MCP central europe

7th International Conference on Mass Customization and Personalization in Central Europe (MCP-CE 2016)

Mass Customization and Open Innovation
September 21-23, 2016, Novi Sad, Serbia

TECHNIQUES FOR 3D HUMAN BODY SCANNING

Miroljub Arbutina(1), **Srdan Mihic**(2), **and Dinu Dragan**(1)
University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Republic of Serbia
DOOB Innovation studio, Novi Sad, Republic of Serbia

Abstract: *3D human scanning is a process of capturing and subsequent digitization of the full body or the body parts into their three-dimensional graphical representation. As 3D human body scanning techniques are becoming ubiquitous, we give an overview of the field with techniques based on photogrammetry, depth sensors, and laser scanning as well as few more promising techniques with emphasis on their application. The purpose of this article is to review 3D body scanning systems currently available, to review different technologies of 3D body data digitization, and to determine the underlying principles that allow these systems to work. Specifications of 3D body scanning systems are compared in order to provide directions for further integration of digitized 3D human bodies. Lastly, paper presents a variety of commercial scanners available in the market.*

Key Words: ***3D human body scanning, 3D body scanning techniques, 3D body scanning systems, 3D body data digitization, 3D scanners commercial products***

1. INTRODUCTION

Lately there has been an increased demand for realistic 3D models of human beings. Even today there are many indications that it will create a revolution in all aspects of our everyday lives. For example, the way people interact (i.e. virtual environments), clothing and styling industry, medical field etc. In this light it is obvious why the demand for 3D scanners of all price ranges has risen. The main topic of this paper will be to give an overview of the techniques widely used.

There are many different devices that can be considered 3D scanners. Any device that measures the physical world using lasers, lights or x-rays and generates dense point clouds or polygon meshes can be considered a 3D scanner. Some common terminology in usage for them includes 3D digitizers, laser scanners, white light scanners, industrial CT, LIDAR, and others. All these devices capture the geometry of physical objects with hundreds of thousands or millions of measurements.

All the different approaches to 3D scanning are stated in Figure 1. Some technologies are ideal for short-range scanning (<1m, very high detail level), while others are better for mid (~m)- or long-range scanning (aerial photography). We will focus on those ranged techniques appropriate for scanning of humans i.e. those that commercially available and widespread. These can be divided into three different groups:

- Laser scanning,
- Projection of light patterns,
- Stereo-vision and image processing.

As far as reflective ranged measurements are concerned, first distinction is between active and passive methods. With active techniques the light sources are specially controlled, as part of the strategy to arrive at the 3D information. Active lighting incorporates some form of temporal or spatial modulation of the illumination. This can be a laser light technology or a visible light pattern projective technology. With passive techniques, on the other hand, light is not controlled or controlled only with respect to image quality. Typically, passive techniques work with whichever reasonable, ambient light available, as is often the case in stereo-vision and image processing methods.

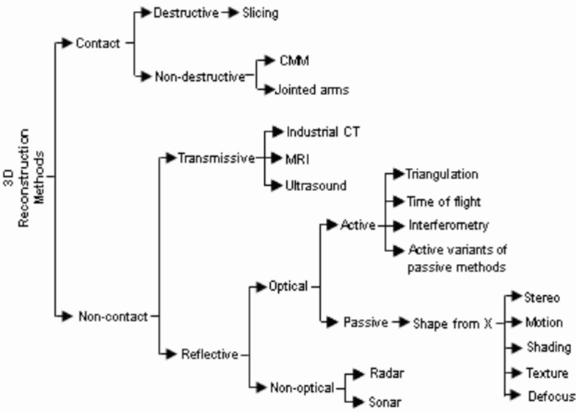

Fig. 1. Classification of 3D reconstruction methods

As this is an overview of relevant 3D acquisition methods, it is not intended to be in-depth nor exhaustive. The techniques that will be presented are are widely studied and used for decades.

The organization of the paper is as follows. Section 2 gives a brief description of the 3D scanning methods belonging to the three aforementioned groups. Section 3 discussed and compares their characteristics. Section 4 lists some of the commectialy available products.. Section 5 concludes the paper.

2. 3D SCANNING TECHNIQUES

Active range sensors acquire distance measurements from a well-known reference coordinate system to points of the object to be reconstructed [1']. They are very common when highly detailed models are required and have been used widely in the industry for decades. In the past ten years there have been an inrush of low-priced products based on scanning techniques to be described in the following section.

2.1. Laser Triangulation 3D Scanners

These scanners use either a laser line or single laser point to scan across an object. A sensor picks up the laser light that is reflected off the object, and using trigonometric triangulation, the system calculates the distance from the object to the scanner (Figure2).

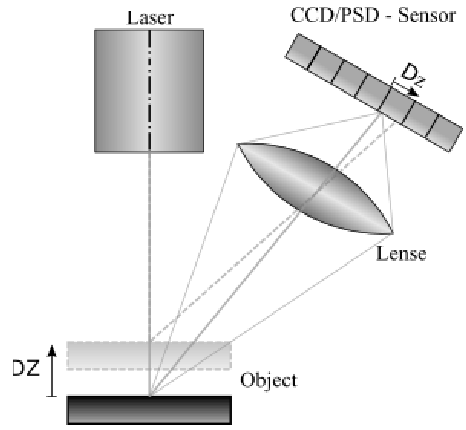

Fig. 2. Laser triangulation principle, image source: Wikipedia

The distance between the laser source and the sensor is known very precisely, as well as the angle between the laser and the sensor. As the laser light reflects off the scanned object, the system can discern what angle it is returning to the sensor at, and therefore the distance from the laser source to the object's surface. Process requires precise mechanical apparatus (e.g., by steering rotating mirrors that reflect the laser light into controlled directions), as well as very precise laser beam. One would also not want the system to take a long time for scanning. Hence, one ends up with the conflicting requirements of guiding the laser spot precisely and fast. These challenging requirements have an adverse effect on the price. Moreover, total time needed to take one image per projected laser spot can be up to seconds or even minutes of overall acquisition time. There are faster, special methods, using super-fast imagers, but again at an additional cost.

In order to remedy this, substantial research has gone into replacing the laser spot by more complicated patterns. For instance, the laser ray can without much difficulty be extended to a plane, e.g., by putting a cylindrical lens in front of the laser. Rather than forming a single laser spot on the surface, the intersection of the plane with the surface will form a curve (Figure 3). Variant of the scanner with several (usually up to four) scanning pillars have been common, without the need to rotate the subject.

2.2. Structured Light 3D Scanners

Structured light scanners also use trigonometric triangulation, but instead of looking at laser light, these systems project a series of linear patterns onto an object (Figure 4). Then, by examining the edges of each line in the pattern, they calculate the distance from the scanner to the object's surface. Essentially, instead of the camera seeing a laser line, it sees the edge of the projected pattern, and calculates the distance similarly.

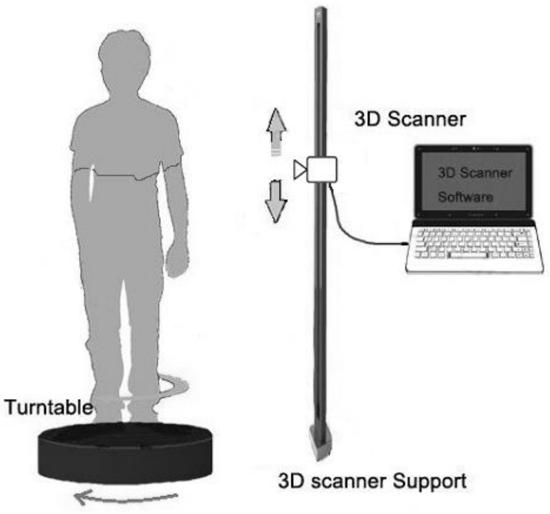

Fig. 3. Laser triangulation system

Two common methods of stripe pattern generation have been established: Laser interference and projection.

The former method allows for the exact and easy generation of very fine patterns with unlimited depth of field. Disadvantages are high cost of implementation and difficulties to provide the ideal beam geometry. The projection method uses incoherent light and basically works like a video projector. A typical measuring assembly consists of one stripe projector and at least one camera.

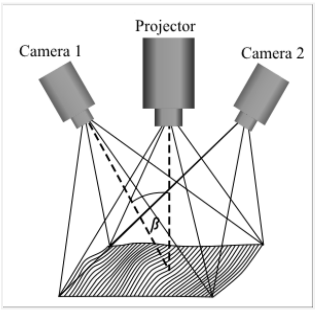

Fig. 4. Structured light system, image source: Wikipedia

There are several depth cues contained in the observed stripe patterns. The displacement of any single stripe can directly be converted into 3D coordinates. For this purpose, the individual stripe has to be identified, which can for example be accomplished by tracing or counting stripes (pattern recognition method). Another common method projects alternating stripe patterns, resulting in binary Gray code sequences identifying the number of each individual stripe hitting the object. An important depth cue also results from the varying stripe widths along the object surface. Stripe width is a function of the steepness of a surface part, i.e. the first derivative of the elevation.

In order to get texture extraction, successive projections of coded and phase-shifted patterns are required to extract a single depth frame, which leads to a lower frame rate. Low frame rate means the subject must remain relatively still during the projection sequence to avoid blurring, however technology is maturing fast so this might not be the case in the near future. The reflected pattern is sensitive to optical interference from the environment; therefore, structured-light tends to be better suited for indoor applications.

Invisible (or imperceptible) structured light uses structured light without interfering with other computer vision tasks for which the projected pattern will be confusing. Example methods include the use of infrared light or of extremely high framerates alternating between two exact opposite patterns. However, in those cases it is usually not possible to extract texture information.

2.3. Time-of-flight (ToF) scanners

Laser pulse-based scanners, also known as time-of-flight scanners, are based on a very simple concept: the speed of light is known very precisely, so if we know how long a laser takes to reach an object and reflect back to a sensor, we know how far away that object is (Figure 5). These systems use circuitry that is accurate to picoseconds to measure the time it takes for millions of laser pulses to return to the sensor, and calculate a distance. By rotating the laser and sensor (usually via a mirror), the scanner can scan up to a full 360 degrees around itself.

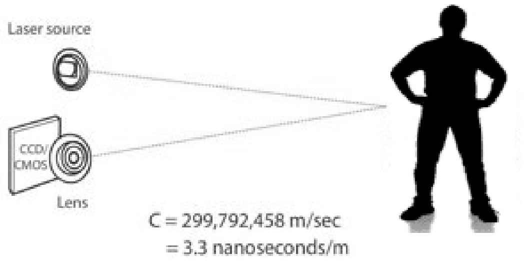

Fig. 5. Time-of-flight scanning system

Laser phase-shift systems are another type of time-of-flight 3D scanner technology, and conceptually work similarly to pulse-based systems. In addition to pulsing the laser, these systems also modulate the power of the laser beam, and the scanner compares the phase of the laser being sent out and then returned to the sensor. For reasons that are beyond this paper's scope, phase shift measurement is more precise.

Figure 6 shows a typical scanning setup where the infrared light wide-angle ToF system is used.

2.4. Multi-view photogrametry systems

Most ubiquitous method in passive image-based 3D re-construction systems use the stereoscopic principles as present in human vision. Here, 3D measurements are not performed, but 3D information is generated and extracted from a sequence of images, acquired from different viewpoints (Figure 7). Multi-view geometry is considered an evolution of stereo-based methods, where only two images were used: first, correspondences between points in both images were established (matching) and then their position in 3D space was determined by triangulation.

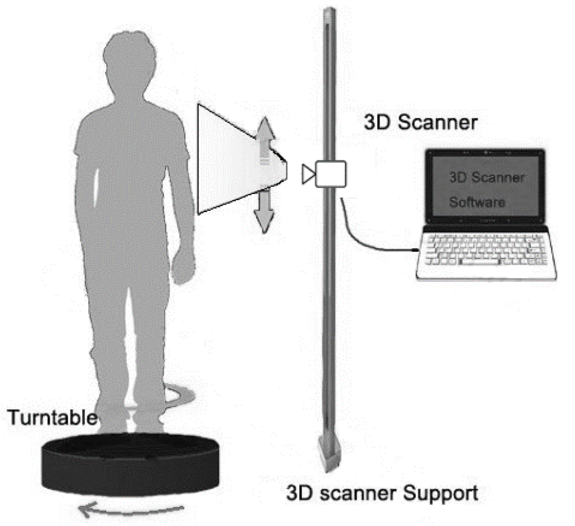

Fig. 6. Typical scanning setup for either wide-angle ToF or structured-light camera systems

Prior to acquiring images, several things can be done to improve reconstruction quality:

- Proper coverage of the subjects from all angles to prevent occlusions, including uniform lightning.
- Cameras should be calibrated prior to acquisition of images to obtain their intrinsic parameters used in triangulation.
- A 3D reference frame defining scale, position and orientation in the form of control points (coded targets), lengths of photo-identifiable objects, and/or angles between photo-identifiable objects.

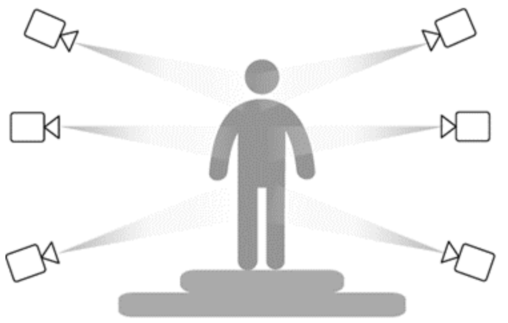

Fig. 7. Multi-view geometry setup

A major challenge in stereo vision photogrammetry is solving the correspondence problem [2’][3'] (Figure 8): given a point in one image, how to find the same point in the other cameras? Until the correspondence can be established, disparity, and therefore depth, cannot be accurately determined. Solving the correspondence problem involves complex, computationally intensive algorithms for feature extraction and matching [4’’]. This area is well studied and there are many mature algorithms [2’][5’] and commercially available software solutions such as Point Cloud Library [6’], Photoscan [7’], Australis [8’] etc. An exhaustive list can be seen here [9’].

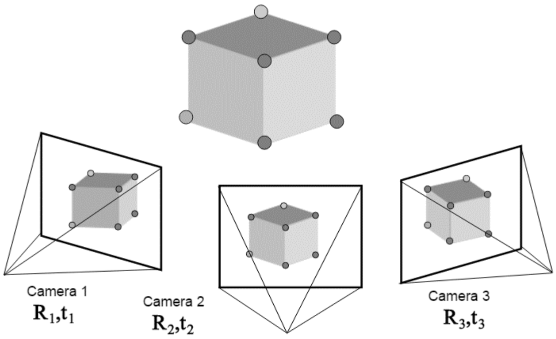

Fig. 8. Multi-view point correspondence problem, image source: Noah Snavely, Cornell University

After the matching phase, triangulation process generates 3D point cloud. Usually, a refining is needed so a bundle adjustment process is performed [10’]. It is an optimization procedure that refines an initial camera and structure parameter estimations for finding the values that most accurately predict the locations of the observed points in the set of available images. After that special interpolation algorithms create dense point cloud followed by polygonal mesh creation that can be exported for printing or other usage.

3. COMPARISON OF 3D SCANNING TECHNIQUES

Time-of-Flight (ToF) laser systems have an advantage in precision when measuring distant objects, however for small distances the error is usually of the order millimeters because of the difficulty in measuring round-trip time precisely. Triangulation systems have limited range but relatively high accuracy of the order of micrometers. but require a good baseline. A disadvantage of ToF systems is that surface texture is not captured and that errors will be substantially larger for dark surfaces, which reflect little of the incoming signal. Missing texture can be resolved by adding a camera, as close as possible to the ToF scanning head. But of course, even then the texture is not taken from exactly the same vantage point. The output is typically delivered as a massive, unordered point cloud, which may cause problems for further processing.

In contrast to stereo vision or triangulation systems, ToF systems are very compact: the illumination is placed just next to the lens, whereas the other systems need a certain minimum baseline. In contrast to laser scanning systems, no mechanical moving parts are needed.

It is a direct process to extract the distance information out of the output signals of the ToF sensor. As a result, this task uses only a small amount of processing power, again in contrast to stereo vision, where complex correlation algorithms are implemented. After the distance data has been extracted, object detection, for example, is also a straightforward process to carry out because the algorithms are not disturbed by patterns on the object.

Time-of-flight cameras are able to measure the distances within a complete scene with a single shot. As the cameras reach up to 160 frames per second, they are ideally suited to be used in real-time applications. High resolution ToF scans, which collect millions of points, can take more than few seconds which can result in distortion from motion [11'].

The advantage of structured-light 3D scanners is speed and precision. Some existing systems are capable of scanning moving objects in real-time. In triangulation based systems, the range and depth variation are limited compared to other methods, but they have a great precision. A major advantage of structured-light is that it can achieve relatively high spatial (X-Y) resolution by using off-the-shelf DLP projectors and HD color cameras.

In image based methods feature extraction and matching also require sufficient intensity and color variation in the image for robust correlation. This requirement renders stereo vision less effective if the subject lacks these variations i.e. when the local appearance is uniform within the neighborhood of each candidate point to be matched. Clothing can have a lack of significant local variation in their appearance or present a repeated pattern. This can, however, be remedied to some extent by placing artificial fiducial markers on the subject. ToF sensing does not have this limitation because it does not depend on color or texture to measure the distance. Secondly, occlusions in the scene make the correspondence between images ambiguous or even impossible, e.g. articulation of the body which leads to self-occlusions.

In stereo vision, the depth resolution error is a quadratic function of the distance while other two methods are much better in this respect. The main advantage of stereo vision is a very low cost of implementation as off-the-shelf cameras can be used.

The main benefit of 3D laser scanning is that it provides higher resolution and detail than photogrammetry. Therefore, facial details and even details on clothing are potentially captured more intricately with a 3D scanner, making the resulting 3D printed figurine more detailed and realistic as well. Scanning can be accomplished in a scanning booth, most often found in retail locations, or by using a 3D scanner in virtually any type of indoor or outdoor setting. A number of 3D laser scanners are available as portable handheld devices, offering quick and easy set up and flexibility in the location of the scan session. Structured light scanners can offer greater speed and accuracy than laser scanners when setup with large fields of view, approaching that of photogrammetry, but usually require several scanners to be synced together in order to capture the required angles without any “shadowing” due to “line of sight” and binocular vision restrictions. These multi-sensor systems, while very effective once set up, can be difficult and sensitive to calibrate.

The downside of using laser scanners for body form applications is that they take a second or more to capture data on any part of a subject during a scan. While that might not seem like much, the living body will move ever so imperceptibly during that second. Even if the movement is only a millimeter, the result will be a double layer of scan data in the same area of the geometry, which will cause problems when creating a mesh. Labour-intensive and often complex adjustment of the scan data is required after scanning in order to prepare the file for the 3D printer. Scanning sessions themselves can take several minutes and often must be repeated in order to properly capture the complete geometry or form [14'].

Many of the modern scanners use these techniques in combination. For example, there are cases of scanners which provide laser scanning precision with photogrammetric texture extraction component. Also as the state of these technologies is in constant movement it is impossible to make definite conclusion.

Wide range of available commercial 3D scanners can be seen in [12']. Table 1 shows the empirical comparison of scanning techniques, in terms of speed and precision, of the same price category. In Table 1, 5 stars represent the best empirical value.

Table 1. *Rough comparison of 3D scanning techniques*

Scanner type:	Speed (5 stars)	Precision (5 stars)
Multi-view	*****	**
Triangulation	***	*****
ToF	**	****

After creation of the point cloud, by some of the described techniques, discarding of erroneous points (i.e. outliers) and noise removal by various statistical methods can be performed. Some of these outliers can be filtered by performing a statistical analysis on each point's neighborhood, and trimming those which do not meet a certain criterion. The sparse outlier removal implementation in the publicly available Point Cloud Library is based on the computation of the distribution of point to neighbors distances in the input data-set. For each point, the mean distance from it to all its neighbors is computed. By assuming that the resulted distribution is Gaussian with a mean and a standard deviation, all points whose mean distances are outside an interval defined by the global distances mean and standard deviation can be considered as outliers and trimmed from the data-set.

Then a dense point cloud is generated using appropriate interpolation algorithm followed by filtering of the points from the result to adjust it and eliminate bad points. After that, creating polygon mesh model can be done for the model to be used in other 3D modeling software. An example of a generated mesh model without added texture is shown below (Fig. 9).

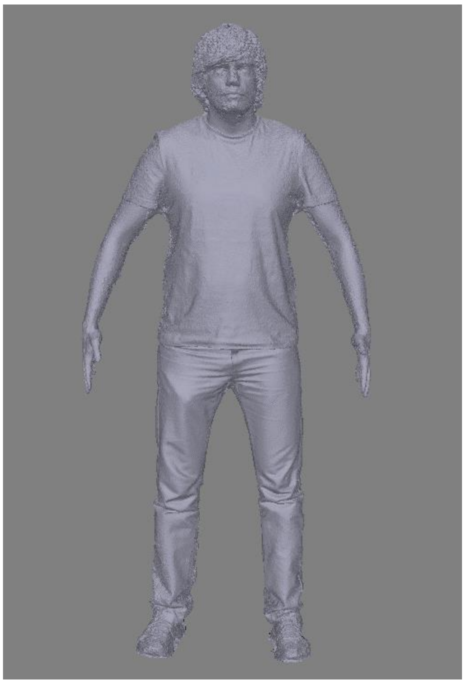

Fig. 9 – Typical 3D mesh model

4. COMMERCIAL PRODUCTS

There are numerous commercial scanners on the market today, with the number and variety constantly increasing. Hand-held 3D scanners are gaining popularity due to the ease of use and portability. Their prices range from 100 euros up to 20000 euros.

Some commercial solutions and the techniques they use:

- Microsoft Kinect 1 uses a pattern of projected infrared points to generate a dense 3D image.
- Microsoft Kinect for Xbox One uses a wide-angle time-of-flight camera.
- Intel RealSense camera (Fig. 10) projects a series of infrared patterns to obtain the 3D structure.

Low cost models like Intel RealSense 3D cameras are quickly finding their way in people’s everyday lives. They are becoming a part of notebooks, phones and handheld devices, replacing the old 2D cameras, or as an add-on accesory These devices are mainly used for the scanning of parts of human body (facial expressions, hand gestures, etc.) for entertainment purposes as their resolution and texture detail tends to be low.

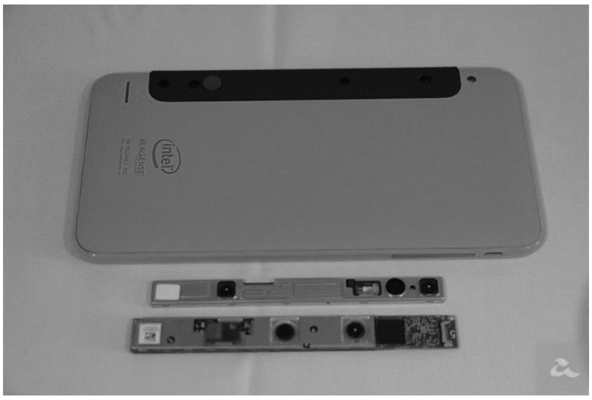

Fig. 10. Intel RealSense 3D cameras and their place on the smartphones –

Fig. 11 – iSense for iPads by 3DSYSTEMS

High-end, metrology-grade, portable scanners like those by Artec 3D (Eva and Space Spider) and Nikon (ModelMaker MMCx on Fig. 12) offer less than 0.5mm resolution, depending on the scanning distance. They are rarely used in full body scanning as the process can take minutes and the obtained level of details is excesive for non-special purposes.

Fig. 12. Nikon ModelMaker MMCx 3D laser scanner

Structured light scanners also have wide price spectrum. In the price range of 3000euros structured-light scanners products have superiority. An example is shown on Fig. 13. Submilimeter resolution and high mesh density for scans that last only several secons has made them videly used. For full body scanning, either a turntable has to be used or a set-up of several scanners.

Fig. 13. David SLS-3-STEREO structured light scanner

An emerging technology is the real time 3D scanning capability. It has applications in animation and medical imaging to capture real facial expressions and body shape changes in real time. Documenting and quantifying anatomical dense surface movement, pose, and expression. 3D camcoders have been on market for some time, and are becoming more affordable. In case of realtime 3D laser scanner they are still not present on the market

5. CONCLUSION

Active sensors provide directly the range data containing the 3D coordinates necessary for the mesh generation phase. Active range scanning technology has been applied to automatically acquire highly accurate geometric data of people. However, generally range based scanners can only capture a single static pose of a person. Passive sensors provide images that need a mathematical model to derive the 3D coordinates. After the measurements, the data must be structured and a consistent polygonal surface generated to build a realistic representation of the modeled objects or scenes. A photo-realistic visualization can afterward be generated by texturing the 3D models with image infor-mation.

Nowadays, when choosing appropriate scanner suiting ones' needs, a wide range of products currently available is rapidly increasing. The old classification of laser scanners being the most expensive is no longer true. Same goes for structured-light scanners. Both methods have products of full body millimeter precision scanners for 2000$ or less [12']. Also small and portable scanners are finding their way to the market, being pushed by large companies with aggressive pricing of several hundred dollars. If one needs a high quality texture model currently the choice boils down to multi-view photogrammetry scanners. However, structured-light scanners might not be too far behind.

The race is on among manufacturers of digital image and geometry capture technology to develop a solution that is affordable, off the shelf, quick and easy to set up, high resolution and fast, in order to accommodate movement. To this end, photogrammetry solution providers are trying to incorporate multi-camera systems and software into one simple, packaged solution. At the same time, makers of 3D laser scanners are trying to develop faster and more automatic geometry capture solutions to accommodate for movement and shorter scanning sessions. There are also major steps being taken by mainstream electronics companies and suppliers to incorporate 3D content capture into mobile phones or tablets, and this is where we are already seeing a big change in the ability of average consumers to generate usable 3D content.

As the need for 3D models of humans will certainly rise in the near future, the market of 3D scanners is becoming a lucrative business and the competition is developing new and innovative solutions at a staggering rate. It is certainly a field where a potential buyer must be up-to-date on a monthly basis.

6. REFERENCES

[1] T.C. Azevedo, "3D Object Reconstruction using Computer Vision: Reconstruction and Characterization Applications for External Human Anatomical Structures", PhD Thesis, Faculdade de Engenharia da Universidade do Porto, pp. 7-32.

[2] S. Gold, A. Rangarajan, C.P. Lu, S. Pappu, E. Mjolsness, "New algorithms for 2D and 3D point matching: Pose estimation and correspondence", *Pattern recognition*, vol. 31, no. 8, pp. 1019-1031 1998.

[3] T. Moons, L.V. Gool, M. Vergauwen, "3D Reconstruction from Multiple Images Part 1: Principles", *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 4, pp. 287–398, 2008.

[4] D. Dragan, S. Mihic, Z. Anisic, I. Lukovic, "Role of background subtraction in creating human body point clouds from photos", The 6th International Conference and Exhibition on 3D Body Scanning Technologies, 27-28 October, 2015, Lugano, Switzerland, pp. 210-217.

[5] M.M. El-gayar and H. Soliman, "A comparative study of image low level feature extraction algorithms", *Egyptian Informatics Journal*, vol. 14, no. 2, pp.175-181, 2013.

[6] R.B. Rusu and S. Cousins, Willow Garage, "3D is here: Point Cloud Library (PCL)", The 2011 IEEE International Conference on Robotics and Automation (ICRA), pp. 1-4.

[7] Agisoft, "Photoscan". Available at: http://www.agisoft.com/ [Accessed 20 June 2016].

[8] Photometrix, "Austrlis. Available at: http://www.photometrix.com.au/australis/ [Accessed 20 June 2016].

[9] "Comparison of photogrammetry software", Wikipedia. Available at https://en.wikipedia.org/wiki/Comparison_of_photogrammetry_software [Accessed 20 June 2016].

[10] B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon, "Bundle Adjustment — A Modern Synthesis", in *International workshop on vision algorithms*, 1999. Springer Berlin Heidelberg. pp. 298-372.

[11] L. Li, "Time-of-Flight Camera – An Introduction", Technical White Paper SLOA190B – January 2014 Revised May 2014, pp. 1-6. Available at: http://css1a0.engr.ccny.cuny.edu/~wolberg/capstone/kinect/ToFBasicsTI14.pdf [Accessed 20 June 2016].

[12] Kira, "The Best 3D Scanners of 2015", www.3ders.org, 2015. Available at: http://www.3ders.org/articles/20151209-best-3d-scanners-2015.html [Accessed 20 June 2016].

[13] N. D'Apuzo, "State of the art of the methods for static 3D scanning of partial or full human body ", Proceedings of Conference on 3D Modelling, June 13-14, 2006, Paris, France, pp. 1-8.

[14] Julie Reece, "The Pros and Cons of Photogrammetry vs. 3D Scanning for 3D Printed Figurines", Available at: http://mcortechnologies.com/the-pros-and-cons-of-photogrammetry-vs-3d-scanning-for-3d-printed-figurines-blog/ [Accessed 20 June 2016].

CORRESPONDENCE

Miroljub Arbutina, Student
University of Novi Sad
Faculty of Technical Sciences,
Trg Dositeja Obradovića 6
21000 Novi Sad, Serbia
m.arbutina@doobinnovation.com

Srdan Mihic, Software Architect
DOOB Innovation studio
Vojvodanskih brigada 28
21000 Novi Sad, Serbia
s.mihic@doobinnovation.com

Dinu Dragan, PhD, Assis. Professor
University of Novi Sad
Faculty of Technical Sciences,
Trg Dositeja Obradovića 6
21000 Novi Sad, Serbia
dinud@uns.ac.rs