Universiteit Leiden

Master Computer Science

FESD: A Fault Estimation Pipeline for Human Pose Estimation

Name:	Leonardo Benedikt Pohl
Student ID:	s3059030
Date:	30/05/2023
Specialisation:	Advanced Computing and Systems
1st supervisor:	Prof. Dr. Erwin Bakker
2nd supervisor:	Prof. Dr. Michael Lew

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

References

[1] Sizhe An, Yin Li, and Umit Ogras. mRI: Multi-modal 3D Human Pose Estimation Dataset using mmWave, RGB-D, and Inertial Sensors. *CVPR*, 2022.

[2] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.

[3] Junjie Bai, Fang Lu, Ke Zhang, et al. ONNX: Open Neural Network Exchange. https://github.com/onnx/onnx, 2019.

[4] Abhijat Biswas, Henny Admoni, and Aaron Steinfeld. Fast on-board 3d torso pose recovery and forecasting*. 2018.

[5] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[6] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In *CVPR*, 2017.

[7] Joao Carreira, Pulkit Agrawal, Katerina Fragkiadaki, and Jitendra Malik. Human Pose Estimation with Iterative Error Feedback. *CVPR*, 2015.

[8] Cristian Sminchisescu Catalin Ionescu, Fuxin Li. Latent Structured Models for Human Pose Estimation. In *International Conference on Computer Vision*, 2011.

[9] Kenny Chen, Paolo Gabriel, Abdulwahab Alasfour, Chenghao Gong, Werner K. Doyle, Orrin Devinsky, Daniel Friedman, Patricia Dugan, Lucia Melloni, Thomas Thesen, David Gonda, Shifteh Sattar, Sonya Wang, and Vikash Gilja. Patient-Specific Pose Estimation in Clinical Environments. *IEEE Journal of Translational Engineering in Health and Medicine*, 6:1–11, 2018.

[10] Qian Chen, Ze Liu, Yi Zhang, Keren Fu, Qijun Zhao, and Hongwei Du. RGB-D Salient Object Detection via 3D Convolutional Neural Networks. 2021.

[11] Qian Chen, Zhenxi Zhang, Yanye Lu, Keren Fu, and Qijun Zhao. 3-D Convolutional Neural Networks for RGB-D Salient Object Detection and Beyond. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–15, 2022.

[12] Yucheng Chen, Yingli Tian, and Mingyi He. Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, 192:102897, 2020.

[13] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[14] Shradha Dubey and Manish Dixit. A comprehensive survey on human pose estimation approaches. *Multimedia Systems*, 29, 08 2022.

[15] Abdessamad Elboushaki, Rachida Hannane, Karim Afdel, and Lahcen Koutti. MultiD-CNN: A multi-dimensional feature learning approach based on deep convolutional networks for gesture recognition in RGB-D image sequences. *Expert systems with applications*, 139:112829, 2020.

[16] Leonhard Euler. Formulae generales pro translatione quacunque corporum rigidorum. *Novi Commentarii academiae scientiarum Petropolitanae*, pages 189–207, 1776.

[17] Martin Fisch and Ronald Clark. Orientation Keypoints for 6D Human Pose Estimation. *CVPR*, 2020.

[18] Kunihiko Fukushima. Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, 1(2):119–130, 1988.

[19] Varun Ganapathi, Christian Plagemann, Daphne Koller, and Sebastian Thrun. Real time motion capture using a single time-of-flight camera. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 755–762, 2010.

[20] Varun Ganapathi, Christian Plagemann, Daphne Koller, and Sebastian Thrun. Real-Time Human Pose Tracking from Range Data. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2012.

[21] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[22] H. Haggag, M. Hossny, S. Nahavandi, and O. Haggag. An adaptable system for RGB-D based human body detection and pose estimation: Incorporating attached props. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 001544–001549, 2016.

[23] Intel. Tensorflow with Intel Realsense camera. https://dev.intelrealsense.com/docs/tensorflow-with-intel-realsense-cameras. Accessed: 18-02-2023.

[24] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, jul 2014.

[25] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2017.

[26] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[27] Fabian Kröger. *Automated Driving in Its Social, Historical and Cultural Contexts*, pages 41–68. Springer, 05 2016.

[28] Laxman Kumarapu and Prerana Mukherjee. AnimePose: Multi-person 3D pose estimation and animation. *Graphics*, 2020.

[29] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on $\mathcal{X}$-transformed points. *CVPR*, 2018.

[30] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft COCO: Common Objects in Context, 2014.

[31] Yunheng Liu. Contour Model and Robust Segmentation based Human Pose Estimation in Images and Videos. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 8:1–10, 03 2015.

[32] Mary McHugh. Interrater reliability: The kappa statistic. *Biochemia medica : časopis Hrvatskoga društva medicinskih biokemičara / HDMB*, 22:276–82, 10 2012.

[33] Sameh Neili Boualia and Najoua ESSOUKRI BEN AMARA. Deep Full-Body HPE for Activity Recognition from RGB Frames Only. *Informatics*, 8:2, 01 2021.

[34] David Pascual-Hernández, Nuria Oyaga de Frutos, Inmaculada Mora-Jiménez, and José María Cañas-Plaza. Efficient 3D human pose estimation from RGBD sensors. *Displays*, 74:102225, 2022.

[35] Bernhard Preim and Monique Meuschke. A survey of medical animations. *Computers and Graphics*, 107, 09 2022.

[36] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015.

[37] Carl F. Sabottke and Bradley M. Spieler. The Effect of Image Resolution on Deep Learning in Radiography. *Radiology: Artificial Intelligence*, 2(1):e190015, 2020. PMID: 33937810.

[38] Bassem Seddik, Sami Gazzah, and Najoua Essoukri Ben Amara. Human-action recognition using a multi-layered fusion scheme of Kinect modalities. *IET Computer Vision*, 11(7):530–540, 2017.

[39] Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. Real-time human pose recognition in parts from single depth images. In *CVPR 2011*, pages 1297–1304, 2011.

[40] Leonid Sigal, Alexandru Balan, and Michael Black. HumanEva: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion. *International Journal of Computer Vision*, 87:4–27, 03 2010.

[41] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. In *CVPR*, 2017.

[42] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014.

[43] Liangchen Song, Gang Yu, Junsong Yuan, and Zicheng Liu. Human pose estimation and its application to action recognition: A survey. *Journal of Visual Communication and Image Representation*, 76:103055, 04 2021.

[44] Jan Stenum, Kendra M. Cherry-Allen, Connor O Pyles, Rachel Reetzke, Michael F. Vignos, and Ryan T. Roemmich. Applications of Pose Estimation in Human Health and Performance across the Lifespan. *Sensors (Basel, Switzerland)*, 21, 2021.

[45] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, Vijay Vasudevan, Wei Han, Jiquan Ngiam, Hang Zhao, Aleksei Timofeev, Scott Ettinger, Maxim Krivokon, Amy Gao, Aditya Joshi, Sheng Zhao, Shuyang Cheng, Yu Zhang, Jonathon Shlens, Zhifeng Chen, and Dragomir Anguelov. Scalability in Perception for Autonomous Driving: Waymo Open Dataset, 2019.

[46] Michal Tölgyessy, Martin Dekan, and Lubos Chovanec. Skeleton Tracking Accuracy and Precision Evaluation of Kinect V1, Kinect V2, and the Azure Kinect. *Applied Sciences*, 11:5756, 06 2021.

[47] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines. In *CVPR*, 2016.

[48] Tianxu Xu, Dong An, Yuetong Jia, and Yang Yue. A Review: Point Cloud-Based 3D Human Joints Estimation. *Sensors*, 21:1684, 03 2021.

[49] Jingxiao Zheng, Xinwei Shi, Alexander Gorban, Junhua Mao, Yang Song, Charles R. Qi, Ting Liu, Visesh Chari, Andre Cornman, Yin Zhou, Congcong Li, and Dragomir Anguelov. Multimodal 3D Human Pose Estimation with 2D Weak Supervision in Autonomous Driving, 2021.

[50] Christian Zimmermann, Tim Welschehold, Christian Dornhege, Wolfram Burgard, and Thomas Brox. 3D Human Pose Estimation in RGBD Images for Robotic Task Learning, 2018.