









Automatic Point Clouds Alignment for the Reconstruction of Upper Limb Anatomy

Alessandro Paoli¹ , Paolo Neri² , Armando V. Razionale³ , Luca Di Angelo⁴ , Paolo Di Stefano⁵  and Emanuele Guardiani⁶ ,

¹Università di Pisa, alessandro.paoli@unipi.it

²Università di Pisa, paolo.neri@unipi.it

³Università di Pisa, armando.viviano.razionale@unipi.it

⁴Università degli Studi dell'Aquila, luca.diangelo@univaq.it

⁵Università degli Studi dell'Aquila, paolo.distefano@univaq.it

⁶Università degli Studi dell'Aquila, emanuele.guardiani@univaq.it

Corresponding author: Alessandro Paoli, alessandro.paoli@unipi.it

Abstract. In the field of optical 3D scanning for healthcare applications, low-cost depth cameras can be efficiently used to capture geometry at video frame rates. However, the complete reconstruction of anatomical geometries remains challenging since different scans, collected from multiple viewpoints, must be aligned into a common reference frame. This paper proposes a fully automatic procedure to align scans of the upper limb patient's anatomy. A 3D optical scanner, obtained by assembling three depth cameras, is used to collect upper limb acquisitions. A relevant dataset of key points on the hand and the forearm geometry is then determined and used to automatically obtain a rough 3D alignment of the different scans. Hand key points are identified through a neural network, which works on RGB images captured by the depth cameras; forearm key points are recognized by directly processing the point clouds through a specifically designed algorithm that evaluates the skeleton line of the forearm. The approach was tested on forearm acquisitions, and the results were compared to alternative alignment methodologies.

Keywords: depth cameras, 3D optical scanning, upper limb anatomy, automatic point clouds alignment.

DOI: <https://doi.org/10.14733/cadaps.2023.S6.158-169>

1 INTRODUCTION

In the last decade, the introduction on the market of consumer-grade depth cameras, such as the Occipital Structure Sensor, the Microsoft® Kinect™ sensors or the Intel® RealSense™ cameras, has supported the development of portable and low-cost dedicated optical 3D body scanners for healthcare applications [6, 13, 15]. The adoption of low-cost sensors can also strongly impact the

opportunity for new telerehabilitation systems that connect therapists and patients with cognitive or motor impairments [1, 12]. In 3D optical scanning, body scanners must face critical issues related to the deformability of human tissues and the non-stationary nature of the scanning target. These issues impose that the scanning time must be reduced as much as possible to minimize scanning artifacts. In this regard, depth cameras can capture geometry (i.e., point clouds) and color information (i.e., RGB images) at video rates (up to 90 fps), thus guaranteeing fast acquisition times. However, the overall reconstruction of a human body part (i.e., lower or upper extremities, chest, head) requires collecting multiple scans from different viewpoints. Two approaches can be followed: 1) simultaneously using multiple scanning devices or 2) collecting multiple acquisitions from the same device, which is appropriately moved around the target surface [10]. The first approach involves a calibration stage before the scanning process, whereas the second requires a strategy to align the sequential scans to the same reference system. The two approaches also differ in terms of the requirements they must meet. The former requires a specific architecture to constrain multiple optical devices together. This architecture may be cumbersome and difficult to move around the patient. For this reason, the overall scanning process should be carried out without planning any repositioning between the scanner architecture and the patient. The second, instead, requires the scanner's portability, which must be easily moved around the patient.

Some examples of the first approach are given in [3, 5, 13]. In [3, 5], multiple Intel RealSense SR300 depth cameras (from four to eight) are arranged on circular rings fixed on a desk. The patient arm is then introduced internally to the rings and housed on specific supports. The relative placement between the depth cameras is evaluated by a calibration stage using a target object with different planes and edges. The cameras acquire the target object, and common features are used to perform a semi-automatic registration of the range maps. In [13], three Kinects are rigidly assembled on a fixed frame to capture different human body parts. Moreover, the patient stands on a 360° rotating platform, allowing the whole-body acquisition in about 30 seconds.

Some examples of the second approach are given in [7, 14]. In [14], the Occipital Structure Sensor has been used to acquire the patient arm to evaluate the volume variations for lymphedema-affected patients. The sensor is mounted on an Apple iPad, which is moved around the target anatomy. The software Skanect is then used to automatically align the point clouds and RGB data to obtain the final polygonal mesh. The same sensor is used in [7] to acquire the hand in different functional postures to create 3D hand parametric models. These approaches require many different placements, which are obtained by freely moving the single scanning sensor around the target anatomy. Commercial software or a great manual effort is required to register all the acquired data in a common reference system. An attempt to overcome this limitation and automatize the scanning process by a single sensor has been proposed in [8] using the Sense 3D Scanner (3DSystems). The sensor is placed on an arm that rotates around the target, thus making possible the acquisition of the patient upper limb from different viewpoints. However, registration issues remain since the circular path of the sensor is obtained by manually acting on a crank handle.

In the present work, a portable 3D optical scanner has been developed to reconstruct the patient's upper limb anatomy. The scanner is based on a multi-sensor layout composed of three Intel® RealSense™ depth cameras assembled on a lightweight circular rig. The three sensors are mutually calibrated to acquire and register acquisitions from three different viewpoints for each scanner pose. A procedure to automatically recognize common key points between scans deriving from different poses of the 3D scanner has been developed. RGB images obtained by the depth cameras are processed through a neural network to detect a dataset of relevant key points of the hand, which are then reprojected onto the point cloud to identify corresponding 3D points. A smaller dataset of relevant key points of the forearm is then added to the hand dataset. This latter dataset is obtained by directly processing the point cloud through a specifically developed method that aims to find matches at the skeleton line of the forearm. These two datasets, detected for different scanner poses, are then used to automatically obtain a rough 3D alignment of the point clouds in a common reference frame. The alignment is finally refined through a fine registration by an iterative closest point (ICP) algorithm. The approach was tested on forearm acquisitions, and the results were compared to alternative alignment methodologies. The research aims to solve the two main issues

arising when a 360° acquisition of an upper limb must be obtained: speeding up the acquisition process and automatically registering point clouds deriving from different scanner poses.

2 HARDWARE SETUP AND CALIBRATION

The scanner has been assembled by integrating three Intel® RealSense™ D415 depth cameras on a circular frame having a diameter of 440 mm. Fig. 1 shows the 3D scanner architecture and a detail of a single sensor housed within a 3D printed cover, which protects from accidental impacts. The D415 camera is a compact sensor (99 mm × 20 mm × 23 mm) equipped with an RGB camera (resolution up to 1920 × 1080 pixels, frame rate 30 fps), an infrared projector, and two IR cameras (resolution up to 1280 × 720 pixels, frame rate up to 90 fps). Due to the high frame rate, the sensor is suitable for capturing human targets, which may undergo small involuntary movements. Moreover, the infrared projector uses a single static pattern to enhance the scene texture, thus avoiding interference problems between different sensors, typical of structured light approaches that use the sequential projection of fringe patterns (i.e., the SR300 Intel® RealSense™ depth camera). The metrological characterization of the D415 depth camera presented in [4] demonstrated that this sensor can be effectively used as a low-cost device for biomedical purposes, even if the device was primarily designed to address applications such as tracking, gaming, or gesture recognition. Each depth camera acquires a point cloud for each scanner pose. The simultaneous use of three different depth cameras increases the field of view since three-point clouds, from different viewpoints, are collected for each scanner pose.

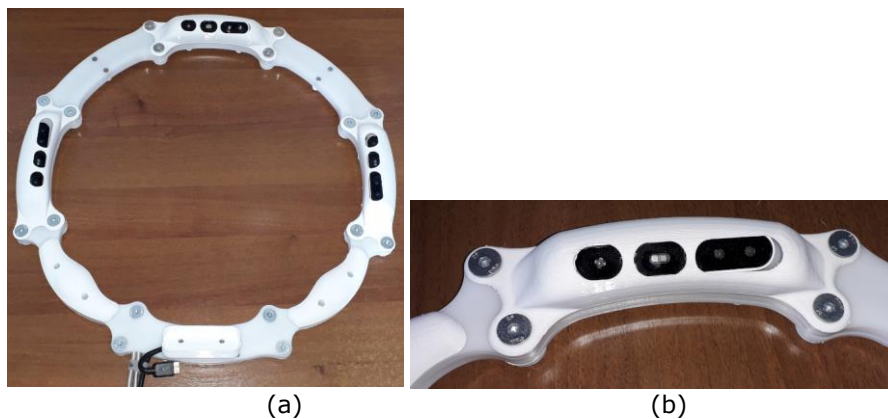


Figure 1: 3D scanner architecture: (a) overview of the three sensors mounted on a circular ring, (b) detail view of a single sensor placed inside the cover.

However, these acquisitions must be aligned into a common reference frame by exploiting a calibration procedure, which relies on determining the relative placement between the sensors. The procedure is possible since the three sensors are connected to the same rigid structure. A 3D shape, obtained by additive manufacturing using PLA (polylactic acid) material and a FFF (fused filament fabrication) machine, is used as a calibration specimen. The specimen is composed of six planar surfaces, differently oriented, and three colored markers attached to the external surfaces (Fig. 2). For each sensor's acquisition, three key points (i.e., the marker's center) can be detected on the captured RGB image and reprojected on the corresponding point cloud using the sensor's factory calibration data. The key points triplets are then used to obtain the rough alignment between the three-point clouds captured for each scanner pose, which is finally refined with an ICP algorithm.



Figure 2: Specimen used to calibrate the mutual placement between the three depth cameras on the circular ring.

3 AUTOMATIC REGISTRATION BETWEEN DIFFERENT SCANNER POSES

Each scanner pose can provide only a limited area of the target surface. For this reason, multiple scans deriving from different scanner viewpoints must be collected. The scans deriving from different poses of the 3D scanner must then be aligned into a common reference frame to obtain a consistent model of the patient's anatomy. In this regard, no calibration procedures can be adopted since the scanner is freely moved around the target by the user. The conventional approach is based on a post-processing phase after the scanning stage, which involves selecting at least three common points between common areas of adjacent point clouds (3-2-1 procedure). This approach, however, is drastically affected by the manual selection of corresponding points on different point clouds. To accomplish this task, each point cloud must be visualized, rotated, compared, and visually inspected to find any relevant feature to identify corresponding points in different clouds. This results in a cumbersome, slow, subjective, and unrepeatable procedure, especially if no markers are attached to the patient's arm. In this scenario, the user mainly looks for recognizable features, such as fingertips, wrist and elbow bones, and knuckles.

Thus, the present work aims to overcome these limitations by defining a fast, reliable, objective, and repeatable automatic procedure. The approach developed to this extent consists of three main steps:

- detection of visible features typical of the hand by a neural network (NN), adopted to replace the human intervention;
- automatic recognition of forearm features to be matched for different acquisitions;
- global registration by an ICP algorithm.

3.1 Neural Network for Hand Key Points Detection

The development of NNs has gained significant interest in recent years thanks to the diffusion of open-source software and NNs training infrastructures. The use of a NN is particularly suitable in this application since each acquired frame provides three-point clouds from different perspectives and the three corresponding RGB images. This information enhances the chances of having a clear view of key points in at least one of the images. The study of hand motion is also of great interest in different fields such as medicine, ergonomics, virtual reality, etc. Hence, several NNs are available which recognize and categorize specific hand features. In particular, the "MediaPipe Hands" NN was selected and used in this work [9]. This NN is implemented in python language, and the availability of open-source code allowed for easy implementation of the routine in the overall scanning workflow, which was developed in MATLAB language. If an RGB image of a hand is selected, the NN provides a list of the detected key points, along with a recognition score. In ideal conditions, the algorithm can detect 21 features of the hand, i.e., four key points on each finger plus one key point on the wrist. The output is always sorted in the same key point order, as reported in the scheme of Fig. 3.

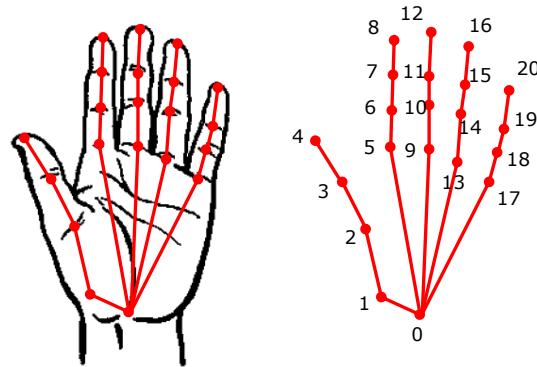


Figure 3: Hand key points detection through the NN.

Since it may happen that one or more key points are not visible in a specific image, the network is implemented to store a not-a-number value for those key points which could not be detected. Thus, the output always consists of 21 points in the same order, thus avoiding misleading feature numbering. This is crucial when the key points detected from different viewpoints must be compared. Due to this numbering strategy, the correspondence of points with the same indexing is always guaranteed.

3.1.1 Neural Network performance in key points detection

An evaluation of the performance of the NN in detecting hand key points on the images stored by the developed scanning device was performed. Several acquisitions of different arms (both left and right) with different poses (e.g., stretched, folded) and from different perspectives were taken. For each scanner pose, the images acquired by three depth cameras were processed by the described NN, and the total number of successfully detected key points was recorded. The results obtained on an almost ideal acquisition are reported, as example, in Fig. 4(a). It is worth nothing that the palm is completely visible by all the three sensors, and the fingers are properly stretched. This condition represents the ideal hand pose for the NN to detect the key points. Indeed, all 21 hand landmarks are correctly detected in the three images and reported as white marks. On the other hand, Fig. 4(b) shows the results of a frequent pose of the hand with respect to the sensors: as can be seen, the palm is visible only for one of the devices (sensor n. 2), while the others capture a lateral view of the hand. In particular, sensor n. 1 has a partial view of the palm; thus, the detected key points are still approximately aligned with the hand. Sensor n. 3, instead, has an entirely lateral view of the hand, which does not allow for a proper landmarks detection resulting in key points completely misaligned with respect to the hand.

These qualitative results were further investigated during the testing campaign by recording for each image whether the detection was successful or not. The campaign involved the arm of 4 different researchers, acquired from ten different viewpoints each, thus having 40 image triplets and a total of 120 images. The NN performance was evaluated from two different perspectives. Firstly, each of the 120 images was treated as a single picture and processed through the NN. The detected key-points were then evaluated by an expert user, who marked whether the detection was successful or not. Thus, the success rate of the NN in hand landmarks detection could be assessed by computing the ratio between the number of images marked as successful and the total number of images, obtaining a value of 67.5 %. Nevertheless, the developed scanning device exploits three synchronous sensors; thus, in this application, detection can be considered successful if at least one of the three sensors detects the key points. The analysis was then repeated on the 40 image triplets, marking as successful any image triplet where at least one image provided a successful detection. A success rate of 92.5 % was obtained. These results demonstrated that the layout of the sensor enhanced the NN performances. It is worth noting that if the key points detection fails in all three images of a

single acquisition, the registration of that point cloud can still be performed with the conventional manual procedure so that no 3D data loss is expected.

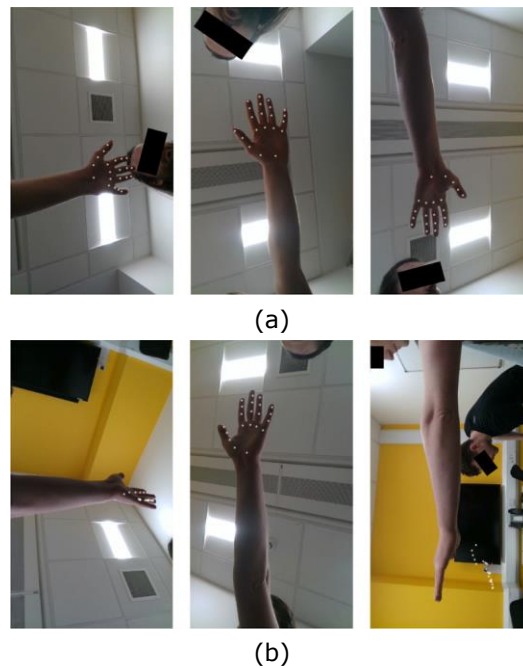


Figure 4: Hand key points detection: (a) almost-ideal hand pose and (b) lateral view.

3.1.2 Hand key points selection

Each acquisition frame includes data from three independent Intel® RealSense™ D415 sensors. The three-point clouds are mutually registered thanks to the calibration of the scanner. Thus, the RGB pictures corresponding to these clouds can be analyzed to detect the hand key point. In principle, this procedure would provide three sets of hand key points, which could be equivalently used to register the clouds to subsequent frames. Indeed, this data redundancy allows for gaining a higher success rate.

Nevertheless, some of the sensors can fail in detecting the hand key points, depending on the hand orientation to the cameras and on the finger pose. In practice, the identification of one sensor can detect a smaller number of key points or can completely fail. Thus, an automatic algorithm is needed to evaluate each frame and to choose the sensor with the best detection among the three available. To this extent, it is worth noting that at least three points are needed to obtain the cloud registration; hence it is not mandatory that all the 21 key points are detected. On the other hand, a larger number of key points provides an over-constrained registration, thus enhancing performance. Following this principle, a key point detection quality factor was defined by multiplying the identification score (provided by NN algorithm for each RGB image) by the number of detected key points in that image. This criterion allows selecting the key points set with many landmarks and a high-quality identification, thus representing the best choice among the three sensors. Once the best 2D key point set is selected, it is possible to project their coordinates into the 3D point cloud by exploiting the D415 factory calibration, thus providing a set of 3D key points that can be stored along with the point cloud and exploited for registration purposes. An overview of the automatic selection algorithm is reported in Fig. 5. It is worth noting that if none of the sensors provided a reliable key points identification for a specific pose, that pose would require manual registration.

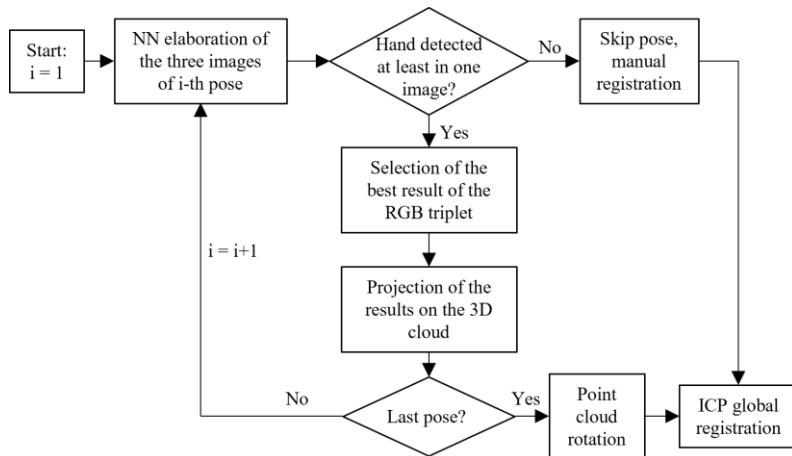


Figure 5. Schematic workflow of the algorithm.

This procedure provides a set of registered 3D landmarks for each cloud. Nevertheless, the set of key points associated with each frame can contain a different number of landmarks, depending on the NN results. This case can be an issue since the rough registration requires two sets of one-by-one corresponding 3D points to achieve cloud alignment. In practice, the most common situation is determined by the NN failing to detect a few random key points of the hand, thus impairing the possibility of directly using the whole set for the rough registration. To overcome this issue, the intersection between the key point lists belonging to the two point-clouds to be registered is considered, and a list of corresponding indexes is stored along with the cloud. When the registration step is performed, only the key points with the selected indexing will be used to compute the alignment roto-translation.

3.2 Feature-Based Algorithm to Detect Forearm Key Points

The proposed approach aims to detect specific points on the patient's forearm that can be matched to drive the global registration of all the acquisition frames. Concerning the registration methods, in literature, typically, one or more feature descriptors can be used to detect potential correspondences among pairs of points of different acquisitions of the human body [2, 11]. These shape descriptors include Simple or Hybrid Shape Measures and the skeleton line. However, they can be used when the geometry is approximated with high-density point clouds as they are based on discrete differential geometry. In the proposed experimental hardware setup, the depth cameras resolution allows for acquisitions, which are affected by the typical staircase effect (Fig. 6); this level of information does not permit the use of the above-mentioned approaches.

Therefore, a specifically designed procedure, whose main steps are described in Fig. 7, is proposed to accomplish this task. The procedure assumes that the external surface of the forearm has a skeleton that can be approximated by a straight line. For each 3D scanner pose, the forearm surface is acquired only for a small angular spanning. For this reason, each section is assumed circular to calculate the skeleton line position more robustly.

Starting from all the acquisition frames (Fig. 7(a)) defined in the global reference system xyz , the method, for each of them, iteratively:

- finds the direction of the skeleton line ξ_{fin} of the forearm;
- detects the set of points φ_{i-th} , which, being aligned along the ξ_{fin} direction, should have a higher potential of a match with the other acquisitions.

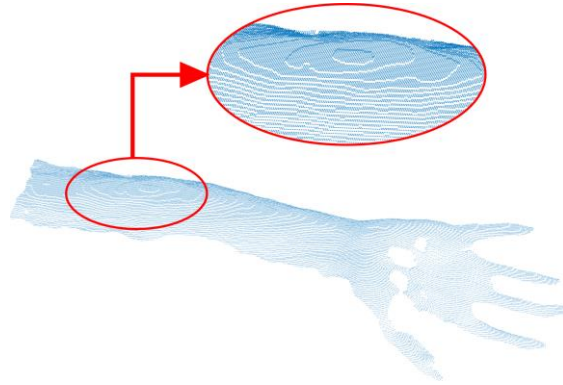


Figure 6. An example of point cloud acquisition with a magnification to highlight the staircase effect.

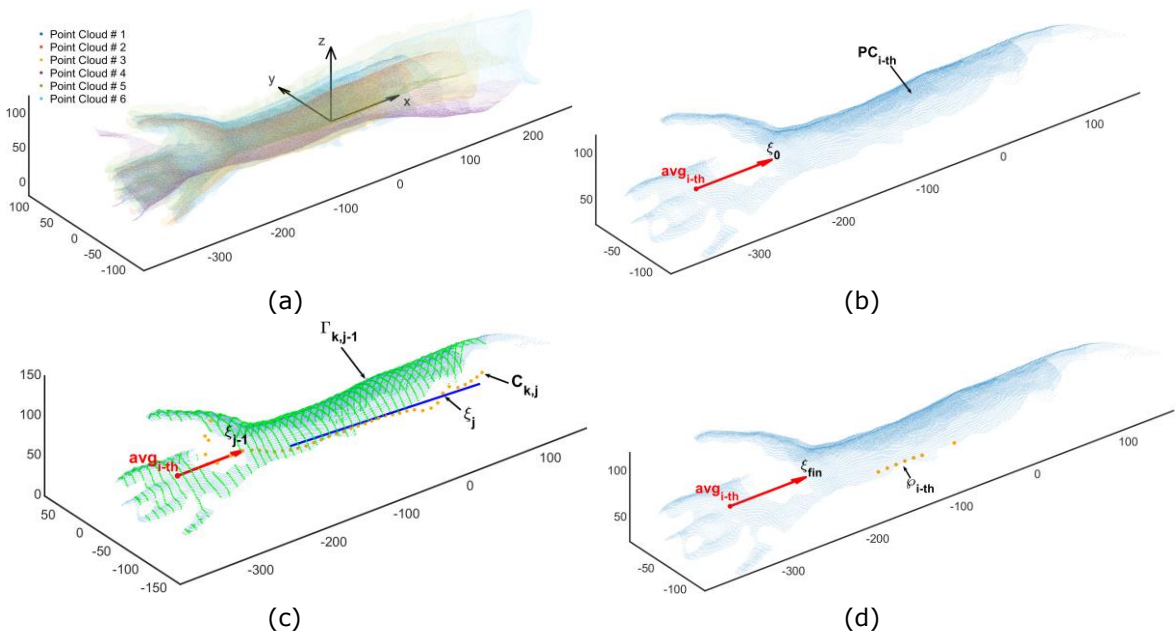


Figure 7. Main steps of the algorithm for the forearm key point detection.

3.2.1 Working principle

For each scanner pose, the processed point cloud is the one obtained by merging the three registered point clouds captured by the three depth cameras. Let PC_{i-th} be the i -th point cloud corresponding to the i -th scanner pose (Fig. 7(b)); the direction of the skeleton line of the forearm of the first attempt (ξ_0) is estimated by evaluating the Principal Component of inertia of the point cloud. In particular, the longitudinal axis ξ_0 is associated with the lower inertia moment of the PC_{i-th} . In Fig. 7(b), ξ_0 is anchored to the barycenter (avg_{i-th}) of the 21 points detected on the hand. Starting from ξ_0 , the final direction is obtained by the following iterative method; at the j -th iteration (Fig. 7(c)):

- Clustering of the points recognized near a set of planes perpendicular to ξ_{j-1} ($\Gamma_{k,j-1}$);
- Approximation of each cluster with a circle whose center is $C_{k,j}$;
- Approximation of the centers $C_{k,j}$ with a line by the RANSAC algorithm that defines the ξ_j .

The iterative algorithm stops when:

$$|\xi_j \cdot \xi_{j-1}| \leq tv$$

where tv is a threshold value specified by the operator.

Evaluated ξ_{fin} , the set of points ρ_{i-th} to be matched with the corresponding of the other point clouds are chosen as the sequence of at least five points most closely aligned (Fig. 7(d)). In order to make the registration process more robust using corresponding points, each element of the ρ_{i-th} set takes the value of its distance along ξ_{fin} from **avg** $_{i-th}$.

3.3 Global Registration

The initial rough alignment obtained by exploiting the key points was finally refined by an ICP global registration to enhance the alignment of all the acquired point clouds. At this stage of the research, this step was carried out by using the Geomagic Studio 2013 (3D System, South Carolina, USA) ICP algorithm. One of the point clouds is fixed to provide a reference, while all the others are kept floating and mutually adjusted. The algorithm iteratively moves all the floating point clouds to obtain the overall best fit of the clouds group. Calculation control options were set as follows: sample size of 2000 points, tolerance value of 0, maximum iterations 100.

4 RESULTS

The proposed automatic method was implemented in an original software and coded in MATLAB. Experimental data were acquired to evaluate the implemented procedure's accuracy, potentialities, and limits. Each acquisition is composed of eighteen distinct point clouds corresponding to six different poses of the 3D scanner around the patient's arm. At this preliminary stage of the research, the method was tested in the registration of the forearm of patients without severe pathologies that would further deteriorate the information from the acquisition process. In particular, in the performed experimentation, patients unable to stretch and control the movements of the upper limb and fingers were discarded. To quantify the performances, the results of the proposed method are compared with those of the gold standard model consisting of the patient's forearm reconstruction obtained using the conventional 3-2-1 manual procedure followed by a global registration.

To highlight the robustness of the method for the recognition of forearm skeleton line points, figure 8 shows the results obtained for the two different cases using four different registration strategies:

- with hand key points only;
- with hand key points followed by a global registration;
- with hand and forearm key points;
- with hand and forearm key points followed by a global registration.

The results highlight that using only the hand key points for the initial alignment of the point clouds, the cloud region close to the hand is pretty well aligned, but a relevant misalignment of the forearms can be noted. Furthermore, this rough alignment is not accurate enough to guarantee the appropriate convergence of the global registration by ICP. On the other hand, adding forearm key points to the hand key points better constrain the first alignment step, which can bring the different point clouds considerably closer. This configuration allows the different point clouds to be correctly registered by running a global ICP registration.

The performances of the proposed method were quantitatively evaluated in Fig. 9, which reports the final model obtained for the two forearms with the proposed method and gold standard one. In the same figure, maps of the distances between the two corresponding models are reported. These distances were computed by using Geomagic Qualify 2013 (3D System, South Carolina, USA).

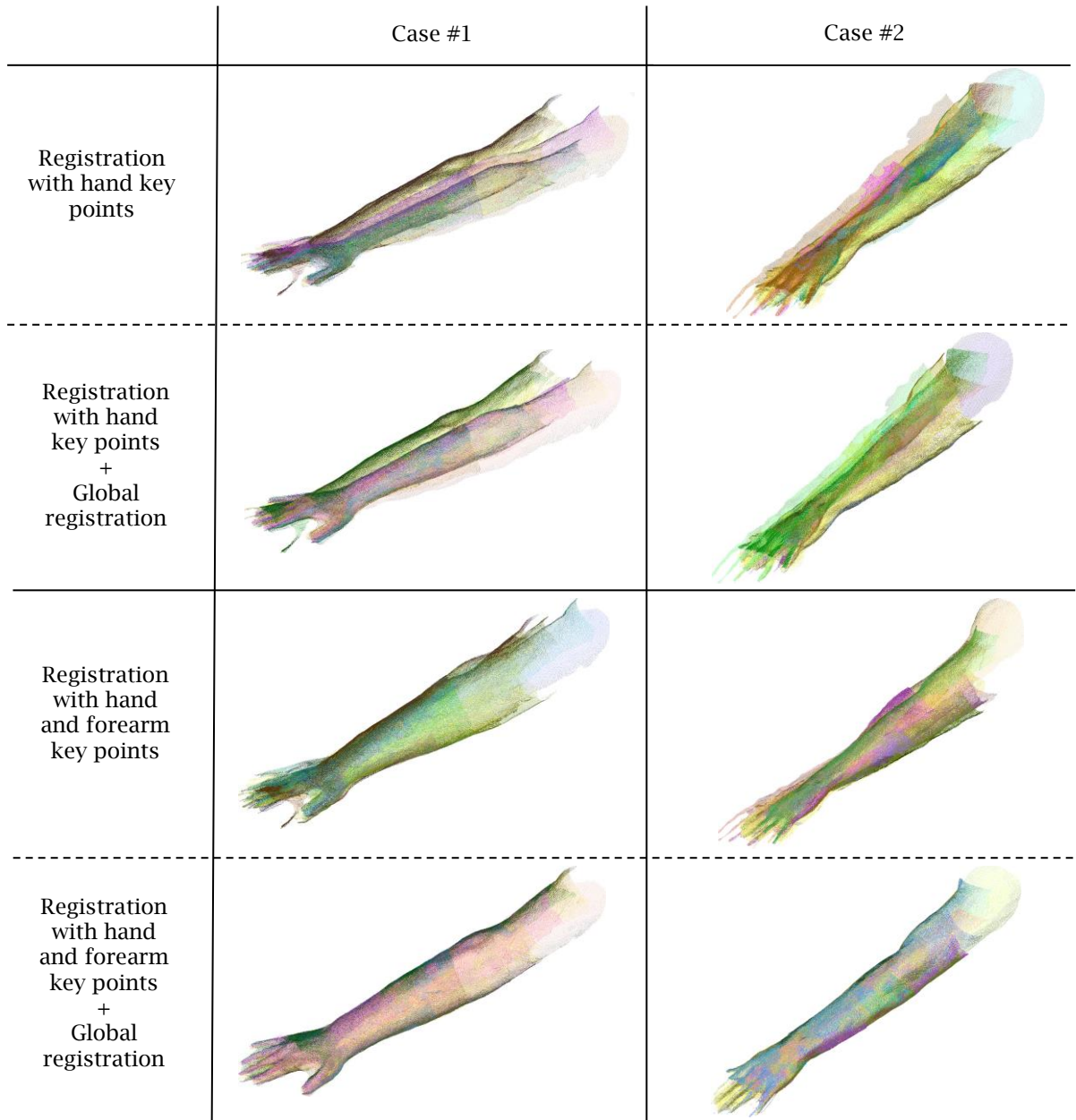


Figure 8: Alignment results obtained by following different strategies for two different cases.

More noisy data can be observed at the fingers due to the optical undercuts which occur during the scanning process. Discrepancies between models obtained by the fully automatic procedure and gold standard models are within the range ± 1 mm (case #1: mean = 0.22 mm, standard deviation = 0.49 mm; case #2: mean = -0.13 mm, standard deviation = 0.44 mm). As expected, greater deviation values are found in the correspondence of the hand region. It is worth noting that soft tissues were measured, and small involuntary movements may occur during scanning, thus 1 mm discrepancy can be considered acceptable for this anatomical application. Additionally, the developed

procedure allowed continuous and closed surfaces, an index of a high-quality 360° acquisition and registration.

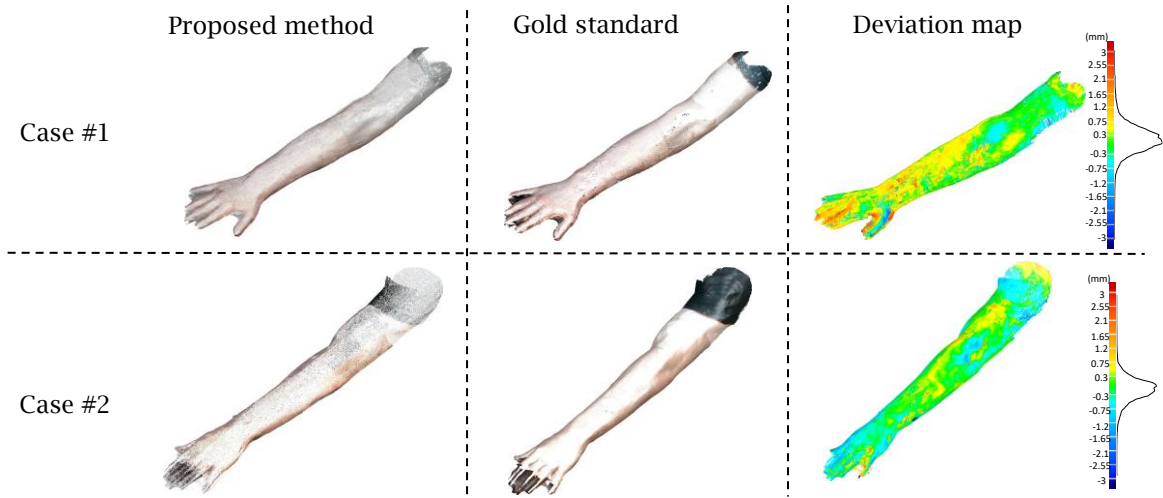


Figure 9: Textured final patient's arm models obtained for the two cases here analyzed by registering all the point clouds with the proposed method, the gold standard one and deviation maps obtained with respect to the gold standard model.

5 CONCLUSIONS

The automatic registration of point clouds captured from different views by an optical scanner is challenging if no calibrated hardware (such as a rotary table or a robotic arm) is used. In this paper, this problem was faced through an automatic algorithm that exploits key point detection to obtain a first rough registration, which is then refined through an ICP algorithm. The methodology firstly processes the RGB images with a neural network to detect up to 21 key points on the hand. Other relevant key points are also detected in the forearm geometry by geometrically deriving an approximation of its axis. This set of key points can be detected on each point cloud, thus providing a 3D point list of the corresponding locations between subsequent clouds acquired by different poses. Therefore, they can be used to obtain roto translation matrixes that provide a rough alignment of the different viewpoints. Finally, the results can be refined with a global registration ICP-based algorithm. This procedure was tested on actual arm scanning performed on human arms. The results obtained with the automatic procedure were compared with those obtained with the conventional manual registration approach. Small discrepancies in the range of ± 1 mm were found and considered acceptable in this specific application field. Future efforts will be focused on experimenting the developed approach on a greater number of cases by also considering patients with a reduced upper limb mobility or with uncontrolled arm movements as those affected by dystonia disorders. Also, the possibility of including the ICP global registration into MATLAB routines will be considered to speed up and fully automatize the overall 3D reconstruction process.

Alessandro Paoli, <https://orcid.org/0000-0002-1918-3033>

Paolo Neri, <https://orcid.org/0000-0003-0730-0893>

Armando V. Razionale, <https://orcid.org/0000-0001-7110-3857>

Luca Di Angelo, <https://orcid.org/0000-0002-5341-0500>

Paolo Di Stefano, <https://orcid.org/0000-0001-5003-2084>

Emanuele Guardiani, <https://orcid.org/0000-0002-1623-1474>

ACKNOWLEDGEMENTS

The authors wish to acknowledge that this research has been supported by the European Commission as part of the H2020 program, under the grant agreement n. 856998 (PRIME-VR2 project).

REFERENCES

- [1] Aruanno, B.; Caruso, G.; Rossini, M.; Molteni, F.; Espinoza, M.C.E.; Covarrubias, M.: Virtual and Augmented Reality Platform for Cognitive Tele-Rehabilitation Based System, *Computers Helping People with Special Needs*, Springer International Publishing, Cham, 2020, pp. 130-137, https://doi.org/10.1007/978-3-030-58796-3_17.
- [2] Berretti, S.; Daoudi, M.; Turaga, P.; Basu, A.: Representation, Analysis, and Recognition of 3D Humans: A Survey, *Acm T Multim Comput*, 14(1), (2018), <https://doi.org/10.1145/3182179>.
- [3] Buonamici, F.; Furferi, R.; Governi, L.; Lazzeri, S.; McGreevy, K.S.; Servi, M.; Talanti, E.; Uccheddu, F.; Volpe, Y.: A practical methodology for computer-aided design of custom 3D printable casts for wrist fractures, *Visual Comput*, 36(2), (2020), 375-390, <https://doi.org/10.1007/s00371-018-01624-z>.
- [4] Carfagni, M.; Furferi, R.; Governi, L.; Santarelli, C.; Servi, M.; Uccheddu, F.; Volpe, Y.: Metrological and Critical Characterization of the Intel D415 Stereo Depth Camera, *Sensors-Basel*, 19(3), (2019), <https://doi.org/10.3390/s19030489>.
- [5] Carfagni, M.; Furferi, R.; Governi, L.; Servi, M.; Uccheddu, F.; Volpe, Y.; McGreevy, K.: Fast and low-cost acquisition and reconstruction system for human hand-wrist-arm anatomy, *Procedia Manuf*, 11, (2017), 1600-1608, <https://doi.org/10.1016/j.promfg.2017.07.306>.
- [6] Chiu, C.Y.; Thelwell, M.; Senior, T.; Choppin, S.; Hart, J.; Wheat, J.: Comparison of depth cameras for three-dimensional reconstruction in medicine, *P I Mech Eng H*, 233(9), (2019), 938-947, <https://doi.org/10.1177/0954411919859922>.
- [7] Chu, C.H.; Wang, I.J.; Sun, J.R.; Liu, C.H.: Customized designs of short thumb orthoses using 3D hand parametric models, *Assist Technol*, (2020), <https://doi.org/10.1080/10400435.2019.1709917>.
- [8] Li, J.; Tanaka, H.: Feasibility study applying a parametric model as the design generator for 3D-printed orthosis for fracture immobilization, *3D Printing in Medicine*, 4(1), (2018), 1, <https://doi.org/10.1186/s41205-017-0024-1>.
- [9] MediaPipe, MediaPipe Hands, 2022. <https://google.github.io/mediapipe/solutions/hands.html>.
- [10] Paoli, A.; Neri, P.; Rationale, A.V.; Tamburrino, F.; Barone, S.: Sensor Architectures and Technologies for Upper Limb 3D Surface Reconstruction: A Review, *Sensors-Basel*, 20(22), (2020), <https://doi.org/10.3390/s20226584>.
- [11] Pickup, D.; Sun, X.; Rosin, P.L.; Martin, R.R.; Cheng, Z.; Lian, Z.; Aono, M.; Ben Hamza, A.; Bronstein, A.; Bronstein, M.; Bu, S.; Castellani, U.; Cheng, S.; Garro, V.; Giachetti, A.; Godil, A.; Isaia, L.; Han, J.; Johan, H.; Lai, L.; Li, B.; Li, C.; Li, H.; Litman, R.; Liu, X.; Liu, Z.; Lu, Y.; Sun, L.; Tam, G.; Tatsuma, A.; Ye, J.: Shape Retrieval of Non-rigid 3D Human Models, *Int J Comput Vision*, 120(2), (2016), 169-193, <http://dx.doi.org/10.2312/3dor.20141056>.
- [12] Siena, F.L.; Byrom, B.; Watts, P.; Breedon, P.: Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research, *J Med Syst*, 42(3), (2018), <https://doi.org/10.1007/s10916-018-0905-x>.
- [13] Tong, J.; Zhou, J.; Liu, L.G.; Pan, Z.G.; Yan, H.: Scanning 3D Full Human Bodies Using Kinects, *Ieee T Vis Comput Gr*, 18(4), (2012), 643-650, <https://doi.org/10.1109/TVCG.2012.56>.
- [14] Vitali, A.; Togni, G.; Regazzoni, D.; Rizzi, C.; Molinero, G.: A virtual environment to evaluate the arm volume for lymphedema affected patients, *Comput Meth Prog Bio*, 198, (2021), <https://doi.org/10.1016/j.cmpb.2020.105795>.
- [15] Zollhofer, M.; Stotko, P.; Gorlitz, A.; Theobalt, C.; Niessner, M.; Klein, R.; Kolb, A.: State of the Art on 3D Reconstruction with RGB-D Cameras, *Comput Graph Forum*, 37(2), (2018), 625-652, <https://doi.org/10.1111/cgf.13386>.