Abnormal Gait Detection with RGB-D Devices using Joint Motion History Features

Alexandros Andre Chaaraoui[1], José Ramón Padilla-López[1] and Francisco Flórez-Revuelta[2]
[1] Department of Computer Technology, University of Alicante, P.O. Box 99, Alicante E-03080, Spain
[2] Faculty of Science, Engineering and Computing, Kingston University, Penrhyn Road, Kingston upon Thames KT1 2EE, UK
alexandros@dtic.ua.es, jpadilla@dtic.ua.es, F.Florez@kingston.ac.uk

***Abstract*— Human gait has become of special interest to health professionals and researchers in recent years, not only due to its relation to a person's quality of life and personal autonomy, but also due to the involved cognitive process, since deviation from normal gait patterns can also be associated to neurological diseases. Vision-based abnormal gait detection can provide support to current human gait analysis procedures providing quantitative and objective metrics that can assist the evaluation of the geriatrician, while at the same time providing technical advantages, such as low intrusiveness and simplified setups. Furthermore, recent advances in RGB-D devices allow to provide low-cost solutions for 3D human body motion analysis. In this sense, this work presents a method for abnormal gait detection relying on skeletal pose representation based on depth data. A novel spatio-temporal feature is presented that provides a representation of a set of consecutive skeletons based on the 3D location of the skeletal joints and the motion's age. The corresponding feature sequences are learned using a machine learning method, namely BagOfKeyPoses. Experimentation with different datasets and evaluation methods shows that reliable detection of abnormal gait is obtained and, at the same time, an outstandingly high temporal performance is provided.**

I. Introduction

Demographic change is one of the main concerns that our current society has to deal with. Ageing and birth rate studies and forecasts have repeatedly shown that the balance between elderly and working population will not be met in a near future. The European Statistical Office projects that by 2060, the ratio between working and retired people will have passed from four-to-one to two-to-one in the EU [1]. Technological innovation and development are key to ease this situation by providing support through assistive and autonomous care services, preferably based on low-cost infrastructure that can be installed at homes or care centres.

Human gait constitutes an essential metric related to a person's health and well-being. Its quality is not only dependent on the strength of the involved muscles but also on a complex mental coordination process [2]. Because of this, abnormal gait detection can also be important for the diagnosis of neurological diseases. Specifically, physical frailty is associated to an increased risk of cognitive impairment [3, 4]. A quantitative evaluation of a person's gait and the detection of deviations from normal gait patterns can support current frailty assessment performed by geriatricians leading to an improved and earlier detection of these diseases. In this sense, the present research aims to apply the recent improvements in markerless human body pose estimation based on low-cost RGB-D devices to gait analysis. Depth cameras such as the Microsoft Kinect, and its second version which has also been employed in this work, provide reliable depth data indoors. Based on the 3D skeletal body pose estimation that can be obtained [5], the human gait can be analysed, since the joint motion is related to the kinematics of the human body [6].

In this work, a novel skeletal-based spatio-temporal feature is presented. The so-called joint motion history feature (JMH) encodes spatial and temporal information providing a feature descriptor that can continuously be obtained across time and compared to previously stored templates, or learned in model-based machine learning algorithms. Similarly to Bobick and Davis' motion history images, both motion location and age are combined into a single feature representation [7]. Skeletons are used to track motion over a segment of frames based on the three-dimensional location of their joints. A volume is employed to set the appropriate 3D coordinates to a value that indicates the recentness of the motion based on the coordinates of the skeleton joints over time. This volume is then projected onto its three orthogonal axes to apply dimensionality reduction. In this way, a spatio-temporal feature is obtained that encodes human motion accurately and at a very detailed level without prohibitively increasing the feature dimensionality. Experimentation on a publicly available dataset from the University of Bristol [8] and one recorded by the authors shows that this proposal enables reliable detection of abnormal gaits. Furthermore, real-time applications are supported since the proposed method requires very low computational power.

The remainder of this paper is organised as follows: Section II summarises existing work in the field and similar methods. Section III details how the proposed feature is extracted as well as the applied dimensionality reduction. Section IV introduces the employed classification technique

Alexandros Andre Chaaraoui and José Ramón Padilla-López acknowledge financial support by the Conselleria d'Educació, Formació i Ocupació of the Generalitat Valenciana (fellowships ACIF/2011/160 and ACIF/2012/064, respectively). This work has been supported by the Ambient Assisted Living Joint Programme and the Technology Strategy Board under project 'BREATHE Platform for self-assessment and efficient management for informal caregivers' (AAL-JP-2012-5-045). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

which is based on the BagOfKeyPoses method. Section V presents the experimental results that have been obtained on two different datasets applying several evaluation methods. Finally, Section VI concludes this paper.

II. Related work

Considering the application of low-cost RGB-D devices, such as the Microsoft Kinect or other similar depth sensors, to gait analysis, a few recent works can be found that address the problem quite differently. Clearly, what can be observed is that the usage of these devices is increasing due to their low cost and low computational requirements in comparison to other 3D body pose estimation methods as marker-based motion capture systems, in addition to the reduced intrusiveness in comparison to wearable sensors. A common application in this field is fall detection. In [9], taking advantage of the provided depth data, fall detection is modelled as two events. First, it is detected that the person is on the ground, and then an ensemble of decision trees is employed to predict if a fall preceded this situation. Experimentation is carried out over extensive long-term data outperforming other state-of-the-art methods. In [10], a gait monitoring method is presented which is based on a robotic system. An autonomous robot follows the person at home at a safe distance and monitors three different metrics to perform the prediction based on skeletal data: stride length, stride duration and centre of mass motion. An extensive comparison to a commercial marker-based motion capture system (Vicon [11]) is carried out, and several advantages are pointed out, such as low cost, ease of use and unlimited capturing volume.

Regarding specifically human gait analysis and monitoring, Stone *et al.* evaluate the accuracy of measuring walking speed, stride time and length based on the depth data from the Kinect, obtaining promising preliminary results [12]. Similarly, in [13] lower and upper body gait parameters, such as stride information and arm kinematics, are extracted from the skeletal data. A machine learning framework is employed to simulate feature signals similar to those that are obtained from gyroscopes. The measured mean difference to the wearable sensors was less than 1%. Abnormal gait detection is targeted in [8], where skeleton data is analysed online. A non-linear manifold learning technique is employed to reduce the dimensionality of the noisy skeleton data. A statistical model is learned from the normal gait samples and the detection is performed based on matching the new observations to the model following Markov assumptions. This method is then applied to detect simulated gait anomalies on subjects which are climbing a flight of stairs. Both event-based and frame-based classification results are reported, where the event-based results are considerably higher due to a refinement stage that detects trend changes. Our method is comparable to this work and, since their dataset is publicly available, a comparison of results is made in the present work.

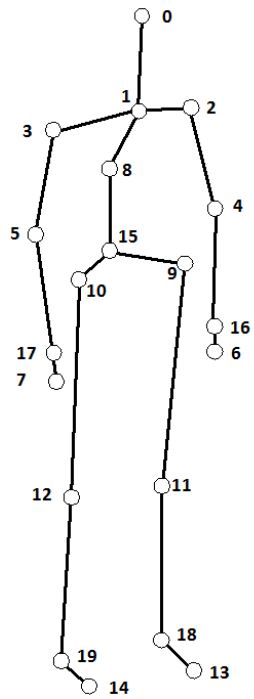

Fig. 1. Sample of a skeleton of $J = 20$ joints obtained with the Microsoft Kinect SDK [14].

III. Joint motion history features

A. Feature extraction

Initially, an input data that consists of skeleton sequences is obtained. For this purpose, an RGB-D camera is employed to capture colour and depth data. Based on markerless human body pose estimation [5], a 3D skeletal representation of J joints is obtained for the person that is being recorded in the field of view (see Fig. 1). These joints are made up of 3D coordinates in real world metrics. Each skeleton S_i is normalised for location, size and rotation as follows:

1) Determine the centroid C_i of the skeleton joints based on the average 3D coordinate of all joints.
2) Determine the normalising length L_i as the average Euclidean distance between the joints and the centroid C_i.
3) Determine the y-axis rotation α_i of the trunk of the skeleton with respect to the camera using the shoulder and hip joints.
4) Set the centroid C_i as origin coordinate.
5) Normalise the size and rotation of the skeleton based on the normalising length L_i and rotation α_i obtained in steps 2 and 3.

These normalised skeletons $\bar{S}_i$ are then accumulated using a sliding window approach. The window has a constant size of τ frames, where the τ involved skeletons are analysed to obtain the joint motion history features.

For each of the skeletons of the window, its normalised joint coordinates are translated to the corresponding 3D coordinates of a volume. This volume is of a size $R_x \times R_y \times R_z$ that should keep the proportions of the skeleton sequences in

order to avoid distortions. The voxel of the equivalent joint coordinates of the volume is set to a value r that indicates the recentness of the motion. The value corresponds to the local index of the skeletons of the sliding window: $\bar{S}_1, \bar{S}_2, \ldots, \bar{S}_\tau$. Therefore, $r = i$, $i \in [1, \tau]$. In this way, the voxels are set to the appropriate values based on the J joints of the skeletons. If a voxel has been hit by multiple joints, the highest r value is maintained. The remaining coordinates of the volume contain a null value that indicates that no motion has been detected. Fig. 2 shows a graphical explanation of this step defined in (1).

$$volume_t(x, y, z) = \begin{cases} \max r & \text{if joint motion is detected,} \\ 0 & \text{otherwise,} \end{cases}$$
$$\forall x \in [0, R_x],\ \forall y \in [0, R_y],\ \forall z \in [0, R_z]. \quad (1)$$

This procedure is performed each λ frames, where $\lambda \in [1, \tau]$. Since τ frames are summarised in each JMH, the feature does not have to be computed for each frame. This makes it possible to set up the appropriate balance between computational cost and temporal feature coverage. Note that higher values of λ do not necessarily lead to worse performance, since also noise and undesired motion details are represented.

B. Dimensionality reduction

At this point, the JMH feature consists of a 3D volume that encodes the skeleton joints' motion over a segment of frames. The dimensionality of this feature is dependent on the chosen scale of the volume. A higher scale will represent motion details accurately, whereas a lower scale will suppress such a level of detail and show more significant posture changes. Since our application of gait analysis and abnormal gait recognition for frailty detection requires an increased level of detail, the scale of the feature may increase dramatically. For that reason, a dimensionality reduction method based on axis projection is proposed for the JMH feature. This is similar to the three orthogonal planes approach (TOP) from [15] for regular video data, but using depth as the third axis instead of time.

Specifically, the previously obtained volume is projected onto its three orthogonal axes in order to obtain a front, a side and a bottom view. Instead of the value of the closest active coordinate to the null value of the axis, the highest value along the summarised axis is employed to obtain the three views, since this allows to prevail the most recent motion and to obtain symmetric views. For instance, the following equation is applied for the front view that is obtained at a temporal instant t:

$$front_t(x, y) = \max_z\ volume_t(x, y, z),$$
$$\forall x \in [1,\ R_x]\ \wedge\ \forall y \in [1,\ R_y], \quad (2)$$

where $z \in [1,\ R_z]$.

The side view $side_t$ and the bottom view $bottom_t$ are obtained similarly, summarising respectively the X and Y axes. Fig. 3 shows an example of the three projections that are obtained using this method.

Finally, the three views are considered as one-dimensional arrays enabling to apply feature fusion based on a concatenation operator in order to generate a single feature descriptor:

$$jmh_t = front_t \parallel side_t \parallel bottom_t. \quad (3)$$

IV. Classification

As classification algorithm, the BagOfKeyPoses method has been chosen. This method handles classification of data with sequential or temporal relation. Any type of feature expressed as an array of double values can be used. This data has to be acquired in an ordered fashion, so that a sequence of features with a meaningful order can be obtained. Therefore, this machine learning method fits into the needs of this work, where sequences of JMH features have to be classified to detect abnormal behaviours.

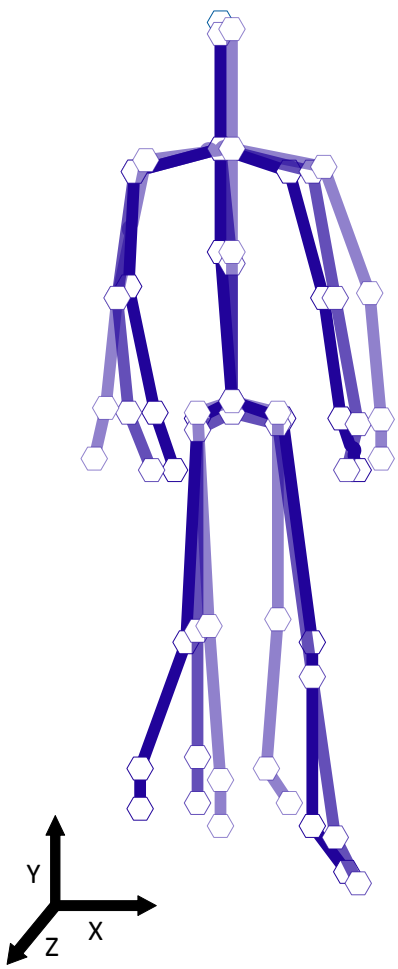

Fig. 2. This figure shows the superposition of normalised skeletons that is applied to obtain the 3D volume that corresponds of a window of τ frames (for visual purposes a low value of $\tau = 3$ has been chosen and the segments between joints are shown).

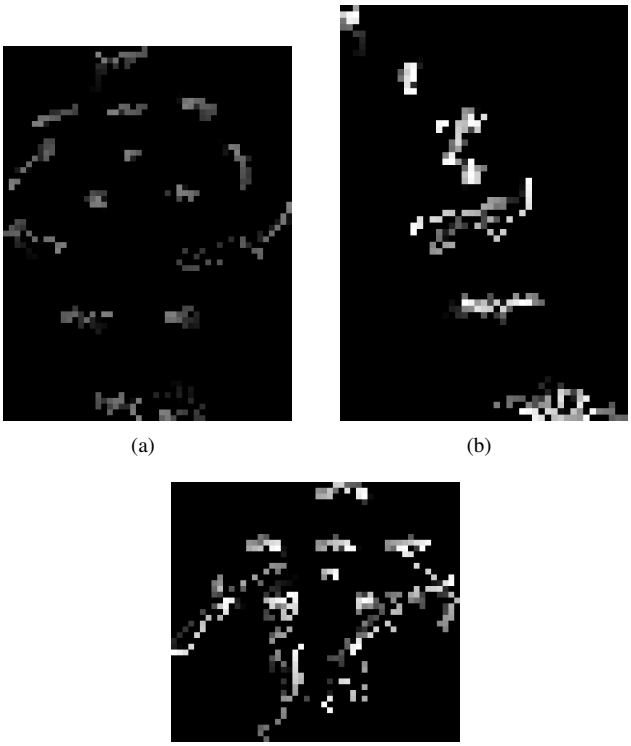

Fig. 3. Sample of front (a), side (b) and bottom (c) views that make up a specific instance of the JMH feature.

The method handles the learning of a bag of key poses model to represent the most common feature instances present in the learning data, similarly to the bag of words (BoW) approach, but without the quantisation step. The temporal relation between key poses is also modelled by means of learning templates of sequences of key poses. For this purpose, the features are substituted with their nearest neighbour key pose out of the bag of key poses model. The test sequences are then aligned to these previously learned sequences using dynamic time warping (DTW) and Manhattan distance for feature comparison. More details can be found in [16] and [17], where it can be observed that this method shows to be proficient for the classification of human behaviour based on spatio-temporal characteristic data as, for instance, for human action or activity recognition. Furthermore, it has also been used with skeletal pose estimations leading to outstanding results [18].

Note that since BagOfKeyPoses is designed as a multiclass classification algorithm, it assumes that the test data belongs to one of the learned classes. In the present work, only the normal human gait is learned and then compared to the test data to detect abnormal behaviours. Since abnormal motion present a great variance, it is preferable to detect deviations of a learned normal model than to aim to learn the different type of anomalies, which may be unknown. The method can easily be adjusted to this type of learning scheme by leaving out the key pose substitution step for unknown sequences. In this way, the normal or abnormal feature sequences from the test set are directly matched to the previously learned normal sequences from the training set and recognition is performed based on thresholding.

V. Experimentation

A. SPHERE Dataset

In order to evaluate this new human motion feature, the publicly available gait analysis dataset [8] from the SPHERE project from the University of Bristol has been employed. It provides a front view of recordings that show an individual walking up a flight of stairs. A set of 17 sequences from six healthy subjects having no injury or disability are provided to learn the normal gait. A different set of another six subjects and a total of 31 sequences are used for testing purposes. Normal, but also abnormal behaviours which are common for knee injuries are provided. Three types of anomalies are included: 1) sequences that include a short stop where the subject stands still and then continues to walk, 2) sequences where the subject walks up the stairs always using the right leg initially (right leg lead), and 3) sequences where each stairstep is started with the left leg (left leg lead). These anomalies are performed by a physiotherapist and five other subjects that followed his instructions. Skeletal data of 15 joints is provided based on the Microsoft Kinect RGB-D camera and the OpenNI skeleton tracker [19].

The JMH parameters have been set empirically to $\tau = 35$ frames and $\lambda = 7$ frames. Based on the typical proportions of normalised skeletons a volume size of $50 \times 65 \times 45$ has been established. The BagOfKeyPoses classification method has been used with its default parameters, $K = 10$ key poses, and relying on the freely available open source library [20]. An abnormal sequence is detected when the DTW distance of the best matching sequence is above the threshold of $dist = 9.2$ per element.

The first results that are stated in [8] are related to the method's ability to generalise normal gait. For that purpose the 14 normal sequences from unlearned actors are evaluated, and with their proposal one false positive is obtained. On the other hand, the different types of abnormal sequences are tested. In this case, four false negatives and two false positives are stated out of 58 abnormal occurrences, *i.e.* 93% of abnormal events are detected.

We follow the proposed evaluation method and the corresponding train and test sets for classification. However, due to our classification method, sequence-based recognitions are obtained. With the detailed parametrisation up to 100% accuracy is achieved detecting both normal and abnormal sequences correctly.

Table I shows a comparison of results based on the F_1-Score. In order to show the contribution of the newly proposed JMH feature, the same test has been carried out using the normalised skeletons $\bar{S}_i$ directly. Although these have been used previously for human action recognition leading to very accurate results, it can be observed that the current proposal presents clear advantages to detect temporal details of human motion, such as the gait cycle changes of the present anomalies.

TABLE I

EXPERIMENTAL RESULTS OBTAINED ON THE SPHERE DATASET USING SEQUENCE-BASED OR EVENT-BASED EVALUATION.

Method	$\boldsymbol{F_1}$**-Score**
Skeleton data + Manifold learning [8]	0.94[a]
Normalised skeletons + BagOfKeyPoses	0.88
JMH + BagOfKeyPoses	**1**

[a] The score has been obtained based on the combination of the evaluation of normal and abnormal sequences.

TABLE II

EXPERIMENTAL RESULTS OBTAINED ON DGD USING SEQUENCE-BASED EVALUATION AND TWO DIFFERENT EVALUATION METHODS. IN THIS CASE, THE JMH PARAMETERS HAVE BEEN ESTABLISHED EMPIRICALLY TO $\tau \in [37, 42]$, $\lambda \in [18, 20]$ AND $dist \in [4.1, 5.3]$

Method	$\boldsymbol{F_1}$**-Score (a)**	$\boldsymbol{F_1}$**-Score (b)**
Normalised skeletons + BagOfKeyPoses	0.95	0.83
JMH + BagOfKeyPoses	**0.98**	**0.85**

B. DGD: DAI gait dataset

In the context of our gait analysis research project, we have also recorded our own dataset for abnormal gait detection based on RGB-D devices. In this case, the recently released Microsoft Kinect 2 camera is employed recording a front view of a corridor, where the person walks towards the camera. Seven subjects were asked to walk normally as well as simulating abnormal gait. For this purpose, we have established two anomalies, a knee injury that implies that the knee cannot be bent but otherwise gait is normal, and a second one where one foot is dragged towards the other, which usually happens when a mobility aid such a crutch or a handrail is employed. These scenarios were performed for the right and the left leg, leading to four different abnormal gait types. These four abnormal samples as well as four normal instances were recorded for each of the seven subjects making up a total of 56 sequences. With the improved Microsoft Kinect 2 sensor, a 1080p colour image and 512×424 depth image are obtained out of which a body pose estimation of 25 joints is generated. We employ 20 of the available joints, discarding the fingers and a redundant joint of the torso (*SpineShoulder*). This dataset is publicly available upon request, however an extended version based on real patients will be released in near future.

Since the recordings are based on individually segmented sequences, a sequence-based evaluation is proposed. Two different evaluation methods have been defined:

(a) A training set of three out of four normal gait sequences of each actor, and a testing set with the remaining normal and abnormal samples.
(b) A training set made up of the normal samples from four actors, and a testing set of normal and abnormal samples of the remaining three actors, performing therefore a cross-subject validation.

The results reported in Table II show that this type of data is handled especially well by the proposed method leading to very high classification scores. As expected, the cross-subject evaluation presents a higher difficulty where three false positives and one false negative are obtained, whereas if the normal gait of the corresponding subjects has been learned previously, only a single false negative is generated. Nonetheless, the cross-subject scenario is closer to a real test scenario, since normally the gait assessment will occur once a possible anomaly is expected or the health of the person is deteriorating, even if a continuous monitorisation is more desirable to detect deviations. It can also be observed, that although the proposed JMH features still outperform the direct usage of skeletal data, the difference is less significant in comparison to the more challenging SPHERE dataset.

C. Temporal evaluation

The method's performance has been evaluated on an Intel i7-2600 CPU at 3.4 (GHz) with 8 (GB) of RAM. Without any parallelisation nor hardware-specific optimisation, a recognition rate of over 9000 (FPS) is reached. The recognition time is especially low due to the temporal integration of data of the JMH feature which is obtained each λ frames. This can be observed if the performance is compared to the rate that is obtained using the frame-wise skeletons. Feature extraction of JMH features takes twice as much time (0.4×), but as a result the recognition time is almost ten times as fast (9.3×) even though the feature dimensionality is significantly higher (400×). This shows that the present approach is particularly suitable for real-time detection of human gait anomalies at people's homes or care centres.

VI. CONCLUSIONS

In this paper, a method for abnormal gait detection has been presented. Based on RGB-D devices and introducing a novel joint motion history feature, reliable detection can be performed using a generic machine learning method that is able to classify temporal feature sequences. The 3D skeletal data which consists of body joints is tracked over a number of frames and combined in a 3D volume prevailing recent over past motion. The spatio-temporal information is then reduced to its three orthogonal projections for dimensionality reduction. Sequences of these spatio-temporal features are learned based on a bag of key poses model. New observations are aligned to the normal gait that has been learned in order to detect deviations. In the proposed scenario, people are asked to walk towards the depth sensor. A real-time classification regarding normal or abnormal gait is then obtained and the thresholded distance to the normal model can be employed for gait assessment related to gait quality and deviation over time. This technique can significantly support geriatricians performing gait analysis in care centres or nurseries. As it has been discussed previously, this is especially interesting for gait monitoring, rehabilitation and even early diagnosis of cognitive impairment. Furthermore, the proposed method presents intrinsic advantages such as low cost, less intrusiveness than wearable devices and simple setup requirements.

For future work, several new research directions can be identified. Regarding the feature proposal, it can be observed that all joints are considered indistinctly. However, feet motion is, for instance, not equivalent to knee motion. Since the origin joint of each motion is known, each joint's motion could be compared individually. This may require to apply further dimensionality reduction techniques (*e.g.* grid histograms can be applied to the projections). Similarly, when comparing joint motion history features, the neighbourhood of each pixel should be considered, since close coordinates can represent the same motion. Low distances should be obtained for similar, but displaced areas, whereas currently high distances are obtained if the motion's location changes slightly. Finally, further experimentation is required. More extensive datasets and evaluation methods have to be applied, and more importantly, data from patients such as injured and elderly people has to be collected. In this sense, the ground truth data from the geriatrician could also include types of anomalies and their associated degrees of frailty, since this would make it possible to significantly improve current abnormal gait detection proposals providing more valuable metrics. Last but not least, joint motion history features can describe very different motion cues depending on the parametrisation. Multi-scale feature extraction could be useful for complex behaviours in order to capture both low-and high-level motion cues. To this extent, JMH can be employed for other applications of human motion analysis, such as the recognition of actions or activities of daily living.

REFERENCES

[1] EC, "Active ageing special eurobarometer 378," tech. rep., DG COMM "Research and Speechwriting" Unit, European Comission, 2012. Conducted by TNS Opinion & Social at the request of Directorate-General for Employment, Social Affairs and Inclusion.

[2] J. Hausdorff, G. Yogev, S. Springer, E. Simon, and N. Giladi, "Walking is more like catching than tapping: gait in the elderly as a complex cognitive task," *Experimental Brain Research*, vol. 164, no. 4, pp. 541–548, 2005.

[3] N. E. Morone, K. Z. Abebe, L. A. Morrow, and D. K. Weiner, "Pain and decreased cognitive function negatively impact physical functioning in older adults with knee osteoarthritis," *Pain Medicine*, vol. 15, no. 9, pp. 1481–1487, 2014.

[4] V. J. Verlinden, J. N. van der Geest, A. Hofman, and M. A. Ikram, "Cognition and gait show a distinct pattern of association in the general population," *Alzheimer's & Dementia*, vol. 10, no. 3, pp. 328 – 335, 2014.

[5] J. Shotton, A. W. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *IEEE Conference on Computer Vision and Pattern Recognition, 2011. CVPR 2011*, pp. 1297 – 1304, 2011.

[6] M. W. Whittle, "Clinical gait analysis: A review," *Human Movement Science*, vol. 15, no. 3, pp. 369 – 387, 1996.

[7] A. Bobick and J. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257 – 267, 2001.

[8] A. Paiement, L. Tao, M. Camplani, S. Hannuna, D. Damen, and M. Mirmehdi, "Online quality assessment of human motion from skeleton data," in *Proceedings of the British Machine Vision Conference*, BMVA Press, 2014.

[9] E. Stone and M. Skubic, "Fall detection in homes of older adults using the Microsoft Kinect," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, pp. 290–301, Jan 2015.

[10] A. Staranowicz, G. R. Brown, and G.-L. Mariottini, "Evaluating the accuracy of a mobile Kinect-based gait-monitoring system for fall prediction," in *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments*, PETRA '13, (New York, NY, USA), pp. 57:1–57:4, ACM, 2013.

[11] "Vicon Motion capture system." http://www.vicon.com/.

[12] E. Stone and M. Skubic, "Evaluation of an inexpensive depth camera for in-home gait assessment," *Journal of Ambient Intelligence and Smart Environments*, vol. 3, no. 4, pp. 349–361, 2011.

[13] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, "Full body gait analysis with Kinect," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 1964–1967, Aug 2012.

[14] "Microsoft Kinect SDK 2.0, Microsoft, Inc.," 2014.

[15] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 915–928, June 2007.

[16] A. A. Chaaraoui, P. Climent-Pérez, and F. Flórez-Revuelta, "Silhouette-based human action recognition using sequences of key-poses," *Pattern Recognition Letters*, vol. 34, no. 15, pp. 1799–1807, 2013. Smart Approaches for Human Action Recognition.

[17] A. A. Chaaraoui, J. R. Padilla-López, F. J. Ferrández-Pastor, M. Nieto-Hidalgo, and F. Flórez-Revuelta, "A vision-based system for intelligent monitoring: Human behaviour analysis and privacy by context," *Sensors*, vol. 14, no. 5, pp. 8895–8925, 2014.

[18] A. Chaaraoui, J. Padilla-Lopez, and F. Florez-Revuelta, "Fusion of skeletal and silhouette-based features for human action recognition with RGB-D devices," in *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pp. 91–97, Dec 2013.

[19] "OpenNI 2 SDK, Structure Sensor," 2014.

[20] A. A. Chaaraoui, P. Climent-Pérez, and F. Flórez-Revuelta, "BagOfKeyPoses Machine Learning Method - Open Source Code (Apache 2.0 License)," Oct. 2014. https://github.com/DAIGroup/BagOfKeyPoses.