

## LOW-COST 3D DEVICES AND LASER SCANNERS COMPARISON FOR THE APPLICATION IN ORTHOPEDIC CENTRES

D. F. Redaelli<sup>1\*</sup>, S. Gonizzi Barsanti<sup>1</sup>, P. Frascini<sup>2</sup>, E. Biffi<sup>2</sup>, G. Colombo<sup>1</sup>

<sup>1</sup> Department of Mechanics, Politecnico di Milano, Italy, (davidefelice.redaelli; sara.gonizzi; giorgio.colombo)@polimi.it

<sup>2</sup> Scientific Institute, IRCCS E. Medea, Via don Luigi Monza 20, 23842 Bosisio Parini, LC, Italy, (paolo.frascini; emilia.biffi)@bp.inf.it

### Commission II, WG 7

**KEY WORDS:** Low-cost 3D sensors; Biomedical applications; 3D Metrology; Resolution; Systematic error.

### ABSTRACT:

Low-cost 3D sensors are nowadays widely diffused and many different solutions are available on the market. Some of these devices were developed for entertaining purposes, but are used also for acquisition and processing of different 3D data with the aim of documentation, research and study. Given the fact that these sensors were not developed for this purpose, it is necessary to evaluate their use in the capturing process. This paper shows a preliminary research comparing the Kinect 1 and 2 by Microsoft, the Structure Sensor by Occipital and the O&P Scan by Rodin4D in a medical scenario (*i.e.* human body scans). In particular, these sensors were compared to Minolta Vivid 9i, chosen as reference because of its higher accuracy. Different test objects were analysed: a calibrated flat plane, for the evaluation of the systematic distance error for each device, and three different parts of a mannequin, used as samples of human body parts. The results showed that the use of a certified flat plane is a good starting point in characterizing the sensors, but a complete analysis with objects similar to the ones of the real context of application is required. For example, the Kinect 2 presented the best results among the low-cost sensors on the flat plane, while the Structure Sensor was more reliable on the mannequin parts.

### 1. OVERVIEW

In the last 30 years, orthopaedic centres started using the CAD-CAM (Computer Aided Design and Manufacturing) tools for the production of orthotics and prosthetics (O&P). Thus, during the past few years, the use of hand-held 3D scanners has become a common practice in the field of O&P.

In the beginning, the standard technique was the laser triangulation; however, because of the high costs of the devices, the 3D acquisition was not considered attractive to orthopaedic centres (Saunders et al., 1989).

Recently, new devices, more affordable thanks to cheaper technologies, have been introduced. The most famous device that signed the turning point is the Microsoft Kinect, which appeared on the market in 2010. It was originally designed for creating an interactive gaming experience, but its technology paved the way to different applications, among which the reverse engineering.

After the Microsoft launch, other triangulation-based low-cost devices were produced, such as Asus Xtion, PrimeSense and Structure Sensor. In 2013, Microsoft presented a new version of the Kinect 2 (Kinect for Xbox One) implemented with a Time of Flight technology.

Among these low-cost devices (budget under 200 €) and the laser scanners (budget over 20K €), other triangulation devices based on white or blue light have been presented. The most representative are Artec Eva and Creaform Go!SCAN 3D. The interest of the medical community in such devices derives from the possibility of acquiring the shape of the patient easily and quickly, with low cost systems that can be transferred from high-specialized institute to small satellite clinical centres spread over the territory. The producers of orthopaedic CAD-CAM solutions started to be interested in all these kinds of devices; therefore, they started adapting them to the biomedical applications, creating both software and new proprietary

hardware devices. Examples are the Rodin4D app to use directly the Structure Sensor with Apple iPad and iPhone, or the Biosculptor and the Vorum systems, composed of laser scanners or cheaper sensors and dedicated proprietary software.

These systems allow not only saving and modifying the models, but also sending them directly to a 3D printer or a CNC (Computer Numerical Control) machine for the creation of a physical cast model for orthoses or prostheses production, passing through their built-in CAM modules.

The aim of this paper is to evaluate and compare the metrological accuracy of a couple of devices currently used for medical purposes in an orthopaedic centre (*i.e.* the Structure Sensor by Occipital and the O&P Scan by Rodin4D) and low cost sensors that may be used for human body surface scan (*i.e.* Microsoft Kinect 1 and 2) and compare the results with a professional triangulation based laser scanner, the Minolta Vivid 9i.

The starting point of the research was to evaluate the systematic error of each device with the use of a calibrated plane in order to highlight which was the best to be used as a reference, as demonstrated by (Guidi et al., 2016). The second step consisted in surveying three different test objects, representative of the biomedical application, and comparing the results with the reference. A polystyrene hand, a chest and a thigh from a mannequin were chosen to replicated the parts of the body of a patient, considering the different level of detail and dimensions. The two Kinects were created as steady units for motion capture, considering the DOF of the objects in front of them. Until now, almost all the research in 3D modelling with low-cost devices regarded only the Kinect, both version 1 and 2. The first Kinect was released in 2010 starting from the Project Natal purposes. From proprietary device, the Kinect was soon opened to different aims, especially the use of the device as a moving unit for the acquisition of a steady environment and the collection of its geometric information as 3D data. The

KinectFusion was the first project analysing the possibility to calculate the 6 DOF of a rigid object with a handheld Kinect, allowing also aligning the single acquisitions for the creation of a 3D model (Newcombe et al., 2011). Several publications dealt with the possibility of using this device as a 3D acquisition instrument considering also its calibration to evaluate its potential as a low cost 3D instrument (Lachat et al., 2015; Kourosh, Elberink, 2012; Pagliari et al., 2014). The Kinect sensors, both version 1 and 2, have been used for several 3D applications: cultural heritage (Wenzel et al., 2012), for robotics (Ayrton et al., 2012) and for human body scanning (Tong et al., 2012).

In medical applications, the 3D scanning of a human body involved mainly highly rated laser scanners for acquiring the parts of interest, in order to improve diagnosis or to facilitate the creation of 3D printed orthosis or prosthesis (Baronio et al., 2016; Telfer, Woodburn, 2010). This device was used as a tracking system (Wang et al., 2012), for rehabilitation (Lange et al., 2011), for foot orthoses (Dombroski et al., 2014) and for improving the design of leg prosthesis (Colombo et al., 2016) with the aim to find a proper way to acquire 3D data using low cost sensors. On the other hand, recently the Structure Sensor was developed for acquiring 3D data of the environment simply connecting it to an iPhone or an iPad. No available researches were found regarding the use of the Structure Sensor in orthopaedic applications. Problems related to motion and deformation of the acquired data have been analysed by (Volonghi et al 2018), proposing and testing a deformable alignment algorithm both on static and real time acquisition with structured light devices.

This paper aims at filling the gap, testing the devices described below in acquiring organic shapes of mannequin parts, representing those human body parts that can be acquired in the orthotics and prosthetics (O&P) applications.

## 2. METHODS AND TOOLS

### 2.1 Devices

Three low-cost general-purpose devices were evaluated in this paper, the Kinect 1 and 2 by Microsoft, the Structure Sensor by Occipital. In addition, the O&P Scan by Rodin4D, specifically created for medical purposes, was also tested. All the four different devices have been compared with a Vivid Minolta 9i, also evaluated with respect to the systematic error.

The Microsoft Kinect 1 was released for XBOX360 and is based on the structured light technology. The device uses a low number of patterns to obtain a depth estimation of the scenery at 30 FPS (Frame Per Second). It is composed of two cameras, a colour RGB (Red, Green and Blue) and a monochrome NIR (Near InfraRed) camera, and a NIR projector with a laser diode of 850 nm wavelength. The baseline between the projector and the NIR camera is 7.5cm. The device uses triangulation technique to compute the depth information, thanks to the known and fixed dot pattern of the NIR projector to illuminate the scene.

The Microsoft Kinect 2 is based on a system originally patented by Canesta (Payne et al., 2014). It involves the modulation of an IR light source with a square wave and a flash camera to determine the distance to the object, by measuring the round trip travel time at each pixel of an amplitude-modulated light going from the source to the target and back.

The time needed by the light for reaching each pixel of the flash camera is evaluated by detecting the phase shift between the square wave and the signal received by each pixel. The device has a  $512 \times 424$  depth image sensor where each  $10 \mu\text{m} \times 10 \mu\text{m}$

pixel incorporates a TOF (Time Of Flight) detector that operates using the Quantum Efficiency Modulation (QEM). As shown in (Bamji et al., 2015; Sell and O'Connor, 2014) this technique uses two different modulating frequencies (80MHz and 100 MHz) for solving possible range ambiguities.

The Occipital Structure Sensor was created to be coupled with Apple iPhone and iPad in order to have cheap and user-friendly hand device to capture 3D environments. It can also be used on a computer through its specific software Skanect. The device is based on triangulation, it projects a speckle pattern of near-IR light on the scene captured by the infrared camera and correlated to a reference pattern of a plane which position is at a known distance from the sensor. For medical purposes, the Structure Sensor can be coupled with Captevia, an application for iPad, iPhone and iPod touch, available for free in the Apple Store that can be used for generating and saving 3D scan files on the Apple devices. The system is made by an IR projector, a  $640 \times 480$  CMOS IR camera dedicated to range sensing and a second  $640 \times 480$  CMOS colour camera for capturing the colour image associated to the depth map (Guidi et al., 2016).

The O&P Scan by Rodin is an handheld 3D digitizer to be used in the medical field for acquiring the parts of the body of the patient to be analysed. The O&P Scan uses a magnetic field to position itself in space, and gets the 3D surface from a laser line and from a camera with an accuracy up to  $\pm 0.5\text{mm}$  on the surface. The device is coupled with the Rodin4D CAD/CAM and the real time display on the PC allows to control all the operations. It is certificated to be able in acquiring data at a distance of maximum 40 cm and is equipped with a mini sensor to be attached to the patient's body in order to correct his movements.

Finally, the Vivid Minolta 9i is based on the principle of triangulation. The device measures  $640 \times 480$  points with one scan, simultaneously acquiring surface shape data and colour image data. It is stated to have an accuracy of  $\pm 0.05 \text{ mm}$  and a precision of  $0.008\text{mm}$ .

Device	Range [m]	Declared accuracy [mm]	Cost [€]
Minolta Vivid 9i (WIDE Lens)	0.5 – 2.5	0.1 – 0.2	~ 80k
Kinect 1	0.5 – 6	Non-available	~ 100
Structure Sensor	0.3 – 5	~ 4	~ 400
Kinect 2	0.4 – 5	Non-available	~ 200
Rodin4D O&P Scan	0.05 – 0.3	0.5	~ 10k

Table 1. Specifications of the tested devices.

### 2.2 Types of error

The International Vocabulary of Metrology (Joint Committee For Guides In Metrology (JCGM), 2008) stated that the measurement uncertainty of each equipment or device is affected by the systematic errors – associated to the concept of accuracy – and unavoidable random errors – associated with the concept of precision –, dependant by unpredictable causes like the electronic noise, that can be only statistically characterized for making the end-user aware of the measurement system intrinsic limitations (Guidi et al., 2016).

The global error can be evaluated by acquiring a certified test object and measuring the deviation of the model from the ideal one. The error can also be divided in spatial or temporal that is observable across multiple consecutive frames (Mallick et al., 2014).

In this work, the attention is focussed on the spatial error, neglecting the temporal error. This choice derives from (Guidi

et al., 2016), that already stated that, in the practical application, a spatial characterization allows to obtain more statistically significant results. Moreover, we noticed that the temporal error should be mitigated by the time averaging that software like Skanect perform. In addition, we could neglect the effects of temperature in time that have been proved to influence the performances of this range of devices over a timespan of several tens of minutes (DiFilippo and Jouaneh, 2015), because of the short time acquisition of the devices (maximum 3 minutes were required by the low cost devices).

### 2.3 Test objects

Four test objects were used to evaluate the different devices: a calibrated flat plane, used to calculate the systematic error (Guidi et al., 2016), and three anatomical parts of a mannequin to simulate the parts of a human body, chosen to investigate different dimensions and detail levels: a hand, a thigh and a chest. The characteristics of the different test objects are summarized in Table 2. The reference plane was made with float glass, which dimensions was 700 mm × 528 mm and a thickness of 11mm. The glass allowed to have a plane with a peak deviation from the theoretical plane in the order of few micrometres, suitable for testing the devices chosen for this research, characterized by measurement uncertainties in the range of millimetres. Due to the transparency of the material, not compliant with an active range sensing device, the surface was painted matt white, with the process used in the car industry. In this way, the painting was uniformly distributed without distorting the geometry of the plane.

The organic shapes of mannequin parts and their dimensions (Table 2) were chosen because representative of the human body segments that are usually acquired in the orthotics and prosthetics (O&P) applications. Despite the mannequin of the laboratory has also the arms with hands (Figure 1 – skin-toned one in the middle), the white polystyrene one was preferred because the geometrical detail level on the backside and on the nails was more realistic. Compared to a real hand (Figure 1 – left side), the tests objects have a more uniform color and are more smoothed. The mannequin chest and thigh are also a little bit more reflective, as visible on the hand in the middle of Figure 1, which has their same properties.



Figure 1. Tested white hand model on the right, compared to a real hand on the left and to the one of the mannequin in the middle.

### 2.4 Acquisition

Each test object was surveyed with all the instruments placed on a tripod at a fixed distance. The O&P Scan was the only exception because it is a handy scan and so it was used only in free hand movement, while the acquisitions with the Structure Sensor used both settings.

The Minolta laser scanner acquired with its proprietary software, the Structure Sensor and the Kinect 1 were connected to the pc by USB cable and used the Skanect software, the

Kinect 2 was connected to the pc by USB cable too, but used the Kinect Fusion Explorer available in the SDK Browser v2.0, and the Rodin4D O&P Scan hand laser scanner worked with its proprietary software.

The objects were all acquired emulating the application environment, that is the orthopaedic centre room with artificial lights (neon) and avoiding the use of markers. The reason of this choice is both because patients have usually difficulties staying still in respect to external references and because applying markers to the target objects could be trivial for the complete acquisition of their geometry.

Test Objects	Material	Dimensions (mm)	Colour
Calibrated Plane	Glass (opaque)	700x528	White
Hand	Polystyrene	200x95 approx.	White
Thigh	Polyethylene	420x130 approx.	Skin tone
Chest	Polyethylene	700x230 approx.	Skin tone

Table 2. Objects used in this project with material, dimension and colour specifications.

#### 2.4.1 Flat plane

The first test object surveyed was the reference plane to evaluate the systematic error for each device tested and to choose the reference for the following comparison. The calibrated flat plane was placed on a table and all the devices were put on a tripod slightly oblique in front of it, to avoid reflections, choosing the distance for the acquisition and all the setting parameters considering the specifics for each instrument. The survey started with the Minolta Vivid 9i, mounted with a middle lens at a distance of 1310 mm from the object. The choice of the lens depended on the dimension of the plane. Given the fact that the surface of the plane is opaque white the parameter of the intensity of the scan was set at 12. The acquisition of the scan was performed using the proprietary software Polygon Editing Tool (PET) that exported the file in \*.cdk format. The scan was then imported in Innometric Polyworks and exported in \*.obj.

For the Structure Sensor and the Microsoft Kinect 1, it was used the Skanect software, imposing a bounding box of 1x1x1 m placing the device at a distance of 1 m. Given the fact that this device acquires at a frame rate of 30/60 FPS, it was decided, after several tests, to stop the acquisition after 15 seconds, the same amount of time necessary for scanning with the Minolta. This setting gave the possibility to average the measurements, also because in this case the device was placed in a fixed position. The obtained 3D models were then exported in \*.obj at high resolution. The same setup was used for the Microsoft Kinect 2, using in this case the SDK software specifically implemented for this device, that allows to set also the parameters for the Voxels per meter and the Volume Voxel Resolution. To better understand the differences in the results changing these parameters, several tests were performed, concluding that the best setting was to fix the Voxels per meter at 768 and, the Volume Voxel Resolution at 512. Also in this case, the acquisition was stopped after 15 seconds, while the distance between the device and the object was set at 900 mm.

For the O&P Scan, the RODIN4D software was used. This device is a handy scan, and the acquisition of the plane was performed doing several strips to cover the entire surface of the plane. All the strips were automatically aligned. To obtain the

final 3D model in obj extension, the software RODIN4D NEO was used, with the lower value of smoothing to avoid deformation.

#### 2.4.2 Hand, thigh and chest.

For the acquisition of the hand, the Minolta Vivid 9i was coupled with a middle lens and placed at a distance of 800 mm for the back part of the hand, 870 mm for the front and at 940 mm for the tilted scans. The parameter regarding the light intensity was set equal to 13, and 28 scans were acquired moving the object in front of the laser scanner, and were then aligned and merged using Innometric Polyworks.

The Structure Sensor and the Microsoft Kinect 1 were placed at a distance from the object of 450 mm and 600 mm respectively setting the bounding box at  $0.3 \times 0.3 \times 0.3 \text{ m}^3$ , following the steps defined during the survey of the flat plane. In this case, the hand was placed on a stool that was rotated in front of the devices with two rounds.

The same setting was used for the acquisition of the object with the Microsoft Kinect 2, using its proprietary software and placing the device at a distance of 750 mm. The best setting was to fix the Voxels per meter at 512 and, the Volume Voxel Resolution at 256.

Given the dimensions of the thigh and the chest, the Minolta Vivid 9i was in this case coupled with a wide lens and placed at a distance of 1100 mm for the acquisition of the thigh and at a distance of 1350 mm for the chest. The intensity was set at 18 because the colour of the object was not so light and the object was rotated in front of the scanner. At the end of the acquisition, the 19 scans of the thigh and the 25 scans of the chest were imported, aligned and merged in Innometric Polyworks to obtain the two final 3D models.

The Structure Sensor and the Kinect 1 were placed, respectively, at 450 mm and 750 mm distance with the bounding box set at  $0.6 \times 0.6 \times 0.6 \text{ m}^3$  for the acquisition of the thigh and at 750 and 850 mm respectively with the bounding box set at  $0.8 \times 0.8 \times 0.8 \text{ m}^3$  for the chest. Also in this case, the object was placed on a rotating stool and two completed rounds were done. The Kinect 2 was placed at 800 mm from the object. For both the thigh and the chest, the best setting was to fix the Voxels per meter at 512 and, the Volume Voxel Resolution at 384.

In addition to the models acquired with Structure Sensor, Kinect 1 and 2 and O&P Scan, another model of each object has been acquired in free hand movement by the orthopaedic technician, with a Structure Sensor mounted on an iPad and using Captevia, the app developed by Rodin4D. The O&P Scan was moved in strips all around the objects, while the Structure Sensor barely in circles.

#### 2.5 Analysis

The first analysis regarded the estimation of the global error and the systematic error for the different devices analysed in this project, following the basic steps used by (Guidi et al., 2016). Considering the results of the analysis on the flat plane, the subsequent comparisons have been performed setting the models acquired with the Minolta laser scanner as reference.

In all the analyses, the models acquired with the Kinect 2 were scaled with a factor of 1000 in order to change the measurement units from meters to millimetres, like all the other models.

For the analysis of the distance errors with the flat plane, a Taubin smooth filter was used according to the previous work by (Guidi et al., 2016). The filter is basically a low-pass filter that does not apply any subdivision to the mesh, *i.e.* the noise, which is commonly known for having high-frequencies, should be removed without altering further the point positions.

#### 2.6 Hardware and software specs for the analyses

The analyses have been performed using CloudCompare, the Open Source Project 3D point cloud and mesh processing software. The version used is the latest stable release 2.9.1, installed on a Windows 10 notebook with i7-6700HQ CPU, 16GB of RAM memory and NVIDIA GeForce GTX 960M (2GB GDDR5 video memory). In addition, Autodesk Meshmixer (version 3.3.15) was used for the manual rough cleaning and MeshLab (version 2016.12) was used for applying the Taubin smooth filter.

#### 2.7 Processing

**2.7.1 Flat plane:** First, the acquired models of the plane were cleaned roughly from the elements that were clearly part of the surrounding environment.

Then in Cloud Compare an ideal plane was created using the "Primitive factory". The dimensions were set equal to the real plane ones, 700 by 528 mm.

After having imported the models, these were roughly aligned and registered using a partial overlap of 80%. This step allowed using only the points related to the flat plane discarding the surrounding.

Afterwards, the models have been cropped using a bounding box of 680 mm by 500 mm by 40 mm, aligned with the plane in order to effectively remove the surroundings and the borders, thus keeping only the points of the plane. A final fine registration has been performed setting the final overlap to 100%. This solution allowed aligning the scanned plane to the ideal plane, determining all the DOF parameters. This is clear for the orthogonal translation and the rotations, but the lateral shifts (in horizontal and vertical directions) have been avoided because the 100% of the remaining model is aligned with the ideal plane that has the same external dimensions. This means that all the points should be as close as possible to the plane  $680 \times 500 \text{ mm}^2$  and thus not be able to shift on an infinite plane.

At this point, the resulting cropped and aligned models were exported, and a Taubin smoothing filter (Taubin, 1995) has been applied onto them before being imported back to the comparison software. This step was performed using MeshLab (version 2016.12), setting  $\lambda=0.95$ ,  $\mu=-0.98$  and 50 iterations. These values have been set in compliance with the limits presented in (Taubin, 1995),  $\lambda>0$ ,  $\mu<-\lambda$  and the threshold  $k_{pb}(=0.0315) > 0$  (preferable between 0.01 and 0.1).

**2.7.2 Hand:** The models have been roughly cleaned removing most of the surrounding environment, each model separately. This step has been also a requirement in order to import all the models in a single comparison file.

Similarly, to the previous analysis, the models have been roughly aligned manually and then finely registered automatically, with the constraint of 90% final overlap.

Subsequently all the models, aligned in respect to the reference one, have been cropped in the region of the forearm at the same location. Finally, a fine automatic registration has been performed using the 100% of the final models.

**2.7.3 Thigh:** The analysis is very similar to the previous one (2.7.2), with the exception of a double cropping of the models, both in the distal and proximal regions of the thigh, and an added intermediate step of alignment, with the 95% of overlap, in order to help the alignment of the models, mainly for the ones of the Kinect and O&P Scan, because of the lower detail level of the objects.

**2.7.4 Chest:** For this object, the steps are exactly the same of the ones used for the hand, but having also here a double cropping: on the neck and on the lower part of the abdomen. These cuts have been due to the positioning of the object for scanning and to the relative position of the scanners in some of the acquisition. On the other hand, the non-realistic couplings for the arms of the mannequin have been left because acquired in all the models.

### 3. RESULTS

In order to understand better the following results, it is important to remember that the global error, obtained analysing the models without any filter, is composed of a systematic component, that can be highlighted applying the Taubin filter (Guidi et al., 2016) because related to low frequency, and a random error, that has high frequency and that could be obtained by subtracting the systematic error from the global one. Moreover, the lower the standard deviation values are, the lower the error is and the better the device performances are.

#### 3.1 Flat plane

The distance cloud/mesh has been computed between the ideal plane and each of the acquired models, both with and without the application of the Taubin filter. The results have been fitted using Gaussian distributions and the values of mesh average dimension and standard deviation are reported below in Table 3. Graphical results also depicted in Figure 2.

Device	Mesh average dimension [mm]	Std Dev [mm]	
		Without any filter	With Taubin filter
Minolta Vivid 9i	1.3	0.341	0.296
Kinect 1	1.9	3.098	3.055
Structure Sensor	1.9	1.809	1.705
Kinect 2	1.56	1.054	1.050
Rodin4D O&P Scan	2	4.492	4.435

Table 3. Values of mesh size and standard deviation of the cloud to mesh distance for the flat plane models before and after the application of the Taubin filter.

#### 3.2 Hand, thigh and chest

The distance cloud/mesh has been computed between each of the acquired models and the one obtained with the Minolta Vivid 9i. The results have been fitted using Gaussian distributions and the values of standard deviation are reported below with the value of average mesh size, Tables 4, 5 and 6. For the comparison of these models, no filter was applied. All the comparisons are represented in Figure 3.

Device	Mesh average dimension [mm]	Std Dev [mm]
Kinect 1	0.6	0.910
Structure Sensor fixed	0.6	0.521
Kinect 2	1.95	2.180
Rodin4D O&P Scan	1	0.710
Structure Sensor free hand	3.0	0.539

Table 4. Values of mesh size and standard deviation of the cloud to mesh distance for the hand models in respect to the one acquired with Minolta Vivid 9i.

Device	Mesh average dimension [mm]	Std Dev [mm]
Kinect 1	1.2	1.675
Structure Sensor fixed	1.2	0.560
Kinect 2	1.95	1.924
Rodin4D O&P Scan	2	1.955
Structure Sensor free hand	3.5	0.456

Table 5. Values of mesh size and standard deviation of the cloud to mesh distance for the thigh models in respect to the one acquired with Minolta Vivid 9i.

Device	Mesh average dimension [mm]	Std Dev [mm]
Kinect 1	1.6	1.678
Structure Sensor fixed	1.6	0.534
Kinect 2	1.95	2.951
Rodin4D O&P Scan	3	3.689
Structure Sensor free hand	4.5	0.628

Table 6. Values of mesh size and standard deviation of the cloud to mesh distance for the chest models in respect to the one acquired with Minolta Vivid 9i.

### 4. DISCUSSION AND CONCLUSIONS

In this chapter, we present some critical consideration regarding the results themselves and the possible conclusions about the devices.

#### 4.1 Flat plane

As it was expected, the Minolta laser scanner has the lowest values of standard deviation, both with and without the application of the Taubin filter (Table 3). Because of these results, it was selected as a reference for the comparison of the test objects. The results regarding the global error of the Structure Sensor and the Kinect 2 are coherent with the ones found by Guidi et al. (2016), considering that in the present tests, the acquisition lasted for 15 seconds and many frames were captured, so that the software applied a time averaging. The O&P Scan instead was expected to have better results than the 2 low-cost devices, but, as visible in Figure 2, it showed its two main disadvantages: the bad behaviour of the magnetic field for determining the relative position of the scanner with respect to the sensor, producing the spherical warping, and the small width of the laser blade, related to the short stand-off distance, that determined the need of multiple strips, comparable to the ones left by a paint brush. Moreover, the survey has to follow specific rules: the strips have to be contiguous with a small overlap and following a straightforward path.

For all the devices, the standard deviation after the application of the Taubin filter still gave high values, meaning that the systematic error is the one affecting more the 3D acquisition. This means that a good calibration would improve the accuracy of the survey and therefore the resulting 3D model. Indeed, as we can notice from the Figure 2, the pattern of the deviation with the Taubin filter (second row) is very similar to the ones without the filter, confirming that the systematic error prevails on the high frequency random noise.

A proper analysis has to be done evaluating the values of the errors of the Structure Sensor and of the Kinect 2, because of time averaging due to the 15 seconds acquisition. In (Guidi et

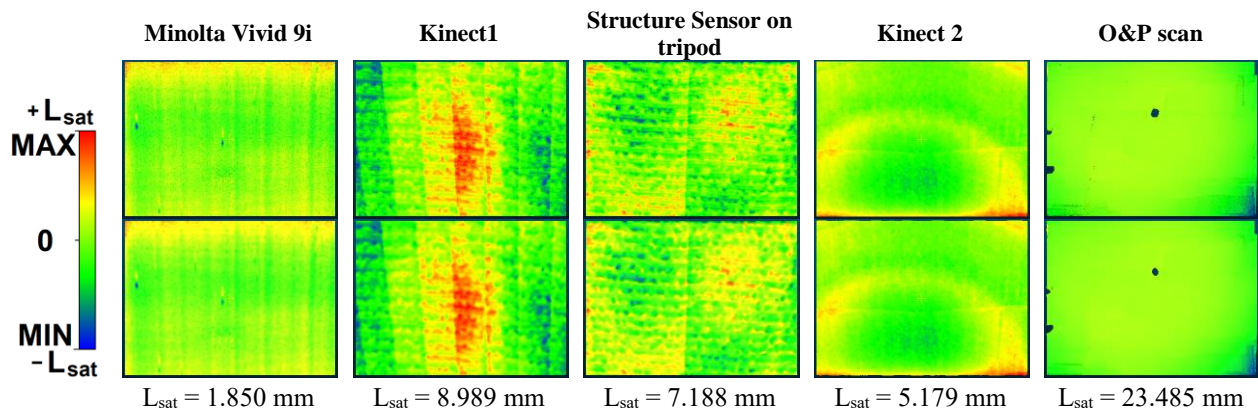


Figure 2. Comparison of the results on the flat plane the different devices, organized by column, in respect to an ideal plane. The first row regards the models without the filter, thus showing the global error. The second row contains the results after the application of the Taubin filter, that show the systematic error. Notice that the colour scale is the same in the column but different between the devices, and the colour saturation limits are indicated below with a single value, since the scale is symmetrical.

al., 2016), on the contrary, the random error on the acquisition of the plane, at more or less the same distance between the device and the object, was about half of the global error, but in that work just one frame was used avoiding the time averaging.

#### 4.2 Hand, thigh and chest

The first consideration regards the mesh dimension: the average mesh size among the different reference models acquired with Minolta Vivid 9i is approximately the same while for the other devices it varied because of the different bounding box settings. In addition, the O&P Scan, differently from what expected, showed an increase in the average mesh dimensions passing from the hand to the chest. The problem could be due both to the faster movement of the operator and the additional step of conversion achieved in the software Rodin4D Neo in order to be able to export the models in a standard non-proprietary file format, as STL file.

The Kinect 1 obtained good results in terms of standard deviation with the objects, better than the one obtained with the flat plane. Comparing the objects, the value of deviation on the hand seems to be better than the ones with the thigh and the chest. However, analysing the models in Figure 3, we can notice that the hand is particularly good only on the wrist, palm and back, but the reconstruction on the fingers is poor.

The Structure Sensor used on a tripod demonstrated to have a stable deviation in all the models with respect to the references. For both the Kinect 1 and the Structure Sensor, the values of the standard deviation showed that the averaging, due to the high number of frames acquired during the scans, and the paths followed in the relative motion object-device, produced better results.

The Kinect 2 showed sensibly higher values of standard deviation with all the 3D objects, in respect to the ones obtained with the flat plane. The main reason was already clear during the real-time view of the acquisition and depends on the very poor alignment achieved by the software. The models are visibly deformed, for example the hand has connected fingers and the thigh is warped, and the acquisitions themselves required slow and smooth movements not to lose the tracking. Moreover, trying to use the maximal resolution, it presented issues in the capturing frequency that decreased from the normal 30FPS down to 2 FPS.

The O&P Scan demonstrated, similarly to the Structure Sensor, that the alignment of the strips on 3D objects performs better

than on the flat plane. However, the device showed an increase in the error with the increase of the dimensions of the objects. This can be related again to alignment problems of the magnetic field technology. Probably the working environment, a room with some metal frames needed to keep the patient in certain postures, could have biased the position tracking of the device around bigger objects.

The Structure Sensor used as a hand-held device by the orthotist, mounted on an iPad with the Rodin4D app performed again quite well, with respect to the higher budget one that the same producer of O&P solution commercialised. The lower spatial resolution is clearly a setting of the app, due to the computational limits of the tablet in respect to a normal computer, but this is still compatible with the goal of virtualizing human parts in order to design and produce patient-specific product, such as an orthosis or a prosthesis.

#### 4.3 Application-related choice of the device

The Minolta Vivid 9i has the best resolution but the type of technology makes it inadequate for human body scanning. It requires multiple scans taking minutes to be completed and it is not easy to move around the patient's body. Thus, the patient would move changing also the geometry of the parts to be acquired, due to both joints and soft tissues movements.

The best device, among the others we tested in this project, resulted to be the Structure Sensor, not only for the low deviation from the reference models but also for the possibility to move around the objects without any wire and without the need of an additional tracking system.

For the Kinect 1, we can state that it could be still a very cheap and acceptable device for acquiring medium to large body parts, depending on the final goal. For example, if we have to develop a new wrist orthosis, the shape could be still similar to the real one but the error would not be acceptable because of the thin skin layer around the bones. On the other hand, the Kinect 1 would still be a valid device for creating a socket for the prosthesis of an above knee amputee, as the case described in (Colombo et al, 2016), because of the thicker layers of soft tissues around the bone that could mitigate the deviation.

Even though the Kinect 2 showed good performance on the flat plane, the poor alignment determined the inability in a realistic acquisition of patients for O&P purposes.

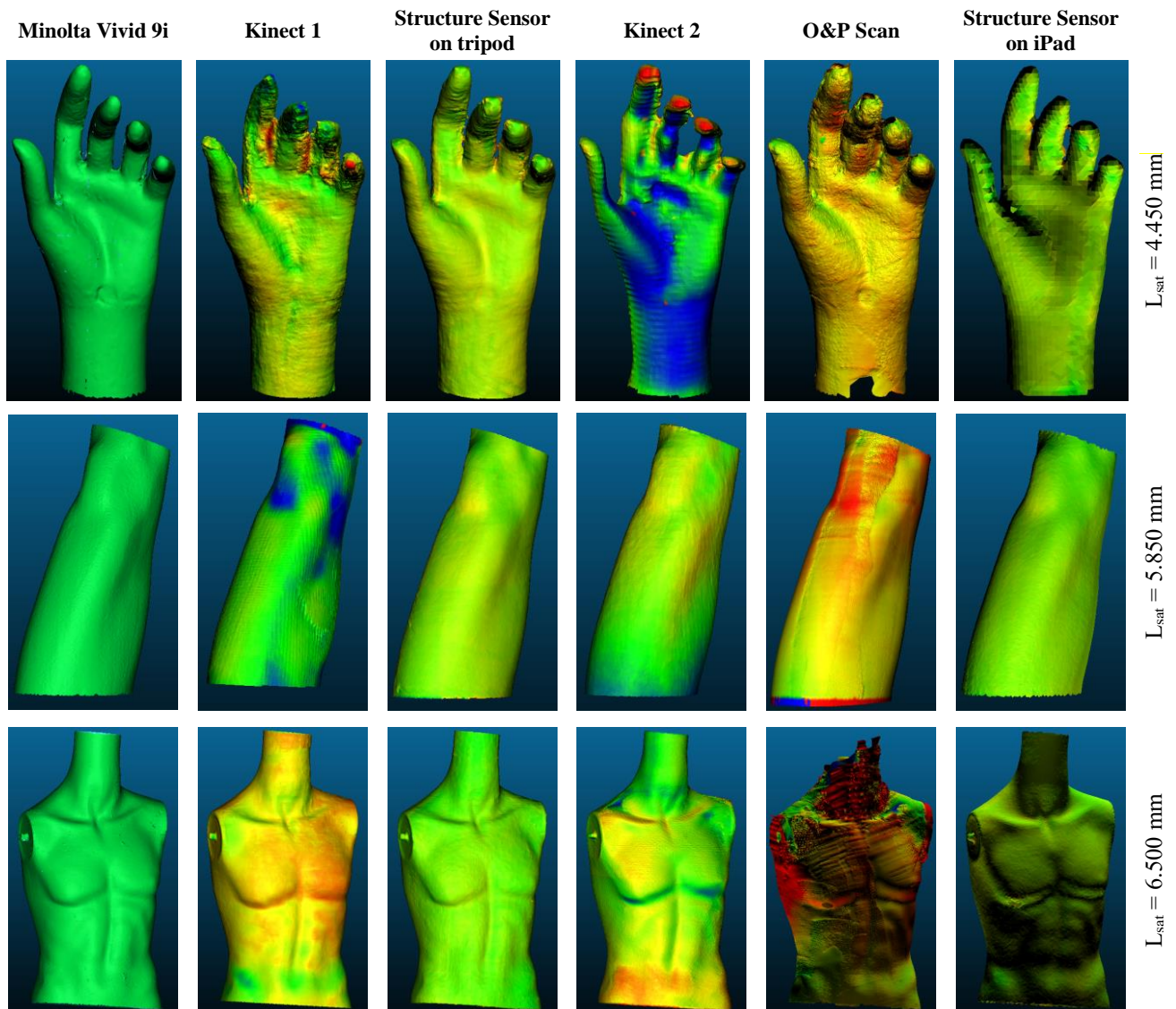


Figure 3. Results obtained by the different devices, organized by column, acquiring the three anatomical objects, organized by row. The models acquired with Minolta, first column, were used as reference in the comparison of all the other devices. The colour scale is the same gradual one depicted in the first column of Figure 2 and it has been set to be the same in the row with the limits for saturation indicated by the value  $L_{sat}$  on the right side, different for each object.

#### 4.4 Final considerations

The analyses showed that a simple test of the devices on a standard object, like the flat plane, is only a preliminary indication of the performances of the sensors, that will change in real application. At the same time, it remains an important step for the definition of the systematic error that could be removed using a proper calibration, thanks to standard testfields.

Therefore, a second test in conditions similar to the final application (*i.e.* scanning the human body in the case of this project) is always needed. Furthermore, depending on the field of use of the devices, it is suggested to create a standardized object that could be used as a specific reference, having also the original CAD model in order to compare the results without the bias due to the error of the device selected as reference.

#### REFERENCES

- Alnowami, M., Alnwaimi, B., Tahavori, F., Copland, M., Wells, K., 2012. A quantitative assessment of using the Kinect for Xbox360 for respiratory surface motion tracking. Proc. SPIE. doi:10.1117/12.911463
- Ayrton Oliver , Steven Kang , Burkhard C. Wünsche , Bruce MacDonald, 2012. Using the Kinect as a navigation sensor for mobile robotics, Proceedings of the 27th Conference on Image and Vision Computing New Zealand, p.509-514, Dunedin, New Zealand [doi>10.1145/2425836.2425932] (November 26-28, 2012)
- Bamji, C.S., O'Connor, P., Elkhatib, T., Mehta, S., Thompson, B., Prather, L.A., Snow, D., Akkaya, O.C., Daniel, A., Payne, A.D., Perry, T., Fenton, M., Chan, V.H., 2015. A 0.13  $\mu\text{m}$  CMOS System-on-Chip for a 512x424 Time-of-Flight Image

- Sensor with Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC. *IEEE J. Solid-State Circuits* 50, 303–319. doi:10.1109/JSSC.2014.2364270
- Baronio, G., Harran, S., Signoroni, A., 2016. A Critical Analysis of a Hand Orthosis Reverse Engineering and 3D Printing Process, *Applied Bionics and Biomechanics*, Vol. 6. doi: 10.1155/2016/8347478
- Colombo, G., Comotti, C., Redaelli, D.F. Regazzoni, D., Rizzi, C., Vitali, A., 2016. A method to improve prosthesis leg design based on pressure analysis at the socket-residual limb interface, *Proceedings of the ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference IDETC/CIE 2016*, Charlotte, North Carolina (August 21-24, 2016)
- Di Filippo, N.M., Jouaneh, M.K., 2015. Characterization of Different Microsoft Kinect Sensor Models. *IEEE Sens. J.* 15, 4554–4564. doi:10.1109/JSEN.2015.2422611
- Dombroski, C.E., Balsdon, M.E.R., Froats, A., 2014. The use of a low cost 3D scanning and printing tool in the manufacture of custom-made foot orthoses: a preliminary study, *Research Notes* 2014, 7:443
- Gonzalez-Jorge, H., Rodriguez-Gonzalvez, P., MartanezSanchez, J., Gonzolez-Aguilera, D., Arias, P., Gesto, M., DiazVilarino, L., 2015. Metrological comparison between Kinect i and Kinect II sensors. *Meas. J. Int. Meas. Confed.* 70, 21–26. doi:10.1016/j.measurement.2015.03.042
- Guidi, G., Gonizzi, S., and Micoli, L., 2016. 3D Capturing Performances of Low-Cost Range Sensors for Mass-Market Applications, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5, 33-40, <https://doi.org/10.5194/isprs-archives-XLI-B5-33-2016>.
- Khoshelham, K., Elberink, S.O., 2012. Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors* 12, 1437–1454. doi:10.3390/s120201437.
- Kouros, K., Elberink, S.O., 2012. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications, *Sensors* 2012, 12, 1437-1454; doi: 10.3390/s120201437.
- Lachat, E., Macher, H., Mittet, M.-A., Landes, T., and Grussenmeyer, P., 2015. First Experiences with Kinect V2 Sensor for Close Range 3D Modelling, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5/W4, 93-100, <https://doi.org/10.5194/isprsarchives-XL-5-W4-93-2015>.
- Lachat, E., Macher, H., Landes, T., Grussenmeyer, P., 2015. Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modelling. *Remote Sensing*, 2015, 7 (10), 13070-13097; doi: 10.3390/rs71013070.
- Lange, B., Chang, C.Y., Suma, E., Newman, B., Rizzo, A.S., Bolas, M., 2011. Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor, *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1831 - 1834.
- Mallick, T., Das, P.P., Majumdar, A.K., 2014. Characterizations of noise in Kinect depth images: A review. *IEEE Sens. J.* 14, 1731–1740. doi:10.1109/JSEN.2014.2309987
- Molnár, B., Toth, C.K., Detrekői, A., 2012. Accuracy Test of Microsoft Kinect for Human Morphologic Measurements. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XXXIXB3, 543–547. doi:10.5194/isprsarchives-XXXIX-B3-543-2012
- Pagliari, D., Menna, F., Roncella, R., Remondino, F., and Pinto, L., 2014. Kinect Fusion improvement using depth camera calibration, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5, 479-485, <https://doi.org/10.5194/isprsarchives-XL-5-479-2014>.
- Payne, A., Daniel, A., Mehta, A., Thompson, B., Bamji, C.S., Snow, D., Oshima, H., Prather, L., Fenton, M., Kordus, L., O'Connor, P., McCauley, R., Nayak, S., Acharya, S., Mehta, S., Elkhatib, T., Meyer, T., O'Dwyer, T., Perry, T., Chan, V.H., Wong, V., Mogallapu, V., Qian, W., Xu, Z., 2014. A 512×424 CMOS 3D Time-of-Flight Image Sensor with Multi-Frequency Photo-Demodulation up to 130MHz and 2GS/s ADC, in: *SolidState Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International.* pp. 134–135. doi:10.1109/ISSCC.2014.6757370
- Saunders, C. G., Bannon, M., Sabiston, R. M., Panych, L., Jenks, S. L., Wood, I. R., Raschke, S., 1989. The CANFIT System: Shape Management Technology for Prosthetic and Orthotic Applications, *JPO: Journal of Prosthetics and Orthotics: April 1989 - Volume 1 - Issue 3* – pp. 122-130.
- Sell, J., O'Connor, P., 2014. The Xbox One System on a Chip and Kinect Sensor. *IEEE Micro* 34, 44–53. doi:10.1109/MM.2014.9
- Taubin, G., 1995. A Signal Processing Approach to Fair Surface Design, in: *Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95.* ACM, New York, NY, USA, pp. 351–358.
- Tong, J., Zhou, J., Liu, L., Pan, Z., Yan, H., 2012. Scanning 3D Full Human Bodies Using Kinects, *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, NO. 4.
- Volonghi, P., Baronio, G., Signoroni, A., 2018. 3D scanning and geometry processing techniques for customised hand orthotics: an experimental assessment, *Virtual and Physical Prototyping*, 13:2, 105-116, DOI: 10.1080/17452759.2018.1426328
- Wang, X.L., Stolka, P.J., Bector, E., Hager, G., Choti, M., 2012. The Kinect as an interventional tracking system. In: *SPIE Medical Imaging*, pp. 83160U–6.
- Wenzel, K., Abdel-Wahab, M., Cefalu, A., and Fritsch, D., 2012. High-Resolution Surface Reconstruction from Imagery fo Close Range Cultural Heritage Applications, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B5, 133-138, <https://doi.org/10.5194/isprsarchives-XXXIX-B5-133-2012>.