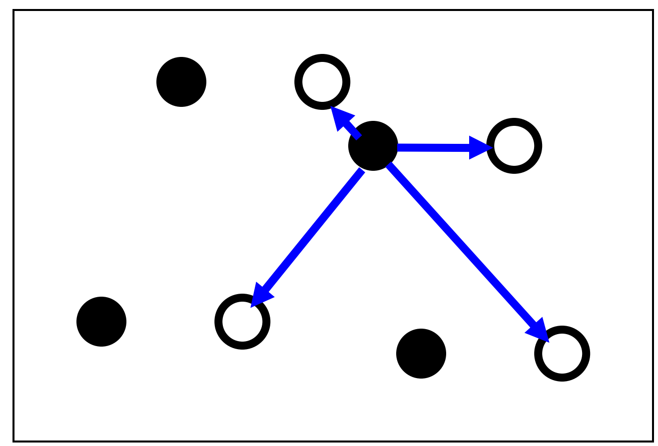
Correspondence Problem

Dot positions in the first frame

Dot positions in the second frame



Stimulus generation
4 dots move coherently



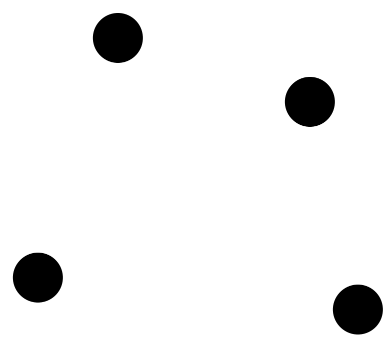
Correspondence uncertainty

Random Dot Kinematograms (RDK)

- Common stimuli used in psychophysics and physiology, Barlow et al. (1979, 1997); Britten et al. (1992, 1995); Newsome et al. (1989, 1990), etc.

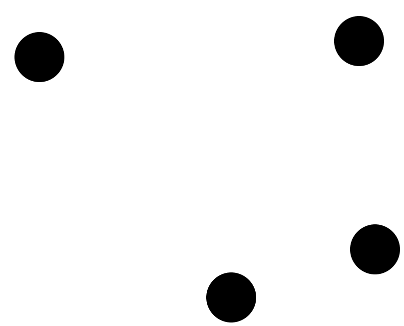
Motion Detection Task

Trial 1, Frame 1



Motion Detection Task

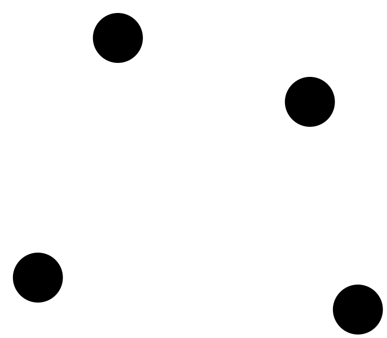
Trial 1, Frame 2



Do you perceive coherent motion?

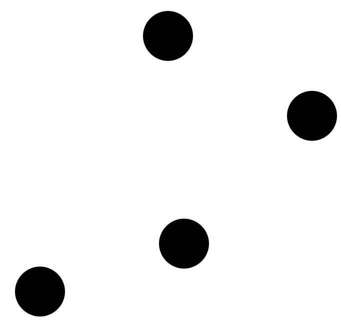
Motion Detection Task

Trial 2, Frame 1



Motion Detection Task

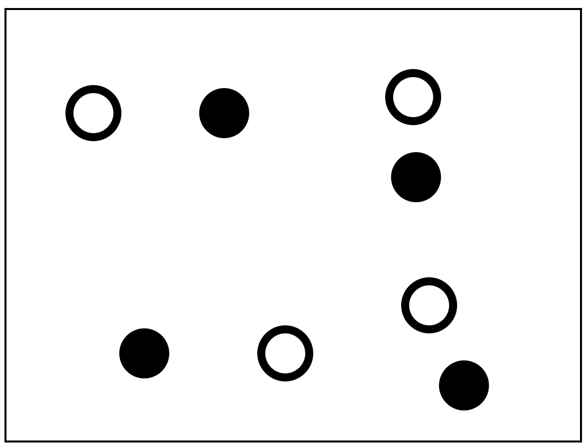
Trial 2, Frame 2



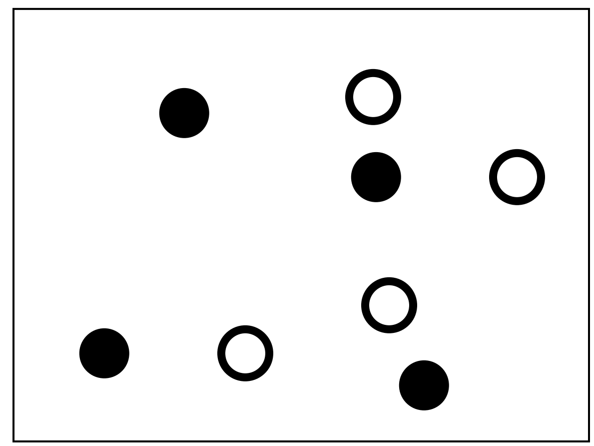
Do you perceive coherent motion ?

Motion Detection Task

Trial 1



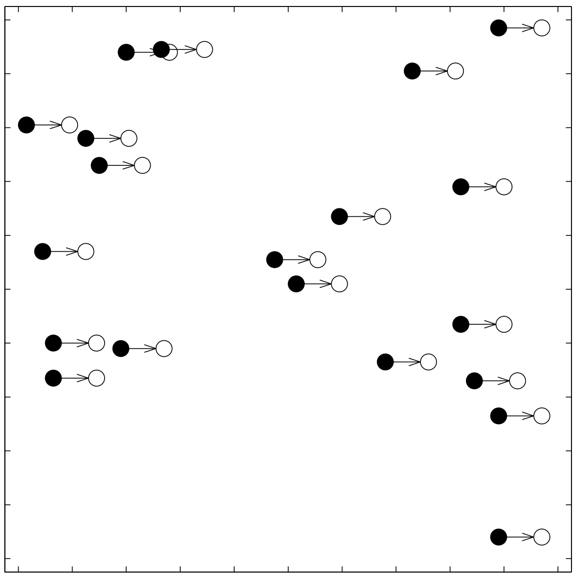
Trial 2



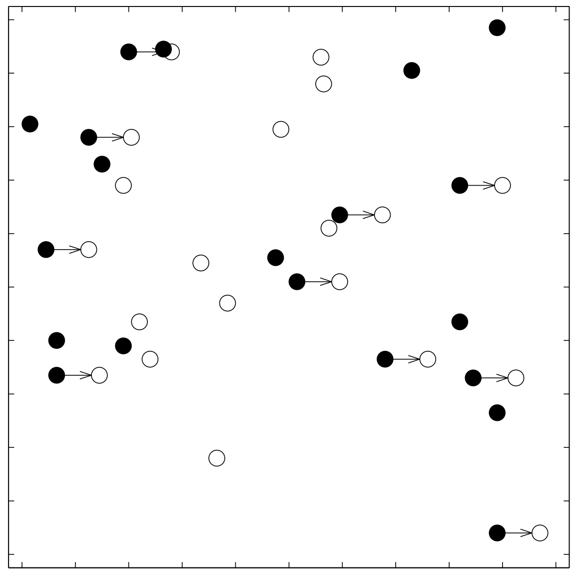
Do you perceive coherent motion?

Random Dot Kinematograms (RDK)

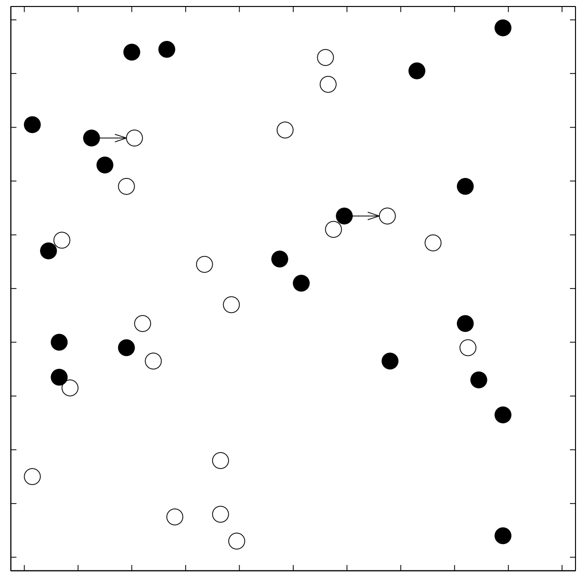
- Common stimuli used in psychophysics and physiology, Barlow et al. (1979, 1997); Britten et al. (1992, 1995); Newsome et al. (1989, 1990), etc.



N=20, C=1.0



N=20, C=0.5



N=20, C=0.1

N: number of dots

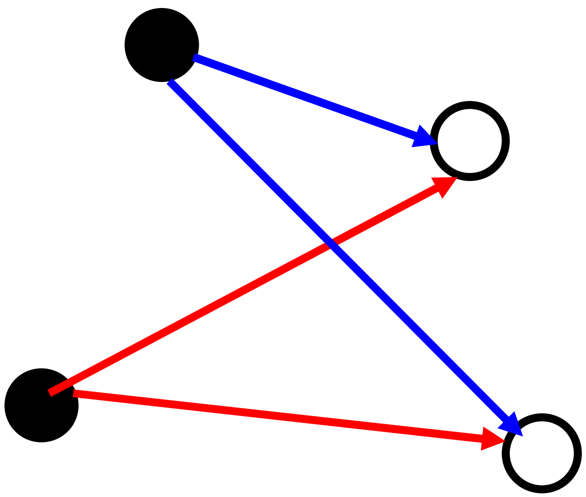
C: coherence ratio, the proportion of dots that move with the same velocity

Psychophysical Experiments

- Manipulate visual stimuli by changing
  - Dot number, N
  - Dot displacement between two frames
  - Number of frames
  - …
- Human performance measurement
  - Accuracy
    - high accuracy indicates good performance)
  - Coherence threshold is the coherence ratio needed to achieve 75% accuracy in a specific task
    - low coherence threshold indicates good performance

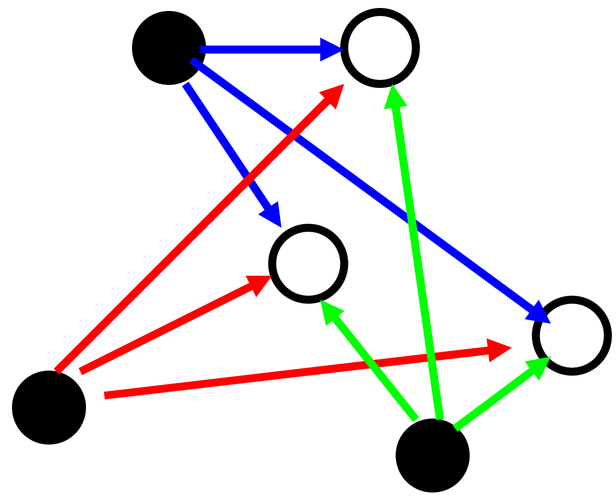
Correspondence Pairs

Two dots



N=2, $\Psi = 4$

Three dots

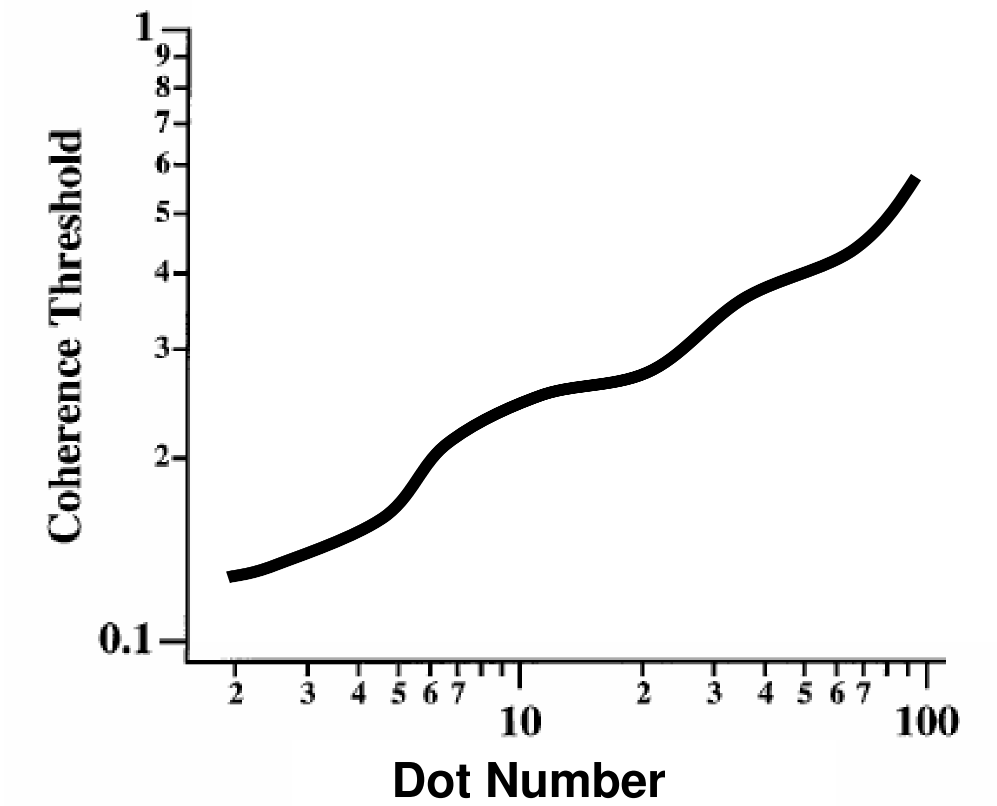


N=3, $\Psi = 9$

# of possible correspondence pairs, $\Psi = N^2$

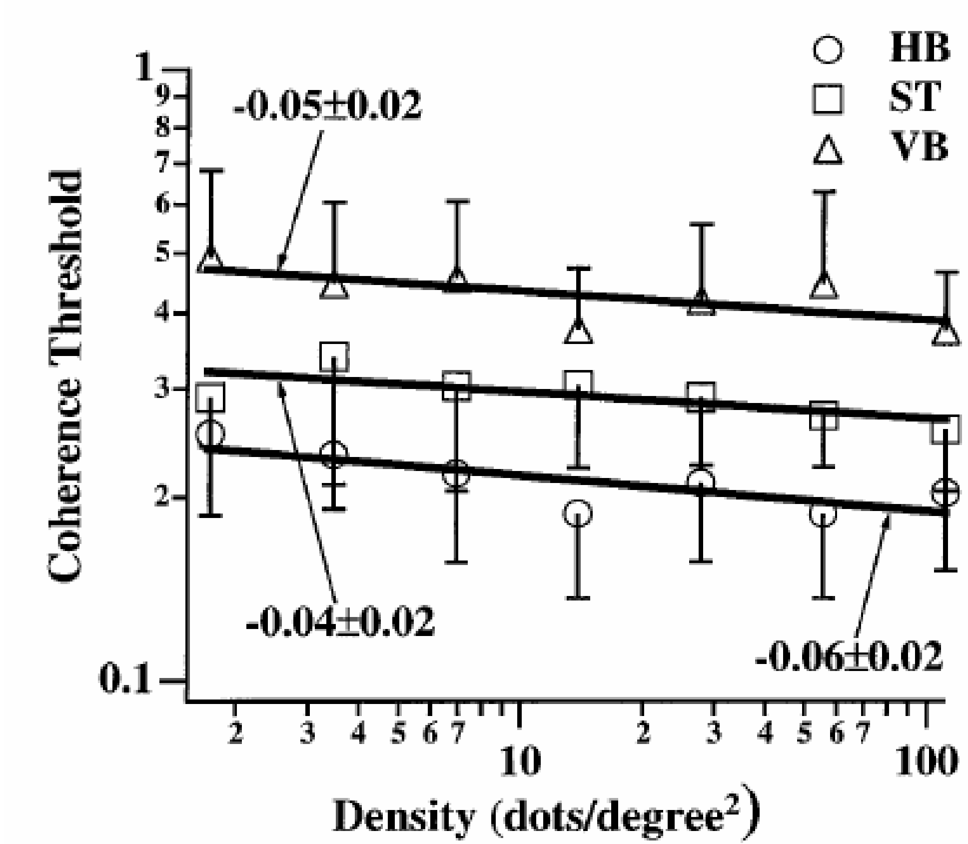
Coherence Threshold

- # of possible correspondence pairs, $\Psi = N^2$
  - Difficulty of coherent motion perception would be expected to increase with the number of dots in RDK
  - i.e. coherence threshold would be expected to increase with the number of dots



Human Coherence Threshold

- Intriguing psychophysical finding (Barlow & Tripathy, 1997)
  - Human coherence threshold is almost constant with the number of dots in RDK



Outline

- Uncertainty in motion stimulus
  - Correspondence problem
- Qualitative fitting using ideal observer models
  - Based on signal detection theory
  - Based on Bayesian framework

Ideal Observer for Detecting Coherent Motion

- Ideal observer model provides optimal performance as a benchmark to evaluate *how well* humans process motion information on a specific task

- Ideal observer using signal detection theory (SDT) (Barlow & Tripathy, 1997)

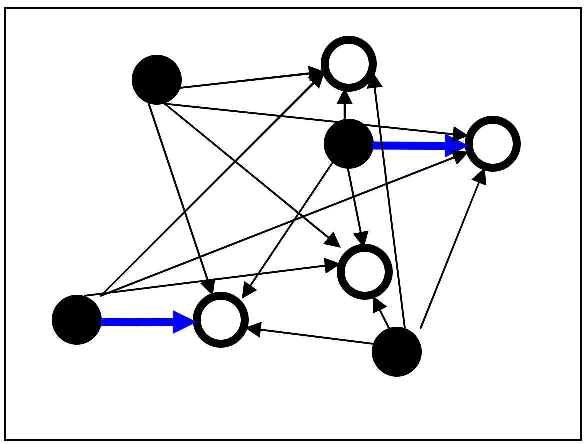
- Ideal observer model uses "experimenter" prior, i.e.
  - the proportion of dots that move coherently, ***C***
  - velocity of coherent motion, ***T***

Ideal Observer Using SDT

- Detection decision is based on
  - Number of matches, ψ , i.e. the number of dots in the first frame that have a corresponding dot with displacement $\boldsymbol{T}$ in the second frame

Ideal Observer Using SDT

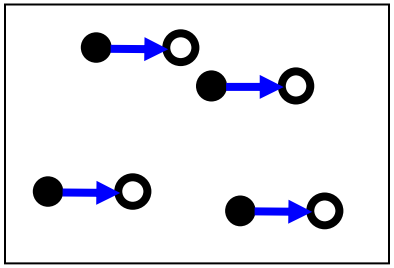
Observation



- Two possibilities
  - Random motion, the two matches are due to chance alignment of randomly moving dots
  - Coherent motion, the two matches are from coherent motion of signal dots

Ideal Observer Using SDT

– Random motion stimulus



N dots in the display.

For each dot in the first frame, the probability of a dot at the corresponding position in the second frame is N/Q. Q is the total number of possible positions in the display

ψ ~ Binomial (N, N/Q)

$$<\psi_0> = N\frac{N}{Q}$$

$$VAR(\psi_0) = N\frac{N}{Q}\left(1-\frac{N}{Q}\right)$$

– Motion stimulus with ***CN*** dots moving coherently

$\psi - CN$ ~ Binomial ((1-C)N, N/Q)

$$<\psi_C> = CN + (1-C)\frac{N^2}{Q}$$

Assume $VAR(\psi_C) = VAR(\psi_0)$

Ideal Observer Using SDT
(Barlow & Tripathy, 1997)

- Compute sensitivity and coherence threshold using SDT

Sensitivity

$$d' = \frac{<\psi_C> - <\psi_0>}{\sqrt{VAR(\psi_0)}} = C\sqrt{Q-N}$$

Coherence threshold for d′ = 1

$$C_{th} = 1 \Big/ \sqrt{Q-N} \approx 1 \Big/ \sqrt{Q} \qquad Q >> N$$

Ideal observer predicts that coherence threshold is constant with the number of dots in RDK

Bayesian Ideal Observer
(Lu & Yuille, 2005)

- Define correspondence variables V_{ia}

  $V_{ia} = 1$, if $y_a = x_i + T$; $V_{ia} = 0$, otherwise

- Generative model for the RDK stimulus
  - Random motion stimulus

    $$P(\{y_a\},\{x_i\} \mid rand) = P(\{y_a\})P(\{x_i\})$$

  - Motion stimulus with ***CN*** dots moving coherently

$$P(\{y_a\},\{x_i\} \mid coh, T) = \sum_{V_{ia}} P(\{y_a\} \mid \{x_i\},\{V_{ia}\},T)P(\{V_{ia}\})P(\{x_i\})$$

with the constraint $\sum_{ia} V_{ia} = CN$

Bayesian Ideal Observer

- Generative model for the RDK stimulus

$$P(\{y_a\},\{x_i\} \mid coh,T) = \sum_{V_{ia}} P(\{y_a\} \mid \{x_i\},\{V_{ia}\},T)P(\{V_{ia}\})P(\{x_i\})$$

  - Prior distribution for the dot positions allows all configurations of the dots to be equally likely

$$P(\{y_a\}) = P(\{x_i\}) = (Q-N)!/Q! \approx 1/Q^N \text{ if } Q >> N$$

  - Prior distribution for correspondence variable is uniform

$$P(\{V_{ia}\}) = \frac{(N-CN)!}{N!}\frac{(N-CN)!(CN)!}{N!}$$

  - Conditional Distribution given dot positions in the first frame

$$P(\{y_a\} \mid \{x_i\},\{V_{ia}\},T) = \frac{(Q-N)!}{(Q-CN)!}\prod_{ia}\left(\delta(y_a, x_i + T)\right)^{V_{ia}},$$

Bayesian Ideal Observer

- Generative model for the RDK stimulus

$$P(\{y_a\},\{x_i\} \mid coh,T) = \sum_{V_{ia}} P(\{y_a\} \mid \{x_i\},\{V_{ia}\},T)\,P(\{V_{ia}\})\,P(\{x_i\})$$

$$= \frac{(Q-N)!}{(Q-CN)!} \sum_{V_{ia}} \prod_{ia} \left(\delta(y_a, x_i + T)\right)^{V_{ia}} \frac{(N-CN)!}{N!} \frac{(N-CN)!(CN)!}{N!} \frac{(Q-N)!}{Q!}$$

$$= \frac{(Q-N)!}{(Q-CN)!} \frac{\psi!}{(\psi-CN)!(CN)!} \frac{(N-CN)!}{N!} \frac{(N-CN)!(CN)!}{N!} \frac{(Q-N)!}{Q!}$$

- Likelihood ratio

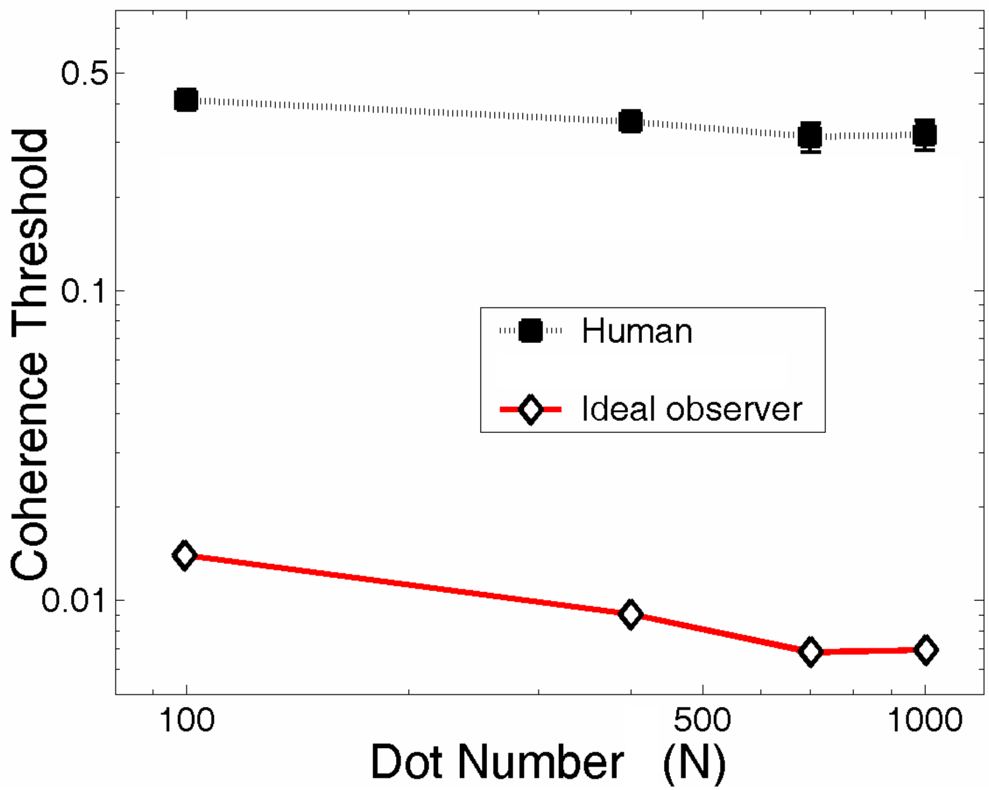
$$\frac{P(Data \mid coh,T)}{P(Data \mid rand)} = \left(\frac{(N-CN)!}{N!}\right)^2 \frac{Q!}{(Q-s)!} \frac{\psi!}{(\psi-s)!}$$

Ideal Observer

- Ideal observer using signal detection theory is an approximation to Bayesian ideal observer

Human Observer vs. Bayesian Ideal Observer Model

- Ideal observer model predicts the qualitative trend of human performance
- Remarkably low human efficiency, 0.3~1%



Human Coherence Threshold:

30%~50%

Ideal Coherence Threshold:

0.5%~1%

Why Low Efficiency?

- Ideal observer model uses prior specific to the experiment, because ideal observer model knows
  - the proportion of dots that move coherently, ***C***
  - velocity of coherent motion, ***T***

$$P(v \mid \{x_i\},\{y_a\}) = \frac{1}{Z} P(\{y_a\} \mid \{x_i\}, v) \underline{P(v)}$$

Ideal observer model uses specific prior knowledge about the RDK stimulus used in the experiment

Why Low Efficiency?

- Ideal observer model uses prior specific to the experiment
- Human observers might use some ***generic prior*** consistent with motion statistics in natural scenes

$$P(v \mid \{x_i\},\{y_a\}) = \frac{1}{Z} P(\{y_a\} \mid \{x_i\}, v) \underline{P(v)}$$

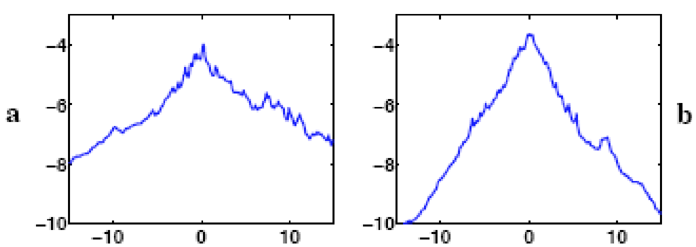
Ideal observer model uses specific prior
Human observers might use generic prior

Outline

- Uncertainty in motion stimulus
  - Correspondence problem
- Qualitative fitting using ideal observer models
  - Based on signal detection theory
  - Based on Bayesian framework
- Quantitative fitting using Bayesian models with priors
  - Bayesian model with generic priors
  - Expectation-Maximization (EM) implementation

Generic Prior: Slow & Smooth

- Prefer slow and spatially smooth motion
  - Slowness: most objects are static
  - Smoothness: object features tend to move together
- Natural statistics of motion flow (Roth & Black, 2005)
  - log-histogram of



Horizontal velocity, u　　　Vertical velocity, v

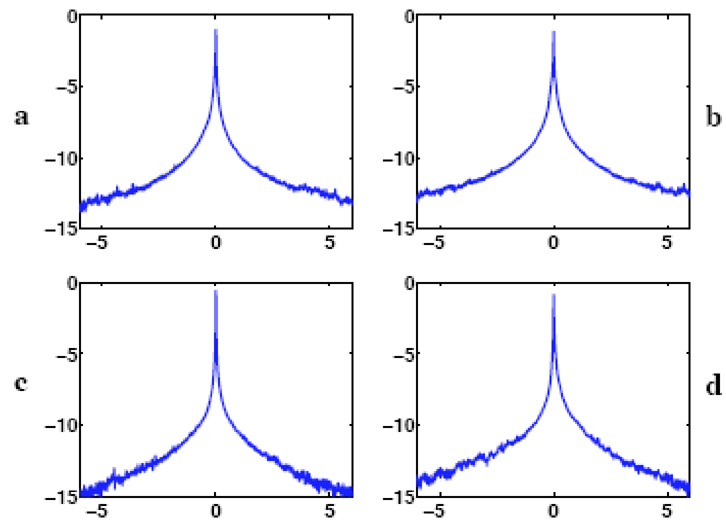


Figure 4. Derivative statistics of optical flow: log-histograms of **(a)** $\partial u/\partial x$, **(b)** $\partial u/\partial y$, **(c)** $\partial v/\partial x$, **(d)** $\partial v/\partial y$.

Generic Prior: Slow & Smooth

- Prefer slow and spatially smooth motion

$$P(\vec{v}) = e^{-\lambda \|L\vec{v}\|^2} = e^{-\lambda \sum_{m=0}^{\infty} \frac{\sigma^{2m}}{m!2^m} \left(\frac{\partial^m \vec{v}}{\partial x^m} \right)^2}$$

Yuille & Grzywacz (1988)

Slowness

$v = 0$

First-order Smoothness

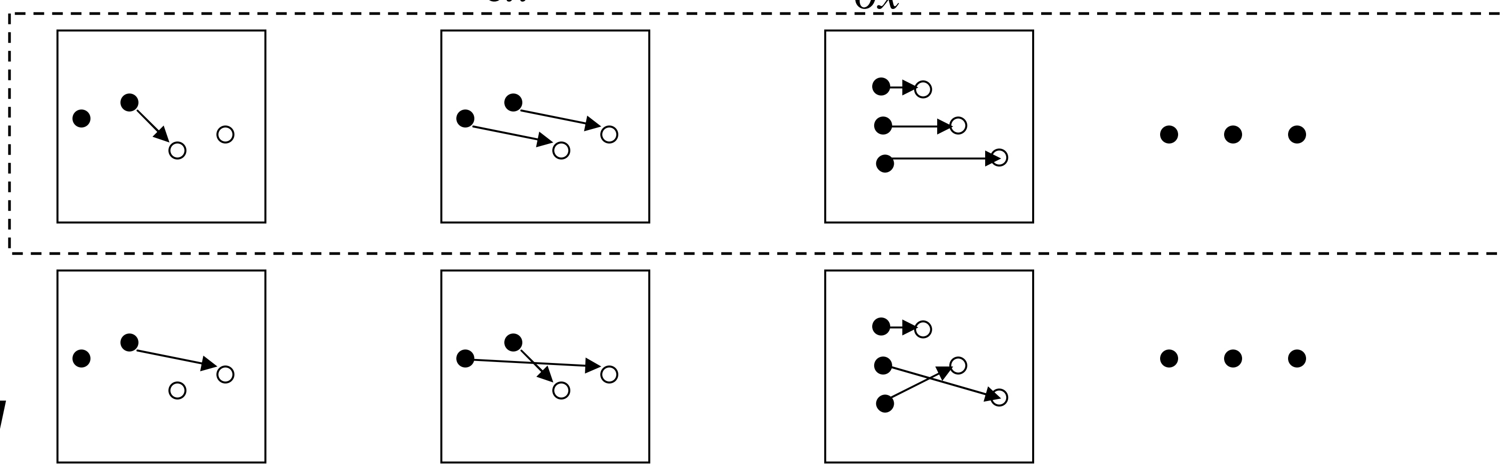
$\frac{\partial v}{\partial x} = 0$

Second-order Smoothness

$\frac{\partial^2 v}{\partial x^2} = 0$

Preferred

Less Preferred



Bayesian Model with Priors

- Ideal observer model uses "experimenter" prior, i.e.
  - the proportion of dots that move coherently, ***C***
  - velocity of coherent motion, ***T***

- Slow & smooth model estimates velocity field without requiring any knowledge about stimulus parameters

Notation in Slow & Smooth Model

- Binary correspondence variable

$$M_{ia} = 1, \text{ if } x_i \text{ corresponds to } y_a, x_i \rightarrow y_a$$

$$M_{ia} = 0, \text{ otherwise}$$

- Binary outlier variable

$M_{i0} = 1,$ if the i-th dot in the first frame is an outlier without a corresponding dot in the second frame

$$M_{i0} = 0, \text{ otherwise}$$

with a constraint that ensures that dot *i* the first frame is either unmatched as an outlier, or is matched to a dot a in the second frame, $\sum_{a=0}^{N} M_{ia} = 1$

- Probabilistic correspondence

$$m_{ia} = P(M_{ia} = 1) \qquad m_{i0} = P(M_{i0} = 1)$$

with the constraint $\sum_{a=0}^{N} m_{ia} = 1$

Bayesian Model with Slow & Smooth Prior

Goal: maximize $P(\vec{v} \mid \{x_i\},\{y_a\}) = \sum_M P(\vec{v},\{M_{ia}\} \mid \{x_i\},\{y_a\})$

with the constraint of $\sum_{a=0}^{N} M_{ia} = 1, \forall i$

Apply Bayes rule

$$P(\vec{v},\{M_{ia}\} \mid \{x_i\},\{y_a\}) = \frac{1}{Z} P(\{y_a\} \mid \{x_i\},\vec{v},\{M_{ia}\}) P(\vec{v}) P(\{M_{ia}\})$$

$$= \frac{1}{Z} e^{-\sum_{i=1}^{N}\sum_{a=1}^{N} M_{ia}(y_a - x_i - v(x_i))^2 / T} \; e^{-\lambda \|Lv\|^2 / T} \; e^{-\xi \sum_{i=1}^{N} M_{i0} / T}$$

Likelihood

SS generic prior on velocity field

Bias against mismatch

Expectation-Maximization (EM)

Dempster, Laird, & Rubin (1977)

- The EM algorithm is a general method of estimating parameters of an underlying distribution from a given data set when the data is incomplete or has missing values
- Two main applications of EM
  - When the data indeed has missing values, due to problems with or limitations of the observation process
  - When maximizing the likelihood or posterior probability is analytically intractable, but it is doable by assuming hidden variables to simplify the distribution function

Expectation-Maximization (EM)

- E-step
  - Replace hidden variables by their expectations conditional on the estimated parameter
  - In the current example, E-step is to evaluate

$$m_{ia}^{(t)} = P(M_{ia}^{(t)} = 1 \mid \{x_i\}, \{y_a\}, v^{(t-1)})$$

    of the hidden variable M_{ia}^{t}, conditional on the observed data and the velocity field from the last M-step

- M-step
  - estimate parameters by maximizing expected log posterior distribution as if hidden variables were observed

$$v^{(t)} = \arg\max_v E_m \left[\log P(v, m_{ia}^{(t)} \mid \{x_i\}, \{y_a\}) \right]$$

EM Implementation

Goal: maximize

$$P(v \mid \{x_i\},\{y_a\}) = \sum_M \frac{1}{Z} e^{-\sum_{i=1}^{N}\sum_{a=1}^{N} M_{ia}(y_a - x_i - v(x_i))^2/T \; -\lambda\|Lv\|^2/T - \xi\sum_{i=1}^{N} M_{i0}}$$

with the constraint of $\sum_{a=0}^{N} M_{ia} = 1, \forall i$

- The classical EM algorithm is equivalent to minimizing the free energy function (Frey & Hinton, 1996)

$$F[m,v] = \sum_{i=1}^{N}\sum_{a=1}^{N} m_{ia}(y_a - x_i - v(x_i))^2 + \lambda\|Lv\|^2 + \xi\sum_{i=1}^{N} m_{i0} + T\sum_{i=1}^{N}\sum_{a=1}^{N} m_{ia}\log m_{ia}$$

with the constraint $\sum_{a=0}^{N} m_{ia} = 1$

EM Implementation

Minimizing free energy

$$F[m,v] = \sum_{i=1}^{N}\sum_{a=1}^{N} m_{ia}(y_a - x_i - v(x_i))^2 + \lambda \|Lv\|^2 + \xi \sum_{i=1}^{N} m_{i0} + T\sum_{i=1}^{N}\sum_{a=1}^{N} m_{ia}\log m_{ia}$$

with the constraint $\sum_{a=0}^{N} m_{ia} = 1$

- E step:
  - minimizing this energy function with respect to m_{ia} while holding v constant
- M step:
  - minimizing this energy function with respect to v while holding m_{ia} constant

EM Implementation

E step: update the correspondences

$$\frac{\partial F}{\partial m_{ia}} = 0 \qquad \Rightarrow \qquad m_{ia} = \frac{e^{(-(y_a - x_i - v(x_i))^2/T)}}{e^{-\xi/T} + \sum_{a=1}^{N} e^{(-(y_a - x_i - v(x_i))^2/T)}}$$

$$\frac{\partial F}{\partial m_{i0}} = 0 \qquad \qquad m_{i0} = \frac{e^{-\xi/T}}{e^{-\xi/T} + \sum_{a=1}^{N} e^{(-(y_a - x_i - v(x_i))^2/T)}}$$

EM Algorithm

M step: update the velocity field

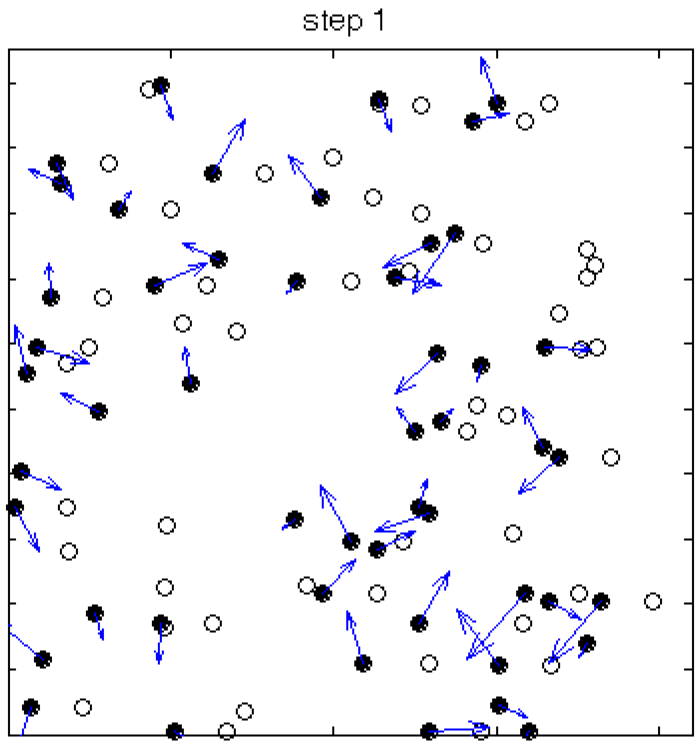
$$v(x_i) = \sum_{j=1}^{N} \beta_j G(x_i, x_j)$$

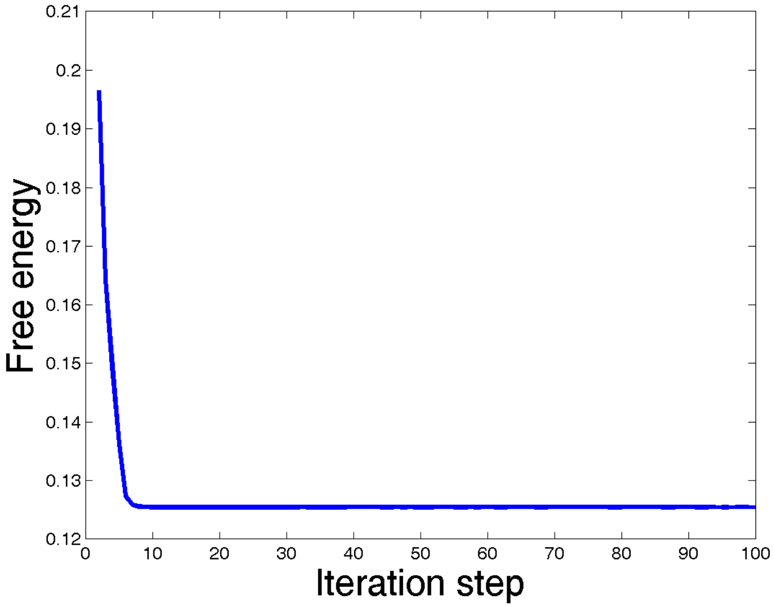
$$\frac{\partial E}{\partial v} = 0 \;\Rightarrow\; \sum_{j=1}^{N}\left(\underset{m}{G(x_i, x_j)} \sum_{i=1}^{N} m_{ia} + \lambda \delta_{ij} \right)\beta_j = \sum_{a=1}^{N} m_{ia}(x_a - x_i)$$

$$where \quad G(x_i, x_j) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-(x_i - x_j)^2}{2\sigma^2} \right)$$

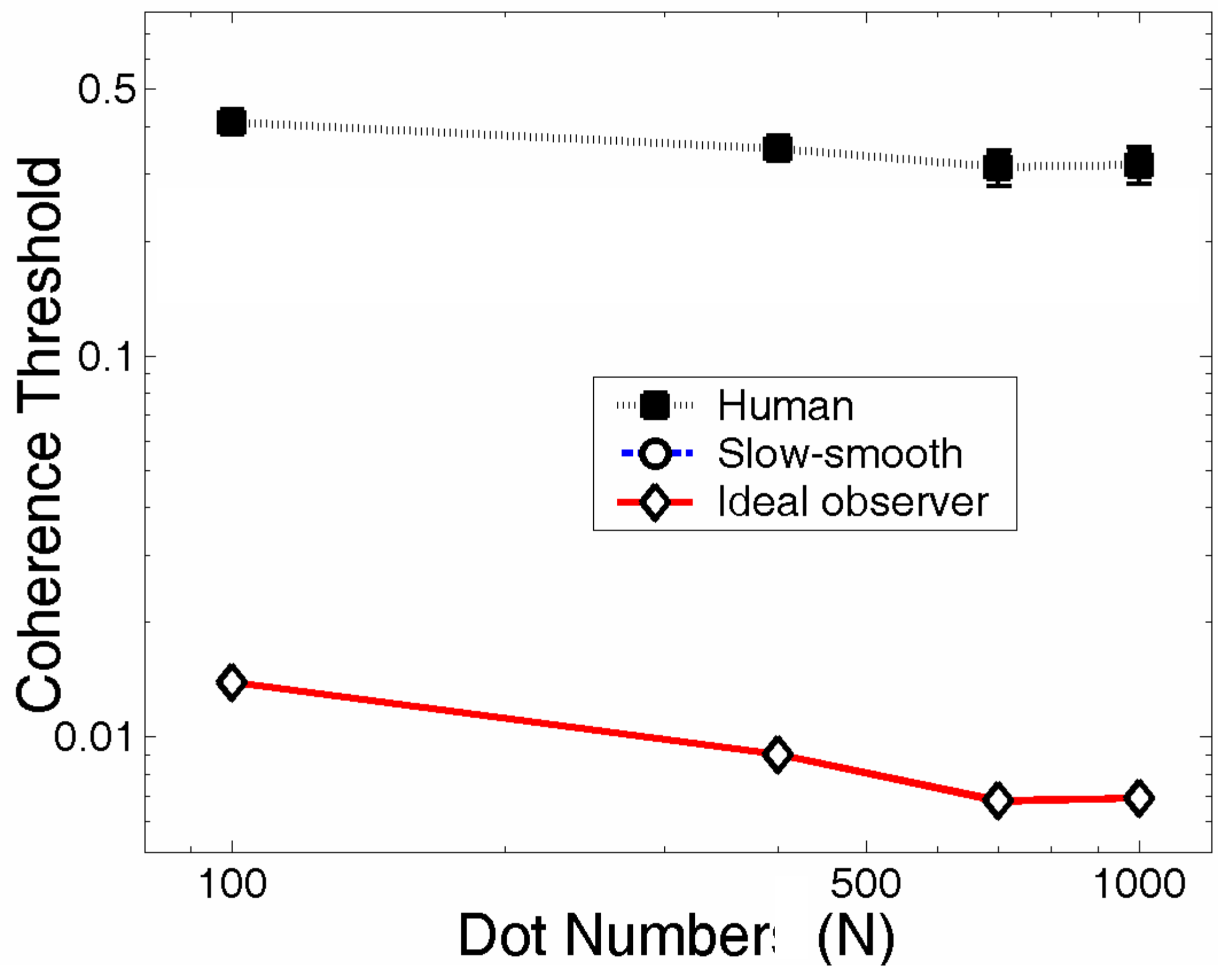
EM Iteration Steps

N=50, C=0.5





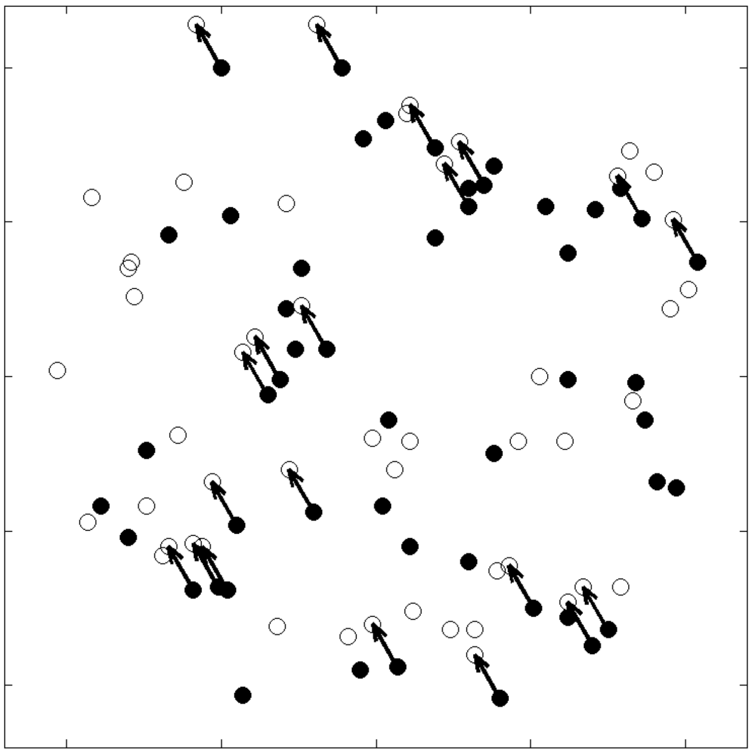
Exp1. Coherence Threshold in Detecting Coherent Motion



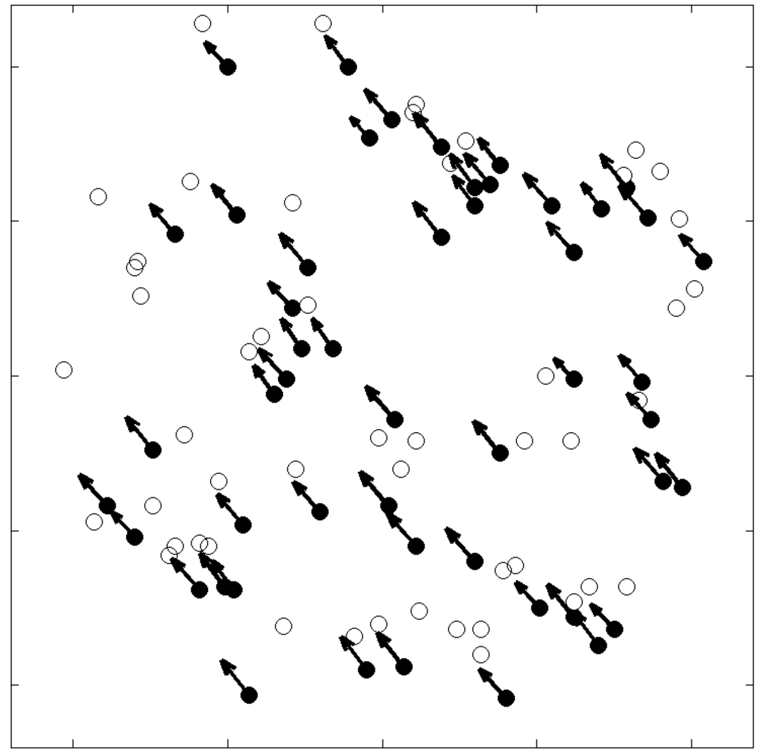
Model Prediction in a Trial

N=50, C=0.4

Stimulus

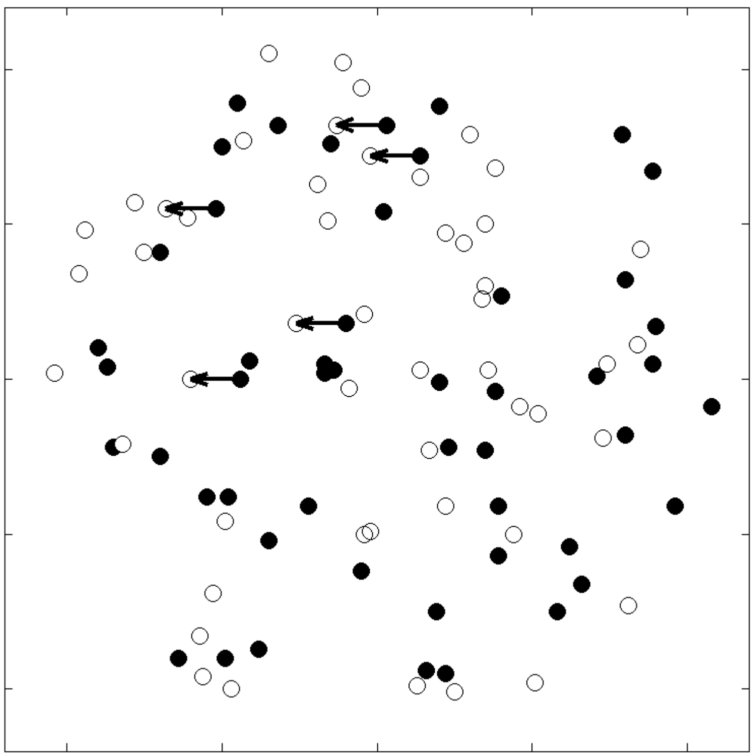


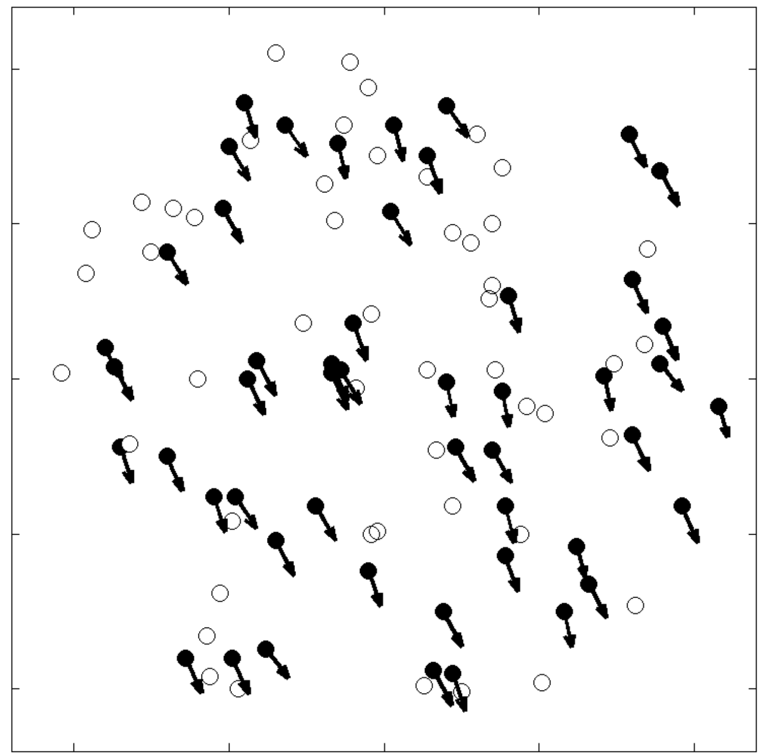
Model Prediction



Interesting Model Prediction

N=50, C=0.1

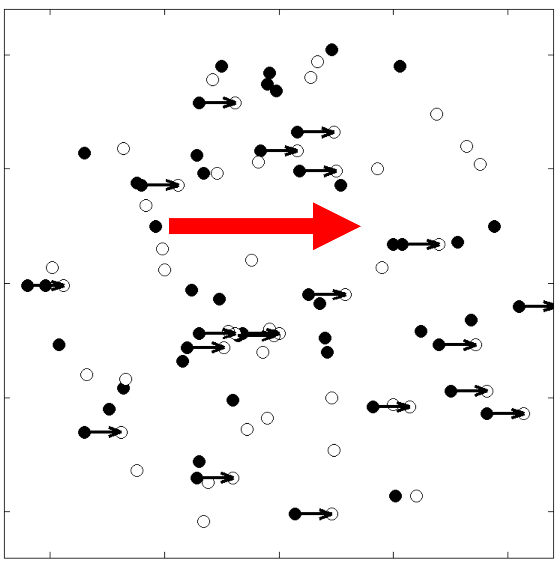




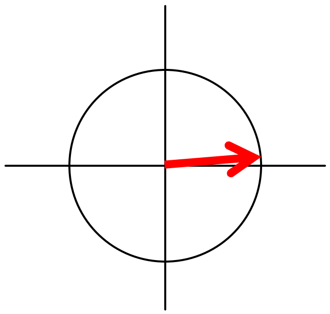
What would humans perceive?

Exp 2: Motion Perception in a Single Trial
N=50, C=0.4

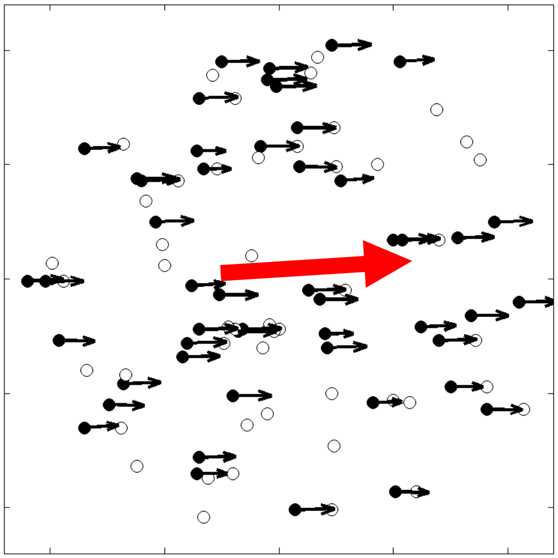
Stimulus



N=50, C=0.4

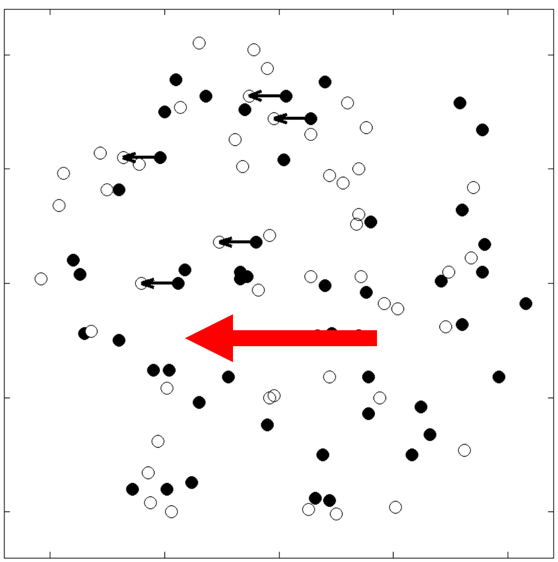


Model predicted global motion direction: 1 °



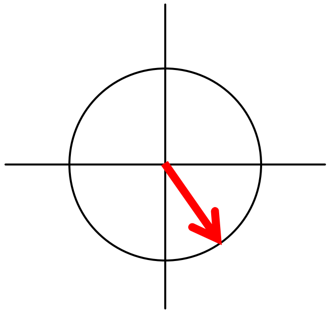
Exp 2: Motion Perception in a Single Trial
N=50, C=0.1

Stimulus

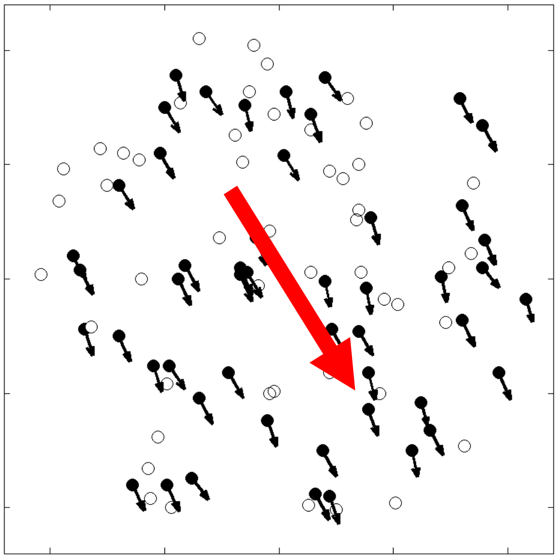


N=50, C=0.1

Human perceived global motion direction: 298°



Model predicted global motion direction: 294°



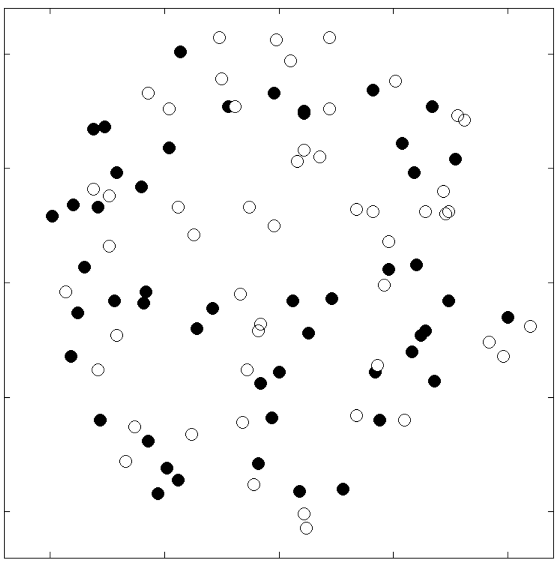
Experimental Test of Generic Priors

- Conjecture

  Perception will be largely dominated by priors when data is extremely noisy, e.g. random motion

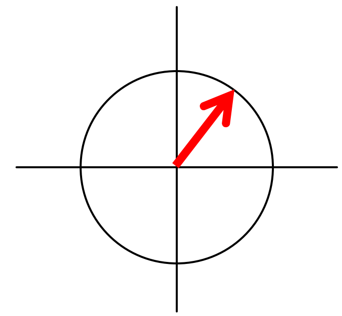
Exp 2: Motion Perception in a Single Trial
N=50, C=0

Stimulus

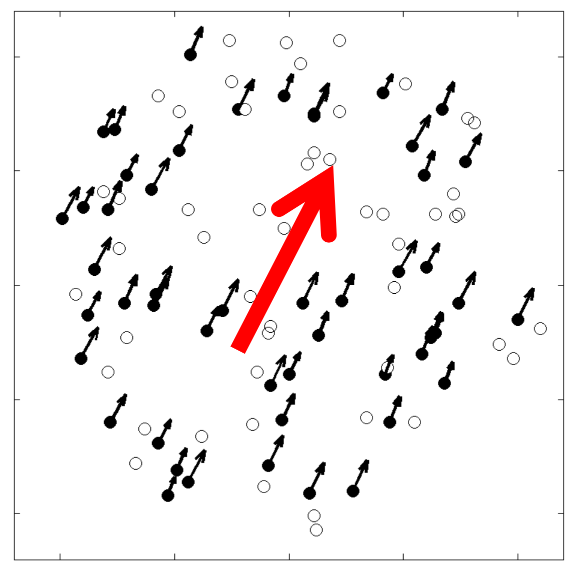


N=50, C=0%

Human perceived global motion direction: 57°

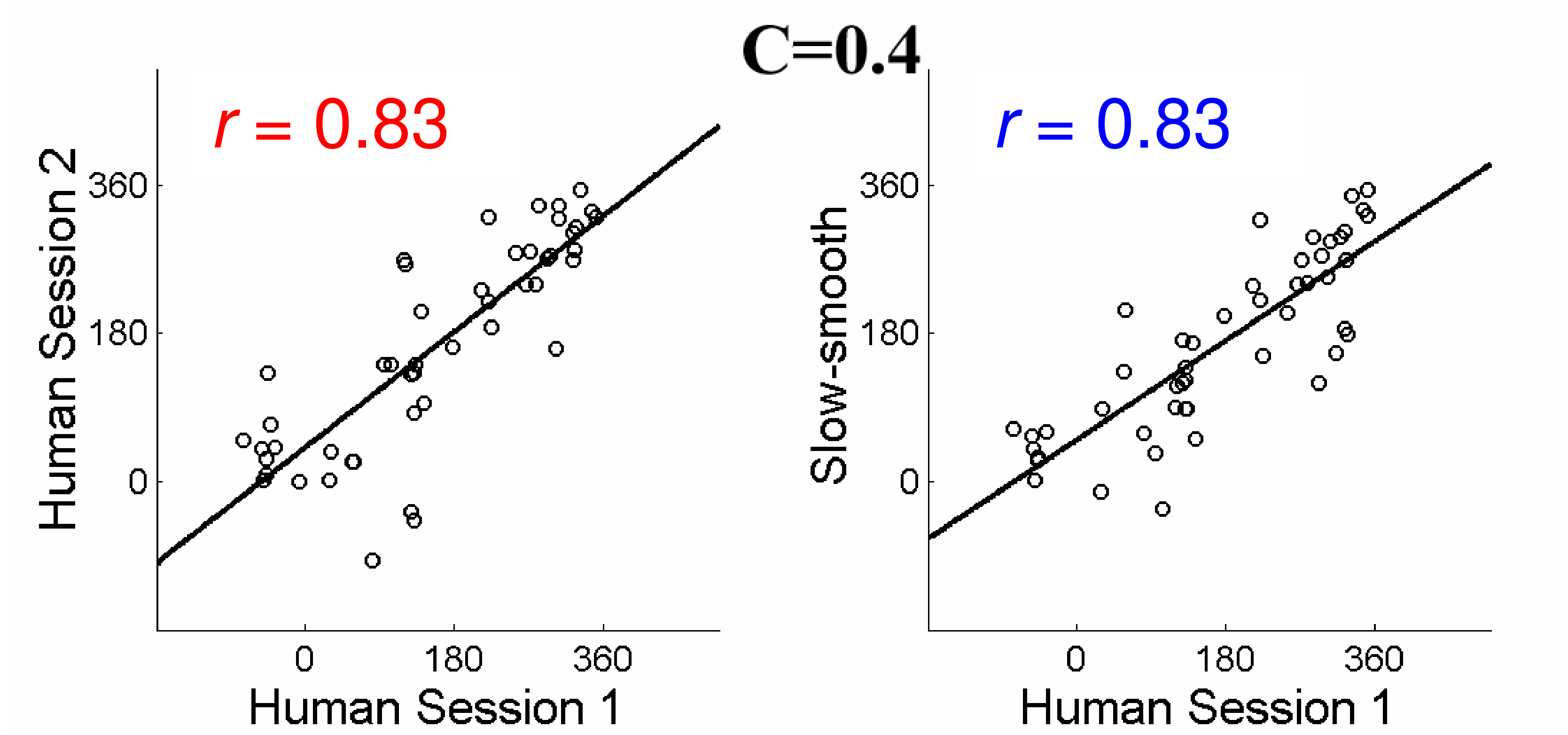


Model predicted global motion direction: 64°



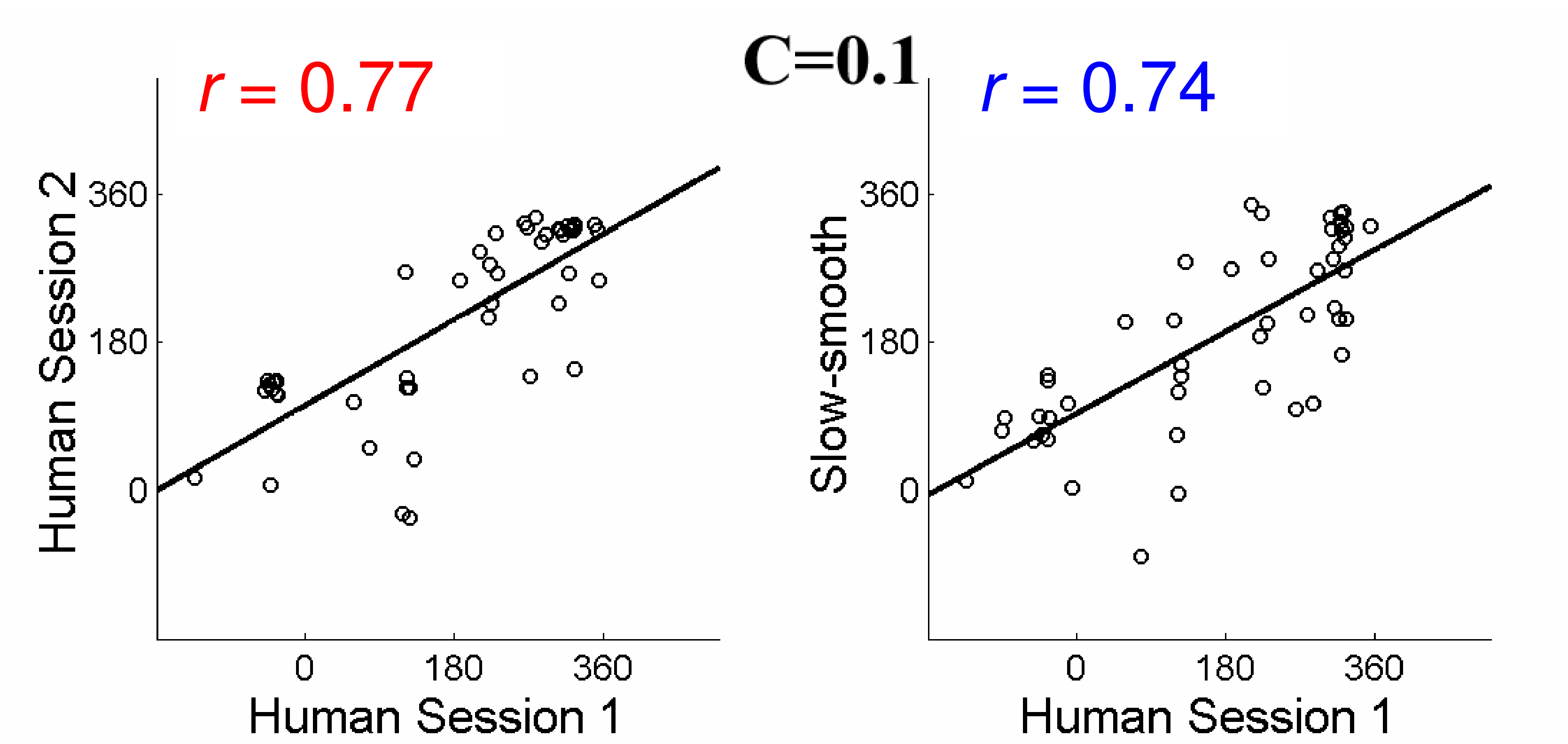
Trial-by-trial Correlation Between Human Perception and Model Prediction

Subject BR



Trial-by-trial Correlation Between Human Perception and Model Prediction

Subject BR

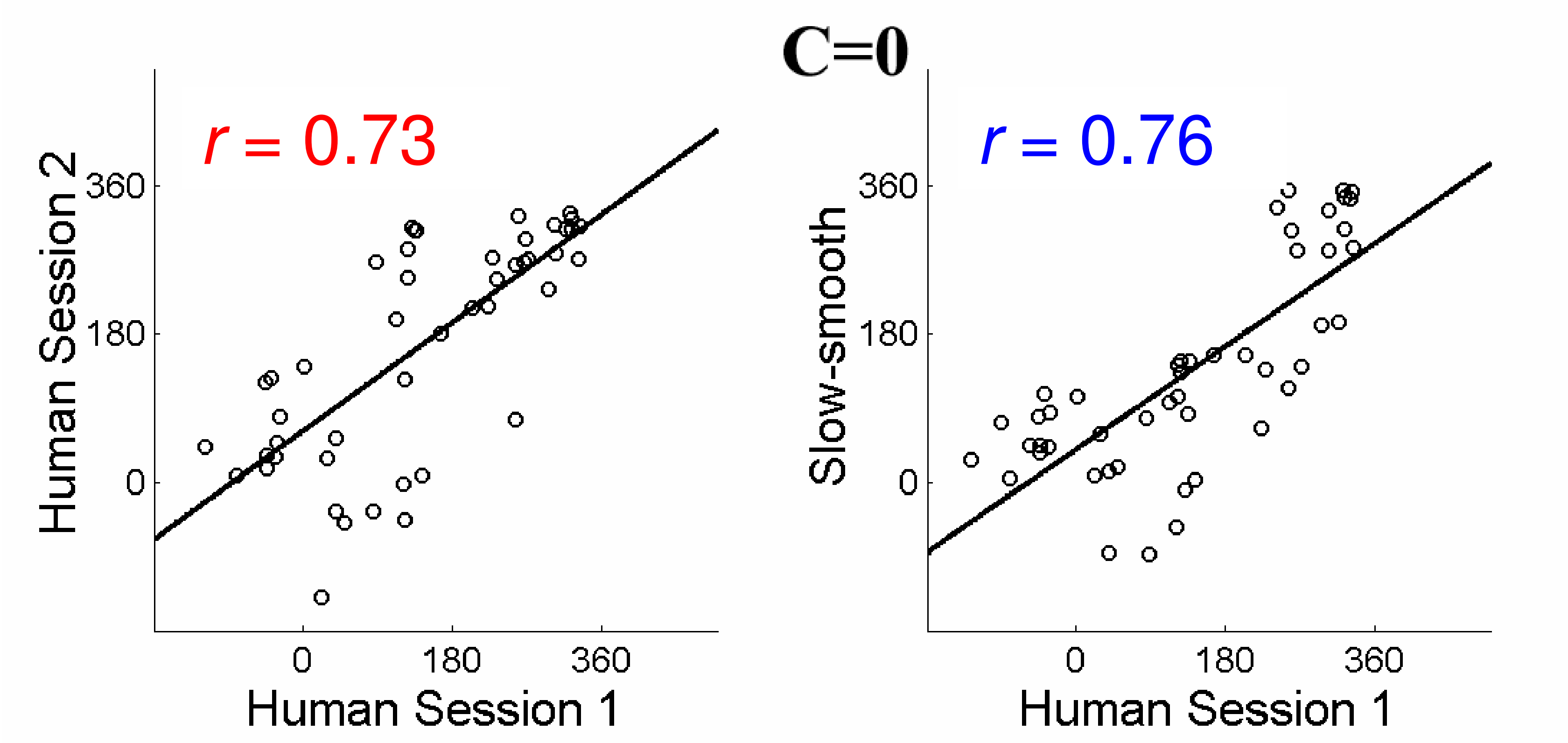


Can humans perceive any structure in random motion?

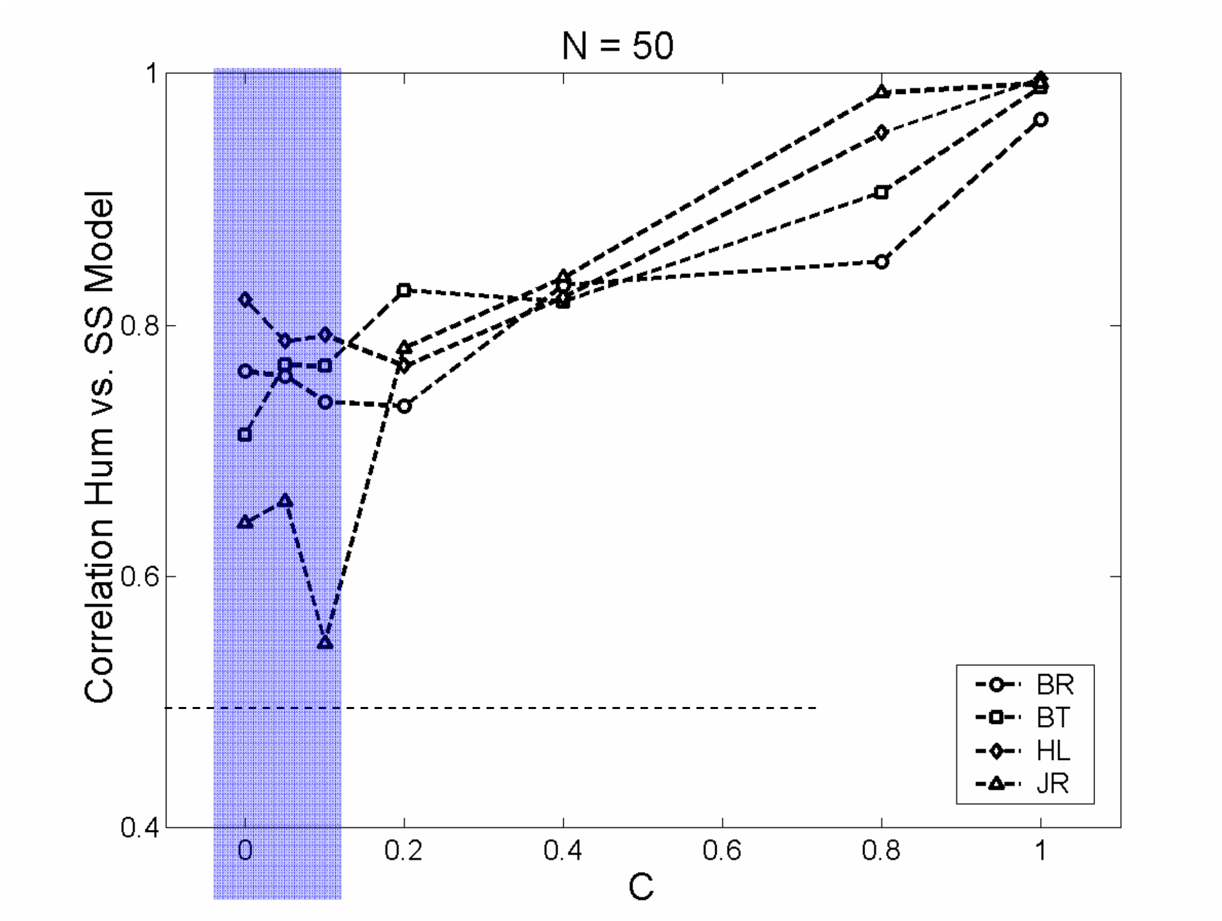
Can Bayesian model with Slow & Smooth prior predict human perception?

Trial-by-trial Correlation Between Human Perception and Model Prediction

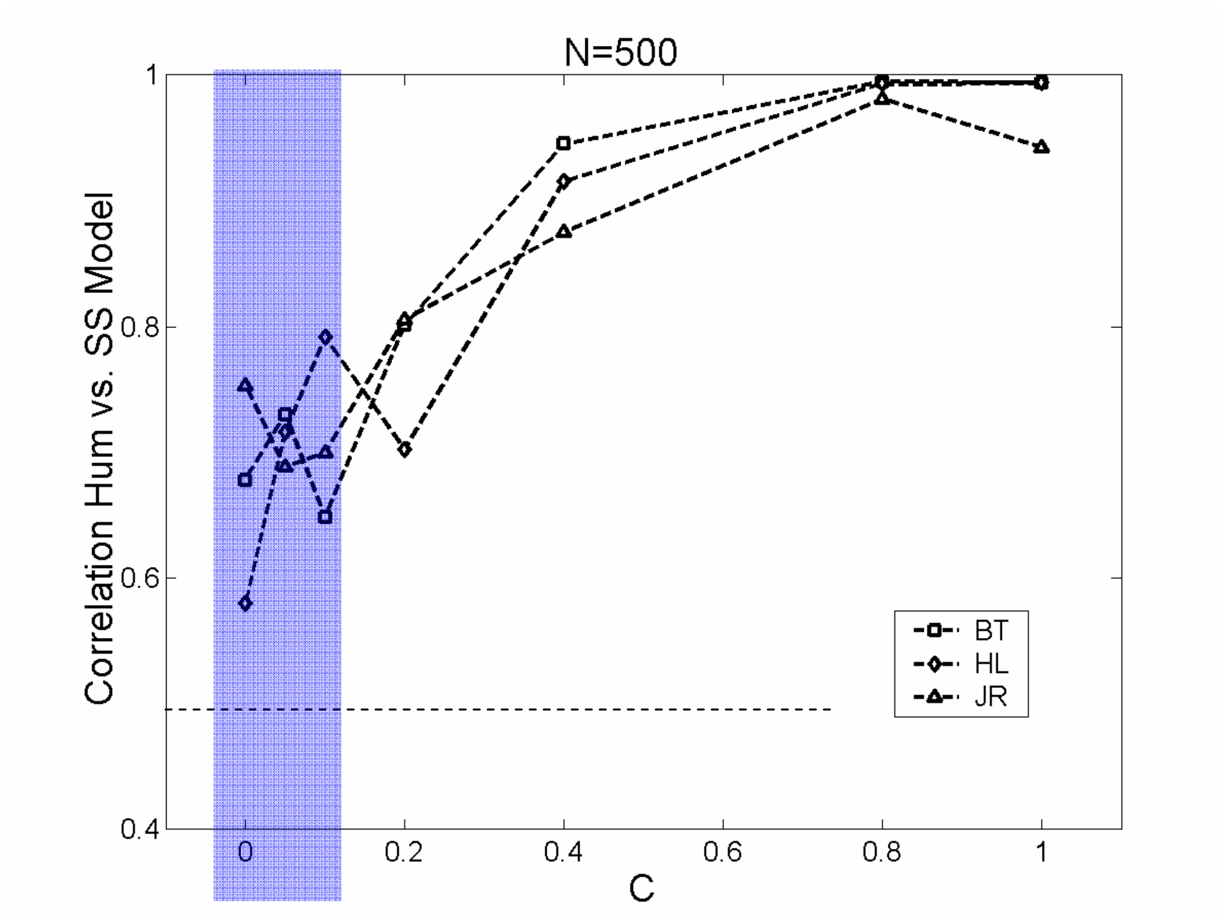
Subject BR



Trial-by-trial Correlation, N=50



Trial-by-trial Correlation, N=500



Conclusions

- Synergy between perceptual experiments and Bayesian framework
- Ideal observer with specific experimental prior predicts human performance qualitatively
- Bayesian model with Slow & Smooth generic prior predicts human performance both qualitatively and quantitatively
- Model predicts new phenomena, including perceiving structure in random motion