Accuracy and Reliability of Kinect Motion Sensing Input Device’s 3D Models: A Comparison to Direct Anthropometry and 2D Photogrammetry

Ahmed Mangoud Badr1*, Wael M. Mubarak Refai2, Mohamed Gaber El-Shal1, Ahmed Nasef Abdelhameed3

1Department of Orthodontics, Faculty of Dentistry, Aswan University, Tingar, Egypt; 2Department of Orthodontics, Faculty of Dentistry, Minia University, Dean of Faculty of Dentistry, Aswan University, Tingar, Egypt; 3Department of Orthodontics, Faculty of Dentistry, Minia University, Minya, Egypt

Abstract

AIM: This study aims to evaluate the accuracy and reliability of Kinect motion sensing input device’s three-dimensional (3D) models by comparing it with direct anthropometry and digital 2D photogrammetry.

MATERIALS AND METHODS: Six profiles and four frontal parameters were directly measured on the faces of 80 participants. The same measurements were repeated using two-dimensional (2D) photogrammetry and 3D images obtained from Kinect device. Another observer made the same measurements for 30% of the images obtained with 3D technique, and interobserver reproducibility was evaluated for 3D images. Intraobserver reproducibility was evaluated. Statistical analysis was conducted using the paired samples t-test, interclass correlation coefficient, and Bland-Altman limits of agreement.

RESULTS: The highest mean difference was 0.0084 mm between direct measurement and photogrammetry, 0.027 mm between direct measurement and 3D Kinect’s models, and 0.016 mm between photogrammetry and 3D Kinect’s. The lowest agreement value was 0.016 in the all parameter between the photogrammetry and 3D Kinect’s methods. Agreement between the two observers varied from 0.999 Sn-Me to 1 with the rest of linear measurements.

CONCLUSION: Measurements done using 3D Images obtained from Kinect device indicate that it may be an accurate and reliable imaging method for use in orthodontics. It also provides an easy low-cost 3D imaging technique that has become increasingly popular in clinical settings, offering advantages for surgical planning and outcome evaluation.

Introduction

After the introduction of the soft tissue paradigm in orthodontics, orthodontic approaches established on dental and skeletal structures [1]. If the soft tissues paradigm was objectively evaluated, efficient treatment planning can be produced, and the patient can be accurately assessed at the end of the treatment [2]. Various methods have been used to measure facial soft tissues such as direct anthropometry [3], two-dimensional (2D) photogrammetry [4], lateral cephalometry [5], cone-beam computed tomography (CBCT), and surface scanning methods (laser scanning, moiré topography, and the three-dimensional (3D) stereophotogrammetric method) [6], [7], [8]. Direct anthropometric measurement is positively a reliable and affordable method. Farkas et al. [3] conducted significant studies to create a large database of direct anthropometric measurements that can be used for facial measurements. Direct anthropometry, though considered as the gold standard for facial measurements, has a number of disadvantages; for example, it consumes time and requires patient compliance [9]. Frontal and profile photographs are generally used for photogrammetric measurement. On the one hand, 2D images (photograph, lateral cephalometry) are a snapshot of a dynamic object; therefore, they require cooperation only during acquisition and are easier to obtain than direct measurement [10]. When using 2D imaging methods, on the other hand, magnification and distortion issues are possible to be found, and many variables can affect the standard of measurement, such as illumination variations and object-camera distance. Another disadvantage to this technique is that significant insufficiencies may occur during the evaluation of the 2D cross-sectional images of a 3D object [11]. The limitations of previously mentioned methods were regulated using methods such as CBCT and laser scanners [12]. Because of the radiation used, computed tomography is a highly expensive and invasive product [13]. In the laser surface scanning method, if the time used for scanning is long, motion artifacts can appear [14]. That's
the reason why, to overcome these limitations, 3D stereophotogrammetry was developed. With this method, 3D images are acquired by combining photographs captured from various angles with synchronous digital cameras. The advantages of this method are the lack of motion artifacts because of the short imaging time, high color resolution, and the opportunity for administration without harming patients for repeated analyses, quick configuration, imaging through advanced software, ease of archiving, and 3D storage of patient images [15], [16]. This method can turn out to be a routine process in orthodontic practice when 3D stereophotogrammetry devices become affordable and accessible. Therefore, the present study aims to compare three measurement methods and evaluate the accuracy of 3D by comparing it with direct anthropometry. Furthermore, the study also intends to evaluate the intraobserver and interobserver reliability of 3D stereophotogrammetric measurements in 10 linear and six angular measurements.

Materials and Methods

After approval was obtained from Ethical Committee of Faculty of Dentistry, Minia University, the participants were informed verbally and in writing before the study, and volunteer consent forms were collected from all participants.

Sample

The present study was conducted with 80 participants between 25 and 45 years old. For an effect size of 0.4 at a 0.05 significance level, there could be more than 90% power with a sample size of 80. The mean age of the participants was 31.7 ± 3.4 years, and there were 48 males (60%) and 32 females (40%). The study included people who had no previous facial surgery, no craniofacial defects, and no specific scar tissue on the face. The 2D and 3D images were acquired within the same day, and direct measurements were made on the day of image acquisition. Measurements of 2D and 3D images were randomly made on a different day from direct measurements to avoid any chances of interference.

Direct measurements

By inspection and palpation, morphologic points required for linear measurements were determined and were then marked on the face. During the determination and measurement of the points, it was made sure that the patients were relaxed and seated with a natural head position and relaxed lips. A digital millimeter caliper (sliding) was applied to directly measure the distance between the four different points in the frontal plane (exocanthion, endocanthion, cheilion, and alare base) and the distance between seven different points in the sagittal plane (tragus, exocanthion, nasion, pronasale, subnasale, stomion, and menton). Measurements were taken under the same room conditions with the same illumination (Table 1).

### Table 1: Anthropometric landmarks and definition of linear distances

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Landmarks</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tr-Ex</td>
<td>Tragus, exocanthion (mm)</td>
<td>Sagittal linear measurement from tragus to exocanthion</td>
</tr>
<tr>
<td>N-Pn</td>
<td>Nasion, pronasale (mm)</td>
<td>Vertical linear measurement from nasion to pronasale</td>
</tr>
<tr>
<td>N-Sn</td>
<td>Nasion, subnasale (mm)</td>
<td>Vertical linear measurement of upper facial dimension as measured from nasion to subnasale</td>
</tr>
<tr>
<td>Sn-St</td>
<td>Subnasale, stomion (mm)</td>
<td>Vertical linear measurement of overall upper labial height from subnasale to stomion</td>
</tr>
<tr>
<td>Sn-Pn</td>
<td>Subnasale, pronasale (mm)</td>
<td>Sagittal linear measurement of nasal tip</td>
</tr>
<tr>
<td>Sn-Me</td>
<td>Subnasale, menton (mm)</td>
<td>Vertical linear measurement of lower facial dimension as measured from subnasale to stomion</td>
</tr>
<tr>
<td>Ex-Ex</td>
<td>Exocanthion right, exocanthion left (mm)</td>
<td>Transverse linear measurement of bilateral (lateral canthi) width from exocanthion right, or Ex (R), to exocanthion left, or Ex (L)</td>
</tr>
<tr>
<td>En-En</td>
<td>Endocanthion right, endocanthion left (mm)</td>
<td>Transverse linear measurement of intercanthal distance from endocanthion right, or En (R), to endocanthion left, or En (L)</td>
</tr>
<tr>
<td>Ch-Ch</td>
<td>Cheilion right, cheilion left (mm)</td>
<td>Transverse linear measurement of mouth width from cheilion right, or Ch (R), to cheilion left, or Ch (L)</td>
</tr>
<tr>
<td>Al-Al</td>
<td>Alare right, alare left (mm)</td>
<td>Transverse linear measurement of nasal width from alare right, or Al (R), to alare left, or Al (L)</td>
</tr>
<tr>
<td>NFA</td>
<td>Nasion, glabella, pronasale (u)</td>
<td>Angular measurement from nasion to glabella to pronasale (nasofrontal angle)</td>
</tr>
<tr>
<td>NLA</td>
<td>Columella, subnasale, labium superior (u)</td>
<td>Angular measurement from columella to subnasale to labium superior (nasolabial angle)</td>
</tr>
<tr>
<td>MLA</td>
<td>Labium inferior, supramental, pogonion (u)</td>
<td>Angular measurement from labium inferior to supramental to pogonion (mentolabial angle)</td>
</tr>
<tr>
<td>Middle 1/3</td>
<td>Nasion, tragus, subnasale (u)</td>
<td>Angular measurement from nasion to tragus to subnasale (angle of medium facial third)</td>
</tr>
<tr>
<td>Lower 1/3</td>
<td>Subnasale, tragus, menton (u)</td>
<td>Angular measurement from subnasale to tragus to menton (angle of inferior facial third)</td>
</tr>
<tr>
<td>Conv</td>
<td>Glabella, subnasale, pogonion (u)</td>
<td>Angular measurement of soft tissue profile from glabella to subnasale to pogonion</td>
</tr>
</tbody>
</table>

Cont Table 1. Anthropometric landmarks and definition of linear distances.

2D photogrammetric acquisition and measurements

All images were obtained by placing the participants in the same position as in the direct measurement, frontal and profile photographs were captured under the same illumination using a professional camera (Nikon D 560018–55 VR) AF-P DX NIKKOR 18–55 mm f/3.5–5.6G VR

A millimeter ruler was used to avoid magnification errors, and the ruler was carefully held in the same plane with the frontal and profile measurement points. All photographs were transferred to a computer, and calibration procedures were performed using View Box 4.0 software (dHAL Software, Kifissia, Greece). Ten linear and six angular measurements were made on the images.

2D Models Photogrammetric Acquisition and Measurements

3D scanner A Microsoft Kinect v2 module was utilized as the 3D scanner is used for gesture recognition and can capture raw 3D scan data. It
has a 1080p camera operating at 30 Hz which can capture a regular 2D image. Its depth sensor operates at 30 Hz and has a 512×424 sensor. The horizontal field of view is 70 degrees, while the vertical is 60 degrees. Hardware configuration uses a good graphics processing capabilities laptop (RAM > 4GB, dual-core, or multicore CPU) to handle the 3D data acquisition and rendering. A Lenovo IdeaPad L340 Gaming laptop was used with an NVIDIA GeForce GTX 1650. Furthermore, a frame rate of 16–20 fps was used. Software Better performance is obtained with using Windows 10 operating system and using a 3D scanning software as Microsoft 3D scan. Kinect software development kit (SDK) is also used for capturing raw 3D data. The 3D builder application is used to create the mesh, and MeshLab is used to clean, repair, and smoothen the mesh. Data were stored in ply format. All participants were placed in same position to capture the images and Kinect one (Microsoft Kinect scanner version 2) motion sensing input device was used to generate the 3D model.

**Statistical analysis of the data**

The collected data were coded, listed, and statistically analyzed using SPSS program (Statistical Package for the Social Sciences) software version 25.

Descriptive statistics were done for parametric (normally distributed) quantitative data by mean, standard deviation (SD), and minimum and maximum of range and for qualitative data by frequency and percentage. Distribution of the data was done by Kolmogorov–Smirnov test. Testing of agreement between each 2 methods was made by Bland-Altman plot followed by paired-samples t-test.

Reliability between different methods and intraobserver and interobserver agreement was made by interclass correlation coefficient (ICC).

The level of significance was taken at p ≤ 0.05.

The results were determined to be clinically acceptable at an arbitrary value of 1 mm between two different measurements. In the 3D images, to evaluate the interobserver agreement, a different observer made 30% of the measurements similar to the first observer. In addition, to indicate the intraobserver agreement, both observers made the same measurements 30 days after the initial measurements. Intraobserver and interobserver agreement for the 3D method were assessed similarly using the paired samples t-test, ICC, and the Bland-Altman limits of agreement. The results were stated to be clinically acceptable at an arbitrary value of 2 mm between observers and within an observer. The statistical significance level was 0.05 in all statistical analyses.

**Results**

**Linear measurements**

a. Tr-Ex

There was no statistically significant difference was found between (Anthro), (2D), and (3D) measurements where (p = 0.345).

b. B) N-Prn

There was a statistically significant difference was found between (Anthro), (2D), and (3D) measurements where (p < 0.001).

A statistically significant difference was found between (Anthro) and each of (2D) and (3D) measurements where (p = 0.015) and (p < 0.001).

Furthermore, a statistically significant difference was found between (2D) and (3D) measurements where (p < 0.001).

c. N-Sn

There was a statistically significant difference was found between (Anthro), (2D), and (3D) measurements where (p = 0.047).

A statistically significant difference was found between (Anthro) and each of (2D) and (3D) measurements where (p = 0.028) and (p = 0.046).

While no statistically significant difference was found between (2D) and (3D) measurements where (p = 0.722).

d. Sn-St

There was no statistically significant difference which was found between (Anthro), (2D), and (3D) measurements where (p = 0.219).

e. Sn-Prn

There was a statistically significant difference which was found between (Anthro), (2D), and (3D) measurements where (p = 0.023).

A statistically significant difference was found between (Anthro) and each of (2D) and (3D) measurements where (p = 0.004) and (p = 0.012).

While no statistically significant difference was found between (2D) and (3D) measurements where (p = 0.094).

f. Sn-Me

There was a statistically significant difference which was found between (Anthro), (2D), and (3D) measurements where (p = 0.032).

A statistically significant difference was found between (3D) and each of (Anthro) and (2D) measurements where (p = 0.025) and (p = 0.021).

While no statistically significant difference was found between (Anthro) and (2D) measurements where (p = 0.247).
There was a statistically significant difference which was found between (Anthro), (2D), and (3D) measurements where (p = 0.010). A statistically significant difference was found between (3D) and each of (Anthro) and (2D) measurements where (p = 0.018) and (p = 0.007).

While no statistically significant difference was found between (Anthro) and (2D) measurements where (p = 0.926).

There was no statistically significant difference was found between (Anthro), (2D), and (3D) measurements where (p = 0.475).

There was no statistically significant difference was found between (Anthro), (2D), and (3D) measurements where (p = 0.537).

There was no statistically significant difference which was found between (Anthro), (2D), and (3D) measurements where (p = 0.464) as shown in Table 2.

No statistically significant difference was found between (2D) and (3D) groups where (p = 0.802).

Middle third
No statistically significant difference was found between (2D) and (3D) groups where (p = 0.365).

Lower third
No statistically significant difference was found between (2D) and (3D) groups where (p = 0.109).

No statistically significant difference was found between (2D) and (3D) groups where (p = 0.582) as shown in Table 3.

Table 3: Relationship between three groups in different angular parameters

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Groups</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>S.D.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>126.700</td>
<td>134.100</td>
<td>130.415</td>
<td>1.242</td>
<td>0.581ns</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>126.700</td>
<td>134.100</td>
<td>130.419</td>
<td>1.239</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLA</td>
<td>92.900</td>
<td>98.000</td>
<td>95.105</td>
<td>1.279</td>
<td>0.043*</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>92.900</td>
<td>98.000</td>
<td>95.126</td>
<td>1.290</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLA</td>
<td>116.500</td>
<td>124.500</td>
<td>120.139</td>
<td>2.224</td>
<td>0.802ns</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>116.500</td>
<td>124.500</td>
<td>120.139</td>
<td>2.223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle third</td>
<td>23.000</td>
<td>29.500</td>
<td>26.554</td>
<td>1.620</td>
<td>0.365ns</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>23.000</td>
<td>29.500</td>
<td>26.530</td>
<td>1.617</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower third</td>
<td>30.500</td>
<td>39.100</td>
<td>34.359</td>
<td>2.340</td>
<td>0.109ns</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>30.500</td>
<td>39.100</td>
<td>34.359</td>
<td>2.340</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conv</td>
<td>163.800</td>
<td>178.900</td>
<td>169.589</td>
<td>3.554</td>
<td>0.582ns</td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td>163.800</td>
<td>178.900</td>
<td>169.594</td>
<td>3.533</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S.D.: Standard deviation. Min.: Minimum value. Max.: Maximum value. *: Significant (p ≤ 0.05), ns: Non-significant (p > 0.05).

Both inter- and intraobserver measurements showed extremely high (ICC) in both linear and angular measurements where (ICC) was (1) in all parameters except for (Sn-Me) which was (0.999), which all states an extraordinarily strong reliability and agreement between readings.

Both inter- and intraobserver measurements showed extremely high (ICC) in both linear and angular measurements where (ICC) was (1) in all parameters except for (NFA) and (lower third) parameters in inter measurements and in (NFA), (NLA), and (lower third) parameters in intra measurements where (ICC) were (0.999), which all states an extraordinarily strong reliability and agreement between readings.

Both inter- and intraobserver measurements showed extremely high (ICC) in both linear and angular measurements where (ICC) was (1) in all parameters,
which all states an extraordinarily strong reliability and agreement between readings.

Discussion

The present study evaluated the accuracy and reliability of 3D models of the Kinect. A very high level of agreement was found between direct anthropometry, photogrammetry, and 3D models since the highest mean difference was 0.5 mm. For intraobserver and interobserver reliability, the mean difference was <1 mm in linear measurements and <2° for angular measurements in 3D images. These values were considered to be clinically insignificant. In addition to clinical evaluation, quantitative evaluations performed in facial soft tissues are necessary for assessing treatment goals and treatment results. CBCT can be used for three dimensions soft tissue analysis despite having several disadvantages [12]. Besides, the measurement difference between CBCT and 3D models appeared to be clinically insignificant [17]. The use of Kinect's 3D models, hence, is deemed adequate in imaging soft tissues and measuring facial soft tissues. The direct anthropometry method is not routinely used in clinical practice; however, it is vital to provide actual measurement results when accurate and careful measurements are made. In this study, the mean difference was found to be 0.5 mm between these two methods. The validity of 3D models was methodologically evaluated using devices of different types and brands. Such studies have found >1 mm mean difference between different methods and between intraobserver and interobserver measurement points [9], [17], [18], [19], [20]. In the literature, deviations are reported to be mostly caused by observer errors during the placement of anthropometric points [21], [22]. Plooij et al. [23] found intraobserver reliability varying between 0.90 and 0.99. They found interobserver agreement above 0.8 in most points and stated that reproducibility was substantially lower than 0.5 mm. Khambay et al. [24] reported that the grand mean of the precision calculated across subjects along all axes for all landmarks was 0.827 mm. Lubbers et al. [25] found a reproducibility error of 0.5 mm. Lu¨bbers et al. examined the precision and accuracy of 3D stereophotogrammetry and declared a mean global error between 0.1 and 0.5 mm. The greatest and most significant difference was 1.42 mm in the N-Pm measurement for interobserver agreement; however, the difference was 2 mm for any parameters. The highest interobserver agreement was in the Al-Al measurement (95% CI, –0.21, 0.56; mean 5 0.17 mm). Agreement between measurements was above 0.9 in all measurements based on ICC results. These findings are consistent with the studies by Aldridge et al. [19], Wong et al. [9], and Schaaf et al. [26]. According to Heike et al. [27] and Junqueira-Júnior et al. [28], the intra-rater reliability correlation coefficients for the 3D stereophotogrammetric images were ≥0.95 for 26 of the 30 measurements and mean absolute differences were 1 mm. The present study found agreement between 2D photogrammetry and other methods, which suggests that 2D measurements can be safely used in images collected with accurate technique and attention. Farkas et al. [29] and Wellens et al. [30] reported that, by the distortions in photogrammetry, the difference between direct anthropometry and 2D photogrammetry may be caused. It seems crucial to make the calibration in the same plane as the measuring points in 2D measurements. In addition, it is possible to reuse such data when necessary and to use such data by comparing it with other methods. For reliability, it is pivotal to have a clear acquisition of the facial regions to be measured. There may be image errors in the ear region (e.g., tragus) in systems that can capture image up to 180u. The Kinect device, which was used in the present study, can be converted into different modules according to the desired imaging area. In the present study, 360u images were acquired using five modular units (front: Two, rear: Two, and top: One) to avoid image loss all over the face, including the ear region. In the case of image distortions, particularly in the ear region, acquisition was repeated depending on the facial morphology. What comes as a priority of contemporary orthodontic diagnosis and treatment planning is evaluating the properties of facial soft tissue clinically and quantitatively. Many imaging techniques have been adopted to assess facial soft tissues [31], but 3D Kinect's models attracted more attention because of the many advantages mentioned. The reproducibility of the points determined for measurements is one of the most important factors determined for measurements. In fact, the low-range variation in the present study is known to have resulted from minor deviations that took place during the placement of the morphologic earlier. This method was concluded to be accurate and reliable to be integrated into orthodontic clinical practice. Furthermore, it is possible to process and analyze 3D Model images in different software in accordance with different clinical or research purposes.

Conclusion

- Measurements using Kinect's 3D models were consistent with both direct anthropometric and 2D photogrammetric measurements
- The high intraobserver and interobserver reproducibility suggest that this method can be reliably used
- Kinect's 3D models provide efficient and low-cost 3D models that have become increasingly popular in clinical settings, offering advantages for surgical planning and outcome evaluation.

https://oamjms.eu/index.php/mjms/index
Due to its superior gained data in relation to the conventional 2D Images, it also saves the chairside time and eliminates the need for high levels of patient compliance.

Research Ethical Approval

This study was made with the approval of the Ethical Committee of the Faculty of Dentistry, Minia University.

References

PMid:10534979

PMid:22889202

PMid:1643057

PMid:17920500

PMid:25095325

PMid:20691358

PMid:24513736

PMid:9456625

PMid:18452351

PMid:12594423

PMid:3421563

PMid:24710215

PMid:14704586

PMid:15071154

PMid:22879898

PMid:20135249

PMid:24182583

PMid:15352857

PMid:16158436

PMid:17980940


