Sensing Behaviour using the Kinect: Identifying Characteristic Features of Instability and Poor Performance during Challenging Balancing Tasks

Daniel Leightley[1], Moi Hoon Yap[2], Brett M. Hewitt[2] and Jamie S. McPhee[3]

[1]Centre for Military Health Research, King's College, London, England, SE5 9RJ

[2]School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, M15 6BH

[3]School of Healthcare Science, Manchester Metropolitan University, M15 6BH

dleightley@ieee.org, m.yap@mmu.ac.uk, b.hewitt@mmu.ac.uk, j.s.mcphee@mmu.ac.uk

Abstract

A framework is proposed to utilise the Microsoft Kinect One, a low-cost, unobtrusive and reliable sensing device to analyse the apprehension that exists in human motion to aid clinical decision making. We extract kinematic features, Centre-of-Mass and Body Movement Zone to determine the spatial-temporal directional changes for balancing tasks. This framework includes a pipeline that obtains the Kinect skeletal data and applies techniques to interpret the information to determine the steadiness or instability, hence providing an objective indicator of the difficulty the participant has completing the task. We conclude with a summary of our initial findings, and highlight areas for further research.

Introduction

In recent years, there has been an interest in digitalised methods for detection, analysis and quantification of human motion. This is due to the increased availability of low-cost multi-modality marker-less capturing devices [1]–[3]. The release of the latest edition in the Microsoft Kinect (henceforth Kinect) series, Kinect One, has enabled a new and improved immersive gaming experience. Kinect can be utilised for additional applications other than gaming or entertainment, most notably within the medical domain. Bigy *et al*. [4] proposed a technique for recognising posture and Freezing of Gait in those with Parkinson's disease to aid in detecting trips and falls within the home. Leightley *et al*. [5] introduced a classification framework to detect motions commonly found in a rehabilitation setting to support patient wellbeing. Yang *et al*. [6] implemented a framework that extracts both depth and colour image data from the Kinect to assess the posture of participants when performing standing balance, the framework allowed for detection of subtle directional changes such as postural sway.

The Kinect series has been utilised in a range of medical applications, to support this several studies have sought to validate the Kinect and its viability in the medical domain. Clark *et al*. [7] (revisited in [8]) captured a large number of participants performing a series of balance motions consisting of single and double limb support. Kinect and marker-based Vicon data were captured concurrently, with data from both systems filtered and synchronised. The Kinect was found to be highly robust and accurate when compared to the '*gold-standard*' of the Vicon. Mentiplay *et al*. [9] assessed the validity of the Kinect in tracking gait and its inter-day reliability when compared to 3DMA marker-based camera system. The authors found that while the Kinect is not suitable for tracking lower body kinematic data, it is capable of measuring spatiotemporal aspects of gait. Gonzalez *et al*. [10] implemented a framework that combined the Kinect and Wii Balance board to extract Centre-of-Mass (CoM), the authors found a strong correlation between the two devices and were able to assess standing and walking motions amongst a small population sample. These works detail the validity and clinical feasibility of the Kinect to assess kinematic strategies related to gait and posture.

In this work, we propose a new framework for *sensing the behaviour* of motion with the Kinect to identify difficulty people have performing balancing tasks. Progressing from previous studies [3], [11], we utilise CoM and Body Movement Zone (BMZ), coupled with the temporal domain to provide a low latency outcome for potential utilisation in *real world* situations. This will provide greater insight for practitioners and allow us to identify when a participant is unsteady, in distress or finding the motion troublesome to perform/execute. This would allow the practitioner to implement coping strategies or develop interventions.

Sensing behaviour with Kinect

Several works have been proposed to measure and analyse behaviour present in human motion. The most prevalent methodology suggests using one or a combination of intrusive sensors, such as body-based accelerometer or reflective-markers. Then, the clinically relevant indicators are extracted by analysing the patterns presented in time series data generated by these data mediums [12]. In recent years, the Computer Science community has proposed an array of solutions. These works have predominantly focused on depth sensor technology, which has been shown sufficiently accurate and responsive for tracking in both in-home and medical settings [5], [10].

These frameworks follow a similar structure. They first seek to identify the human motion, using motion classification techniques, and then undertake quantitative analysis and feature extraction techniques on the motion to provide greater understanding. Depth-sensors have been used for the assessment of balance by extracting gait-based features such as CoM from a skeletal stream to provide mobility indicators (*e.g.* [6], [8], [13]–[16]). For example, Dolatabadi *et al*. [14] proposed a home-based system for assessing changes in gait and balance using kinematic features such as CoM. The authors utilise a Microsoft Kinect sensor to observe gait recovery in a participant that had undergone surgery. They were able to track the gait changes over a number of weeks, helping to inform clinical judgements based on the information extracted. Typical gait patterns require stability standing on one-leg which is a task that becomes increasingly difficult for people with movement impairments, yet very little research has focussed on sensing stability.

The aim of the present study was to develop automated methods to sense the instability in human motion using depth sensor technology, CoM and body-movement-zone including a wide range of participant characteristics. Our work relies on several key computer vision and pattern recognition techniques such as stereovision, pose estimation, and feature extraction/representation.

Identifying Behaviour and Kinect: Framework

An overarching framework for analysing human motion obtained from the Kinect to identify unease in motion is proposed (see Figure 1). This section will introduce the framework methodology, utilisation and implementation for possible *real world* deployment. The analysis framework has been implemented in Matlab 2014a and all source code is openly available at (Link upon acceptance).

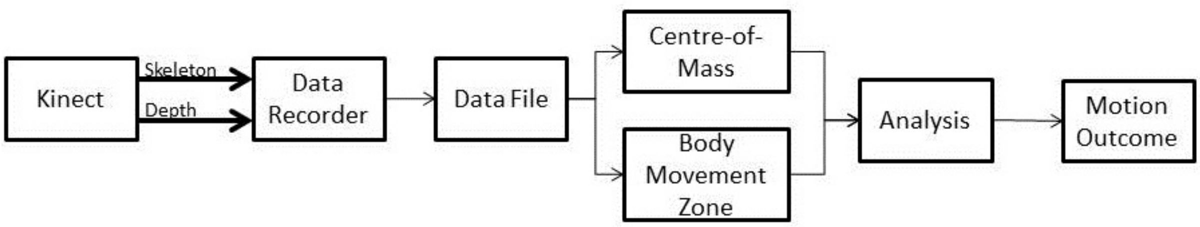

Figure 1 - Overarching framework for analysing human motion collected from the Microsoft Kinect to identify movement hesitation.

Table 1: Participant demographics used in this study, extracted from the K3Da Dataset [17].

Variables	Young (≤59yrs, n=5)	Old (≥60yrs, n=5)	Athletic Old (≥60yrs, n=5)
Age	25.6 (7.03)	73.4 (3.36)	64.8 (6.34)
Height	178.6 (8.26)	172.8 (4.54)	172.8 (4.54)
Weight	75.2 (12.07)	79.4 (21.04)	62 (10.93)
Sex (male %)	70%	40%	60%
BMI	20.7 (3.63)	26.46 (7.12)	21.9 (1.56)

In this work, the newly released K3Da Dataset [17] was employed. The dataset consists of more than 50 participants performing a range of clinically relevant movements captured by a Kinect One. The dataset is composed of three participant groups, young (≤59yrs), old (≥60yrs) and athletic old (≥60yrs). In this work, fifteen participants were randomly selected (see Table 1 for demographic information), five from each group with the following motions extracted: One-Leg Balance (Eyes Open) and One-Leg Balance (Eyes Closed), providing a total of 52 motions. Each trial was independently assessed by two coders (using skeletal and video streams) and categorised as follows: "Stable and successful", "unstable and successful" and "unstable and unsuccessful", providing ground truth information to validate the proposed framework. The following definitions were used when determining category:

- *Stable and successful*: The participant had very little body movement and was able to complete the balancing task.
- *Unstable and successful*: The participant had high amounts of extraneous body movement, but was able to complete the balancing task.
- *Unstable and unsuccessful*: The participant had high amounts of extraneous body movement and was unable to complete the balancing task.

The first coding iteration resulted in an inter-reliability of 0.84. In the second iteration, the coders consulted one-another to reach final agreement on the rating of trials. The final categorisation composition can be found in Table 2.

Table 2: Categorisation labelling for each participant group.

Categorisation	Young	Old	Athletic Old
Stable and successful	9	4	6
Unstable and successful	5	2	3
Unstable and unsuccessful	1	13	9

To identify the level of stability, each motion sequence was analysed **independently**, hence only the motion itself is evaluated and not dependent on any other motion that the participant (or others) had performed. While the Kinect provides a skeletal stream, similar to motion capture from a marker-based system (such as Vicon), it does so without the need for placing markers on the participants body. The Kinect utilises depth sensor technology to identify key anatomical landmarks on the body that are tracked over time (up to 30Hz). In this work, we decompose the Kinect skeletal stream into two time-series parameters, CoM and BMZ to describe the participants' behaviour at each time period (frame). This enables an overall determination of the behaviour and performance of the participant.

Using three joint locations provided by the Kinect (centre of the hip, left shoulder and right shoulder), while on their own they can provide descriptive information about the motion we seek to derive a single variable that is powerful enough to describe the motion. The CoM [10], [13] is encoded for each frame as the average point between the aforementioned joints. The change in position between consecutive frames is the directional change over time. Therefore, for each period of time we are able to identify the position, and the temporal directional movement of the CoM.

In this work, we introduce the BMZ, a parameter that describes the total volume of space occupied by the participant when in motion (see Figure 2 for a visual example). This is computed by identifying the total space covered (or occupied) by the participant's skeleton when undertaking a motion per frame using standard volume calculations. For example, if the participant is stable, with little movement, the BMZ variable is small, whereas with large variations in motion such as raising of the arm the size of the overtime BMZ increases.

To determine the behaviour exhibited by the participant, the CoM and BMZ are firstly assessed individually, and then assessed together to determine an outcome. For each parameter, the intra-framed variation is computed (*e.g.* difference between frame 1 and frame 2), using this we are able to generate the percentage difference between the current and last frame. Hence, we can represent the CoM and BMZ parameters with a percentage change value over time (*e.g.* if we have 300 frames, for each parameter we will have 300 percentage values). Therefore, we are able to identify sudden and large motion variations that are typically present when the participant has encountered difficultly when undertaking a motion. Each frame (represented by the percentage value for both CoM and BMZ) is categorised as follows, based on our experience working in the field and other similar works, if the percentage is:

1. Less than 30% it is labelled as ***Stable and successful***.

2. Between 31% and 70% labelled as ***Unstable and successful***.

3. Greater than 71% labelled as ***Unstable and unsuccessful***.

Having represented each parameter by a set of labels representing the state of the motion at a specific time period, these are aggregated; the category with the most "votes" is identified as the classification for the motion.

Figure 2 - Example of a skeleton within a Body Movement Zone. Defined as the space occupied by the participant for each frame. Observe the entire sequence (represent by blue) and the surrounding box (represented by black), which is the total space occupied by the participant.

Identifying Behaviour

The ground truth labels derived in this work are used to evaluate the proposed framework for the task of detecting unease in motion performance across a range of participants. For evaluation, random samples of participants were identified, representing a large range of demographics profiles. The results were repeated five times to ensure reproducibility in the coding – no anomalies were identified. Each motion was assessed individually and compared to the ground truth labelling to determine the success of the framework.

Table 3: Classification accuracy for each participant group based on categorisation.

	Young (accuracy)	**Old (accuracy)**	**Athletic Old (accuracy)**
Stable and successful	9 (0.96)	4 (0.89)	6 (0.93)
Unstable and successful	5 (0.97)	2 (0.91)	3 (0.94)
Unstable and unsuccessful	1 (0.90)	13 (0.93)	9 (0.92)
Average:	**0.94**	**0.91**	**0.93**

The framework provided a high degree of accuracy when identifying mobility of participants who had been identified as finding the motion difficult during ground truth labelling (see Table 3). Overall, the framework performed robustly with a classification accuracy of 0.92, which is considerably high for the type of data being classified. Confusion observed across the participant range low, with a high rate of true positives, and a small rate of false positives and a robust overall sensitivity of 0.94, specificity of 0.97 and Matthew Correlation Coefficient score of 0.97. There was no significant difference between One Leg Balance (Eyes Open) and One Leg Balance (Eyes Closed), with an overall accuracy of 0.95 and 0.95 respectively.

Discussion and conclusion

There is little doubt to the benefits of utilising the Kinect sensor for use in monitoring, quantify and evaluating human motion. This work has relied upon the K3Da dataset, which has provided a large number of clinically relevant motions and a dives range of participants, however it is not without its challenges. First, range of the sensor is limited – only capering motions within a 4 metre by 4 metre areas. Second, Joint occlusion and sensor accuracy varied for each participant, and the types of motion being undertaken. However, when the motions were coded the coders had access to the video stream and not the skeleton alone – allowing for greater interpretation. Conversely, the framework only utilised skeletal information yet was able to identify motion stability with a high degree of accuracy.

In this work, we have sought to propose a framework for identifying instability in motion, allowing for a practitioner to intervene and identify what may lie behind the instability. Only a single outcome is provided which is able to provide an overview of what is being observed. However, it would be more helpful to identify specific temporal dynamics to improve the clinical decision making process. Further, we have utilised two motion groups, One-Leg Balance (Eyes Open) and One-Leg Balance (Eyes Closed). Future work will explore the possibility of adapting the framework to provide detailed joint-level temporal kinematic insight; hence practitioners can focus their efforts in improving confidence in these areas.

References

[1] Y.-J. Chang, S.-F. Chen, and J.-D. Huang, "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities," *Res. Dev. Disabil.*, vol. 32, no. 6, pp. 2566–2570, 2011.

[2] A. Ejupi, M. Brodie, Y. J. Gschwind, S. R. Lord, W. L. Zagler, and K. Delbaere, "Kinect-Based Five-Times-Sit-to-Stand Test for Clinical and In-Home Assessment of Fall Risk in Older People," *J. Gerontol.*, 2015.

[3] A. Kargar, A. Mollahosseini, T. Struemph, W. Pace, R. Nielsen, and M. Mahoor, "Automatic measurement of physical mobility in Get-Up-and-Go Test using kinect sensor," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014.

[4] A. Amini Maghsoud Bigy, K. Banitsas, A. Badii, and J. Cosmas, "Recognition of postures and freezing of gait in Parkinson's disease patients using Microsoft Kinect Sensor," in *IEEE Conference on Neural Engineering*, 2015, pp. 731–734.

[5] D. Leightley, J. Darby, B. Li, J. S. Mcphee, and M. H. Yap, "Human Activity Recognition for Physical

Rehabilitation," in *International Conference on Systems, Man and Cybernetics*, 2013.

[6] Y. Yang, F. Pu, Y. Li, S. Li, Y. Fan, and D. Li, "Reliability and Validity of Kinect RGB-D Sensor for Assessing Standing Balance," *Sensors J.*, pp. 1633–1638, 2014.

[7] R. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. Webster, L. Denehy, and A. Bryant, "Validity of the Microsoft Kinect for assessment of postural control," *Gait Posture*, vol. 36, no. 3, pp. 372–377, Jul. 2012.

[8] R. Clark, Y.-H. Pua, C. Oliveira, K. Bower, S. Thilarajah, R. McGaw, K. Hasanki, and B. Mentiplay, "Reliability and concurrent validity of the Microsoft Kinect V2 for assessment of standing balance and postural control," *Gait Posture*, 2015.

[9] B. F. Mentiplay, L. G. Perraton, K. J. Bower, Y. H. Pua, R. McGaw, S. Heywood, and R. A. Clark, "Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables," *J. Biomech.*, vol. 48, no. 10, pp. 2166–2170, 2015.

[10] A. Gonzalez, M. Hayashibe, and P. Fraisse, "Estimation of the center of mass with Kinect and Wii balance board," in *International Conference on Intelligent Robots and Systems*, 2012, pp. 1023–1028.

[11] W. van Teijlingen, E. L. van den Broek, R. Konemann, and J. G. M. Schavemaker, "Towards Sensing Behavior Using the Kinect," in *Proceedings of Measuring Behavior*, 2012.

[12] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, "Sensing, Understanding, and Shaping Social Behavior," *IEEE Trans. Comput. Soc. Syst.*, vol. 1, no. 1, pp. 22–34, 2014.

[13] B. Lange, C.-Y. Chang, E. Suma, B. Newman, A. S. Rizzo, and M. Bolas, "Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor," in *International Conference of Engineering in Medicine and Biology Society*, 2011, pp. 1831–1834.

[14] E. Dolatabadi, B. Taati, G. S. Parra-Dominguez, and A. Mihailidis, "A Markerless Motion Tracking Approach To Understand Changes In Gait And Balance: A Case Study," in *ehabilitation Engineering and Assistive Technology Society of North America*, 2013.

[15] N. Vernadakis, V. Derri, E. Tsitskari, and P. Antoniou, "The effect of Xbox Kinect intervention on balance ability for previously injured young competitive male athletes: A preliminary study," *Phys. Ther. Sport*, vol. 15, no. 3, pp. 148–155, 2014.

[16] O. Celiktutan, A. C. Burak, C. Wolf, and B. Sankur, "Graph-Based Analysis of Physical Exercise Actions," in *ACM Multimedia Workshop on Multimedia Indexing and Information Retrieval for Healthcare*, 2013.

[17] D. Leightley, M. H. Yap, Y. B. J. Coulson, and J. S. Mcphee, "Benchmarking Human Motion Analysis Using Kinect One: an open source dataset," in *IEEE Conference of Asia-Pacific Signal and Information Processing Association*, 2015.