Leveraging Structure from Motion to Learn Discriminative Codebooks for Scalable Landmark Classification
– Supplementary material –

Alessandro Bergamo
Dartmouth College
Hanover, NH, USA
aleb@cs.dartmouth.edu

Sudipta N. Sinha
Microsoft Research
Redmond, WA, USA
sudipsin@microsoft.com

Lorenzo Torresani
Dartmouth College
Hanover, NH, USA
lorenzo@cs.dartmouth.edu

1. Additional Results

The supplementary material contains additional results involving the following experiments.

- **Comparison of DAISY and SIFT descriptors:** We have compared DAISY and SIFT descriptors in preliminary experiments using the *HKM+BoW* method on LANDMARK-3D, as reported in Section 4.2 of the paper. These results are shown in Figure 1.

- **Comparison of classifiers:** In Sec. 4.2 of the paper, we also reported that the *Crammer and Singer* SVM formulation, used by *Li et al.* [ICCV, 2009], produced comparable results to one-vs-all SVMs. The results from the preliminary experiments comparing the two classifiers using the *HKM+BoW* method on LANDMARK-3D are shown in Figure 2.

- **Comparison of Information Gain estimators:** In Figure 3, we show classification accuracies using our random forest *RFT* codebooks, where we switched the standard information gain estimator with the one proposed by *Nowozin et al.* [ICML 2012].

- *RFT* **codebooks and forest parameters:** In Table 1, we report the classification accuracies obtained using *RFT* codebooks with *BoW* encoding on LANDMARK-3D, where the tree-count and tree height are set to different values to obtain codebooks of different sizes.

- **SfM reconstruction statistics**: The number of downloaded images, the number of images selected after outlier removal using SfM, and relevant statistics of reconstructions obtained from SfM pre-processing on the training image sets, are shown in Table 2 for a few selected landmarks.

- **Results on popular landmarks**: For ten landmarks discussed in the paper, we report the top five landmarks they are most frequently confused with, in Table 3.
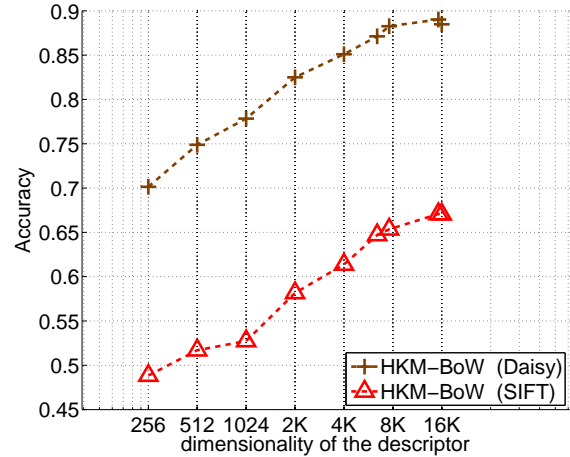

Figure 1: **DAISY-vs-SIFT**: The classification accuracy on LANDMARK-3D using the *HKM+BoW* method and different codebook sizes: DAISY considerably outperforms SIFT for all descriptor dimensions.
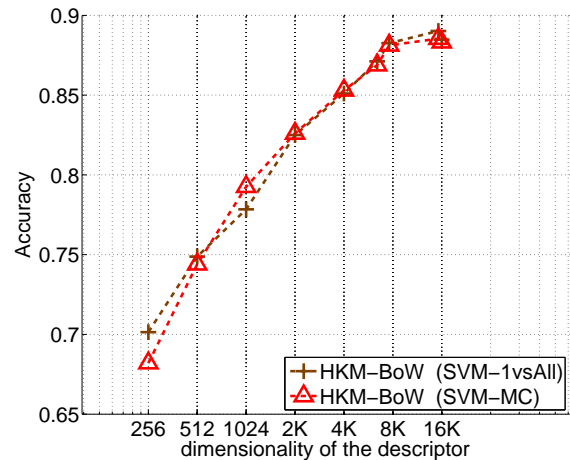

Figure 2: **Type of classifier**: The classification accuracies on LANDMARK-3D with *HKM+BoW* and different SVM formulations was comparable for different codebook sizes.
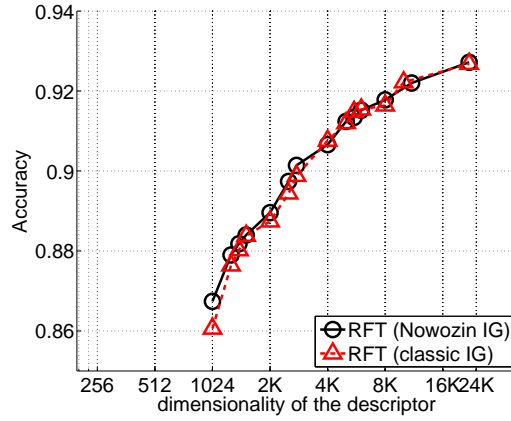
Figure 3: **Comparing Information Gain estimators**: The accuracies obtained using *RFT+BoW* with *RFT* codebooks trained using the standard information gain estimator as well as the improved estimator proposed by *Nowozin et al.* [ICML 2012].

T ╱ h	8	9	10	11	12
8	0.86 (1024)	0.88 (2048)	0.89 (4096)	0.90 (8192)	0.91 (16384)
10	0.86 (1280)	0.89 (2560)	0.90 (5120)	0.91 (10240)	0.91 (20480)
11	0.87 (1408)	0.89 (2816)	0.91 (5632)	0.91 (11264)	0.92 (22528)
12	0.88 (1536)	0.89 (3072)	0.91 (6144)	0.91 (12288)	0.92 (24576)

Table 1: **Tree parameters**: The classification accuracies obtained using various settings for parameters t (tree-count), and h (tree height) for *RFT* codebooks using *BoW* encoding on the LANDMARK-3D dataset. All entries in a column have the same tree height whereas each row corresponds to a fixed tree-count. The final descriptor dimensionality is indicated in brackets.

Ten Popular Landmarks	nImgsRaw	nImgsSel	nPts	nCams	nComps
Trevi Fountain, Rome	3201	2493	722232	2374	3
Tyn Church, Prague	3307	2199	352978	2078	9
Piazza Navona, Rome	3013	1191	194713	1046	8
Chichen Itza, Yucatan	3434	1402	271026	1286	20
Arc de Triomphe, Paris	3255	1759	329112	1622	14
Taj Mahal, Agra	3338	1701	321878	1562	14
Times Square, New York	3391	367	16915	237	14
Pike Place, Seattle	3432	707	30898	595	13
Trafalgar Square, London	3926	589	34533	452	19
Pantheon, Rome	3293	1885	446712	1766	6
Overall	nImgsRaw	nImgsSel	nPts	nCams	nComps
Average	2844	991	161194	864	14
Minimum	251	242	2119	119	1
Maximum	4063	2593	1552838	2467	32

Table 2: **LANDMARK-620 SfM reconstructions**: This table reports SfM reconstruction statistics for ten popular landmarks (ones discussed in the main paper) as well as statistics computed over all the 620 landmarks. The columns denote – (*nImgsRaw*) the number of images originally downloaded; (*nImgsSel*) the number of images selected after using SfM for outlier removal; (*nPts, nCams, nComps*) the number of 3D points, cameras and connected components in the SfM reconstructions obtained after running SfM on the training image sets.

Ten popular landmarks	Five landmarks most frequently confused with				
Trevi Fountain (77%)	Castelsantangelo, Rome [3]	Fuente de Cibeles, Madrid [2]	Barberini Fountain, Rome [2]	Rialto Canal, Venice [1]	Spanish steps, Piazza di Spagna, Rome [1]
Tyn Church (82%)	Belfry of Bruges [2]	Xian Bell Tower [1]	Galata Tower Istanbul [1]	Sagradafamilia, Barcelona [1]	Space Needle, Seattle [1]
Piazza Navona (55%)	Spanish steps, Piazza di Spagna, Rome [3]	Barberini Fountain, Rome [3]	Salzburg Cathedral [2]	Uspenski Cathedral, Helsinki [1]	Grossmunster, Zurich [1]
Chichen Itza (71%)	Chariot Temple, Hampi [2]	Lime Kiln Lighthouse, San Juan Island [2]	Opera House, Sydney [1]	Disneyworld castle, Orlando [1]	Powis Castle, Wales [1]
Arc de Triomphe (75%)	Cloud Gate Bean, Chicago [1]	Coit Tower, San Francisco [1]	Seville Cathedral [1]	Palacio de Comunicaciones, Madrid [1]	Hagia Sophia Mosque, Ayasofya [1]
Taj Mahal (89%)	Tower Bridge, London [1]	Coit Tower, San Francisco [1]	Public Library, Boston [1]	Ellis Island, New York [1]	Mormon Temple, Salt Lake City [1]
Times Square (47%)	Luxor Hotel Casino, Las Vegas [4]	Coit Tower, San Francisco [1]	Spanish steps, Piazza di spagna, Rome [1]	New York casino, Las Vegas [1]	Dam Square Amsterdam [1]
Pike Place, Seattle (88%)	Luna Park, Sydney [1]	New York casino, Las Vegas [1]	Moulin Rouge, Paris [1]	Luxor Hotel casino, Las Vegas [1]	Strasbourg Cathedral, Alsace [1]
Trafalgar Square (51%)	Piazza San Carlo, Torino [3]	City Hall, Belfast [2]	Marble Arch Park, London [2]	Pike Place, Seattle [1]	Place de la Concorde, Paris [1]
Pantheon (83%)	Speyer Kaiserdom [2]	Campanile UC Berkeley [1]	Vatican Globe [1]	Transamerica Pyramid, San Francisco [1]	Philadelphia Museum of Art [1]

Table 3: **Ten popular landmarks and top five misclassifications:** For each of ten popular landmarks, the top five incorrect landmark predictions are shown. The leftmost columns indicates the name of the landmark and the classification accuracy within brackets. The other five columns show some of the incorrect predictions and the numbers in square brackets refer to the number of misclassifications. These results on LANDMARK-620 were obtained using the proposed *RFT+VLADSa* method at dimensionality = 8192. Recall that the test set of 62K images contains 100 images of each landmark.