Semi-automatic Point Clouds Registration for Upper Limb Anatomy

Paolo Neri(✉), Beatrice Aruanno, Sandro Barone, Alessandro Paoli, and Armando Viviano Razionale

University of Pisa, Largo L. Lazzarino 1, 56122 Pisa, Italy
paolo.neri@unipi.it

Abstract. In this paper, a semi-automatic procedure to perform point clouds registration is presented. The method was developed for upper limb 3D scanning. During the acquisition, several frames are acquired from different points of view, to obtain a full 360° acquisition of the arm. Each frame stores both the point clouds coordinates and the corresponding RGB image. During post-processing, the RGB image is elaborated through a neural network, to detect relevant key points of the hand, which are then projected to the point clouds. The corresponding key points detected from different acquisitions are then used to automatically obtain a rough 3D rotation that aligns the point clouds corresponding to different perspectives in a common reference frame. Finally, the registration is refined through an iterative closest point algorithm. The method was tested on actual arm acquisitions, and the registration results are compared with the conventional fully manual 3-2-1 registration procedure, showing promising results of the proposed method.

Keywords: Semi-automatic registration · Upper limb 3D scan · Neural network

1 Introduction

In recent years, a great interest in the field of upper limb rehabilitation arose around the possibility of producing custom medical devices [1–3]. The use of bespoke devices can drastically enhance the patient's comfort and treatment effectiveness. This is a two-fold peculiarity of custom devices: the optimized design guarantees higher performances, and at the same time, the increased comfort enhances the patients' engagement in the therapy.

The first step in creating a bespoke device is usually the 3D scanning of the patient's anatomy to obtain a CAD model that will guide the design process of the device. Nowadays, many 3D scanning techniques exist based on passive (such as photogrammetry) or active (such as structured light) methods [4]. Regardless of the adopted technique, acquiring complex 3D shapes (such as human anatomy) always requires acquiring several frames from different viewpoints to achieve a full 360° acquisition of the geometry. Since the raw acquired data (i.e., point clouds) for each frame are measured in the scanner reference frame, the acquisitions from different perspectives need to be aligned in a common reference frame to reproduce the whole geometry [2]. This process is generally

S. Gerbino et al. (Eds.): JCM 2022, LNME, pp. 733–742, 2023.
https://doi.org/10.1007/978-3-031-15928-2_64

referred to as point clouds registration. In many applications, the scanning procedure is performed on a plaster reproduction of the studied geometry [5]. In this scenario, several robust approaches can be adopted to speed up and automatize the registration process, such as the use of rotary tables, which guarantee a precise and known placement of the object with respect to the scanning device.

On the other hand, many practical situations do not allow for the use of plaster reproductions, which would be time-consuming and would negatively impact the patient's comfort during the process. In these cases, real-time scanning devices are adopted to directly measure the patient's anatomy since they can minimize the presence of artifacts due to involuntary movements. In this context, the introduction of consumer-grade depth cameras (RGB-D camera) [6] has led to a significant impulse in the definition of affordable hand-held 3D scanners for biomedical applications [7–9]. These devices are moved around the target anatomy to collect 3D data from different viewpoints, which are not generally known a-priori [10]. The alignment between the different acquisitions is then generally achieved through the so-called 3-2-1 registration process, which is based on a first manual registration phase: at least three corresponding points are manually selected on two partially overlapping point clouds, to obtain a rough registration. The registration is then refined through automatic algorithms, such as the Iterative Closest Point (ICP) algorithm, which are based on processing the overall point cloud data set [11]. The manual point selection process requires skilled users, it is time-consuming and can result tedious when many patients must be treated. Thus, this step is generally automatized by gluing some markers on the target object, which can then be automatically detected by a software tool to determine corresponding points between overlapping areas of adjacent scans. The main drawback of this approach is the difficulty of automatically recognizing the different markers, which are manually glued on unknown locations of the target surface. Additionally, in the case of direct human anatomy measurements, the use of markers can result invasive for some fragile patients, thus any contact with the patient's body should be avoided.

This paper aims at automatizing the registration process in the specific case of upper limb anatomy 3D scanning. The research, which was developed in the framework of PRIME-VR2 European Project, exploits the information retrieved from different sensors to obtain a first rough registration, thus replacing the manual selection of corresponding points on different point clouds. This is achieved by using a neural network, which automatically detects the position of relevant key points of the hand shown in the acquired images. These key points are then reprojected on the corresponding point clouds, to obtain the reference points which are required for the initial rough alignment. This rough alignment is then further refined through conventional ICP algorithms to obtain a global registration. The developed method was tested on actual patients scanned at a Living Lab and compared with the conventional fully manual registration process. The activity has shown promising results in terms of semi-automatic clouds registration.

2 Scanning Equipment

In this research, a hand-held 3D scanner is considered. The scanner was developed by exploiting D415 Intel RealSense sensors [12]. The overall cost of the adopted hardware was about 600 €, which is much lower with respect to other commercial devices for upper limb acquisition, as deeply detailed in [4]. Each sensor integrates one RGB camera to acquire the texture, two IR calibrated cameras to measure the depth, and an IR laser light source to project a pattern on the target surface. Three sensors, placed at 90°, are mounted on a three-layered polymethylmethacrylate (PMMA) circular frame, in order to obtain a larger field of view with a single shot. The sensors are placed with both a radial and an axial orientation, to allow for a frontal acquisition with respect to the patient. During the scanning, the two diametral sensors acquire a wide-angle of the arm, while the third sensor is oriented behind the thumb to avoid undercuts. A schematic view of the placement of the sensors is reported in Fig. 1: Fig. 1(a) shows the front view, while Fig. 1(b) shows the lateral view. On the other hand, Fig. 1(c) shows a picture of the device.

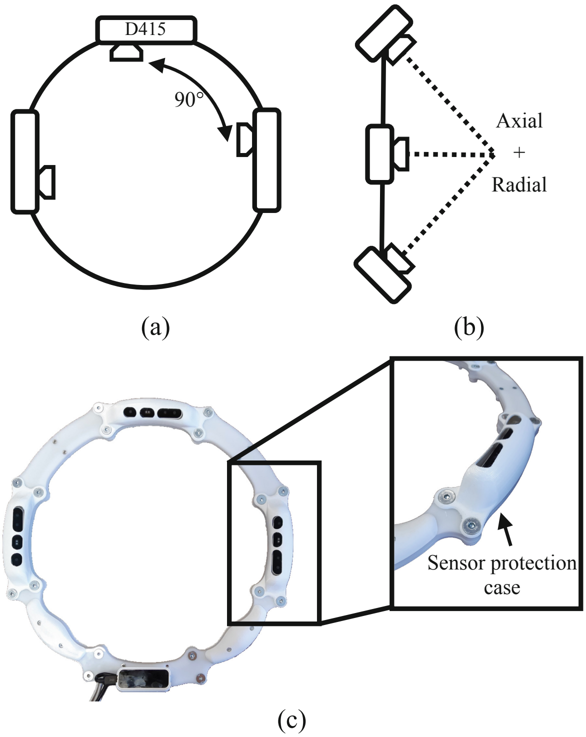

Fig. 1. Schematic layout of the scanning device: a) front view and b) lateral view. (c) Picture of the scanning device.

As can be seen, the sensors are protected by 3D printed covers, while all the cables are managed in the middle layer of the PMMA structure. Two ergonomic handles were added to the structure to facilitate the grip of the device during acquisition. Finally, a pedal switch is connected to the USB hub of the device, acting as a trigger to save the frames during acquisition. The three sensors are preliminarily calibrated, using the calibration procedure described in [12], thus allowing to align the scans provided by each sensor in a common reference frame.

The control software of the device was developed in Matlab® language, featuring a graphical user interface (GUI) that allows setting the main parameters, such as the total number of frames to be acquired. Additionally, a large area of the GUI shows in real-time the point clouds as acquired by the three sensors. An automatic algorithm computes the distance between the point cloud centroid and the device, highlighting with different colors if the distance is suitable for the acquisition, too far or too close. In particular, three colors are used for the point clouds depending on their distance with respect to the device: blue if the distance is lower than 450 mm (too close), green if the distance is between 450 and 550 mm (optimal), and red if the distance is larger than 550 mm (too far). This expedient enhances the usability of the device even for un-qualified users, such as medical operators who are not experts in 3D scanning.

Practically, the protocol to obtain a full 360° scan of the patient's arm requires:

1. asking the patient to stretch the arm, laterally or frontally depending on his comfort;
2. placing the scanning device in correspondence of the first location at the optimal distance;
3. acquiring the first frame;
4. moving the device around the arm to the next location, guaranteeing a field-of-view overlap with respect to the previous acquisition, and acquiring the corresponding frame;
5. repeating the previous step until all the required frames have been collected.

The preliminary testing campaign proved that a number between six and ten different placements of the acquisition device is generally enough to obtain a full 360° scanning of the patient's arm. The time required to complete the described procedure is in the range of 50–80 s, depending on the overall number of frames to be acquired. It is worth noting that each saved frame stores three distinct point clouds (in terms of 3D coordinates) and the corresponding RGB images. Additionally, since the RGB camera is calibrated with respect to the IR cameras, it is possible to correlate each 3D point with the corresponding pixel of the RGB image.

3 Conventional Manual Registration Procedure

The described procedure allows the acquisition of a certain number of point clouds, each corresponding to a portion of the arm acquired from a different viewpoint. These point clouds must then be registered into a common reference frame, to obtain the final 3D model of the arm. This is generally performed through the manual 3-2-1 registration process. Given a pair of point clouds, the operator must select at least three corresponding points on each cloud. Since three points allow for the definition of a rigid rotation between two different reference systems, this preliminary procedure allows for the rough registration of the two clouds. It is worth noting that the presented approach was developed for rehabilitation patients, and all the clinical centres cooperating in the research agreed that the use of markers could reduce patients' comfort and engagement in the process, and would increase the overall scanning time due to the preparation stage. Thus, since no markers are used on the patient to avoid discomfort, the manual selection of the corresponding points is based on geometrical and anatomical considerations of the operator, or on natural skin features. This subjective procedure, based on the experience of the operator, is cumbersome, time-consuming, unreliable, and not repeatable. For this reason, the manual alignment is usually followed by an automatic registration refinement, which is generally performed through an iterative closest point (ICP) algorithm. This algorithm automatically looks for corresponding points on overlapping regions of the clouds to be registered, iteratively adjusting the relative positioning to reduce the surface distance. While the described 3D scanning procedure can be completed in the range of one minute, this registration procedure can be time-consuming, taking from 5 to 10 min depending on the number of point clouds and on the extension of the overlapping regions. An example of the above-described process is shown in Fig. 2 in the case of two distinct viewpoints. Figure 2(a) shows the corresponding points as selected on the two clouds, Fig. 2(b) shows the resulting rough registration while Fig. 2(c) shows the final registration obtained through the ICP algorithm.

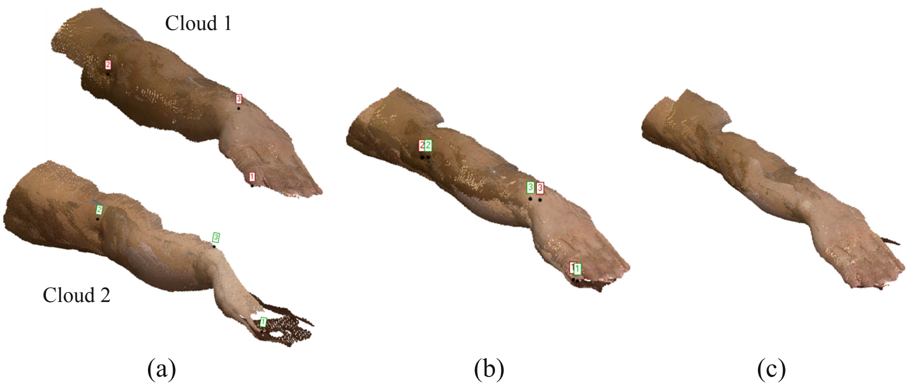

Fig. 2. Manual 3-2-1 registration process: (a) selection of three corresponding points, (b) rough initial alignment and (c) fine registration through ICP algorithm.

4 Semi-automatic Registration Strategy

The bottleneck of the described conventional 3-2-1 procedure is represented by the manual selection of corresponding points. Indeed, this selection is to be repeated for any couple of clouds, which must be selected, rotated, visually compared to identify corresponding points. Since no markers are present on the patient's arm, the choice generally falls on easily recognizable locations, such as fingertips, phalanges, wrist bones, and elbow. Following this principle, the application of a neural network (NN) to detect key points on RGB images was found to have a great emphasis in the literature. Indeed, a properly trained NN could be able to automatically detect the aforementioned locations. Indeed, since the specific task (i.e. arm/hand acquisition) can be of interest in several different fields, several NNs are available, which detect some specific locations of the hand. In particular, the "MediaPipe Hands" NN was used [13]. The main advantages of this specific NN rely on being open-source and implemented also in Python. This allows for easy integration of the automatic key-points detection in the overall workflow. Additionally, the possibility to customize the output simplifies the subsequent data processing, ensuring that the same notation and key-points order is always adopted. The key points detected by the NN will then be used to obtain a rough alignment of the different viewpoints, which will be fine-registered through an ICP algorithm.

4.1 Key-Points Detection Through Neural Network

The described NN requires as input an RGB image, and returns as output a list of the detected key points and a detection score. The results obtained on a sample hand image are reported in Fig. 3. As can be noted, in ideal conditions, the NN infers 21 landmarks on the hand, four for each finger plus one on the wrist. These key points are always stored in the same order. In the case that some points are not detected on a specific image, a NaN value is returned, so that the specific missing key point can be skipped avoiding the risk of erroneously assigning the numbering. This aspect is crucial since the correspondence between key points detected from different viewpoints is achieved by comparing the indexing value.

Fig. 3. Results of the NN on a sample hand image.

4.2 Key-Points Selection and Rough Cloud Alignment

As stated, each frame corresponds to one point of view and is composed by the acquisition of the three D415 sensors, i.e. three point clouds (already registered exploiting the calibration stage) and three RGB images. Theoretically, each of these three images could be equivalently used to detect the key points, since the clouds are already registered. Nevertheless, this is not feasible in practice, since the success of the NN in detecting the 21 key points strongly depends on the hand orientation with respect to the camera and the visibility and pose of the fingers. For these reasons, the results obtained on the three images of the same frame can be largely different, since in some images a few key points could be missing or the whole identification could fail. It is worth noting that three points are sufficient to obtain a rigid rotation, thus not all 21 key-points need to be detected for all the poses. Nevertheless, a higher number of key points guarantees data redundancy and better registration performances, thus the best results are detected by multiplying the identification score provided by the NN by the number of detected key points to compute the total score. The image in the RGB triplet which has the maximum total score is selected for the registration, and the 2D pixel coordinates of the key points are projected on the corresponding 3D locations on the point clouds and stored for further processing. An overview of the procedure is schematized in Fig. 4.

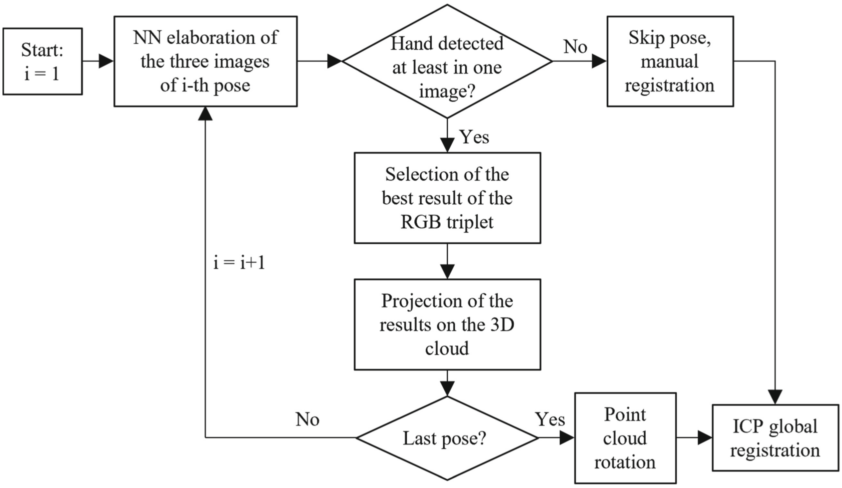

Fig. 4. Schematic workflow of the algorithm.

In the unlikely event of NN failing in detecting the key points in all the three images of a specific pose, that pose is skipped and processed through a conventional manual 3-2-1 procedure. Anyway, in the performed trials, the NN was always able to detect at least three key points in at least one of the RGB images for all the acquired poses. Once each pose has been independently processed through the NN algorithm, the rough rigid transformations can be computed. To this extent, each pose is registered with respect to the previous one. Anyway, the key points detected by the NN for two different poses may not have the same indexing. For example, for pose n.1 all 21 key-points may be detected, while for pose n.2 key-points 5 and 6 may be missing, thus impairing the possibility to register the clouds. To solve this issue, only the intersection between the key points detected in the two poses is considered for the alignment. After the common key points are detected, the rough alignment is computed by finding the roto-translation which minimizes the mean square distance between corresponding points. After this procedure is completed for all the acquired poses, all the clouds are roughly aligned into a common reference frame.

4.3 Registration Refinement

The final step of the alignment procedure is the registration refinement. The described algorithm allows obtaining a rough initial alignment, which is affected by errors due to the NN precision in detecting the key points. Moreover, only a maximum of 21 key points belonging to the hand are considered, which are a drastically reduced subset with respect to the thousands of points of the whole point cloud. The overall point clouds data set is then processed through an ICP algorithm to minimize the alignment error. This algorithm can automatically detect the overlapping regions between adjacent point clouds, and iteratively compute a roto-translation that minimizes the distance. The final result is the registered 3D full acquisition of the upper limb anatomy, which can be further processed through conventional point cloud elaboration pipelines (e.g. noise filtering, point sampling).

5 Results

The described procedure was tested on actual measurements taken by a clinician on a patient's arm, which are representative of realistic data in operational environment. The acquired data were both processed through the conventional 3-2-1 procedure and the presented algorithm, in order to compare the results. Figure 5(a) shows an example of the surface resulting from the conventional manual procedure. As can be noted, a complete 360° view of the arm was obtained, without holes. More noisy data were obtained in the fingers region, because of the undercut and the occlusion during the scanning process. On the other hand, Fig. 5(b) shows the surface resulting from the described automatic procedure. As can be seen, differences between the two geometries are not easily detectable at a first glance. Finally, Fig. 5(c) shows the deviations between the surfaces obtained with the manual procedure and with the proposed procedure. All the values are comprised in the range ± 2 mm, and the main deviations are found in the correspondence of the hand region where more noisy data are expected. Nevertheless, the deviation histogram on the right of Fig. 5(c) evidences that most of the points deviate less than 0.1 mm between the two surfaces, thus demonstrating that the two procedures achieve equivalent results. It is worth noting that, in the described application, deviations in the millimetres order of magnitude are considered acceptable.

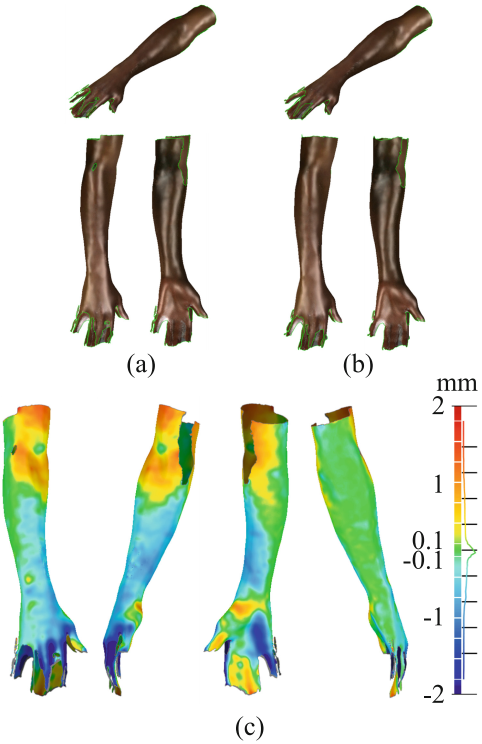

Fig. 5. Comparison between scanning elaboration: (a) manual procedure, (b) proposed procedure and (c) deviation map between manual and automatic procedure surfaces

6 Conclusions

This work focuses on the registration of point clouds acquired through 3D scanning of upper limb anatomy. A Neural Network (NN) was applied to RGB images to define a semi-automatic registration procedure, thus overcoming the conventional and manual 3-2-1 registration procedure. The NN automatically detects up to 21 key points on the hand, which are then reprojected on the 3D point cloud and used to obtain a first rough registration. This step practically replaces the tedious and unreliable manual selection of corresponding points on different point clouds, substantially reducing the registration time. Data post-processing is then refined through a conventional pipeline, i.e. fine registration (Iterative Closest Point algorithm), denoising, surface extraction.

The proposed procedure was then validated through experimental analysis of human arm acquisition. The same raw data were processed both using the conventional manual procedure and the proposed automatic procedure. The surfaces resulting from the two procedures are not distinguishable by the naked eye. Furthermore, a deviation analysis was performed, demonstrating that most of the deviations are in the range ± 0.1 mm, while the worst cases are in the range ± 2 mm, which is considered acceptable for the biomedical application. Finally, a further enhancement of the proposed approach can be achieved by extending the NN to the detection of additional keypoints on the arm, instead of only considering the hand, to better constraint the registration process.

Acknowledgement. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 856998.

References

1. Buonamici, F., et al.: A practical methodology for computer-aided design of custom 3D printable casts for wrist fractures. Vis. Comput. **36**(2), 375–390 (2019). https://doi.org/10.1007/s00371-018-01624-z
2. Volonghi, P., Baronio, G., Signoroni, A.: 3D scanning and geometry processing techniques for customised hand orthotics: An experimental assessment. Virtual Phys Prototy **13**, 105–116 (2018)
3. Asanovic, I., Millward, H., Lewis, A.: Development of a 3D scan posture-correction procedure to facilitate the direct-digital splinting approach. Virtual Phys. Prototy. **14**, 92–103 (2019)
4. Paoli, A., Neri, P., Razionale, A.V., Tamburrino, F., Barone, S.: Sensor architectures and technologies for upper limb 3D surface reconstruction: A review. Sensors-Basel **20**, 6584 (2020)
5. Barrios-Muriel, J., Romero-Sánchez, F., Alonso-Sánchez, F.J., Salgado, D.R.: Advances in orthotic and prosthetic manufacturing: A technology review. Materials **13**(2), 295 (2020). https://doi.org/10.3390/ma13020295
6. Zollhofer, M., et al.: State of the art on 3D with RGB-D Cameras. Comput. Graph. Forum **37**, 625–652 (2018)
7. Redaelli, D.F., Barsanti, S.G., Fraschini, P., Biffi, E., Colombo, G.: Low-cost 3D Devices and Laser Scanners Comparison for the Application in Orthopedic Centres. Int Arch Photogramm Remote Sens Spatial Inf Sci **XLII–2**, 953–960 (2018)
8. Oranges, C.M., et al.: Three-dimensional assessment of the breast: Validation of a novel, simple and inexpensive scanning process. In Vivo **33**, 839–842 (2019)
9. Dessery, Y., Pallari, J.: Measurements agreement between low-cost and high-level handheld 3D scanners to scan the knee for designing a 3D printed knee brace. PLoS ONE **13**, e0190585 (2018)
10. Li, J., Tanaka, H.: Feasibility study applying a parametric model as the design generator for 3D–printed orthosis for fracture immobilization. 3D Printing Med. **4**(1), 1–15 (2018). https://doi.org/10.1186/s41205-017-0024-1
11. Zhu, H., et al.: A review of point set registration: From pairwise registration to groupwise registration. Sensors **19**(5), 1191 (2019). https://doi.org/10.3390/s19051191
12. Neri, P., Barone, S., Paoli, A., Razionale, A.V., Tamburrino, F.: A depth-camera based system for the real-time scanning of upper limb anatomy. In: Rizzi, C., Campana, F., Bici, M., Gherardini, F., Ingrassia, T., Cicconi, P. (eds.) ADM 2021. LNME, pp. 245–255. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-91234-5_25
13. MediaPipe Hands: Google LLC (2020). https://google.github.io/mediapipe/solutions/hands