Configural processing of 2D shape

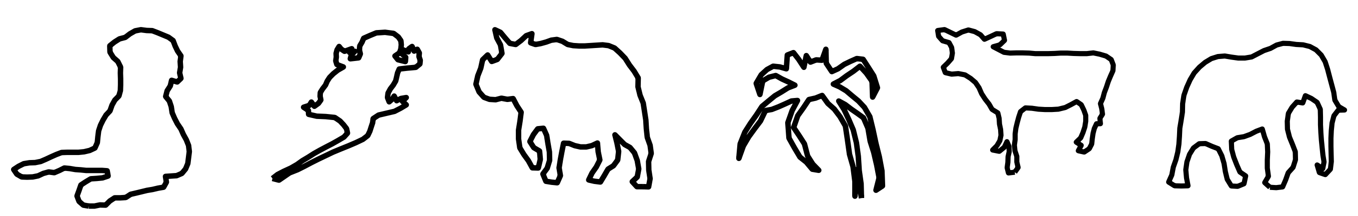
Shaiyan Keshvari, Ingo Fründ, and James H. Elder

QUESTION

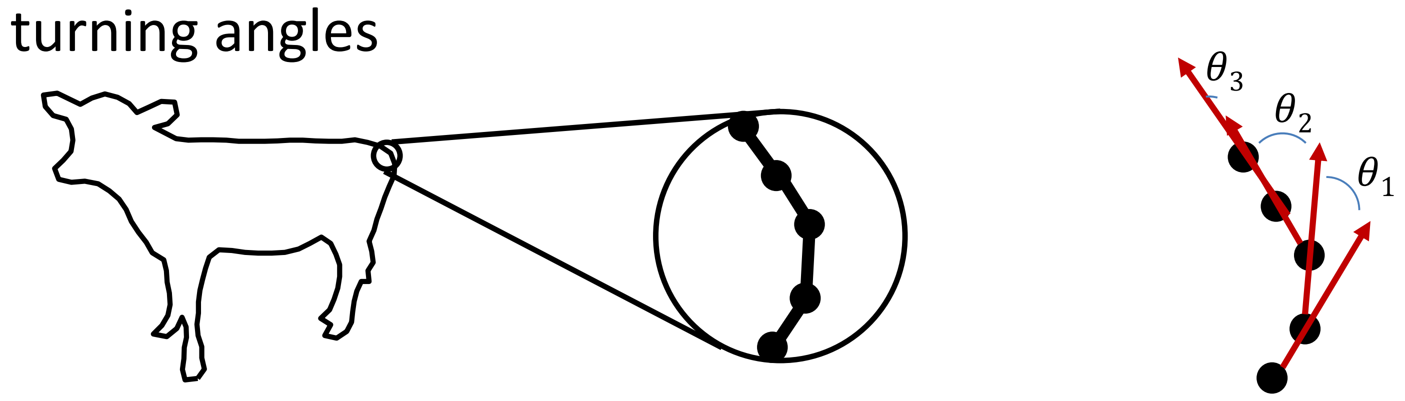
- Recent work suggests that object recognition in deep networks is based upon pooling of local features rather than global configural processing [1-3].
- Humans, while sensitive to local shape features (e.g., curvature), can easily discriminate global natural shapes from 'metamers' with matched curvature statistics [4,5].
- Does this human ability to discriminate global shapes emerge from a **pooling** of more complex local shape features (as in deep network models) or **configural** processing?

STIMULI

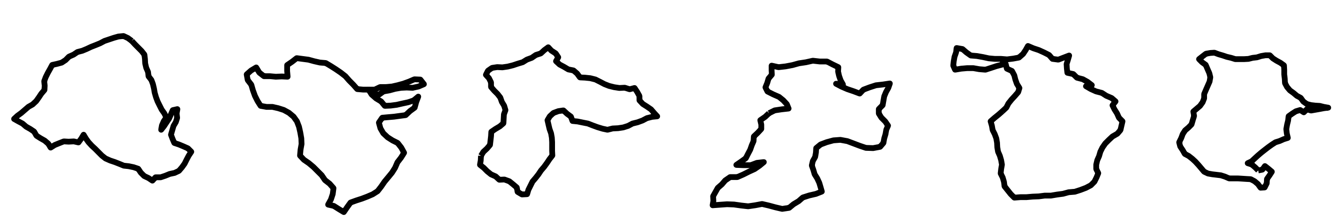
- Animal shapes are approximated as 120-point equilateral polygons.



- Discrete curvature is represented by the sequence of turning angles

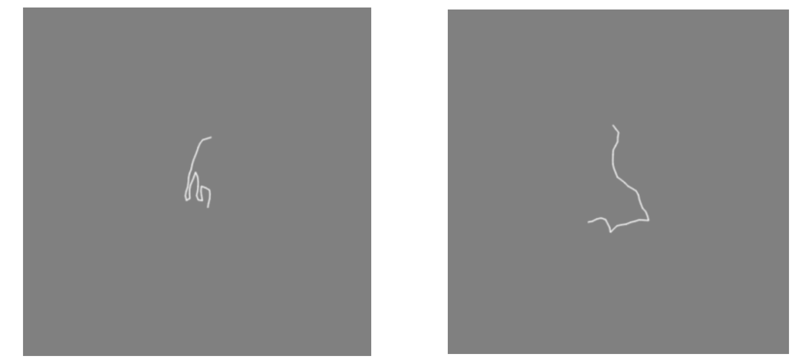


- An MCMC method is used to create a set of local curvature metamers: simple closed contours with the same marginal distribution of turning angles [4,5].

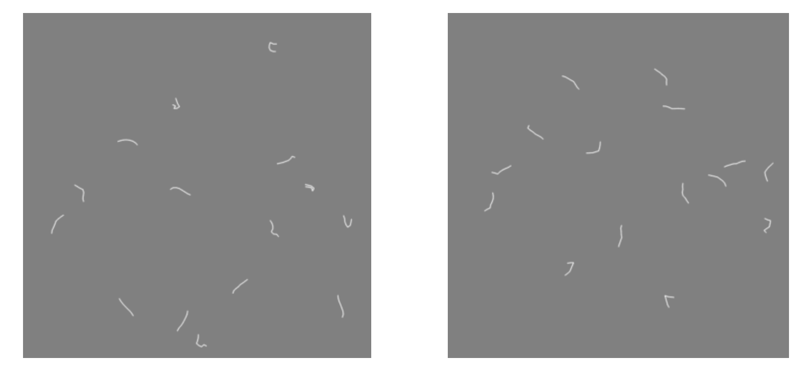


PROCEDURE

- In a 2IFC task, observers are shown random fragments of animal contours and random fragments of metamer contours and asked to identify which are from an animal.
- **Single-fragment condition:** Each stimulus consists of a single fragment. Length varied from 2 to 120 segments.

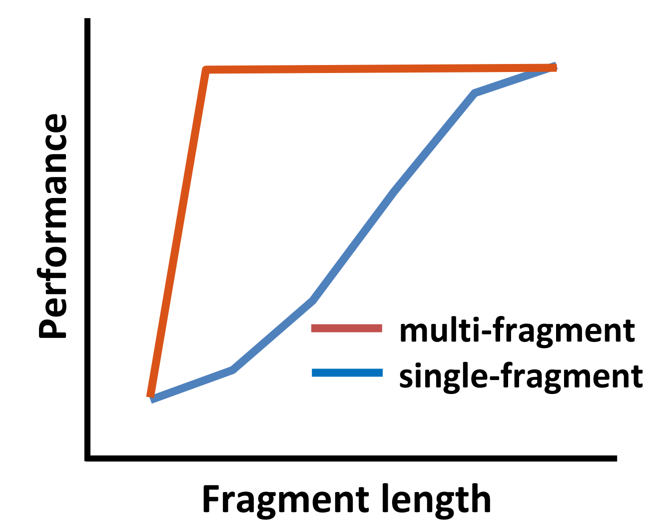


- **Multi-fragment condition:** Each stimulus consists of 1 to 120 fragments drawn from different shapes. Length of fragments varied inversely so that each stimulus has a total of 120 turning angles.

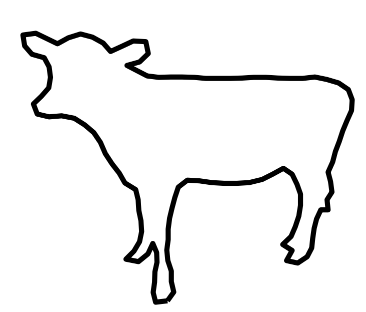


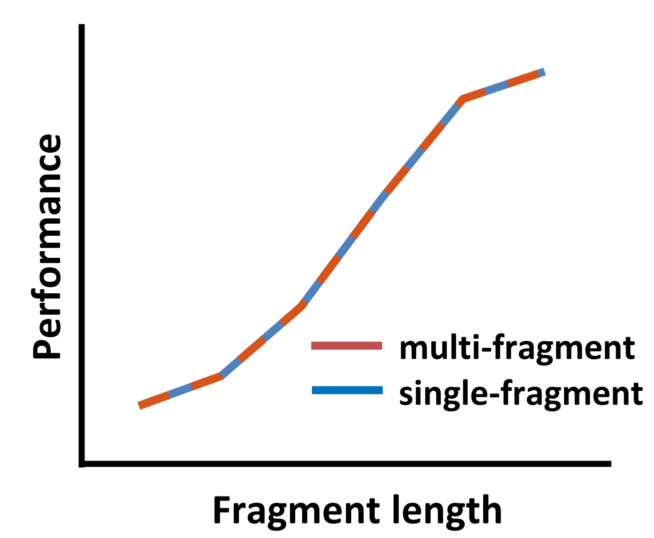
PREDICTIONS

- **Pooling model:** Predicts higher performance for the multi-fragment condition, since there are more local features.

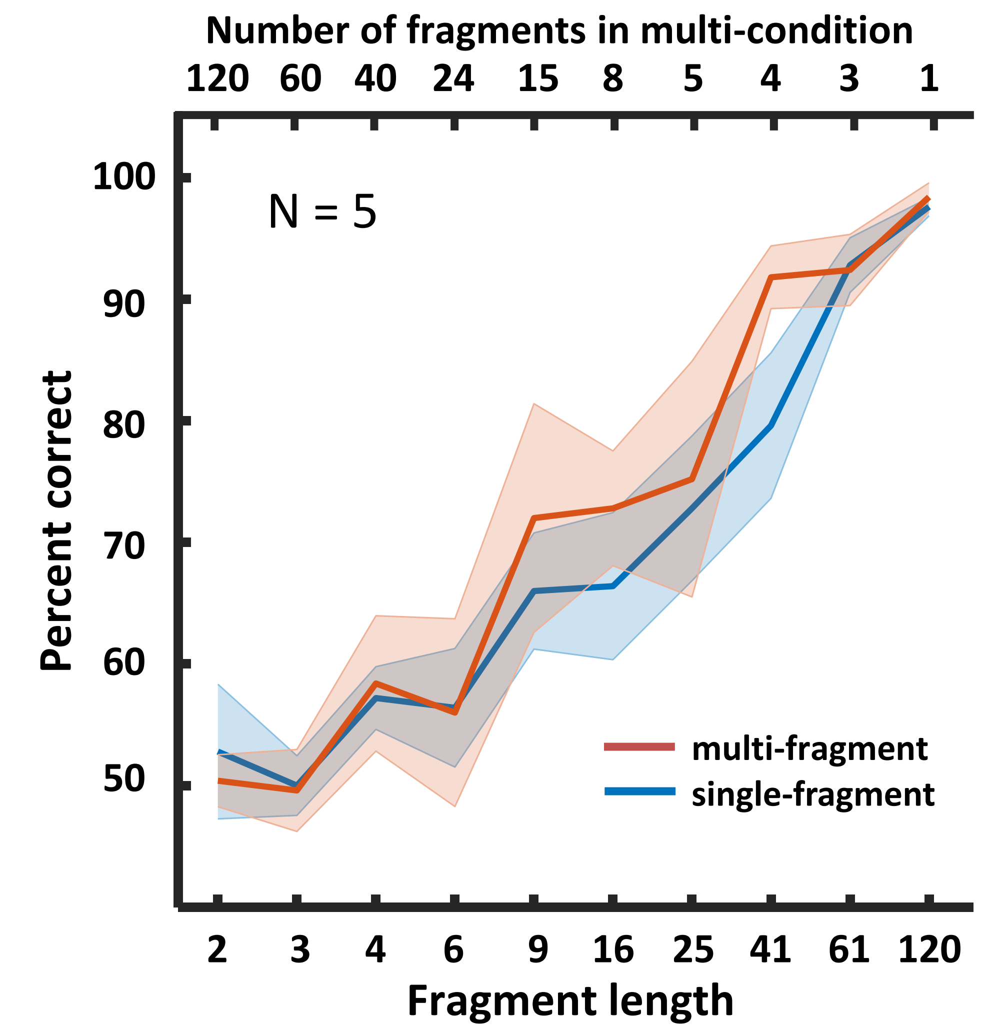


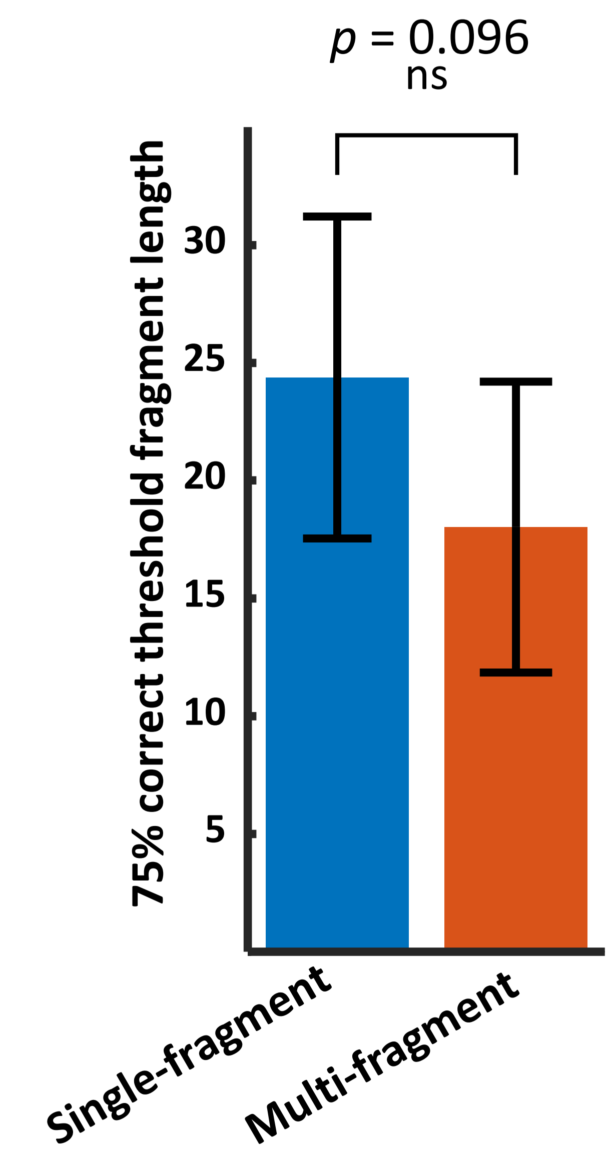
- **Configural model:** Predicts similar performance for single- and multi-fragment conditions, since the maximum fragment length is the same.





RESULTS





Main Findings:

- Human performance similar in single- and multi-fragment conditions, indicating dominance of configural processing.
- No indication of asymptote in multi-fragment condition, indicating global processing.
- This differs dramatically from the pooling behavior of deep network models.

References

[1] Baker, N., Lu, H., Erlikhman, G., & Kellman, P.J. (2018). *PLoS computational biology 14*(12), e1006613.
[2] Brendel, W., Bethge, M. (2019). *International Conference on Learning Representations (ICLR)*.
[3] Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2019). *International Conference on Learning Representations (ICLR)*.
[4] Fründ, I. & Elder, J.H. (2015). *Computational and Systems Neuroscience (Cosyne)*.
[5] Elder, J.H., Oleskiw, T.D. & Fruend, I. (2018). *Journal of Vision*, *18*(12):14.