

Use of the image and depth sensors of the Microsoft Kinect for the detection of gait disorders

Aleš Procházka · Oldřich Vyšata · Martin Vališ ·
Ondřej Ťupa · Martin Schätz · Vladimír Mařík

Received: 18 October 2014 / Accepted: 15 January 2015 / Published online: 3 February 2015
© The Natural Computing Applications Forum 2015

Abstract This paper presents a novel method of gait recognition that uses the image and depth sensors of the Microsoft (MS) Kinect to track the skeleton of a moving body and allows for simple human–machine interaction. While video sequences acquired by complex camera systems enable very precise data analyses and motion detection, much simpler technical devices can be used to analyze video frames with sufficient accuracy in many cases. The experimental part of this paper is devoted to gait data acquisition from 18 individuals with Parkinson’s disease and 18 healthy age-matched controls via the proposed MS Kinect graphical user interface. The methods designed for video frame data

processing include the selection of gait segments and data filtering for the estimation of chosen gait characteristics. The proposed computational algorithms for the processing of the matrices acquired by the image and depth sensors were then used for spatial modeling of the moving bodies and the estimation of selected gait features. Normalized mean stride lengths were evaluated for the individuals with Parkinson’s disease and those in the control group and were determined to be 0.38 and 0.53 m, respectively. These mean stride lengths were then used as features for classification. The achieved accuracy was >90 %, which suggests the potential of the use of the image and depth sensors of the MS Kinect for these applications. Further potential increases in classification accuracy via additional biosensors and body features are also discussed.

A. Procházka (✉) · O. Ťupa · M. Schätz
Department of Computing and Control Engineering,
Institute of Chemical Technology in Prague,
166 28 Prague 6, Czech Republic
e-mail: A.Prochazka@ieee.org

O. Ťupa
e-mail: Ondrej.Tupa@vscht.cz

M. Schätz
e-mail: Martin.Schatz@vscht.cz

A. Procházka · V. Mařík
Czech Institute of Informatics, Robotics and Cybernetics,
Czech Technical University in Prague,
166 36 Prague 6, Czech Republic

V. Mařík
e-mail: marik@ciirc.cvut.cz

O. Vyšata · M. Vališ
Department of Neurology, Faculty of Medicine in Hradec
Králové, Charles University in Prague,
500 05 Hradec Králové, Czech Republic
e-mail: Vysatao@gmail.com

M. Vališ
e-mail: Valismar@seznam.cz

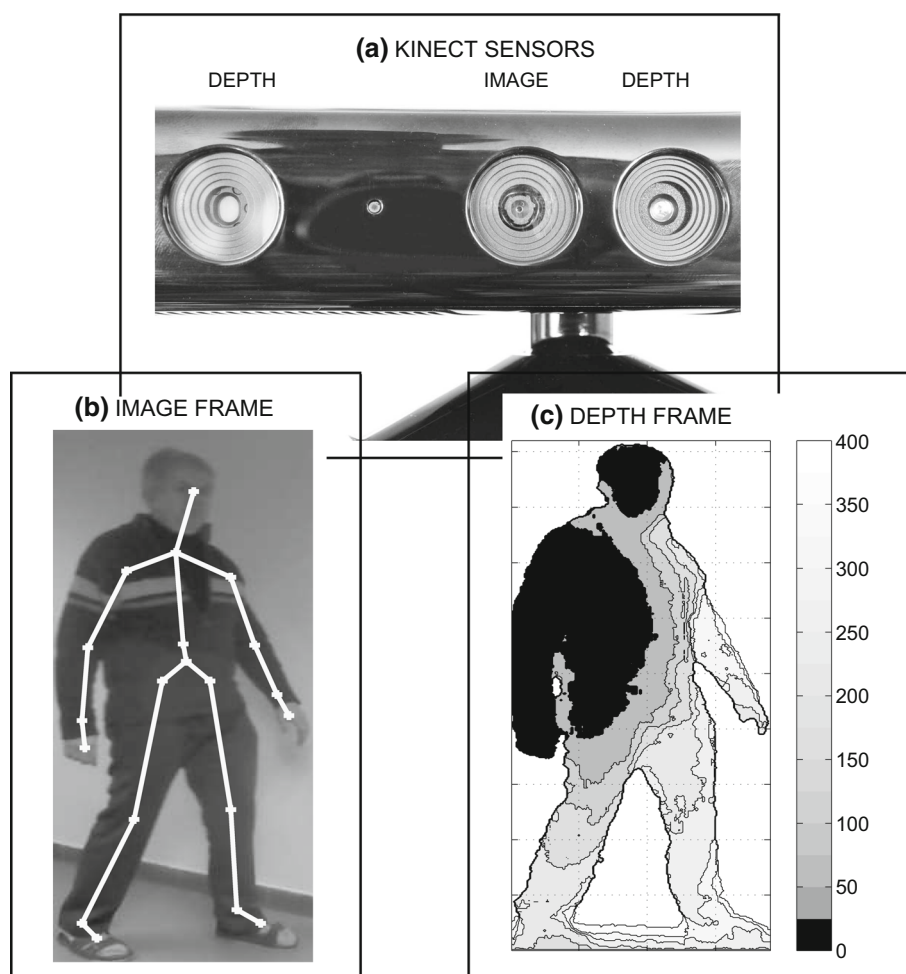
Keywords Image and depth sensors · Gait disorders ·
Motion features · MS Kinect · Video processing ·
Parkinson’s disease

1 Introduction

Systems that enable human–machine interactions and spatial modeling have a wide range of applications in modern engineering, robotic and biomedical devices [1].

While complex synchronized video camera systems [5–7] represent precise but expensive technical solutions, it is possible to use much cheaper systems that employ depth sensors to acquire data with sufficient accuracy for many applications. The Microsoft (MS) Kinect™ [2–4, 8] allows the recording of such data sets via its image and depth sensors (illustrated in Fig. 1a) and the transfer of these data to appropriate mathematical environments, such as

Fig. 1 An example frame recorded by the MS Kinect including: **a** the MS Kinect's RGB camera for image video recording and depth sensors, which includes a projector and a receiver, for the acquisition of depth frame matrices, **b** the image frame matrix combined with the skeleton estimate and **c** the *contour* plot of the depth frame matrix with distances from the selected plane



MATLAB, for further processing. The acquired data sets can then be used to propose methods and algorithms for movement analyses, scene modeling [9], gesture and body recognition [10–12], rehabilitation [13] and posture reconstruction [14–16]. These new devices, combined with motion sensors and specific control units, are also often used in robotic systems control [11, 17–19].

This article is devoted to the use of the MS Kinect system for movement data acquisition, detection of gait features and analysis of gait disorders [20–24] via selected digital signal and image processing methods. The proposed graphical user interface was used to acquire data in a clinical environment from patients with Parkinson's disease [25–27] and from healthy individuals to form a reference dataset. Specific algorithms were then designed and used for motion tracking, gait feature evaluation and the classification of the observed sets of individuals. The obtained results were evaluated from both the engineering and neurological perspectives.

The proposed methods show how modern sensors can be used to acquire video frames of matrices that enable

human–machine interaction. The application discussed here is related to the analysis of gait disorders but could be further extended to other areas, including rehabilitation engineering and robotic systems control.

2 Methods

2.1 Data acquisition

The information related to the body motions of the selected individuals was recorded with the MS Kinect sensors, which are illustrated in Fig. 1a. The RGB camera in the middle of the device recorded video image frames with a resolution of 480×640 pixels and a frequency of 30 frames/s. The depth sensor consists of an infrared projector (on the left) and an infrared camera (on the right) that use the structured light principle [28, 29] to detect the distances of image pixels with a precision of 4–40 mm depending upon the distance from the sensor. The resulting matrix has a size of 480×640 .

Figure 1b, c present portions of selected frames that were recorded by the image and depth sensors. The selected image presented in Fig. 1b was combined with the skeleton projection and the estimated positions of the joints. Figure 1c illustrates information from the depth sensor. The contour plot in Fig. 1c presents the distances of the individual pixels from a selected plane that was at a distance of 2,200 mm from the MS Kinect. Structure of the whole system for data acquisition is presented in Fig. 2.

The proposed graphical user interface (GUI) was used to record MS Kinect data from the observed individuals and to process the information in the MATLAB (version 2014b) environment [30]. The GUI was designed to allow the simple recording of video frames in clinical environments via the following steps:

1. the recording of the name and surname of the patient;
2. an initialization of the MS Kinect system (connected through a USB); and
3. the recording or interruption of data acquisition.

Additional functions of the graphical user interface included the selection of further parameters for recording and the previewing of data sets including the options of previewing image and depth sensor data from the database.

The skeleton tracking algorithm, which processes the data, also provides information about the locations of joints as specified in Table 1. The joint numbering and the connection map are presented in Table 1 as well.

Both the RGB camera and the depth sensors store information in 640×480 element matrices according to the schematic diagram of the system presented in Fig. 2. A histogram of differences in static portions of consecutive image frames illustrates the accuracy of the MS Kinect image sensor. Figure 2 also presents a similar histogram of these differences for the static portions of the matrices that include the depth sensor data. The accuracy of the system is fundamental for spatial data modeling [31–34] and, as expected, was in the range of -50 to 50 mm.

The experimental portion of this study was devoted to the analyses of the gaits of the following two sets of individuals: (1) 18 patients with Parkinson’s disease (PD) and (2) 18 healthy age-matched individuals (Norm). The MS Kinect that was used for data acquisition was installed approximately 60 cm above the floor. Each individual repeated a straight walk of approximately 4 m (five steps) back and forth 5 times. Each video record acquired with the sampling rate of 30 frames/s contained useful information about the direct walks but also unnecessary frames recorded during the turns.

Fig. 2 The coordinate system for the MS Kinect data acquisition and distribution of image and depth sensors errors based on differences of subsequent submatrices in their static regions for selected data frames

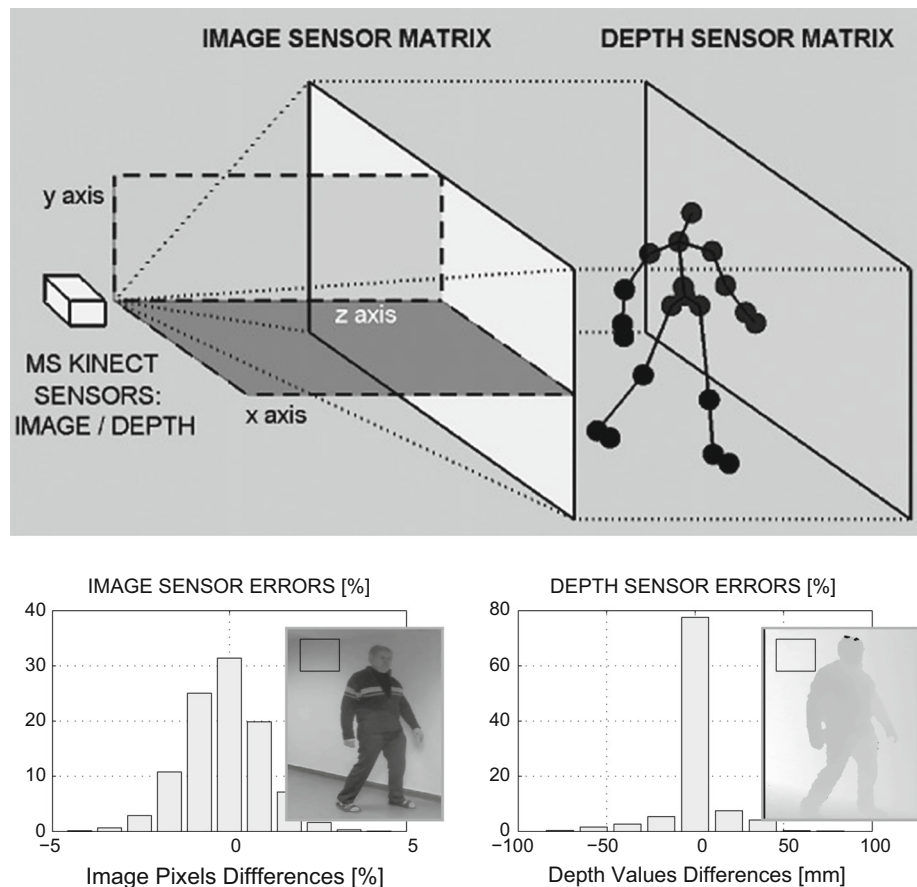
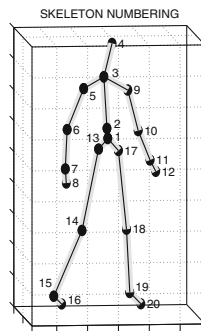


Table 1 Skeleton positions and connection map of an individual used for data acquisition and processing of video records

Skeleton positions			
Joint	No.	Joint	No.
Hip center	1	Wrist right	11
Spine	2	Hand right	12
Shoulder center	3	Hip left	13
Head	4	Knee left	14
Shoulder left	5	Ankle left	15
Elbow left	6	Foot left	16
Wrist left	7	Hip right	17
Hand left	8	Knee right	18
Shoulder right	9	ankle right	19
Elbow right	10	Foot right	20



Connection map

Part	Connection vectors
Spine	[1 2], [2 3], [3 4]
Left hand	[3 5], [5 6], [6 7], [7 8]
Right hand	[3 9], [9 10], [10 11], [11 12]
Left leg	[1 13], [13 14], [14 15], [15 16]
Right leg	[1 17], [17 18], [18 19], [19 20]

2.2 Skeleton tracking and stride length estimation

The skeleton tracking algorithm processed data matrices from the image and depth sensors. The algorithm also provided coordinates that specified the spatial locations of all joints in the selected coordinate system, as illustrated in Fig. 2, by utilizing the joint numbering and connection maps defined in Table 1.

The steps of proposed algorithm for gait feature detection that utilized the MS Kinect can be summarized as follows:

1. a preprocessing of the skeleton data to remove gross errors and finite impulse response (FIR) filtering of the joint positions to minimize observation errors;
2. the rejection of frames with substantial errors based on the temporal evolution of the centers of mass (COM) as evaluated based on joints 1, 2 and 3 (i.e., the hip-center, spine, and shoulder-center joints, respectively) and presented in Fig. 3a, b;
3. the extraction of the gait segments in one direction and the rejection of distorted frames;
4. the evaluation of the positions of the leg centers from joints 15, 16 and 19, 20 (i.e., the ankle and foot joints of the left and right legs, respectively) in each segment

(Fig. 3a) and estimation of the average step length via the evaluation of their Euclidian distances (Fig. 3c) and the detection of their maxima;

5. the evaluation of the lengths of the legs of each individual based on the spatial positions of joints 13, 14, 15 and 17, 18, 19 (i.e., the hip, knee and ankle joints of the left and right legs, respectively) and with the results presented in Fig. 4a, the averaging for each individual and normalization of the step lengths; and
6. the estimation of the stride features via the averaging of the normalized step lengths for each individual in each segment of a straight walk.

Further stride features, including walking speed, can be detected in a similar manner using data from the image and depth sensors of the MS Kinect.

3 Results

Table 2 presents descriptions of the data sets from the 18 individuals (12 men and 6 women) with the Parkinson's disease and the 18 controls (7 men and 11 women). All patients met UK Parkinson's Disease Society brain bank clinical diagnostic criteria, and they were monitored in the movement disorder unit.

Using the proposed algorithm after the reduction in the observation errors, the numerical results obtained from the data (acquired with the MS Kinect at a sampling rate of 30 frames per second) are presented in Table 2. The resulting average stride lengths (SLs) suggest that this system could be used to classify individuals with Parkinson's disease (SL = 0.38 m, SD = 0.06) using the age-matched individuals (SL = 0.53 m, SD = 0.05) as a reference set. As expected the average stride length was shorter for the PD group than for the reference set.

The leg lengths were analyzed separately to normalize the stride lengths of the specific individuals and to increase the reliability of the gait features. The joint positions in three-dimensional space were estimated with the MS Kinect system and used to evaluate the Euclidean distances between the hip and knee and the knee and ankle. The sums of these values for each leg were used to estimate leg lengths. The average difference in the lengths of the legs of each of the 36 individuals was 11 mm with a standard deviation of 8 mm, which was within a normal range according to a long history of clinical observations. The average values of the leg lengths of each individual were calculated (across all subjects, the average leg length was 0.784 m with a standard deviation of 0.011 m) and used for the stride length normalization. Figure 4a presents these results for the reference set of subjects.

Fig. 3 Detection of the right and left legs movement illustrating: **a** temporal evolution of the centers of mass and leg centers in the three-dimensional space, **b** temporal evolution of the centers of mass positions above the horizontal plane and **c** the distances between the leg centers for a selected walk segment

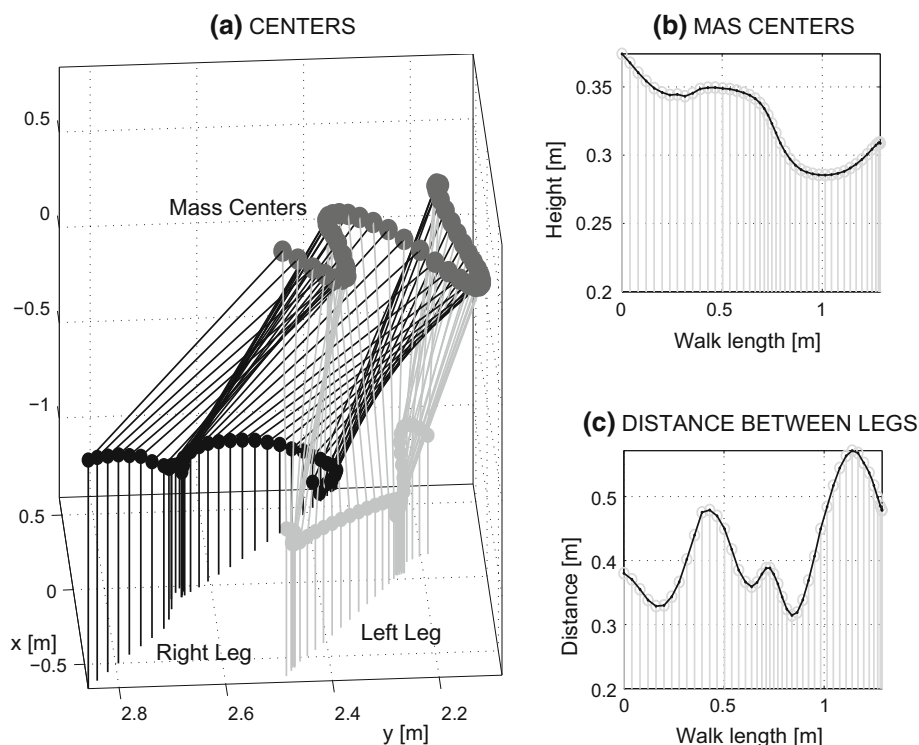


Fig. 4 Selected features presenting: **a** individual leg lengths evaluated from MS Kinect joints detection of the reference set and age dependence of stride lengths for **b** the individuals with Parkinson’s disease (PD set) and **c** the age-matched controls (reference set)

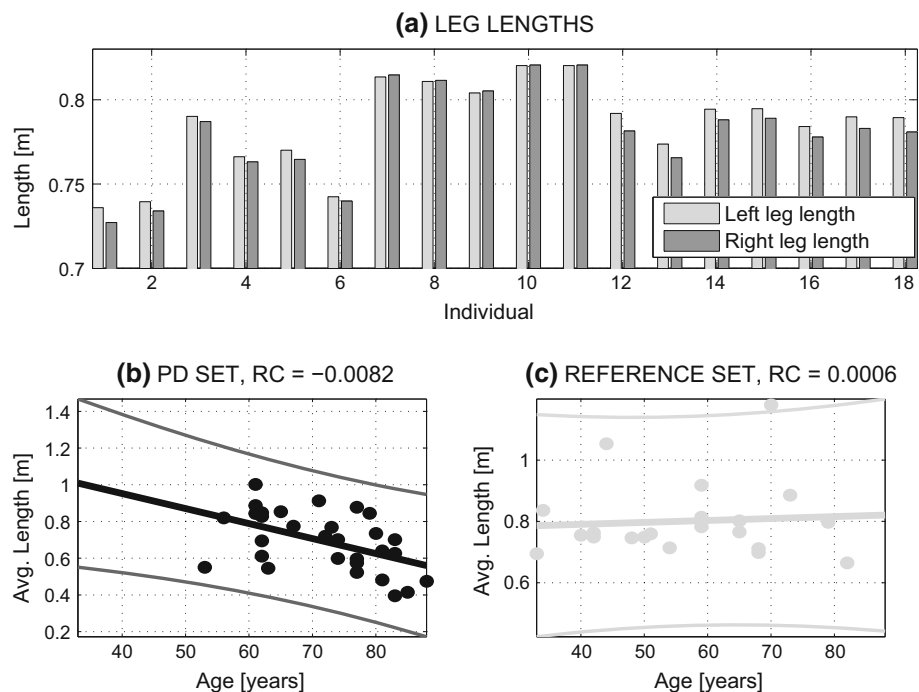


Figure 4b, c present the age-dependent stride lengths for the individuals with Parkinson’s disease (positive set) and the age-matched controls (negative set). While for the positive set, it is possible to observe the decrease in the stride lengths with the regression coefficient

$RC = -0.0082$ m/year, no such dependence exists for the reference set. For this reason, no normalization to the age of individuals was performed in this study.

Figure 5 presents the distributions of stride lengths for the individuals with Parkinson’s disease (positive set) and

Table 2 Stride length results for the two sets of individuals (PD: Parkinson’s disease, Norm: control) with the standard deviations (SD)

Group size	Subjects age (years)			Stride length (m)	
	Range	Mean	SD	Mean	SD
PD: 18	52–87	73.6	9.2	0.38	0.06
Norm: 18	32–81	55.0	14.5	0.53	0.05

the age-matched controls (negative set), and these data correspond to those in Table 2.

A more detailed analysis of the data involving sensitivity and specificity [35] is presented in Fig. 5. As shown in Fig. 5a, it was possible to determine the true negative (TN) and false positive (FP) rates in the negative set (i.e., the controls) and the true positive (TP) and false negative (FN) rates in the positive set (i.e., the patients with

Fig. 5 Stride length analysis for the individuals with Parkinson’s disease (positive set) and the age-matched controls (negative set) presenting: **a** the distributions of true and false results across criterion values, **b** the accuracy achieved with the optimal criterion value and **c** the sensitivity/specificity plots

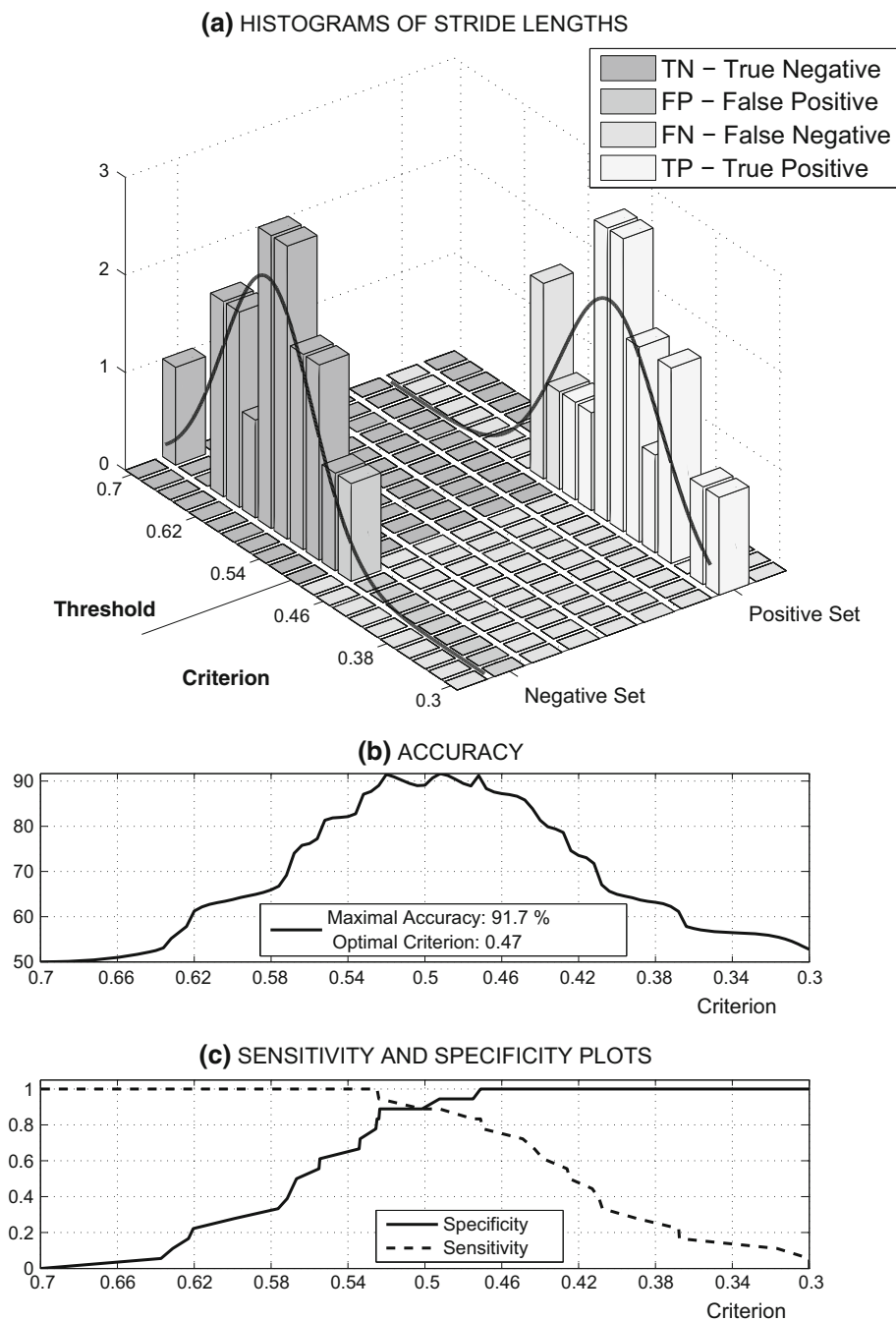
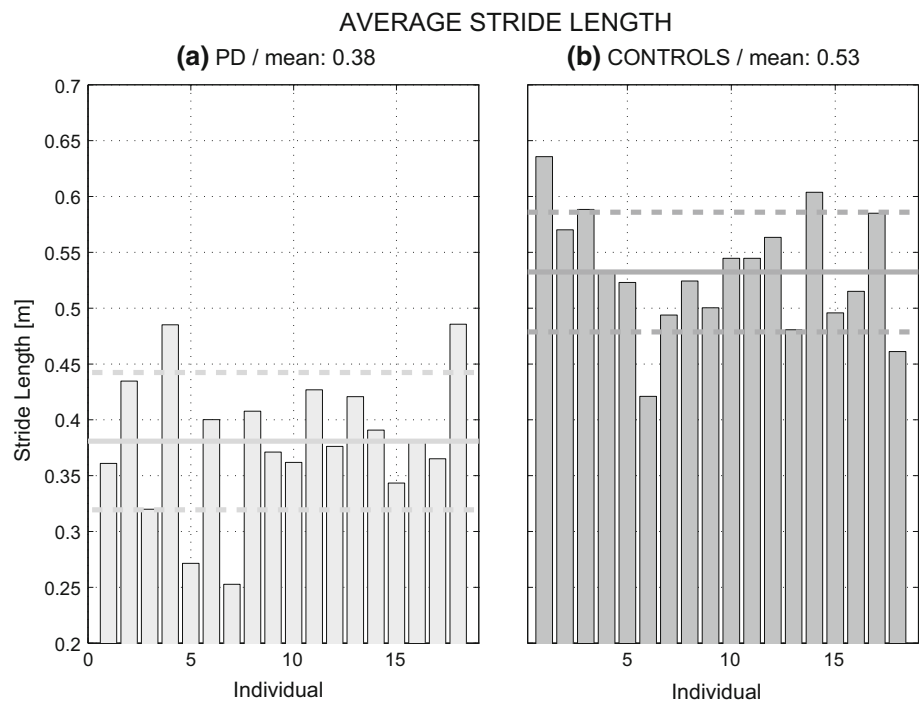


Fig. 6 The average stride lengths (with standard deviations) of **a** the patients with Parkinson’s disease and **b** the reference set of individuals



Parkinson’s disease) for any given threshold step length value. Next, the accuracy (ACC), sensitivity (SE) and specificity (SP) were evaluated according to the following relations:

$$ACC = \frac{TN + TP}{TN + FP + TP + FN}, \tag{1}$$

$$SE = \frac{TP}{TP + FN}, \tag{2}$$

$$SP = \frac{TN}{TN + FP} \tag{3}$$

with respect to the selected criterion (step length) value. The resulting plot in Fig. 5b shows that it was possible to achieve an accuracy of 91.7 % at the optimal step length threshold of 0.47 m.

Figure 6 presents the average stride lengths (with standard deviations) of the individuals in the positive and negative sets and the results of the accuracy analyses. These results suggest that stride length could be a useful feature for the classification of individuals into Parkinson’s disease and non-Parkinson’s disease groups.

4 Conclusion

Human–machine interaction and computer intelligence are key components of the rapidly developing interdisciplinary field that combines sensor technology, data fusion, computer vision, image processing, control engineering and

robotics. Numerous papers have been devoted to the identification and detection of motion features [36–38] with applications in biomedical signal processing and the diagnoses of gait disorders [39, 40]. The latest research [41, 42] related to wearable and non-wearable systems indicate the increasing interest in portable systems and specific body sensors for gait analysis.

This motion analysis and Parkinson’s disease recognition can be performed by specialized and expensive camera systems with specific sensors. These systems are commonly used for the detections of movement with high accuracy. This paper presents a new approach to the analysis of gait disorders that utilizes the relatively inexpensive MS Kinect. The MS Kinect has a depth sensor accuracy of 4–40 mm, which is sufficient for many applications. The results obtained suggest that the MS Kinect can be used for the detection of gait disorders and for the recognition of Parkinson’s disease. The maximum accuracy observed in the present study was 91.7 %.

Further work will be devoted to the study of more extensive data sets and the evaluation of a higher number of parameters to increase the classification accuracy of motion features. We assume that combination of data from an increased number of biosensors will produce pattern matrices that can be used for more accurate classifications across a wide range of criteria values and provide tools for remote diagnostics and wireless data processing.

Acknowledgments Authors would like to thank all patients who signed the informed consent to participate in the project with all the

procedures approved by the local ethics committee as stipulated by the Helsinki Declaration. All data were kindly provided by the Movement Disorders Center of the Faculty Hospital in Hradec Králové, Czech Republic.

References

1. Brscic D, Kanda T, Ikeda T, Miyashita T (2013) Person tracking in large public spaces using 3-D range sensors. *IEEE Trans Hum Mach Syst* 43(6):522–534
2. Han J, Shao L, Xu D, Shotton J (2013) Enhanced computer vision with Microsoft Kinect sensor—a review. *IEEE Trans Cybern* 43(5):1318–1344
3. Procházka A, Vyšata O, Schätz M, Ťupa O, Yadollahi M, Vališ M (2014) The MS Kinect image and depth sensors use for gait features detection. In: *IEEE international conference on image processing (ICIP)*
4. Ťupa O (2014) Multi-dimensional data modelling and analysis using MS Kinect. Master thesis, Institute of Chemical Technology in Prague
5. Ringer M, Lasenby J (2002) Multiple hypothesis tracking for automatic optical motion capture. *Lecture Notes in Computer Science* 2350, Springer-Verlag Berlin, pp 524–536
6. Lasenby J, Stevenson A (2000) Using geometric algebra for optical motion capture. In: Bayro-Corrochano E, Sobczyk G (eds) *Applied Clifford algebras in computer science and engineering*. Birkhauser, Boston
7. Procházka A, Kubíček M, Pavelka A (2006) Multicamera systems in the moving body recognition. In: *IEEE 48th international symposium on ELMAR*
8. Fortino G, Giannantonio R, Gravina R, Kuryloski P, Jafari R (2013) Enabling effective programming and flexible management of efficient body sensor network applications. *IEEE Trans Hum Mach Syst* 43(1):115–133
9. Camplani M, Mantecon T, Salgado L (2013) Depth-color fusion strategy for 3-D scene modeling with Kinect. *IEEE Trans Cybern* 43(6):1560–1571
10. Lu J, Tan YP (2013) Ordinary preserving manifold analysis for human age and head pose estimation. *IEEE Trans Hum Mach Syst* 43(2):249–258
11. Yanik P, Manganeli J, Merino J, Threatt A, Brooks J, Green K, Walker I (2014) A gesture learning interface for simulated robot path shaping with a human teacher. *IEEE Trans Hum Mach Syst* 44(1):41–54
12. Li Y (2012) Multi-scenario gesture recognition using Kinect. In: *17th international conference on computer games*, pp 126–130
13. Strbac M, Markovic M, Popovic DB (2012) Kinect in neurorehabilitation: computer vision system for real time hand and object detection and distance estimation. In: *IEEE 11th symposium on neural network application in electrical engineering (NEUREL 2012)*, USA, pp 127–132
14. Milovanovic M, Minovic M, Starkevic D (2012) New gait recognition method using Kinect stick figure and CBIR. In: *Proceedings of the telecommunications forum (TELFOR)*, pp 1323–1326
15. Shum H, Ho E, Jiang Y, Takagi S (2013) Real-time posture reconstruction for Microsoft Kinect. *IEEE Trans Cybern* 43(5):1357–1369
16. Choudry M, Beach T, Callaghan J, Kulic D (2013) A stochastic framework for movement strategy identification and analysis. *IEEE Trans Hum Mach Syst* 43(3):314–327
17. Zhang C, Xu J, Xi N, Jia Y, Li W (2012) Development of an omni-directional 3D camera for robot navigation. In: *2012 IEEE/ASME international conference on advanced intelligent mechatronics*, USA, pp 262–267
18. Wang Y, Yang C, Wu X, Xu S, Li H (2012) Kinect based dynamic hand gesture recognition algorithm research. In: *4th international conference on intelligent human-machine systems and cybernetics*, pp 274–279
19. Machida E, Cao M, Murao T, Hashimoto H (2012) Human motion tracking of mobile robot with Kinect 3D sensor. In: *SICE annual conference*, Akita University, Japan, pp 2207–2211
20. Clark RA, Pua YH, Fortin K, Ritchie C, Webster KE, Denehy L, Bryant AL (2012) Validity of the Microsoft Kinect for assessment of postural control. *Gait Posture* 36:372–377
21. Gabel M, Renshaw E, Schuster A, Gilad-Bachrach R (2012) Full body gait analysis with Kinect. In: *International conference of the IEEE engineering in medicine and biology*
22. Preis J, Kessel M, Werner M, Linnhoff-Popien C (2012) Gait recognition with Kinect. In: *1st international workshop on Kinect in pervasive computing*, pp P1–P4
23. Weber I, Koch J, Meskemper J, Friedl K, Heinrich K, Hartmann U (2012) Is the MS Kinect suitable for motion analysis? *Biomed Eng* 57(8):664
24. Lee S, Shin S (2014) Gait signal analysis with similarity measure. *Sci World J* 2014:1–8
25. Gelb D, Oliver E, Gilman S (2011) Diagnostic criteria for Parkinson disease. *JAMA Neurol* 56(1):33–39
26. Venkat I, De Wilde P (2011) Robust gait recognition by learning and exploiting sub-gait characteristics. *Int J Comput Vis* 91(1):7–23
27. Yogev G, Giladi N, Peretz C, Springer C, Simon ES, Hausdorff JM (2005) Dual tasking, gait rhythmicity, and Parkinson's disease: which aspects of gait are attention demanding? *Eur J Neurosci* 22(5):1248–1256
28. Zhang Z (2012) Accuracy and resolution of Kinect depth data for indoor mapping applications. *IEEE Multimed* 19(2):4–10
29. Qin S, Zhu X, Yang Y (2014) Real-time hand gesture recognition from depth images using convex shape decomposition method. *J Signal Process Syst* 74:47–58
30. *Image Acquisition Toolbox: User's Guide* (2013) The MathWorks, Natick
31. Khoshelham K, Elberink SO (2012) Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors* 12:1437–1454
32. Obdrzalek S, Kurillo G, Ofi F, Bajcsy R, Seto E, Jimison H, Pavel M (2012) Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. In: *34th annual international conference of the IEEE EMBS, IEEE, USA*, pp 1188–1193
33. Smisek J, Jancosek M, Pajdla T (2012) 3D with Kinect. In: *IEEE international conference on computer vision, IEEE, USA*, pp 1154–1160
34. Essmaeel K, Gallo L, Damiani E, De Pietro G, Dipanda A (2012) Temporal denoising of Kinect depth data. In: *Eighth international conference on signal image technology and internet based systems. IEEE Computer Society, IEEE, USA*, pp 47–52
35. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874
36. BenAbdelkader C, Cutler R, Davis L (2002) Person identification using automatic height and stride estimation. In: *Proceedings of the 16th international conference on pattern recognition*, vol 4, pp 377–380
37. Munsell B, Temlyakov A, Qu C, Wang S. (2012) Person identification using full-body motion and anthropometric biometrics from Kinect videos. In: *ECCV 2012 ARTEMIS workshop*
38. Dockstader S, Bergkessel K, Tekalp A (2002) Feature extraction for the analysis of gait and human motion. In: *Proceedings of the 16th international conference on pattern recognition*, vol 1, pp 5–8
39. Vysata O, Schatz M, Valis M, Prochazka A, Simunek L (2013) Quantitative gait analysis in Parkinson's disease using MS

- Kinect: entropy and fractal dimension. *J Neurol Sci* 333(Suppl. 1):e67
40. Chen Y, Cho C, Lin S, Lai H, Lo Y, Chen S, Chang Y, Huang W, Chen C, Jaw F, Tsang S, Tsai S (2012) A vision-based regression model to evaluate Parkinsonian gait from monocular image sequences. *Expert Syst Appl* 39:520–526
41. Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A (2014) Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors* 14:3362–3394
42. Galna B, Barry G, Jackson D, Mhiripiri D, Olivier P, Rochester L (2014) Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait Posture* 39:1062–1068