

# Spatial Accuracy of a Low Cost High Resolution 3D Surface Imaging Device for Medical Applications

Berthold Shin<sup>1</sup>, Rajkumar Venkatramani<sup>2,3</sup>, Priya Borker<sup>4</sup>, Arthur Olch<sup>5,6</sup>, John Grimm<sup>7</sup>,  
Kenneth Wong<sup>5,8</sup>

<sup>1</sup>Touro College of Osteopathic Medicine, New York, USA

<sup>2</sup>Division of Hematology/Oncology, Children's Hospital Los Angeles, Los Angeles, USA

<sup>3</sup>Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, USA

<sup>4</sup>Case Western Reserve University School of Medicine, Cleveland, USA

<sup>5</sup>Radiation Oncology Program, Center for Cancer and Blood Diseases, Children's Hospital Los Angeles, Los Angeles, USA

<sup>6</sup>Department of Radiation Oncology, Keck School of Medicine, University of Southern California, Los Angeles, USA

<sup>7</sup>Department of Radiology, Children's Hospital Los Angeles, Los Angeles, USA

<sup>8</sup>Department of Radiation Oncology, Geffen School of Medicine, UCLA, Los Angeles, USA

Email: [bshin@student.touro.edu](mailto:bshin@student.touro.edu)

Received February 22, 2013; revised March 2, 2013; accepted April 15, 2013

Copyright © 2013 Berthold Shin *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

The Kinect is a low-cost motion-sensing device designed for Microsoft's Xbox 360. Software has been created that enables user to access data from the Kinect, enhancing its versatility. This study characterizes the spatial accuracy and precision of the Kinect for creating 3D images for use in medical applications. Measurements of distances between surface features on both flat and curved objects were made using 3D images created by the Kinect. These measurements were compared to control measurements made by a ruler, calipers or by a CT scan and using the ruler tools provided. Measurements on flat surfaces matched closely to control measurements, with average differences between the Kinect and control measurements of less than 2 mm and percent errors of less than 1%. Measurements on curved surfaces also matched control measurements but errors up to 3mm occurred when measuring protruding surface features or features along lateral boundaries of objects. The Kinect is an alternative to other 3D imaging devices such as CT scanners, laser scanners and photogrammetric devices. Alternative 3D meshing algorithms and combining images from multiple Kinects could resolve errors made when using the Kinect to measure features on curved surfaces. Medical applications include craniofacial anthropometry, radiotherapy patient positioning and surgical planning.

**Keywords:** Kinect; Anthropometric; Craniofacial Measurement; Surface Imaging

## 1. Introduction

A variety of 3D surface imaging devices is available for medical use. These include computed tomography (CT) scanners with surface rendering software, laser scanners specifically for radiotherapy applications (C-Rad, Uppsala, Sweden) or for general use (David 3D Laserscanner [1], Koblenz, Germany), photogrammetry devices such as the 3dMD imaging systems (3dMD, Atlanta, GA), and video surface imaging systems for radiotherapy use (Vision RT, Columbia, MD). CT scanners are ubiquitous in the US and are a highly accurate and sensitive source of 3D spatial information but also expose the patient to ionizing radiation and are expensive. Laser scanners such as the David 3D Laserscanner [1] (David Vision Systems, Koblenz, Germany) provide another source of 3D infor-

mation and have become increasingly available to the general public due to reductions in cost and size, but scan quality can be heavily influenced by the skill of the scanner and patient movements [2]. In some applications, such as radiotherapy patient positioning, the distance from the subject to the camera system must be at least 1 m to not interfere with the treatment process, a limitation that must be considered when evaluating the low cost solutions. Sophisticated photogrammetric systems such as the 3dMD system (3dMD, Atlanta, GA) are widely used by plastic surgeons and dentists to assess and plan treatments [3]. Though fast, these systems can be costly, while the use of older systems can be time consuming and error prone unless they are combined with other technologies (such as a laser scanner). The laser and

video-based surface imaging systems for radiotherapy patient positioning from C-rad and Vision RT cost well over \$100,000 US. The Kinect (Microsoft, Redmond, WA) was introduced in November 2010 as a low-cost (about \$200 US) motion-sensing input device for Microsoft's Xbox 360 gaming console. Since then, amateur and professional developers have recognized its versatility and have been able to harness its capabilities by creating software that enables the user to extract raw data from the Kinect's camera and depth sensor. The ability to access this data broadens the Kinect's usefulness, making it suitable in a variety of applications beyond that of a simple gaming device. The potential benefits for low-cost, high resolution 3D surface scanning extend to a variety of other medical fields besides radiation therapy such as maxillofacial surgery and surgical oncology [4]. Applications in maxillofacial surgery and surgical oncology include the documentation of facial growth, craniofacial anomalies, surgical treatment planning, and to measure and track linear distances between relevant facial features over time. Although surface area and Volumetric measurements using 3D imaging systems have been obtained for reconstructive surgery of the breast [5], linear distance measurements between standardized facial locations have been the mainstay for anthropometry [6,7]. Due to the low cost and wide availability of the Kinect, it is an attractive device for use in these medical applications where distances between anatomical features need to be measured, especially when having a stored image that can be analyzed conveniently and repeatedly is desirable.

The purpose of this study was to characterize the spatial accuracy and precision of the Kinect system for linear distance measurements from static 3D surface images. Although the Kinect is designed for motion sensing, and this would be a valuable additional function, spatial accuracy is the prerequisite for any use of this device. To this end, we designed a series of experiments comparing the accuracy of linear distance measurements made by the Kinect of flat and curved surfaces to several standard measuring devices including a ruler, calipers and CT scanner-based measurements.

## 2. Materials and Methods

The Kinect is composed of an Aptina MT9M112 CMOS (complementary metal-oxide semiconductor) RGB camera with a resolution of  $1280 \times 1024$  @10 FPS (frames per second) and  $640 \times 480$  @30 FPS, an Aptina MT-9M001 CMOS IR (infrared) camera at the same resolution, class 1,830 nm IR laser, tilt motor, accelerometer and microphone [8]. The Kinect uses structured light to determine the distance from the camera to an object, or depth. The IR laser projects a speckle pattern onto the

scene, which is detected by the IR camera. The Kinect compares the dimensions of the speckle pattern with known dimensions stored in its memory, and uses this comparison to calculate the depth at various points that comprise the scene [9]. No calibration process of the Kinect was performed during this study.

PrimeSense (PrimeSense, Cary, NC), the creator of the Kinect's camera technology has released hardware drivers, the NITE middleware and has joined with other developers to create the OpenNI framework application programming interface (API) [10]. The OpenNI framework serves as a way to connect the lower-level hardware drivers with the NITE middleware [11] that processes the data and higher-level applications that are accessed by the user. The primary application used to access data via the OpenNI framework was RGB-Demo v.0.5.0 [12]. RGB-Demo is an open source application which provides the user with RGB camera video, depth maps, screen captures and the ability to export scene information as a point cloud in .ply format. Once the scene information is exported, MeshLab 64 bit v.1.3.0 [13] is used to remove extraneous scene elements, rescale image dimensions, flip axes, and calculate vector normals at each vertex. The meshing algorithm, Robust Implicit Moving Least Squares (RIMLS) [14], was then used to create a 3D mesh. Attributes such as color and geometry are then transferred from the original point cloud to the mesh to create a recognizable 3D image.

To experimentally assess the precision of the Kinect we first defined a 3D coordinate system whose origin is at a point on the outer surface of the Kinect's IR camera with the positive Z-axis as a vector projecting outward from the camera parallel to the IR beam and the XY plane perpendicular to the Z-axis consisting of a horizontal X-axis and vertical Y-axis. We then conducted a series of linear distance measurements on flat surfaces to determine precision along the X-, Y- and Z-axes and then on curved surfaces of anatomically realistic head phantoms. The medical application relevant to Section 2.3 below is the anthropometric measurement of facial landmarks for surgical planning or follow up after surgery, or for documentation of facial variation over time or from normative values.

### 2.1. Graph Paper Measurements

To determine accuracy along the X- and Y-axes, we traced concentric squares of dimensions  $10 \times 10$  cm,  $20 \times 20$  cm, and  $25 \times 25$  cm on graph paper and placed Beekley CT-Spots<sup>®</sup> (2.3 mm artifact-free pellets) (Beekley Corporation, Bristol, CT) at the vertices of each square. The graph paper was placed 50 cm, 70 cm, 1 m, and 2m from the Kinect perpendicular to its IR beam and imaged. Photographs were taken for each view and point clouds and 3D meshes were created. Measurements of the length

and widths of each square were made using MeshLab's measuring tool by five observers and recorded and compared to the known distances.

## 2.2. Depth Accuracy Measurements

To determine accuracy along the Z axis, we placed blocks at distances 50 cm, 60 cm, 70 cm, 80 cm, 90 cm and 1 m as measured by the Kinect depth sensor. Each distance was then measured by five observers using a ruler and recorded.

## 2.3. Curved Surface Measurements

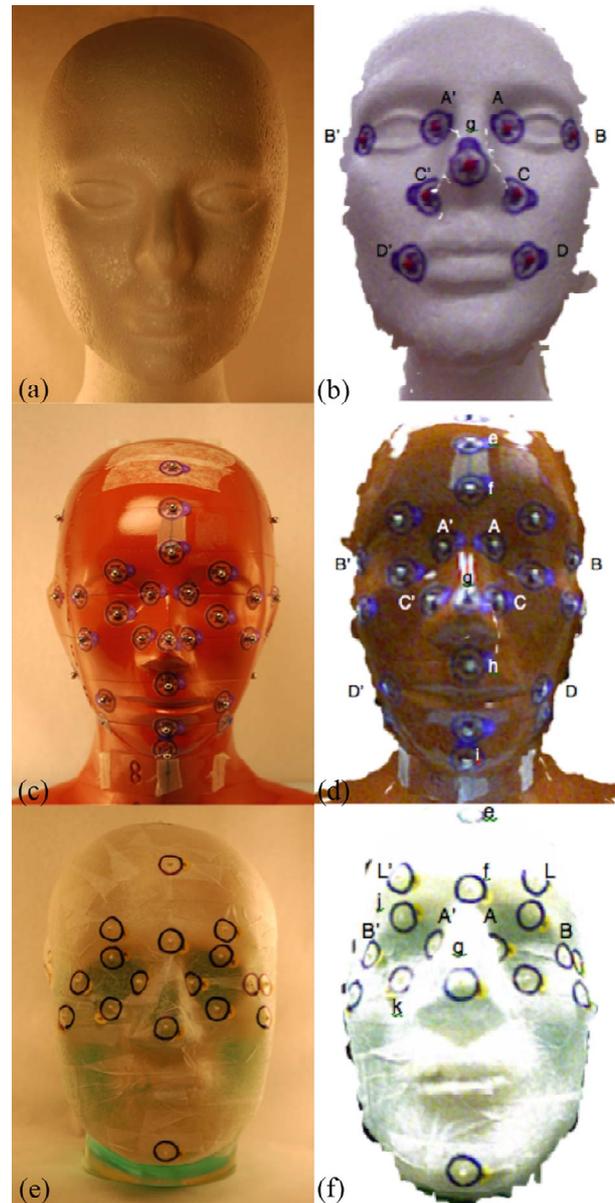
Following the measurements on flat surfaces, we sought to determine the accuracy of the Kinect when conducting linear distance measurements on curved surfaces resembling the human head of three different sizes (see **Figure 1**). We first placed painted ball bearings (BBs) (Beekley Z-Spots<sup>®</sup> 4.3 mm radiopaque pellets, Beekley Corporation, Bristol, CT) on a Styrofoam model head of dimensions similar to a child's head (head circumference of 48.5 cm). Measurements of these facial features were then taken using Mitutoyo calipers (Mitutoyo America Corporation, Aurora, IL) to a precision of 0.05 mm by five observers. Single en-face Kinect images of the Styrofoam model were then taken at 50 cm. Point clouds and 3D meshes were created, and measurements of the linear distances between the facial feature markers using the Kinect were made by five observers, recorded, and compared to the the caliper-measured distances.

4.3 mm Beekley Z-Spots were also placed on an anthropomorphic "red phantom" head of dimensions similar to an adolescent's head (head circumference of 52.2 cm). The red phantom was then scanned using a GE Lightspeed<sup>®</sup> (GE Healthcare, Little Chalfont, UK) CT scanner (120 KVp, 150 mA) with a slice spacing of 2.5 mm and imaged using the Kinect at 50 cm. Distance measurements between pairs of BBs demarcating various facial features were then taken by 5 observers using the CT scan images as well as the 3D meshes created using the Kinect. We repeated this procedure on an anthropomorphic "clear phantom" head, using Beekley CT-Spots 2.3 mm artifact-free pellets and a CT scan with slice spacing of 1.25 mm. Since the clear phantom consisted of a human skull surrounded by translucent material, we covered the surface of the phantom with surgical tape to provide an opaque surface that would be easily recognized by the Kinect cameras. Measurements of facial features were taken by five observers and recorded.

## 3. Results

### 3.1. Graph Paper Measurements

Measurements made on meshes of flat surfaces oriented



**Figure 1.** From top down, styrofoam model, red and clear phantoms ((a), (c), (e)) with corresponding mesh images ((b), (d), (f)). Facial features demarcated by BBs are indicated.

perpendicular to the camera axis (Z-axis) were most accurate at distances less than one meter, with average differences between Kinect measurements and actual dimensions being less than 1 mm and with percent differences of less than 1% and within 2 mm for both X and Y dimensions up to an imaging distance of 2 m (see **Table 1**). Kinect measurements consistently underestimated widths with negative mean differences at all distances, whereas measurements of length were overestimated mostly with positive mean differences. Mean absolute differences and mean percent differences of width increased with increasing distance, reflecting a decline in accuracy, whereas mean absolute and percent differences of length

**Table 1. Mean errors and standard deviations of Kinect measurements at 4 distances for markers on the corners of a 10 cm x 10 cm, 20 cm x 20 cm, and 25 cm x 25 cm square. All dimensions are in millimeters.**

Distance	Widths				Lengths			
	Mean Standard Deviation	Mean Difference	Mean Absolute Difference	Mean Percent Difference	Mean Standard Deviation	Mean Difference	Mean Absolute Difference	Mean Percent Difference
500	0.65	-0.2	0.5	0.41%	0.79	0.7	0.9	0.46%
700	0.59	-0.3	0.7	0.45%	0.82	0.1	0.7	0.44%
1000	0.95	-1.0	1.0	0.65%	1.44	-0.7	0.9	0.50%
2000	1.53	-0.1	1.4	0.82%	2.18	0.2	1.0	0.47%

fluctuated but did not increase significantly with distance. The mean standard deviation of measurements of both width and length increased with distance, which reflects a general decline in precision.

**3.2. Depth Accuracy Measurements**

Measurements of distance from the Kinect to objects at distances from 50 cm to 100 cm matched measurements made by a ruler to within 2 mm. These data are plotted along with the linear regression trend line that indicates excellent agreement over the range of distances measured (see **Figure 2**).

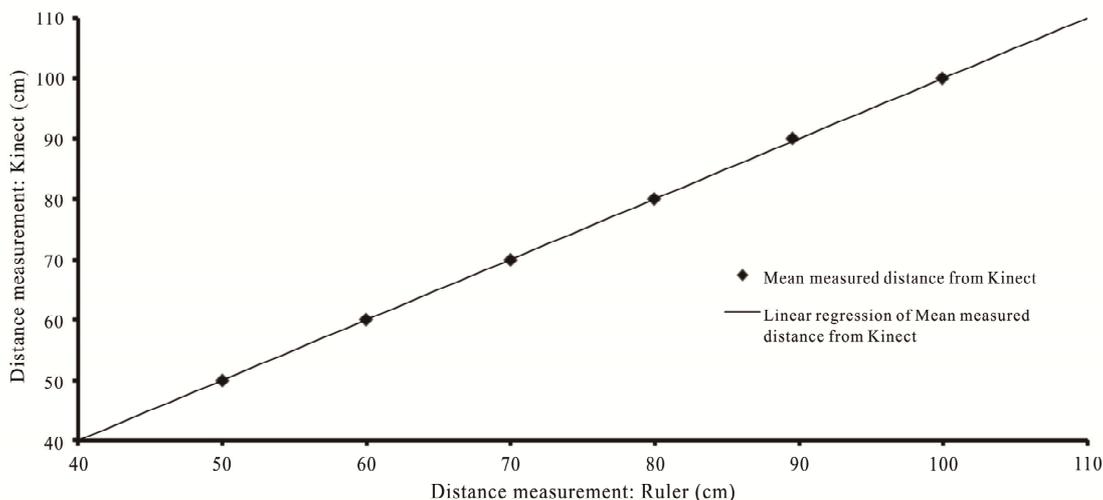
**3.3. Curved Surface Measurements**

The results for linear distance measurements made on curved surfaces are presented in **Figures 3-5**. These figures show the average distance measurements and 95% confidence intervals (CIs) for various anatomical features using the Kinect and calipers in the case of the styrofoam model, and the Kinect and CT scanner in the case of the red and clear phantoms. On the styrofoam model, Kinect measurements matched measurements made by calipers

to within 1 - 2 mm. The CIs were sub millimeter for 7 of the 9 measurements for both the Kinect and the caliper. Similarly, Kinect measurements made on the clear and red phantoms matched measurements made using the CT scanner within 1 - 3 mm. CIs were sub millimeter in 15 of 18 and 18 of 18 measurements for the Kinect and CT, respectively. The largest CI for Kinect was 1.45 mm. Kinect and CT measurements differed when measuring wide features whose landmarks were either at the lateral boundaries of the curved surfaces or that protruded significantly from a surrounding surface due to the Kinect’s single enface camera position.

**4. Discussion**

As our results show, the Kinect offers the user a low cost reasonably accurate and precise 3D imaging device for measuring linear distances in the facial region. Menna *et al.* provided equations for the theoretical XY and Z precision of Kinect measurements. At a distance of 500 mm, this gives 0.7 to 0.8 mm, which is in agreement with our CIs. Aldridge *et al.* used a commercial 3D camera system for measuring distances between facial landmarks on 15



**Figure 2. Measured distances of blocks from camera: Kinect vs Ruler.**

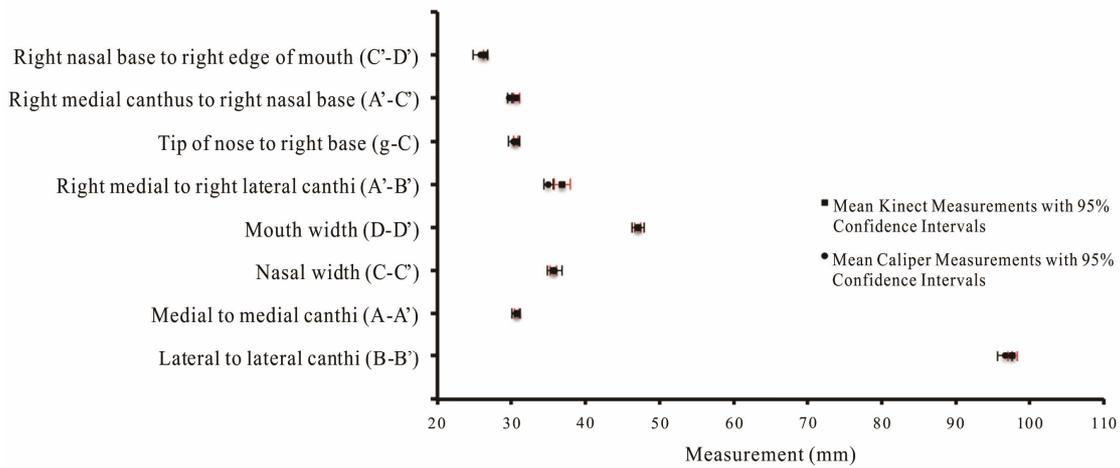


Figure 3. Measurements of anatomical features on the Styrofoam Model.

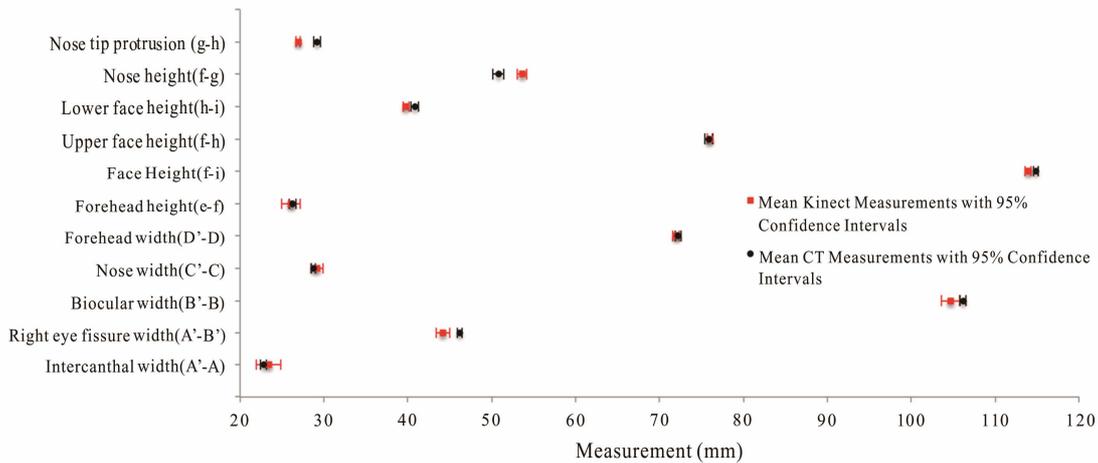


Figure 4. Measurements of anatomical features on the Red Phantom.

subjects. They reported a precision of about 1 mm for measurements with the 3D camera but did not compare secondary measures of the distances to determine accuracy. Krimmel *et al.* used a 3D surface imaging system to measure 21 standard anatomical dimensions in children with cleft lip and compared them to normative values. They relied on the manufacturer’s stated accuracy of 1 mm. In our study, compared to CT or caliper measurements of anthropomorphic head phantoms, a single Kinect camera was accurate to within 1 - 2 mm in most situations, with 95% confidence intervals generally less than 1mm. This degree of accuracy is sufficient for many applications, and would likely improve with multicamera images. Unlike CT and laser scanners, the Kinect does not expose the subject to ionizing radiation or potentially high intensity laser light. The entire apparatus is also relatively small and lightweight, making it portable and easy to setup.

Errors in measurement were likely due to discontinuities in the 3D meshes produced as a result of protruding

surface landmarks (such as the nose) and at the lateral edges of curved surfaces. The sudden change in distance between a landmark protruding from a curved surface or the lateral edges of a surface and objects in the background produces a “shadow” around the near surface. This shadow results in a lack of surface data around the discontinuity and results in a hole or empty space in the 3D mesh, introducing a source of error in surface measurements. This difficulty can be resolved using meshing techniques such as Poisson reconstruction [15] yielding a smoothed, “watertight” mesh, but one whose features seem blurred. A more satisfactory solution would be the combination of several images of the subject taken by multiple Kinects from different angles. Each image would effectively fill the holes left by the other images, yielding a more fully formed 3D image.

Further limitations of the Kinect involve the software designed to control it. The software used in this study was based on the OpenNI framework, an open source collaboration of several companies as well as amateur

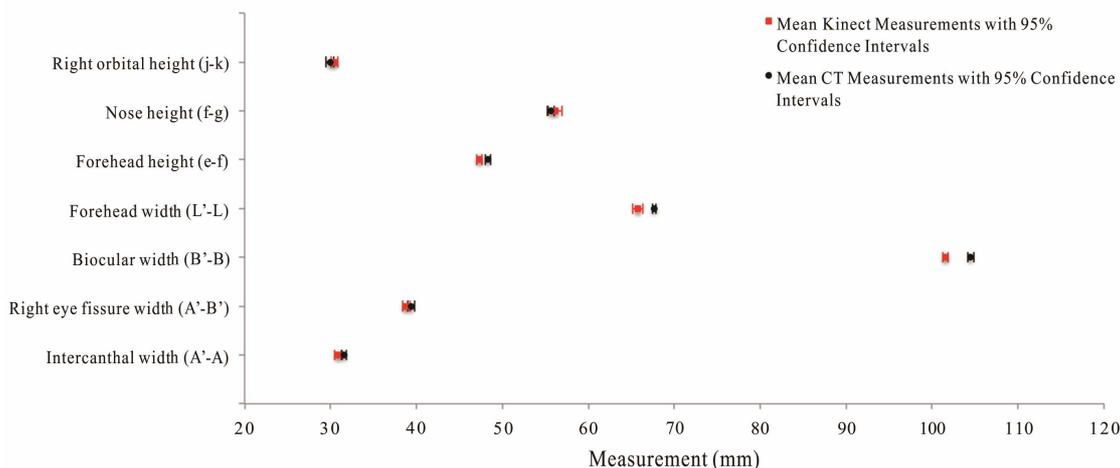


Figure 5. Measurements of anatomical features on the Clear Phantom.

developers. Accessing the data produced by the Kinect and putting it to use required the installation of a variety of software of unknown quality and stability with limited documentation and support. The release of Microsoft's Kinect for Windows SDK (software development kit) beta is a potential boon for the development of further applications for the Kinect, but could also result in the obsolescence of much of the open source software currently available.

Given the continuing development of novel applications for the Kinect and the continuous improvement of current applications, there are many possible directions in which research could proceed. One limitation of our study is that we validated the Kinect as providing accurate linear distance measurements for an area up to 25 cm × 25 cm at a distance of up to 2 m and on objects similar to the human head at distances of 0.5 m. The accuracy of the Kinect should be determined if used outside these parameters or for sizes and shapes not measured in this study.

The ability to access raw data from the Kinect for Xbox 360 has transformed the Kinect from a peripheral video gaming device into a versatile 3D imaging device with diverse applications. This study shows that the Kinect possesses sufficient accuracy and precision for it to serve as a safe, low-cost alternative to other more traditional medical 3D surface imaging systems. These results for linear distance measurements from static 3D images warrant further research into the accuracy of Kinect for Volumetric measurements, motion studies, and other medical applications requiring imaging parameters not covered in this study.

## REFERENCES

- [1] S. Winkelbach, "David 3D Scanner," 2011. <http://www.david-laserscanner.com>
- [2] S. B. Sholts, S. K. T. S. Wärmländer, L. M. Flores, K. W. P. Müller and P. L. Walker, "Variation in the Measurement of Cranial Volume and Surface Area Using 3D Laser Scanning Technology," *Journal of Forensic Sciences*, Vol. 55, No. 4, 2010, pp. 871-876. [doi:10.1111/j.1556-4029.2010.01380.x](https://doi.org/10.1111/j.1556-4029.2010.01380.x)
- [3] J. Y. Wong, A. K. Oh, E. Ohta, A. T. Hunt, G. F. Rogers J. B. Mulliken and C. K. Deutsch, "Validity and Reliability of Craniofacial Anthropometric Measurement of 3D Digital Photogrammetric Images," *Cleft Palate Craniofacial Journal*, Vol. 45, No. 3, 2008, pp. 232-239. [doi:10.1597/06-175](https://doi.org/10.1597/06-175)
- [4] C. H. Kau, S. Richmond, A. Incrapera, J. English and J. J. Xia, "Three-Dimensional Surface Acquisition Systems for the Study of Facial Morphology and Their Application to Maxillofacial Surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, Vol. 3, No. 2, 2007, pp. 97-110. [doi:10.1002/rcs.141](https://doi.org/10.1002/rcs.141)
- [5] A. Losken, H. Seify, D. D. Denson, A. A. Paredes Jr. and G. W. Carlson, "Validating Three-Dimensional Imaging of the Breast," *Annals of Plastic Surgery*, Vol. 54, No. 5, 2005, pp. 471-476. [doi:10.1097/01.sap.0000155278.87790.a1](https://doi.org/10.1097/01.sap.0000155278.87790.a1)
- [6] K. Aldridge, S. A. Boyadijev, G. T. Capone, V. B. DeLeon and J. T. Richtsmeier, "Precision and Error of Three-Dimensional Phenotypic Measures Acquired from 3dMD Photogrammetric Images," *American Journal of Medical Genetics Part A*, Vol. 138A, No. 3, 2005, pp. 247-253. [doi:10.1002/ajmg.a.30959](https://doi.org/10.1002/ajmg.a.30959)
- [7] M. Krimmel, S. Kluba, M. Bacher, K. Dietz and S. Reinert, "Digital Surface Photogrammetry for Anthropometric Analysis of the Cleft Infant Face," *Cleft Palate Craniofacial Journal*, Vol. 43, No. 3, 2006, pp. 350-355.
- [8] F. Menna, F. Remondino, R. Battisti and E. Nocerino. "Geometric Investigation of a Gaming Active Device," *SPIE Proceedings*, Vol. 8085, 2011, p. 15. [doi:10.1117/12.890070](https://doi.org/10.1117/12.890070)
- [9] A. Shpunt and Z. Zalevsky, "Three-Dimensional Sensing Using Speckle Patterns," U.S. Patent No. 0096783, 2009.
- [10] "OpenNI," 2011. <http://www.openni.org>

- [11] "NITE Middleware, PrimeSense Natural Interaction," 2011. <http://www.primesense.com/en/nite>
- [12] N. Burrus, "Kinect RGBDemo v0.5.0, Nicolas Burrus Homepage," 2011. <http://nicolas.burrus.name/index.php/Research/KinectRgbDemoV5>
- [13] "MeshLab," 2011. <http://meshlab.sourceforge.net>
- [14] A. C. Öztireli, G. Guennebaud and M. Gross, "Feature Pre-Serving Point Set Surfaces Based on Non-Linear Kernel Regression," *Computer Graphics Forum*, Vol. 28, No. 2, 2009, pp. 493-501. [doi:10.1111/j.1467-8659.2009.01388.x](https://doi.org/10.1111/j.1467-8659.2009.01388.x)
- [15] M. Kazhdan, M. Bolitho and H. Hoppe, "Poisson Surface Reconstruction," *Proceedings of the 4th Eurographics Symposium on Geometry Processing*, Cagliari, 26-28 June 2006, pp. 61-70.