

Recognizing Scene Viewpoint using Panoramic Place Representation

Supplementary Material: Performance Table

Jianxiong Xiao
jxiao@mit.edu

Krista A. Ehinger
kehinger@mit.edu

Aude Oliva
oliva@mit.edu

Antonio Torralba
torralba@mit.edu

Massachusetts Institute of Technology

Abstract

*This document contains an extended version of Table 1 and 2 in the paper to show the performance by category. We show the alignment performance and testing performance for each of the 14 categories with symmetry Type I, II and III. For the 12 categories with symmetry Type IV, we only evaluate their category classification accuracy, shown as average “Place” accuracy with a * in the last two tables.*

Category	HOG-S	HOG-L	Tiny	HOG	COM	Final	Chance
beach	24.4%	27.8%	33.9%	54.9%	69.8%	84.4%	8.3%
church	67.2%	76.1%	34.2%	88.2%	90.7%	95.5%	8.3%
hotel room	62.5%	60.0%	38.1%	82.7%	88.5%	88.5%	8.3%
street	57.4%	46.8%	28.4%	65.8%	66.7%	76.6%	8.3%
subway station	46.3%	48.8%	30.7%	67.7%	75.4%	75.6%	8.3%
theater	58.6%	69.0%	23.6%	83.3%	93.1%	100.0%	8.3%
train interior	55.0%	57.5%	39.8%	62.9%	55.8%	53.3%	8.3%
wharf	38.5%	59.0%	19.0%	45.3%	66.2%	82.1%	8.3%
corridor	43.9%	40.1%	19.6%	43.9%	51.0%	48.0%	8.3%
living room	30.6%	32.3%	12.9%	53.0%	53.5%	53.5%	8.3%
coast	17.9%	41.0%	22.6%	26.9%	31.0%	51.3%	8.3%
lawn	18.8%	14.6%	14.6%	26.4%	25.7%	41.7%	8.3%
plaza courtyard	24.4%	21.1%	13.2%	40.1%	41.0%	44.1%	8.3%
shop	40.0%	36.0%	32.3%	48.5%	63.0%	76.0%	8.3%
average	41.8%	45.0%	25.9%	56.4%	62.2%	69.3%	8.3%

Table 1. Extension of Table 1 in the paper: Accuracy of automatic panorama alignment for each of the 14 categories with Symmetry Types I, II and III. HOG-S and HOG-L are two baseline algorithms for comparison (Section 8.2 in the paper). Tiny, HOG and COM are the results that do NOT make use of symmetry and view sharing (Section 4 in the paper), using various features presented in Section 7 in the paper. Final is our complete model with symmetry and view sharing (Section 5 in the paper).

Category	HOG-S	HOG-L	Tiny	HOG	COM	Final	Chance
beach	72.3°	38.0°	43.1°	24.7°	18.1°	7.7°	90.0°
church	54.6°	40.3°	87.8°	17.8°	12.8°	3.6°	90.0°
hotel room	27.8°	36.8°	53.6°	15.6°	6.5°	6.5°	90.0°
street	60.0°	75.6°	80.5°	51.9°	54.7°	40.5°	90.0°
subway station	63.7°	79.8°	82.0°	56.2°	42.1°	43.9°	90.0°
theater	59.0°	36.2°	84.8°	16.7°	9.9°	0.0°	90.0°
train interior	76.5°	76.5°	84.8°	65.2°	78.4°	84.0°	90.0°
wharf	58.5°	37.7°	65.2°	45.8°	29.1°	12.3°	90.0°
corridor	79.1°	90.8°	84.5°	87.6°	76.0°	86.9°	90.0°
living room	69.7°	82.3°	83.5°	48.4°	46.3°	46.3°	90.0°
coast	71.5°	47.7°	48.3°	55.9°	47.1°	22.3°	90.0°
lawn	70.0°	80.6°	88.4°	61.3°	49.3°	45.6°	90.0°
plaza courtyard	83.4°	84.9°	84.1°	65.6°	64.3°	61.9°	90.0°
shop	72.0°	68.4°	57.5°	59.8°	44.4°	27.6°	90.0°
average	65.6°	62.5°	73.4°	48.0°	41.4°	34.9°	90.0°

Table 2. Extension of Table 1 in the paper: Angle deviation of automatic panorama alignment for each of the 14 categories with Symmetry Types I, II and III. HOG-S and HOG-L are two baseline algorithms for comparison (Section 8.2 in the paper). Tiny, HOG and COM are the results that do NOT make use of symmetry and view sharing (Section 4 in the paper), using various features presented in Section 7 in the paper. Final is our complete model with symmetry and view sharing (Section 5 in the paper).

Test Set	Manual Alignment				Automatic Alignment							
	Accuracy 1		Accuracy 2		Accuracy			Angle Deviation				
	Place	Both	Place	Both	Place	View	Both	I	II	III	IV	
beach	74.4%	42.9%	76.2%	40.1%	76.2%	53.1%	42.6%	29.0°	24.6°	85.1°	90.0°	
church	81.0%	72.6%	71.4%	61.9%	71.4%	85.8%	62.3%	25.5°	20.0°	87.5°	90.0°	
hotel room	63.7%	44.4%	46.9%	31.0%	46.9%	64.2%	29.8%	20.9°	18.9°	85.6°	90.0°	
street	45.1%	29.9%	38.8%	26.3%	38.8%	59.2%	25.3%	65.6°	64.1°	83.9°	90.0°	
subway station	50.6%	41.1%	40.2%	30.5%	40.2%	77.2%	35.6%	40.7°	33.9°	82.4°	90.0°	
theater	41.4%	34.8%	36.0%	27.1%	36.0%	74.4%	27.7%	30.9°	23.2°	83.8°	90.0°	
train interior	56.9%	27.5%	56.0%	29.2%	56.0%	50.8%	28.7%	80.4°	76.9°	85.2°	90.0°	
wharf	28.2%	13.5%	33.3%	12.0%	33.3%	47.6%	13.7%	50.4°	46.5°	82.8°	90.0°	
corridor	30.8%	14.0%	23.0%	8.8%	23.0%	40.4%	8.8%	84.8°	85.8°	85.2°	90.0°	
living room	25.7%	13.4%	28.1%	12.3%	28.1%	34.8%	12.2%	66.9°	64.4°	83.8°	90.0°	
coast	23.2%	6.1%	31.4%	9.0%	31.4%	28.3%	11.2%	51.4°	43.2°	89.0°	90.0°	
lawn	28.0%	7.1%	35.1%	9.4%	35.1%	20.6%	8.7%	73.2°	65.1°	88.7°	90.0°	
plaza courtyard	70.8%	21.7%	73.0%	22.4%	73.0%	28.6%	20.5%	75.2°	76.1°	90.3°	90.0°	
shop	27.5%	13.8%	32.8%	13.5%	32.8%	37.7%	12.2%	74.9°	73.5°	85.5°	90.0°	
average	48.4%	27.3%	51.9%	23.8%	51.9%	50.2%	24.2%	55.0°	51.2°	85.6°	90.0°	
chance	3.8%	0.3%	3.8%	0.3%	3.8%	8.3%	0.3%	90.0°	90.0°	90.0°	90.0°	

Table 3. Extension of Table 2 in the paper: Testing accuracy and average viewpoint deviation for the panorama generated test set. We compare the performance of our automatic alignment algorithm with manual alignments using two algorithms (Section 8.3 in the paper): an 1-step algorithm (Accuracy 1) and a 2-step algorithm (Accuracy 2). For each algorithm, “Place” is the accuracy of place classification, “View” is the accuracy of viewpoint prediction, and “Both” is the accuracy requiring both correct place category prediction and correct viewpoint prediction, and is the final result. The “Place” average accuracy (with a *) is an average accuracy over all 26 categories with symmetry Type I, II, III and Type IV, because place classification is meaningful for all categories. The “View” and “Both” average accuracy are only over the 14 categories with symmetry Type I, II and III.

Test Set	Manual Alignment				Automatic Alignment						
	Accuracy 1		Accuracy 2		Accuracy			Angle Deviation			
	Place	Both	Place	Both	Place	View	Both	I	II	III	IV
beach	51.6%	29.0%	60.0%	29.4%	60.0%	53.5%	32.4%	25.0°	23.4°	49.9°	90.0°
church	31.5%	25.6%	25.2%	18.1%	25.2%	69.7%	19.7%	23.3°	21.8°	19.5°	90.0°
hotel room	5.6%	3.6%	9.7%	5.6%	9.7%	43.6%	4.1%	47.5°	57.8°	67.1°	90.0°
street	8.6%	4.5%	11.5%	4.2%	11.5%	66.9%	6.0%	16.5°	16.2°	19.7°	90.0°
subway station	29.5%	24.7%	19.9%	14.0%	19.9%	76.9%	16.4%	19.7°	13.8°	32.0°	90.0°
theater	14.8%	10.6%	27.4%	14.8%	27.4%	52.8%	16.1%	58.3°	48.7°	74.5°	90.0°
train interior	13.8%	11.8%	7.6%	5.0%	7.6%	72.6%	5.3%	17.3°	20.0°	20.0°	90.0°
wharf	13.8%	8.0%	20.7%	9.6%	20.7%	41.5%	11.2%	36.9°	37.0°	37.8°	90.0°
corridor	49.6%	44.4%	33.8%	31.6%	33.8%	69.2%	25.6%	25.7°	35.9°	25.9°	90.0°
living room	7.0%	4.8%	7.1%	4.7%	7.1%	60.1%	4.6%	21.4°	31.7°	38.3°	90.0°
coast	24.1%	12.5%	33.9%	17.0%	33.9%	63.4%	22.3%	27.1°	19.0°	33.5°	90.0°
lawn	22.7%	9.1%	29.5%	6.8%	29.5%	25.0%	9.1%	30.0°	31.4°	49.1°	90.0°
plaza courtyard	23.4%	7.8%	25.9%	10.1%	25.9%	40.5%	10.2%	30.3°	28.3°	31.4°	90.0°
shop	15.2%	6.8%	24.5%	11.0%	24.5%	43.6%	11.3%	29.6°	30.4°	31.5°	90.0°
average	22.2%	14.5%	24.1%	13.0%	24.1%	55.7%	13.9%	29.2°	29.7°	37.9°	90.0°
chance	3.8%	0.3%	3.8%	0.3%	3.8%	8.3%	0.3%	90.0°	90.0°	90.0°	90.0°

Table 4. Extension of Table 2 in the paper: Testing accuracy and average viewpoint deviation for SUN test set. We compare the performance of our automatic alignment algorithm with manual alignments using two algorithms (Section 8.3 in the paper): an 1-step algorithm (Accuracy 1) and a 2-step algorithm (Accuracy 2). For each algorithm, “Place” is the accuracy of place classification, “View” is the accuracy of viewpoint prediction, and “Both” is the accuracy requiring both correct place category prediction and correct viewpoint prediction, and is the final result. The “Place” average accuracy (with a *) is an average accuracy over all 26 categories with symmetry Type I, II, III and Type IV, because place classification is meaningful for all categories. The “View” and “Both” average accuracy are only over the 14 categories with symmetry Type I, II and III.