

Local contour symmetry facilitates the neural representation of scene categories in the PPA.

John Wilder (jdwilder@cs.toronto.edu)

Department of Psychology and Department of Computer Science, University of Toronto
100 St. George Street, Toronto ON, Canada

Morteza Rezanejad (morteza@cim.mcgill.ca)

School of Computer Science, Centre for Intelligent Machines, McGill University
3480 University Street, Montreal, PQ, Canada

Kaleem Siddiqi (siddiqi@cim.mcgill.ca)

School of Computer Science, Centre for Intelligent Machines, McGill University
3480 University Street, Montreal, PQ, Canada

Allan Jepson (jepson@cs.toronto.edu)

Department of Computer Science, University of Toronto
6 King's College Rd, Toronto ON, Canada
Samsung Toronto AI Research Center
101 College St., Toronto, ON, Canada

Sven Dickinson (sven@cs.toronto.edu)

Department of Computer Science, University of Toronto
6 King's College Rd, Toronto ON, Canada
Samsung Toronto AI Research Center
101 College St., Toronto, ON, Canada

Dirk B. Walther (bernhardt-walther@psych.utoronto.ca)

Department of Psychology, University of Toronto
100 St. George Street, Toronto ON, Canada

Abstract

Human observers can rapidly classify real-world scenes into their natural categories (e.g. beaches, mountains). It is unclear what neural mechanisms underlie this rapid processing of scenes. In a previous behavioral study, we demonstrated that local ribbon symmetry facilitates scene classification. Here we manipulate the ribbon symmetry content of line drawings of real-world scenes and then decode scene categories from patterns of voxel activities of the observers obtained via fMRI. We can decode scene categories from the parahippocampal place area (PPA) more easily from symmetric scenes than asymmetric scenes. In earlier visual areas the decoding accuracy for symmetric and asymmetric scenes was not significantly different. This suggests that the benefit for symmetric scenes in both behavior and fMRI is not solely driven by a lower-level preference for symmetry. Instead, ribbon symmetry may be uniquely informative for scene categorization.

Keywords: Scene perception; Symmetry; Gestalt; perceptual grouping; fMRI

Introduction

As soon as a person opens their eyes, their visual experience is that of a cohesive world, composed of individual objects and

surfaces. The human visual system must transform low-level image properties into conscious percepts of objects and continuous surfaces. Visual processing takes the responses from individual photo-receptors and groups them through a hierarchical process, where the elements are grouped into larger structures, and finally into objects and scenes. An observer can classify a scene when it is presented very briefly (Potter & Levy, 1969; Thorpe, Fize, Marlot, et al., 1996; VanRullen & Thorpe, 2001). Human observers can rapidly classify line drawings of real-world scenes, even though they do not contain the richness of a photograph (Biederman, Mezzanotte, & Rabinowitz, 1982; Walther, Chai, Caddigan, Beck, & Fei-Fei, 2011). The speed at which photographs or line drawings can be classified demonstrates that the grouping process carried out by the visual system occurs very rapidly.

The Gestalt psychologists proposed qualitative grouping rules describing what image features influence this process. While many of these grouping rules have been implemented in concrete, quantitative algorithms, a rigorous algorithm for the cue of symmetry has been lacking. We recently described an algorithm for measuring contour symmetry in a line drawing of a scene (Wilder et al., 2019), and showed that the presence of symmetry greatly facilitates scene categorization.

Here we look for the neural correlates of this local symmetry



effect, using fMRI. Previous work on global mirror symmetry as well as axiomaticity of shapes (related to the local symmetry we measure here) has found an influence of symmetry in V3, V4, and LOC. Local contour symmetry is distinct from these previously studied types of symmetry, so it is not clear if we should expect an effect of symmetry in these areas. The parahippocampal place area has been shown to have activity consistent with behavioral scene classification as well as 3D scene geometry (Ferrara & Park, 2016; Lescroart & Gallant, 2019), so we hypothesize that PPA will be influenced by local symmetry in line drawings of natural scenes. To test this hypothesis we showed intact, symmetric, and asymmetric scenes to observers and decoded scene categories from voxel activity patterns in several ROIs (V1, V2, V3, LOC, PPA, OPA, and RSC).

Methods

Scenes dataset

We showed participants line drawings of six categories of real-world scenes (Walther et al., 2011; Wilder et al., 2019): Beaches, Forests, Mountains, Cities, Highways, and Offices. The line drawings we generated by artists tracing the important lines from a set of color photographs. There are 475 images across all scene categories.

Scoring Symmetry

Our measure of symmetry is based upon the medial axis transform, which is a way to transform a shape from the pixels along its boundary to the central axis in each part of the shape. The medial axis marks where in the shape the contour is equidistant on both sides. Details of our method can be found in (Wilder et al., 2019). We will briefly describe the process here.

We compute symmetry by starting with a line drawing of a scene, and taking the Euclidean distance transform, which, for each pixel in the image, measures the distance from that pixel to the nearest contour. From this distance image we can determine the gradient at each point. The medial axis lies at the points where this gradient flows outward. From the distance transform, we also know the distance from a medial axis point to the contour in each direction. We call this the radius function. As we move along the medial axis, we measure how the radius function changes. If there is no change in the radius, we are in a locally ribbon symmetric region, meaning that the contours on either side are parallel, as a curving ribbon of constant width. Before scoring the contours, we must score the axis points. Our score is based on the behavior of the radius function within a local region. We count the number of times the radius changes between neighboring axis points within that local region. If this number is high, we give a low symmetry score, because the contours are curving independently of one another. If the number of radii changes between neighboring axis points in this region is low, we give a high symmetry score. Once the axis is scored, we find the contour points that this axis point flows to (using the gradient of the

distance transform). Those contour points receive the score of their corresponding axis point. As our method finds the medial axis in all white-space regions of an image, there are two skeletons that correspond to each contour point, one on either side. The score we assign is the maximum score of the two axis scores.

Concretely, we consider a window of $2K + 1$ medial axis points centered at our target medial axis point Q . These points are Q_{-K}, \dots, Q_K , and $Q_0 = Q$. At medial axis point Q , the score $S(Q)$ is

$$S(Q) = \frac{\#\{Q_i | \forall i = \{-K, \dots, K-1\} \text{ where } \frac{|\mathcal{R}(Q_i) - \mathcal{R}(Q_{i+1})|}{\max(\mathcal{R}(Q_i), \mathcal{R}(Q_{i+1}))} \leq \tau\}}{2K}$$

where $\mathcal{R}(Q_i)$ represents the radius value at Q_i , and τ is a marginal threshold.

Stimuli

In our study we manipulated the amount of symmetric content in the stimuli we showed to our observers. We used either the original, intact line drawing, or we used a half-image, containing exactly half of the contour pixels of the intact image. To create the half-images we first rank ordered all contour pixels based upon their symmetry score. For the *Symmetric* images, we used the highest ranking half of the pixels, and for the *Asymmetric* images, we used the lowest ranking half of the pixels. Thus, the symmetric and asymmetric images combine to create the intact image, and they contain an equal number of contour pixels. You can see an example of a scored image, along with the accompanying splits in Fig. 1.

Each of the 475 images was transformed into three conditions, resulting in 1425 total images that may be shown to participants.

Participants

Twenty-two participants were recruited from University of Toronto Facebook groups, and were paid for their participation. Only 19 participants' data were used because three participants fell asleep during multiple runs of the experiment.

Design and Procedure

We used a block design fMRI experiment. Participants were scanned in nine runs. Each run contained 18 blocks, one block for each combination of scene category and image condition. In each block, participants were shown 8 scenes from the same category and image condition. Each image was shown for 800 ms, with a 200 ms gap between images. There was a 10s fixation period prior to the first block and after each block. Participants were instructed to maintain fixation for the entire run, and a fixation mark was always on the screen to aid the participants.

Separate localizer scans were conducted in order to allow for the determination of areas V1-V3, LOC, PPA, OPA, and RSC, for each individual participant.

Results

Multivariate Analysis

Our main analysis is a multivariate analysis where we train linear support-vector machines (SVMs) to classify scene categories from the activity patterns of voxels elicited by participants viewing the scene images. The classifiers were trained using leave-one-run-out cross-validations. Separate classifiers were trained for each image condition.

In all ROIs decoding accuracy was above chance, for each image condition. We were specifically interested in the relative performance between the symmetric and asymmetric conditions. For each ROI we conducted a paired t-test on the decoding accuracies from the symmetric and asymmetric conditions. We found significantly better performance in the symmetric condition for both PPA and OPA ($p < 0.05$). In RSC we found significantly better performance in the asymmetric condition ($p < 0.05$).

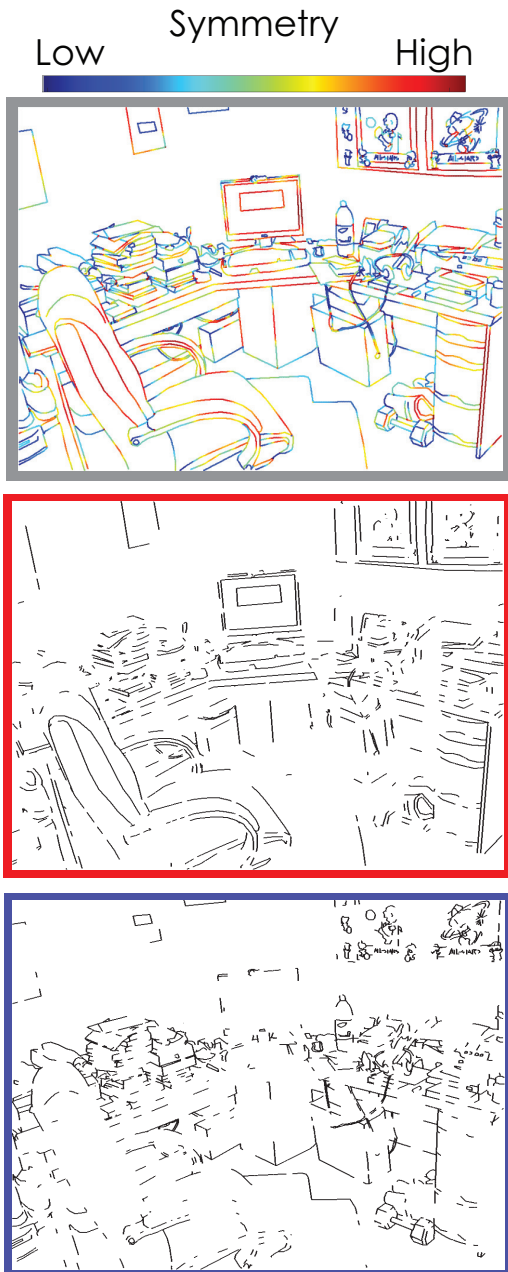


Figure 1: Example Office line drawing. The top image (gray border) shows the intact image with the symmetry score overlaid on top of the contour pixels. After the image is scored the pixels are rank ordered, and the top half is used to create the symmetric image (in red) and the bottom half is used to create the asymmetric image (in blue). Each of these images contain exactly half of the contour pixels of the intact image, and combine to form the intact image.

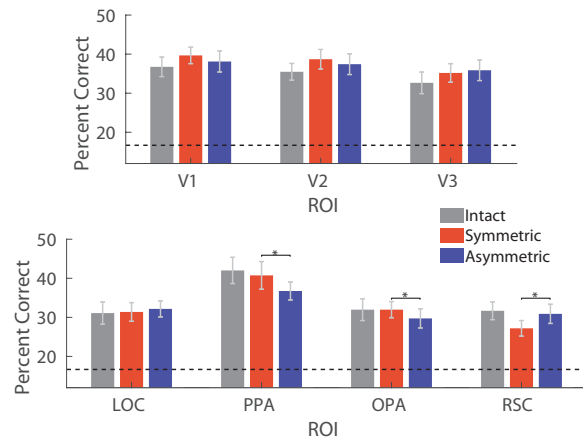


Figure 2: Decoding accuracy for each ROI (V1, V2, V3, LOC, PPA, OPA, RSC) for intact (gray), symmetric (red), and asymmetric (blue) scenes. * denotes $p < 0.05$

Univariate Analysis

We also performed a univariate analysis in each ROI, for each image condition. In all ROIs, the activity for the intact condition was the lowest. In V1, V2, and V3 there was significantly more activity in the asymmetric condition than the symmetric condition (V1 and V2 $p < 0.001$, V3 $p < 0.05$). Also, in the OPA, there was significantly more activity for the symmetric condition than the asymmetric condition ($p < 0.05$). No differences were found in LOC, PPA, and RSC.

Discussion

As hypothesized, we found a significant effect of symmetry versus asymmetry in scene selective ROIs.

Previous work has suggested that the axial structure of shapes are represented in V3 (Lescroart & Biederman, 2013). Additionally, single cell recordings in early visual cortex have

shown that some neurons respond highly when their receptive field is centered on the medial axis (Lee, 1996). Our scoring of symmetry is based upon the medial axis, thus these areas could have been involved in processing our stimuli, but we failed to find any significant results. Our line drawing scene stimuli are quite different from the stimuli of the previous studies and are not isolating single medial axes. This could be why we found no preference for symmetry in these areas. Additionally, we restrict our symmetry measure to ribbon symmetry. While many objects have ribbon symmetry in the real world, when projected into the image plane, due to perspective foreshortening, they taper in the image. We may have found no influence of ribbon symmetry in earlier areas that encode object and shape parts because our method does not score tapering regions highly.

In addition to the significant difference between decoding the scene categories for symmetric scenes and asymmetric scenes in PPA, we found a similar effect in OPA, and the opposite effect in RSC. All three of these areas are involved in processing scenes. Recent studies have characterized the ways in which the functions for these areas differ (Dillon, Persichetti, Spelke, & Dilks, 2018; Persichetti & Dilks, 2018). The PPA was reported to be highly activated during categorization but not for a navigation task, while OPA was more active for navigation than categorizations (Persichetti & Dilks, 2018). RSC showed low activity in both tasks. Both PPA and OPA were sensitive to changes in scene layout, but RSC was not (Dillon et al., 2018). Our work is consistent with those findings; manipulating symmetry information, which is related to scene layout, affects both PPA and OPA in a consistent manner, while RSC had the opposite effect.

Conclusion

Human rapid scene classification is aided by local symmetry. Here we show that scene selective cortex represents local symmetry in its voxel patterns. Area PPA mirrors that of categorization behavior, where symmetric scenes are easier to classify than asymmetric ones. Ribbon symmetry may be uniquely informative for scene categorization.

Acknowledgments

This work was supported by NSERC, SSHRC, Samsung, and Sony.

References

- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143–177.
- Dillon, M. R., Persichetti, A. S., Spelke, E. S., & Dilks, D. D. (2018). Places in the brain: bridging layout and object geometry in scene-selective cortex. *Cerebral Cortex*, *28*(7), 2365–2374.
- Ferrara, K., & Park, S. (2016). Neural representation of scene boundaries. *Neuropsychologia*, *89*, 180–190. doi: <https://doi.org/10.1016/j.neuropsychologia.2016.05.012>

- Lee, T. S. (1996). Neurophysiological evidence for image segmentation and medial axis computation in primate v1. In *Computation and Neural Systems: Proceedings of the Fourth Annual Computational Neuroscience Conference*.
- Lescroart, M. D., & Biederman, I. (2013). Cortical representation of medial axis structure. *Cerebral Cortex*, *23*(3), 629–637.
- Lescroart, M. D., & Gallant, J. L. (2019). Human scene-selective areas represent 3d configurations of surfaces. *Neuron*, *101*(1), 178–192.
- Persichetti, A. S., & Dilks, D. D. (2018). Dissociable neural systems for recognizing places and navigating through them. *Journal of Neuroscience*, *38*(48), 10295–10304.
- Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10.
- Thorpe, S., Fize, D., Marlot, C., et al. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.
- VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception*, *30*(6), 655–668.
- Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences*, *108*(23), 9661–9666.
- Wilder, J., Rezanejad, M., Dickinson, S., Siddiqi, K., Jepson, A., & Walther, D. B. (2019). Local contour symmetry facilitates scene categorization. *Cognition*, *182*, 307–317. doi: <https://doi.org/10.1016/j.cognition.2018.09.014>