

# Fits Like a Glove: Rapid and Reliable Hand Shape Personalization

## Supplementary Materials

David Joseph Tan<sup>1,2</sup>    Thomas Cashman<sup>1</sup>    Jonathan Taylor<sup>1</sup>    Andrew Fitzgibbon<sup>1</sup>  
 Daniel Tarlow<sup>1</sup>    Sameh Khamis<sup>1</sup>    Shahram Izadi<sup>1</sup>    Jamie Shotton<sup>1</sup>  
<sup>1</sup>Microsoft Research    <sup>2</sup>Technische Universität München

### 1. Dexter dataset

In the paper, we show the performance of our hand tracker in its template and personalized variants for all frames in the Dexter dataset, apart from a small number at the start of each sequence that Sridhar *et al.* recommend are excluded for evaluation (in the archived results that are currently available from the authors’ web page [2]). Unfortunately, the results in the paper are not directly comparable to those published by Sridhar *et al.* [1], as theirs are filtered to remove any frame with average error greater than 100mm. Because of the way the Dexter dataset is constructed, one effect of this filtering is to exclude any frames where a finger tip is occluded and hence has no ground truth annotation.

In this supplementary material, we attempt to replicate this same error metric, by also excluding any frame with occlusion and any where our tracker exhibits average error greater than 100mm. Table 1 shows that apart from 2 sequences, this does result in a comparable evaluation. Fig. 1 and 2 show the results of our tracker on this metric. The graphs shown in Fig. 2 replicate the per-sequence accuracy results published by Sridhar *et al.* [1].

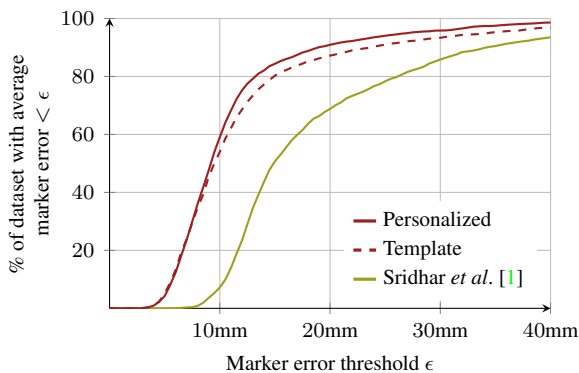


Figure 1: Results for marker localization error on Dexter dataset using a similar metric to Sridhar *et al.* [1]. The results for this dataset have been normalized so that each of the 7 sequences has equal weight.

Sequence	# frames (Sridhar <i>et al.</i> [1])	# frames (ours in Fig. 1/2)	# frames (ours in paper)
adbadd	357	357	398
fingercount	310	310	330
fingerwave	267	461	507
flexex1	477	477	488
pinch	305	305	313
random	398	402	481
tigergrasp	402	402	414
<b>Total</b>	<b>2516</b>	<b>2714</b>	2931

Table 1: Number of frames used to calculate error on the Dexter dataset, in this supplementary material and in the paper. The differences between Sridhar *et al.* and our attempt to replicate the same error metric are marked in red.

### 2. FingerPaint dataset

For the FingerPaint dataset, we include here further images (see Fig. 3) augmenting Figure 10 in the paper with the remaining subjects of the FingerPaint dataset.

### References

- [1] S. Sridhar, F. Mueller, A. Oulasvirta, and C. Theobalt. Fast and robust hand tracking using detection-guided optimization. In *Proc. CVPR*, pages 3213–3221, 2015. 1, 2
- [2] <http://handtracker.mpi-inf.mpg.de/projects/FastHandTracker/>. Accessed 2015-11-13. 1

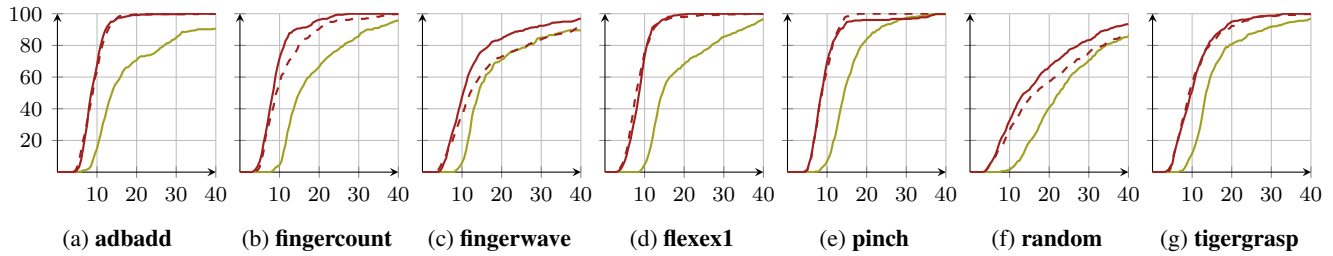


Figure 2: Per-sequence results for marker localization error on Dexter dataset using a similar metric to Sridhar *et al.* [1]. Each plot shows results for our tracker using the template (---) and a personalized model (—) alongside the results published by Sridhar *et al.* (—).

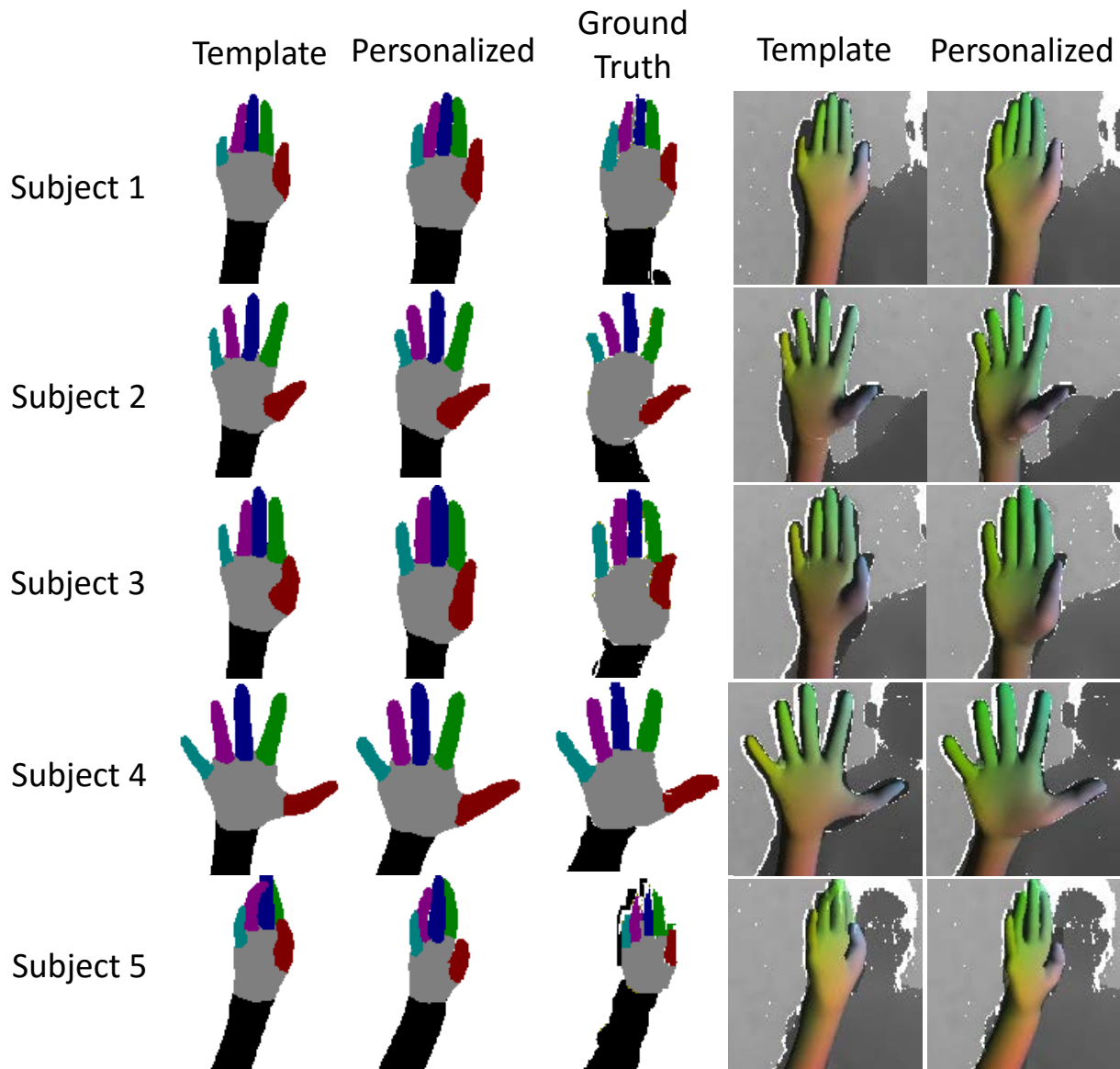


Figure 3: Example frames where each of the five subjects in the FingerPaint dataset exhibit ‘high five’ poses. From left to right the columns show the labels inferred by the template and by the personalized model, the ground truth labels, and the fit of the model to the data for the template and the personalized model.