

Evaluation of Positioning Techniques for mm-Wave Portable Scanners

Jaime Laviada¹, Miguel López-Portugués¹, Ana Arboleya², Fernando Las-Heras¹

¹Departamento de Ingeniería Eléctrica, Universidad de Oviedo, Gijón (Asturias), Spain, jlaviada@tsc.uniovi.es

²Electronics for Connected Objects, Université Nice Sophia Antipolis, Nice, France, ana.arboleya@unice.fr

Abstract—Hybrid methods that combine electromagnetic and optical 3D imaging can be used to scan objects or bodies at mm-wave frequencies by merging multiview information. The first goal of the optical model is to obtain a 3D representation of the target under test. The second goal lies in the estimation of positions and attitude angles of the electromagnetic scanner in order to properly combine the data collected at each acquisition view. In this contribution, two different positioning techniques are evaluated. The first one is based on a conventional optical camera, whereas the second one relies on a *depth camera*. Finally, an example of application to security screening is used to analyze the accuracy of both methods.

Index Terms—millimeter-wave imaging, RGB-D, depth camera, structured light, portable scanner, photogrammetry.

I. INTRODUCTION

In the last years, micro- and mm-wave camera and portable systems have been developed [1], [2], [3]. In addition, obtaining a real-time 3D model of a given target may now be feasible by means of using portable optical cameras — such as those attached to smartphones [4]— or novel low-cost *depth cameras*. In this manner, hybrid methods that combine two different but complementary 3D imaging methods are emerging.

In [5], the authors proposed the *hybridization* of a mm-wave camera and an optical camera to obtain a portable 3D scanner, in which a Synthetic Aperture Radar (SAR) is used to scan a certain object or body at mm-wave frequencies by merging multiview information. The information is obtained from multiple angles in order to achieve a complete model scan. In addition, both position and attitude of the scanner at each acquisition view are estimated by taking pictures (using an optical camera) at the same positions as the mm-wave acquisitions and then processing them with Structure from Motion techniques [6]. Once the scanner rotation angles and positions are found, the appropriate multiview merge algorithms can be applied.

On the one hand, positioning techniques based on optical cameras —such as the scale-invariant feature transform (SIFT) used in [5]— usually require finding several common key points from different views in order to obtain the position by triangulation. Thus, the aforementioned techniques are prone to fail or lose accuracy when considering objects with flat textures or under insufficient lighting conditions.

On the other hand, the use of depth cameras [7], [8], also known as RGB-D cameras, may overcome the limitations of

optical cameras at the price of a more complex (and less widespread) hardware. Depth cameras are categorized into two types, those built upon *structured light* and those based on *time-of-flight*. The former project an infrared pattern on the object in order to estimate the shape of the object from the pattern deformation [9], whereas the latter relies on measuring the round-trip time that it takes for light to travel from the sensor to a certain scene point [10].

This contribution presents a comparison between the positioning technique based on a conventional optical camera proposed by the authors in [5] and a depth-camera based technique, both of them applied to the concept of mm-wave portable scanners. The goal is to study the advantages and drawbacks of each approach and analyze their accuracy.

II. POSITIONING TECHNIQUES FOR MM-WAVE SCANNERS

The concept of mm-wave portable scanner proposed by the authors in [5] is based on a small-sized aperture in order to facilitate the motion around the object under test for the purpose of performing a multiview scan. In addition, both mm-wave and optical (or depth) camera acquisition can be performed in parallel.

The acquired scattered field is processed by using standard range-migration techniques [11] in order to obtain the reflectivity on a regular grid. The estimation of the scanner poses and positions, and the generation of the optical 3D model, may be solved using two different approaches that are detailed below. Finally, the reflectivity and optical data are merged yielding a double 3D model: the first one associated with the external layer of the object and the second one related to inner layers.

A. Photogrammetry

In the photogrammetry-based technique, pictures are taken from the same positions from which the scattered field is acquired. In addition, it is advisable to take some in-between pictures to increase the overlap among them. The set of pictures is processed by using standard Structure from Motion techniques [6] as follows:

- 1) The relevant points of each image are identified and descriptors are computed by means of SIFT.
- 2) The descriptors between sequential images are compared in order to find matching points.

- 3) The global 3D position of the matching points is calculated by using triangulation, so that a sparse reconstruction of the 3D optical model is found. Camera poses and positions are also calculated at this step.

In those cases in which the object under test is composed of a flat texture, the number of keypoints may be insufficient, so the positioning system could fail. Furthermore, common optical cameras may be unable to operate properly with poor illumination conditions.

B. Depth camera

In this technique, the positioning does not rely on finding key points but on comparing *point clouds* (PCs) obtained with a depth camera. The RGB-D camera used in this work is based on structured light. Thus, in this particular case, an infrared projector illuminates an object with a set of predefined patterns. The patterns projected are warped by the object surface and then reflected back to the camera, where they are captured using an infrared sensor. Finally, the infrared pixels are processed to obtain a depth frame, which may be combined with the data gathered with a RGB sensor to generate a full RGB-D frame (i.e., a 3D color image).

In this work, the Iterative Closest Point algorithm (ICP) [12] implementation from the Point Cloud Library (PCL) [13] is used to obtain the scanner positions and generate the optical 3D model. The procedure may be divided in three steps:

- 1) *Acquisition* of a new PC.
- 2) *Registration* of the current PC by calculating the relative translation and rotation matrix from the previous PC (relative camera pose and position).
- 3) *Integration* of the new points into the model.

An initial model mesh is reconstructed from the PC that is registered at the first position. Then, the following registered PCs from the different positions are merged with the model, one by one, in order to obtain a complete optical 3D model. Merging is performed by searching for the nearest neighbors from the PC to the current model mesh. If the angle between the normals of a new and an old point is smaller than a predefined threshold, then both points are considered the same one. On the contrary, if the distance between a new and an old point is higher than a given threshold, that point is considered a new point and added to the model mesh. Finally, if a point exceeds a maximum *age* (time since a point has been recorded for the last time) without reaching a minimum *visibility confidence* (number of unique directions from which a point has been recorded) then it is removed from the mesh.

III. RESULTS

In order to analyze the accuracy of the above-mentioned positioning techniques applied to mm-wave portable scanners, the measurements presented in [5], in which a mannequin torso with an attached knife is considered, are revisited. The torso of the mannequin is wrapped in aluminium foil to provide a fair

Fig. 1. Detail view of the measurement setup comprising a mannequin with an attached knife (without the raincoat).

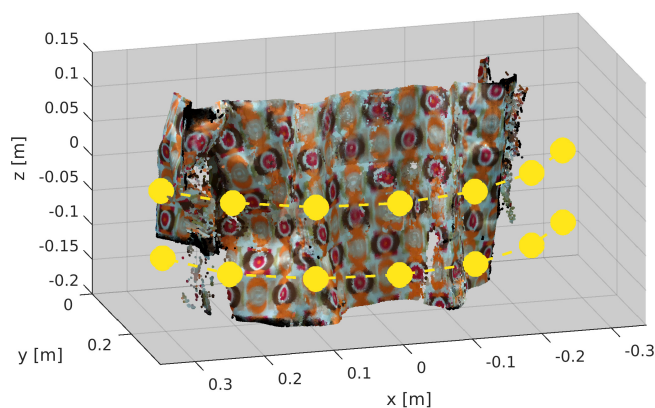


Fig. 2. Scanner positions superimposed on the depth-camera based 3D-model (left-to-right and top-to-bottom).

approximation of the human skin at mm-waves (see Fig. 1), and dressed with a raincoat to conceal the bladed weapon.

A setup which is equivalent to moving the scanner along 14 different positions is considered (see Fig. 2). These points of view correspond to two arcs ($\phi \in [120^\circ, 240^\circ]$, $\Delta\phi = 20^\circ$) with a radius of 35.23 cm that are separated 10 cm along the z -axis. In order to estimate the scanner positions, two different setups are analyzed:

- 1) A smartphone with a customized image acquisition software (photogrammetry).
- 2) An Intel® RealSense™ SR300 RGB-D camera that operates at VGA resolution and relies on structured light (depth camera).

When the optical camera is used, 3 intermediate positions are considered (i.e., rotation steps of 5°) in order to increase the overlap among pictures, which yields a total of 50 pictures. In addition, 3 extra pictures separated by 1 cm are also taken to estimate the scale factor of the model.

When using the depth camera, a bounding box of $30 \times 55 \times 20$ cm is introduced to filter out any external object that could deteriorate the positioning accuracy. In addition, 3 intermediate positions are also considered to increase the overlap between consecutive PCs. The minimum allowed overlap is set to 70% to ensure the proper functioning of the ICP algorithm.

The real positions of the scanner, as well as the positions estimated with photogrammetry and by using the depth camera, are shown in Table I. Additionally, in Table II, real scanner attitude angles (yaw, pitch, roll) and estimated angles are provided.

When the first arc is considered ($z = 0$ cm), the error between two consecutive positions using photogrammetry ranges from 4.2 mm to 6.4 mm with a mean value of 5.2 mm, which implies a relative error between 3.4% and 5.2% with a mean value of 4.2%. Under the same circumstances, the error among consecutive positions using the RGB-D camera ranges from 7.4 mm to 1.3 cm with a mean value of 9.6 mm, which supposes a relative error between 6% and 10.8% with a mean value of 7.8%. For the second arc ($z = -10$ cm), the positioning error between two consecutive acquisitions using photogrammetry ranges from 4.3 mm to 1.1 cm with a mean value of 6.8 mm, which supposes a relative error between 3.5% and 9.3% with a mean value of 5.5%. In the same arc, the error among consecutive positions using the RGB-D camera ranges from 7.4 mm to 1 cm with a mean value of 8.6 mm, which implies a relative error between 6% and 8.4% with a mean value of 7%. Furthermore, the accumulated error over a scanning path of 73.8 cm (from the beginning to the end of each arc) is 2.1 cm using photogrammetry and 2.3 cm using the RGB-D camera.

Finally, several steps for the cumulative reflectivity are shown in Fig. 3, in which the partial images are merged using the calculated scanner positions to yield a global mm-wave model where the presence of the knife is clearly noticeable. More details about the scattered field acquisition and the reflectivity calculation can be found in [5].

IV. CONCLUSIONS

In this contribution, the authors analyze two positioning techniques applicable to mm-wave portable scanners, one of them based on photogrammetry and another one relying on a RGB-D camera. On the one hand, optical cameras used in photogrammetry may be considered commodity hardware, so a portable positioning system based on a smartphone camera can be an inexpensive solution. Nevertheless, positioning based on optical cameras may fail (or lose accuracy) when scanning objects with flat textures or under poor lighting conditions. On the other hand, RGB-D cameras are texture independent and can operate even with bad illumination conditions.

As shown in the results section, if the conditions (i.e., textures and illumination) are appropriate, both positioning techniques achieve good accuracy, with relative errors under 11% between consecutive positions, yielding a resolution finer than 1.3 cm. As a consequence, the mm-wave images can be

merged accurately, so the portable scanner concept proposed by the authors may become a very useful tool to detect hidden threats.

ACKNOWLEDGMENT

This work has been supported by the “Ministerio de Ciencia e Innovación” of Spain/FEDER under projects TEC2014-55290-JIN; by the “Gobierno del Principado de Asturias” (PCTI)/FEDER-FSE under project GRUPIN14-114; and by “Ayudas Fundación BBVA a Investigadores y Creadores Culturales 2016”.

REFERENCES

- [1] M. T. Ghasr, M. A. Abou-Khousa, S. Kharkovsky, R. Zoughi, and D. Pommerenke, “Portable real-time microwave camera at 24 GHz,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 2, pp. 1114–1125, Feb 2012.
- [2] A. T. Mobashsher, A. Mahmoud, and A. M. Abbosh, “Portable wide-band microwave imaging system for intracranial hemorrhage detection using improved back-projection algorithm with model of effective head permittivity,” *Scientific Reports*, vol. 6, p. 20459, 2016.
- [3] M. T. Ghasr, M. J. Horst, M. R. Dvorsky, and R. Zoughi, “Wideband microwave camera for real-time 3-D imaging,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 1, pp. 258–268, Jan 2017.
- [4] P. Tanskanen, K. Kolev, L. Meier, F. Camposco, O. Saurer, and M. Pollefeys, “Live metric 3D reconstruction on mobile phones,” in *2013 IEEE International Conference on Computer Vision*, Dec 2013, pp. 65–72.
- [5] J. Laviada, A. Arboleya-Arboleya, Y. Álvarez, B. González-Valdés, and F. Las-Heras, “Multiview three-dimensional reconstruction by millimetre-wave portable camera,” *Scientific Reports*, vol. 7, p. 6479, 2017.
- [6] C. Wu, “Towards linear-time incremental structure from motion,” in *Proceedings of the 2013 International Conference on 3D Vision*, ser. 3DV '13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 127–134.
- [7] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, *RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 477–491.
- [8] K. Khoshelham and S. O. Elberink, “Accuracy and resolution of kinect depth data for indoor mapping applications,” *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.
- [9] R. A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissano, “Structured light using pseudorandom codes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 322–327, Mar 1998.
- [10] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, “High-quality scanning using time-of-flight depth superresolution,” in *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, June 2008, pp. 1–7.
- [11] J. M. Lopez-Sanchez and J. Fortuny-Guasch, “3-D radar imaging using range migration techniques,” *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 5, pp. 728–737, May 2000.
- [12] P. J. Besl and N. D. McKay, “A method for registration of 3-D shapes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb 1992.
- [13] R. B. Rusu and S. Cousins, “3D is here: Point Cloud Library (PCL),” in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 9-13 2011.

TABLE I
THEORETICAL SCANNER POSITIONS (x, y, z) AND ESTIMATED POSITIONS (\hat{x}_p FOR PHOTOGRAMMETRY AND \hat{x}_d FOR DEPTH CAMERA).

x [cm]	\hat{x}_p [cm]	\hat{x}_d [cm]	y [cm]	\hat{y}_p [cm]	\hat{y}_d [cm]	z [cm]	\hat{z}_p [cm]	\hat{z}_d [cm]
30.910	30.007	28.914	18.308	17.606	17.458	0	-0.367	0.548
22.833	22.193	21.452	27.502	26.562	26.180	0	-0.122	-0.069
12.098	11.775	11.393	33.379	32.216	31.765	0	0.039	-0.159
0	0.090	0.090	35.230	33.672	33.670	0	0.105	0.100
-12.001	-11.395	-10.979	32.832	31.204	31.742	0	0.150	0.926
-22.458	-21.334	-20.857	26.474	24.965	26.010	0	0.217	1.115
-30.110	-28.409	-28.235	16.922	15.683	17.354	0	0.224	1.302
30.910	29.758	28.969	18.308	17.401	18.984	-10.000	-10.160	-10.065
22.833	22.753	21.079	27.502	26.984	27.484	-10.000	-10.304	-9.891
12.098	11.833	10.849	33.379	32.232	32.673	-10.000	-10.299	-9.822
0	0.114	-0.585	35.230	33.580	34.097	-10.000	-10.295	-9.964
-12.001	-11.469	-11.801	32.832	31.291	31.635	-10.000	-10.304	-10.640
-22.458	-21.355	-21.432	26.474	24.932	25.550	-10.000	-10.289	-10.954
-30.110	-28.440	-28.532	16.922	15.654	16.482	-10.000	-10.233	-11.195

TABLE II
THEORETICAL SCANNER ATTITUDE ANGLES DESCRIBED BY YAW (γ), PITCH (α) AND ROLL (β), AND ESTIMATED ANGLES ($\hat{\gamma}_p$ FOR PHOTOGRAMMETRY AND $\hat{\gamma}_d$ FOR DEPTH CAMERA).

γ [deg]	$\hat{\gamma}_p$ [deg]	$\hat{\gamma}_d$ [deg]	α [deg]	$\hat{\alpha}_p$ [deg]	$\hat{\alpha}_d$ [deg]	β [deg]	$\hat{\beta}_p$ [deg]	$\hat{\beta}_d$ [deg]
60.000	60.440	60.116	90.000	88.770	88.314	0	0.458	5.539
40.000	40.151	40.045	90.000	89.533	88.431	0	0.068	3.113
20.000	20.037	19.903	90.000	89.873	89.435	0	-0.045	1.200
0	0	0	90.000	90.000	90.000	0	0	0
-20.000	-20.124	-19.219	90.000	90.074	90.852	0	0.076	-0.005
-40.000	-40.253	-38.929	90.000	90.214	91.947	0	0.214	-0.065
-60.000	-60.268	-58.310	90.000	90.230	93.358	0	0.324	0.685
60.000	60.167	57.667	90.000	90.449	87.096	0	-1.164	3.379
40.000	40.401	37.673	90.000	90.470	88.032	0	-0.421	1.724
20.000	20.029	17.852	90.000	90.423	89.084	0	-0.056	0.986
0	0.017	1.881	90.000	90.413	90.067	0	0.110	0.177
-20.000	-20.256	-21.845	90.000	90.408	91.099	0	0.280	-0.062
-40.000	-40.320	-41.439	90.000	90.392	92.083	0	0.445	0.066
-60.000	-60.295	-61.255	90.000	90.289	93.674	0	0.571	1.589

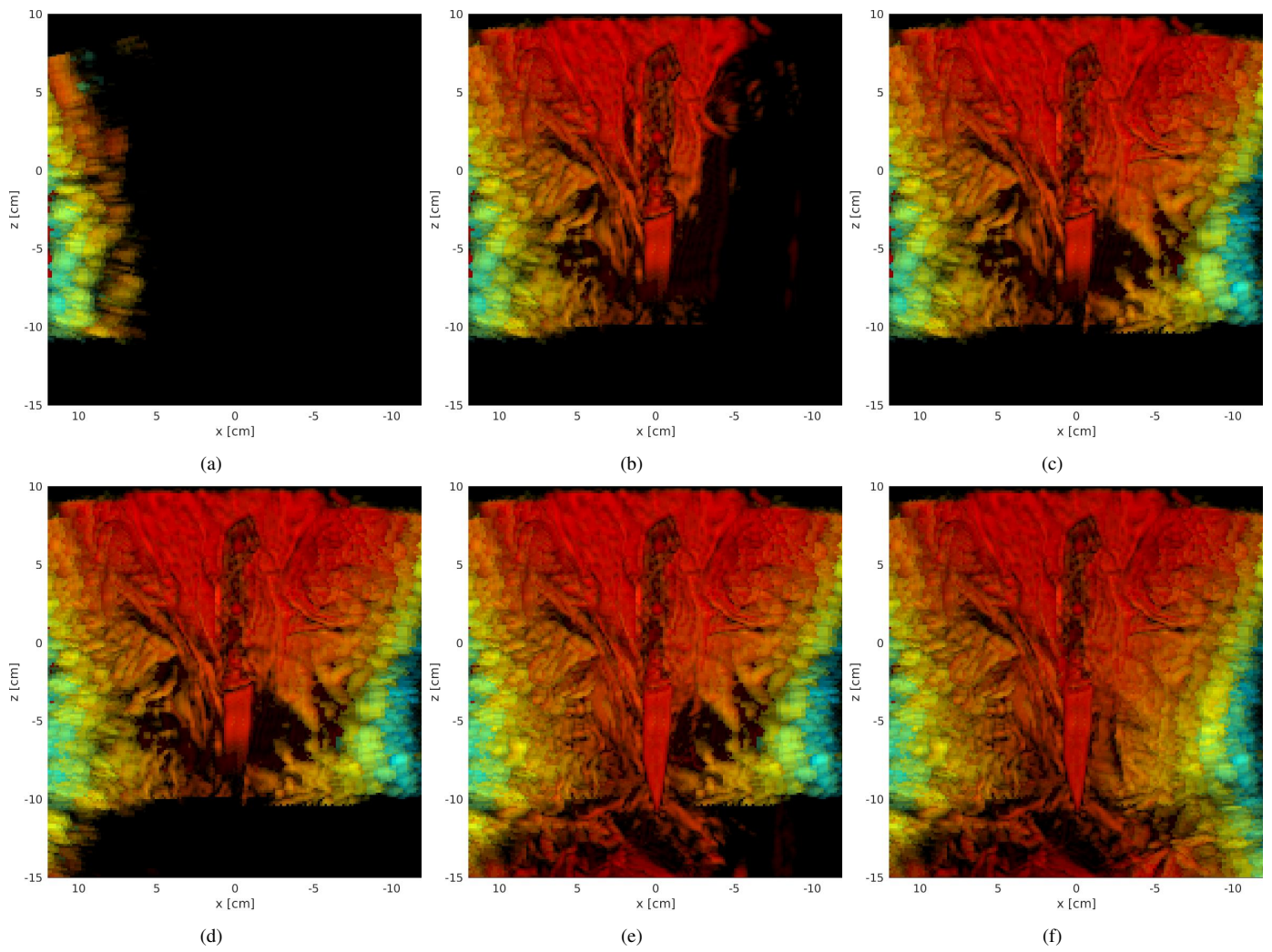


Fig. 3. Mannequin with a concealed knife under a raincoat. mm-wave image considering 14 scanner positions. Cumulative image after (a) 1, (b) 4, (c) 7, (d) 8, (e) 11, and (f) 14 positions.