

Skeleton tracking for serious games and real-time medical diagnosis

Mohamed Adjel^{*,†,‡}, Antoine Seilles^{*}, Denis Mottet[†]
and Guillaume Tallon^{*}

^{*}Euromov, 700 Avenue du Pic Saint-Loup, 34090 Montpellier, France
denis.mottet@umontpellier.fr

[†]NaturalPad, 700 Avenue du Pic Saint-Loup, 34090 Montpellier, France
antoine@naturalpad.fr, guillaume.tallon@naturalpad.fr

[‡]Polytech Marseille, 163 Avenue de Luminy, 13009 Marseille
mohamed-ala-eddine.adjel@etu.univ-amu.fr

Abstract

Physical rehabilitation of people with reduced mobility implies to monitor the movements of the patients during the rehabilitation sessions, so to individualize the therapy patient by patient. A serious-games company, NaturalPad (NP), would like to develop a cheap real-time markerless skeleton tracking device ensuring diagnosis assistance of neuromuscular and articular pathologies among reduced mobility persons such as elderly, post-stroke and persons affected by disability. In this way, the goal of this device is to precisely assess 3D body joints coordinates in real-time, that will be used to format accurate indicators about articular capacities of the patient during a physiotherapy session. These indicators, such as the Range of Motion (ROM) of each articulation, will be printed on a Graphical User Interface (GUI), so the physiotherapist can monitor the

How to cite this book chapter:

Adjel, M., Seilles, A., Mottet, D. and Tallon, G. 2020. Skeleton tracking for serious games and real-time medical diagnosis. In: Loizides, F., Winckler, M., Chatterjee, U., Abdelnour-Nocera, J. and Parmaxi, A. (eds.) *Human Computer Interaction and Emerging Technologies: Adjunct Proceedings from the INTERACT 2019 Workshops*. Pp. 361–368. Cardiff: Cardiff University Press. DOI: <https://doi.org/10.18573/book3.au>. License: CC-BY 4.0.

evolution of the patients pathologies. After giving details about related studies, we will explicit technological requirements and project constraints. Last we will define a benchmark process of existing skeleton tracking algorithms and cheap motion capture devices. The results will allow us to evaluate if there is an enough accurate camera/algorithm combination to deal with our issues.

Keywords

Real-time skeleton tracking · Medical diagnosis · Joint angles estimation

1 Related works and technological requirements

1.1 Device and algorithm requirements

The Kinect v2¹ and Unity3D² are interesting tools to develop real-time interaction games for physical rehabilitation [1–2]. Actually, a lot of movements are supported by NPs serious games platform, such as steps, chest inclination, hands movement and squatting series. We are able to have precise enough skeleton data to improve functional autonomy among older adults [2–3]. Nevertheless, due to Kinect imprecision, we are unable to correctly recognize head inclinations, ankle/chest rotations and center of mass approximation. Moreover, the Kinect skeleton tracking algorithm doesn't take into account osseous and articular constraints of human body, so it's not precise enough for detailed articular angles analysis [2]. Yet, our device must respect skeletal constraints to be used for joints angles assessment.

Thus, the device needs to fulfill 3 technological constraints: 1) being able to handle RGB and/or depth data for real-time interaction in games, 2) being able to extract accurate enough 3D coordinates of the patients articulations to estimate articulations angles for diagnosis purposes, and 3) being sensor agnostic. The combination sensor/algorithm and hardware configuration have to be as cheap as possible, because of commercial constraints. The final device must be easy to set up, thereby the physiotherapist won't have to reconfigure and calibrate the set up between two sessions.

1.2 Clinical uses of depth cameras

Many studies [5–12] have examined the Kinect for assessment and balance control. They indicate that for relatively slow movements, the Kinect can give enough accurate skeleton data to perform dynamic tests as functional reach, sit

¹ <http://www.arzapstudio.com/kinect-for-windows/>.

² <https://unity.com/>.

to stand and timed up and go. Encouraging results also shown that the Kinect sensor can be useful for medical diagnosis and monitoring of patients suffering from Parkinsons, scoliosis and post-stroke diseases [13–15]. However, as we want to estimate articular capacities, were not sure if Kinect is worthwhile given the inaccuracy associated with some of the variables extracted by the sensor [5]. This is particularly true considering the turning movement, as the Kinect cannot accurately record postural movement when the patient performs the turn. Very recent studies [16–19] already worked on medical diagnosis with depth sensors. In [16], a serious games platform is conceived for home-based rehabilitation after the hospitalisation period, with automated evaluation of the patient during the training. Clinical indicators are extracted, such as neglected body areas during session or errors in limbs trajectory. In [16–18] its shown that depth-sensors can be useful for post-stroke rehabilitation serious games and motor function diseases diagnosis among elderly. The study in [19] also demonstrates very encouraging results in clinical data assessment using Intel RealSense depth cameras.

1.3 Joint angles estimation

As a reminder, the device we want to develop should be able to precisely extract 3D coordinates of skeleton joints in real-time, accurately estimate joint angles during a physiotherapy session, and, in the ideal case, during a serious-game session. Several studies [20–23] used markerless motion capture systems to estimate joint angles and compared these values with ground truth to estimate the accuracy of the depth sensor for such task. Studies [20–22] assessed the joint angles estimation accuracy of the Kinect for clinical uses. Marker based motion capture systems were used in [20–21] as ground truth, while [22] used an IMU device. [20] concluded that the Kinect system is not yet suitable for clinical assessments while [21] concluded the opposite. This contrast is explained by the fact that [20] uses a VICON system as ground truth, yet, studies [24–25] shown that there can be interferences between VICON and Kinect that slant the joints coordinates assessment of the Kinect [21] used a jig as guinea pig, instead of real humans, which can distort the results [22] demonstrates that the Kinect is efficient in knee joint angle estimation, which is not sufficient as we want accuracy on all body joints. As far as we know, there is no markerless device that aims accurately estimating joint angles in real-time to deduce articular capacities of the patient in the context of medical diagnosis support.

1.4 Real-time human pose estimation

Several papers [26–31, 43] tackled the 3D real-time human pose estimation issue. Even if [27] estimates only 2D joint coordinates, we will keep it for our benchmark, for several reasons: 1) It works in real-time on cheap hardware

2) We can deduce 3D coordinates from 2D [32] 3) It uses a monocular RGB camera, which is cheap 4) We want to verify if its accurate for joint angles estimation. We already eliminated [26, 28–31] because of License requirements (cost or lack of documentation). Implementations remaining for the benchmark are [27], Nuitrack, Kinect SDK v2 and Orbbec SDK. First of all, we aim to determine the level of accuracy we can obtain with state-of-the-art real-time skeleton tracking algorithms and a single sensor. We will benchmark different combinations of algorithms/sensors to determine which couple is the best, using the constraints mentioned in 2)a) as criteria to select the best combination camera/algorithm. Then, physiotherapists will assess the clinical relevance of the best combination in selected use cases, for diagnosis and physical rehabilitation.

2 Benchmarking process

For this benchmark, we will test the skeleton tracking algorithms mentioned above and following markerless motion capture devices: Microsoft Kinect v2, Orbbec Astra, Intel RealSense D435i, Regular webcam (for [27]). We will assess the error of pose estimation for each sensor/algorithm combination. To calculate the estimation error for each device, we will use a state of the art motion capture system (VICON, Oxford metrics)³ as reference system. We will record the movements with the two systems, and will compare the 3D coordinates values given by the VICON with 3D coordinates values given by the tested device/algorithm combination. As mentioned above, the VICON system infrared waves can interfere with markerless sensors. A protocol is defined in [24] to minimize this noise, so we will reduce the number of markers and the distance between Kinect and the volunteer. Then, we will implement the following steps: 1) Collecting skeleton data from Vicon and Tested Camera 2) Synchronize the data in Time, as Vicon and cameras have different frequencies 3) Compute each joints angles thanks to cosines law and calculate the angles assessment error of each camera/algorithm combination 4) Selection of the optimal combination. The last step will consist in comparing the joints angle assessment accuracy of the optimal solution with the accuracy needed by physiotherapists.

3 Further researches

Actually Convolutional Neural Networks (CNNs) are widely used for monocular 3D human pose estimation and show the most accurate results [46]. Nevertheless, both top-down and bottom-up approaches don't take into account the human skeleton constraints. Our approach will consist in using

³ <http://www.vicon.com>.

Denavit-Hartenberg (D-H) convention to model geometric and kinematic skeleton [45]. Coupled with CNN algorithm we theoretically will be able to have an accurate and constrained skeleton [43] shows we can enhance joints coordinates estimation with integrating skeleton constraints during the training process. Moreover, without going into the algorithm technicalities, we can add human skeleton constraints in two other ways:

- Refine identified human silhouettes (works for top-down approaches) with D-H before recognizing joints on the silhouette
- Refine the pose with D-H after joints coordinates assessment (works for both bottom-up and top-down approaches)

We think both of these refinement steps would enhance accuracy and realism of extracted skeleton coordinates. However, this will have an impact on computational speed that we will have to keep reasonable. Depending on obtained results, we will remove some cost constraints to compute a heavy algorithm on costly hardware device.

References

1. Bakhti K, Mottet D, Mélia G, Laffont I, *Low cost objective diagnosis of learned nonuse of the paretic arm after a stroke using Kinect technology*. Annals of Physical and Rehabilitation Medicine, 58, e11–e12. (2015)
2. Tallon G, Seilles A, Mélia G, Andary S, Di Loreto I, Blain H, *Intérêt du serious game Hammer & Planks pour améliorer l'autonomie fonctionnelle chez la personne âgée institutionnalisée*. ICEPS, 19–21 mars, Montpellier (2015)
3. Tallon G, Seilles A, Mélia G, Andary S, Bernard P, Di Loreto I, Blain H. *Effects of the serious game Medimooov on the functional autonomy of institutionalized older adults*. Annals of Physical and Rehabilitation Medicine – Proceedings of the 30th Annual Congress of the French Society of Physical and Rehabilitation Medicine 58, e113e114.(2015)
4. K. Bakhti, I. Laffont, M. Muthalib, J. Froger and D. Mottet, *Kinect-based assessment of proximal arm non-use after a stroke*, (2018)
5. Ross A Clark, F. Mentiplay, E. Hough, Y. Pua, *Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives*, 2018
6. A. Clark, Y. Pua, K. Fortin, C. Ritchie, E. Webster, L. Denehy, L. Bryant, *Validity of the Microsoft Kinect for assessment of postural control*, (2012)
7. B. Dehbandi, A. Barachant, H. Smeragliuolo, J. Long, S. Bumanlag, V. He, A. Lampe, D. Putrino, *Using data from the Microsoft Kinect 2 to determine postural stability in healthy subjects: A feasibility trial*, (2017)
8. M. Eltoukhy, C. Kuenze, J. Oh, F. Signorile, *Validation of Static and Dynamic Balance Assessment using Microsoft Kinect for Young and Elderly Populations*, (2018)

9. S. Galen, V. Pardo, D. Wyatt, A. Diamond, V. Brodith, A. Pavlov, *Validity of an Interactive Functional Reach Test*, (2015)
10. T. Ramei, Y. Orito, H. Funaya, K. Ikeda, Y. Okada, T. Shibata, *Kinect-based posturography for in-home rehabilitation of balance disorders*, (2015)
11. Y. Yang, F. Pu, Y. Li, S. Li, Y. Fan, D. Li, *Reliability and Validity of Kinect RGB-D Sensor for Assessing Standing Balance*, (2014)
12. A. Clark, Y. Pua, C. Oliveira, J. Bower, T. Rebekah, McGaw, K. Hasanki, F. Mentiplay, *Reliability and concurrent validity of the Microsoft Kinect V2 for assessment of standing balance and postural control*, (2015)
13. R. Torres, M. Huerta, R. Clotet, R. Gonzlez, L.E. Snchez, D. Rivas and M. Erazo, *A Kinect Based Approach to Assist in the Diagnosis and Quantification of Parkinsons Disease*, (2016)
14. Vincent Bonnet, Takazumi Yamaguchi, Arnaud Dupeyron, Sebastien Andary, Antoine Seilles, Philippe Fraisse, *Automatic Estimate of Back Anatomical Landmarks and 3D spine curve From a Kinect Sensor*, (2016)
15. Stephanie Vernon, Kade Paterson, Kelly Bower, Jennifer McGinley, Kimberly Miller, Yong-Hao Pua, and Ross A. Clark, *Quantifying individual components of the timed up and go using the kinect in people living with stroke*, (2014)
16. Matteo Morando, Serena Ponte, Elisa Ferrara and Silvana Dellepiane, *Definition of Motion and Biophysical Indicators for Home-Based Rehabilitation through Serious Games*, (2018)
17. Robert J. DaweI, Lei Yu, Sue E. Leurgans, Timothy Truty, Thomas Curran, Jeffrey M. Hausdorff, Markus A. Wimmer, Joel A. Block, David A. Bennett, Aron S. Buchman, *Expanding instrumented gait testing in the community setting: A portable, depth-sensing camera captures joint motion in older adults*, (2019)
18. Mengxuan Ma, Rachel Proffitt, Marjorie Skubic, *Validation of a Kinect V2 based rehabilitation game*, (2018)
19. Francesco Luke Siena, Bill Byrom, Paul Watts & Philip Breedon, *Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research*, (2018)
20. Shayan Bahadori, Philip Davenport, Tikki Immins & Thomas W. Wainwright, *Validation of joint angle measurements: comparison of a novel low-cost marker-less system with an industry standard marker-based system*, (2019)
21. Anne Schmitz, Mao Ye, Robert Shapiro, Ruigang Yang, Brian Noehren, *Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system*, (2014)
22. Nazeeh Alothmany, Afzal Khan, Majdi Alnowaimi, Ali H. Morfeq, Ehab A Hafez, *Accuracy of joint angles tracking using markerless motion system*, (2014)
23. Franziska Schlagenhaut, Siddarth Sreeram and William Singhose, *Comparison of Kinect and Vicon Motion Capture of Upper-Body Joint Angle Tracking*, (2018)

24. Mreza Naeemabadi, Birthe Dinesen, Ole Kseler Andersen and John Hansen, *Influence of a marker-based motion capture system on the performance of Microsoft Kinect v2 skeleton algorithm*, (2018)
25. MReza Naeemabadi, Birthe Dinesen, Ole Kseler Andersen, John Hansen, *Investigating the impact of a motion capture system on Microsoft Kinect v2 recordings: A caution for using the technologies together*, (2019)
26. Zhe Cao Tomas Simon Shih-En Wei Yaser Sheikh, *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, The Robotics Institute, Carnegie Mellon University, (2017)
27. Daniil Osokin, *Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose*, (2018)
28. Muhammed Kocabas, M. Salih Karagoz, Emre Akbas, *MultiPoseNet: Fast Multi-Person Pose Estimation using Pose Residual Network*, (2018)
29. Dushyant Mehta, Srinath Sridhar, Oleksandr Sotnychenko, Helge Rhodin, Mohammad Shafiei, Hans-Peter Seidel, Weipeng Xu, Dan Casas, Christian Theobalt, *VNect: Real-time 3D Human Pose Estimation with a Single RGB camera*, (2017)
30. Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, Cewu Lu, *RMPE: Regional Multi-Person Pose Estimation*, (2018)
31. Riza Alp Guler, Natalia Neverova, Iasonas Kokkinos, *DensePose: Dense Human Pose Estimation In The Wild*, (2018)
32. Sandika Biswas, Sanjana Sinha, Kavya Gupta and Brojeshwar Bhowmick, *Lifting 2d Human Pose to 3d: A Weakly Supervised Approach*, (2019)
33. Kwok-Yun Yeung, Tsz-Ho Kwok, and Charlie C. L. Wang, *Improved Skeleton Tracking by Duplex Kinects: A Practical Approach for Real-Time Applications*, (2013)
34. Marco Carraro, Matteo Munaro, Jeff Burke and Emanuele Menegatti, *Real-time marker-less multi-person 3D pose estimation in RGB-Depth camera networks*, (2017)
35. Sungphill Moon, Youngbin Park, Dong Wook Ko, Il Hong Suh, *Multiple Kinect Sensor fusion for human skeleton tracking using Kalman Filtering*, (2017)
36. Ziren Wang, Guoliang Liu, Guohui Tian, *Human Skeleton Tracking Using Information Weighted Consensus Filter in Distributed Camera Networks*, (2017)
37. Guoliang Liu, Member, IEEE, Guohui Tian, Junwei Li, Xianglai Zhu, and Ziren Wang, *Human Action Recognition Using a Distributed RGB-Depth Camera Network*, (2018)
38. Sungjin Hong and Yejin Kim, *Dynamic Pose Estimation Using Multiple RGB-D Cameras*, (2018)
39. Madhura P. Pathegama, Dileepa M. Marasinghey, Kanishka Wijayasekaraz, Ishan Karunanayakex, Chamira U. S. Edussooriya, Pujitha Silva k, and Ranga Rodrigo, *Moving Kinect-Based Gait Analysis with Increased Range*, (2018)

40. Filippo Basso, Riccardo Levorato and Emanuele Menegatti, *Online Calibration for Networks of Cameras and Depth Sensors*, (2014)
41. Semih Dinc, Farbod Fahimi, and Ramazan Aygun, *Mirage: an $O(n)$ time analytical solution to 3D camera pose estimation with multi-camera support*, (2017)
42. A. Elhayek, C. Stoll, N. Hasler, K. I. Kim, H.-P. Seidel, C. Theobalt, *Spatio-temporal Motion Tracking with Unsynchronized Cameras*, (2012)
43. David Pavllo, David Grangier, Michael Auli, *QuaterNet: A Quaternion-based Recurrent Model for Human Motion*, (2019)
44. William Dyce, Nancy Rodriguez, Antoine Seilles, Benoit Lange, Sebastien Andary, *Tabu search for human pose recognition*, (2014)
45. Cheng Xu, Jie He, Xiaotong Zhang, Cui Yao, Po-Hsuan Tseng, *Geometrical kinematic modeling on human motion using method of multi-sensor fusion*, (2018)
46. Xiao Sun, Chuankang Li, and Stephen Lin, *An Integral Pose Regression System for the ECCV2018 PoseTrack Challenge*, (2018)