RingIt: Supplementary Material

Dataset			Refined			Not Refined		
Name	#img	#ppl	SURF	SIFT	IDSC	SURF	SIFT	IDSC
HEADSTAND	11	11	0	0	-	1	0	-
THIGH-STAND	11	11	0	4	-	0	4	-
BOTTOM-UP	14	3	5	16	-	5	16	-
BIKE&STATUE	21	1	0	1	-	39	4	-
FRONT HANG	21	21	1	1	-	10	2	-
SIT-ON-FEET	21	21	4	8	-	6	7	-
CANDLESTICK	21	21	5	5	-	7	4	-
STAND-ON-FEET	21	21	1	1	-	1	1	-
BACK BALANCE	31	21	3	3	-	4	5	-
ARM BALANCE	31	21	5	7	-	18	17	-
FRONT SWAN	35	21	0	3	-	7	5	-
MOTORCYCLE	71	1	0	4	-	15	17	-
ARABESQUE	10	10	-	-	5	-	-	4
BALLERINA	15	1	-	-	6	-	-	5
BACKBEND	15	1	-	-	8	-	-	37
FLAMINGO	15	1	-	-	5	-	-	8
BRIDGE	17	2	-	-	1	-	-	1
DEVANT	20	10	-	-	5	-	-	9

Table 1: **Descriptor Evaluation**. The entries display the swapping distance to the ground truth. The refined distances are with NRDC (for the feature-based approach) or L1 distances (for the contour-based approach). As SURF is generally more successful than SIFT, it is our default basic descriptor.

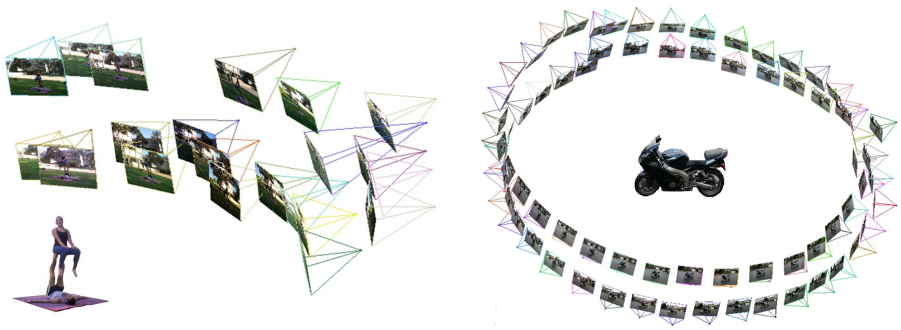

Figure 1: **SfM Reconstruction**: A visualization of the reconstruction of our SIT-ON-FEET and MOTORCYCLE sets. These are the only two sets which were fully recovered by a SfM technique. The success is attributed to the density of the camera distribution.

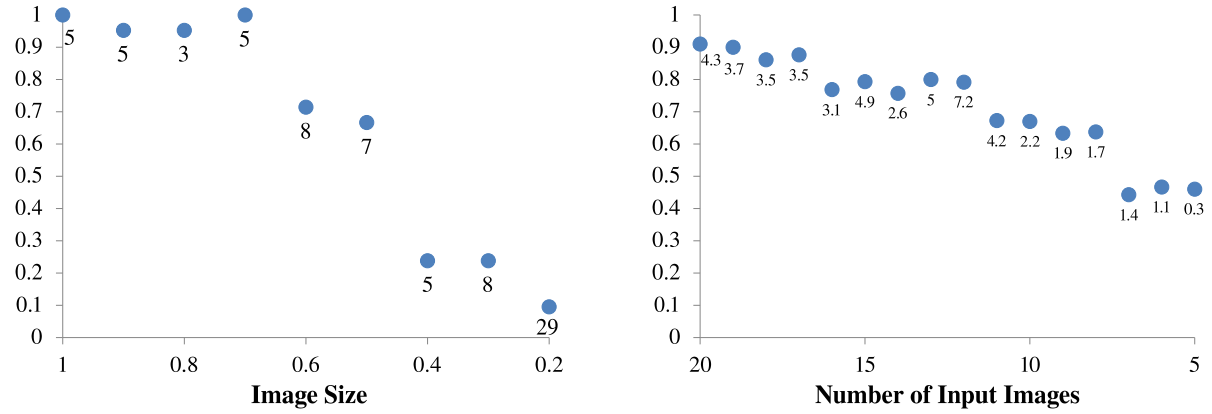

Figure 2: A comparison between *RingIt* and SfM on our densely captured SIT-ON-FEET set, as a function of the image size (on the left) and the number of images in the set (on the right). The blue dots illustrate the percentage of input views that were reconstructed by VisualSFM. Recall that a ring-ordering of the input images via an SfM reconstruction requires that all the input views are reconstructed into one consistent model. The numbers below the blue dots correspond to the number of swaps between the *RingIt* output and the ground-truth permutation. Left: The quality of the input images degrades as a function of the image size. Notice that a significant drop in SfM performance occurs roughly around a resize factor of 0.6. In comparison, our performance remains roughly constant until the images are scaled down by a factor 0.1. Right: Random image subsets, varying from 20 images picked out of the 21 set images to subsets containing only 5 images, are evaluated. The numbers are an average of ten random subsets that were evaluated at each level. As the graph demonstrates, the performance of SfM is clearly affected by the image density. Our method, on the other hand, yields a low number of swaps regardless of the number of input images.