

3D technologies to acquire and visualize the human body for improving dietetic treatment †

Andrés Fuster-Guilló ^{1,*}, Jorge Azorín-López ², Juan Miguel Castillo Zaragoza ³, Luis Fernando Pérez Pérez ⁴, Marcelo Saval-Calvo ⁵ and Robert B Fisher ⁶

¹ University of Alicante; fuster@ua.es

* Correspondence: fuster@ua.es; Tel.: +34-965903400

² University of Alicante; jazorin@ua.es

³ University of Alicante; jmcastillo1@dtic.ua.es

⁴ University of Alicante; lfpp1@alu.ua.es

⁵ University of Alicante; m.saval@ua.es

⁶ University of Edinburgh; rbf@inf.ed.ac.uk

† Presented at the 13th International Conference on Ubiquitous Computing and Ambient

Abstract: This research aims to improve the adherence to dietetic-nutritional treatment by using state-of-the-art technology of RGB-D sensors and virtual reality (VR). Recent studies show that adherence to treatment can be improved by using multimedia technologies which impact on the body awareness of patients. However, there are no studies published to date using 3D data and VR technologies for this purpose. This paper describes a system capable of obtaining the complete 3D model of a body with high accuracy and a realistic visualization of it for 2D and VR devices in order to be used for studying the effect of 3D technologies on the adherence to obesity treatment.

Keywords: obesity, adherence, 3D vision, RGBD sensors, virtual reality

1. Introduction

The prevalence of overweight and obesity has increased globally, tripling over the last three decades in the countries of the European Union. Overweight and obesity contribute to the emergence of chronic diseases (hypertension, type II diabetes, cancer, etc.) and the development of neurodegenerative pathologies (Alzheimer's or dementias) [1], [2]. The treatment and follow-up care required by these patients has a high impact on the costs of health services [3]–[5]. Adherence to the treatment of obesity has been considered one of the factors causing the failure of intervention programs [6]–[8]. Given this evidence, improving adherence would contribute to the outcome of treatments and their maintenance over time, leading to lower health costs.

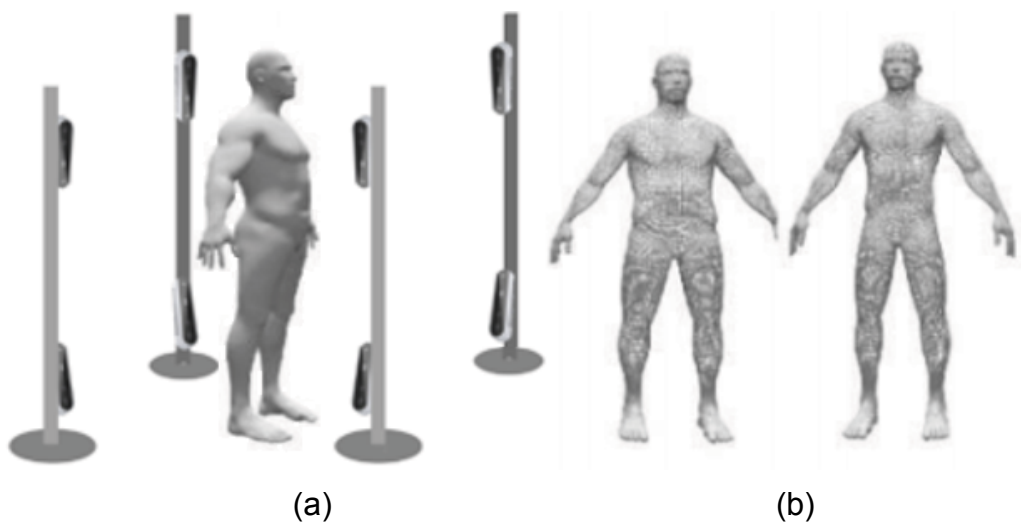
Some authors have suggested that nutritional interventions that reinforce the follow-up of therapies through the use of technologies achieve beneficial effects over time [8], [9], [10]. The results show how adherence to treatment can be increased by incorporating techniques based on the use of (2D) images of the patient's evolution in dietary treatment that enhance their cognitive experience [11]. However, to the best of our knowledge there are no studies that take advantage of and quantify the potential of using realistic 3D images and virtual reality techniques to reinforce adherence to treatment. This work provides the necessary framework for specialists to address this type of studies.

Nowadays, 3D modelling of the human body is transforming our ability to accurately measure and visualize it, showing great potential for health applications from epidemiology to diagnosis or patient monitoring [12], [13]. Recent works [14], [15] deal with the acquisition of human body from RGB-D cameras and video sequences providing 3D models with texture and avatars, but they do not provide

the accuracy necessary in health applications. There are systems on the market for body model acquisition: Fit3D [16] uses a single camera but needs a complex and expensive device; Body labs acquired by Amazon also developed technology to capture in 3 dimensions the human body, but it does not consider any time parameter; Naked labs [17] provides body visualization and tracking functions for the fitness market.

There are different types of 3D sensors with different characteristics that can be used to capture the human body. Devices based mainly on laser, such as Lidar, have high accuracy but they only provide depth information and do not provide color data. Stereo sensors use two color cameras to infer the depth, which usually means high cost and difficult portability since both cameras must be calibrated. Recently, RGB-D cameras (such as Microsoft Kinect or Intel RealSense) integrate color and depth in a single device, and they use different technologies to estimate the depth (structured light, ToF, active stereo). The characteristics of these RGB-D devices including accuracy, portability, capture frequency, etc., are causing their popularization and integration in mobile consumer devices [18]. For these reasons, this work makes use of these devices to capture the 3D model. Moreover, the use of RGB-D device networks is proposed to meet the required quality levels for 3D representations.

Virtual reality has also experienced an important growth in recent years. This technology simulates realistic 3D interactive environments using HMDs (Head Mounted Displays), also known as virtual reality glasses. At first, these devices had high prices due to their complexity. However, recent developments such as "Google Cardboard" allowed any smartphone to be converted into a HMD with very low costs. Currently, there are many works that use the potential of virtual reality. The immersive experience provided by virtual reality stands out, improving concentration in the training process [19]. It is hoped that the use of these technologies can be effective in improving adherence to the treatment of overweight patients [10].



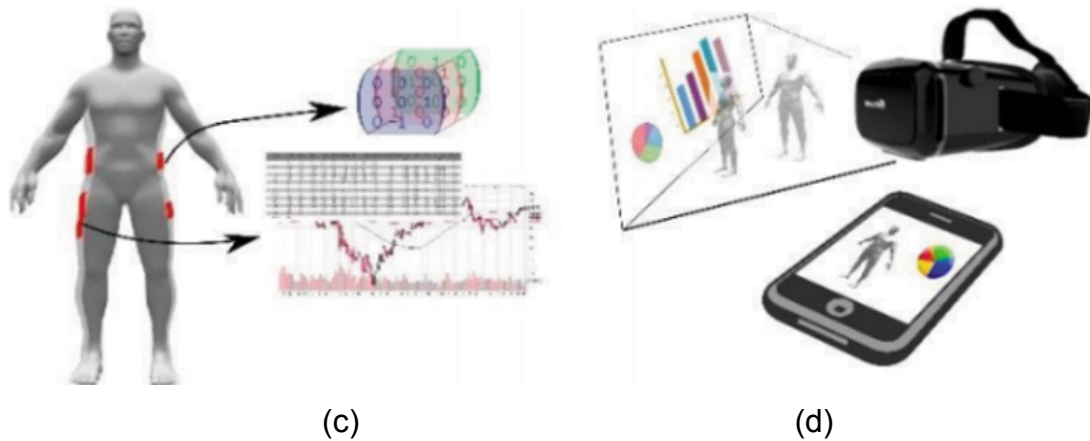


Figure 1 Acquisition system based on RGB-D camera network (a). Body models: point clouds and mesh (b). Shape evolution analysis (c) and textured representation with VR devices (d).

Classical treatments for obesity patients have shown limited effectiveness in resolving chronicity. The use of technologies for 3D reconstruction of the human body is sufficiently mature in different fields of application. The field of virtual reality is also evolving remarkably, finding low-cost systems and successful application experiences in different fields. However, to the best of our knowledge, there are no works that combine RGB-D acquisition devices and virtual reality, analyzing shape changes over time (4D) to improve adherence to obesity treatments.

There are important scientific challenges on the development of computational methods for the study of changes in the shape of the human body using 3D/4D vision techniques to improve obesity treatment processes. In this paper, we provide a system capable of obtaining the 3D model of the body with color texture representation and a realistic visualization for 2D and virtual reality devices (see Figure 1) in order to be used, in the future, for studying the effect of 3D technologies on the adherence to obesity treatment.

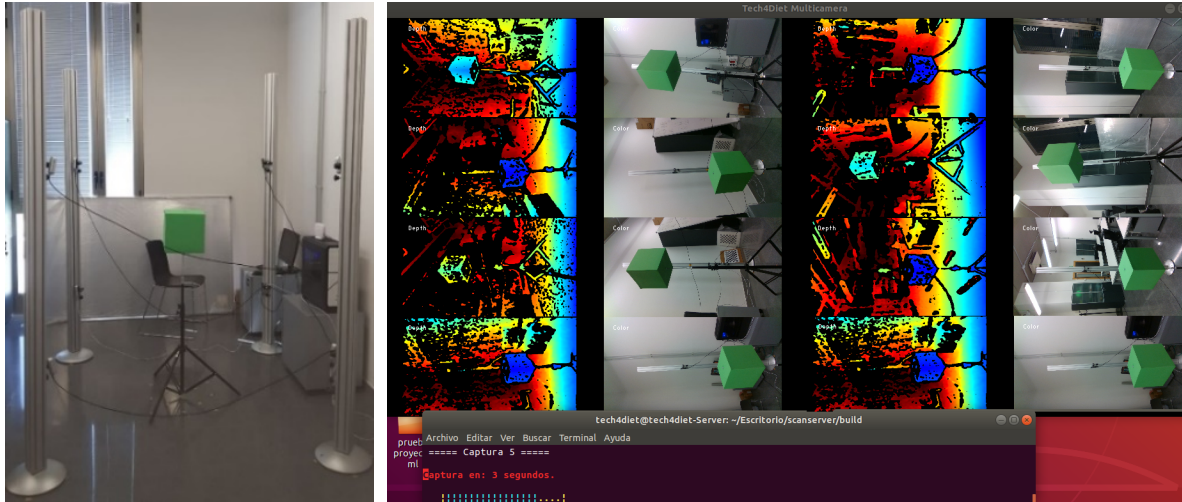
For the development of this research the following specific objectives have been addressed:

- **Obtaining the 3D/4D model:** 3D acquisition of the human body from low cost RGB-D camera network, obtaining the 3D geometric model and the texture representation of sequences of bodies over time (4D).
- **Visualization of the 3D body:** From the 3D models captured over time, realistic visualizations of the body evolution are generated using virtual reality.

The rest of the paper is organized as follows: Section 2 describes the system for reconstructing the human body from RGB-D cameras, Section 3 details the system of visualization of the human body and, finally, Section 4 presents the conclusions and summarizes the nutritional intervention plans and their impact and expected results to be obtained using the system.

2. 3D reconstruction of human body from multiple RGB-D views

In order to obtain the 3D model of the human body, a network of RGB-D cameras has been used to meet the quality requirements of the system. The network is composed of 8 Intel RealSense RGB-D cameras located on 4 aluminum masts of 2200 x 80 x 80 mm distributed around the capture area. Figure 2 shows the setup to capture a human body (left) and an example of 8 depth and RGB color images obtained by the system (right).



(a)

(b)

Figure 2. (a) Experimental set-up 8 RGB-D cameras. (b) Acquisition results for the 8 cameras (color and depth images). The green cube has been used for extrinsic calibration.

The pipeline used to obtain the 3D textured model from different RGB-D sensors has five stages (Figure 3) from (a) to (e): acquisition, pre-processing, registration, mesh generation and texture projection. Calibration of the individual sensors and estimation of the relative sensor positions is explained but not included in the pipeline as it is part of the set-up process.

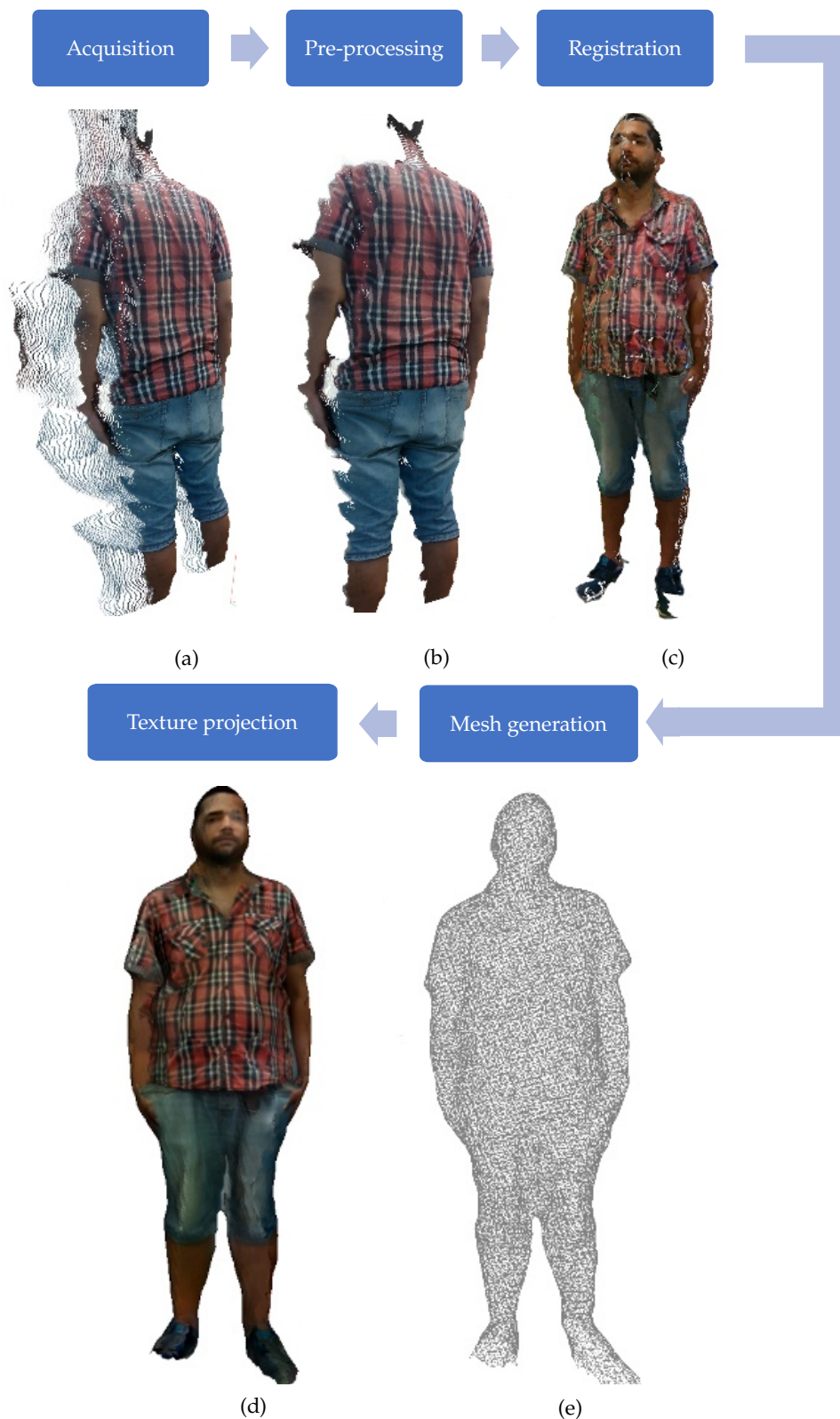


Figure 3. Pipeline of 3D body reconstruction. The system is able to acquire several images from cameras (a) that are preprocessed in order to improve the quality of the acquisition (b). The set of points are registered in a unique origin of coordinates (c). Finally, in order to obtain the 3D model of the body, the 3D points are converted into a mesh (e) and the images are projected on it (d)

2.1 Calibration

To correct the distortions of the images caused by the lens, an intrinsic calibration is carried out using the provided Intel RealSense SDK. The calibration method requires a minimum of six samples of the provided chessboard marker from different points distributed in the capture area.

Since we are using a network of RGB-D cameras, it is necessary to carry out an extrinsic calibration to unify the views locating the different point clouds in the same coordinate space. To obtain the transformation matrices of each of the point clouds we carried out an extrinsic calibration based on 3D markers, specifically we have used two types of 3D markers, spherical and cubic [20].

2.2 Acquisition

The network is composed by 8 Intel RealSense RGB-D cameras D435 with the appropriate characteristics (FOV, color and depth resolution...). Intel's SDK for RealSense has been used as the basis for the development of the acquisition software. Once the system has been calibrated, the acquisition process requires the synchronization of all the cameras in the network to perform the capture, so that the captured data from the 8 cameras all come from the same time. Semaphore management has been used to address the synchronization. The semaphores act in a similar way to a barrier and assure that all threads begin the acquisition practically at the same time, and they do not finish the acquisition of the frame until the rest of the cameras have finished.

2.3 Preprocessing

At this stage, some noisy point clouds from the different RGB-D sensors have been obtained, so it is necessary to apply different methods to improve their quality (see Figure 3-b). First, the point cloud is truncated in the z-axis (depth) to remove the points that are beyond the center of the capture area. After that, three filters are applied: median, bilateral and statistical outlier removal (SOR). The median filter is able to reduce noise and it is very efficient in terms of processing time, as it requires a single pass over the cloud [21]. The bilateral filter smooths edges and areas with high curvatures while preserving sharpness by a non-linear combination of values from nearby areas [22]. The SOR filter removes edge noise and outliers using neighborhood statistics [23]. Finally, the normal vector for each point in the cloud is calculated. The problem of determining the normal at a point on the surface is approached by the problem of estimating the normal of a plane tangent to the surface [24].

2.4 Registration

In order to align the different point clouds in a single 3D coordinate system, the transformation matrices T obtained from the extrinsic calibration are applied to the data extracted from each camera (see Figure 3-c). Each camera has its own reference system with the origin (0,0,0) located in itself. With the extrinsic parameters from the calibration we assume one camera as reference and transform the rest of point clouds to this one and obtain a unified dataset [25].

2.5 Mesh generation

Different methods such as Greedy Projection or Marching Cubes were tested, obtaining the best result with the Poisson surfacing algorithm [26]. It is able to reconstruct a triangle mesh from a set of oriented 3D points by solving a Poisson system. This is, solving a 3D Laplacian system with positional value constraints.

The method approaches the problem of surface reconstruction using an implicit function framework. Computes a 3D indicator function that returns as greater than 0 the points inside the surface and as less than 0 the points outside it. This function can be found because there is a relationship between the orientation of the points and the function itself. Specifically, the gradient of the function to be found is a vector field with a value of 0 in all points except those close to the surface which takes the value of the surface normal oriented towards the interior (see Figure 3-e). After that, the reconstructed surface is obtained by extracting an appropriate isosurface.



Figure 4. Different views of the same textured body model.

2.6 Texture projection

The method proposed by Callieri et al. [27] has been used to carry out the raster projection and texture generation. This method generates parameters of the mesh in relation to its vertices and generates the texture based on the projection of the different images considering the position and orientation of the different cameras. Figure 4 shows the same body represented by the mesh with the projected texture seen from 4 different points of view.

3. 4D visualization of the human body using virtual reality for obesity treatment improvement

The second objective of this work is to provide a visualization system of the generated 3D models. This system allows interaction with the acquisition subsystem, management of patient data and a realistic visualization of the human body models over time. This system will be used both by the medical specialist to assist research in the field of obesity treatment, and by the patient to improve their adherence to treatment. Thus, the system is composed by two subsystems: the specialist 4D visualization system for obesity treatment and the virtual reality visualization system.

3.1 Specialist 4D image visualization system for obesity treatment

The visualization system allows the medical specialist several options. From the system, they can perform a scan of the patient's body since the visualization system communicates with the acquisition system.

In addition, the system provides the possibility to connect the virtual reality system to the specialist so the patient can see himself with the VR glasses while the expert is acting with the virtual model. On the other hand, the system can visualize all the body models generated allowing to change the point of view, compare body models and navigate through information from different sessions (*Figure 5*). Finally, the system allows the specialist to register new patients or view the history of sessions.



Figure 5. 4D visualization of a sequence of 3D models.

3.2 Virtual reality system

The second subsystem is composed by a virtual reality (VR) system contained in a mobile application that allows patients to see their progress in a more immersive way with the goal to improve the adherence to treatments. The VR system, as well as the visualization system, has been developed using Unity [28].

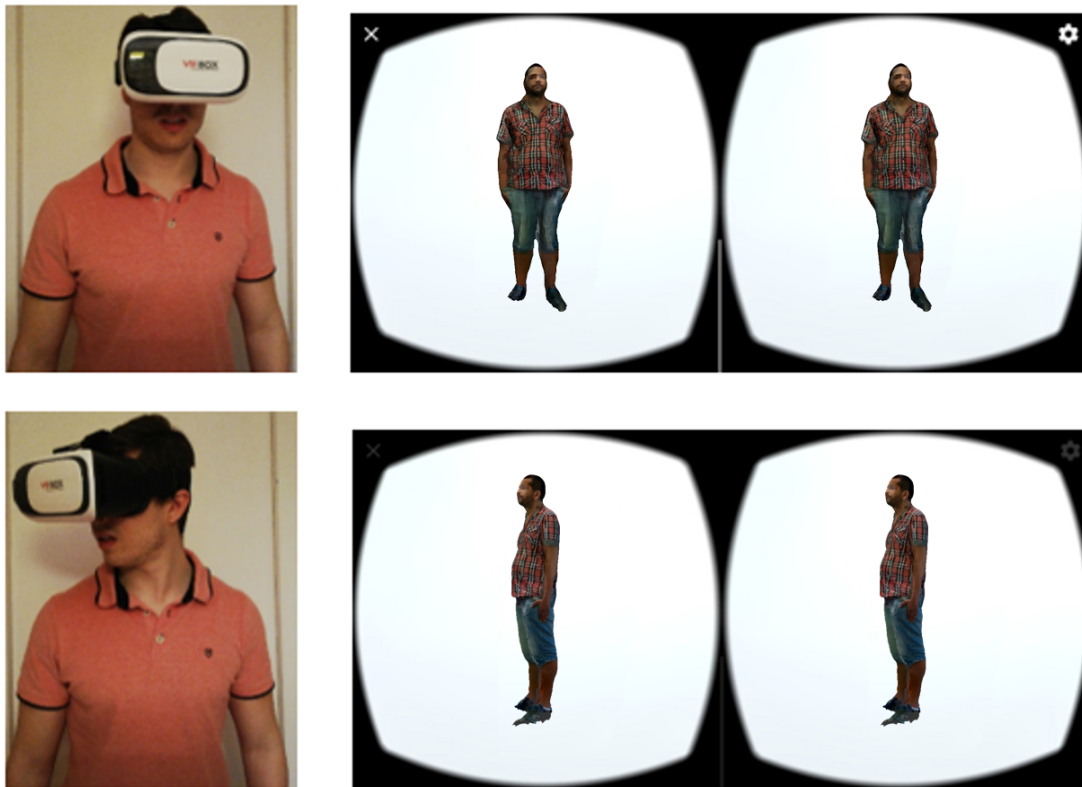


Figure 6 Transmission of motion to the 3D model (left) head positions (right) stereo images

Two main functionalities have been developed for this subsystem. First, the VR system is synchronized in real time with the computer specialist visualization system allowing the interaction between both systems. Second, the movement of the user's head is transmitted to the rotation of the 3D model being visualized (Figure 6).

4. Conclusions

This work focuses its contribution at developing technologies for health. Specifically, 3D cameras (RGB-D) and virtual reality technologies are intended to improve obesity treatment. Future research will investigate the positive effect of realistic 3D representations on the body awareness and psychological well-being of the individual. The proposed system has been developed to give the medical specialists a tool to study problems related to adherence to dietetic-nutritional treatment. The proposed methodology allows to obtain a 3D visual model of human body over time to analyze the morphological progress due to changes by obesity treatments. 3D models are obtained from the acquisition of multiple views through a network of RGB-D cameras. These views are filtered and aligned to obtain a mesh model on which the texture is projected generating realistic 3D models. Sequences of realistic 3D models allow the generation of the 4D models used in the visualizations, which we hope will provide a powerful method to improve adherence to obesity treatments. Different scientific challenges in the area of computer vision have been addressed, from the problem of calibrating RGB-D camera networks, improving the quality of point clouds as well as the realistic 3D representation. Furthermore, from a technological point of view, a system based on low-cost, adaptable and portable technologies, and intuitive environments based on realistic 3D representations is designed. The potential of these technologies,

together with the purpose of improving the treatment of obesity, allow predicting a high social and economic impact. The system is planned to be installed in three Primary Health Care Centers in order to obtain and study the psychological results. Subsequently, it is planned to transfer the operation to health institutions in the area of Alicante (Spain).

Supplementary Materials: The webpage shows a video of the system capable of capture the body and the visualization subsystems: <http://tech4d.dtic.ua.es/media/>

Author Contributions: Conceptualization and Methodology, Andrés Fuster-Guilló. Jorge Azorín-López; Software, Juan Miguel Castillo Zaragoza. Luis Fernando Pérez Pérez; Validation, Marcelo Saval-Calvo; Writing-Review, Andrés Fuster-Guilló. Jorge Azorín-López; Supervision, Robert B Fisher.

Funding: This work has been partially funded by the Spanish Government TIN2017-89069-R grant supported with Feder funds

References

- [1] C.-J. Boraxbekk *et al.*, "Diet-Induced Weight Loss Alters Functional Brain Responses during an Episodic Memory Task," *Obes. Facts*, 2015.
- [2] J. Drigny *et al.*, "Effect of interval training on cognitive functioning and cerebral oxygenation in obese patients: A pilot study," *J. Rehabil. Med.*, 2014.
- [3] T. Lehnert, D. Sonntag, A. Konnopka, S. Riedel-Heller, and H.-H. König, "Economic costs of overweight and obesity," *Best Pract. Res. Clin. Endocrinol. Metab.*, 2013.
- [4] D. Withrow and D. A. Alter, "The economic burden of obesity worldwide: A systematic review of the direct costs of obesity," *Obesity Reviews*. 2011.
- [5] A. Sicras-Mainar, J. Gil, T. Mora, and J. Ayma, "Prevalencia e impacto económico de la obesidad en adultos durante el periodo 2003-2010," *Medicina Clinica*. 2012.
- [6] S. De Geest and E. Sabaté, "Adherence to long-term therapies: Evidence for action," *European Journal of Cardiovascular Nursing*. 2003.
- [7] A. Boeka and K. Lokken, "Neuropsychological performance of a clinical sample of extremely obese individuals," *Arch. Clin. Neuropsychol.*, 2008.
- [8] M. G. Roseman, M. C. Riddell, and J. N. Haynes, "A Content Analysis of Kindergarten-12th Grade School-based Nutrition Interventions: Taking Advantage of Past Learning," *J. Nutr. Educ. Behav.*, 2011.
- [9] W. N. Ajie and K. M. Chapman-Novakofski, "Impact of computer-mediated, obesity-related nutrition education interventions for adolescents: A systematic review," *J. Adolesc. Heal.*, 2014.
- [10] M. Ferrer-Garcia, J. Gutiérrez-Maldonado, and G. Riva, "Virtual Reality Based Treatments in Eating Disorders and Obesity: A Review," *J. Contemp. Psychother.*, vol. 43, no. 4, pp. 207–221, Dec. 2013.
- [11] I. Kuzmar, M. Rizo, and E. Cortés-Castell, "Adherence to an overweight and obesity treatment: how to motivate a patient?," *PeerJ*, 2014.
- [12] P. Treleaven and J. Wells, "3D body scanning and healthcare applications," *Computer (Long. Beach. Calif.)*, 2007.
- [13] J.-D. Lin, W.-K. Chiou, H.-F. Weng, Y.-H. Tsai, and T.-H. Liu, "Comparison of three-dimensional anthropometric body surface scanning to waist-hip ratio and body mass index in correlation with metabolic risk factors," *J. Clin. Epidemiol.*, 2002.
- [14] T. Alldieck, M. A. Magnor, W. Xu, C. Theobalt, and G. Pons-Moll, "Detailed Human Avatars from Monocular Video," *undefined*, 2018.
- [15] T. Yu *et al.*, "DoubleFusion: Real-time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor," *undefined*, 2018.
- [16] "Fit 3D Body Scanners." [Online]. Available: <https://fit3d.com/>. [Accessed: 14-Jul-2019].

- [17] "Naked – The World's First Home Body Scanner." [Online]. Available: <https://nakedlabs.com/>. [Accessed: 14-Jul-2019].
- [18] V. Villena-Martínez *et al.*, "A quantitative comparison of calibration methods for RGB-D sensors using different technologies," *Sensors (Switzerland)*, vol. 17, no. 2, 2017.
- [19] M. Hussein and C. Nätterdal, "The benefits of virtual reality in education-A comparison study," no. June, 2015.
- [20] M. Saval-Calvo, J. Azorin-Lopez, A. Fuster-Guillo, and H. Mora-Mora, "μ-MAR: Multiplane 3D Marker based Registration for depth-sensing cameras," *Expert Syst. Appl.*, vol. 42, no. 23, pp. 9353–9365, 2015.
- [21] PCL Team, "Point Cloud Library (PCL): pcl::MedianFilter< PointT > Class Template Reference," 2013. [Online]. Available: http://docs.pointclouds.org/1.7.1/classpcl_1_1_median_filter.html. [Accessed: 27-May-2019].
- [22] PCL Team, "Point Cloud Library (PCL): pcl::BilateralFilter< PointT > Class Template Reference," 2019. [Online]. Available: http://docs.pointclouds.org/trunk/classpcl_1_1_bilateral_filter.html. [Accessed: 27-May-2019].
- [23] PCL Team, "Point Cloud Library (PCL): pcl::StatisticalOutlierRemoval< PointT > Class Template Reference," 2013. [Online]. Available: http://docs.pointclouds.org/1.7.1/classpcl_1_1_statistical_outlier_removal.html. [Accessed: 27-May-2019].
- [24] Radu Bogdan Rusu, "Documentation - Point Cloud Library (PCL)." [Online]. Available: http://pointclouds.org/documentation/tutorials/normal_estimation.php. [Accessed: 19-May-2019].
- [25] M. Saval-Calvo, J. Azorin-Lopez, and A. Fuster-Guillo, "Model-Based Multi-view Registration for RGB-D Sensors," in *Advances in Computational Intelligence*, vol. 7903, I. Rojas, G. Joya, and J. Cabestany, Eds. Springer Berlin Heidelberg, 2013, pp. 496–503.
- [26] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson Surface Reconstruction," 2006.
- [27] M. Callieri, P. Cignoni, M. Corsini, and R. Scopigno, "Masked photo blending: Mapping dense photographic data set on high-resolution sampled 3D models," *Comput. Graph.*, vol. 32, no. 4, pp. 464–473, Aug. 2008.
- [28] Unity, "Unity User Manual (2018.3) - Unity Manual," 2018. .



© 2018 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).