Home 3D Body Scans from Noisy Image and Range Data

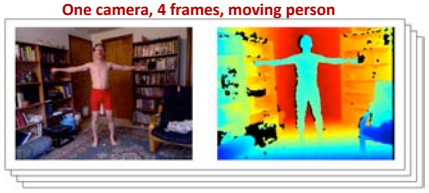
Alexander Weiss* **David Hirshberg†** **Michael J. Black†**

*Department of Computer Science, Brown University, Providence, RI, USA

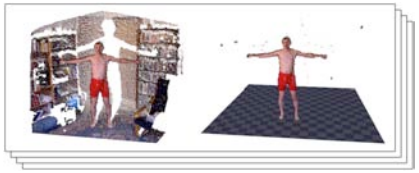
†Perceiving Systems Dept., Max Planck Institute for Intelligent Systems, Tübingen, Germany, http://is.mpg.de/ps

aweiss@cs.brown.edu david.hirshberg@is.mpg.de black@tuebingen.mpg.de

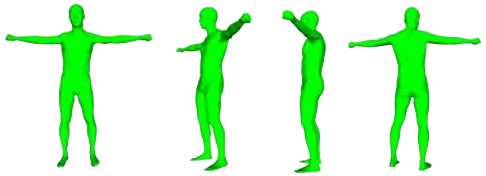
One camera, 4 frames, moving person



Foreground segmentation



SCAPE model fit with consistent shape and varying pose



Model reposed

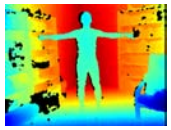
Goal

Accurate body shape from an inexpensive sensor

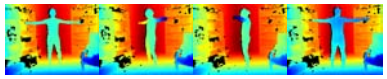
$

$$$$

Problem: Single view, low resolution

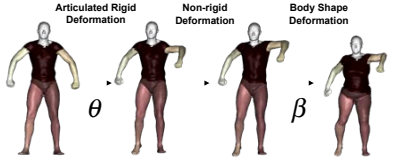


Solution: Let them move! Many frames = Many views



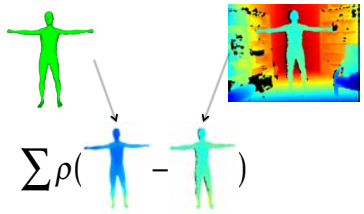
Problem: Humans are non-rigid, shape changes with pose

Solution: Use a SCAPE model to factor shape and pose, fit one shape and many poses



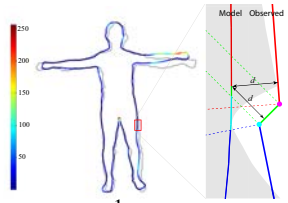
Objective Function

Depth – Depth map difference in the overlapping region



$$E_d(\theta,\beta;U)=\frac{1}{|U|}\sum_{(x,t)\in U}\rho\left(D_t(\theta,\beta)-\tilde{D}_x\right)$$

Silhouette – Correspondence free contour matching



$$E_{uni}(A,B)=\frac{1}{|\partial A|}\int_{\vec{x}\in\partial A}\min_{\vec{y}\in\partial B}\rho\left(\|\vec{x}-\vec{y}\|\right)$$

$$E_s\left(S(\theta,\beta),\tilde{S}\right)=\tfrac{1}{2}E_{uni}\left(S(\theta,\beta),\tilde{S}\right)+\tfrac{1}{2}E_{uni}\left(\tilde{S},S(\theta,\beta)\right)$$

Optimization

Pixel/Triangle correspondences

$$U(\theta,\beta)=\left\{\left(x_1,t_{x_1}(\theta,\beta)\right),\ldots\middle|\,x\in S(\theta,\beta)\cap\tilde{S}\right\}$$

alternating

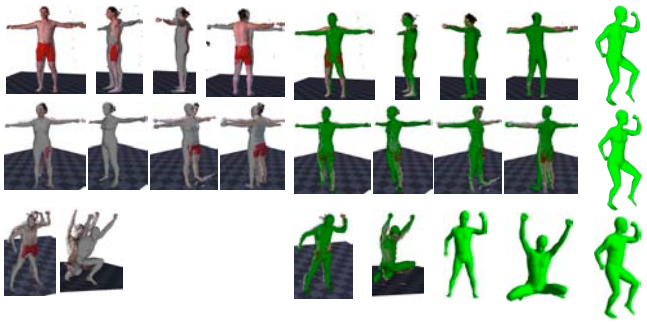
Quasi-Newton local optimization

$$\Theta_i,\beta_i=\arg\max_{\Theta,\beta}E\left(\Theta,\beta;\Theta_{i-1},\beta_{i-1}\right)$$

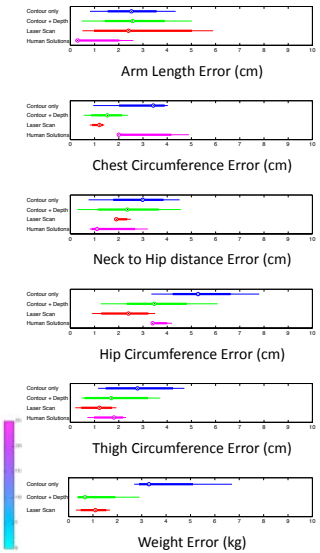
$$E(\Theta,\beta;\Theta_{i-1},\beta_{i-1})=\sum_f E_d\left(\theta_f,\beta;U\left(\theta_{f\,i-1},\beta_{i-1}\right)\right)$$
$$+\lambda_1\sum_f E_s\left(S\left(\theta_f,\beta\right),\tilde{S}_f\right)+\lambda_2\sum_f E_{pose}\left(\theta_f\right)$$

Results

Initialization **Result** **Reposed**

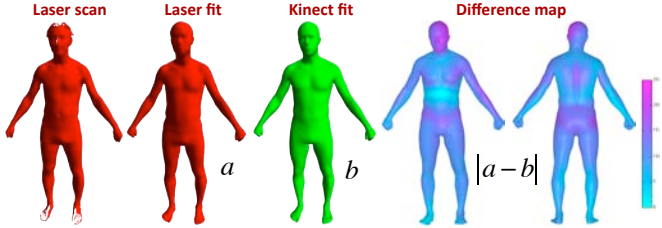


Accuracy



Comparison to Laser Scan

Laser scan **Laser fit** **Kinect fit** **Difference map**



Contributions

1) A system for at home body scanning.
2) Combination of multiple low-resolution, noisy, monocular views to estimate a consistent 3D body shape with varying pose.
3) Correspondence free, bidirectional, differentiable method for matching to silhouettes.
4) Simple method to predict 3D body measurements from SCAPE model.
5) Comparison with commercial state-of-the-art solution for scanning and measuring bodies.

Limitations

1) Requires tight fitting clothing.
2) Not fast enough to be interactive (65min per fit).

References

[1] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. SCAPE: Shape completion and animation of people. *ACM Trans. Graph.*, 24(3):408–416, 2005.
[2] A. O. Balan and M. J. Black. The naked truth: Estimating body shape under clothing. *ECCV*, pp. 15–29, 2008.
[3] M. de La Gorce, N. Paragios, and D. Fleet. Model-based hand tracking with texture, shading and self-occlusions. *CVPR*, pp. 1–8, 2008.
[4] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. *CVPR*, 2011.

Acknowledgements

This work was supported in part by NIH EUREKA award 1R01NS066311-01 and NSF award IIS– 0812364. We thank Loretta Reiss for her measurement expertise and Lisa Wang for mathematical discussions.

ICCV 2011 – Barcelona, Spain Nov. 8-11 2011