Semantic Shape Editing Using Deformation Handles – Supplemental Material

M.E. Yumer, S. Chaudhuri, J.K. Hodgins, and L.B. Kara, ACM TOG (Proceedings of SIGGRAPH 2015), 34(4)

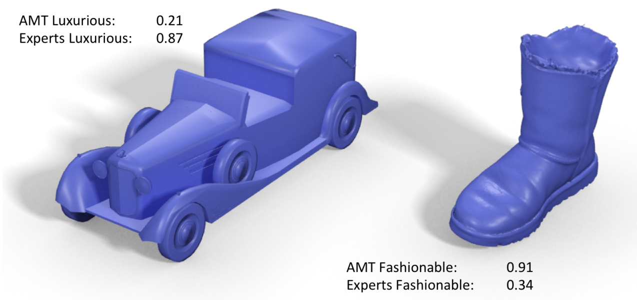

Figure 1: *Most notable disagreements between experts and the general public.*

Pairwise Attribute Data Collected from Amazon Mechanical Turk Users vs. Experts

We built our system and presented the results in our paper using the Amazon Mechanical Turk (AMT) data. We also collected pairwise attribute data from experts for the *Cars* and *Shoes* datasets. To compare the two groups, we train a scoring function (Section 5.2.2 in the paper) individually for each group using leave-one-out approach. We compare the score of the left out model. We compute average difference between the two user data for each attribute a as follows:

$$d_a = \frac{1}{|\mathcal{D}|} \sum_{i \in \mathcal{D}} |{}_{AMT}a_i - {}_{EXP}a_i| \tag{1}$$

Table 1: *Percent average score difference between AMT and EXP, p_a, for all datasets and attributes.*

Cars attribute	p_a	Shoes attribute	p_a
Luxurious	1.69%	Fashionable	1.54%
Sporty	1.01%	Comfy	0.42%
Compact	0.32%	Feminine	0.74%
Muscular	0.71%	Active	0.95%
Modern	0.69%	Durable	0.65%

where d_a is the average difference in attribute score for all shapes i in dataset $\mathcal{D}$. ${}_{AMT}a_i$ and ${}_{EXP}a_i$ is the attribute score computed for shape i with training the system on AMT user data and experts data, respectively.

Because the scoring function (Equation 3 in the paper) is scaled so that the attribute scores are between 0 and 1, percent average difference is: $p_a = 100d_a$. Table 1 summarizes p_a for the *Cars* and *Shoes* datasets.

As seen in Table 1, experts and the general public (AMT) data are significantly consistent. Relatively different perspectives between the experts and the general public is observed in the *luxurious* attribute of the *Cars* dataset, and the attribute in the *Shoes* dataset. Figure 1 illustrates models for which the highest disparity between the experts' and AMTs' ratings is observed. In the shoe set, this corresponds to the UGG[1] boot. While the general public thought that it was very fashionable, the experts disagreed. The most disparity between the AMT and expert users for the car set is the 1930s Cadillac[2]. The experts thought that this was a very luxurious car, while the general public disagreed. There may be several factors leading to this outcome, one being that the experts may have recognized that it was a Cadillac associated the brand with luxury while for the general public this particular model may not have triggered any particular brand association. For both examples, Figure 1 shows the average attribute scores obtained from experts as well as AMT users.

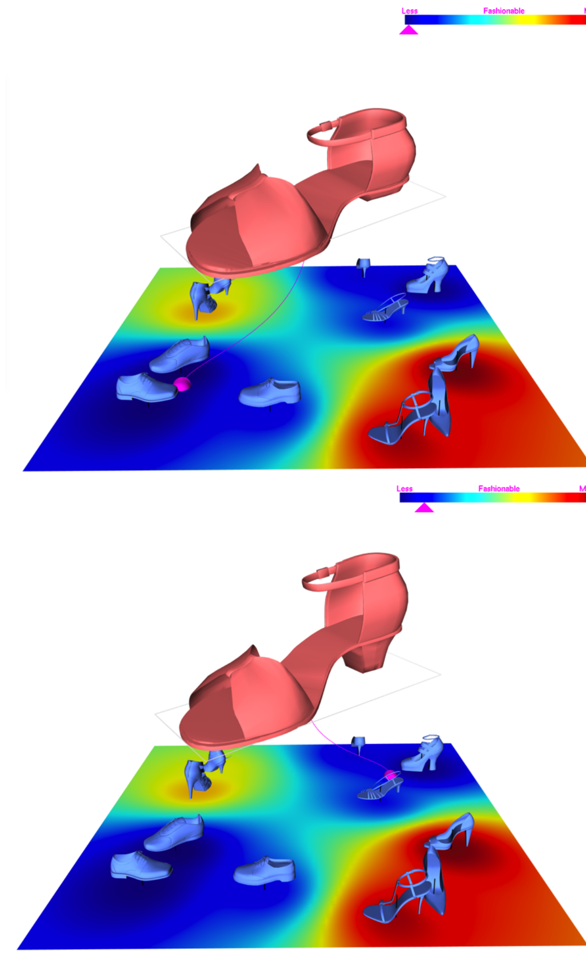

Figure 2: *Low scores for* fashionable *can be achieved by placing the shoe in the proximity of the sports shoes (top), or in the proximity less fashionable mid-heeled shoes (bottom).*

We make our datasets publicly available, as well as the AMT attribute data associated with these datasets[3].

Map exploration example.

Figure 2 illustrates an example where similar attribute levels can be attained at different geometric configurations of the same model.

[1]UGG is a trademark of UGG Australia.

[2]Cadillac is a trademark of General Motors LLC.

[3]http://www.meyumer.com