

Canonical views of scenes depend on the shape of the space

Krista A. Ehinger & Aude Oliva

Introduction

Preference for “canonical” views of objects in recognition, depiction, and imagery (Palmer, Rosch, & Chase, 1981). Are there canonical views of scenes? What determines the canonical view of a scene?

Experiment

1084 panoramic photos, each shown to 10 different workers on Amazon Mechanical Turk

On each trial, workers performed two tasks:

1. Name the location shown in the image (eg, “classroom”)
2. Rotate the image in a 360-degree viewer to show the “best view” of the location

Select the best view of a location.
Requester: cvd Reward: \$0.01 per HIT HITs Available: 100 Duration: 20 minutes
 Qualification Required: Total approved HITs is not less than 100, Location in US, HIT approval rate (%) is not less than 95

Select the best view of a location

This HIT uses a Flash interface -- in order to do it, you need a recent version of the Flash player. If you do not see an image below, or if you cannot interact with the image using your mouse, please SKIP THIS HIT.

1. Name the location shown below (e.g., kitchen, garden, stadium)

2. Image that you are a photographer, and you have been asked to take a photo of this location. Click and drag in the image below to rotate the view to get the best possible “snapshot” of this location. When you are happy with your view, click the submit button below.

Task window. The image appeared in an interactive viewer: users could rotate the view shown to simulate looking around in the scene.

Results

Agreement was generally high (Rayleigh’s test of nonuniformity returned $p < 0.01$ for 538 images (50% of images), $p < 0.05$ for 694 images (64% of images))

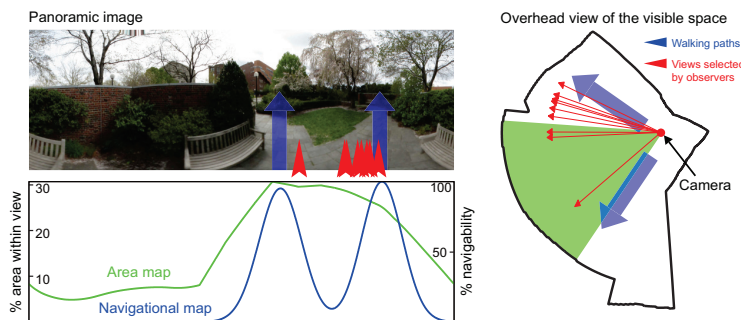
Examples (high agreement to low agreement):

<p>Panoramic image</p>	<p>Selected views</p>
------------------------	-----------------------

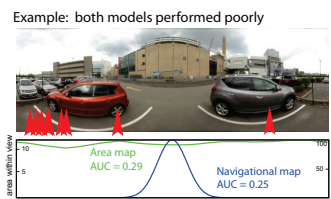
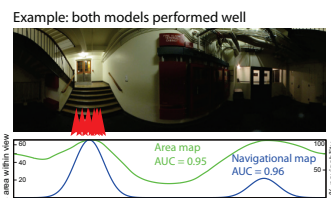
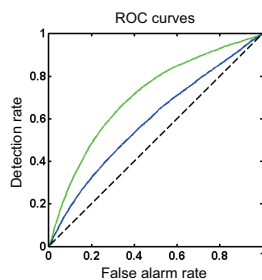
“Best” views selected by observers

Modeling the shape of the space

The boundaries of the space were obtained by outlining the ground plane and calculating the area around the camera. Navigational paths were marked by Mechanical Turk workers.

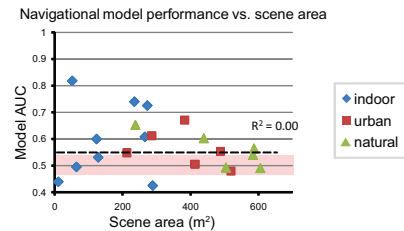
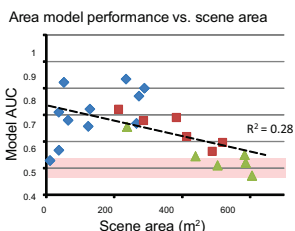
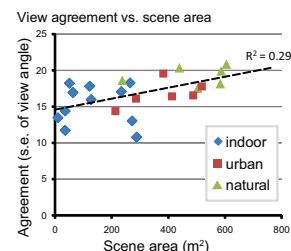


The area map represents the percent of the space visible in each direction. The navigational map represents navigability in each direction.



Model performance and scene area

View agreement was highest in small, indoor / man-made spaces and lowest in natural scenes. The area model also performed better in smaller, indoor spaces. The navigational model’s performance was not related to scene area.



Conclusion

There is high agreement on the “best view” of a scene. The best view of a scene is the one that shows as much of the space as possible, not necessarily the functional view for navigating in that space.

References

Palmer, S., Rosch, E., Chase, P., (1981). Canonical perspective and the perception of 40 objects. In Attention and Performance IX, Ed. J. Long, A. Baddeley (Hillsdale, NJ: Lawrence Erlbaum), pp. 135-151.