



Bayesian inference as a unifying model of auditory-visual integration and segregation



U. R. Beierholm¹, S. R. Quartz¹, L. Shams^{1,2}

1. Computation and Neural Systems, Caltech, Pasadena 2. Dept. of Psychology, UCLA, Los Angeles

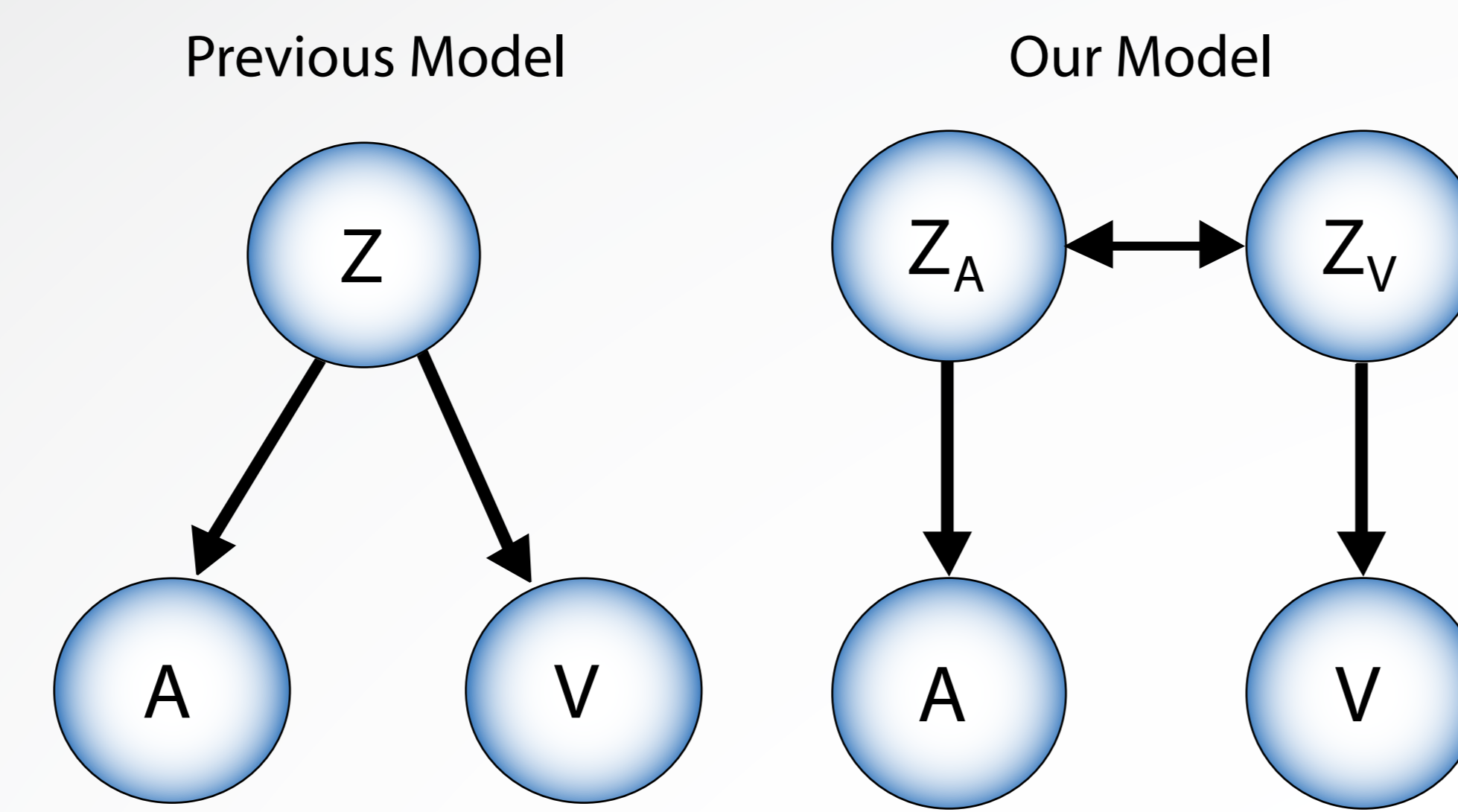
email: beierh@caltech.edu, steve@hss.caltech.edu, ladan@psych.ucla.edu

Abstract: To optimally react to its surroundings, the nervous system must decide what the relations are between stimuli of different modalities. Given the statistics of a variable environment, the best statistical strategy to decide whether and how to integrate or segregate sensory information is a Bayesian one that optimally combines the available evidence (likelihood) with previous knowledge (prior).

To examine whether humans perform in accordance with such a model, we used auditory-visual spatial perception as a test paradigm. Subjects were presented with audio and/or visual stimuli, and were asked to report both their perceived locations in each trial. Using the unisensory data (to calculate the likelihood functions) and the priors (estimated from a different dataset and with no free parameters), we can predict how an observer using Bayesian inference would perform when presented with auditory and visual stimuli simultaneously, and we can compare these predictions with the data obtained from the observers in the auditory-visual conditions.

The human participants' responses were remarkably consistent with the predictions of the Bayesian ideal observer ($R^2=0.93$). An important feature of Bayesian inference is the statistical independence of the prior from the likelihoods. In a second experiment we tested whether this feature held for our test subjects by manipulating the stimulus parameters (such as contrast of the visual stimulus) to vary the likelihood functions and recalculated the priors. While participants performance was significantly influenced by the changes in the auditory-visual stimuli, indicating a change in their likelihoods, the estimated prior remained constant. The independence of the priors from the likelihoods, together with the excellent consistency between the model and subjects' responses, suggests that human auditory-visual integration/segregation indeed follows Bayesian inference.

Models of cue integration



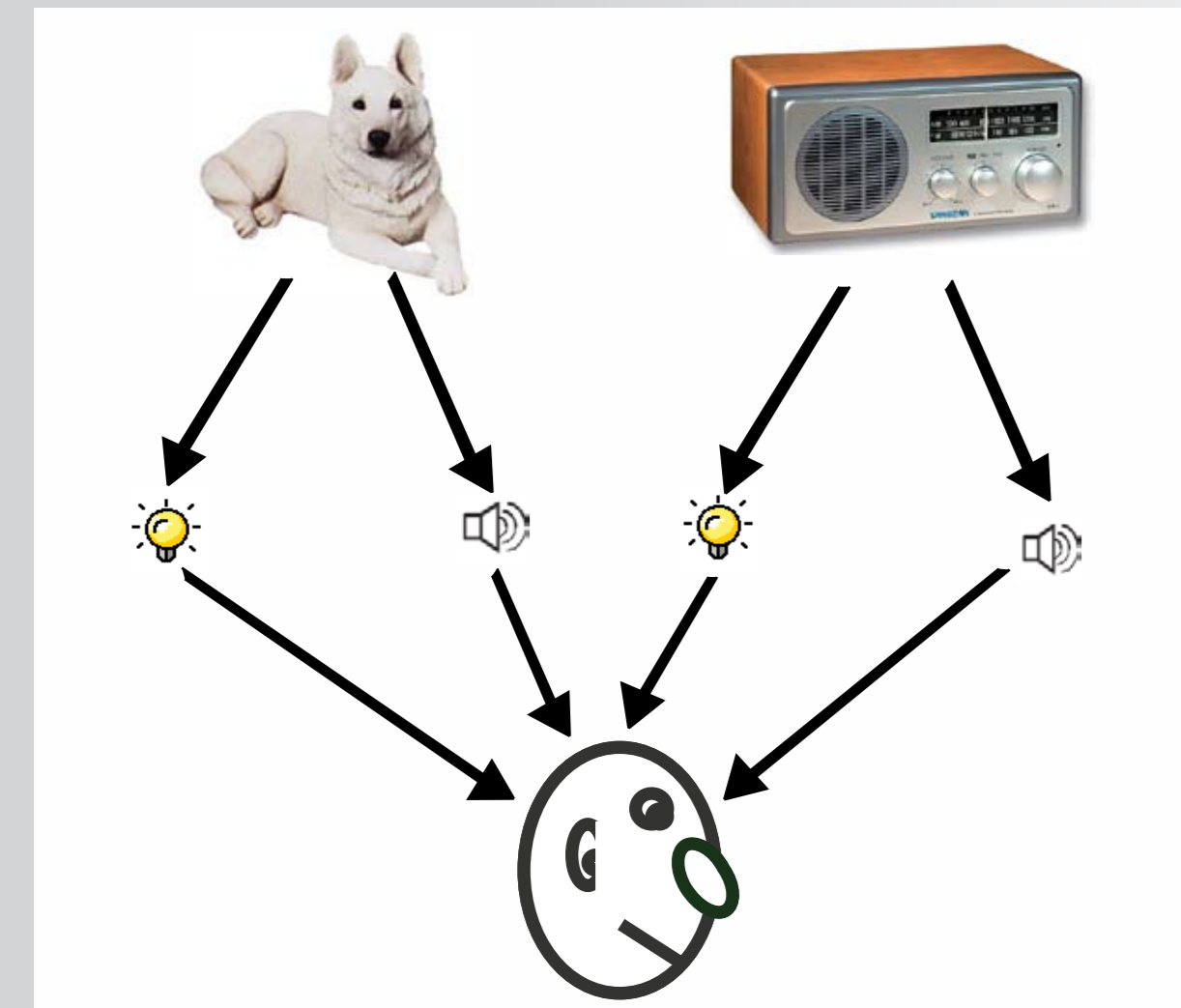
- * Previous model assumes fusion of cues and is equivalent to a Maximum Likelihood model (refs. 1,2).
- * Our full Bayesian model does not assume fusion.
- * Z, Z_A and Z_V are causes in the world giving rise to aud. and vis. sensations A and V.

$$P(Z_A, Z_V | A, V) = P(A, V | Z_A, Z_V) * P(Z_A, Z_V) / P(A, V)$$

$$P(Z_A, Z_V | A, V) = \frac{P(A | Z_A) * P(V | Z_V) * P(Z_A, Z_V)}{P(A, V)}$$

Motivating Questions:

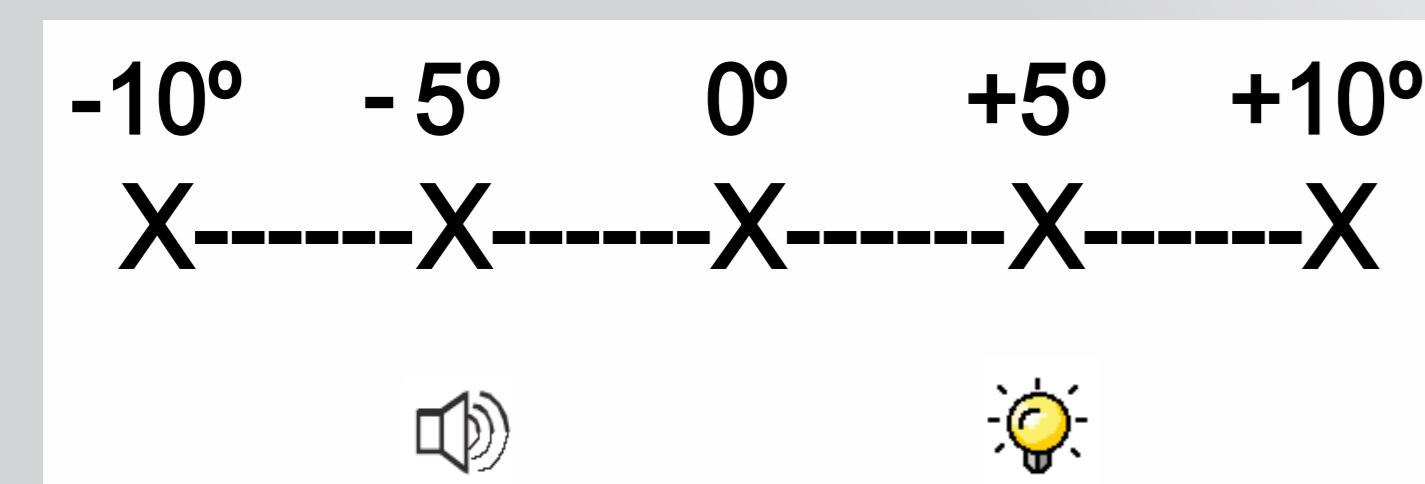
- 1) When should the nervous system combine information?
- 2) How does it combine information?



How does the nervous system know to combine the audio from the dog with the visual from the dog, and not with visual from the radio?

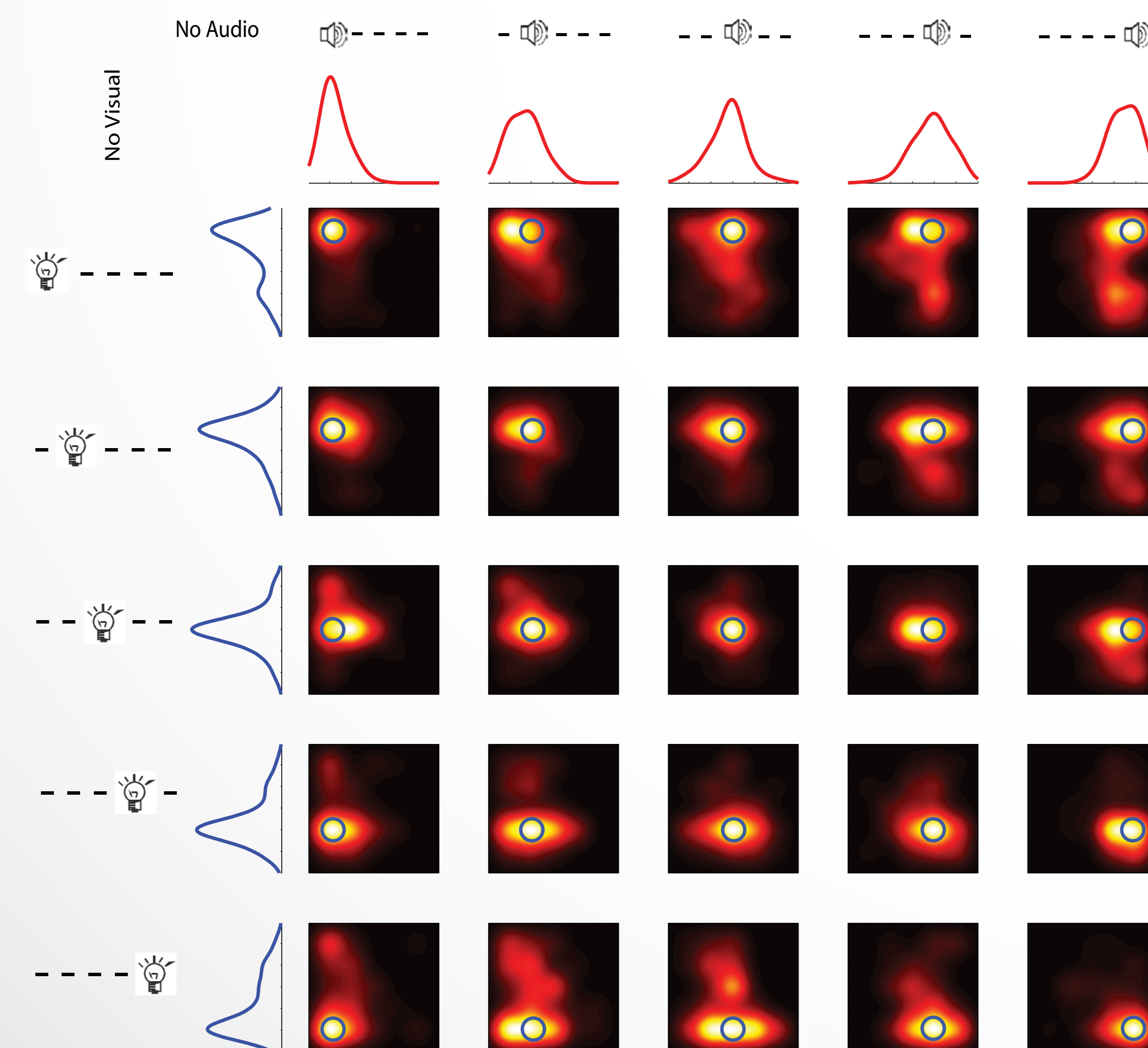
Dual task: report audio and visual location

Visual and auditory stimuli were presented at one of 5 locations, subjects were asked to report both perceived locations



- * Visual stimuli: low contrast Gabors on noisy background
- * Auditory stimuli: HRTF filtered (ref. 3) white noise
- * Presented at one of five locations, together and separately (5*5+5+5=35 conditions)
- * 15 trials of each condition per subject
- * 16 subjects split in two groups

Experimental data



Responses from 8 subjects are shown, 5 visual conditions combined with 5 auditory conditions plus null-conditions (first row and column). Color in each condition represents probability density of the response. A gaussian with SD of 2.5 deg was used as a Kernel density estimator.

Computing the model predictions

* We estimate the likelihoods from unimodal data

Likelihoods

$$P(A | Z_A) = P(Z_A | A) P(A) / \sum_A P(Z_A | A) P(A)$$

$$P(V | Z_V) = P(Z_V | V) P(V) / \sum_V P(Z_V | V) P(V)$$

* We estimate the prior from bimodal data from a **different** group of subjects

Prior

$$P(Z_A, Z_V) = \sum_{A, V} P(Z_A, Z_V | A, V) P(A, V)$$

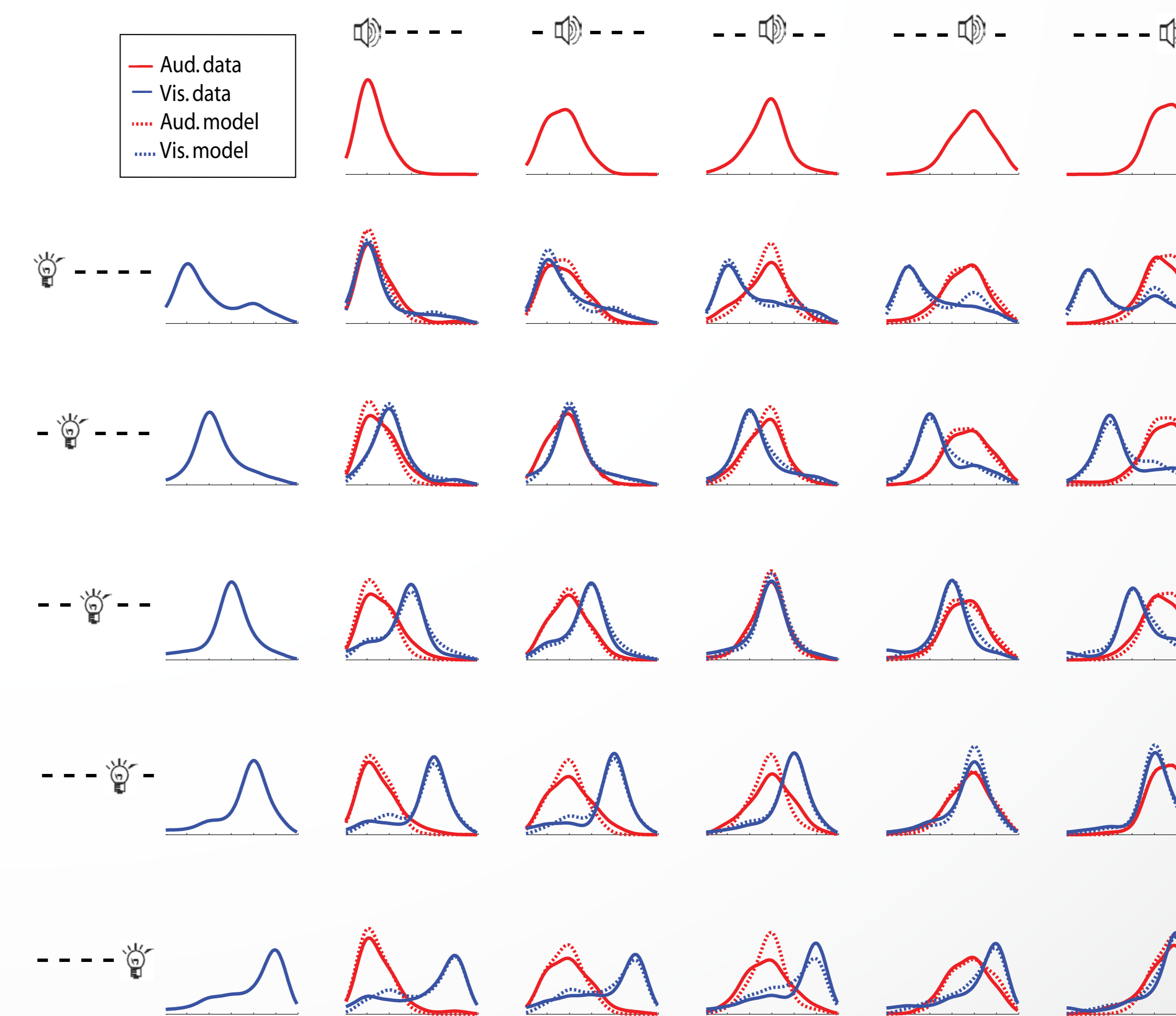
* We assume P(A, V) uniform.

Comparing the model to subjects responses

Estimating the prior from one group of subjects (N=8) and applying the model to the data from another group of subjects, we find a good correlation between the experimental data and the model (see below).

$$R^2 = 0.93$$

Experimental data vs model

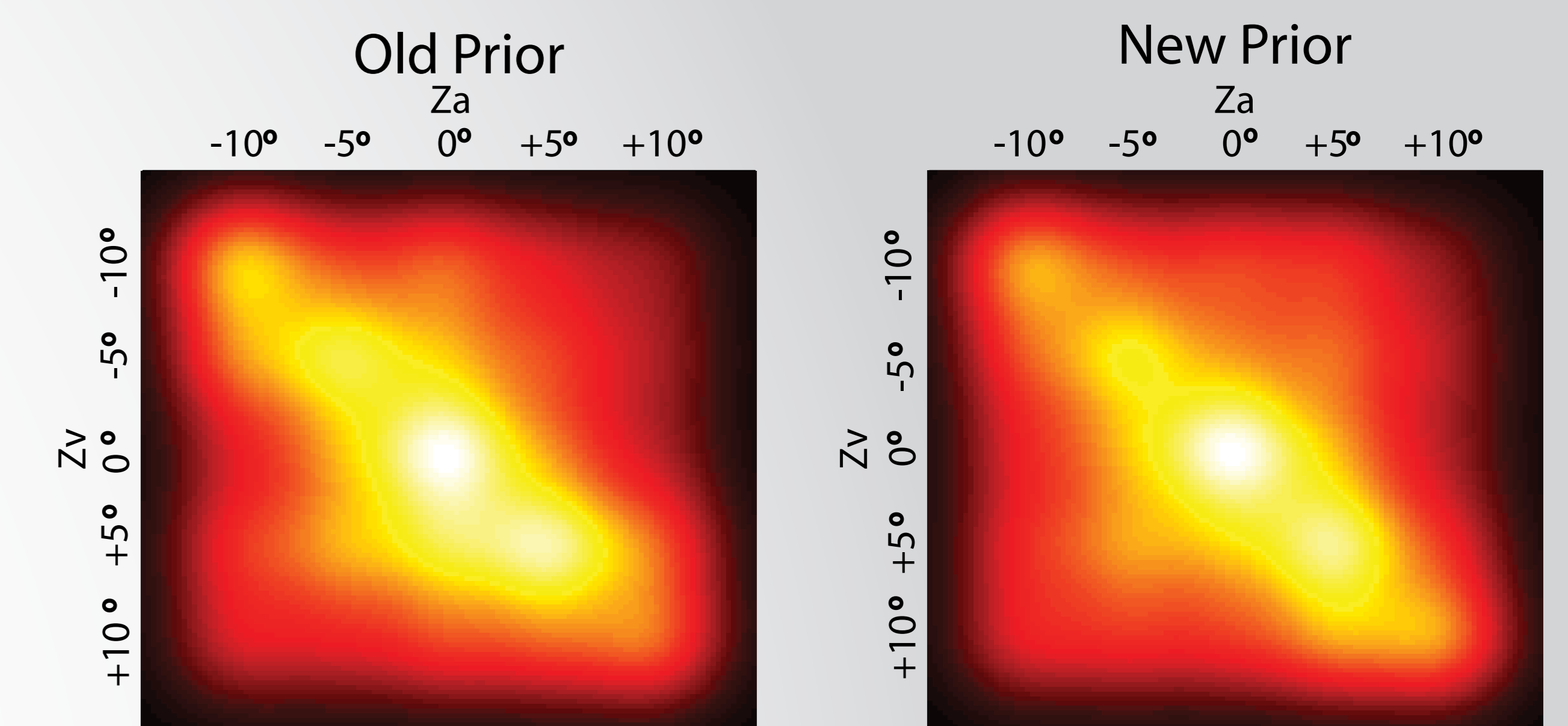


Comparison of marginalised posteriors for data and model. The correspondence is excellent for most conditions ($R^2=0.93$ for all data).

Are the likelihood and priors independent?

In a Bayesian model the prior should be independent of the current evidence. We tested this by performing the same experiment on the same subjects with a higher visual contrast.

- * Subject visual posteriors changed by ~20%, whereas
- * Priors only changed by 3%,
- * Suggests that priors are indeed independent of the stimulus.



Priors from same 15 subjects for 4.5% visual contrast and 5.5% visual contrast. There were three months between experiments.

Conclusion:

- * Human performance on task is Bayes optimal
- * Model explains wide range of discrepancies, from none to large, arguably explaining fusion and segregation
- * The Bayesian prior as estimated in our experiment is independent of the evidence, providing further support for the hypothesis of Bayesian inference
- * Given our previous results (ref. 4), this suggests that human audio-visual integration/segregation is optimal in general

References:

1. Battaglia, P.W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of Optical Society of America* 20, 1391-1397.
2. Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14, 257-262.
3. <http://sound.media.mit.edu/KEMAR.html>
4. Shams, L. & Ma, W. J., VSS 2004 and IMRF 2005 & Shams, L., Ma, W.J., Beierholm, U. R. (2005) *Neuroreport*, 16: 1923-1927.

This research was partly funded by a grant from the David and Lucile Packard Foundation.