Supplementary Material for Paper 230 ACCV 2016

Here we provide additional material that complements our submitted paper.

1 On Binary Classification of Facade versus Non-facade

Given the feature set F used for facade classification (Section 2.2 in the submitted ACCV paper 230) where $F = [P, A, E, S, T, O]$ (*P: Probability, A: Area, E: Entropy, S: A-Score: T: Texel, O: Occupancy*), we carried out an exhaustive search for the best feature subset from 2^6 possible subsets. We have found that the most discriminative feature set for this binary classification problem is when $|F| = 6$. In Figure 1 we display the data distribution in all the possible 3-feature subspaces, a total of 20 (6 choose 3), and Figure 2 shows the data projection from the 6-dimensional space to 2D using the *tSNE* Method [2].

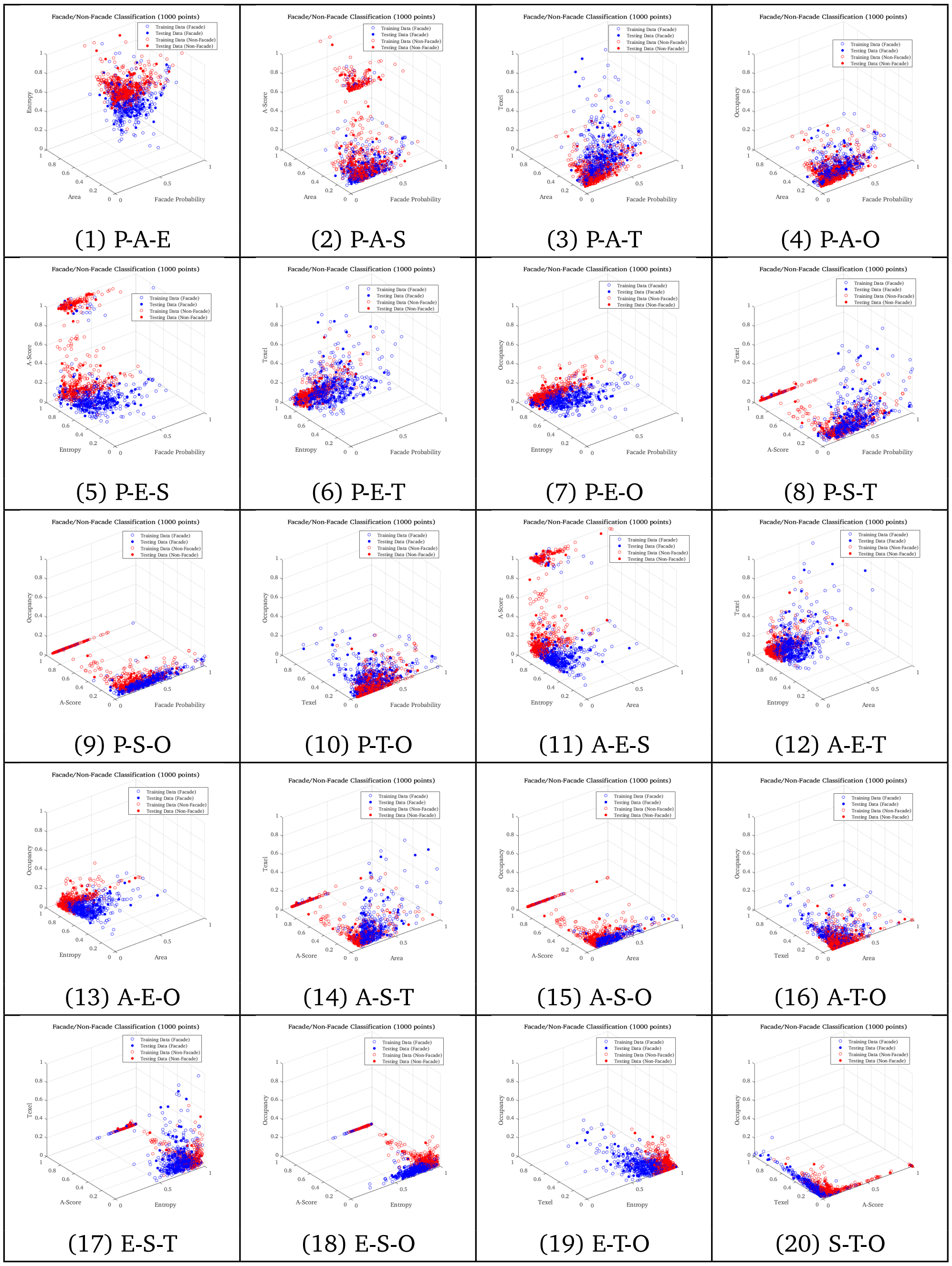

Figure 1: When the total number of samples is 1,000 ($N = 500$ for each class), we visualize the data distribution in all possible 3-feature subspaces.

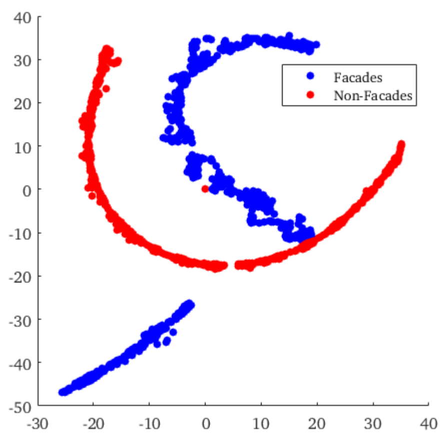

Figure 2: Projecting the data points from its 6-Feature space to 2D using tSNE [2] method.

Finally, we demonstrate the quantitative improvement our facade classification method made over the raw input from [3] using area (A) as threshold for data sets.

	$A \geq 1000$	$A \geq 500$	$A \geq 200$	$A \geq 100$	$A \geq 50$	$A \geq 20$	$A \geq 10$	$A \geq 0$
NYC	0.69	0.70	0.67	0.60	0.53	0.48	0.48	0.48
ROME	0.10	0.20	0.23	0.21	0.20	0.20	0.17	0.17
SF	0.49	0.44	0.34	0.28	0.24	0.24	0.21	0.21
NYC	0.87	0.87	0.84	0.82	0.84	0.84	0.84	0.84
ROME	0.50	0.67	0.81	0.57	0.71	0.83	0.83	0.46
SF	0.90	0.90	0.90	0.89	0.89	0.89	0.86	0.90

Table 1: Improvement facades classification over the raw output of [3]

2 On 3-City Classification (NYC, Rome, SF)

For 3-city classification visualization, we use one-versus-all ECOC framework. [1] (The true label is 0, the false label is -1.)

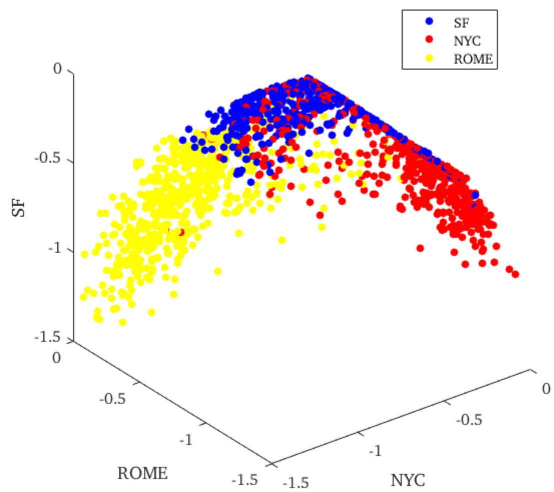

Figure 3: 3 City Classification Visualization

Here we compare 3-City Classification results between (1) the output from our binary classification algorithm. #F denotes the top selected facades (rank by the number of times classified as facade after 50 iterations) when N true plus N false facades are used, versus (2) the ground truth (GT) facades (hand labeled facades) of $2N$ true and false facades respectively. The results show that they are comparable. Using GT input does a better job in Rome and SF.

	$N = 500$ # F: 100	$N = 800$ # F: 200	$N = 1000$ # F: 300	$N = 2000$ # F: 400	$N = 3000$ # F: 500	$N = 4000$ # F: 600
Precision	0.97	0.96	0.93	0.93	0.88	0.85
Recall	0.90	0.90	0.94	0.93	0.91	0.90
Precision	0.90	0.78	0.74	0.88	0.82	0.85
Recall	0.90	0.90	0.98	0.95	0.91	0.89

Table 2: Comparing rates of GT (bottom) for facade detection in NYC

	$N = 500$ # F: 100	$N = 800$ # F: 200	$N = 1000$ # F: 300	$N = 2000$ # F: 400	$N = 3000$ # F: 500	$N = 4000$ # F: 600
Precision	0.90	0.92	0.93	0.95	0.90	0.88
Recall	0.85	0.87	0.83	0.86	0.83	0.80
Precision	0.95	0.91	0.98	0.95	0.97	0.95
Recall	0.95	1.00	0.90	1.00	0.94	0.95

Table 3: Comparing rates of GT (bottom) for facade detection in ROME

	$N = 500$ # F: 100	$N = 800$ # F: 200	$N = 1000$ # F: 300	$N = 2000$ # F: 400	$N = 3000$ # F: 500	$N = 4000$ # F: 600
Precision	0.95	0.96	0.90	0.93	0.85	0.85
Recall	0.86	0.86	0.82	0.85	0.81	0.82
Precision	0.90	0.96	0.96	0.97	0.93	0.91
Recall	0.90	0.83	0.83	0.85	0.86	0.88

Table 4: Comparing rates of GT (bottom) for facade detection in SF

References

[1] Sergio Escalera, Oriol Pujol, and Petia Radeva. Error-correcting ouput codes library. *The Journal of Machine Learning Research*, 11:661–664, 2010.

[2] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(2579-2605):85, 2008.

[3] Jingchen Liu and Yanxi Liu. Local regularity-driven city-scale facade detection from aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3778–3785, 2014.