Annals of Biomedical Engineering, Vol. 45, No. 12, December 2017 (© 2017) pp. 2784–2793
DOI: 10.1007/s10439-017-1911-8

The Automated Assessment of Postural Stability: Balance Detection Algorithm

ALESSANDRO NAPOLI,[1] STEPHEN M. GLASS,[2] CAROLE TUCKER,[1,2] and IYAD OBEID[1]

[1]Department of Electrical & Computer Engineering, Temple University, Philadelphia, PA 19122, USA; and [2]Department of Physical Therapy, Temple University, Philadelphia, PA 19140, USA

(*Received 17 July 2017; accepted 27 August 2017; published online 30 August 2017*)

Associate Editor Karol Miller oversaw the review of this article.

Abstract—Impaired balance is a common indicator of mild traumatic brain injury, concussion and musculoskeletal injury. Given the clinical relevance of such injuries, especially in military settings, it is paramount to develop more accurate and reliable on-field evaluation tools. This work presents the design and implementation of the automated assessment of postural stability (AAPS) system, for on-field evaluations following concussion. The AAPS is a computer system, based on inexpensive off-the-shelf components and custom software, that aims to automatically and reliably evaluate balance deficits, by replicating a known on-field clinical test, namely, the Balance Error Scoring System (BESS). The AAPS main innovation is its balance error detection algorithm that has been designed to acquire data from a Microsoft Kinect® sensor and convert them into clinically-relevant BESS scores, using the same detection criteria defined by the original BESS test. In order to assess the AAPS balance evaluation capability, a total of 15 healthy subjects (7 male, 8 female) were required to perform the BESS test, while simultaneously being tracked by a Kinect 2.0 sensor and a professional-grade motion capture system (Qualisys AB, Gothenburg, Sweden). High definition videos with BESS trials were scored off-line by three experienced observers for reference scores. AAPS performance was assessed by comparing the AAPS automated scores to those derived by three experienced observers. Our results show that the AAPS error detection algorithm presented here can accurately and precisely detect balance deficits with performance levels that are comparable to those of experienced medical personnel. Specifically, agreement levels between the AAPS algorithm and the human average BESS scores ranging between 87.9% (single-leg on foam) and 99.8% (double-leg on firm ground) were detected. Moreover, statistically significant differences in balance scores were not detected by an ANOVA test with alpha equal to 0.05. Despite some level of disagreement between human and AAPS-generated scores, the use of an automated system yields important advantages over currently available human-based alternatives. These results underscore the value of using the AAPS, that can be quickly deployed in the field and/or in outdoor settings with minimal set-up time. Finally, the AAPS can record multiple error types and their time course with extremely high temporal resolution. These features are not achievable by humans, who cannot keep track of multiple balance errors with such a high resolution. Together, these results suggest that computerized BESS calculation may provide more accurate and consistent measures of balance than those derived from human experts.

Keywords—Mild traumatic brain injury, Concussion detection, Field-expedient balance test, Automated BEES, Automatic balance error scoring detection, Kinect, Return-to-duty evaluation, On-field automatic balance detection.

INTRODUCTION

The incidence of mild traumatic brain injury (mTBI), concussion, and musculoskeletal injuries has increased in the patient population of the Department of Defense (DoD) and the Veterans Health Administration (VHA) as a result of injuries in military and combat operations.[9,17,34] Such injuries cause a substantial number of deaths and can lead to temporary or permanent disability. Despite their clinical relevance, many injuries are still unreported and this matter is further complicated by the limited sensitivity and reliability of current on-field clinical tests.[23,31] With the increased attention on recognition of neuromusculoskeletal injuries, there is a strong need for new assessment tools to help evaluate these injuries onsite, in non-clinical environments, more effectively and in a timely manner.

In on-field situations, balance is a commonly used indicator of mild traumatic brain injury (mTBI), concussion, and musculoskeletal injury.[14,16,32] To measure balance, a number of standardized screening tools are

Address correspondence to Alessandro Napoli, Department of Electrical & Computer Engineering, Temple University, Philadelphia, PA 19122, USA. Electronic mail: a.napoli@temple.edu

becoming prevalent in sideline/on-field balance assessment, replacing routine physical and clinical exams.[6,15,29] These on-field evaluations aim to provide a relatively brief assessment for determining whether a potentially injured service member or athlete is suitable to return to duty

The most commonly used clinical balance assessment tool following concussion is the Balance Error Scoring System (BESS).[16,28] The BESS test measures static postural stability and it is typically administered by trained medical personnel who must observe and count on a 0–10 scale, specific behaviors corresponding to deficits in postural control while simultaneously spotting the subject to prevent falls. The subject under test is required to maintain balance with eyes closed and hands on hips in three stance conditions: double-leg, single-leg and tandem stance. Each stance is performed on two surface types, hard ground (DS, SS, TS) and on a foam pad (DF, SF, TS). The standardized BESS defines the subject's balance errors, which must be counted:

- Moving the hands off the hips.
- Opening the eyes.
- Step, stumble or fall.
- Abduction or flexion of the hip beyond 30°.
- Lifting the forefoot or heel off the testing surface.
- Remaining out of the proper testing position for longer than 5 s.

Although fast and inexpensive, the BESS test presents a series of limitations that are intrinsically related to its subjective and manual scoring method. The BESS has been reported to have modest and widely ranging test sensitivity due to scoring inaccuracies and observer bias.[11,13] For instance, in Ref. 11 it has been reported that the inter-rater and intra-rater minimum detectible change for the total BESS score were respectively 9.4 and 7.3 points. These changes are in the same range as BESS score differences between baseline and testing in concussed subjects. It has been found that the average BESS score after concussion is 17 errors (range 15–19 errors), compared with ten errors at baseline (range 8.4–12.7 errors).[5] Further BESS limitations are the need for properly trained medical personnel to administer the test and its susceptibility to fatigue and practice effects.[5,18]

Given these limitations, numerous research efforts have aimed at improving balance evaluation for the BESS test. Such efforts can be divided into two main groups: (1) modifying and optimizing the human-based version of the test to make it more sensitive and to reduce the effects of fatigue and practice[3,18,19,33]; (2) Instrumenting the BESS, using either inertial measurement units[2,20,21] or force platforms,[1,8] to make the test more reliable and accurate, and reduce variability due to human bias.

MATERIALS AND METHODS

In order to overcome the BESS test limitations, we developed the automated assessment of postural stability (AAPS) system to quantify balance control in military field conditions. The AAPS' objective is to evaluate postural deficits due to concussion and musculoskeletal injuries commonly seen in active duty military personnel for return-to-duty assessment. The AAPS system utilizes an inexpensive Microsoft Kinect v2.0 motion capture sensor[5,18] and a custom designed software suite for Microsoft Windows to objectively track kinematics during Balance Error Scoring System (BESS) testing.

The AAPS system has been developed in the C# programming language in the Microsoft Visual Studio 2015 programming environment and the .NET 4.5 framework. The system requires minimal set-up time and no dedicated calibration time. Furthermore, it includes a comprehensive graphical user interface (GUI) to guide the operator and the subjects through the BESS test. The GUI also provides user controls, data management features, a real-time display of the detected body and intuitive visual feedback on the AAPS tracking capabilities. The system is user-friendly and its use only requires minimal training and experience. These characteristics facilitate the AAPS' integration and deployment in military practices. In addition to collecting information regarding the subject's joint positions and eye status (open/closed), the system provides real-time visual feedback to the operator. These characteristics help non-medical operators to properly position the subject in the field of view and guarantee that the necessary parameters are tracked correctly before starting the test.[27] The eye status tracking feature has been implemented using a face tracking library developed by Microsoft and available *via* the Kinect SDK 2.0. The face tracking combines HD color and infrared video streams to detect the location and status of the subject's eyes.[25]

This paper focuses on the balance error detection algorithm that has been implemented in the AAPS system to evaluate postural stability and provide a reliable and automated BESS score starting from raw Kinect sensor data. The algorithm has been designed to track balance errors as they are defined in the BESS standard.

This research was approved by the Temple University Institutional Review Board. All subjects provided written, informed consent prior to participating. A total of 15 healthy subjects (7 male, 8 female) each performed two complete BESS tests. Every trial was simultaneously tracked by a Kinect 2.0 sensor and a professional-grade motion capture system (Qualisys AB, Gothenburg, Sweden). High definition (HD) videos with BESS trials were scored off-line by three experienced observers. Further details regarding the experimental setup can be found in Ref. 26.

The AAPS system can be divided into three operating blocks: the Kinect sensor that detects human bodies using the depth data stream, the Microsoft proprietary skeletal tracking algorithm that converts the 3D camera images into 3D body joint coordinates, and the AAPS software that processes these coordinates and records balance errors. A top level block diagram of the system is shown in Fig. 1.

The Microsoft Kinect™ sensor is a low-cost, portable, and marker-less motion tracking system developed for video game applications. Despite being a mass-produced commodity, this 3D depth camera combined with the Microsoft proprietary skeletal tracking algorithm has the potential to be used as an alternative to laboratory-grade motion tracking systems.

Over the past few years, numerous performance comparisons between the Kinect and professional motion tracking systems have been carried out.[7,10,22,24,30,35] It is not surprising that such an inexpensive off-the-shelf commodity item cannot reach the levels of performance of professional-grade systems. Consequently, some important signal processing challenges must be overcome when using Kinect in applications such as the AAPS. For instance, the Kinect sensor frequently shows inaccuracies and oscillations when tracking body joints in both static and dynamic conditions. Such inaccuracies are affected by various parameters such as room conditions and geometry, the subject's body type, their distance from the sensor, and/or their clothing. Other types of tracking error can be due to quantization noise or missing information in the sensor data stream. Furthermore, the Kinect skeletal tracking presents two types of inaccuracy due to: (1) relatively small levels of white noise caused by detection imprecisions; and (2) temporary spikes in noise levels caused by joint tracking inaccuracies on a frame-by-frame basis. In real-world applications, as a result of these challenges, skeletal tracking can be carried out with precision levels in the range of a few centimeters.[12,26] Thus, without a strict rigid body model to be superimposed onto the Kinect raw data, it is not possible to compensate completely for the sensor-related errors.[4] This work demonstrates how dedicated signal processing techniques can mitigate these errors.

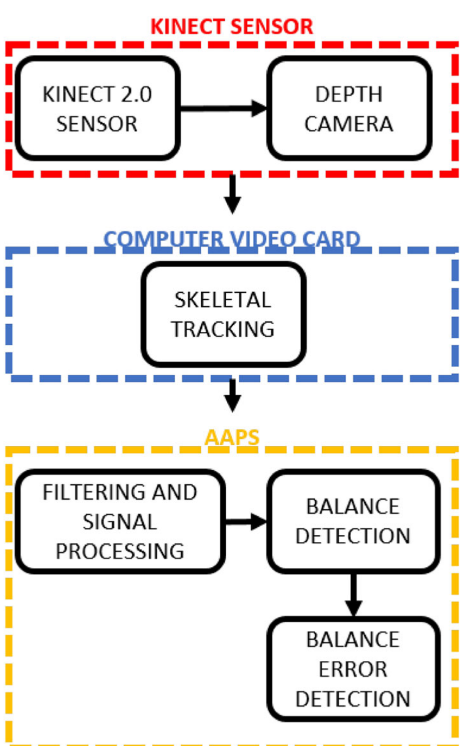

FIGURE 1. AAPS top level block diagram.

The first task of the AAPS software is to visualize and store the Kinect sensor raw data output, which is composed of HD video, infra-red video, three-dimensional depth data, joint position and orientation. Subsequently, as shown in Fig. 2, the AAPS extracts human body joint coordinates and locates the floor plane in real-time. The floor plane is used to identify the position and tilt of the sensor with respect to the subject. The joint coordinates are multiplied by a rotation matrix to compensate for sensor tilt and positioning. Next, the data frame rate is set to a constant value of 30 frames per second using linear interpolation. This is necessary because the Kinect provides data at a variable frame rate that depends on the instantaneous operating conditions of the acquisition computer (hardware/software) and data collection environment conditions such as lighting, room geometry, type and number of objects in the sensor field of view. To further account for the potentially large variability in the Kinect sensor frame rate (5–30 fps), the AAPS software was designed to perform real-time frame rate checks. If during a trial, the instantaneous frame rate drops below a certain value (10 fps in this application), an error message is displayed and the user is notified that the acquisition needs to be repeated. This is a fundamental feature in an automated system to guarantee acceptable performance levels in any condition. Based on our data collection sessions with the AAPS system, the ideal value of 30 fps tends to drop to 15 fps a few times per minute, while lower values are less frequent and usually occur once every 50 trials.

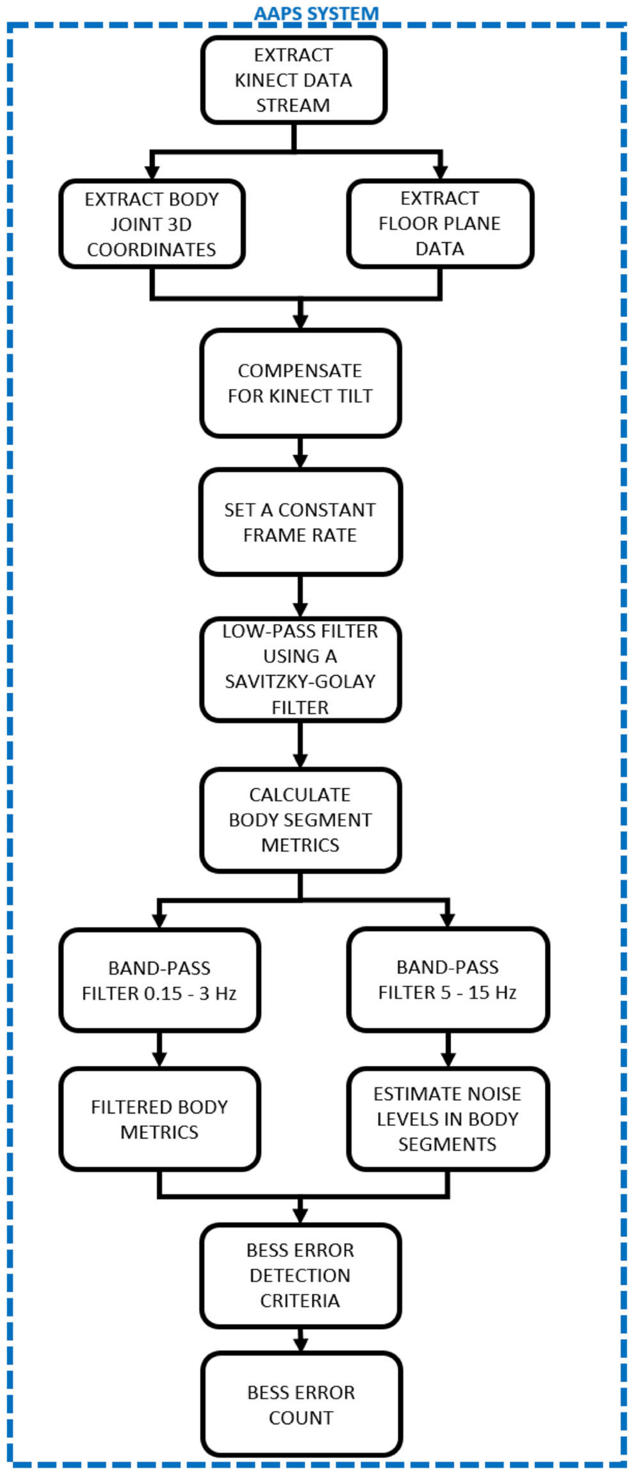

FIGURE 2. Detailed block diagram of the AAPS error detection algorithm.

Next the extracted body joint 3D coordinates are filtered using a Savitzky–Golay filter. This is a smoothing filter with minimal signal distortion that operates by fitting low-order polynomial approximations to consecutive signal time windows using a least-squares approach. A filter with a third order polynomial approximation and a time window duration of 0.166 s was used. At a constant sampling frequency of 30 Hz, such a window length corresponds to selecting five data points for each step of the least-squares approximation. As discussed above, filtering Kinect data with a smoothing filter is necessary to attenuate the effects of the Kinect inaccuracy and variability in estimating the joint positions of a tracked human body, even when subjects stand perfectly still in the sensor field of view. With the signal adequately smoothed, body metrics are calculated on a frame-by frame basis. The metrics that have been used in the AAPS algorithm to detect balance errors during BESS trials are listed in Table 1.

In order to detect errors in the subject's pose during balance trials, the algorithm uses a 1 s calibration window to estimate the reference subject's stance and the current levels of noise in the Microsoft skeletal tracking algorithm. The calibration is necessary to assess data variability due to changes in both subject-specific poses and sensor-specific body estimations. Subsequently, the metrics are bandpass filtered (using a second order Butterworth filter) between 0.15 and 3 Hz to emphasize signal components that are related to subject motion and to minimize other sources of variability (noise).

Additionally, sensor tracking inaccuracy is estimated by measuring the standard deviation of the noise in the calibration window. Specifically, the raw metrics are band-pass filtered with a second order band-pass Butterworth filter with passband set to 5–15 Hz. This frequency range was selected to emphasize the signal components that are mainly due to measurement noise.

During the 20-s long BESS trials, the estimated calibration stance and the current subject's position are continuously compared. The comparison is carried out using a threshold that is set using the estimated standard deviation of the noise and the mean of the metric obtained during calibration. Balance errors are flagged each time the metrics cross such a threshold. Specifically, in the ith frame, a balance error E_i is detected if the absolute difference between the calibration metric M_{cal} and the current metric $\hat{M}_i$ exceeds the threshold, set to ϵ times the estimated standard deviation σ_{cal} of the noise. The list of the kinematic metrics (M) that have been used to calculate the respective BESS errors (E_i) is presented in Table 1.

Mathematically, three categories of balance errors are detected:

(1) *Unilateral single threshold errors* estimated from low-noise and unilateral metrics.

$$E_i = \left(\left|\hat{M}_{n_i} - M_{n_{\mathrm{cal}}}\right| > \epsilon_n \times \sigma_{n_{\mathrm{cal}}}\right),$$

where the subscripts i, n and cal indicate respectively the frame number, the type of metric and the calibration window.

(2) *Bilateral errors* estimated from low-noise bilateral metrics. An error is detected if the threshold is crossed on either side of the body.

$$E_i = \begin{matrix} \left(\left|\hat{M}_{n_{\mathrm{left}_i}} - M_{n_{\mathrm{left}_{\mathrm{cal}}}}\right| > \epsilon_n \times \sigma_{n_{\mathrm{left}_{\mathrm{cal}}}}\right) \\ \mathrm{OR} \\ \left(\left|\hat{M}_{n_{\mathrm{right}_i}} - M_{n_{\mathrm{right}_{\mathrm{cal}}}}\right| > \epsilon_n \times \sigma_{n_{\mathrm{right}_{\mathrm{cal}}}}\right) \end{matrix}$$

where the subscripts left and right indicate from which side of the body the metrics were derived.

(3) *Double threshold errors* to improve detection performance, errors, that are estimated using low-accuracy metrics, are detected using two correlated metrics and corresponding thresholds. An error is detected only if both metrics cross the threshold.

$$\begin{aligned} E_i = &\left(\left|\hat{M}_{n_i} - M_{n_{\mathrm{cal}}}\right| > \epsilon_n \times \sigma_{n_{\mathrm{cal}}}\right) \\ &\text{and} \quad \left(\left|\hat{M}_{m_i} - \hat{M}_{m_{\mathrm{cal}}}\right| > \epsilon_m \times \sigma_{m_{\mathrm{cal}}}\right) \end{aligned}$$

where the subscripts n and m indicate different metrics.

The above error types can be combined for improved balance detection precision. The different error types detected on a frame-by-frame basis are then converted into BESS scores, namely the total error count per trial, with two important caveats: (1) at most one error type can be detected within a pre-defined

TABLE 1. Calculated metrics extracted from Kinect raw data that are tracked during BESS tests.

Joints of interest	Metric (M)	Detected balance error (E_i)
Left hand—left hip	3D distance	Hands off hips
Right hand—right hip	3D distance	Hands off hips
Left elbow—left hip	3D distance	Hands off hips
Right elbow—right hip	3D distance	Hands off hips
Left knee—right knee	3D distance	Foot movement
Left hip—left ankle	3D distance (single-leg stance)	Hip flexion
Right hip—right ankle	3D distance (single-leg stance)	Hip flexion
Ankles	3D position (tandem stance & single-leg stance)	Foot movement
Frontal plane spine angle	Angle	Spine frontal motion
Sagittal plane spine angle	Angle	Spine sagittal motion

time window (set to 2 s); (2) a BESS error is recorded only if the infraction remains above the threshold for a pre-defined time duration (set to 110 ms). A detailed block diagram of the AAPS algorithm is shown in Fig. 2.

In order to validate the results of the error detection algorithm, we simultaneously collected data using a Kinect sensor and a 12-Camera Qualisys system. Qualisys data have been post-processed using Opensim with a modified plug-in-gait model. After running inverse kinematics on the trajectory data, three-dimensional body joint positions were derived. The Kinect and Qualisys derived joint coordinate time series were time-synchronized using a large movement performed at the beginning of each trial and then fed into the BESS error detection algorithm as described above. Finally, scores obtained from the two systems were compared against scores from three human experts reviewing video footage of the BESS tests.

RESULTS

The AAPS algorithm was tested, using data derived from both Qualisys and Kinect systems, on 15 healthy subjects, each performing the BESS test twice. These subjects' balance was also evaluated by three expert observers using the gold standard BESS method. In the algorithm performance analysis, the average human scores have been chosen as ground truth (Reference) for the correct error count.

Figure 3 shows the differences in the scores obtained using the different evaluation techniques: AAPS vs. Reference, Qualisys vs. Reference, AAPS vs. Qualisys, Human 1 vs. Reference, Human 2 vs. Reference, and Human 3 vs. Reference. The comparison of AAPS vs. Qualisys was carried out to investigate potential differences in performance due to the two different optical acquisition systems. Variations in scores have been quantified by calculating the signed average difference between each technique and the reference. Differences can range between −10 and 10 points, where low error levels are indicated by values close to zero. Standard deviations are presented as error bars.

Table 2 reports the overall level of agreement for the different groups, where values close to 100% (high agreement) correspond to differences in BESS scores close to zero. The values in the table are calculated by taking the percentage complement of the normalized absolute average differences in the scores. The absolute differences were normalized using the BESS full scale (10 points per trial).

To evaluate the statistical significance of the observed score variations a multiple comparison one-way ANOVA test was implemented (α set to 0.05). The results are shown in Fig. 4, where the means (filled circles) and 95% confidence intervals (horizontal lines) of condition-based balance scores are presented. The gray vertical dotted lines represent the 95% confidence intervals with respect to the Reference group. No statistically significant differences were found between any of the balance scoring methods and the Reference (average human scores, in blue). The multiple comparison ANOVA results emphasize that although differences in the scores are non-significant, the Kinect-based AAPS reaches its lowest performance in the single-leg on foam condition, as also highlighted by the lowest agreement levels reported in Table 2.

Although single-leg on foam was identified as the condition with lowest agreement levels between observers, there was no significant difference in performance between the Kinect-based and the Qualisys-based AAPS. This finding suggests that the AAPS software algorithm provides satisfactory performance levels using raw data from both motion capture systems; BESS error detection performance is not significantly affected by the acquisition hardware.

DISCUSSION

As shown in the RESULTS section, the ANOVA analysis did not reveal any significant difference in the scores. It is worth noting that the lowest AAPS performance levels are detected in single-leg and tandem stances on foam. In such conditions, despite the Qualisys-based AAPS system performing more closely to humans than the Kinect-based one, statistical analysis shows no significant difference in performance. This result demonstrates that the AAPS, built around an inexpensive, general-purpose 3D single-camera sensor, is viable for use in on-field applications.

The lowest agreements between both the AAPS systems and human observers are seen in the single-leg and tandem stances on foam condition. We hypothesize that lower agreement levels might be due to the higher levels of subjective evaluation that this condition requires to detect BESS errors. Specifically, we identified three main factors. First, the presence of the foam complicates balance evaluation, because the foot on which the subjects stand is partially obscured by the foam. Secondly, this condition is arguably the most challenging, and consequently more motion is expected. This results in multiple errors and subjects having more difficulty to find and maintain their balance when trying to go back into the right position. In these cases, we found that human observers tend to use their "judgment" to count errors rather than strictly relying upon the BESS rules for balance error count. Finally, in single-leg on foam conditions, the auto-

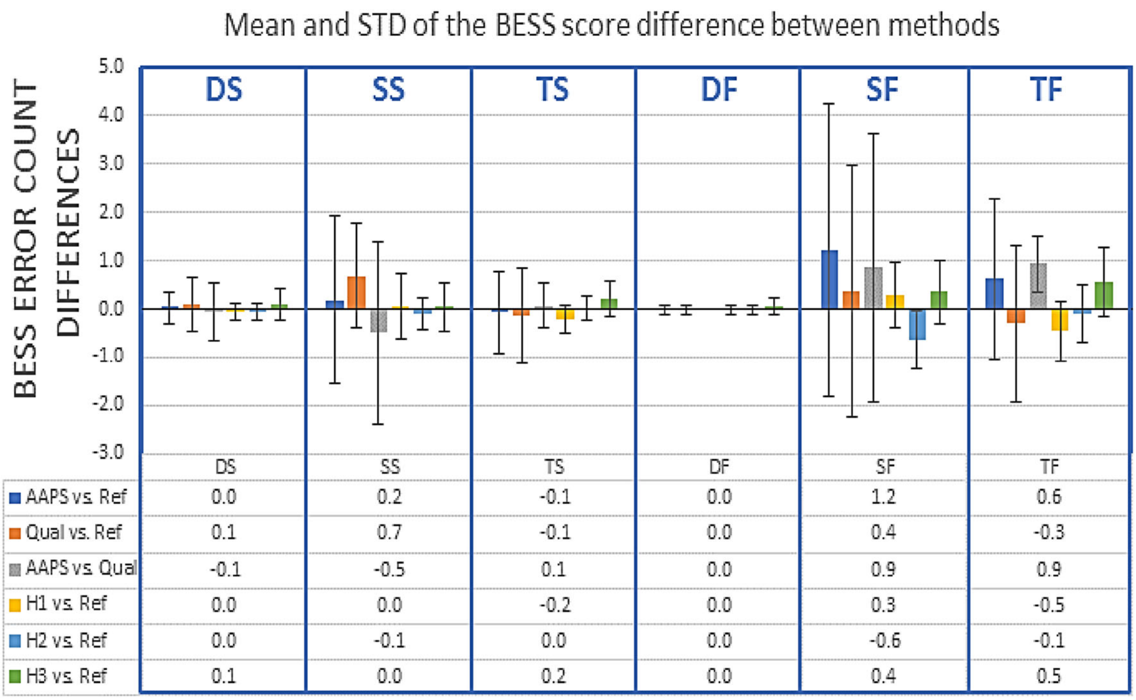

FIGURE 3. Means and standard deviations of the score differences calculated for each balance scoring method and grouped by stance condition. Bottom: mean error values for each group and condition. The tested stance conditions are: double leg (DS), single leg (SS) and tandem stance (TS) on firm ground; double leg (DF), single leg (SF) and tandem stance (TF) on foam pad. The blue, orange, grey, yellow, light blue and green bars represent different balance evaluations derived respectively for AAPS vs. Reference, Qualisys vs. Reference, AAPS vs. Qualisys, Human 1 vs. Reference, Human 2 vs. Reference, and Human 3 vs. Reference.

TABLE 2. Average differences expressed as percentage of agreement between different balance evaluation systems in detecting BESS scores, grouped by condition.

Percentage of agreement						
Condition	AAPS vs. Ref	Qual vs. REF	AAPS vs. Qual	H1 vs. Ref	H2 Vs. Ref	H3 vs. Ref
DS	99.8	99.0	99.3	99.5	99.5	99.0
SS	98.1	93.1	95.0	99.5	99.0	99.5
TS	99.3	98.6	99.3	97.9	100.0	97.9
DF	99.8	99.8	100.0	99.8	99.8	99.5
SF	87.9	96.4	91.4	97.1	93.6	96.4
TF	93.8	96.9	90.7	95.5	99.0	94.5

matic system seems to be operating at the limits of agreement between humans and AAPS systems because of the low sensitivity of the BESS test. This limitation has been reported in previous studies in which the modest sensitivity of the BESS is explained by the large variance in performance during the stances on foam. Over 53% of the variance in errors can be attributed to the single-leg and tandem conditions on foam.[21]

The BESS only focuses on static postural control tasks and lacks assessment of more dynamic postural tasks. Thus, the choice of filtering the kinematic metrics between 0.15 and 3 Hz to emphasize relevant data was deemed appropriate. The Kinect, and consequently the AAPS capabilities will be tested at their operational limit when introducing dynamic testing with the aim of capturing "faster" human movements. In such conditions, although the motion of large human body segments rarely exceeds a few Hertz, the filter high cut-off frequency needs to be increased to avoid signal's distortion and artifacts. However, based on our preliminary data during dynamic trials, the AAPS seems to perform at acceptable levels when compared to the Qualisys lab-grade performance.

Testing only for static stability may not capture other important domains of balance, including dy-

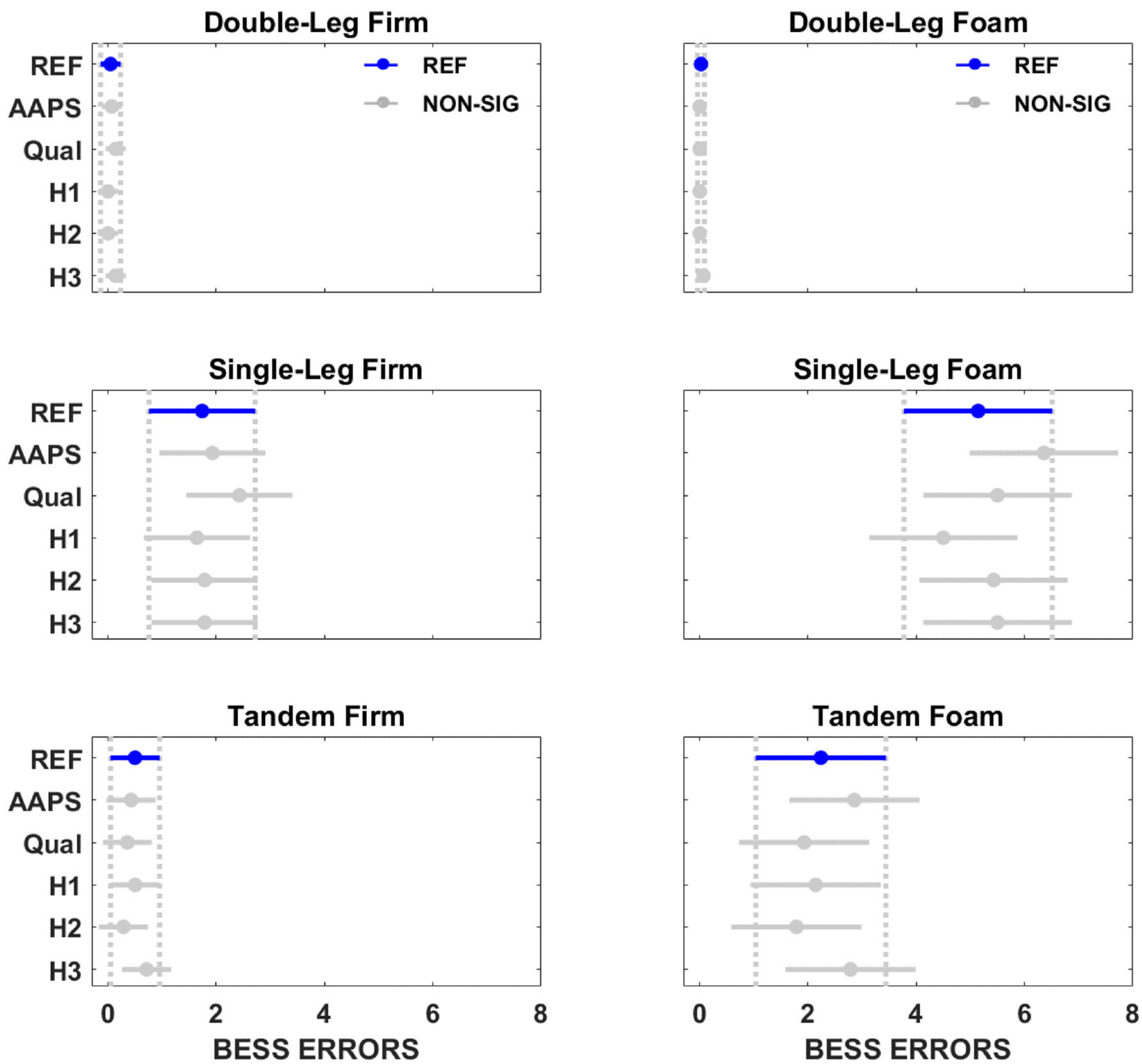

FIGURE 4. Results of a multiple comparison ANOVA test on the BESS scores. BESS errors derived using the AAPS, the Qualisys and three different human observers are compared to the average human scores, used as reference. Reference groups are in blue; vertical dotted lines are 95% confidence intervals for the Reference group. None of the differences with respect to Reference are statistically significant.

namic or multi-task postural control aspects.[13] It is worth noting that the AAPS capability of detecting balance deficits had to be reduced to a single error count number per trial for the purpose of the comparison presented here.

These limitations derive from human administration of such testing protocols, wherein some information (e.g., error type, time, and magnitude) must be sacrificed in order to accommodate the capacity of a human observer. We hypothesize that an improved automated balance test, in which dynamic conditions and more reliable proxy kinematic variables are used, can be readily implemented by exploiting the existing capabilities of the AAPS system. The use of such a system to detect, track and quantify balance deficits in the field will provide the opportunity to go beyond traditional balance testing protocols that only rely on human visual observations reported with manual annotations. This will facilitate more informed and data-driven clinical decision making in non-clinical settings.

Despite some level of disagreement between human and AAPS-generated scores, the use of an automated system yields important advantages over currently available human-based alternatives. A computer scoring system is by definition deterministic, meaning that it eliminates variability during repeated evaluations, the same criterion does not apply to human scoring. Moreover, the AAPS can record specific error types with extremely high temporal resolution, it can detect multiple error types on a frame-by-frame basis and record their time course progression. These features are not achievable by humans, who cannot keep track of all those variables with such a high time resolution. Together, these results suggest that computerized BESS calculation may provide more accurate and consistent measures of balance than those derived from human experts.

Our results show that the AAPS error detection algorithm presented here can accurately and precisely detect balance deficits with performance levels that are comparable to those of experienced medical personnel. Specifically, our results show agreement levels between the AAPS algorithm and the human average BESS scores ranging between 87.9% (single-leg on foam) and 99.8% (double-leg on firm ground). In addition, statistically significant differences were not detected by an ANOVA test with significance level set to 0.05. Moreover, significant performance deficits were not detected when the less expensive, portable and markerless AAPS was compared to a lab-grade system, with agreement levels between the two different motion capture systems ranging between 90.7% (tandem on foam) and 100% (double-leg on foam). These results underscore the value of using the Kinect-based AAPS, which can be quickly deployed in the field and/or in outdoor settings with minimal set-up time.

In future work, we plan on expanding the AAPS with new features, such as introducing criteria to account for balance error characteristics and fine-grained evaluation of dynamic and static postural control strategies using kinematic variables rather than trying to capture complex motion performance with an arbitrary summary scale. Such a system will also implement functional dynamic protocols that can be customized to a specific subject and application. These new dynamic posture screening tools combined with the ability to derive real-time meaningful postural metrics will help us develop innovative automated tools for more effective and comprehensive on-field postural strategy assessment. Furthermore, the AAPS capabilities will be tested in clinical populations, such as individuals suffering from low-extremity injuries and concussion.

ACKNOWLEDGMENTS

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Neurosensory and Rehabilitation Research Award Program under Award No. W81XWH-15-1-0045. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office.

REFERENCES

[1]Alsalaheen, B. A., J. Haines, A. Yorke, K. Stockdale, and S. P. Broglio. Reliability and concurrent validity of instrumented balance error scoring system using a portable force plate system. *Phys. Sports Med.* 43(3):221–226, 2015.

[2]Amick, R. Z., S. D. Jansen, A. Chaparro, N. A. Hakansson, J. A. Patterson, and M. J. Jorgensen. Comparison of the SWAY balance mobile application to the abbreviated balance error scoring system. *Athl. Train. Sport Health Care* 7(3):89–96, 2015.

[3]Azad, A. M., S. Al Juma, J. A. Bhatti, and J. S. Delaney. Modified balance error scoring system (M-BESS) test scores in athletes wearing protective equipment and cleats. *BMJ Open Sport Exerc. Med.* 2(1):e000117, 2016.

[4]Azimi, M. Skeletal Joint Smoothing White Paper. In: Microsoft Developer Network. https://msdn.microsoft.com/en-us/library/jj131429.aspx. [Accessed: 01-Jan-2017].

[5]Bell, D. R., K. M. Guskiewicz, M. A. Clark, and D. A. Padua. Systematic review of the balance error scoring system. *Sport Health A* 3(3):287–295, 2011.

[6]Broglio, S. P., S. N. Macciocchi, and M. S. Ferrara. Sensitivity of the concussion assessment battery. *Neurosurgery* 60(6):1050–1057, 2007.

[7]Clark, R. A., K. J. Bower, B. F. Mentiplay, K. Paterson, and Y. H. Pua. Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *J. Biomech.* 46(15):2722–2725, 2013.

[8]Clark, R. A., A. L. Bryant, Y. Pua, P. McCrory, K. Bennell, and M. Hunt. Validity and reliability of the Nintendo Wii balance board for assessment of standing balance. *Gait Posture* 31(3):307–310, 2010.

[9]Clinical Practice Guideline. Management of Concussion/mild Traumatic Brain Injury. The Management of Concussion/mTBI Working Group. VA/DoD Evidence Based Practice. April, 2009.

[10]Dolatabadi, E., B. Taati, and A. Mihailidis. Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. *Med. Eng. Phys.* 38:952–958, 2016.

[11]Finnoff, J. T., V. J. Peterson, J. H. Hollman, and J. Smith. Intrarater and interrater reliability of the balance error scoring system (BESS). *PM R* 1(1):50–54, 2009.

[12]Galna, B., G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait Posture* 39(4):1062–1068, 2014.

[13]Giza, C. C., J. S. Kutcher, S. Ashwal, J. Barth, T. S. D. Getchius, G. A. Gioia, G. S. Gronseth, K. Guskiewicz, S. Mandel, G. Manley, D. B. McKeag, D. J. Thurman, and R. Zafonte. Summary of evidence-based guideline update: evaluation and management of concussion in sports. *Neurology* 80(24):2250–2257, 2013.

[14]Guskiewicz, K. M. Postural stability assessment following concussion: one piece of the puzzle. *Clin. J. Sport Med.* 11(3):182–189, 2001.

[15]Guskiewicz, K. M., B. L. Riemann, D. H. Perrin, and L. M. Nashner. Alternative approaches to the assessment of mild head injury in athletes. *Med. Sci. Sports Exerc.* 29(7 Suppl):S213–S221, 1997.

[16]Guskiewicz, K. M., S. E. Ross, and S. W. Marshall. Postural stability and neuropsychological deficits after concussion in collegiate athletes. *J. Athl. Train.* 36(3):263–273, 2001.

[17]Helmick, K. M., C. A. Spells, S. Z. Malik, C. A. Davies, D. W. Marion, and S. R. Hinds. Traumatic brain injury in the

US military: epidemiology and key clinical and research programs. *Brain Imaging Behav*. 9:358–366, 2015.
[18]Hunt, T. N., M. S. Ferrara, R. A. Bornstein, and T. A. Baumgartner. The reliability of the modified balance error scoring system. *Clin. J. Sport Med*. 19(6):471–475, 2009.
[19]Iverson, G. L., and M. S. Koehle. Normative data for the modified balance error scoring system in adults. *Brain Inj*. 27(5):596–599, 2013.
[20]King, L. A., F. B. Horak, M. Mancini, D. Pierce, K. C. Priest, J. Chesnutt, P. Sullivan, and J. C. Chapman. Instrumenting the balance error scoring system for use with patients reporting persistent balance problems after mild traumatic brain injury. *Arch. Phys. Med. Rehabil*. 95(2):353–359, 2014.
[21]King, L. A., M. Mancini, P. C. Fino, J. Chesnutt, C. W. Swanson, S. Markwardt, and J. C. Chapman. Sensor-based balance measures outperform modified balance error scoring system in identifying acute concussion. *Ann. Biomed. Eng*. 2017. doi:10.1007/s10439-017-1856-y.
[22]Lim, D., C. Kim, H. Jung, D. Jung, and K. Chun. Use of the microsoft kinect system to characterize balance ability during balance training. *Clin. Interv. Aging* 10:1077–1083, 2015.
[23]McCrory, P., W. Meeuwisse, M. Aubry, B. Cantu, J. Dvořák, R. Echemendia, L. Engebretsen, K. Johnston, J. Kutcher, M. Raftery, A. Sills, B. Benson, G. Davis, R. Ellenbogen, K. Guskiewicz, S. A. Herring, G. Iverson, B. Jordan, J. Kissick, M. McCrea, A. McIntosh, D. Maddocks, M. Makdissi, L. Purcell, M. Putukian, K. Schneider, C. Tator, and M. Turner. In: Consensus Statement on Concussion in Sport. The 4th International Conference on Concussion in Sport held in Zurich, November 2012, *Phys. Ther. Sport*, 14(2):250–258, 2013.
[24]Mentiplay, B. F., 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*. 48(10):2166–2170, 2015.
[25]Microsoft. Face Tracking. In: Kinect for Windows SDK 2.0—Programming Guide. https://msdn.microsoft.com/en-us/library/dn782034.aspx. [Accessed: 01-Jan-2017].
[26]Napoli, A., S. Glass, C. Ward, C. Tucker, and I. Obeid. Performance analysis of a generalized motion capture system using microsoft kinect 2.0. *Biomed. Signal Process. Control* 38:265–280, 2017.
[27]Napoli, A., C. R. Ward, S. M. Glass, C. Tucker, and I. Obeid. In: Automated assessment of postural stability system. Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6090–6093, 2016.
[28]Onate, J. A., B. C. Beck, and B. L. van Lunen. On-field testing environment and balance error scoring system performance during preseason screening of healthy collegiate baseball players. *J. Athl. Train*. 42(4):446–451, 2007.
[29]Peterson, C. L., M. S. Ferrara, M. Mrazik, S. Piland, and R. Elliott. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clin. J. Sport Med*. 13(4):230–237, 2003.
[30]Pfister, A., A. M. West, S. Bronner, and J. A. Noah. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *J. Med. Eng. Technol*. 1902(August):1–7, 2014.
[31]Putukian, M. The acute symptoms of sport-related concussion: diagnosis and on-field management. *Clin. Sports Med*. 30(1):49–61, 2011.
[32]Riemann, B. L., K. M. Guskiewicz, and E. W. Shields. Relationship between clinical and forceplate measures of postural stability. *J. Sport Rehabil*. 8(2):71–82, 1999.
[33]Starling, A. J., D. F. Leong, J. M. Bogle, and B. B. Vargas. Variability of the modified balance error scoring system at baseline using objective and subjective balance measures. *Concussion* 1:1–8, 2015.
[34]Stroupe, K. T., B. M. Smith, T. P. Hogan, J. R. St, T. Andre, M. L. Pape, E. Steiner, Z. Huo Proescher, and C. T. Evans. Healthcare utilization and costs of veterans screened and assessed for traumatic brain injury. *J. Rehabil. Res. Dev*. 50(8):1047–1068, 2013.
[35]Yeung, L. F., K. C. Cheng, C. H. Fong, W. C. C. Lee, and K.-Y. Tong. Evaluation of the Microsoft Kinect as a clinical assessment tool of body sway. *Gait Posture* 40(4):532–538, 2014.