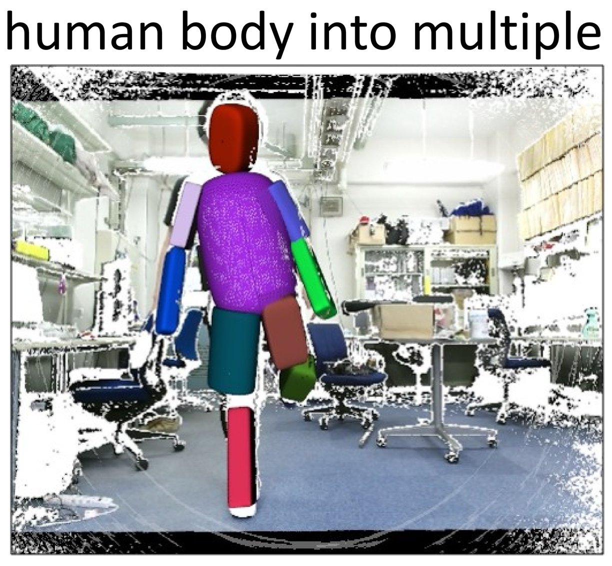
Volumetric Representation of Human Body Parts Using Superquadrics

Ryo Hachiuma, Hideo Saito
Keio University
{ryo-hachiuma, hs}@keio.jp

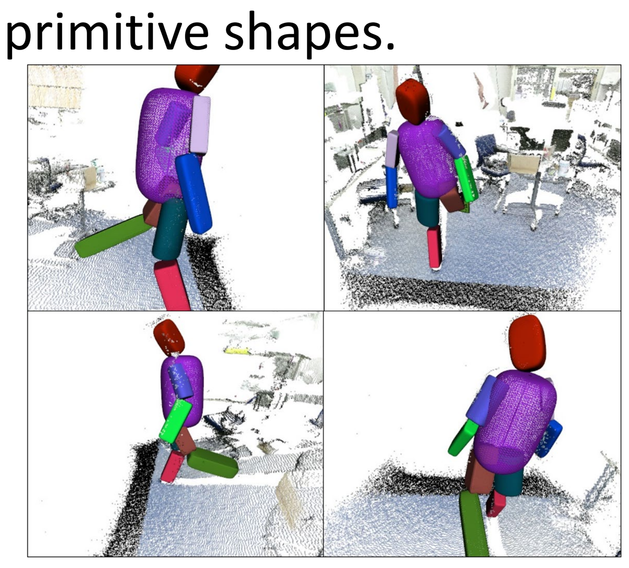
Abstract

The idea of automatically sensing the 3D human body has been of interest in computer vision. The 3D human body can be represented in many ways, such as meshes, parametric models and joint skeletons.
We represent the human body into multiple 3D primitive shapes.

RGB-D Image



Primitive shaped representation



Rendered from different viewpoints

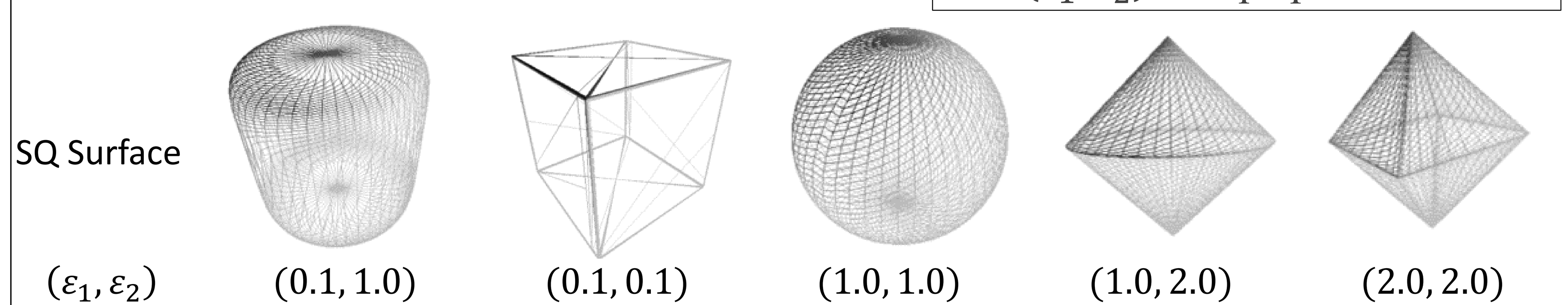
Future application: 1. Outfit size estimation, 2. Free viewpoint image generation, 3. 3D virtual avatar creation.

Superquadrics

Superquadrics [1] can represent various types of primitive shapes with three scale parameters and two shape parameters.

$$f(x_q, y_q, z_q) = \left\{ \left(\frac{x_q}{s_1}\right)^{\frac{2}{\varepsilon_2}} + \left(\frac{y_q}{s_2}\right)^{\frac{2}{\varepsilon_2}} \right\}^{\frac{\varepsilon_2}{\varepsilon_1}} + \left(\frac{z_q}{s_3}\right)^{\frac{2}{\varepsilon_1}} = 1$$

$\boldsymbol{s} = (s_1, s_2, s_3)$: Scale parameters
$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$: Shape parameters

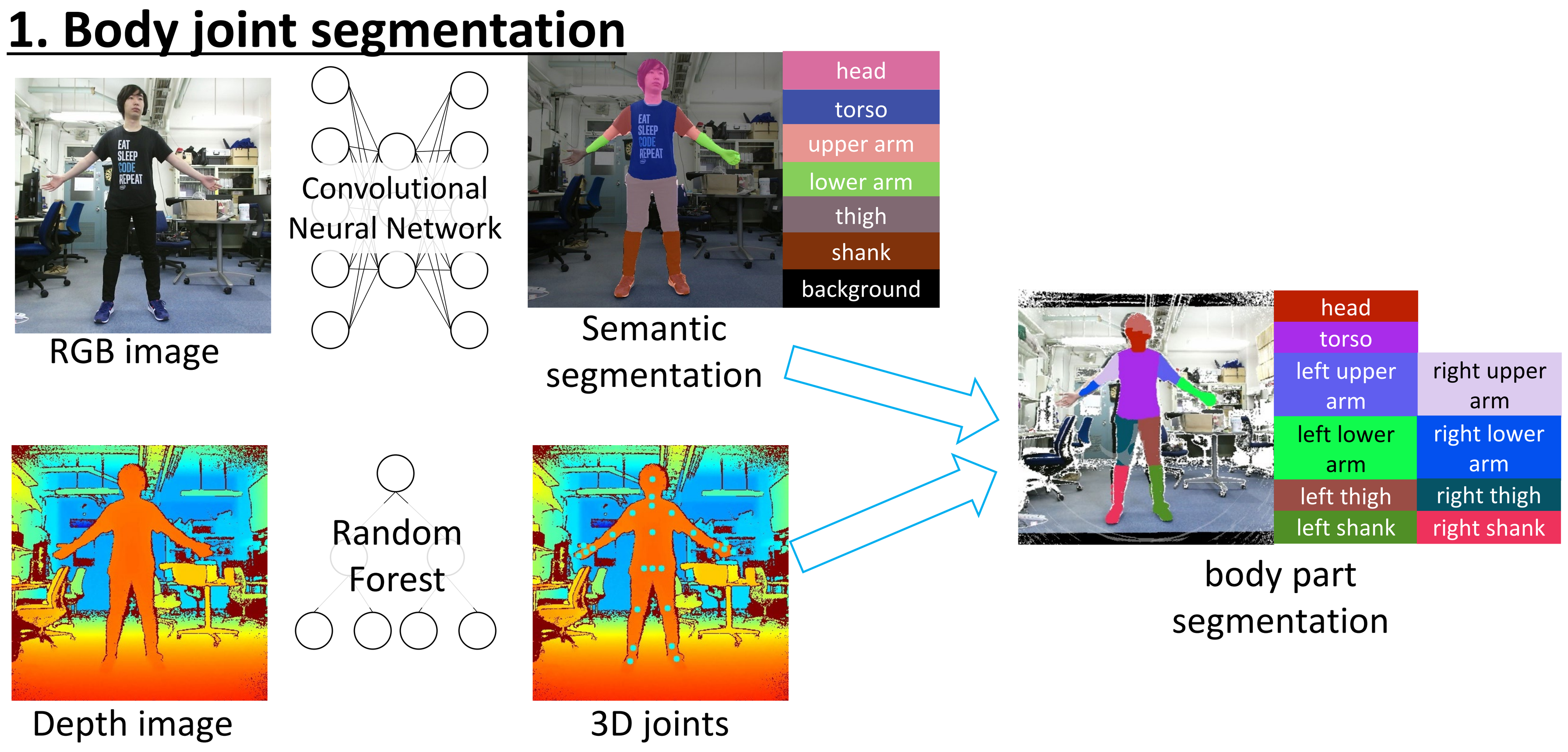


Superquadrics is employed to approximate shape for object shape understanding [2], object grasping [3], collision detection [4].

Method

To represent each joint by superquadrics, the point cloud of each joint must be extracted. Therefore, we first semantically segment human body into joint parts. Next, we estimate superquadric parameters of joints.

1. Body joint segmentation



We employed Light-weight RefineNet [5] trained on PASCAL Person-Part dataset [6] . To segment left/right limbs, we extract 3D joints by the method proposed by Shotton et.al [7].
For each pixel labeled as a limb, we calculate 3D Euclidean distance between 3D joint and each pixel. If the closest joint is left limb, the pixel is labeled as left part, and vice versa.

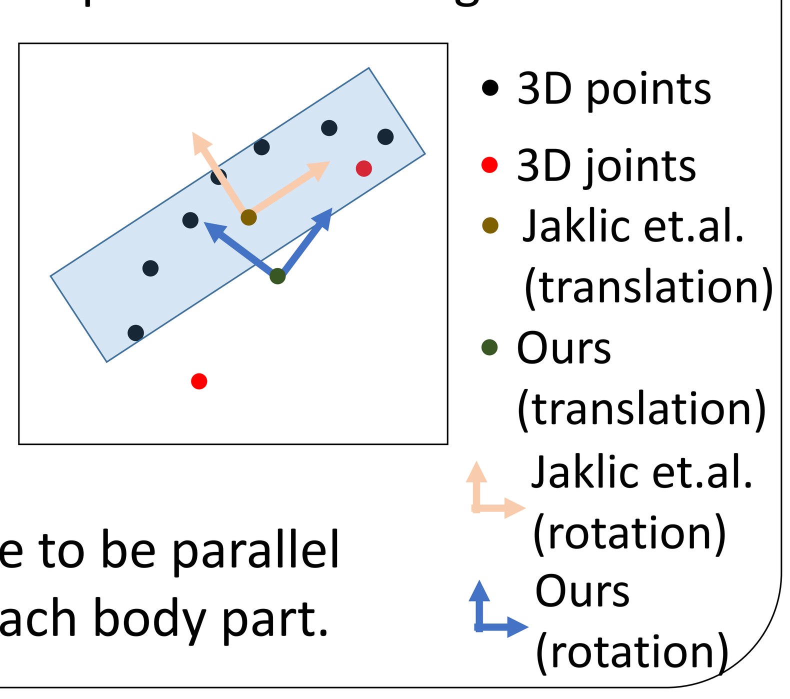
2. Superquadric estimation

Superquadric and its pose parameters $\Lambda(\varepsilon_1, \varepsilon_2, s_1, s_2, s_3, t_1, t_2, t_3, \theta_1, \theta_2, \theta_3)$ are estimated from object point cloud $\{(x_k, y_k, z_k) | 0 \leq k \leq M\}$ by Levenberg-Marquardt algorithm.

$$\min_{\Lambda} \sum_{k=0}^{M} [\sqrt{s_1 s_2 s_3}(f(x_k, y_k, z_k; \Lambda) - 1)]^2$$

As the minimization function is not a convex function, the initial parameters determine which local minimum the minimization converges to.

We propose a novel approach to estimate initial parameters using 3D skeleton joints to find optimal parameter.



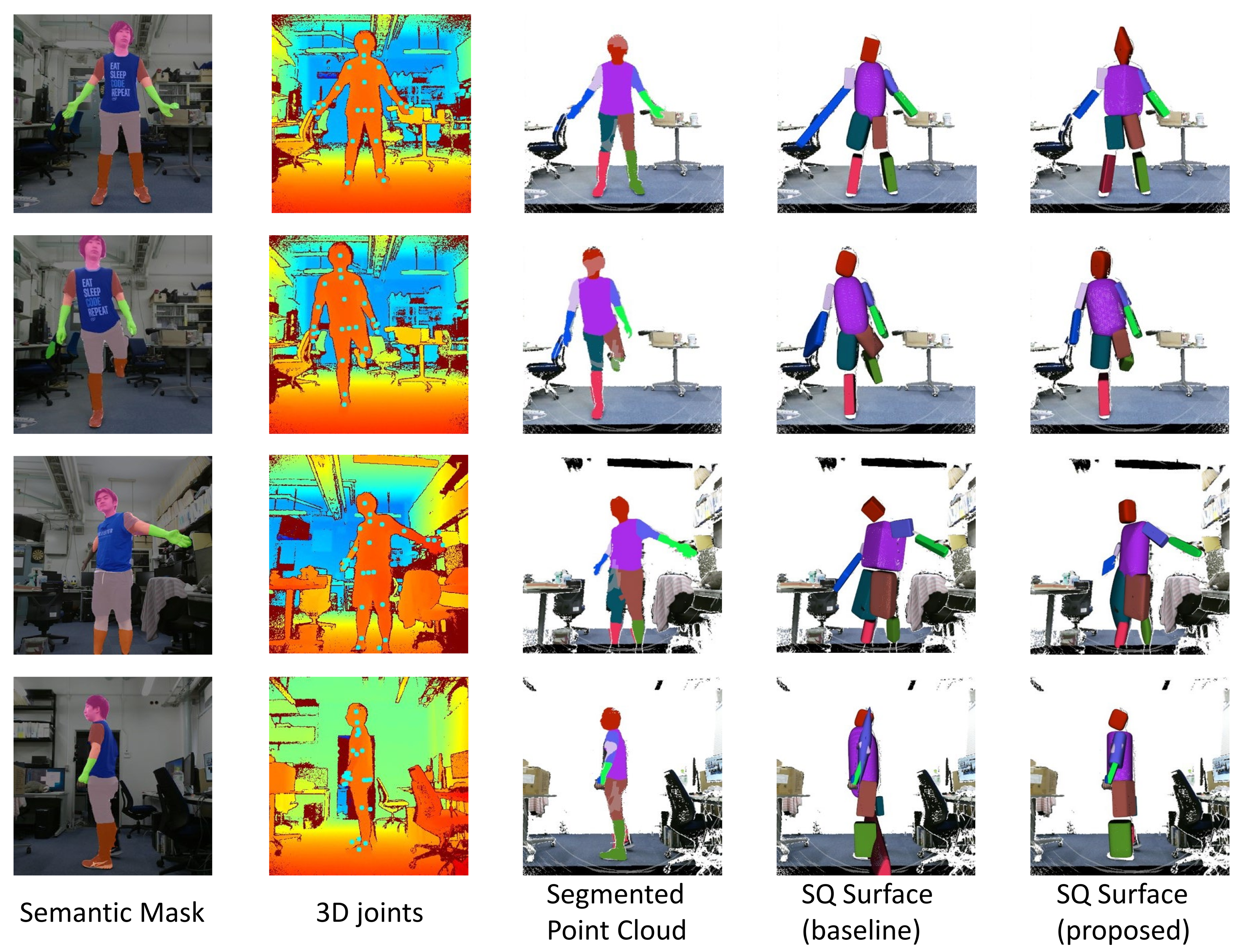
Translation:
Jaklic et.al. [8]: the 3D centroid of 3D points.
Ours: the 3D centroid of 3D joints.

Rotation:
Jaklic et.al. [8]: Eigen vectors of 3D points.
Ours: aligns the z-axis of superquadric surface to be parallel to the vector of two connected 3D joints in each body part.

Result

We record three sequences using Kinect v.2. In each sequence, the person is standing in front of the camera.
Baseline Method: We compared the estimation results with the previous work [8]. This previous method has been widely used for estimating the initial parameter estimation.

1. Qualitative results



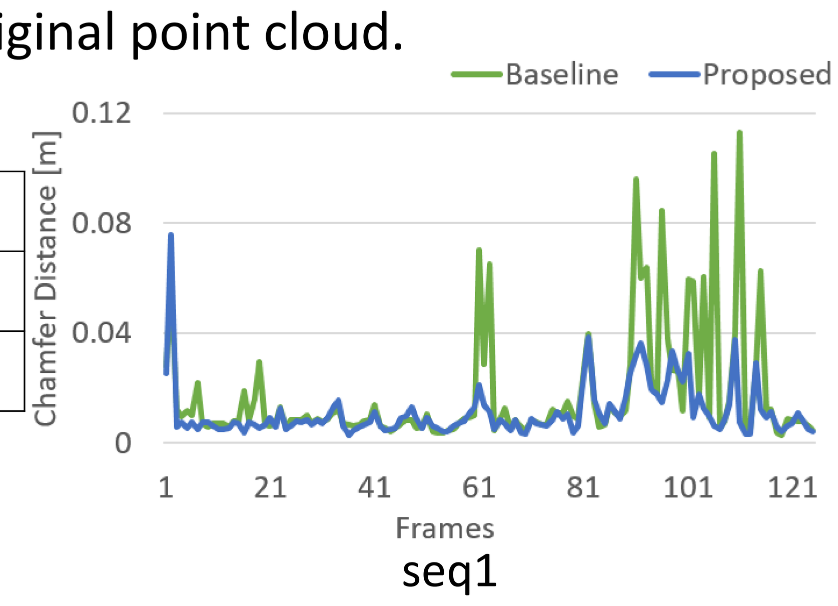
Our proposed method successfully estimated multiple superquadric parameters which approximate the point cloud of each body part.

2. Quantitative results

We employed Chamfer distance as an evaluation metric if the estimated superquadric surfaces represented the original point cloud.

Chamfer distance↓ [m]

	seq1	seq2	seq3	Average
Baseline	1.735	1.671	1.535	1.647
proposed	**1.105**	**1.533**	**1.349**	**1.329**



seq1

We can verify that our novel initial parameter estimation method found more optimal parameters than the previous method [8].

References

[1] Barr, A.H.: Superquadrics and angle-preserving transformations. IEEE Comput. Graph. Appl. 1(1), 11–23 (1981)
[2] Paschalidou, D., Ulusoy, A.O., Geiger, A.: Superquadrics revisited: learning 3D shape parsing beyond cuboids. In: CVPR (2019)
[3] Makhal, A., Thomas, F., Gracia, A.P.: Grasping unknown objects in clutter by superquadric representation. CoRRabs/1710.02121 (2017).
[4] Vaskevicius, N., Pathak, K., Birk, A.: Recognition and localization of sacks for autonomous container unloading by fitting superquadrics in noisy, partial views from a low-cost rgbdsensor. Journal of Intelligent & Robotic Systems 88(1)
[5] Nekrasov, V., Shen, C., Reid, I.D.: Light-weight refinenetfor real-time semantic segmentation. In: BMVC (2018)
[6] Chen, X., Mottaghi, R., Liu, X., Fidler, S., Urtasun, R., Yuille, A.L.: Detect what you can: detecting and representing objects using holistic models and body parts. In: CVPR, pp. 1979–1986 (2014)
[7]Shotton, J., et al.: Real-time human pose recognition in parts from single depth images. In: CVPR, pp. 1297–1304, June 2011
[8] A. Jaklic, A. Leonardis, and F. Solina, "Segmentation and recovery of superquadrics," Computational imaging and vision, vol. 20, 2000.