Automated System for Balance Error Scoring

Paarth Dave[1], Iyad Obeid[1] and Carole Tucker[2]
[1]*Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, U.S.A.*
[2]*Department of Physical Therapy, Temple University, Philadelphia, PA, U.S.A.*

Keywords: BESS, Kinect, Traumatic Brain Injury, Concussion, Automated.

Abstract: The Balance Error Scoring System (BESS) test is a commonly used tool for assessing static postural stability after concussion that quantifies compensatory arm, eye and trunk movements. However, since it is scored by clinician observation, it is potentially susceptible to biased and inaccurate test scores. It is further limited by the need for properly trained clinicians to simultaneously administer, score and interpret the test. Such personnel may not always be available when concussion testing is needed such as at amateur sporting events or in military field situations. In response, we are creating a system to automatically administer and score the BESS in field conditions. The system is based on the Microsoft Kinect, which is an inexpensive commodity motion capture system originally developed for gaming applications. The Kinect can be interfaced to a custom-programmed laptop computer in order to quantitatively measure patient posture compensations for preventing balance loss such as degree of hip abduction/flexion, heel lift, eye opening, and hand movement. By (a) removing the need for an adequately trained clinician, (b) improving accuracy, and (c) using rugged off-the-shelf system components, it will be possible to administer better, more accurate concussion assessments outside of standard clinical settings.

1 INTRODUCTION

Concussion is highly prevalent and the cases of multiple concussions are increasing exponentially due to the lack of reliable diagnostic techniques of post concussive symptoms. Concussion, also referred to as mild traumatic brain injury (mTBI), is recognized as a clinical syndrome of biochemically induced alterations of brain function, typically affecting memory and orientation, which may involve loss of consciousness (American Academy of Neurology, 2013). A concussion occurs when the head hits or is hit by an object with sufficient force to cause temporary loss of function in the higher centers of the brain.

The Center for Disease Control and Prevention, estimates 1.5 million traumatic brain injuries in the United States each year, of which up to 75% are concussions. A separate epidemiological study has estimated 1.6 to 3.8 million sport-related concussions in the US alone (Covassin, 2013). Concussions are also seen in the military and recently estimates suggest that anywhere from 8-23% of military personnel who have deployed to Iraq and Afghanistan may have sustained a traumatic brain injury (TBI) with the increased risk of exposure to concussive injuries secondary to explosions and other military-related accidents (Bryan, 2013).

With the proliferation of concussions, which are difficult to diagnose in a timely fashion without a medical professional, there is therefore a need to identify a novel way to reliably diagnose and recognize concussive symptoms in order to prevent and improve interventions. Clinical strategies for the diagnosis and management of concussion have evolved considerably over the past decade (Lovell, 2006). All the methods developed require administration of the tests by clinicians in order to detect and document the post concussive symptoms. This introduces the chance of human errors during the administration and test scoring, potentially resulting in less reliable conclusions. A recent international consensus statement recommends that several aspects of concussion can be evaluated, such as dizziness, headache, poor sleep, and emotional problems; the group also reaffirmed that the balance component is a reliable and valid addition to the assessment of concussions (McCrory, 2009).

The Balance Error Scoring System (BESS) is

Dave P., Obeid I. and Tucker C..
Automated System for Balance Error Scoring.
DOI: 10.5220/0004951403290333
In *Proceedings of the International Conference on Biomedical Electronics and Devices* (TPDULL-2014), pages 329-333
ISBN: 978-989-758-013-0
Copyright © 2014 SCITEPRESS (Science and Technology Publications, Lda.)

one of the most commonly used clinical tests to assess concussion, and measures standing posture and balance related impairments. This test requires the clinician to count balance "errors" that include eye opening, arm movements, trunk leaning and stepping while simultaneously protecting the patient against balance loss. This multitasking on the part of the clinician can increase the chances of human errors in counting the balance errors. A means to administer the BESS automatically may reduce the error in scoring by individual clinicians and also allow them to focus on patient safety during the test.

The purpose of this research is to replace the manual administration of the BESS test with a fully automated version by using emerging motion capturing and image processing technologies in combination with custom software. This would not only increase the reliability of the test results, but would also provide an easier way of conducting different tests and recording the results into the database for future comparisons. Resulting research will not only aid in detecting post concussive symptoms, but also help in preventing the risks of multiple concussions by improving the reliability of return-to play decisions.

The concept of administrating the BESS using machine vision can be achieved using cameras capable of measuring depth such as laser-based time-of-flight cameras, structured light systems and camera-based triangulation systems which may cost ~\$100k USD. Alternatively, we explore the use of emerging gaming technology such as the Kinect for Windows which costs less than \$300 USD, opening up the use of depth cameras in a wide range of applications (Choppin, 2013). As a research tool, the Kinect can be controlled and accessed through computer and driver software easily (Kinect for Windows Programming Guide, 2013).

2 METHODS

The purpose of this research is to create a system for inexpensively and accurately quantifying post-concussive symptoms by administering a computer automated version of the standard BESS test. Whereas the standard BESS test is scored by a highly trained human clinician, our system will use the Microsoft Kinect, a commodity motion capture system, to track patient movement and to score the exam. This system will be valuable because it will facilitate the measurement of concussion especially in situations where a trained clinician is not readily available such as at amateur sporting events or in active military environments. By improving the determination of concussion symptoms, our system will facilitate return-to-play and return-to-duty decisions, thereby improving clinical outcomes for patients.

2.1 Overview

Our system is comprised of just two hardware elements, both of which are readily available commodity items requiring no physical alteration or modification. The Kinect (Microsoft, Redmond, WA, USA) is a relatively inexpensive motion capture system originally developed for gaming applications. A built-in software layer tracks human body movement in real-time, expressed as x-y coordinates for 20 key body joints. The second hardware element is a standard Window-based personal computer. Using an open-source software development kit, custom software is written for the PC that can quantify the relevant measurements of the BESS test (trunk angle, foot lift, etc) using the skeleton coordinates returned from the Kinect. The Kinect is also used to detect eyeblinks. For ease of development, software is written using Matlab (Mathworks, Natick MA, USA), although an eventual production version of the system would be coded in C/C++ or Java. The system is self-contained, portable, and can be easily administered by a technician with no medical training.

2.2 Microsoft Kinect

Launched in November 2010, Microsoft Kinect is a sensor suite based around the PrimeSense design, which allows it to provide depth, RGB, infrared and audio information to the end user of the product (Boulos, 2011). The sensor has an RGB (red-green-blue) camera for color video, and an infrared emitter and camera that measure depth (in millimeter). Through its depth camera it is able to capture point cloud data at 30Hz, effectively scanning a surface as it does so. Proprietary algorithms developed by PrimeSense and Microsoft are not only able to use the depth cloud to recognize human users within the field of view but also to calculate joint positions and segment angles for the purposes of gesture recognition and command (Choppin, 2013).

The Kinect can see a usable range from zero to five meters in front of the sensor. The field of view is 57° horizontal and 43° vertical. The motorized tilt of the sensor allows for ±28° of movement in the vertical axis. Image data is captured at 1280x1024 but the algorithm operating within the Kinect compresses it down to 640x480 to allow for transmission at 30fps. It is also possible to capture information from the Kinect using OpenNI (Open Natural Interaction) backend libraries. OpenNI is a non-profit organization that provides certification and improves the feasibility of natural user interface and organic user interface for natural interaction devices like Microsoft Kinect. OpenNI framework is an open source software development kit (SDK) used for the development of natural interactions libraries and applications (OpenNI, 2013). These libraries allow for the collection of the full 1280x1024 pixel images but at slowed frame rate of only 10Hz. The maximum resolution of the Kinect is to resolve 1mm to a single pixel starting at a range of 0.8m. The depth image is mapped to a 320x240 pixel image further reducing the allowable resolution and capping how precise the Kinect is. This 320x240 pixel image is then mapped to the provided 640x480 pixel RGB/IR image before passing to the end user (OpenNI Programming Guide, 2013).

2.3 BESS Test

The BESS is a neuropsychological test that is commonly used to evaluate balance. It is especially valuable in cases of concussion since static balance involves feedback from the somatosensory, visual, and vestibular systems to achieve steadiness (Guskiewicz, 2011), all of which may be effected by brain injury. The BESS consists of three stances: double-leg stance (hands on hip and feet together), single leg stance (standing on the non-dominant leg with hands on hips), and a tandem stance (non-dominant foot behind the dominant foot) in a heel-to-toe fashion. The stances are performed on a firm surface and on a foam surface with eyes closed. The purpose of the foam pad is to create an unstable surface and a more challenging task. Each stance is held for a single 20-second trial and is scored by counting the errors or deviations from the proper stance. An error is counted when any of the following occur:

- Moving the hands off the iliac crests (hips)
- Opening the eyes
- Step stumble or fall
- Abduction or flexion of the hip beyond 30 degrees
- Lifting the forefoot or heel off of the testing surface.
- Remaining out of the proper testing position for greater than 5 seconds.

The maximum total number of errors for any single condition is 10.

Subjects that are unable to maintain the testing procedure for a minimum of five seconds are assigned the highest possible score of ten points for that testing condition. A positive test is a score that is 25% above the patient's baseline score and indicates cerebral dysfunction.

The purpose of the software under development in this project is to automatically detect each of the six error conditions in order to properly score the BESS test without needing a trained clinician.

3 PRELIMINARY WORK

Our system is designed using MATLAB owing to its ease of development and its rich libraries of image processing functions as well as its integration with the Kinect. In particular, the MATLAB Image Acquisition Toolbox contains many built-in tools to facilitate detection of the six BESS errors.

3.1 Spine Angle

Spine angle detection uses the joint indices metadata from the skeletal viewer for *Hip Center*, *Shoulder Center*, and *Head* in order to calculate the angle measurements. The logic behind the spine angle detection is to obtain the respective joint indices pixel values for each frame captured and to subtract it from the standard fixed joint indices pixel values, resulting in values that reflect the total change in movement. The values obtained can be converted into angles using a standard four-quadrant arctangent function built into MATLAB.

3.2 Foot Lift

There are two methods for detecting foot lift using Kinect. Both methods are being implemented in MATLAB and are being compared for accuracy. Foot lifts can be detected either by touch sensing technology or skeleton data generated due to knee bending. In the touch sensing method, the application of depth-sensing cameras to detect touch has been explored (Wilson, 2010). A novel interactive surface and touch screen technology has been presented that uses image processing techniques to produce a touch image useful for many gesture-based and perceptual computing scenarios (Wilson, 2004). These techniques presented the use of depth sensor and infrared sensor in Kinect to emulate touch screen sensor technology. Usually, the depth-sensing camera reports distance to the nearest surface at each pixel. The offset infrared camera is used to calculate the precise manner in which the fixed pattern of infrared light is distorted as a function of the depth of the nearest physical surface. Using this concept, a design can be programmed in a similar way to detect foot lift. But, this technique can be inaccurate because the depth calculations are based on triangulating features in the image which decreases the depth precision as the distance from the camera to subject increases (Wilson, 2010).

The other method is to detect foot lift (or foot drop) by detecting the change in knee angle and angle between two legs. To implement such method we can consider various test positions and the probability of stumble in any directions can be evaluated. Logical reasoning in the detection due to movement of leg can play an important role in detecting a foot lift efficiently. For example, during standing position the knee angle can be considered to be at 0° and any change in this angle can be assumed to be foot lift.

3.3 Eye Opening Detection

A built-in function in MATLAB "`vision.CascadeObjectDetector`" is used to perform eye detection. The function uses the Viola-Jones algorithm. The cascade object detector can detect faces, noses, eyes, mouth, or upper body. The Viola-Jones object detection framework is the first object detection framework to provide real-time object detection (Viola, 2001). The features employed by the detection framework universally involve the sums of image pixels within rectangular areas. The value of any given feature is always simply the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles. The eye tracking also allows to further detection of opening and closing of the eyes. The eye detection technique results in a bounding box that gives access to obtain the pixel values in that box.

Eye and pupil detection can be implemented using many methods if the eyes are clearly visible within an image. However, with the Kinect, the field of view is an obstacle that may hinder the efficacy of the Viola-Jones method. An alternate approach is under development and will be compared for efficacy. The concept is to calculate the width of the eyes in real-time. Logically, the width of the eyes increases when eyes are opened compared to when they are closed. The built-in image processing toolboxes can be used to detect changes in eye-width within a bounding box area in order to assess eye-opening events in real-time.

4 FUTURE WORK

Once the baseline system has been fully evaluated (by comparing its performance to that of professional clinicians), a number of important improvements can be made to enhance its functionality.

4.1 Automatic Calibration

In real-world settings, testing conditions are typically not as well controlled as in the research lab. Differences in ambient lighting, background image clutter, sensor tilt, distance to patient could all affect the accuracy of the system. Since our goal is for the system to be of value in practical contexts (such as sporting venues or even military theatres), these issues present a concern.

In response, we anticipate the need for developing an automatic calibration algorithm that will correct for these issues. The first step will be to work with our clinical partners to test the system in various field situations to determine which issues are most prevalent and contribute most to system inaccuracies. Anticipated solutions include (a) auto-tilt adjustment using a 3D accelerometer, (b) lighting adjustment using real-time color correction, (c) background clutter subtraction and (d) distance-dependent calibration scales.

In a wider context, we envision using our system to develop an expanded BESS test that is capable of measuring not only static postures but also dynamic movement tasks. This would require enhancing the system by creating an on-screen avatar whose movements the user would have to mimic; real-time on-screen feedback would show how well the subject can match those movements. Expanding the BESS to quantify a subject's ability to perform tasks such as pointing, leaning, and crouching may lead to more salient return-to-play or return-to-duty assessments since dynamic tasks may be more indicative of concussion severity than static poses.

In conclusion, our system will improve the way concussion is assessed by removing the need for a trained clinician to provide the first level of screening. By interfacing the Microsoft Kinect to a suite of custom software, it will be possible to automate the standard-of-care BESS test and to deploy it in non-clinical environments.

REFERENCES

American Academy of Neurology, 2013. https://www.aan.com/, accessed 12/2/2013.

Boulos, M.N.K., Blanchard, B.J., Walker, C., Montero, J., Tripathy, A., Gutierrez-Osuna, R., 2011. Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. In *Int. J. Health Geogr.*, 10(1): 45.

Bryan, C.J., 2013. Multiple traumatic brain injury and concussive symptoms among deployed military personnel. In *Brain Inj.*, 27(12): 1333–7.

Choppin, S., Wheat, J., 2013. The potential of the Microsoft Kinect in sports analysis and biomechanics. In *Sports Technology*, 6(2): 78–85.

Covassin, T., Moran, R., Wilhelm, K., 2013. Concussion Symptoms and Neurocognitive Performance of High School and College Athletes Who Incur Multiple Concussions. In *Am. J. Sports Med.*, 41(12): 2885-9.

Guskiewicz, K.M., Clark, M.A., Padua, D.A., Bell, D.R., 2011. Systematic Review of the Balance Error Scoring System. In *Sports Health: A Multidisciplinary Approach*, 3(3): 287–295.

Kinect for Windows Programming Guide, 2013. http://msdn.microsoft.com/en-us/library/hh855348.aspx, accessed 12/2/2013.

Lovell, M.R., Iverson, G.L., Collins, M.W., Podell, K., Johnston, K.M., Pardini, D., Pardini, J., Norwig, J., Maroon, J.C., 2006. Measurement of symptoms following sports-related concussion: reliability and normative data for the post-concussion scale. In *Appl. Neuropsychol.*, 13(3): 166–74.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., Cantu, R., 2009. Consensus statement on Concussion in Sport--the 3rd International Conference on Concussion in Sport held in Zurich, November 2008. In *J. Sci. Med. Sport*, 12(3): 340–51.

OpenNI, 2013. http://www.openni.org/, accessed 12/2/2013.

OpenNI Programming Guide, 2013. http://www.openni.org/openni-programmers-guide/, accessed 12/2/2013.

Viola, P., Jones, M., 2001. Robust Real-time Object Detection. *Intl J Computer Vision*.

Wilson A.D., 2010. Using a depth camera as a touch sensor. In *ACM Int. Conf. Interact. Tabletops Surfaces - ITS '10*.

Wilson, A.D., 2004. TouchLight : An Imaging Touch Screen and Display for Gesture-Based Interaction. In *ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces*.

SCIENCE AND TECHNOLOGY PUBLICATIONS