Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 11 (2017) 1600 – 1608

Procedia MANUFACTURING

27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy

Fast and low cost acquisition and reconstruction system for human hand-wrist-arm anatomy

Monica Carfagni[a], Rocco Furferi[a*] , Lapo Governi[a], Michaela Servi[a], Francesca Uccheddu[a], Yary Volpe[a], Kathleen Mcgreevy[b]

[a] *Department of Industrial Engineering, Via di Santa Marta 3, Firenze 50139, Italy*
[b] *Research, Innovation and International Relations Office, Meyer Children's Hospital, Viale Gaetano Pieraccini, 24, Firenze 50139, Italy*

Abstract

3D body scanners are nowadays used in a range of applications spanning from health, fashion and fitness to reverse engineering applications for robotics and computer vision. Nowadays very good performances are achievable when using commercial 3D body scanners; however, focusing on relative complex shape of some body details, the results still lack precision and acceptable accuracy. Such critical issue remains unsolved also when dealing with the instantaneous acquisition of the hand-wrist-arm (HWA) anatomy. In this paper, we present a new approach that leverages the emerging 3D depth cameras technologies to design a compact low cost 3D dedicated HWA scanner system capable of delivering almost instantaneous full 3D measurement.

Peer-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing

Keywords: 3D body scanner; hand-wrist-arm anatomy; depth sensors; calibration; 3D surface

1. Introduction

Dedicated 3D body scanners are paramount to deliver the exact measures of a human body to be used in a range of applications dealing with health, fashion and fitness as well as the myriad of reverse engineering applications for robotics or industrial engineering in general [1-5].

* Corresponding author.
E-mail address: rocco.furferi@unifi.it

Peer-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing
doi:10.1016/j.promfg.2017.07.306

Human oriented 3D scanners pose new challenges in the panorama of existing optical measurement systems; in fact, the agile and deformable human nature, imposes the acquisition to be specifically fast to avoid movement artefacts. Nowadays very good results are possible with the existing body scanners (both professionals and consumer devices) however, focusing on relative complex shape of some body details, the results still lack precision and acceptable accuracy. This is particularly true when the human body (of some of its parts) is acquired by using low-cost devices such as, for instance, the Microsoft® Kinect™ [6] and the Intel® Realsense [7] i.e. camera systems able to provide both color and dense depth image. These sensors have been recently applied in a huge number of industrial and biomedical applications. Compared with traditional 3D scanners, these sensors can capture depth and RGB data at video rate. Despite both the RGB quality and the depth resolution are limited, the major benefit comes from the overall acquisition speed and the near IR pattern that allows the acquisition in poor lighting conditions as well as on dark surfaces. Consequently, the use of such devices rapidly grew especially regarding the human motion analysis [8] but also for creating low-cost body scanners. An early publication dealing with the use of Kinect™ as a 3D body scanner was performed in 2012 demonstrating the effectiveness of approaches based on the use of low-cost device for this kind of application [9-11]. Unfortunately, several drawbacks remain in developing full body scanners based on RGB-D sensors.

First, the depth data captured is of extreme low quality (low X/Y resolution and depth accuracy) especially when dealing with non-static bodies (like living human bodies are) that should be measured instantaneously [12]. Furthermore, RGB-D sensors require relatively long computational time to reconstruct a complete model from the scan data (and the obtained model is often unreliable). Eventually, to increase the resolution of the acquired body/body part requires the contemporary use of multiple devices with several problems related to registration, superimposition of projected IR patterns, data flow management etc. Such critical drawback remains unsolved also when dealing with the instantaneous acquisition of the hand-wrist-arm (HWA) anatomy. Required in a variety of applications related to robotics, medical devices (cast and orthosis), tailor made jewelry as well as sport and fashion apparel, the peculiar geometry of human HWA remains one of the most challenging part to measure.

In this paper, we present a new approach that leverages the emerging 3D depth cameras technologies to design a compact low cost 3D dedicated HWA scanner system capable of delivering, almost instantaneously, a full 3D measurement. Especially suited for medical purposes (i.e. the acquisition of paediatric patients) the dedicated scanner consists of a set of four RGB-D scanners arranged on a circular ring to acquire in a single "shot", the entire geometry of interest under the condition that the patient remains, during the scanning process, as static as possible. The proposed system is provided also with a procedure for the semi-automatic reconstruction of a parametric CAD from the acquired point cloud. This procedure allows to obtain a mesh to be eventually used for several applications (e.g. design of orthosis and ergonomic analysis).

2. Design of a 3D HWA scanner

As mentioned above, obtaining a complete 3D model of the hand-wrist-arm anatomy is a complex task: besides technical issues arising from the need of using more than a single device, other practical problems (mainly due to the agile and deformable human nature and to the variety of dimensions) may arise. This is particularly true, especially when it comes to growing children (see Fig. 1).

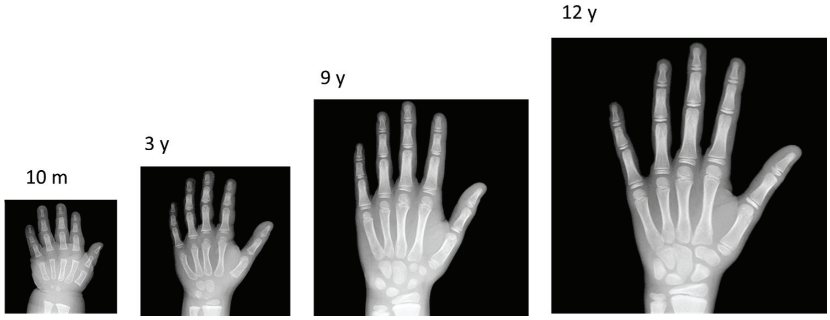

Fig. 1. As an example of the wide variety of the hand-wrist-arm segment across different patient (age/sex/individual), the developmental stage differences between (from left to right) 10 months, 3 years, 9 years and 12 years of age is reported.

For this kind of application, in fact, it is necessary to focus our attention on: (1) which kind of 3D sensor is the more suitable for acquiring the HWA anatomy; (2) how many cameras are needed to acquire the entire geometry and how to arrange the acquisition devices to scan bigger and smaller arms; (3) how to acquire the 3D data as fast as possible. Only once these issues are solved, and the resolution of obtained point clouds is sufficient to fully reconstruct the acquired target, we can process the acquired data to obtain the final surface.

1.1. Choice of the 3D acquisition sensor

As mentioned in the introductory section, a number of 3D scanners are available in the market, roughly divided into two main categories: professional 3D scanners (e.g. Romer Absolute Arm [13], Konica Minolta Range7[14], Aicon 3D System StereoScan [15]) and low-cost devices such as, for instance RGB-D cameras (e.g. Kinect [6], Intel RealSense [7], Occipital Structure [16]). Even if this last typology of devices was specifically conceived to address topics related to face analytics and tracking, scene segmentation, hand and fingers tracking, gaming and Augmented Reality [17-19], they prove to have the potential to be used as a low-cost 3D scanner [20, 21]. For this reason, together with the need of scalability of the proposed system, we chose this kind of devices for the design of a new HWA 3D scanner. In particular, among commercial 3D scanners, it emerged in terms of performances and versatility the new Intel RealSense SR300 frontal camera [7]. In Fig. 2 a single shot acquisition of a mannequin HWA is depicted. The "real" geometry is estimated by measuring it with a Romer Absolute Arm 7520 SI/SE (Hexagon Metrology SpA) that has an overall accuracy of ±0.063 mm (i.e. a sufficiently reliable "reference measure").

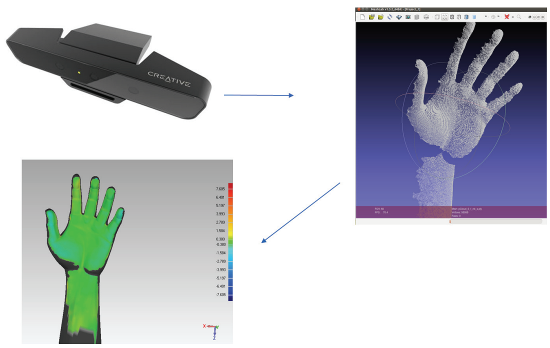

Fig. 2. Single scan acquisition of a mannequin hand. The SR300 depth camera shows a very high accuracy in representing the HWA geometry.

SR300 is a short-range camera that uses coded light 3D imaging system. The camera (see Fig. 3) is composed of an infrared laser projector, a Fast VGA infrared camera and a 2-Mpixel RGB color camera with an integrated image signal processor.

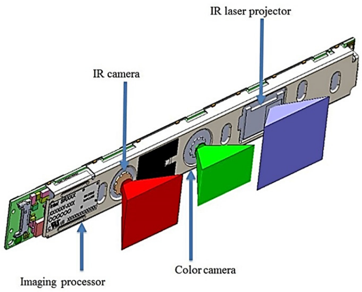

Fig. 3. 3D drawing of the SR300 internal components.

Fast VGA depth mode reduces exposure time and allows dynamic motion up to 2 m/s. This new feature leads to a less-noisy result when the scene is not static. The effective range of depth is optimized from 0.2 m to 1.5 m for use indoors. The dimensions of the device are approximately 110 mm x 12.6 mm x 3.8–4.1 mm; its weight is 9.4 grams, which makes it very compact and well-suited for a wide range of 3D imaging applications.

The overall technical specifications of Intel RealSense SR300 are summarized in Table 1. The SR300 implements the so called "Temporal plus Range Multiplexing", thus meaning that in addition to projecting a sequence of frames, each single channel characterized by multiple grey levels is considered. Temporal multiplexing samples the information in time, this limits the system capability of acquiring depth information of scenes with temporal frequency variations. If the scene is not static during the acquisition of the N projected frames, shadow artefacts can occur in the estimated depth map.

Table 1. Technical specifications of Intel RealSense SR300.

Technology	Coded Light, Fast VGA 60 fps
Color Camera	Up to 1080p 30 fps, 720p 60 fps
Indoor Range	20-150 cm
Depth FOV (D x W x H)	88 mm x 71.5 mm x 55 mm
Depth fps	30/60
IR fps	30-300

1.2. Number of cameras and arrangement

With the aim of building an accurate HWA scanner a crucial step is to state the correct number and the positioning of cameras needed to acquire the entire 3D geometry while preserving the hardware complexity of the whole scanner. As a matter of fact, optical depth sensors acquisitions are often prone to a borders degradation, depending on the underlying camera technology; in fact, when the optical ray is tangent with respect to the object to be measured, a little variation of a pixel on the sensor leads to a considerable difference in the detected pixel coordinates [22]. For this reason, the post processing step includes an edges removal step thus shrinking the actual working area. A virtual simulation of the entire setup has been performed with a 3D computer graphics software (Blender [23]) taking into account also the wasted border points.

In Fig. 4 the hardware setup is depicted and in Fig. 5 the virtualization of the scanning overlapping regions in case of three and four cameras.

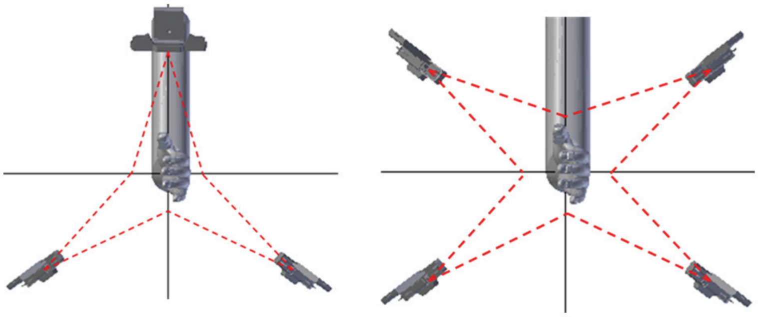

Fig. 4. Hardware set up. The three cameras configuration on the left, and the four cameras configuration on the right.

In case of three cameras, the red, green and blue colors are associated to the area framed by each of them. The black part (between the index and the thumb) are uncovered by the scanner; the mixed colors are the overlapping parts. In case of four cameras the area covered by the couple of opposite cameras are represented in green and blue. The case of four cameras presents more overlapped area and it fully cover the hand/wrist/arm part thus not showing black parts in the area of interest.

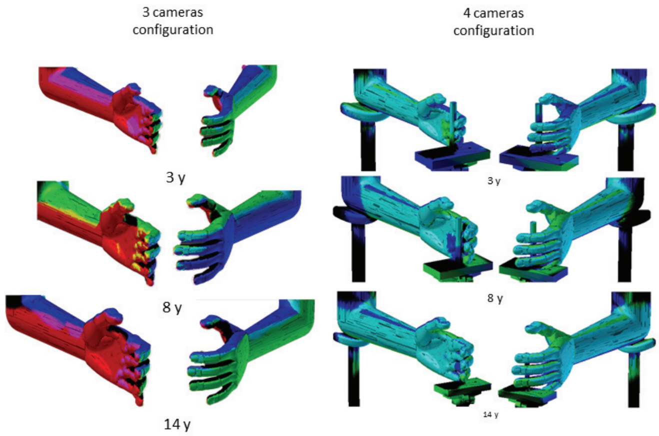

Fig. 5. The virtualization of the scanning overlapping regions in case of three and four cameras.

Due to the large variety of the pediatric HWA sizes, with less cameras the over imposition areas may be not enough extended to cover the whole geometry, and the final reproduction may present some holes. The circular rig configuration allows an easy scalability of the system: if a cameras repositioning is requested, or one or more devices are needed to expand the 3D scanner, three additional support are placed over the ring, maintaining a constant distance between adjacent supports. This circular disposition in fact allows to choose the observing angle keeping the camera-subject distance constant.

Having analyzed different configurations, it emerged that four cameras, arranged on a circular ring of diameter 720 mm, are the best trade-off between hardware complexity and scanner accuracy. With such an arrangement, the HWA anatomy is acquired by each camera from a distance of approximately 245-255 mm (depending on the size of pediatric patient). A double ring coupling is finally devised to cover the entire variability of arm size in the range 3-14 years old patients. The final lay-out of the acquisition system is in Fig. 6.

Fig. 6. The final layout of the HWA 3D scanner. On the left, the layout with one ring hosting the four cameras. In the middle the frame composed by two rings for a modular implementation of the scanner. On the right a rendering of the final configuration.

1.3. 3D scene acquisition

Data acquisition was carried out using a dedicated software, developed by means of the open source library Librealsense [24], which is a cross-platform library that permits to access to the Intel RealSense device.
The HWA scanning process requires a very fast acquisition time to minimize artefacts due to subject movement; in fact, it is very hard to hold the hand still, even for few seconds.

Although Intel® SR300 allows a fast single acquisition, using more devices at the same time is not feasible due to the interference problem. In fact, since the underlying technology of the SR300 cameras is based on the projection for each camera, of an IR pattern in the correspondent field of view, if one or more camera share the same field of view the patterns interfere and the resulting acquisition is fully degraded. For this reason, we exploited the functionalities of these devices to manually fine tune some camera parameters (e.g. accuracy, motion range, filter option). The proposed approach is based on adjusting the projector laser pattern intensity of the four cameras via software and turning them on in sequence; such approach allows to reach a full acquisition time of approximately 1.2 seconds. For each obtained cloud, a region of interest is automatically selected based on the previous analysis on acquisition angles and acquired scene dimensions such to cut the unessential acquired data. A contour filter is also applied to compensate wrong values around edges.

The 3D reconstruction process requires an initial cameras calibration in which the relative positions of all the four cameras equipping the scanner are determined. In fact, even if every part of the object is acquired by at least one camera, each camera captures the scene with respect to its own reference system. A proper calibration procedure needs to be made to move all the point clouds in the same reference system with a roto-translation matrix. To estimate the four matrices, we designed a target object having different planes and edges in several directions such to define robust efficient features [25] (see Fig. 7).

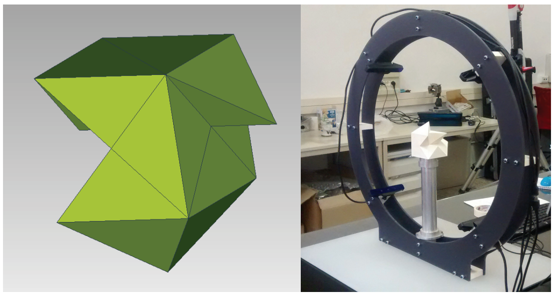

Fig. 7. Target object having different planes and edges in several directions such to define robust efficient features for clouds registration

A simple and intuitive software interface has been implemented both for launching the acquisitions and for calibrating the cameras, developed using PCL and Qt libraries [26]. For each camera, the user selects at least three corresponding points on the acquired cloud and on the target cloud; then, the calibration is carried out with the traditional two-steps procedure comprising a first rigid alignment based on the selected points, and a subsequent iterative fine alignment based on iterative closing points (ICP) algorithm [27].

This procedure is performed every time the system is installed or moved. The 4 roto-translation matrices obtained by correctly aligning the 4 acquired point clouds on the actual 3D model of the target are stored and finally applied to each of the newly acquired data from the 4 cameras. Consequently, any further acquisition allows a direct automatic clouds registration resulting in a 360° view of the arm (see Fig. 8).

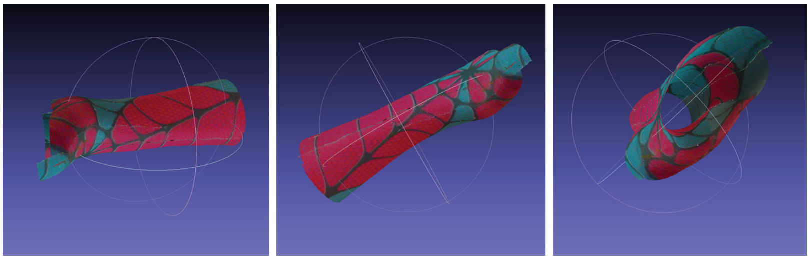

Fig. 8. Example of a 360° acquisition of the wrist handle arm anatomy.

3. Post processing and surface extraction

The HWA district surface reconstruction comprises several operations that can be automatized. The steps of the procedure for creating the reconstructed surface from the four meshes have been implemented with the VTK library [29] and depicted in Fig 9.

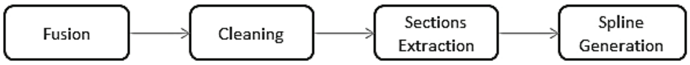

Fig. 9. Scheme of necessary steps to obtain the surface starting from the four triangulated point clouds.

- *Fusion*: the four point clouds are first triangulated and then translated to cover the whole hand-wrist-arm district; the meshes are then combined in a single object.
- *Cleaning*: each mesh is cleaned by removing non-connected-components and boundary vertices, and by filling up the generated holes. A trade-off between mesh cleaning and points removal has been investigated through the analysis of different arm sizes in the age range of 6-14 years. Three times iteration of the cleaning step turned out to be the optimal compromise to delete almost all the artifacts.
- *Mesh sectioning*: a parallel sheaf of planes having as axis the inertial axis of the global mesh is use to obtain parallel sections of the arm and hand. Each section is a set of points representing the intersection with the 3D model (Fig. 10);
- *Spline generation*: the points of each model section are ordered according to the polar coordinates criteria to preserve the circular shape of the arm and exported in IGES format which will be given as an input to the CAD/CAM software for the surface reconstruction. The final model can be used, for instance, to design orthosis or other medical devices [29].

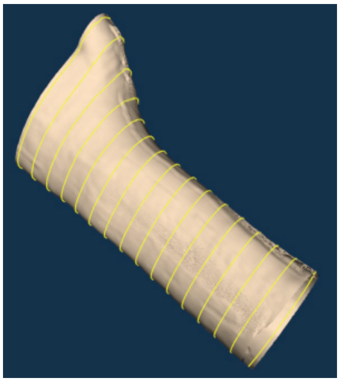

Fig. 10. Example of mesh sectioning: a parallel sheaf of planes having as axis the inertial axis of the global mesh is use to obtain parallel sections of the arm and hand.

4. Conclusions

In this paper a fast and low cost 3D scanner for human hand-wrist-arm anatomy has been proposed. Starting from a first hardware analysis of accuracy and speed requirements imposed by the agile and deformable human nature, the Intel Realsense SR300 was selected. To achieve a reliable structure able to capture the hole 360° geometry and at the same time minimize the overall acquisition time, an architecture consisting on arranging four cameras in a circular frame was designed. With this configuration pattern interferences are avoided and the overall acquisition time is about 1.2 seconds. Finally, for the HWA surface reconstruction a procedure has been designed and fully automatized with a dedicated software. Future works will be addressed towards the implementation of the system in real-life scenarios.

Acknowledgements

Authors wish to thank the Ospedale Pediatrico Meyer of Florence (Italy) for supporting this work by providing hand-wrist-arm patient statistics and medical specifications.

References

[1] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, “Real-time human pose recognition in parts from a single depth image”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.,2011, pp. 1297–1304
[2] S. Barone, A. Paoli, A.V. Razionale, “A coded structured light system based on primary color stripe projection and monochrome imaging”, Sensors. 2013; 13: 13802-19.
[3] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, “KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera”, Proc. ACM Symp. User Interface Softw. Technol.,2011, pp. 559–568.
[4] L. Di Angelo, P. Di Stefano, A. Spezzaneve, "A method for 3D detection of symmetry line in asymmetric postures.", Computer methods in biomechanics and biomedical engineering 16.11 (2013): 1213-1220.
[5] L. Di Angelo, P. Di Stefano, "A computational method for bilateral symmetry recognition in asymmetrically scanned human faces.", Computer-Aided Design and Applications 11.3 (2014): 275-283.
[6] Microsoft Kinect™: https://msdn.microsoft.com/en-us/library/jj131033.aspx. Accessed 2017
[7] Intel RealSense, www.intel.com/realsense (Accessed in 2017)
[8] C. Chen, K. Liu, and N. Kehtarnavaz. "Real-time human action recognition based on depth motion maps." Journal of real-time image processing 12.1 (2016): 155-163.
[9] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, “Scanning 3D full human bodies using kinects”, IEEE Trans. Visual. Comput. Graphics, 18 (2012), pp. 643–650.
[10] S. Song, S. Yu, and W. Xu, “Study on 3D body scanning, reconstruction and measurement techniques based on Kinect”, J. Tianjin Polytech. Univ., 31 (2012) 34–37.

[11] R. Wang, J. Choi, and G. Medioni, “Accurate full body scanning from a single fixed 3D camera”, in: Proc. – Jt. 3DIM/3DPVT Conf.: 3D Imaging, Model., Process., Vis. Transm., 3DIMPVT, 2012, pp. 432–439.
[12] R. Furferi, L. Governi, F. Uccheddu, and Y. Volpe, “A RGB-D based instant body-scanning solution for compact box installation” in Advances on Mechanics, Design Engineering and Manufacturing. Springer International Publishing, 2017. 387-396.
[13] http://www.hexagonmetrology.it/ROMER-Absolute-Arm-con-scanner-integrato_1005.htm Accessed 2017
[14] https://www.konicaminolta.eu/it/strumenti-di-misura/prodotti/misurazioni-3d/range-7/introduzione.html Accessed 2017
[15] http://aicon3d.com/products/aicon-scanner/stereoscan-neo/at-a-glance.html Accessed 2017
[16] https://structure.io/ Accessed 2017
[17] Q. Cai, D. Gallup, C. Zhang, and Z. Zhang, “3D Deformable Face Tracking with Commodity Depth Camera”, Computer Vision – ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part III, pp 229-242
[18] N. Silberman, and R. Fergus, "Indoor scene segmentation using a structured light sensor," 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, 2011, pp. 601-608
[19] I. Oikonomidis, N. Kyriazis, and A. A. Argyros. "Efficient model-based 3D tracking of hand articulations using Kinect." BmVC. Vol. 1. No. 2. 2011.
[20] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer, “Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling”, Remote Sens.2015, 7(10), 13070-13097; doi:10.3390/rs71013070
[21] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, "Scannig 3D full human bodies using Kinects”, IEEE Trans Vis Comput Graph. 2012 Apr;18(4):643-50. doi: 10.1109/TVCG.2012.56.
[22] W. Boehler, M.B. Vicent, and A. Marbs, “Investigating laser scanner accuracy”, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, (2003), 34(Part 5), 696-701.
[23] Blender: https://www.blender.org/ Accessed 2017.
[24] Librealsense: https://github.com/IntelRealSense/librealsense Accessed 2017.
[25] Rusinkiewicz, Szymon, and Marc Levoy. "Efficient variants of the ICP algorithm." 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on. IEEE, 2001.
[26] R.B. Rusu and S. Cousins, “3d is here: Point Cloud Library (PCL)”. In: Robotics and Automation (ICRA), 2011 IEEE International Conference on (pp. 1-4). IEEE.
[27] Z. Zhang, “Iterative point matching for registration of free-form curves and surfaces”, International journal of computer vision13.2 ,1994,: pp 119-152.
[28] VTK: http://www.vtk.org Accessed 2017.
[29] A Critical Analysis of a Hand Orthosis Reverse Engineering and 3D Printing Process (2016) Applied Bionics and Biomechanics, 2016, art. no. 8347478.