Precision of Gait Indices Approximation by Kinect Based Motion Acquisition

Agnieszka Michalczuk[1], Damian Pęszor[1], Henryk Josiński[2], Adam Świtoński[2], Romualda Mucha[3], and Konrad Wojciechowski[1,2]

[1] Polish-Japanese Academy of Information Technology,
Aleja Legionów 2, 41-902 Bytom, Poland
{amichalczuk,dpeszor,kwojciechowski}@pjwstk.edu.pl
[2] Silesian University of Technology, Institute of Informatics,
Akademicka 16, 44-100 Gliwice, Poland
{Henryk.Josinski,Adam.Switonski,Konrad.Wojciechowski}@polsl.pl
[3] Medical University of Silesia, Batorego 15, 41-902 Bytom, Poland
romam28@wp.pl

Abstract. Step length constitutes one of the important gait indices. The research described in the present paper was focused on the method of determination of the step length by means of the Kinect device. Gait sequences recorded in the Human Motion Laboratory of the Polish-Japanese Academy of Information Technology using the Vicon system played the role of the reference data. The group of six subjects participated in the experiments. Conclusions from the comparative analysis of the results of both approaches summarize the paper.

Keywords: step indices, motion capture, Kinect, gait.

1 Introduction

The gait is one of the fundamental elements of human motor skills. It depends on age, gender, coordination predisposition and construction of the human body. Exercise of motor skills of the human body leads to increase mobility, which is fundamen-tal for athletes and for people with the musculoskeletal system disorders. Mobility improvement is often associated with supervision specialist who controls the correct course of exercises. Recording enables data analysis of their correctness and error correction. One of the most effective data recording method is motion capture technique. It is based on the cameras working in the near infrared light. Cameras trace the cloud of points defining the position of an actor in the 3D space by recording the light reflected from the markers. Markers are placed in the anthropometric points on the human body. Cameras operate in high frequencies (e.g. 100 Hz) and in high resolution. In order to correctly identify a single marker it needs to be visible by two cameras at least. Thus, system efficiency depends on the number of cameras. In addition, system is designed for permanent installation in the premises, since its portability is troublesome.

Another most popular method of digital recording is the use of video cameras. These cameras are relatively inexpensive and their sizes allow recording video in almost any place. Using a single camera limits the visibility of the subject because of occlusions. Another limitation is the noise in the recorded image. During recording an interesting object, the camera also records its neighborhood. This makes it difficult to extract the object from the background, especially when they are recorded in a similar color scheme.

Combining both methods of digital recording: motion capture system and video cameras constitutes motion sensor for Xbox 360 produced by Microsoft named Kinect. It comprises an emitter and a camera recording the reflected light to determine the distance of the objects in the camera range. Motion sensor is also equipped with a standard RGB camera, which allows easy acquisition of the scene in colors scheme proper for human. Although Kinect controller is quite handy device it is in fact stationary device and each position change requires new calibration.

Analysis of mobility of the human body is dependent on the method of registering the object. The technique of motion capture provides information about the location of the markers in 3D space and joint angles in the sagittal, frontal and transverse planes, among others. In [1] capture motion technique was used to determine the indices of gait for patients with Parkinson's disease. On the basis of the position of the markers indices such as the step length or wrist trajectory were determined for each patient. The second group of indices – the arm deflexion angle and the Decomposition Index which requires computation of time interval when selected joints are in motion – is based on the Euler angles. In addition to publications where the authors focus on the analysis of a single person, capture motion technique is also used to analyze a group of people. In [2] authors considered a human identification problem based on joints movement expressed by the Euler angles.

Digital recording using the video camera doesn't provide accurate output data hence their analysis is difficult. Depending on considered issue, authors put up with some inaccuracy, or they try to project objects from 2D to 3D space. In [3] the authors apply the particle swarm optimization algorithm which creates a model of human embedded in three-dimensional space by synchronization of images from four video cameras. Data processed in this way were recorded for 22 people and gave 90% success rate in their identification. For the purpose of person identification the authors used the distance between the ankles, the person's height and the angles in selected joints. Digital recording of objects using the Kinect sensor allows research of human gait identification. In [4] the authors explored the problem of classification of gender gait quality. Data were collected using the Kinect sensor. In the data preprocessing authors used tracking motion method to get a human skeleton in 3D space. Application of 3D version of the Discrete Wavelet Transform allowed obtaining 83.75% effectiveness in the problem of classification.

In [5] the authors present a method which detects phases of the gait cycle (stride): stance and swing. The authors compared the parameters obtained from the 3D virtual skeleton from the Kinect sensor and from the in-shoe pressure sensors and a gyro-scope attached to the wrist via straps. The average difference between the duration measured by the pressure sensor and the duration measured by the regression model is less than 1%.

Step length, step width and gait variability, which can provide information about mobility and balance of human, have particular importance in the gait analysis. Lack of reproducibility during gait and step length are associated with the fall risk [6] which, in particular for older people, may lead to permanent disability. Analyzing spatial gait variability parameters authors registered the biggest correlation between the step length and the level of medications. This may indicate that the step length is an important factor that reflects the loss of mobility in humans causing a fall risk. In addition, elevated serum anticholinergic activity has been associated with significant slowing of both gait speed and simple response time.

In [7] authors compared two web-cameras and two Kinects with Vicon. Estimation of the stride length was better with web-cameras than with Kinect. On the other hand, the Kinect sensors were better than video cameras in estimation of stride time and its velocity.

A Kinect sensor as a non-invasive and relatively cheap equipment which, compared to other image recording devices, could be used at home where user's behavior is natural which leads to objective and reliable data. However, data collected by a Kinect sensor may be more or less accurate. Thus, laboratory studies should identify causes of inaccurate measurements which could be, among others, location of the sensor, method of computation of the step length.

This work presents the analysis of step length parameter, based on the methods of digital recording that can be used at home. For this purpose, a Kinect sensor was used to record digital data. To verify the accuracy of recorded data, we used an additional source of recordings based on the motion capture technique.

2 Dataset

The recordings occurred in the Human Motion Laboratory of the Polish-Japanese Academy of Information Technology in Bytom [8]. The laboratory is equipped with the Vicon motion capture system based on 10 cameras operating in the near-infrared light and 4 reference cameras operating in the visible light range. In addition, we used a Kinect sensor. Actors walked on a straight line on the path of 5m length which allows performing about 8 steps by healthy people. According to the instructions on the manufacturer's website (`http://support.xbox.com/`), sensor was located at a height of 55 cm above the floor at the end of the path on which subjects walked. The study included six actors with not identified abnormalities in the musculoskeletal system. Actors performed four walking trials along the path.

Gait has a cyclical nature and is characterized by the repetition of events such as foot contact with the ground or forward leg movement [9], therefore it is assumed that every step should be similar to every other step. Limited area of sensor‘s visibility and restricted length of walking path used by actors to walk, caused that only the middle step was chosen to analyze as a gait representative. In addition, the middle step is not perturbed by the start phase or the stop phase. It should be determined separately for each side of the human body, because for the given side a step begins with the first contact of the opposite foot with the ground and ends with the first contact of the given foot with the ground.

3 Preprocessing

The data collected from the reference system (motion capture) are stored in binary .c3d files. In the preprocessing phase Java implementation based on the j3d.org library was used. The markers placed at the ankles (RHLL marker for the right ankle and LHLL marker for left ankle) have been selected for the analysis purposes. Events needed to determine the step length, that's events in which foot contact with the ground occurs, are defined as FootStrike in the Vicon system. On the basis of the coordinates of the position of the foot in 3D space the Euclidean distance was used to determine the step length.

The Kinect Studio application developed by Microsoft, which is designed for recording motion data, does not allow directly export of collected data. Therefore, based on Kinect for Windows Software Development Kit v. 1.7 an application was created, which allows the registration of movement and their export. Kinect Sensor captures image and audio data. An image contains: video, depth map and the simplified human body skeleton. Unlike video data and depth map, which record a scene, in order to record a human body skeleton the registered object must be seen for a longer time on the camera so that a sensor can detect the skeleton. As mentioned earlier, analysis of human gait should not include a start or stop phase. Therefore, the actors started and finished walk aside from a camera visibility. This distance was too large for a sensor to recognize a skeleton. For this reason, a step length was analyzed only on the basis of the depth map and video. Stride length was analyzed from the camera located across a path on which actors were walking. For this purpose, video data were synchronized with the depth map. Events representing foot contact with the ground were selected in the video data to get their coordinates. Based on collected coordinates a depth image value of pixel was read from the depth map. In the preprocessing phase the depth map has been scaled to 8-bit blue color component. Thus, distance between two points on the depth map is determined by the difference of their values for the blue channel. To determine the real distance between the points, the distance calculated from the depth map for six selected calibration marks 1 was compared with real distance between them. The change rate of the depth image value from the depth map to the real image is equal to 42.857 mm. This remark can be expressed by the following equation: $world\_depth\_difference = depth\_map\_difference * 42.857$.

The change of values ranges on the depth map which depends on the registered values of the calibration marks was also tested. The graph on the depth

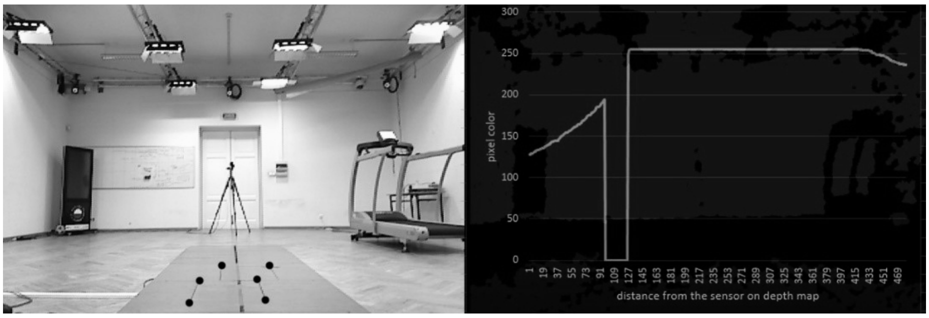

Fig. 1. Six calibration marks determining the real distance. Relationship between a depth image value and the distance.

map in Fig. 1 shows linear increase of values on the depth map for a subject located farther away from the Kinect sensor. Black areas and a sharp drop in the graph illustrate the place for which the sensor has not registered the depth map.

4 Results

From recordings captured by the Kinect sensor three frames were selected. In each of them the first foot contact with the ground is visible, which is marked as a circle in Fig. 2.

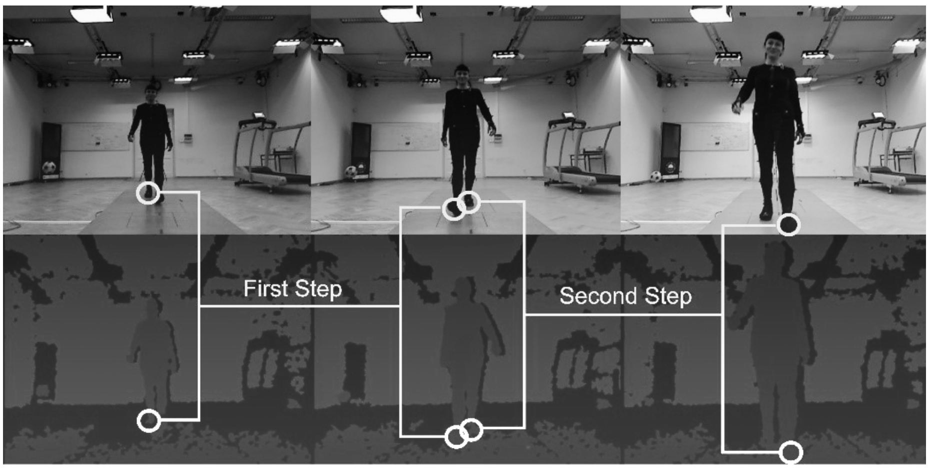

Fig. 2. Determination of steps.

Identification of a single step requires use of depth values from two images. It is worthwhile to mention that presented images are reflections of the real images. Hence, the first step constitutes a step made by the left foot and the second one – a step made by the right foot.
Since the analysis is based on images from one sensor with a resolution of 640x480, where depending on a foot orientation only its part is visible, the depth values were sampled from different locations. In the first experiment, depth map values were taken from the places chosen subjectively as locations where the foot is most visible on the image. Experiment was repeated three times for each actors walking.

In the second experiment values were collected from characteristic locations such as toes, ankles or heels (depending on the actor's gait) and the central part of an ankle. In the last experiment, the values taken from characteristic locations were averaged.

The step lengths (in millimeters) computed for each of the above described experiments are shown for a left and right foot in Table 1 and Table 2, respectively. Step length was averaged for each actor over his four walks.

Table 1. Step length of the left foot.

		Depth Map				
	Motion capture	Most visible place	toes	ankles or heels	central part of an ankle	average
Actor1	1450,68	1532,14	1060,71	1285,71	1564,29	1342,86
Actor2	1290,61	1107,14	782,14	1060,71	1060,71	967,86
Actor3	1333,55	1333,55	964,29	1425,00	1542,86	1328,57
Actor4	1427,24	1617,86	1125,00	1596,43	1607,14	1442,86
Actor5	1330,79	1414,75	1157,14	1135,71	1467,86	1289,29
Actor6	1385,14	1574,66	1200,00	1510,71	1521,43	1410,71
correlation		0,87	0,58	0,61	0,74	0,74

The correlation between depth map data and data obtained using the motion capture system was examined for each variant of the depth map data collection method. For the left side the highest correlation 0.86 was obtained for data collected from the most visible places on human foot. In contrast, the lowest correlation 0.58 was related to the data collected from the toes.

For the right side the highest correlation 0.86 was obtained for data collected from the toes. The lowest correlation 0.47 was related to the data collected from both the ankle and the heel.

5 Conclusion

The Kinect sensor is designed to record movements in a restricted area of the scene. Thus, the research becomes difficult when the subject goes beyond this

Table 2. Step length of the right foot.

		Depth Map				
	Motion capture	Most visible place	toes	ankles or heels	central part of an ankle	average
Actor1	1397,53	1589,29	1467,86	1360,71	1639,29	1489,29
Actor2	1266,73	1178,57	1125,00	1125,00	1178,57	1142,86
Actor3	1374,53	1374,53	1617,86	1414,29	1746,43	1592,86
Actor4	1497,95	1628,57	1832,14	1553,57	1917,86	1767,86
Actor5	1364,17	1514,69	1660,71	1767,86	1778,57	1735,71
Actor6	1364,59	1634,62	1596,43	1028,57	1778,57	1467,86
correlation		0,77	0,87	0,48	0,84	0,81

area. Nevertheless, analysis of the step length collected using the Kinect sensor revealed high correlation of 86% between recordings of different actors and recordings using motion capture technology. The relative difference of step length between two discussed approaches of movement recording averaged over all actors is equal to 13%. The method of retrieving data from three places of the feet obtained quite high correlation for both left and right side. Thus, the use of additional calibration marks could increase correlation, at the same time improving the ability to identify unnatural steps. Perhaps the more accurate results could be obtained using the newest version of the Kinect sensor because according to the producer the device should be charac-terized by higher fidelity depth map.

Acknowledgments. This project has been supported by the National Centre for Research and Development, Poland.(UOD-DEM-1-183/001 "Intelligent video analysis system for behavior and event recognition in surveillance networks").

References

1. Stawarz, M., Kwiek, S., Polański, A., Janik, Ł., Boczarska-Jedynak, M., Przybyszewski, A., Wojciechowski, K.: Algorithms for Computing Indexes of Neurological Gait Abnormalities in Patients after DBS Surgery for Parkinson Disease Based on Motion Capture Data. Machine Graphics and Vision. v. 20, No. 3, pp. 299–317. (2011)
2. Świtoński, A., Josiński, H., Michalczuk, A., Pruszowski, P., Wojciechowski, K.: Feature Selection of Motion Capture Data in Gait Identification Challenge Problem. ACIIDS, pp. 535–544. (2014)
3. Krzeszowski, T., Michalczuk, A., Kwolek, B., Świtoński, A., Josiński, H.: Gait recognition based on marker-less 3D motion capture. AVSS 2013: pp. 232–237. (2013)
4. Arai, K., Asmara, R.A.: Human Gait Gender Classification using 3D Discrete Wavelet Transform Feature Extraction. International Journal of Advanced Research in Artificial Intelligence. 3(2), pp. 12–17. (2014)
5. Gabel, M., Gilad-Bachrach, R., Renshaw, E., Schuster, A.: Full body gait analysis with Kinect Engineering in Medicine and Biology Society (EMBC). 2012 Annual International Conference of the IEEE, pp. 1964–67. (2012)

6. Nebes, R.D., Pollock, B.G., Halligan, E.M. Kirshner, M.A. Houck, P.R.: Serum Anti-cholinergic Activity and Motor Performance in Elderly Persons. The Journals of Gerontology. Series A, v. 62, pp. 83–85. (2007)
7. Stone, E., Skubic, M. Passive.: In-Home Measurement of Stride-to-Stride Gait Variability Comparing Vision and Kinect Sensing. 33rd Annual International Conference of the IEEE EMBS, IEEE 2011, pp.6491–94. (2011)
8. Website of the Human Motion Laboratory of the Polish-Japanese Academy of Information Technology: `http://hml.pjwstk.edu.pl/`
9. Barlett, R.: Introduction to sports biomechanics, Analyzing Human Movement Patterns. Taylor & Francis Group (2007)