International Journal of Electrical and Electronic Engineering & Telecommunications Vol. 7, No. 1, January 2018

Heuristic Method for Calculating the Translation of Human Body Recordings Using Data from an Inertial Motion Capture Costume

Tomasz Hachaj[1] and Marek R. Ogiela[1,2]
[1] Pedagogical University of Krakow, Institute of Computer Science and Computer Methods, Krakow, Poland.
[2] AGH University of Science and Technology, Cryptography and Cognitive Informatics, Research Group, Krakow, Poland.
Email: tomekhachaj@o2.pl; mogiela@agh.edu.pl; marek.ogiela@up.krakow.pl

***Abstract*—The heuristic method presented in this paper is capable of calculating the translation of human body recordings carried out using an inertial motion capture (MoCap) costume. As an input, the method takes only the movement described with the hierarchical kinematic model and acceleration data but without a translation of the root joint. We have compared the results obtained through our method with those generated using commercial software. In the test dataset, we used eight MoCap recordings of karate katas performed by two black belt karate masters. Based on our experiment, the proposed heuristic method enables the production of valuable data that closely corresponds to the direction of movements of the person wearing a costume. The method can be used, for example, to generate the correct visualization of various movements relative to a global reference frame.**

***Index Terms*—Motion capture, inertial sensors, translation calculation, karate**

I. Introduction

Motion capture (MoCap) is a technology that enables precise three-dimensional (3D) high-frequency real-time measurement of human movements. There are two basic types of MoCap: the first are vision-based systems that directly track the coordinates of markers placed over the observed person's body. This system uses high-frequency cameras that have to be calibrated in a dedicated laboratory. The second type of MoCap technology uses inertial sensors that are attached to a costume. The sensors are often composed of 3-axis accelerometers, gyroscopes, and magnetometers. A set of sensors indirectly measures the body kinematics, for example, by a double integration of the acceleration data [1]. Although vision systems are more precise than inertial systems, modern inertial costumes provide an unobtrusive and affordable way to generate high-quality recordings [2]. Inertial MoCap does not need external cameras, emitters, or markers and can be used outdoors as well as indoors; there are no restrictions for lighting and it does not suffer from occlusion or missing markers [3], [4]. Many methods have been proposed to increase the accuracy of inertial technology. Researchers have used statistical body models that include anthropometric constraints [5], [6], artificial neural networks and nearest neighbor search [7], and a hybrid tracker that combines videos with a small number of inertial units [8]. To calculate the position of the entire body in a global coordinates frame, the sensor fusion scheme can be seamlessly integrated with various types of aiding sensors, such as Global Positioning System, RF-based local positioning sensors, a barometer, a camera, and pressure and/or force sensors [3], [9] or a laser scanner [10].

The aim of this paper is to propose a novel method that enables the calculation of body position in a global frame. The heuristic method presented in this paper is capable of calculating the translation of the human body recording carried out using an inertial motion capture costume. As an input, the method takes only the movement described by the hierarchical kinematic model and acceleration data. We have compared the results obtained through our method with those generated using commercial software. As the test dataset, we used eight MoCap recordings of karate katas performed by two black belt karate masters.

II. Material and Methods

To gather MoCap data, we have used the Shadow 2.0 system. This comprises 17 inertial measurement units, each of which contains a 3-axis accelerometer, gyroscope, and magnetometer. The tracking frequency of the device was set to 100 Hz with 0.5 degree static and 2 degrees dynamic accuracy. The software provided by the hardware manufacturer enables one to calculate the kinematics of the tracked body and to present it with a hierarchical kinematic model (we present this model in Fig. 1).

The Shadow software also allows one to calculate the movement relative to a global reference frame. The frame coordinate calculation method is similar to other internal solutions [11] in that it acquires information regarding the location of the experiment and estimates the magnetic field of the Earth. The rest of that algorithm works as follows:

- It uses the sensor rotations to compute the skeleton pose
- Next, it uses the skeleton pose along with measurement data (pressure data gathered with insoles and inertial data) to compute the contact points

Manuscript received October 2, 2017; revised December 10, 2017.

doi: 10.18178/ijeetc.7.1.26-29

- The contact points and a dynamic simulation are used to calculate the translation/movement of the body.

However, this method has some major drawbacks. It is possible that the translation vector generated by it has an opposite direction than the correct aim of the movement. This means that the tracked body moves backward instead of forward. Moreover, while a person is rapidly bending a knee, the original algorithm may interpret it as climbing down due to the ground level not being preserved during the whole recording. A similar situation is also present while jumping—the body very rarely returns to the same ground level.

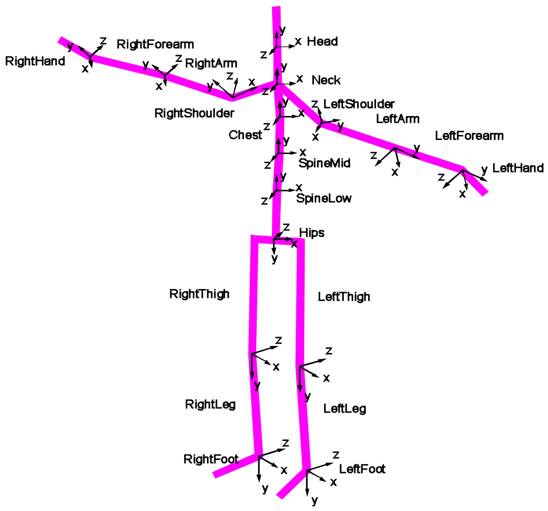

Figure 1. The hierarchical kinematic model; the root of the model is the hip joint.

In this paper, we propose a novel, heuristic approach that allows some of the above limitations to be overcome under certain constraints:

- As input data, we take the body kinematics in the form of a hierarchical kinematic model (however, without the translation data of the root joint) and acceleration measurements of the feet joints.
- The algorithms work under the assumption that during body translation, at least one foot resides on the floor (i.e., a tracked person does not jump).
- Our method assumes that there is no sliding motion of the foot that is on the ground and the movement occurs on a fixed horizontal level (i.e., on a flat surface).
- Either the magnitude of the acceleration vector of the foot that is moving is higher than the magnitude of the acceleration of the other foot or the foot that is moving has a higher vertical coordinate than the other one.

Our algorithm can be formulated as follows:

Let us calculate the magnitude of acceleration of the left $\left(\overrightarrow{a_{lf_i}}\right)$ and the right $\left(\overrightarrow{a_{rf_i}}\right)$ foot (n is number of samples):

$$\text{Lf} = \left(\left\|\overrightarrow{a_{lf_i}}\right\|, i = 1, \ldots, n\right) \tag{1}$$

$$\text{Rf} = \left(\left\|\overrightarrow{a_{rf_i}}\right\|, i = 1, \ldots, n\right) \tag{2}$$

Let us also perform a Gaussian low pass filter to smooth the signals:

$$\text{Lf}_{\text{smooth}} = G \otimes \text{Lf} \tag{3}$$

$$\text{Rf}_{\text{smooth}} = G \otimes \text{Rf} \tag{4}$$

where $\otimes$ is a convolution operator and G is a Gaussian kernel.

Next, we loop through the samples of all the recordings and check if the magnitude of the acceleration of the left (right) foot in the sample a and $a + 1$ is greater than that of the right (left) foot. If so, we change the coordinates of all the body joints in a way that the left (right) foot did not change its position. This results in the translation of the whole body in a particular direction (see Pseudocode 1.1).

Pseudocode 1.1: Calculation of the body translation using the acceleration data.

```
for a in 1 to n−1
  if (Rf_smooth[a] > Lf_smooth[a]
  and Rf_smooth[a + 1] > Lf_smooth[a + 1])
    diff := Lf_smooth[a + 1]−Lf_smooth[a]
    #subtract diff from all MoCap signals from a+1 to nth sample
  all[ a+1:n] := all[a+1:n]−diff
  else if (Rf_smooth[a] < Lf_smooth[a]
  and Rf_smooth[a + 1] < Lf_smooth[a + 1])
    diff := Rf_smooth[a + 1] - Rf_smooth[a]
  all[ a+1:n] := all[ a+1:n]−diff
  end if
end for
```

However, it is possible that the heuristic in Pseudocode 1.1 will not give satisfactory results, for example, one foot has a higher magnitude of acceleration than the other due to a rotation movement but it remains in the same place. This happens, for example, when a person is rotating on their toes. In this situation, the foot that remains on the ground should have a lower vertical position than the second one. This can be modeled by following heuristic:

Let us calculate the vertical (y-axis) position of the left $\left(\overrightarrow{a_{lf_i}}[y]\right)$ and right $\left(\overrightarrow{a_{rf_i}}[y]\right)$ foot:

$$\text{Lfy} = \left(\overrightarrow{a_{lf_i}}[y], i = 1, \ldots, n\right) \tag{5}$$

$$\text{Rfy} = \left(\overrightarrow{a_{rf_i}}[y], i = 1, \ldots, n\right) \tag{6}$$

Let us also perform a Gaussian low pass filter to smooth the signals:

$$\text{Lfy}_{\text{smooth}} = G \otimes \text{Lfy} \tag{7}$$

$$\text{Rfy}_{\text{smooth}} = G \otimes \text{Rfy} \tag{8}$$

Further, we loop through the samples of all the recording and check if the difference in the vertical position between left and right foot is above a certain threshold, *eps*. If so, we change the coordinates of all the body joints in a way that the left (right) foot does not change its position (this approach is similar to the one in Pseudocode 1.1). This results in the translation of the whole body in a certain direction (see Pseudocode 1.2).

The role of the last step in our method is to vertically translate the right and/or left foot to remain on the ground while not moving. This is done using the procedure presented in Pseudocode 1.3.

Pseudocode 1.2: Calculation of the body translation using the relative vertical position of the left and the right foot.

```
for a in 1 to n-1
  if ||Lfy_smooth[a] − Rfy_smooth[a]|| > eps
    if (Lfy_smooth[a] > Rfy_smooth[a]
    and (Lfy_smooth[a + 1] > Rfy_smooth[a + 1])
      #subtract diff from all MoCap signals from a+1 to nth sample
      diff := Lfy_smooth[a + 1]−Lfy_smooth[a]
      all[a+1:n] = all[a+1:n]−diff
    else if (Lfy_smooth[a]<Rfy_smooth[a]
    and (Lfy_smooth[a + 1] < Rfy_smooth[a + 1])
      diff := Rfy_smooth[a + 1] − Rfy_smooth[a]
      all[a+1:n] = all[a+1:n]−diff2
    end if
  end if
end for
```

Pseudocode 1.3: Translation of the feet to remain on the ground (the ground level has vertical coordinates held in variable *groundPosition*).

```
for a in 1 to n-1
    if (Rfy_smooth[a] ≥ Lfy_smooth[a])
      diff := Lfy_smooth[a]- groundPosition
      #subtract diff from all MoCap signals
      all[a] := all[a] −diff
    else
      diff := Rfy_smooth[a]- groundPosition
      all[a] := all[a] −diff
    end if
end for
```

III. RESULTS

The movements that satisfy the constraints of our algorithm are, for example, selected karate techniques (such as some karate katas). Karate kata is a set of forms (techniques) that are ordered in sequence that demonstrate the application of various defensive and attacking movements. Because katas are standardized, it is easy to follow their MoCap and judge if the translations are calculated in the correct directions. As the dataset for our experiment, we have used MoCap recordings of two black belt karate masters from two schools. The Shorin–ryu master (adult man) performed katas Fukyugata Ichi, Fukyugata Ni, Pinian Shodan, and Pinian Nidan. The Shotokan master (adult woman) performed Heian Shodan, Heian Nidan, Heian Sandan, and Heian Yondan. We have recorded those forms with the MoCap system presented in Section 2. We have compared the processing results of our algorithm with the commercial Shadow algorithm. Using our method, we have obtained the correct translation of the tracked body. The commercial algorithm introduced many errors that were absent in our approach. Because this paper has a limited number of pages, we will present a detailed evaluation only of Heian Shodan kata. The source code of our algorithm and the input data can be downloaded from [12]. The 3D renderings of our method can be viewed online [13].

In Fig. 2, Fig. 3, and Fig. 4, we present the MoCap data before and after processing at various stages of our heuristic. Fig. 5 compares the trajectories of the hip joints calculated using the original Shadow algorithm with our heuristic.

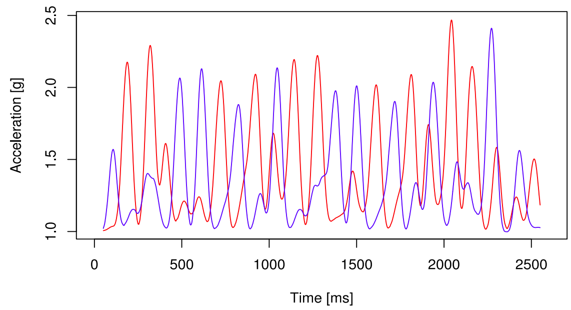

Figure 2. The magnitude of the smoothed acceleration vectors of the right foot (red signal) and left foot (blue signal).

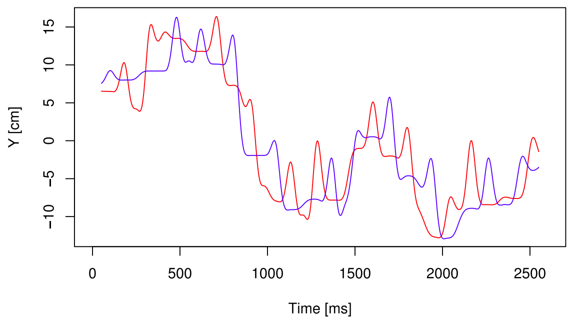

Figure 3. The value of the Y (vertical) smoothed coordinates of the right foot (red signal) and the left foot (blue signal) after applying the original Shadow algorithm to the input data.

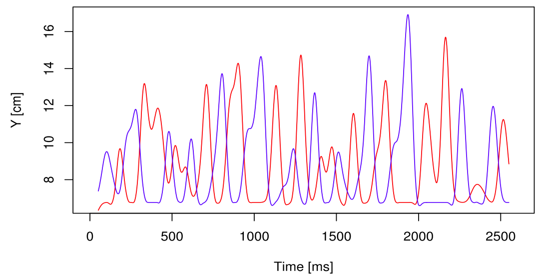

Figure 4. The value of the Y (vertical) smoothed coordinate of the right foot (red signal) and the left foot (blue signal) after applying Pseudocode 1.3 to the original data.

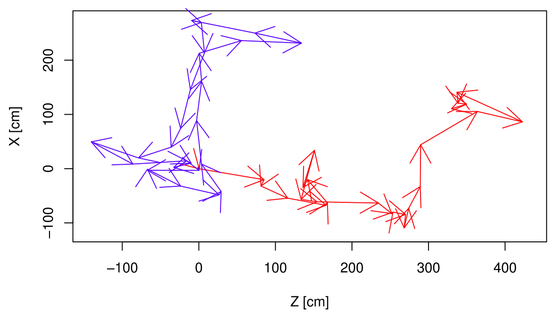

Figure 5. Translation of hip joints while performing Heian Shodan; the red arrows represent the translations calculated by the original Shadow algorithm while the blue arrows show the translations calculated using our approach.

IV. Discussion

As can be seen in Fig. 2, the assumption that a foot that moves has a higher acceleration than the other one is nearly always correct. However, it is possible that this assumption is false. In our dataset, this happened in a part of the Hein Yondan kata (we did not present the visualization of this situation in this paper); the right foot was rotated around the y-axis without changing its space coordinates while the left foot was slightly moving forward. This situation has been corrected by the algorithm presented in Pseudocode 1.2. As can be seen in Fig. 5, the trajectories obtained using our method may significantly differ from those calculated by the original software. Although the fact that the constraints of our algorithm seems to be bold (especially due to the assumption of the lack of sliding motion), the results obtained using our algorithms correspond to the correct directions of the karate movements.

V. Conclusions

On the basis of our experiment, we can conclude that the proposed heuristic method for calculating the translation of human body recordings using data from an inertial MoCap costume enables one to produce valuable results that correspond to the direction of movements of a person wearing the costume. At this moment, the method can be used, for example, to generate the correct visualization of various movements relative to a global reference frame [14]. In our future research, we would like to solve the problem of correctly modeling jumping and moving up and down (without a fixed ground level).

Acknowledgment

This work has been supported by the National Science Centre, Poland, under project number 2015/17/D/ST6/04051. We kindly acknowledge the support of this study by a Pedagogical University of Krakow Statutory Research Grant.

References

[1] P. Neto, P. J. Norberto, and A. P. Moreira, "3-D Position estimation from inertial sensing: Minimizing the error from the process of double integration of accelerations," in *Proc. Industrial Electronics Society, IECON 2013 - 39th Annual Conf. of the IEEE*, 2013.

[2] A. Baak, T. Helten, M. Müller, G. Pons-Moll, B. Rosenhahn, and H. P. Seidel, "Analyzing and evaluating markerless motion tracking using inertial sensors," *Lecture Notes in Computer Science*, vol. 6553, 2012.

[3] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors," XSENS Technologies - Version April 3, 2013.

[4] S. Gupta and R. Mandal, "A survey on image enhancement techniques," *Int. Journal of Electrical and Electronic Engineering & Telecommunications*, vol. 4, no. 1, pp. 47-54, January 2015.

[5] A. Baak, T. Helten, M. Müller, G. Pons-Moll, B. Rosenhahn, and H. P. Seidel, "Analyzing and evaluating markerless motion tracking using inertial sensors," in *Proc. 11th European Conference on Trends and Topics in Computer Vision*, 2010, pp. 139-152.

[6] B. Taetz, G. Bleser, and M. Miezal, "Towards self-calibrating inertial body motion capture," in *Proc. 19th Int. Conf. on Information Fusion*, Heidelberg, Germany - July 5-8, 2016.

[7] F. J. Wouda, M. Giuberti, G. Bellusci, and P. H. Veltink, "Estimation of full-body poses using only five inertial sensors: An eager or lazy learning approach?" *Sensors*, vol. 16, 2016.

[8] G. Pons-Moll, A. Baak, T. Helten, M. Müller, H. P. Seidel, and B. Rosenhahn, "Multisensor-Fusion for 3D full-body human motion capture," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 663-670.

[9] A. Dwivedi and P. Saxena, "Android phone security using wi-fi positioning system," *Int. Journal of Electrical and Electronic Engineering & Telecommunications*, vol. 4, no. 1, pp. 1-6, 2015.

[10] J. Ziegler, K. Kretzschmar, C. Stachniss, G. Grisetti, and W. Burgard, "Accurate human motion capture in large areas by combining IMU- and laser-based people tracking," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2011.

[11] K. Lebel, P. Boissy, M. Hamel, and C. Duval, "Inertial measures of motion for clinical biomechanics: Comparative assessment of accuracy under controlled conditions - effect of velocity," *PLoS One*., vol. 8, no. 11, p. e79945, 2013.

[12] The GitHub repository with source code of our method. [Online]. Available: https://github.com/browarsoftware/ShadowMoCap

[13] Website of GDL project and motion capture repository. [Online]. Available: http://gdl.org.pl/

[14] T. Hachaj and M. R. Ogiela, "The adaptation of GDL motion recognition system to sport and rehabilitation techniques analysis," *J. Med. Syst*., vol. 2016, no. 40, p. 137, 2016.

Dr. Tomasz Hachaj received his MSc degree in Computer Science from the Tadeusz Kosciuszko Krakow University of Technology, Poland; Ph.D. degree in Computer Science from AGH University of Science and Technology, Poland; and D.Sc. in Computer Science from Wroclaw University of Science and Technology, Poland, in 2006, 2010, and 2017, respectively. He is a Deputy Head of the Institute of Computer Science at the Pedagogical University in Krakow, Poland. He has authored more than 60 papers on pattern recognition, signal processing, and image analysis. He specializes in human actions recognition, social networks analysis, and medical image processing and visualization.

Prof. Dr. Marek R. Ogiela Professor of Computer Science; cognitive scientist and cryptographer; and head of Cryptography and Cognitive Informatics Laboratory. Professor Marek R. Ogiela works at AGH University of Science and Technology and Pedagogical University in Krakow. In 1992, he graduated from the Mathematics and Physics Department at Jagiellonian University. In 1996, for his honors doctoral thesis on the syntactic methods of analysis and image recognition, he was awarded the title of Doctor of Control Engineering and Robotics at the Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering of AGH University of Science and Technology.

In 2001, he was awarded the title of Doctor Habilitated in Computer Science for his research on medical image automatic analysis and understanding.

In 2005, he received the title of a Professor in technical sciences.

He is a member of numerous world scientific associations (SPIE-Fellow Member, IEEE-Senior Member, SIIM etc.) as well as the Forecast Committee "Poland 2000 Plus" of Polish Academy of Science and member of Interdisciplinary Scientific Committee of the Polish Academy of Arts and Sciences (Biocybernetics and Biomedical Engineering Section in years 2003–2011). He has authored more than 340 scientific international publications on pattern recognition and image understanding, artificial intelligence, information technology (IT) systems, and biocybernetics. He is also the author of recognized monographs in the field of cryptography and IT techniques and the author of an innovative approach to cognitive medical image analysis and linguistic threshold schemes. For his achievements in these fields, he has been awarded many prestigious scientific honors.