An Iterative Method for 3D Body Registration Using a Single RGB-D Sensor

Victor Villena-Martinez, Department of Computer Technology, University of Alicante, Alicante, Spain

Andres Fuster-Guillo, Department of Computer Technology, University of Alicante, Alicante, Spain

Marcelo Saval-Calvo, Department of Computer Technology, University of Alicante, Alicante, Spain

Jorge Azorin-Lopez, Department of Computer Technology, University of Alicante, Alicante, Spain

ABSTRACT

In this paper, the problem of 3D body registration using a single RGB-D sensor is approached. It has been guided by three main requirements: low-cost, unconstrained movement and accuracy. In order to fit them, an iterative registration method for accurately aligning data from single RGB-D sensor is proposed. The data is acquired while a person rotates in front of the camera, without the need of any external marker or constraint about its pose. The articulated alignment is carried out in a model-free approach in order to be more consistent with the real data. The iterative method is divided in stages, contributing to each other by the refinement of a specific part of the acquired data. The exploratory results validate the proposed method that is able to feed on itself in each iteration improving the final result by a progressive iteration, with the required precision under the conditions of affordability and unconstrained movement acquisition.

KEYWORDS

3D Body Registration, Articulated Registration, Depth Sensor, Microsoft Kinect, RGB-D

1. INTRODUCTION

Nowadays, there are several techniques to obtain a 3D model of the human body. This model is useful in many applications of different scopes: medical, textile, shoes, etc. The tandem of 3D technology and medicine has a long trajectory, 3D models have been used to help experts study patients and make decisions (Treleaven & Wells, 2007), e.g.: dietetic treatment, child growth monitoring, podiatry, orthopedics, among others. There are several solutions to obtain a high definition 3D model of the human body, but they are expensive and bulky, making its portability more difficult, with a prohibitive cost. The usefulness of these models has made more attractive the development of inexpensive systems with few restrictions for the subject pose during the acquisition. Consumer RGB-D, low-cost RGB-D, or just RGB-D sensors have become popular in fitting those requirements (Lai, Bo, Ren, & Fox, 2013) due to their combination of affordability and portability. However, the accuracy of them is not high enough for some applications.

Therefore, in this paper we face the problem of providing a 3D body representation using RGB-D sensors. The proposal has to fit some requirements: (1) affordability and flexibility for broad transferring solutions, (2) unconstrained method in terms of free movement of the subject in front of the sensor, (3) accurate alignment of the data to create an accurate representation.

Low-cost RGB-D sensors are considered a new type of sensors due to the large usage they are receiving in the scientific and industrial community. This kind of sensors combine color and depth information, the later one estimated using Time of Flight (ToF) or structured-light techniques. In this work, we propose a method for RGB-D devices, due to the low-cost requirement, and their accuracy are appropriate for body modeling since they were originally developed for body movement tracking (Saval-Calvo et al., 2017). However, it could be possible to use our approach with any other sensor which provides color and depth information.

To develop a full model of bodies using RGB-D data, our proposal faces the registration or alignment of 3D points. The registration is the process to align one data set onto one or more data sets. It could be performed rigidly and non-rigidly, which refer to transform the data all with the same transformation, or by applying different transformations to each data point (Saval-Calvo, Azorin-Lopez, Fuster-Guillo, & Mora-Mora, 2015). Registration of 3D data is a widely studied problem. (Henry, Krainin, Herbst, Ren, & Fox, 2014) made a 3D reconstruction of indoor environments with an RGB-D sensor. The work of (Lovato, Bissolo, Lanza, Stella, & Giachetti, 2014) carried out an accurate 3D registration of the foot using a PrimeSense rotating around it. These sensors could be used in combination with augmented reality markers in the scene in order to estimate more accurately the transformation to obtain the 3D model (Mihalyi, Pathak, Vaskevicius, Fromm, & Birk, 2015).

The 3D registration of the human body is complex due to its articulated nature, and the impossibility to keep the same pose in different instants of time. These inconveniences could be approached by different acquisition approximations to obtain a model:

- Single view, performing a partial reconstruction of the subject of interest.
- Adding sensors to acquire the subject from different angles in the same instant of time, avoiding any movement of the subject.
- Controlled environment, using external elements to the sensor, like augmented reality markers.
- Using articulated/isometric techniques for the registration. These techniques take into account the movements performed by the subject during the acquisition.

This work is focused on the fourth approach in order to provide a registration of the human body fitting the requirements above mentioned of affordability, flexibility and accuracy.

The rest of the paper is organized as follows: a study of the background is presented in section 1.1; an explanation of the proposed method is done in section 2; the experiments are presented in section 3; finally, the conclusion is presented in 4.

1.1. Related Works

Various proposals for body scanning could be found, either related to the sensor system or to the registration. There exist commercial body scanners, like *Vitus BodyScan* or the *Cyberware Whole Body Scanner*. However, these systems are expensive. Some works carried out the reconstruction of body introducing multiple RGB-D devices (Tong, Zhou, Liu, Pan, & Yan, 2012; Shapiro et al., 2014; Lin, Chen, Ralph, & Cheng, 2016). Nevertheless, the affordability and flexibility are required in this work. One restriction is the use of a single RGB-D sensor, avoiding complex instrumentation, but forcing the rotation of the body in front of the camera. Then, the use of articulated or isometric techniques are mandatory (Pellegrini, Schindler, & Nardi, 2008) to overcome misalignment problems.

The articulated alignment is currently under intense study. In human bodies, it consists of non-rigid alignment of the different parts in the body, taking into account the joints, allowing pose variations

during the acquisition. It divides the body into different parts defined by pairs of joints, and rigidly registers each part. To determine each part the work (Charles & Everingham, 2011) differentiates each one from a single view obtained with an RGB-D sensor. Shotton et al. (2013, p. 116) proposed a method to estimate the parts of the body and its pose in real time from a single frame using random forest. Schwarz et al. (2012, p. 217) estimated the pose from depth data, using geodesic distances and optical flow for occlusions. Other works use Convolution Neural Networks (CNNs) for pose and skeleton estimation (Chen et al., 2016; Oliveira, Valada, Bollen, Burgard, & Brox, 2016; Wei, Ramakrishna, Kanade, & Sheikh, 2016).

Focused on the articulated 3D reconstruction of human body, two approaches have been used, model-free and model-based techniques (Bogo, Black, Loper, & Romero, 2015). The former needs multiple depth frames from different view-points to get information of the whole body and performs the 3D reconstruction. The latter techniques fit the data points on a generic model as (Weiss, Hirshberg, & Black, 2011), which uses the SCAPE (Anguelov, Srinivasan, & Koller, 2005).

Model-free approaches use different techniques to get the 3D model. Barmpotis et al. (2013, p. 1347) carried out a real-time reconstruction of a body of a subject in front of the camera through parametrization of cylindrical objects using Cartesian tensor and b-splines. Cui et al. (2013, p. 133) proposed a method to obtain a full body model using a single depth camera with a Kinect sensor and employing an articulated model to describe de non-rigid motion of the body forcing the subject placed in "T" pose. Wang et al. (2012, p. 432) propose a cylindrical representation of the body model, performing the registration in a top-bottom-top manner and using the cylinders to simplify the registration to a bi-dimensional problem.

Our main goal is to develop a system for 3D body registration of subjects. The system aims to be flexible, accurate, and affordable. In order to accomplish this objective, we propose to use a single RGB-D sensor, due to its affordability and flexibility, being the subject who rotates in front of the camera. Furthermore, the registration is performed in a model-free approach using all the acquired information from a large set of frames. To the best of our knowledge, there exist the need to propose a system which accurately registers full body fitting the requirements above mentioned. To achieve the accuracy of the problem, we propose an iterative method for 3D reconstruction of the human body, based on an initial estimation of the skeleton. This method is divided in three main stages, each one contributing to the others by the refinement of a specific part of the data, performing an iterative strategy fitting the accuracy of the problem.

2. METHOD OVERVIEW

In this paper, an iterative method for 3D registration of the human body is proposed. This method is based on the initial estimation of a 3D medial axis (i.e. a skeleton of the subject), which is iteratively improved by the registration process itself. Initially, the proposed method expects as input a set of views, being each one a 3D point cloud aligned with the color information of the scene. In these set of views, the subject is rotating by itself in front of the camera in a "tired T" pose (see Figure 1) that is appropriate for perceiving the whole body along the different views.

Figure 1 shows the outline of the method composed by three stages that are fed back in each iteration in order to improve the final result.

The first stage estimates the skeleton of the subject (position of body joints related to the 3D point set). Moreover, in each iteration, the information of previous skeletons and registered points are used to overcome possible problems (see Section 2.1). At the labeling stage, the different parts of the body are detected and each point in the view is accordingly labeled (forearm, chest, etc.). The points labels are used in the isometric alignment of the data. Finally, the registration stage performs the articulated alignment of the 3D data having as output a set of aligned views that are, moreover, used as feedback for the next iteration of the process.

Figure 1. Iterative scheme of the method

Table 1. Set of joints model used in this paper

| N	Joint	N	Joint	N	Joint	N	Joint
1	Head	7	Left wrist	13	Right hand	19	Left knee
2	Neck	8	Left elbow	14	Right thumb	20	Left ankle
3	Spine shoulder	9	Left shoulder	15	Right hand tip	21	Right hip
4	Left hand tip	10	Right shoulder	16	Middle spine	22	Right knee
5	Left thumb	11	Right elbow	17	Spine base	23	Right ankle
6	Left hand	12	Right wrist	18	Left hip		

2.1. Skeleton Estimation

The skeleton model to be estimated is composed by a set of joints and bones. The joints are a set of 3D points in the location of the body joints related to the input point set. The bones of the skeleton are segments connecting the joints keeping the shape model of a human body. As a case of study in this paper, the skeleton model used is composed by 23 joints (see Table 1) and 25 bones but any other approach can be used.

The skeleton estimation is a challenging problem, mainly when the subject is not facing the camera and becoming very difficult when the body is positioned from the side with respect to the camera. In order to correct the estimation errors, different refinements have been proposed:

- Re-estimating left and right joints and bones in case the subject back to-wards the camera (see 2.1).
- Detecting and discarding wrongly estimated joints by overlapping (see 2.1). For example, when the subject is seen from the side, both arms are detected in the same position (see Figure 2).
- Re-estimating joints placed in wrong locations, discarded, or those that have not been estimated (see 2.1).

Figure 2. Overlapped joints problem. Left: Wrong joints in an arm. Both are detected in the same side. Right: Skeleton correction after correcting wrong joints
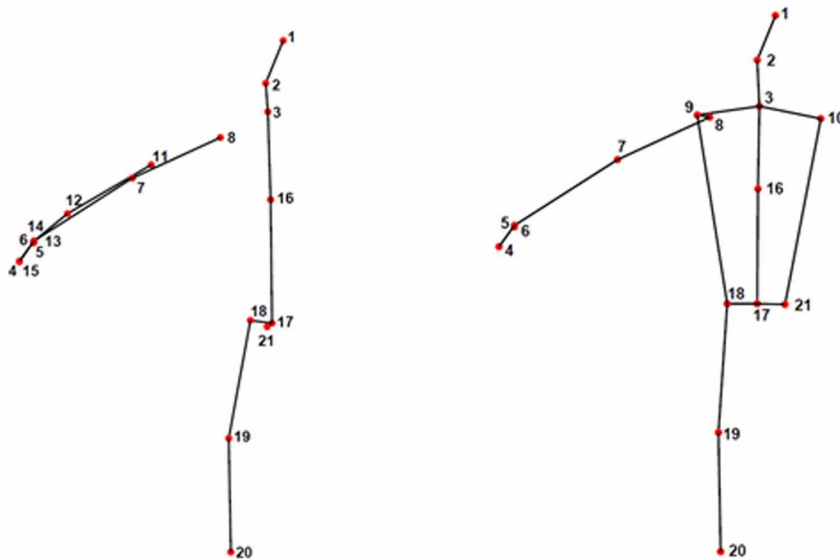
2.1.1. Backward Correction

One of the usual problems in body skeleton estimation algorithms become when the subject is facing backwards to the camera due to the calculated joints are not correct being exchanging the left and right ones.

   The proposed method checks distances between the same joints in different views. The joints are clustered into four groups: right arm (10-14), left arm (4-9), right leg (21-23) and left leg (18-20). Then, the view where the subject is facing the camera is taken as reference (i.e. view with best skeleton). For each view, the clusters are analyzed so as each joint is compared to its corresponding and its opposite (other side). If the distance is larger in the corresponding, an exchange is performed to correct it. It is important to remark the need of a prior alignment of the torso (the most static part of the body).
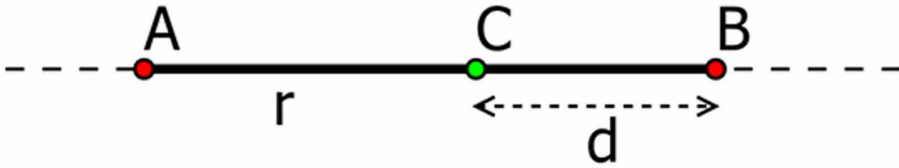
2.1.2. Overlapped Joints

Another common problem in skeleton estimation algorithms is that, in some cases, the joints are overlapped. Figure 2 shows the estimation of the skeleton when the subject is seen from the side with respect to the camera. Figure 2-left shows the problem, the joints of both arms (4-8 and 11-15 from Table 1) are overlapped. This problem will result in a wrong labeling of the body parts. Again, this correction also needs that the torso is previously registered.

   First, the method improves the estimation of joints by positioning those from the reference frame into the remaining views once pre-aligned. Then, the overlapped joints are detected assuming that the Euclidean distance between hands should be greater than the distance between shoulders (joints 9 and 10).

   Next, it is necessary to decide which of the two groups of joints (left or right) is wrong. Hence, we calculate the distance between the hand and both its corresponding shoulder and the opposite one. If the former is larger, those joints are incorrect and a re-estimation is required as explained in Section 2.1.3.

Figure 3. Segment from A, hand, to B, shoulder, to find C elbow or wrist

2.1.3. Re-Estimation

After the correction of the overlapped joints, the skeleton model may have some non-detected joints, or wrongly positioned ones. For example, in Figure 2-right image, the wrist joint is placed at the elbow, and the elbow joint at the shoulder, producing errors in the labeling stage. To solve this problem, the first step is to detect the frames where the joints are wrong estimated by calculating the standard deviation of the distances between the joints showed in Table 2. Then, the distances for each frame are calculated and compared with the reference frame, checking the following condition: |*distance - reference distance*| > *std*, where *std* is the standard deviation of all the distances. In those frames were the condition is true, a re-estimation of the wrist and elbow joints is needed.

Assuming a "tired T" pose of the subject, the problem could be solved as basic geometry problem. Given two points A and B, being the hand and shoulder as shown in Figure 3, and the segment which connects them, we find a third point C (elbow or wrist) that belongs to the segment and is placed at a certain distance from A or B. These distances, from the hand to the wrist and from the shoulder to the elbow, are recalculated taking into account those from the rest of views.

2.2. Part Labeling

The articulated registration of the human body needs to know to which part of the body each point of the point cloud belongs to. These parts are defined taking into account the joints of the skeleton and the bones, which are the segments defined by pairs of joints following human body shape. With the information of the previous stage, each 3D point of the point cloud per view is classified in a part of the body. In the experiments, 12 parts have been considered (Table 3) but this could be different depending on the requirements of the problem.

To classify the points, the distance from each 3D point to each bone of the skeleton is computed. The closest bone to a specific point determines the part in which will be classified. The distance from a point to a line is given by the orthogonal projection of the point in the line, so it could be calculated with basic three-dimensional geometry.

Next, the topology defined by a part is validated, checking the connectivity of the points under the same label. Each part or label only contains connected points. Sometimes, if two parts are close, points from one could be labeled as the other. In this case, the clusters are analyzed selecting one of them as correct, and classifying again the remaining groups.
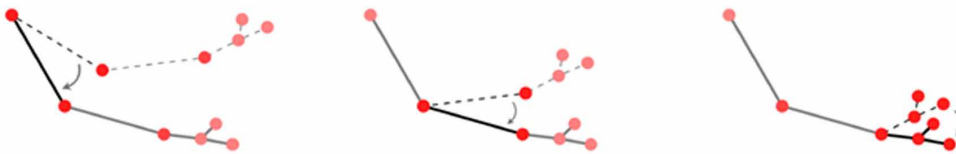
2.3. Registration

The registration stage performs the alignment of different views to obtain the whole 3D model. This alignment is usually performed with rigid registration techniques. Rigid registration algorithms could be classified in terms of (Salvi, Matabosch, Fofi, & Forest, 2007): precision (coarse and fine) and registration strategy (pair-wise and multiview). However, rigid registration techniques are not enough, such as in the problem of body modeling due to the acquisition of the different viewpoints are not simultaneously taken and the subject is in motion. For that reason, it is necessary to take advantage of the topological information provided in previous stages, i.e. joints, bones and body parts labeling.

Table 2. Joint distances calculated for correct position of joint in the arms

Joint	
6 – 7	Left hand – wrist distance
9 – 8	Left shoulder – elbow distance
10 – 11	Right shoulder – elbow distance
12 – 13	Right hand – wrist distance

Figure 4. Arm articulated registration process. From left to right, first the biceps are registered propagating to the rest. Second, the forearm is registered propagating to the hand. Thirdly, the hand is registered

In this paper, the articulated registration is approached by dividing the model in parts and aligning each one independently with a rigid registration method. Then, the transformation has to be propagated to the dependent parts, e.g.: in case of the arm (Figure 4), a transformation to the biceps must be applied to the forearm and hand, and so on. The dependencies between the different parts of the body are defined by the graph in Figure 5.

Hence, the registration is divided into different steps (see Figure 6). First, the registration of the torso is performed to coarsely align the different views of the subject. Second, the views are filtered to determine which one could be used. Finally, the registration of the limbs is carried out.

2.3.1. Torso Registration

The torso is fine registered using Iterative Closest Point (ICP) (Rusinkiewicz & Levoy, 2001) with a pair-wise strategy. The main objective of this step is to perform the alignment of all views based on the torso, due to the remaining parts depend on it and it is the most static part of the body. This implies that a transformation of the torso must be applied to other parts. In this case, it is possible to use a fine approximation without a previous alignment due to the closeness between consecutive frames having a minimum variation between consecutive views. If the views are not close each other, a pre-alignment could be used.

2.3.2. Filtering Views

The views are filtered to select and remove the parts that are not used to be registered. The method discard the parts that lie in one of the next situations: the part is occluded (side views); the number of points of a part is widely different that the average in the rest of views; and, the distance between joints in the skeleton is widely different that the same in the rest of frames.

2.3.3. Limbs Registration

In this step, each part is aligned individually, taking into account the dependencies with other parts in order to apply the transformations (see Figure 5). The head usually has few points and a uniform geometry mainly in the back, which difficult its alignment. To register this part, first, a coarse alignment
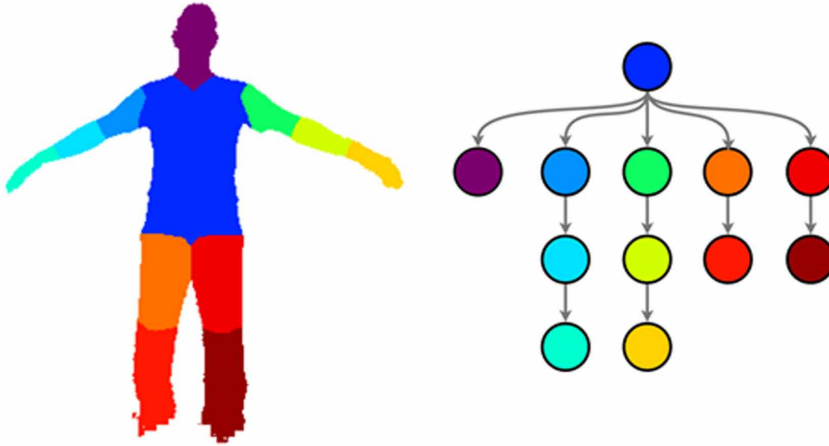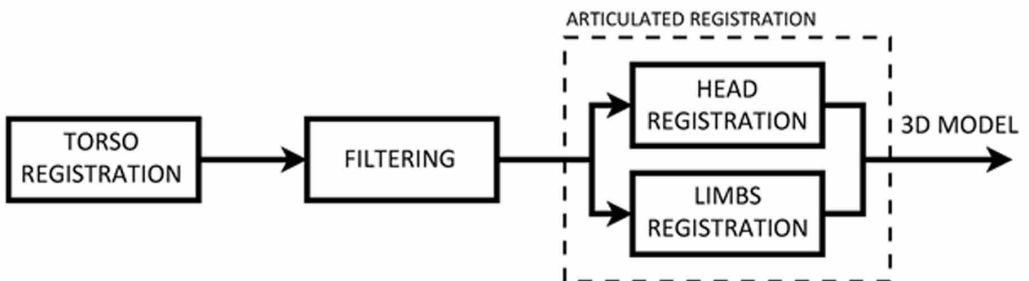
Figure 5. Dependencies between body parts

Table 3. Labels and the associated part of the body

Part		Part		Part	
1	Head	5	Left hand	9	Left thigh
2	Torso	6	Right biceps	10	Left calf
3	Left biceps	7	Right forearm	11	Right thigh
4	Left forearm	8	Right hand	12	Right calf

Figure 6. Schematic of articulated registration

is performed using Random Sample Consensus (RANSAC) (Fischler & Bolles, 1981) with visual features as SIFT and SURF. Then, a fine alignment it is done with ICP.

The arms and legs are composed by different parts, so the transformations are applied according to the dependencies of each part (see Figure 4). In this step, a pair-wise strategy is used between consecutive frames with a coarse registration based on Principal Component Analysis (PCA) and ICP as fine refinement. After that, a multiview strategy is applied to each part for improving the final result. This strategy consists in align each view with the whole model in order to polish the imperfections produced by the registration of consecutive views with a pair-wise strategy.

3. EXPERIMENTS

In this section, a set of experiments have been carried out to verify the results provided by our method. This exploration results are evaluated using visual inspection. The setup includes an RGB-D sensor Microsoft Kinect V2 and the Microsoft Kinect SDK. The SDK provides the skeleton of the subject, estimated with a variation of the method proposed by (Shotton et al., 2013) that provides joints for thumbs tracking, which are not useful for this method.

The experimentation is divided into three parts: overlapped joints detection and correction; joints re-estimation when they are wrongly estimated; 3D registration of body subjects with the proposed iterative method.

3.1. Overlapped Joints

The proposed method can solve problems where both arms have been detected in the same side, shown in Figure 7a. This problem has direct consequences in the labeling and registration stages, as can be seen in Figure 7b which shows the result of the labeling stage with the previous wrong skeleton where the arm is divided into regions that do not correspond to any correct human model. With the proposed method in Section 2.1, the skeleton is corrected by removing the wrong joints, and re-estimating some of them as can be seen in Figure 7c. Then, the 3D data can be properly labeled (Figure 7d) to provide useful information for the articulated registration step.

3.2. Joints Re-Estimation

The second experiment studies the joints re-estimation, when the skeleton estimation missed some joint. For example, Figure 8a shows the skeleton where the joint 9, corresponding with the left shoulder, has not been detected. This problem produces erroneous detection of body parts, in this case the left biceps and shoulder are equally labeled (Figure 8b). The proposed method can solve this situation by providing the skeleton showed in Figure 8c making the left biceps distinguishable in the labeling stage.

3.3. Articulated Registration

In order to evaluate the articulated registration, a sequence of frames of a man facing the camera making free movements with his arms is used. A subset of these frames is showed in Figure 9. Figure 9-bottom-right depicts the 3D points of all frames superimposed. After applying the proposed method, the articulated alignment achieves an accurate registration of the views represented in Figure 10.

Figure 7. 7b shows the wrong joints estimation due to the arms overlapping, and 7a shows the problems in labeling of the arm. 7c shows the joint correction and 7d shows the adequate labeling after applying the proposed method

(a) Wrong skeleton   (b) Wrong labeling   (c) Correct skeleton   (d) Correct labeling
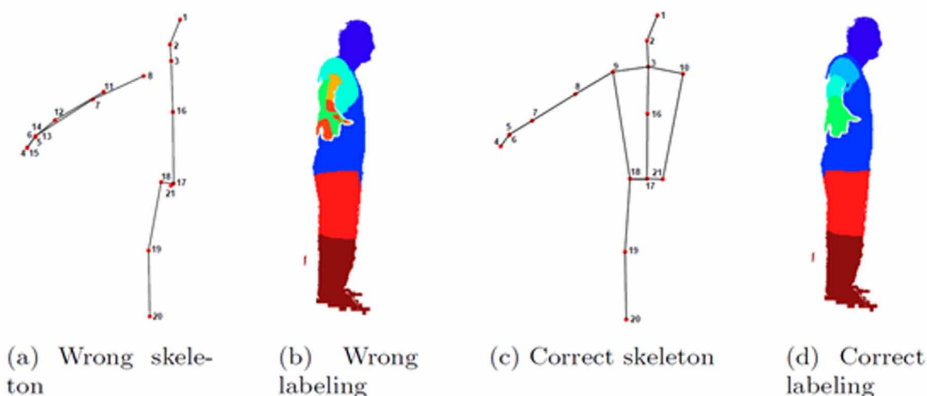
3.4. 3D Body Model

The purpose of the proposed method is to obtain a 3D registered model of a human body. To evaluate the overall performance of the proposal, various experiments have been carried out. Because of the similarity between the results, only one example is shown in this paper, but similar accuracy has been achieved for the rest of models. The model is obtained from a sequence of 281 images of the subject rotating by itself in front the camera. Then, the whole method has been applied to those frames, obtaining an accurate result, as shown in Figure 11 in frontal and profile view.

Figure 8. Figure 8a shows the skeleton model in which the joint 9 has not been estimated; 8b shows the wrong labeling of the data where shoulder and biceps are the same part; 8c presents the corrected skeleton with the proposed method, which leads to an adequate labeling shown in 8d
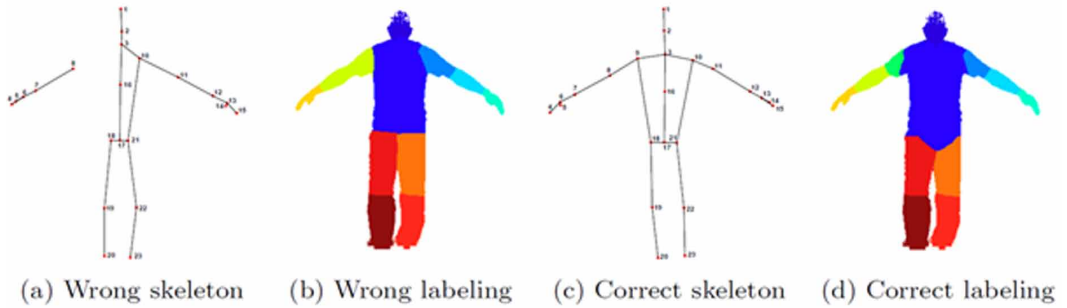

Figure 9. First row and three left figures in bottom row are a subset of frames of free movements with the arms. Bottom right shows 3D data overlapped of the movement

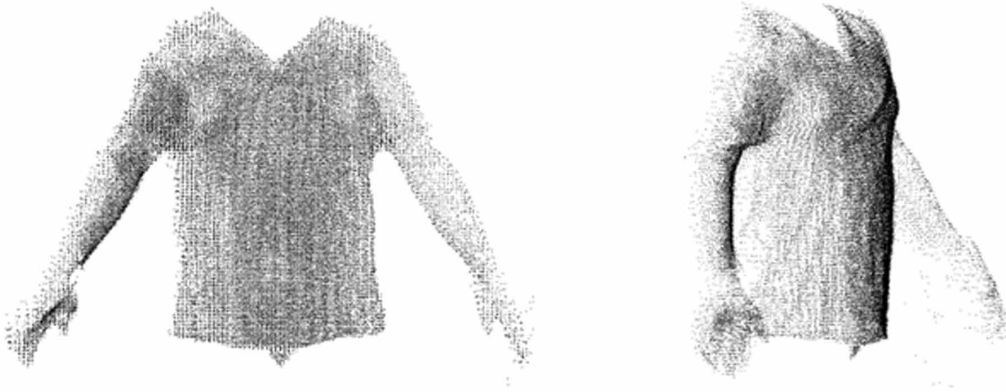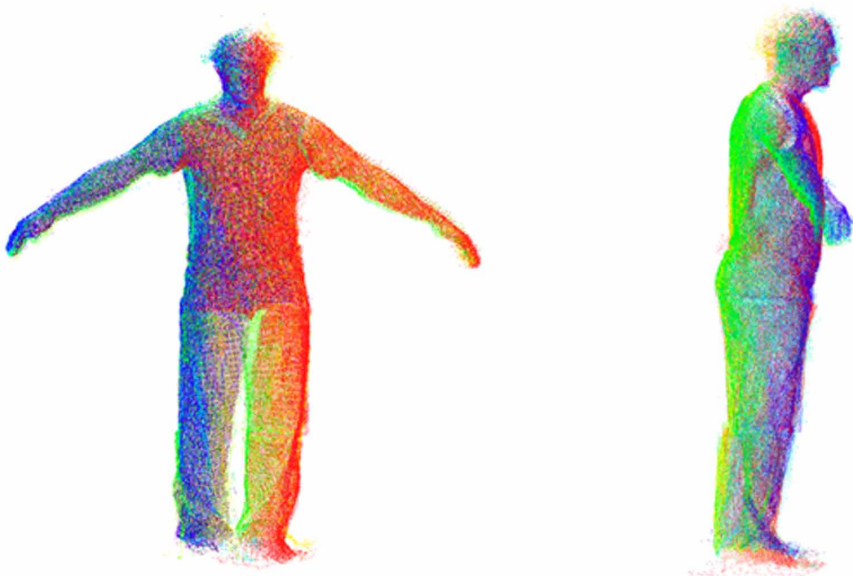**Figure 10. Result of the articulated alignment**

Figure 11. 3D model obtained with the proposed iterative method

4. CONCLUSION

In this paper, an iterative method for 3D body registration is presented aiming to be low-cost, dealing with unconstrained movements and accuracy. In line of these objectives, RGB-D sensors have been selected for the acquisition of 3D data to fit the low-cost requirement. The acquisition of this data is performed while the subject rotates in front of the camera, without the need of any external device or forcing the subject to be in a specific pose. The alignment of the different views is performed in a model-free approach in order to be more consistent with the real data.

The iterative method for 3D registration of the human body is based on an initial estimation of the skeleton. The method is divided in three stages, contributing to each other by the refinement of a specific part of the data. Moreover, the method feeds back in each iteration to improve the final result.

The experimentation with real data allows to validate the proposed method. It has been shown how it can progressively improve the acquired data to achieve the more accurate registration. The articulated alignment has been tested performing free arm movements, proving that the method can take into account those movements aligning all views into a specific pose. Finally, the main purpose of the method has been carried out performing the 3D registration of a body rotating in front of the camera satisfying the requirements of low-cost, unconstrained movements and accuracy.

As future lines, it is considered to introduce non-rigid registration methods to study the shape variation of subjects by obtaining 3D models in different instants of time in order to quantify that variation.

ACKNOWLEDGMENT

REFERENCES

Anguelov, D., Srinivasan, P., & Koller, D. (2005). Scape: shape completion and animation of people. *ACM Transactions on Graphics, 24*(3), 408-416.

Barmpoutis, A. (2013). Tensor body: Real-time reconstruction of the human body and avatar synthesis from RGB-D. *IEEE Trans. on Cybernetics, 43*(5), 1347-1356. doi:10.1109/TCYB.2013.2276430

Bogo, F., Black, M. J., Loper, M., & Romero, J. (2015). Detailed Full-Body Reconstructions of Moving People from Monocular RGB-D Sequences. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2300-2308). doi:10.1109/ICCV.2015.265

Charles, J., & Everingham, M. (2011). *Learning shape models for monocular human pose estimation from the Microsoft Xbox Kinect. In Proceedings of the 2011 IEEE Int. Conf. on Computer Vision Workshops* (pp. 1202–1208). IEEE. doi:10.1109/ICCVW.2011.6130387

Chen, W., Wang, H., Li, Y., Su, H., Tu, C., Lischinsk, D., Chen, B. (2016). Synthesizing Training Images for Boosting Human 3D Pose Estimation. doi:10.1109/3DV.2016.58

Cui, Y., Chang, W., Nöll, T., & Stricker, D. (2013). Kinect Avatar: Fully automatic body capture using a single Kinect. In Asian Conference on Computer Vision, LNCS (Vol. 7729, pp. 133-147). doi:10.1007/978-3-642-37484-5-12

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381–395. doi:10.1145/358669.358692

Henry, P., Krainin, M., Herbst, E., Ren, X., & Fox, D. (2014). RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments. In *Experimental Robotics (pp. 477-491)*. Springer. doi:10.1007/978-3-642-28572-1

Lai, K., Bo, L., Ren, X., & Fox, D. (2013). Consumer Depth Cameras for Computer Vision. In Consumer depth cameras for computer vision (p. 167). doi:10.1007/978-1-4471-4640-7

Lin, S., Chen, Y., Lai, Y. K., Martin, R. R., & Cheng, Z. Q. (2016). Fast capture of textured full-body avatar with RGB-D cameras. *The Visual Computer, 32*(6-8), 681–691. doi:10.1007/s00371-016-1245-9

Lovato, C., Bissolo, E., Lanza, N., Stella, A., & Giachetti, A. (2014). A low cost and easy to use setup for foot scanning.

Mihalyi, R.-G., Pathak, K., Vaskevicius, N., Fromm, T., & Birk, A. (2015). Robust 3D object modeling with a low-cost RGBD-sensor and AR-markers for applications with untrained end-users. *Robotics and Autonomous Systems, 66*, 1–17. doi:10.1016/j.robot.2015.01.005

Oliveira, G. L., Valada, A., Bollen, C., Burgard, W., & Brox, T. (2016, May). Deep learning for human part discovery in images. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1634-1641). IEEE. doi:10.1109/ICRA.2016.7487304

Pellegrini, S., Schindler, K., & Nardi, D. (2008). A Generalisation of the ICP Algorithm for Articulated Bodies. In *Proc. of the BMVC*. doi:10.1.1.165.4314 doi:10.5244/C.22.87

Rusinkiewicz, S., & Levoy, M. (2001). Efficient variants of the ICP algorithm. In *Proceedings third international conference on 3-d digital imaging and modeling* (pp. 145-152). IEEE. doi:10.1109/IM.2001.924423

Salvi, J., Matabosch, C., Fofi, D., & Forest, J. (2007). A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing, 25*(5), 578–596. doi:10.1016/j.imavis.2006.05.012

Saval-Calvo, M., Azorin-Lopez, J., Fuster-Guillo, A., Garcia-Rodriguez, J., Orts-Escolano, S., & Garcia-Garcia, A. (2017). Evaluation of sampling method effects in 3d non-rigid registration. *Neural Computing & Applications, 28*(5), 953–967. doi:10.1007/s00521-016-2258-z

Saval-Calvo, M., Azorin-Lopez, J., Fuster-Guillo, A., & Mora-Mora, H. (2015). μ-mar: Multiplane 3d marker based registration for depth-sensing cameras. *Expert Systems with Applications, 42*(23), 9353–9365. doi:10.1016/j.eswa.2015.08.011

Schwarz, L. A., Mkhitaryan, A., Mateus, D., & Navab, N. (2012). Human skeleton tracking from depth data using geodesic distances and optical flow. *Image and Vision Computing*, *30*(3), 217–226. doi:10.1016/j.imavis.2011.12.001

Shapiro, A., Feng, A., Wang, R., Li, H., Bolas, M., Medioni, G., & Suma, E. (2014). Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds*, *25*(3-4), 201–211. doi:10.1002/cav.1579

Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., & Moore, R. (2013, January). Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, *56*(1), 116. doi:10.1145/2398356.2398381

Tong, J., Zhou, J., Liu, L., Pan, Z., & Yan, H. (2012). Scanning 3D full human bodies using Kinects. *IEEE Transactions on Visualization and Computer Graphics*, *18*(4), 643–650. doi:10.1109/TVCG.2012.56 PMID:22402692

Treleaven, P., & Wells, J. (2007, July). 3D Body Scanning and Healthcare Applications. *Computer*, *40*(7), 28–34. doi:10.1109/MC.2007.225

Wang, R., Choi, J., & Medioni, G. (2012). Accurate full body scanning from a single fixed 3D camera. In *Proceedings of the 2nd Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission, 3DIMPVT '12* (pp. 432-439). doi:10.1109/3DIMPVT.2012.57

Wei, S.-E., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016). Convolutional Pose Machines. In *Proc. of CVPR* (pp. 4724-4732). doi:10.1109/CVPR.2016.511

Weiss, A., Hirshberg, D., & Black, M. (2011). Home 3D body scans from noisy image and range data. In Proceedings of *Computer Vision '11* (pp. 1951–1958). doi:10.1109/ICCV.2011.6126465

Victor Villena-Martinez is a PhD Student at the University of Alicante. He received his Master's Degree in Automation and Robotics in June 2016 and his Bachelor's Degree in Computer engineering in June 2015. He has collaborated in the project "Acquisition and modeling of growing plants" (GV/2013/005). His main research is focused on the calibration of RGB-D devices and the reconstruction of the human body using the same devices.

Andres Fuster-Guillo received the BS degree in Computer Science Engineering from Polytechnic University of Valencia (Spain) in 1995 and the PhD degree in Computer Science at the University of Alicante (Spain) in 2003. Since 1997, he has been a member of the faculty of the Department of "Computer Science Technology and Computation" at the University of Alicante, where he is currently a professor. He was Deputy Coordinator of the Polytechnic School at the University of Alicante for seven years and Director of the Secretariat for Information Technology at the University of Alicante for four years. During this period, he has coordinated and participated in several strategic technology projects: Open University (transparency portal and open data), UACloud, Smart University, among others.

Marcelo Saval-Calvo obtained a PhD in Computer Science in 2015, as well as Computer engineer at University of Alicante in 2010 and Master of Science in 2011. His interests include computer vision problem in human behaviour understanding, 3D shapes changes analysis and modelling, and 3D sensing and mapping. Moreover, he has expertise in CUDA programming for GPGPUs. Marcelo has published various paper in these areas and has attended international conferences and specialist courses. He worked as a senior researcher in a National Funded project in Advanced Driver Assistance Systems accelerating computer vision algorithms using heterogeneous platforms. He is a postdoc visitor in the School of Informatics, University of Edinburgh, in the computer vision group led by Prof. Robert B. Fisher.

Jorge Azorin-Lopez received a degree in Computer Engineering in 2001 and a PhD degree in Computer Science at the University of Alicante (Spain) in 2007. Since 2001, he has been a faculty member of the Department of Computer Science Technology and Computation at the same university, where he is currently an Associate Professor and the Deputy Director of Research. He was awarded the Post-Doctoral Research Fellowship "Automatic visual inspection of shape defects on specular surfaces. Methods to compensate low sensitivity of 3D image acquisition and reconstruction techniques" by the Spanish Ministry of Science and Education for research at University of Edinburgh. He has worked in 14 research projects and has published more than 40 papers on computer vision and computer architecture in several journals, conferences and book chapters. He has served as a reviewer to numerous scientific journals and international conferences.