

# Hongwei Yi

Max Planck Institute for Intelligent Systems  
Perceiving Systems  
Max Ring 4, Tuebingen, Germany  
Phone: (+49) 015237672470  
Email: hongwei.yi@tuebingen.mpg.de



## Education

Ph.D Student      Researching on understanding/capturing/generating human motion and interaction with 3D scenes, and 3D scene understanding.      since 2020.09-

**MPI-IS, Supervised by Michael J. Black**

M.A.      Computer Application Technology      2017.09-2020.07

**Peking University**

B.S.      Computer Science and Technology      Top 10% (30/300) of Grade      2013.09-2017.07

**Beijing University of Posts and Telecommunications**

## Publications

### *Proceedings*

*TADA! Text to Animatable Digital Avatars. (In ArXiv)*

Tingting Liao\*, **Hongwei Yi\***, Yuliang Xiu, Jiaxiang Tang, Yangyi Huang, Justus Thies, Michael J. Black (\* denotes equal contribution.)

*TeCH: Text-guided Reconstruction of Lifelike Clothed Humans. (In ArXiv)*

Yangyi Huang\*, **Hongwei Yi\***, Yuliang Xiu\*, Tingting Liao, Jiaxiang Tang, Deng Cai, Justus Thies (\* denotes equal contribution.)

*Real-time Monocular Full-body Capture in World Space via Sequential Proxy-to-Motion Learning. (In arXiv)*

Yuxiang Zhang, Hongwen Zhang\*, Liangxiao Hu, **Hongwei Yi**, Shengping Zhang, Yebin Liu1\*(\* denotes corresponding authors.)

*GraMMaR: Ground-aware Motion Model for 3D Human Motion Reconstruction. (ACMMM2023)*

Sihan Ma, Qiong Cao, **Hongwei Yi**, Jing Zhang, Dacheng Tao

*DECO: Dense Estimation of 3D Human-Scene COntact in the Wild. (ICCV2023 ORAL)*

Shashank Tripathi\*, Agniv Chatterjee\*, Jean-Claude Passy, **Hongwei Yi**, Dimitrios Tzionas, Michael J. Black (\* denotes equal contribution.)

*One-shot Implicit Animatable Avatars with Model-based Priors. (ICCV2023)*

Yangyi Huang\*, **Hongwei Yi\***, Weiyang Liu, Haofan Wang, Boxi Wu, Wenxiao Wang, Binbin Lin, Debing Zhang, Deng Cai (\* denotes equal contribution.)

*MIME: Human-Aware 3D Scene Generation. (CVPR2023)*

**Hongwei Yi**, Chun-Hao P. Huang, Shashank Tripathi, Lea Hering, Justus Thies, Michael J. Black.

*Generating Holistic 3D Human Motion from Speech. (CVPR2023)*

**Hongwei Yi\***, Hualin Liang\*, Yifei Liu\*, Qiong Cao, Yandong Wen, Timo Bolkart, Dacheng Tao, Michael J. Black. (\* denotes equal contribution.)

*SLOPER4D: A Scene-Aware Dataset For Global 4D Human Pose Estimation In Urban Environments. (CVPR2023)*

Yudi Dai, YiTai Lin, XiPing Lin, Chenglu Wen, Lan Xu, **Hongwei Yi**, Siqi Shen, Yuexin Ma, Cheng Wang

*High-fidelity Clothed Avatar Reconstruction from a Single Image. (CVPR2023)*

Tingting Liao, Xiaomei Zhang, Yuliang Xiu, **Hongwei Yi**, Xudong Liu, Guo-Jun Qi, Yong Zhang, Xuan Wang, Xiangyu Zhu, Zhen Lei

*Human-Aware Object Placement for Visual Environment Reconstruction. (CVPR2022)*

**Hongwei Yi**, Chun-Hao P. Huang, Dimitrios Tzionas, Muhammed Kocabas, Mohamed Hassan, Siyu Tang, Justus Thies, Michael J. Black.

*Capturing and Inferring Dense Full-Body Human-Scene Contact. (CVPR2022)*

Chun-Hao P. Huang, **Hongwei Yi**, Markus Höschle, Matvey Sefroshkin, Tsvetelina Alexiadis, Senya Polikovsky, Daniel Scharstein, Michael J. Black

*Pyramid Multi-view Stereo Net with Self-adaptive View Aggregation. (ECCV2020)*

**Hongwei Yi\***, Zizhuang Wei\*, Mingyu Ding, Runze Zhang, Yisong Chen, Guoping Wang, Yu-Wing Tai. (\* denotes equal contribution.)

*Dense Hybrid Recurrent Multi-view Stereo Net with Dynamic Consistency Checking. (ECCV2020)*

Jianfeng Yan\*, Zizhuang Wei\*, **Hongwei Yi\***, Mingyu Ding, Runze Zhang, Yisong Chen, Guoping Wang, Yu-Wing Tai. (\* denotes equal contribution.)

*Learning Depth-Guided Convolutions for Monocular 3D Object Detection. (CVPR2020)*

Mingyu Ding, Yuqi Huo, **Hongwei Yi**, Zhe Wang, Jianping Shi, Zhiwu Lu, Ping Luo.

*SegVoxelNet: Real-time 3D Object Detection from Point Cloud. (ICRA2020)*

**Hongwei Yi**, Shaoshuai Shi, Mingyu Ding, Jiankai Sun, Kui Xu, Sheng Li, Guoping Wang, Hui Zhou, Zhe Wang.

*MMFace: A Multi-Metric Regression Network for Unconstrained Face Reconstruction. (CVPR2019)*

**Hongwei Yi**, Chen Li, Qiong Cao, Xiaoyong Shen, Sheng Li, Guoping Wang, Yu-Wing Tai.

## Journals

Wei, Z.; Wang, Y.; **Yi, H.**; Chen, Y.; Wang, G. Semantic 3D Reconstruction with Learning MVS and 2D Segmentation of Aerial Images. Appl. Sci. 2020, 10, 1275.

## Research Experience

0.1 *PhD Student. Perceiving Systems. Max Planck Institute for Intelligent Systems. 2020.09*

0.2 *Master Student. Graphics & Interactive Lab. Peking University. 2017.09-2020.07.*

**Researching on Multi-view Stereo.** From August 2019 till now, I have been focusing on the research of deep-learning-based multi-view stereo matching (MVS) method, which is one of the most important parts of 3D reconstruction of large scale scenes. Firstly, I propose an effective and efficient pyramid

multi-view stereo (MVS) net for accurate and complete dense point cloud reconstruction, namely PVA-MVSNet. Secondly, I propose an efficient and effective dense hybrid recurrent multi-view stereo net with dynamic consistency checking, namely  $D^2$ HC-RMVSNet, for accurate dense point cloud reconstruction. Both two network architectures are organized into two papers submitted to ECCV2020.

**Benchmark** our PVA-MVSNet establishes a new state-of-the-art on the *DTU* dataset with significant improvements in the completeness and overall quality, and has strong generalization by achieving a comparable performance as the state-of-the-art methods on the *Tanks and Temples* benchmark. Our  $D^2$ HC-RMVSNet ranks 1<sup>st</sup> on the complex outdoor *Tanks and Temples* benchmark over all the methods, and exhibits competitive performance to the state-of-the-art method while dramatically reduces memory consumption, which costs only 19.4% of R-MVSNet memory consumption.

### 0.3 *Research Intern. Auto-driving Group. SenseTime Group Limited. 2018.12-2019.08.*

**Researched on 3D Object Detection.** I proposed a unified framework called SegVoxelNet that incorporates semantic segmentation information into 3D object detection, where semantic context becomes active guidance for 3D object detection. Moreover, an efficient multi-branch detection head is designed for objects with different depths in autonomous driving scenarios. My SegVoxelNet achieves new state-of-the-art results on the challenging 3D detection benchmark of KITTI dataset with real-time efficiency.

**3D Object KITTI Benchmark:** rank 5 (including anonymous submit methods).

### 0.4 *Research Intern. YouTu-XLAB. Tencent Technology Company. 2017.12-2018.11.*

**Researched on 3D face reconstruction and 3D face alignment.** My tutor and I proposed the Multi-Metric regression network to construct accurate 3D aligned face model from one single image. First we used volumetric network to extract 3D voxel feature map from one image, then implied parametric network to regress 3D morphable model parameters such as shape, expression and pose through 3D voxel feature map, furthermore we used ICP to align 3D morphable model on 3D face voxels for more precise pose information. Our method both got the state-of-the-art results in 3D face reconstruction and 3D face alignments comparing with recent methods in CVPR, ECCV and ICCV.

**Researched on face detection and face recognition.** I had already re-implemented MTCNN face detector on Caffe using C++ and spent a short period research on face recognition methods include ArcFace.

### 0.5 *Deep Learning and Computer Vision Engineer Intern. Beijing DeepGlint Technology Company. 2016.10-2017.06.*

**Developed and Optimized online tracking algorithm of Video Analysis System.** First I used SSD detector to detect car, cyclist and person. Then I integrated single object tracker, motion information, localization and HOG features to track different objects.

**Developed a MOT Demo.** Using a delicate SSD, high level person-reid features and MDP, I had got the highest tracking results, though with some sacrifice on time cost.

**Developed a complete efficient person-reid Label Tool System.**

## Professional Skills

Language: Chinese (native), English (fluent, IELTS-7, CET-6)

Code Languages: PYTHON > C++ > C > MATLAB > JAVA > JSP > JavaScript > HTML+CSS

Deep Learning Architecture: Pytorch > Torch > Caffe > TensorFlow

## Awards and Scholarships

2019. Tiehan Scholarship

2016. First Prize in the NXP Cup National University Students Intelligent Car Competition

2016. Second Prize of American Graduate Mathematical Modeling Contest

2014-2016. The First/Second Prize Scholarship