https://doi.org/10.2352/ISSN.2470-1173.2021.18.3DIA-118

Verification and Regularization Method for 3D-Human Body Pose Estimation based on Prior Knowledge

Julia Denecke, Christian Jauch; Fraunhofer Institute for Manufacturing Engineering and Automation IPA; Stuttgart, Germany; Norbert Link; Inferics GmbH; Karlsruhe, Germany

Abstract

Applications used in human-centered scene analysis often rely on AI processes that provide the 3D data of human bodies. The applications are limited by the accuracy and reliability of the detection. In case of safety applications, an almost perfect detection rate is required. The presented approach gives a confidence measure for the likelihood that detected human bodies are real persons. It measures the consistency of the estimated 3D pose information of body joints with prior knowledge about the physiologically possible spatial sizes and proportions. Therefore, a detailed analysis was done which lead to the development of an error metric that allows the quantitative evaluation of single limbs and in summary of the complete body. For a given dataset an error threshold has been derived that verifies 97% persons correctly and can be used for the identification of false detections, so-called ghosts. Additionally, the 3D-data of single joints could be rated successfully. The results are usable for relabeling and retraining of underlying 2D and 3D pose estimators and provides a quantitative and comparable verification method, which improves significantly a reliable 3D-recognition of real persons and increases hereby the possibilities of high-standard applications of 3D human-centered technologies.

Motivation

The detection of humans with imaging systems has become a standard application in the last years. Thanks to the developments in object detection with AI-techniques and huge open source image datasets, the implementation of person and rough pose detection in 2D image is working in adequate quality e.g. for gaming and interaction systems. Complex scene analysis and safety applications, like emergency detection, need a higher reliability and have to ensure that all relevant situations are detected by reducing false alerts at the same time.
The authors are working on the detection of human pose and activity using the 3D embedded vision systems. In this work, the aim of scene understanding is focused on situations like safety at home, especially for elderly people, public safety, meaning the automated analysis of conventional video surveillance and the support and acquisition of retail experience. An embedded system demands frugal solutions, why the 2D images are processed by with efficient AI-methods for human activity detection. 2D image inference is biased by training data and sometimes lost without scene knowledge. Estimators of persons poses from image data based on deep neural networks sometimes confuse image background structures with body key points (joints, eyes, ears, nose) and generate so-called ghosts or pose errors (Figure 2). This problem leads to the limitation of existing algorithms for applications of interactive products where the frequency of such errors is tolerable.
In the presented work, the basic idea is to use prior knowledge about the physiologically possible spatial arrangements of the bones and anatomical body characteristics in general. This is achieved by assigning the 3D coordinates from an aligned depth image to 2D image coordinates of the body key points, deriving the corresponding body anatomy in 3D, and comparing it to constraints known from the prior knowledge. A verification and regularization method for 3D-human body pose estimation based on anatomical prior knowledge is developed.

In the first section related work of human body estimation and previous ideas of the authors are presented. Then the findings from the physiology literature with respect to anatomical characteristics are presented. Next, the evaluation of promising characteristics with respect to our purpose is described. This is followed by the development of our verification method and the determination of an error metric based on an experimental setup. The evaluation of the proposal and first results are shown and in the end, a summary and outlook is given.

Related Work

The optical detection and localization of persons, as well as the analysis of their activity and movement patterns, is increasingly used for the people counting, entrance control and activity classification. Usually AI methods like [1][2][3] are employed for the recognition of situations and activities, which are based on end-to-end learning and directly processing RGB-d image data streams, such as RNN (e.g. LSTM). Especially RNN or LSTM and the typical sensory 3D detection is carried out by means of stereo cameras. 3D & 4D Body Scan and 3D Motion Capturing systems are used for special generation of models of human movement. The activity detection can be realized by classifying temporal image or pose sequences and/or the analysis of the optical flow, also mostly method LSTM. In current research, the focus is on the topics of 4D body scanning for direct motion digitization (e.g. 4D Dynamic Scanner [4], [5]), on various AI solutions for activity detection, "Human in Context", e.g. detection of person-object interaction [6] as well as the creation of specialized data sets e.g. hand gestures or torso movement. Outstanding works in recent years include the development of realistic data sets for human modelling from various technologies, in particular Meshcapade e.g. FAUST Dataset [7], biometrical human modeling especially with focus on ergonomic aspects [8], [9], the open library OpenPose for real-time capable 2D multi-person [10] and pose recognition with Part Affinity Fields and "Human in Scene Context", in particular contact with objects [11].

Most of the approaches need a high performance and thereby expensive hardware setup and are not focused on the application. The Inferics company is working on software solutions for person activity detection based on optimized embedded hardware systems. Especially the Hemispherical Trinocular 3D Vision Plattform [12] allows the single use of one sensor centered on the ceil by acquiring complete room 2,5D-point cloud with a FOV 180° x 180°, RGB-D resolution of 1200 x 1200 px and onboard application development with edge computing (NVIDIA Jetson TX2), Figure 1.

Figure 1. Resulting data (left) of the Hemispherical Trinocular 3D Vision Plattform and the sensor system (right)

But the state-of-the-art development in human detection does not deal intensely with overhead placement of the optical sensor. In conclusion, the direct transfer of existing AI-detectors is not possible. In addition, the data include more occlusions of body parts than front cameras. Because of the embedded process, the privacy can be guaranteed, but the system performance is restricted. To get close to real-time detection, only 2D-information is taken into account and a skeleton model is matched into the image. However, the existing, corresponding 3D-data points provides the necessary information to reveal false detection. In Figure 2, a ghost and a half body is estimated wrong in 2D.

Figure 2. False detection by patterns of a ghost (left), partly false detection of the pose (right)

The weakness of 2D detection can be faced with the 3D information. The presented scientific approach aims for a verification method if detected human bodies are real by using and comparing the 3D information of joints and prior knowledge about the physiologically possible spatial sizes of the limb. The challenge is to identify and integrate relevant knowledge based on anatomical and statistical restrictions like proportions or possible human sizes correctly into the AI process in order to evaluate both, "real" human body and correct joints in 3D detection.

Evaluation of adequate characteristics

Before preknowledge can be used, it is necessary to identify what kind of data and sizes exist about human physiology, how are the natural ranges and are they measureable with the described 2,5D-joint model gained from ceiling camera position. Inclusion of knowledge about human anatomy generally offers the following possibilities:

- Verification whether it is a real person
- Verification that the pose is correct
- Adjustment the confidence of detection for each joint
- Possible correction of individual joint positions
- Differentiation between child and adult

Types of knowledge about human anatomy are absolute sizes, proportions, movement types and poses. For example, we can find some knowledge about body proportions of adults (from literature) like the average size which is up to 7.5 head sizes large, idealized 8, with following partitions:

- from the apex to the chin (head length)
- from there to the middle of the breast (approximately at the height of the nipples)
- from there to the belly button
- from there to the pubic area
- from there to the middle of the thigh
- from there to just below the knee
- from there to the middle of the calves
- from there to the sole of the foot

More relevant proportions are identified like:

- The pubic area is located in the middle of the body.
- The lower leg is just as long as the thigh.
- Hanging arms are so long that the fingertips reach the center of the thighs.
- The wingspan of the arms (from the fingertip of the middle finger to the tip of the finger) corresponds to the total height.
- The foot length is about as long as the forearm without the hand.

An absolute size is the complete body height, where the natural range can be extracted from statistical records:
Height (Adults): largest human 2.72 m and smallest human 0.59 m, Average Germany: 1.79 m Men, 1.66 m Women and 1.72 m total [13].

We evaluated the transferability of human anatomical and physiological characteristics like these static sizes proportions and motion degrees of freedom. We also looked into the acquisition of quantifiable parameters including the limitations of the degrees of freedom. Therefore, a dataset of detected bodies in the 2D images with defined pose classes were used. For every detected joint the corresponding 3D position was taken and based on the spatial information the proportions and sized were calculated and analyzed if possible. The correspondence of the anatomical and the joint model can be derived directly e.g. arm and leg parts or has to be calculated e.g. wingspan is the sum of all arm parts and length between shoulders. The results show that the high variation of sizes often leads uncertain data. Especially the more bones and joints are included in the calculation, the less the meaning of the characteristic, e.g. the wingspan. But we managed to get the proof of concept for the body size, which is suitable as characteristic for person verification. This means the determination of person height from all bones with help of constant scale factors and using them as estimated body size. We observed a strong association with the current pose (interaction must be regarded) due to the ceiling position. Although the uniformity and proportions of bones vary greatly, the outliers can be used to identify false detected joints. Additionally, for the quantification of natural limitation movements, we focus on the methods of physiotherapy. Here, the neutral-null-method [14] is used to classify the mobility of every joint and body extremity with fix ranges based on a defined neutral position and body axis:

➔ "Neutral position" or zero position: The man stands upright, the arms are relaxed downwards, the thumbs facing forward and the feet stand parallel.

➔ Description of 3-angle mobility for each possible body axis:
  - Angle 1: the deflection in the distance from the body
  - Angle 2: Normal case 0° (= neutral position)
  - Angle 3: the deflection in the near-body direction

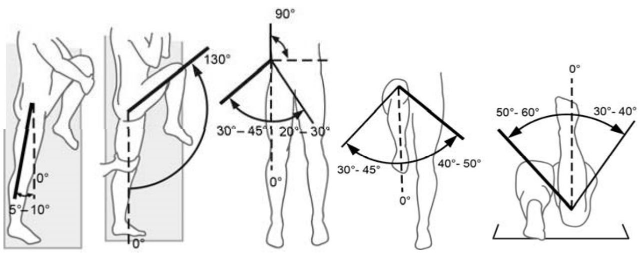

Figure 3: Example of agility measurement with neutral null method: Hip Joint Angle values for Stretching / Diffraction: 10-0-130 and for Abduction / Adduction: 45-0-30, Rotation outward / inward: 45-0-50, Rotation outward / inward (hip joint stretched): 40-0-60 [15]

Proposed Method

We use a 3D trinocular stereo vision system that allows complete room analysis from a single centered overhead position with fisheye camera imaging. First, we elaborate measures that correspond to detected 3D body arrangements. Here, the person size calculated out of statistical proportions for every limb shows a very good behavior considering verification of detection. In addition, relative proportions like leg or arm lengths are considered. Empirical analyses lead to the development of an error metric that allows the quantitative evaluation of every single limb and in summary for the complete body. Here, the AI given confidence values and the image person and pose detection results are taken into account.

Based on the evaluation of calculable sizes and characteristics, solution strategies for the following three problems have been developed:

1. "Recognized person is not a real person" (ghost or image),
2. "Single joints are incorrect", this could be due to incorrect detections or an incorrect spatial assignment from the 2D image point to the corresponding 3D-position.
3. "Angle between bones is anatomically not possible" (based on movement characteristics)

To verify the usability of the prior knowledge, an empirical approach is derived. Therefore, an evaluation data set consists of RGB-D image data and ground-truth data of the respective poses ("lie", "sit", "stand") and the 2D-AI based detected skeleton with confidence value for every Joint, is used as experimental setup.

In the first step, a detailed calculation and evaluation of the respective parameters is done: body sizes out of 3D-bone length and neutral-null-method for poses were determined. This includes the consideration of tolerances and limitations. In the next step, we developed a special error metric for the classification of complete bodies, single joints and poses (joint angles). This error metric needs adequate thresholds, which were estimated experimentally.

The proposed procedure:

1)
   a) Problem 1 and 2: Implementation of body size calculation (including a corresponding visualization for fast visual classification) based on anatomical ratios and determined proportions out of the 3D-joint information
   b) Proposed procedure for problem 3: Implementation of the neutral-null-method (for upper and lower extremities): range of motion of a joint in angular degrees around a certain axis incl. standard values

2) Detailed evaluation of the characteristic body size for upper body and shoulder lengths and extremities depending on confidence and pose
3) Consideration of distributions of body size based on individual bones also in the context of anatomical limits
4) Derivation of suitability and information for the entire body model
5) Development of a metric for error assessment for the entire body as well as individual joints
6) Test of the exemplary evaluation with determined threshold values

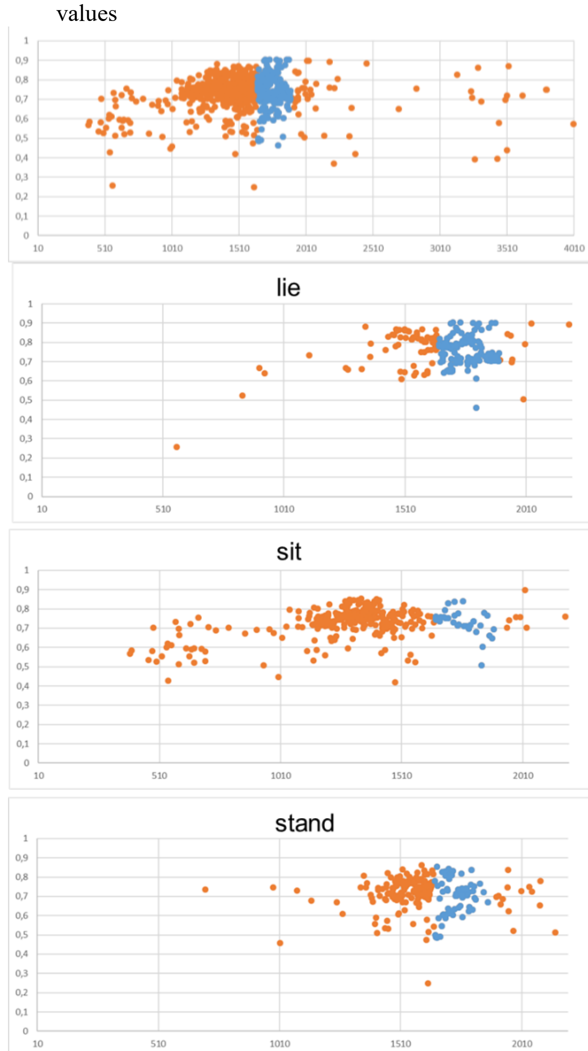

*Figure 4: Example torso: Body height calculated as 3.2 * Distance (Neck – MidHip, Y-axis: Confidence, X-axis: height in mm, Blue dots: Body height in range 1.65 m - 1.90 m (all data in upper plot, and separated by pose below)*

First empirical results of implementation lead to some necessary adjustments of the scale factors that are used to extrapolate the body height from bone length. In Figure 4, the estimated body height and corresponding confidence value of the bone is plotted. We define the lowest confidence value of both joint ends as the confidence of the bone. The underlying data set consists of detection of the same person with the known body height of 1,80 m. According to the population statistic, an acceptable range for the person verification between 1.40 m and 2.20 m has proven reasonable. The separation of the poses (lying, sitting, standing) shows different clusters of the calculated body height. Especially the results for torso length and legs depend on the pose. In tendency, arms and shoulders are independent of the pose. In general, the least dispersion is observed if the torso length is used for body height estimation, although this size tends to be too small in sitting poses. Here, as torso length the distance between detected joint of the neck and the middle hip is defined. Besides the estimated values for the absolute body height, the confidence value is taken into account. Regarding the evaluation results a confidence above 0.5 is plausible, but only considered as sufficient condition.

Experimental Determination of Error Metric

Based on the initial evaluation an error metric is determined with the aim of a measurable and comparable size for the validation of detected persons and single joints. All considerations and observation are summarized and formulated.

As feature values for the validity both mentioned sizes are needed, the body height h_b and the lower value of the both end joint confidences c_l. The body size h_b is calculated from the bone length multiplied with the corresponding scale factor gained from natural proportions and adjusted for the used skeleton model.

Validity ranges for the body height are the upper limit h_{ul}, lower limit h_{ll} and a tolerance range Δ which is an excess range above h_{ul} and below h_{ll}. For the confidence values, the lower limit c_{ll} is the only limitation.

So, it is defined that a bone is valid, if h_b is within the tolerance range and if c_l is above c_{ll}. It is also valid, if h_b is within the body height limits only. The validity flag v is then one and otherwise zero:

$$v = \begin{cases} 1 \; if \; ((h_b > h_{ul} + \Delta) \vee (h_b < h_{ll} - \Delta)) \wedge (c_l < c_{ll}) \\ 1 \text{ if } ((h_b \leq h_{ul}) \wedge (h_b \geq h_{ll})) \\ 0 \text{ else} \end{cases}$$

The bones with $\boldsymbol{v = 1}$ form the set of valid bones $\boldsymbol{V}$ with cardinality $\boldsymbol{n_V}$.

For a valid bone the error e_v is measured as following:

$$e_v \;=\; 0, \; if \; (h_b \leq h_{ul}) \; \wedge \; (h_b \geq h_{ll}) \; \wedge \; (c_l \geq c_{ll})$$

Otherwise, the error is determined depending on the quality of the matching conditions. The complete error metric is

$$e_v = \begin{cases} \dfrac{h_b - h_{ul}}{\Delta}, \; if \; h_b > h_{ul} \\ \dfrac{h_{ll} - h_b}{\Delta}, if \; h_b < h_{ll} \\ 0.5 - c_{ll}, if \; c_{ll} < 0.5 \\ 0, else \end{cases}$$

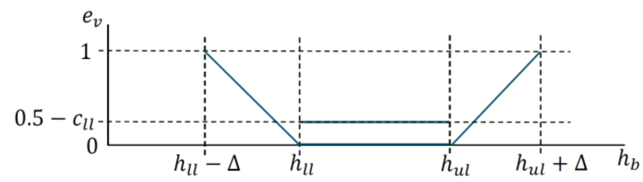

Figure 5: Error metric depending on the estimated height and bone confidence

The error measure for the complete body $\boldsymbol{E_{Body}}$ is the average of the joint errors:

$$\boldsymbol{E_{Body}} = \frac{\boldsymbol{1}}{\boldsymbol{n_v}} \sum\nolimits_{v} \boldsymbol{e_v}$$

If $\boldsymbol{n_V} < \boldsymbol{N_{thresh}}$, then the body is not valid. So a minimum of valid bones is necessary to declare the body as valid. Finally, a threshold $\boldsymbol{E_{thresh}}$ can be used to distinguish between valid and invalid bodies. If $\boldsymbol{E_{Body}} > \boldsymbol{E_{thresh}}$, then the body is not valid. The validation of individual joints can be taken directly by the error value. Here, a classification can be done by ranking the joints, if they are valid with $e_v = 0$, valid in tolerance with $e_v > 0$ or invalid $e_v > 1$ (definition of invalid in implementation). An example of this error for torso length characteristic for the complete data set is presented in Figure 6. A very good correspondence for poses 'lie' and 'stand' is obvious while high error values for the pose 'sit' are visible.

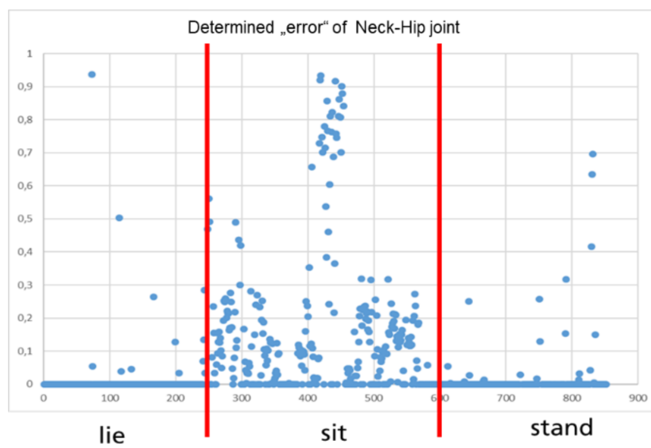

Figure 6: Determined error for the torso bone for every data, Y-axis: calculated error, X-axis: nb of data

To determine the optimal error threshold, we evaluated how many detections pass as valid persons, since it is known that all of them are valid persons. This is done for every bone as well as for the complete body.

Conclusions of comparison of classification of valid persons are that the total value for the complete body with threshold 0.3 achieve over 97%. Best results of a single bone are observed for upper body/torso. The value for the upper arm and forearm reach max. 75%, which is not sufficient for a validity determination. Also, the shoulder reaches only at high thresholds high proportion of detected persons and is not singular usable for the validation of a person.

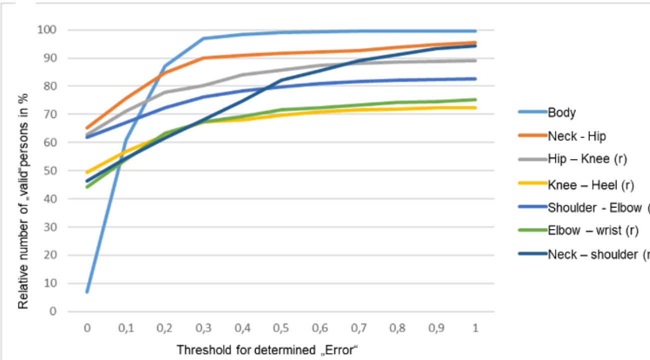

Figure 7: Comparison of classification of valid persons for body and single joints by error value

Agility Measurements with Neutral-Null-Method

The implementation of the neutral null method bases on the adequate definition of the "Neutral Position" or a fixed coordinate system of the body with exiting joints. The least flexible body area is the torso. Therefore, the actual body coordinate system for every detected skeleton is defined separately like this:

- Zero position is the right hip joint
- X-axis: direction from right hip joint j_{rhip} to left hip joint j_{lhip}: $\vec{x} = \overrightarrow{j_{lhip}} - \overrightarrow{j_{rhip}}$
- Y-axis: $\vec{y} = \vec{x} \times \overrightarrow{z_{help}}$, where $\overrightarrow{z_{help}} = \overrightarrow{j_{rshoulder}} - \overrightarrow{j_{rhip}}$

using the joint of the right shoulder $j_{rshoulder}$

- Z-axis: $\vec{z} = \vec{x} \times \vec{y}$

We are aware that there exist various possibilities of body distortions that corrupts this coordinate definition. However, regarding usual behavior and movement for the examined poses sit, lie and stand, this definition is sufficient for the proof of concept.
In the next step, the procedure is exemplary tested with the right leg and hip agility. Therefore, the joint of the right knee is projected on the yz-plane represented by $j_{rkneeProj}$. Then, the angle is calculated between $\vec{z}$ and $\overrightarrow{j_{rkneeProj}} - \overrightarrow{j_{rhip}}$. As result, the angular determination of hip joints corresponds roughly with stretch / diffraction limitation of the neutral null definitions with 10-0-130. Again, a confidence over 0.5 is assumed. In Figure 8 the histograms show, that the results correspond to expected values and ranges.
Accordingly, the developed error metric can be adapted for this pose validation. Sincerely, the ground truth data set was not applicable, because the actual angles vary extremely (e.g. pose lying with straight legs or in bent pose). A special data set is needed for the experimental threshold estimation.

Results and Evaluation

For the evaluation, a data set of 852 successfully detected persons and correct poses were regarded. The experimental setup is again fixed with the described hardware setup. $\boldsymbol{E_{thresh}}$ is set to 0.3 and the minimum number of valid bones $\boldsymbol{N_{thresh}}$ is set to 3. The tolerance ranges are $h_{ll} = 1.40$, $h_{ul} = 2.20$ and $\Delta = 1.00$. In conclusion, a detected person is valid, if $\boldsymbol{E_{Body}} < 0.3$ and more than 3 valid bones exist. For the single bones the following visualization of classification is used:

- Green, if height is between 1.40 m and 2.2 m
- Blue, if in the tolerance band of 1 m
- Red, if invalid

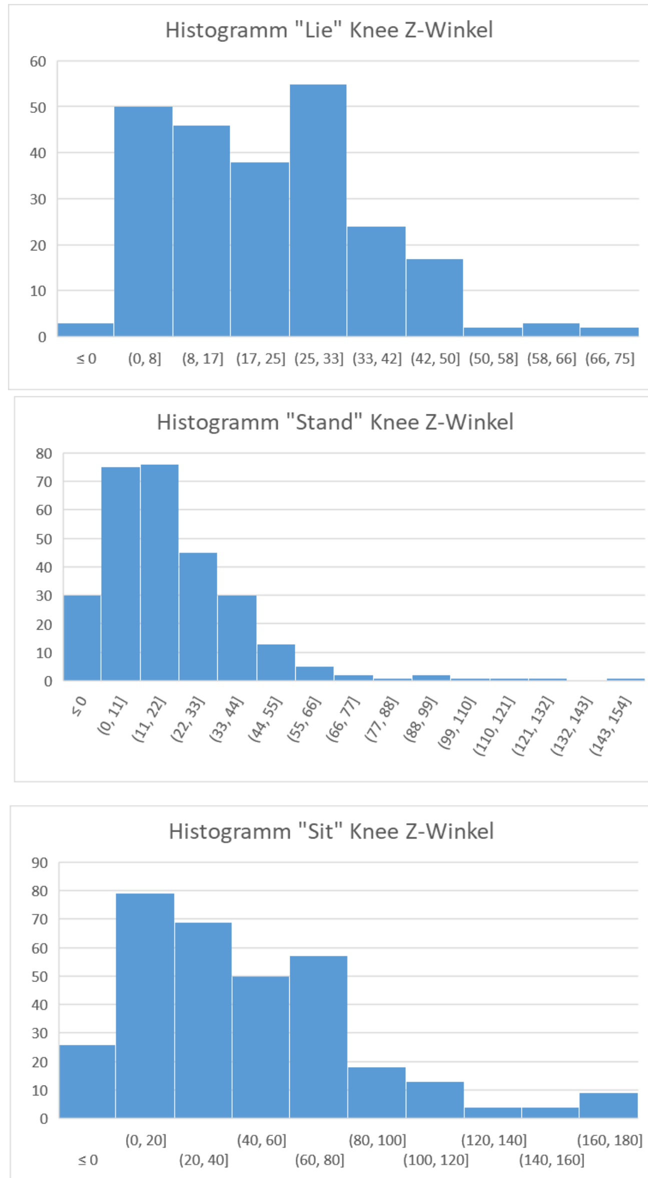

Figure 8: Pose dependent histograms of the calculated hip-knee angle

As result, 27 persons are declared invalid of 852 (of which 2 lying, 8 sitting, 17 standing). The differences of the poses can be explained by the ceiling position of the camera. In Figure 9 some examples of bone evaluation are visualized. Although the person and (rough) pose are valid, some single bones are not matched correctly. In conclusion, further investigations of the detected skeleton could fail or deliver incorrect results. This includes the angle determination for agility evaluation. Otherwise, the proposed method for angle estimation with neutral null method can point out single in correct bones. In Figure 10 some examples of determined angle between hip and knee are calculated and compared to the classification based on the error metric. The observation shows that values often seem plausible and appropriate. The deviations from expected values for non-valid joints are related to invalid bones, but in reverse a statement is currently not possible. In summary, false spatial poses are partially detected, but more false pose data is needed for detailed evaluation.

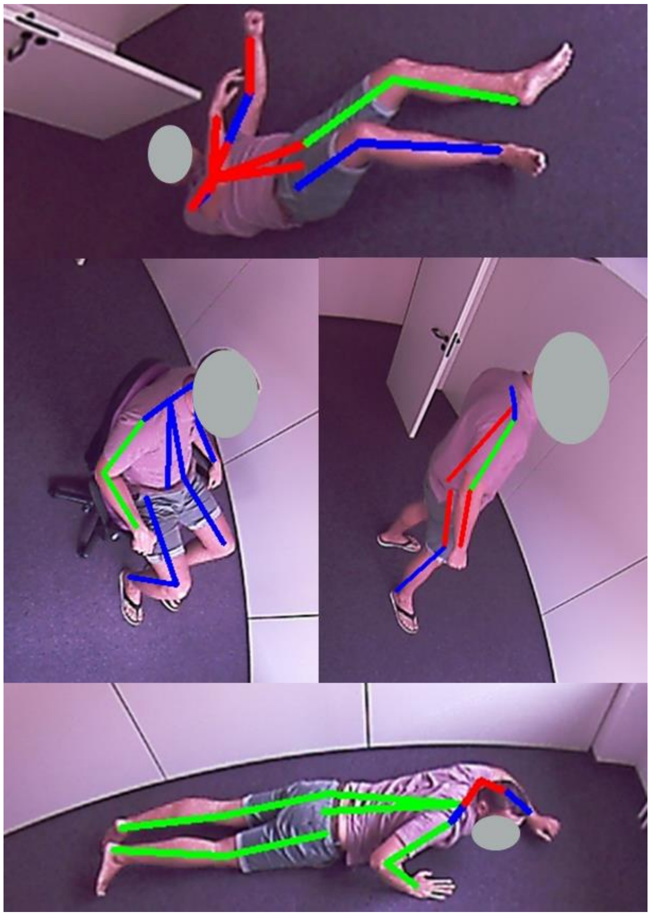

Figure 9: Examples of the visualization of joint classification

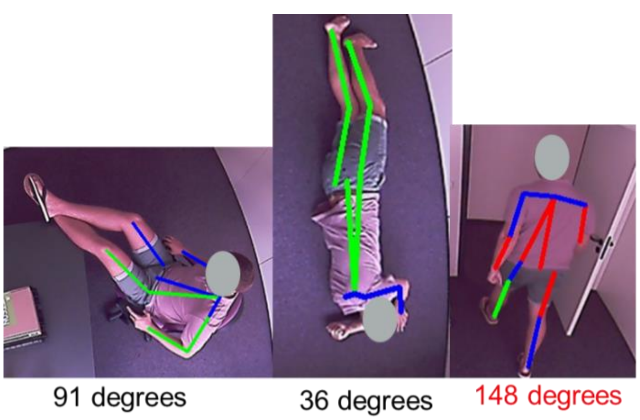

Figure 10: Examples of determined angle between hip and knee are calculated and compared to the classification based on the error metric

Besides the evaluation of the data set of true positives, we also examined some examples of false detection, the ghosts. This data has come from real applications of the sensor system, see Figure 11. From 9 examples 8 were directly classified as invalid and 1 has the calculated error of 0,65. So, all ghosts were classified correctly with the error metric and the determined thresholds, So the proof of concept is done, but more data is necessary.

Figure 11: Image clips in which ghosts were detected

Summary and Outlook

In this work, the evaluation of the transferability of human anatomical/physiological characteristics was presented. The estimated body size out of 3D-Joint lengths leads to best results for person verification and joint detection quality. Solution strategies for the 3 typical problems “ghost” detection, incorrect bones and incurred pose were developed. Empirical evaluations lead to various conclusions, like the torso values deliver best results, but the combination of all joint sizes is relevant and improve the classification quality. The confidence should be over 0.5.
We introduce an error metric for complete body and additionally usable for single joint classification: valid or invalid. The determination of the threshold of 0.3 lead to detection rate of 97% of real persons and marking all ghost examples as invalid.
Our metric for agility and correct pose classification by using neutral-null-method as base shows that first results are plausible, but no final statement of usability was possible.
In summary, a fast and easy implementable method for the verification and regularization of 3D-human body pose estimation has been developed especially for the use in embedded sensor systems.

The methods additionally allow post-labeling with new classification results and retraining to optimize 2D and 3D detection, which includes

→ Reassessment of 3D pose recognition possible
→ Post-labeling and retraining to optimize 2D detection
→ Reassessment of 2D pose detection possible

This could improve the person detection quality directly and offers a wide range of applications, e.g. safety, fall detection, etc. Our approach provides an efficient validation of NN-based human detection, also applicable in embedded systems. Besides the already explained advantage of a quantitative and comparable analysis, the results can be used for relabeling and retraining of underlying 2D and 3D pose estimators by re-rating the detection quality. These optimizations support significantly a reliable 3D recognition of real persons, which is necessary to widen the spectrum of high-standard applications of 3D human-centered technologies.

References

[1] Shotton, J.; Sharp, T.; Kipman, A.; Fitzgibbon, A.; Finocchio, M.; Blake, A. et al. (2013): Real-time human pose recognition in parts from single depth images. In: Commun. ACM 56 (1), S. 116. DOI: 10.1145/2398356.2398381.

[2] Wang, Keze; Lin, Liang; Jiang, Chenhan; Qian, Chen; Wei, Pengxu (2020): 3D Human Pose Machines with Self-Supervised Learning. In: IEEE transactions on pattern analysis and machine intelligence 42 (5), S. 1069–1082. DOI: 10.1109/TPAMI.2019.2892452.

[3] WEI, Shih-En, et al. Convolutional Pose Machines. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. S. 4724-4732.

[4] https://ps.is.tuebingen.mpg.de/pages/4d-capture, (last view 12.1.2021).

[5] F. Bogo and J. Romero and G. Pons-Moll and M. J. Black, „Dynamic FAUST: Registering Human Bodies in Motion“, Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017.

[6] H. Wang, S. Pirk, E. Yumer, V. G. Kim, O. Senner, S. Sridhar, L. J. Guibas, Learning a Generative Model for Multi-Step Human-Object Interactions from Videos, Computer Graphics Forum (Proceedings of Eurographics), 2019.

[7] F. Bogo and J. Romero and M. Loper and M. J. Black, “FAUST: Dataset and evaluation for 3D mesh registration”, Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.

[8] Bullinger-Hoffmann AC, Mühlstedt J. Homo sapiens digitalis-virtuelle Ergonomie und digitale Menschmodelle. Springer; 2016.

[9] Wolf A, Miehling J, Wartzack S. Challenges in interaction modelling with digital human models-A systematic literature review of interaction modelling approaches. Ergonomics 2020: 1–17.

[10] http://openpose.org, (last view 12.1.2021).

[11] Hassan M, Choutas V, Tzionas D, Black MJ. Resolving 3D Human Pose Ambiguities With 3D Scene Constraints. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2019.

[12] https://3dvisionlabs.com/hs-dk1,(last view 12.1.2021).

[13] https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-koerpermasse.html, (last view 12.1.2021).

[14] K.-D. Thomann et al. (Hrsg.): *Orthopädisch-unfallchirurgische Begutachtung - Praxis der klinischen Begutachtung*. Elsevier, Amsterdam, 2008.

[15] Messblatt untere Gliedmaßen (dguv.de), https://www.dguv.de/medien/formtexte/aerzte/f_4224/f4224.pdf, (last view 12.1.2021).

Author Biography

Julia Denecke studied computer science at the University of Stuttgart. Since 2007 she has been a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation in the department of Image Processing and Signal Analysis. 2013 she finished her PhD in the topic of volume data processing. Since 2016 she is the group leader of the scene analysis and focusses on 2D and 3D applications for dynamic detection in scene context.

Norbert Link is co-founder (year 2018) and managing director of Inferics GmbH, Karlsruhe, Germany, which focuses on intelligizing ambient objects via creating and embedding dedicated AI. He is also professor in computer science at the University of Applied Sciences Karlsruhe, teaching machine learning, and was supervisor with the graduate school 1483 at the Karlsruhe Institute of Technology (KIT). His research activities are focused on intelligent systems and on computer vision. Norbert is also dedicated to developing new technologies based on latest research results and implementing them in the market. In the year 2000 he co-founded Vitracom AG where he was haed of research and chairman of the board until 2009. Results of his scientific and technology work can be found in numerous scientific publications and patents.

Christian Jauch has been a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart, Germany since 2015 and studied technical cybernetics at the University of Stuttgart. He works in the Image and Signal Processing department and is part of the Scene Analysis group. His work focuses on industrial applications, more specifically on hand pose estimation and hand gesture recognition in manual assembly scenarios.