A Brief Review on the Validity and Reliability of Microsoft Kinect Sensors for Functional Assessment Applications

Pablo Rodrigo DIAZ-MONTERROSAS, Ruben POSADA-GOMEZ, Albino MARTINEZ-SIBAJA, Alberto Alfonso AGUILAR-LASSERRE, Ulises JUAREZ-MARTINEZ, Juan Carlos TRUJILLO-CABALLERO
División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Orizaba, Tecnológico Nacional de México, Orizaba, Veracruz, México.
pdiazm@ito-depi.edu.mx

***Abstract*—Kinect sensors are Human Computer Interaction devices oriented to entertainment, but have rapidly spread to several fields such as health care, physical therapy, and training. Their multiple advantages place them at present in a competitive situation compared to traditional solutions. On the other hand, their accuracy and precision for sensitive human applications are still under critical examination. This paper presents a brief literature review on the validity and reliability of the first and the second generation Kinect sensors to get an idea of the feasibility of their propagation as measuring devices in functional assessment applications. Results are difficult to compare because they depend largely on the type of measured elements, the angle of view of the measurement, the distance to the sensor, and even the diversity of human motion features. Nonetheless, they suggest that Kinect sensors are capable of properly identifying posture and motion, but not body or joint rotations, unusual postures, or occlusions.**

***Index Terms*—computer vision, human computer interaction, pervasive computing, reviews, statistical analysis.**

I. INTRODUCTION

The measurement and recording of the position and orientation of the human body in motion is called Motion Capture (MoCap), a process that has been carried out by means of diverse technologies and used in a wide range of applications or as a research tool. Regazzoni, deVecchi, and Rizzi [1] collect from the literature one of the most common classifications that groups MoCap according to their working principles into: optical, mechanical, magnetic, and inertial systems. However, their use in a given field is frequently restricted by their typical drawbacks: markers in most optical systems, heavy suits in mechanical systems, or sensitive sensors in both magnetic and inertial systems.

Recently, the improvement of depth-based optical technology (RGB-D) has achieved a competitive position with traditional technologies. For example, optical multicamera systems are usually composed of 4 to 32 cameras capturing 30 to 2000 fps, by attaching a set of markers at strategic points of human body and by triangulation of images taken from different cameras they can extrapolate the location of joint centers; in turn, RGB-D systems perform the same task by measuring the distance of the user from the sensor, that is, the depth of the scene, usually by means of infrared light projection, which allows them to be invariant to color and tolerate very low lighting levels, in addition to avoiding the use of markers, an uncomfortable and slow process, unsuitable for people with motor impairments. RGB-D technology has achieved an affordable price, portability, and availability with the release of Microsoft™ Kinect™ sensors, in contrast to the high cost of multicamera systems and their orientation to controlled scenarios in specialized facilities. On the other hand, since their technology has been widely used for years, the validity and reliability of multicamera systems have been thoroughly analyzed, and although their accuracy and particularly their precision are debatable for the estimation of joint centers and relative segment orientations, their estimation of the positions of the markers is highly accurate [2], consequently, they are frequently used as a reference in the assessment of the feasibility of RGB-D systems for applications involving human motion analysis, for example, in pose estimation [3], tremor analysis (as in Parkinson's disease) [4], gait analysis [5], activity recognition [6], and hand gesturing [7], among other clinical assessments. In this context, the purpose of this paper is to provide an overview of the outcomes of some recent studies on the validity and reliability of Kinect™ sensors for functional assessment applications, and thus to gain an idea of the feasibility of their propagation as measuring devices.

II. THE MICROSOFT™ KINECT™ SENSORS

The first generation Kinect™ sensor (Kinect™ v1) was released in November 2010 as part of the Xbox 360 video game console, it is based on software developed by Rare, a company bought by Microsoft™ in 2002, and the technology of the PrimeSense cameras, now owned by Apple, that developed a system to interpret specific gestures without the need to touch anything, by means of an infrared camera and projector and a microchip to track 3D objects in motion. In 2011, Microsoft™ released the first Software Development Kit (SDK) with tools, drivers, APIs, and code samples to enable developing of Kinect™ applications, integrating gesture, facial, voice, and 20 body joints recognition (similarly, OpenNI and OpenKinect emerged as an alternative to the official source). The device contains an RGB camera with 8-bit VGA resolution (640x480 pixels), a monochrome depth sensor with 11-bit VGA resolution (320x240 pixels), and a four-element microphone array. The depth sensor consists of an infrared laser projector emitting a pattern of 307,200 points as a mesh and receiving the

Digital Object Identifier 10.4316/AECE.2018.01016

reflected pattern by means of a monochrome CMOS. This Structured Light application allows the device to measure the depth of each point by means of a triangulation technique. In order to automatically determine the body reference points (joints) in near real time, the sensor uses a Random Decision Forest algorithm [8]. For each joint, the sensor indicates whether or not its position has been directly observed, in which case it may have been inferred from previous positions and assumptions of body geometry [9].

The second generation of this device (Kinect™ v2) was released in November 2013 with the Xbox One gaming console, but started selling separately from mid-2014 with an electrical adapter, allowing its use for computing applications development. Instead of Structured Light, it is based on Time of Flight technology from Canesta and 3DV Systems, now Microsoft™ assets. Kinect™ v2 consists of a 1080p RGB camera and a 512x424 pixels depth sensor with higher fidelity to see small objects. It now locates up to 6 skeletons at a time, each consisting of 25 body joints.

Since Kinect™ v1 release, there has been a large number of scientific contributions, including several papers performing literature reviews [10-16], related to the features and applications of the Kinect™ sensors due to their multiple potential advantages compared to other similar technologies, as shown in [17]. On the other hand, there are also drawbacks that are being examined to define their usability in real situations and scenarios.

III. Literature Review

Clark *et al.* [8] performed three postural control tests with Kinect™ v1: forward reach, lateral reach, and single leg with eyes closed standing balance, using a Vicon system as a reference. All tests were performed three times, two trials for data analysis and the last one in case of an error in the collection of data. Both systems showed comparable relative and absolute inter-trial reliability (Intraclass Correlation Coefficient difference = 0.05; range, 0.00-0.16. Ratio Coefficient of Variation difference ≤11.6%) and excellent concurrent validity with Pearson's correlation >0.90 for most measurements (r = 0.96 ± 0.04; range, 0.84-0.99). Ordinary Least Products (OLP) analyzes showed proportional biases for some outputs related to the pelvis and sternum, which implies that an increasing amount of movement tends to extend the difference between the measurements of the systems. The authors concluded that the Kinect™ v1 is comparable to the reference system when assessing anatomical landmark positions and angular displacement during clinical tests of postural control.

Clark, Bower, Mentiplay, Paterson, and Pua [18] evaluated the validity of anatomical information of the Kinect™ v1 to examine the spatiotemporal characteristics of gait using a Vicon system as a reference. A set of tests were performed in which the participants walked at a self-selected pace along a walkway, starting at 3.8 m and decelerating at 1.3 m from the sensor to achieve a full gait cycle. Information on ankle and shoulder center positions was also recorded. Gait speed, step length and time, stride length and time, and foot swing velocity measures were acquired through a supervised automated analysis. The results indicated excellent relative and absolute agreement between the devices for gait speed, step length, and stride length (r and r_c values >0.90 and percentage error <8%). For foot swing velocity, the relative agreement was excellent (r =0.93), but the absolute one was modest (r_c =0.54) with percentage error of 13%. For step and stride times, the relative agreement was 0.82 and 0.69 respectively, however, the absolute agreement was poor, 0.23 and 0.14 respectively with high percentage error, 16% and 19%. The linearity assessment for landmark detection accuracy regarding distance from the sensor was excellent. In summary, the authors concluded that the Kinect™ v1 is valid for some spatiotemporal components of gait: measurements that did not rely heavily on accurate identification of event-specific synchronization (step and stride length and gait speed) were sufficiently valid, but those that required accurate identification of discrete time points (step and stride time) were not sufficiently valid.

Bonnechère *et al.* [2] performed shoulder and hip abduction and elbow and knee flexion tests along the anatomical planes in two sessions, using as a reference a Vicon MXT40S system. Their results indicated important discrepancies in the agreement of the angles acquired from both systems - comparing Range of Motion (ROM). Although outputs were excellent for shoulder abduction, according to their analyzes, Kinect™ v1 had difficulties estimating hip, knee, and especially elbow joint centers, even if the limb segments were correctly identified, but in fact the detected segment length was inconsistent over time (maximum difference of 12º and standard deviation of 5º for the elbow and maximum difference of 5º and standard deviation of 3º for the knee). Such errors are independent of the exact motion because Kinect™ v1 works frame by frame. On the other hand, Intraclass Correlation Coefficients (ICC) for both systems in the two sessions were similar, varying from moderate to good, showing good reproducibility (Kinect >0.66, Vicon >0.61). Errors were less than 1% for shoulder abduction, 6% for elbow flexion, 7% for hip abduction, and 9% for knee flexion. In general, the paired t-tests to compare the output data showed an intra-device and inter-device significance level of p <0.05 for all tests. In summary, the results of this work indicated that Kinect™ v1 has comparable reproducibility with the reference system, however, in terms of validity, the estimation of the joint center locations presented discrepancies, except for shoulder abduction.

Instead of the human body, Schmitz, Ye, Shapiro, Yang, and Noehren [19] used a testing jig simulating a lower limb (thigh and shank), to assess the validity and reliability of a Kinect™ v1 for 7 sessions, one for accuracy and six to accumulate precision data, recalibrating the systems before each test. A surface model template of the jig was built using KinectFusion in order to acquire data from the Kinect™ v1, including the position of 14 markers of a Motion Analysis Corp multicamera system used as the reference. In addition, a digital inclinometer was used as ground-truth to measure the primary angle manipulated in each configuration. The accuracy measurement consisted of six static postures (flexed, extended, abducted, adducted, externally rotated, and internally rotated) and was quantified as the difference between the mean calculated angle of three entities: Kinect-MotionAnalysis, Kinect-inclinometer, and MotionAnalysis-inclinometer. In addition, a paired t-test was used to statistically compare the accuracy of the two MoCap systems, with significance defined as $p < 0.05$. The precision measurement was performed only for three configurations: flexed, adducted, and internally rotated. For each configuration, the coefficient of repeatability, bias, and limits of agreement (LOA) of each system were calculated. The authors report that the flexion-extension and abduction-adduction angles of both systems deviated from the inclinometer measurements by less than 0.5°, being the multicamera system more accurate in abduction, whereas the Kinect™ v1 was more accurate in adduction. The agreement between the multicamera system and the Kinect™ v1 was <0.5º for the flexion-extension and abduction-adduction angles, and <2º for axial rotation. Likewise, the repeatability coefficient for both systems was <0.5º for all angles and their average test-retest difference was <1º. In general, the authors concluded that, since the differences in validity and reliability between the two systems were small, the Kinect™ v1 sensor may be a viable tool for calculating lower limb joint angles and sensitive enough to detect clinically relevant differences.

Galna *et al.* [20] assessed the validity of the Kinect™ v1 for the measurement of clinically relevant movements of people with Parkinson's disease, using a Vicon system as the reference. They performed a set of tests consisting of standing still, reaching forward and sideways, stepping forward and sideways, and walking on the spot, including elements of the Unified Parkinson's Disease Rating Scale: hand clasping, finger tapping, foot tapping and leg agility, sit-to-stand from a chair, and hand pronation. It is worth mentioning that in addition to 9 participants, 10 healthy control subjects were evaluated. An interesting fact of this study is that Kinect™ v1 failed to acquire coherent data on hand clasping of four participants with Parkinson's disease and on finger/toe tapping of most participants. However, the overall temporal accuracy was good. Pearson's and intraclass correlations were excellent, >0.9 for all movements. There was no significant bias between the systems and the LOA were below 10% of the group mean, indicating very good absolute agreement, except for hand clasping and hand pronation/supination in both groups of participants and lateral trunk flexion for the control group. Regarding spatial accuracy, the authors pointed out that the Kinect™ v1 significantly underestimated ROM for lateral flexion, hip kinematics during stepping, and vertical knee height during leg agility movements. Likewise, arm kinematics for shoulder flexion/abduction and elbow flexion movements was overestimated. In general terms, absolute agreement was poor, especially for hand clasping and walking on the spot. Poor Pearson's and intraclass correlations were also noted for arm pronation in participants with Parkinson's disease, but not in controls. Therefore, the authors concluded that the Kinect™ v1 is able to measure with accuracy the timing and gross spatial characteristics of clinically relevant movements but not with the same spatial accuracy for smaller movements, such as hand clasping or finger/toe tapping.

The aim of Yang *et al.* [21] was to determine the validity and reliability of Kinect™ v1 to assess a set of standing balance tests: double limb stance with feet apart about shoulder width, double limb stance with feet together, and single limb stance. However, they were also interested in ascertaining if body's center of mass (COM) could be accurately acquired from the kinematic data delivered by the Kinect™ v1 and the reference system, an Optotrak Certus. Standing balance tests showed no significant difference between the two trials. As regards COM, the authors calculated it using a 15-segment body model, the weight of each segment was calculated according to the subject's weight and anthropometric data, the 15-segment COM was computed based on the 3D coordinates of the positions acquired by both systems, and the whole body COM was calculated by weighting the relative positions of all body segments. Results showed excellent and comparable test-retest reliability for COM parameters of both systems (ICC >0.75). Likewise, concurrent validity for each calculated COM was excellent (ICC >0.88), although the average calculation velocity of the Kinect™ v1 was significantly lower than the reference system. The authors indicated that there was a significant linear relationship between the two systems ($p < 0.001$, $r > 0.930$), which means that biases could be corrected using linear calibration equations. They concluded that if the parameters of COM measured by the Kinect™ v1 are properly calibrated, it can be considered as a convenient device to assess standing balance.

Clark *et al.* [22], using Kinect™ v2, complemented the research carried out in [8] and [18] by assessing a set of static and dynamic standing balance tests, with a Vicon system as reference. Static tests consisted of single and double limb standing with eyes open and eyes closed. Dynamic tests consisted of forward and lateral reach movements and a limits of stability (LOS) assessment performed by shifting the weight of the participant as far as possible forward, backward, left, and right, starting from a straight posture. Authors interpreted their results for trunk angle during the reach and LOS dynamic tests as excellent, with a Pearson's correlation of $r > 0.75$, as well as for anterior-posterior range and path length measurements in all static balance tests. In contrast, for all medial-lateral range and path length measurements the validity was poor with $r < 0.40$, except for the single leg eyes closed balance test. ICC values for both systems had modest to excellent reliability (ICC ≥0.70) in the LOS, reach, and anterior-posterior axis with both limbs tests. In general, for balance tests, results of path length were more reliable than those of range, likewise, reliability values in the anterior-posterior plane were similar between both systems. On the other hand, Kinect™ v2 did not deliver good results in the medial-lateral plane for double limb tests. In summary, these results indicated that Kinect™ v2 achieves higher validity and reliability in anterior-posterior measurements than in medial-lateral ones, furthermore, that although reliability values were variable, they are usually comparable between both systems.

Huber, Seitz, Lesser, and Sternad [23] measured four shoulder joint angles that are representative in rehabilitation exercises: flexion and abduction to 90°, flexion to maximum, and external rotation to maximum at 0° of abduction. In addition, they measured flexion to 90° and flexion to maximum once more from a sagittal view to determine if the data previously acquired were correct, since the shoulder joint is occluded with these postures in the frontal view. They used a StarTrack electromagnetic system and a goniometer as references. Their results indicated a clinically significant discrepancy of ±5° in all postures, only the abduction to 90º achieved a reasonable accuracy comparing with any of the two reference systems. From the frontal view, ICC values for the Kinect™ v1 showed good to very good relative reliability (Kinect >0.76) and small values for Standard Error of the Measure (SEM) and Minimal Detectable Change (MDC), indicating good absolute reliability, except for the occluded flexion to 90°. From the sagittal view, the flexion to 90º showed very good relative and absolute reliability, but very low for flexion to maximum. These results represented a low accuracy, although a good overall precision for the measured shoulder joint angles, so the authors concluded that further assessment of more shoulder angles is needed before considering the use of Kinect™ v1 for clinical applications.

Xu and McGorry [24] pointed out that the joint centers identified by either the Kinect™ technology or the reference system are each relative to their own global coordinate system (GCS), which would explain some location errors. To avoid such situation, they proposed a method to align the coordinate systems of the Kinect™ v1 or the Kinect™ v2 to the GCS of the reference system, an Optotrak Certus, by measuring the points of a wooden wheel with eight spokes and a cross at the tip. After the alignment process, the accuracy of the sensors for joint center location was assessed by measuring a set of standing and sitting postures, concluding that the accuracy for the location of the joints of the upper limbs is better than for the lower limbs, whereas the accuracy for the identification of standing postures is better than for sitting postures (the average error among all participants for all joint centers identified in standing poses was 76 mm and 87 mm for the Kinect™ v1 and v2 respectively). These findings might be explained, according to the authors, due to the videogame native orientation of the Kinect™ technology, so the built-in algorithms might not work properly for sitting activities. The authors indicated that the statistical significances for both the Kinect™ v1 and v2 were $p < 0.0001$ after a paired *t*-test with the pooled error of the joint centers, furthermore, they mentioned that although the Kinect™ v2 has better resolution than the first generation, the accuracy of the location of the joint centers has not been significantly improved.

Darby, Sánchez, Butler, and Loram [25] evaluated the accuracy of the Kinect™ v2 to estimate the rotational and translational components of the three-dimensional posture of the head, in the context of posture control, using a Vicon system as a reference. The ROM of the head was assessed, including left-to-right yaw, up-and-down pitch, and side-to-side roll. The same set of tests was repeated under different conditions: calculating the rotations relative to the resting posture of each participant; with shoulders and torso free to move to increase the range; with hands over mouth to evaluate occlusions; previously performing the face shape calibration procedure of the Kinect™ v2; showing the whole body to the sensor; and calculating the rotations at a 45° angle from the sensor. The results indicated a larger number of missed frames in yaw rotations than in the rest of tests. There was a significant decrease of the mean error in the measurement of pitch rotations ($p < 0.001$) relative to each participant. There was an increase of missed frames ($p < 0.001$) in the free-to-move tests. The number of missed frames increased rapidly beyond 35° for yaw and pitch rotations, reaching a 100% error beyond 65°, on the other hand, the error for side-to-side roll motion was much lower, reaching 100% error beyond 85º. For the occlusion tests, the number of missed frames was also high ($p < 0.001$). Interestingly, there was no significant difference with the previous tests after calibration or with the whole body. Finally, there was a significant increase of missed frames in the 45° angle tests. Such results led the authors to conclude that the visibility of facial features is a determining factor for the correct estimation of head rotation with Kinect™ v2, and they added that a multikinect configuration could cover the full ROM, decreasing the percentage error.

IV. DISCUSSION

Before considering the use of Kinect™ technology in tasks involving human health, it is mandatory to ensure it has the required spatiotemporal accuracy and precision to meet valid and reliable results. Related research is extensive and has rendered interesting outcomes. For example, in [26] the authors concluded that the Kinect™ v1 and Asus Xtion sensors have higher metrical potential for low range applications and indoor environments with moderate accuracy requirements, as the uncertainty in their measurements increased linearly with distance. This is a well-known feature and agrees with other metrological assessments performed in [27] and [28], adding the latter (which compares both Kinect™ sensors) that Kinect™ v2 is much more stable in this respect. However, in [29] a metrological characterization of the Kinect™ v2 also found an uncertainty in the measurements directly related to distance. In this regard, Schmitz *et al.* [19] added that the error in the measurements increases and the resolution decreases as the sensor distance increases, reaching up to 4 cm of error at the maximum range of Kinect™ v1. In contrast, in assessing human body gait with Kinect™ v1, Clark *et al.* [18] showed that the detection of points with respect to distance had an excellent accuracy, but it is evident that such an outcome is directly related to the robustness of the skeleton tracking algorithm and not to the accuracy of the sensor itself.

Most studies performed so far have used a reputable MoCap system as a Gold Standard reference. Moreover, some of them have compared both technologies against precision measuring instruments, as in [19] and [23] an inclinometer and a goniometer respectively were used. Table I presents a summary of the technologies and tests addressed in this analysis. The results have been variable, and agreeing with [2], they are difficult to compare because they depend largely on the measured elements, the angle of measurement, and even the diversity of human motion features, in addition to the distance, as aforementioned. For example, Galna *et al.* [20] concluded that gross elements are measured with better accuracy than fine elements, Clark *et al.* [8] pointed out the inability to assess the internal and external rotation of the joint centers of the peripheral limbs, and Xu and McGorry [24] noted that measurement errors outside the measured range may vary because the accuracy of the sensor varies across the whole range of work.

TABLE I. SUMMARY OF THE STUDIES DISCUSSED IN THIS REVIEW

Paper	Technology / Distance	Reference System	Participants / Tests	Kinect V&R Conclusion
Clark-2012 [8]	Kinect™v1-SDK-LabView at 2.5m.	Vicon MX, 12 cameras.	20 healthy adults. Postural control: forward reach, lateral reach, and single-leg eyes-closed standing balance. 3 repetitions.	Valid and reliable to assess clinical kinematic strategies of postural control. Pelvis and sternum biases.
Clark-2013 [18]	Kinect™v1-SDK-LabView at 1.3m-3.8m.	Vicon MX, 12 cameras.	21 healthy adults. Gait speed, step length and time, stride length and time, and foot swing velocity.	Valid for some spatiotemporal components of gait. Poorly valid for discrete time points.
Bonnechère-2014 [2]	Kinect™v1-SDK at 2m.	Vicon MXT40S, 8 cameras (31 markers).	48 healthy adults. Shoulder abduction, elbow flexion, hip abduction, and knee flexion. 2 sessions, 10 repetitions.	Comparable to reference (reasonably valid, reliable). Measured ROM (joint centers) discrepancies.
Schmitz-2014 [19]	Kinect™v1-KinectFusion. -	MotionAnalysis Corp, 10 cams (14 markers). Inclinometer.	A jig (lower limb simile). 6 static postures: flexed, extended, abducted, adducted, externally rotated, internally rotated. 7 sessions.	Valid (2º) and reliable (1.1º) for single joint angle.
Galna-2014 [20]	Kinect™v1-SDK at 3m.	Vicon MX3+, 10 cameras.	9 adults with Parkinson's disease, 10 controls. Quiet standing, multidirectional reaching, stepping and walking on the spot, hand clasping, finger tapping, foot, leg agility, chair rising, and hand pronation.	Valid for timing and gross spatial characteristics of clinically relevant movements but not with the same spatial accuracy for smaller movements.
Yang-2014 [21]	Kinect™v1-SDK at 2.5m.	Optotrak Certus (20 markers).	9 healthy adults. Double limb stance with feet apart, feet together, and single limb stance. 2 repetitions.	Valid and reliable for standing balance when COM parameters are properly calibrated by linear equations.
Clark-2015 [22]	Kinect™v2-SDK at 2.5m.	Vicon, 9 cameras.	30 healthy adults. Static standing balance: eyes open vs closed and single vs double limb. Dynamic balance: forward reach, lateral reach, and limits of stability. 2 sessions.	Reasonably valid for some anterior-posterior tasks, poorly valid for medial-lateral tasks. Variable though comparable reliability.
Huber-2015 [23]	Kinect™v1-SDK. -	TrackStar. Goniometer.	10 healthy adults. Shoulder joint angle, four frontal static poses: 90º flexion/abduction, max flexion, external max rotation. Two sagittal poses: 90º flexion, max flexion. 2 repetitions.	Same as [21]. Poorly valid (significant discrepancies of ±5º) though reliable for shoulder joint angles.
Xu-2015 [24]	Kinect™v1-SDK at 2m. Kinect™v2-SDK at 2.5m.	Vicon MX, 12 cameras.	20 healthy adults. Standing postures: upright standing, left and right lateral bending, trunk flexion, left and right foot raise, stoop, and squat. Sitting postures: Normal sitting, left, right, and both feet under seat, elbows on knees, left leg over right, and right leg over left.	Higher validity for standing postures and upper limb joint center location than for sitting postures and lower limb joint center location.
Darby-2016 [25]	Kinect™v2-SDK.	Vicon MX, 10 cameras.	8 adults. Head ROM: left-to-right yaw, up-and-down pitch, side-to-side roll, under different situations.	Valid and reliable for head pose estimation where facial features are visible.

In general, in movements without occlusion the Kinect™ v1 showed high accuracy and precision, for example in the shoulder abduction tests in [2] and [23] and the jig tests in [19]. Consistently, their lowest results occurred in occluded movements; in [23] the tests were repeated from a sagittal view to avoid occlusion, but there was no improvement in the results. Apparently, this problem inherent to vision-based MoCap systems has not been solved in Kinect™ v2, according to the results in [22] and [25]; however, an improvement, comparable to traditional systems, is reported in the measurement of non-occluded movements. Another significant result from the literature review, in agreement with [24], is that upper limb location and standing posture identification are usually more accurate than lower limb location and sitting posture identification.

V. CONCLUSION

This brief literature review summarizes some studies specifically oriented to the validity and/or reliability of any of the two Kinect sensors, emphasizing their approach on functional evaluation. There is a general consensus in scientific community about the undeniable advantages of the Kinect™ sensors over traditional systems: their low cost of only a fraction by comparison and their size and weight less than a laptop. Their most attractive advantage may be that they do not need the use of cumbersome and/or bulky elements or a long calibrating time to acquire data from the human body. These qualities give great potential to Kinect™ technology in human motion analysis fields, especially for clinical assessment or rehabilitation support in some motor impairment conditions. In addition, they make it possible for the first time to perform such tasks at non-dedicated facilities, for example, the patient's home.

Results achieved from this analysis suggest that the Kinect™ sensors, especially the second generation, are capable of identifying with sufficient accuracy and precision the posture and movement of the human body. In contrast, poor results are reported in fine motor elements, particularly in Kinect™ v1, as well as in body or joint rotations, in unusual postures, or in occlusions or self-occlusions, among others. A comprehensive review of literature is needed to strengthen conclusions. However, the evidence indicates that the validity and reliability of the Kinect™ technology depends to a large extent on the particular application. For functional assessment applications, a feasible conclusion is that the information provided by the Kinect™ sensors is clinically comparable to that of traditional systems, with minimal discrepancies but with significant advantages.

REFERENCES

[1] D. Regazzoni, G. De Vecchi, and C. Rizzi, "RGB cams vs RGB-D sensors : Low cost motion capture technologies performances and limitations," Journal of Manufacturing Systems, vol. 33, no. 4, pp. 719–728, 2014. doi:10.1016/j.jmsy.2014.07.011.

[2] B. Bonnechère *et al.*, "Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry," Gait and Posture, vol. 39, no. 1, pp. 593–598, 2014. doi:10.1016/j.gaitpost.2013.09.018.

[3] P. Plantard *et al.*, "Pose Estimation with a Kinect for Ergonomic Studies: Evaluation of the Accuracy Using a Virtual Mannequin," Sensors, vol. 15, no. 1, pp. 1785–1803, Jan. 2015. doi:10.3390/s150101785.

[4] O. Geman and H. Costin, "Automatic Assessing of Tremor Severity Using Nonlinear Dynamics, Artificial Neural Networks and Neuro-Fuzzy Classifier," Advances in Electrical and Computer Engineering, vol. 14, no. 1, pp. 133–138, Feb. 2014. doi:10.4316/AECE.2014.01020.

[5] B. Müller *et al.*, "Validation of enhanced kinect sensor based motion capturing for gait assessment," PLOS ONE, vol. 12, no. 4, p. e0175813, Apr. 2017. doi:10.1371/journal.pone.0175813.

[6] G. Blumrosen *et al.*, "A Real-Time Kinect Signature-Based Patient Home Monitoring System," Sensors, vol. 16, no. 11, p. 1965, Nov. 2016. doi:10.3390/s16111965.

[7] Z. Ren *et al.*, "Robust part-based hand gesture recognition using kinect sensor," IEEE Transactions on Multimedia, vol. 15, no. 5, pp. 1110–1120, 2013. doi:10.1109/TMM.2013.2246148.

[8] R. a Clark *et al.*, "Validity of the Microsoft Kinect for assessment of postural control.," Gait & posture, vol. 36, no. 3, pp. 372–7, Jul. 2012. doi:10.1016/j.gaitpost.2012.03.033.

[9] S. Obdrzalek *et al.*, "Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population," in 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012, pp. 1188–1193. doi:10.1109/EMBC.2012.6346149.

[10] L. Chen, H. Wei, and J. Ferryman, "A survey of human motion analysis using depth imagery," Pattern Recognition Letters, vol. 34, no. 15, pp. 1995–2006, 2013. doi:10.1016/j.patrec.2013.02.006.

[11] R. Lun and W. Zhao, "A Survey of Applications and Human Motion Recognition with Microsoft Kinect," International Journal of Pattern Recognition and Artificial Intelligence, vol. 29, no. 5, p. 1555008, Aug. 2015. doi:10.1142/S0218001415550083.

[12] T. Mallick, P. P. Das, and A. K. Majumdar, "Characterizations of Noise in Kinect Depth Images: A Review," IEEE Sensors Journal, vol. 14, no. 6, pp. 1731–1740, Jun. 2014. doi:10.1109/JSEN.2014.2309987.

[13] P. K. Pisharady and M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review," Computer Vision and Image Understanding, vol. 141, pp. 152–165, Dec. 2015. doi:10.1016/j.cviu.2015.08.004.

[14] T. D'Orazio *et al.*, "Recent trends in gesture recognition: how depth data has improved classical approaches," Image and Vision Computing, vol. 52, pp. 56–72, Aug. 2016. doi:10.1016/j.imavis.2016.05.007.

[15] D. Webster and O. Celik, "Systematic review of Kinect applications in elderly care and stroke rehabilitation," Journal of neuroengineering and rehabilitation, vol. 11, p. 108, 2014. doi:10.1186/1743-0003-11-108.

[16] J. Han *et al.*, "Enhanced computer vision with Microsoft Kinect sensor: A review," IEEE Transactions on Cybernetics, vol. 43, no. 5, pp. 1318–1334, 2013. doi:10.1109/TCYB.2013.2265378.

[17] T. Breuer, C. Bodensteiner, and M. Arens, "Low-cost commodity depth sensor comparison and accuracy analysis," in Electro-Optical Remote Sensing, Photonic Technologies, and Applications, 2014, vol. 9250, p. 92500G. doi:10.1117/12.2067155.

[18] R. A. Clark *et al.*, "Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables," Journal of Biomechanics, vol. 46, no. 15, pp. 2722–2725, 2013. doi:10.1016/j.jbiomech.2013.08.011.

[19] A. Schmitz *et al.*, "Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system," Journal of Biomechanics, vol. 47, no. 2, pp. 587–591, 2014. doi:10.1016/j.jbiomech.2013.11.031.

[20] B. Galna *et al.*, "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease," Gait & posture, vol. 39, no. 4, pp. 1062–8, Apr. 2014. doi:10.1016/j.gaitpost.2014.01.008.

[21] Y. Yang *et al.*, "Reliability and Validity of Kinect RGB-D Sensor for Assessing Standing Balance," IEEE Sensors Journal, vol. 14, no. 5, pp. 1633–1638, May 2014. doi:10.1109/JSEN.2013.2296509.

[22] R. A. Clark *et al.*, "Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control," Gait and Posture, vol. 42, no. 2, pp. 210–213, 2015. doi:10.1016/j.gaitpost.2015.03.005.

[23] M. E. Huber *et al.*, "Validity and reliability of Kinect skeleton for measuring shoulder joint angles: A feasibility study," Physiotherapy (United Kingdom), vol. 101, no. 4, 2015. doi:10.1016/j.physio.2015.02.002.

[24] X. Xu and R. W. McGorry, "The validity of the first and second generation Microsoft Kinect for identifying joint center locations during static postures," Applied Ergonomics, vol. 49, pp. 47–54, 2015. doi:10.1016/j.apergo.2015.01.005.

[25] J. Darby *et al.*, "An evaluation of 3D head pose estimation using the Microsoft Kinect v2," Gait & Posture, vol. 48, pp. 83–88, 2016. doi:10.1016/j.gaitpost.2016.04.030.

[26] H. Gonzalez-Jorge *et al.*, "Metrological evaluation of Microsoft Kinect and Asus Xtion sensors," Measurement: Journal of the International Measurement Confederation, vol. 46, no. 6, pp. 1800–1806, 2013. doi:10.1016/j.measurement.2013.01.011.

[27] M. Bueno *et al.*, "Metrological evaluation of Kinect Fusion and its comparison with Microsoft Kinect sensor," Measurement: Journal of the International Measurement Confederation, vol. 73, pp. 137–145, 2015. doi:10.1016/j.measurement.2015.05.018.

[28] H. Gonzalez-Jorge *et al.*, "Metrological comparison between Kinect i and Kinect II sensors," Measurement: Journal of the International Measurement Confederation, vol. 70, pp. 21–26, 2015. doi:10.1016/j.measurement.2015.03.042.

[29] A. Corti *et al.*, "A metrological characterization of the Kinect V2 time-of-flight camera," Robotics and Autonomous Systems, vol. 75, pp. 584–594, Jan. 2016. doi:10.1016/j.robot.2015.09.024.