Citation: I. Cuiral-Zueco, G. López-Nicolás. RGB-D sensing of challenging deformable objects. Workshop on Managing deformation: A step towards higher robot autonomy (MaDef) in conjunction with the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 31-36. October 25-29, 2020, Las Vegas, NV, USA (Virtual).

RGB-D sensing of challenging deformable objects

Ignacio Cuiral-Zueco and Gonzalo López-Nicolás

***Abstract*— The problem of deformable object tracking is prominent in recent robot shape-manipulation research. Additionally, texture-less objects that undergo large deformations and movements lead to difficult scenarios. Three RGB-D sequences of different challenging scenarios are processed in order to evaluate the robustness and versatility of a deformable object tracking method. Everyday objects of different complex characteristics are manipulated and tracked. The tracking system, pushed out the comfort zone, performs satisfactorily.**

I. INTRODUCTION

Tracking deformable and dynamic objects is an essential task when it comes to providing proper feedback on robot-based shape control tasks [1]. Sensing objects with 2D visual sensors may hinder task execution and planning in 3D space, where most deformations happen. Regarding 3D-information-based trackers, systems like [2] tackle 3D object-analysis using RGB-D sensors. Works such as [3] track deformable objects with point clouds and the use of probabilistic methods while other 3D deformable object trackers accomplish efficiency specialising in specific target objects like faces [4], hands [5] or full human bodies [6]. However, in many scenarios, target objects lack texture or specific rigid sub-structure. Seeking generality we presented [7], an RGB-D based object tracking method for texture-less objects that undergo large deformations and movements. In this document the robustness and versatility of the deformable object tracking system is tested with three challenging scenarios for which the method has not been specifically designed. In [7], the method performs properly in base case scenarios. Experimental results show that the method successfully overcomes deformable object tracking when limits are pushed in significantly complex experiments. In particular it performs adequately when facing discontinuities in texture, shape and variety of manipulation actions such us stretching, bending, smashing and spreading.

II. OBJECT TRACKING METHOD OVERVIEW

Method in [7] makes an extensive use of depth information provided by RGB-D cameras and performs a supervoxel [8] over-segmentation of the scene, reducing the number of nodes which must be considered for inference and thus allowing faster run times. The method encloses the target object information in supervoxel graphs that discretise the object's surface position and orientation. The object's state update is performed as a local analysis using supervoxels' colour and normals along with graph's node information. Unlike pure global object shape-based analysis, a local surface analysis allows dealing with a large diversity of deformations. During the method development, special importance was given to achieving generality. However, the method happened to be more suited for convex and mainly continuous box-shaped objects. The object tracking system is combined with a SLAM system thus allowing real time camera location. In [7] a camera position computation method for optimal deformable object perception is also introduced.

*This work was supported by the Spanish Government/EU through project PGC2018-098719-B-I00 (MCIU/AEI/FEDER, UE), and project COMMANDIA SOE2/P1/F0638 (Interreg Sudoe Programme, ERDF).

Instituto de Investigación en Ingeniería de Aragón. Universidad de Zaragoza, Spain. Email: `ignaciocuiral@unizar.es`, `gonlopez@unizar.es`.

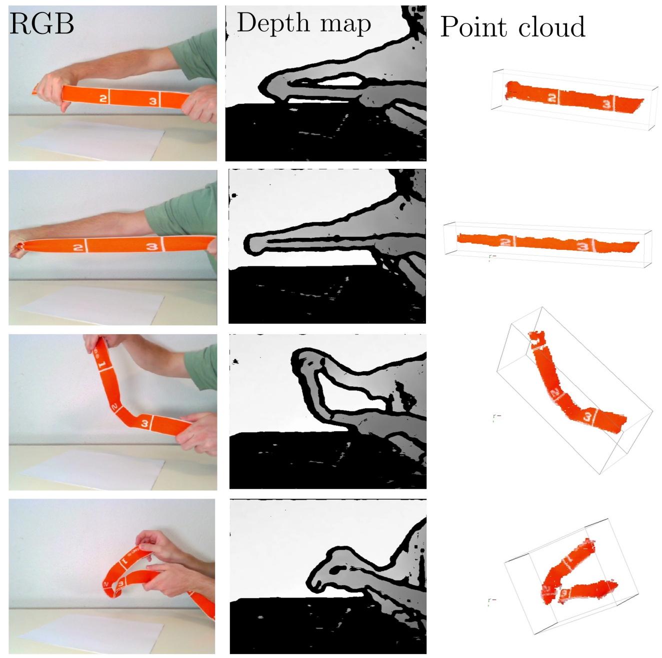

Fig. 1: Experiment 1: elastic band. An orange elastic band is stretched and bent.

III. CHALLENGING OBJECT EXPERIMENTS

Three experiments with challenging objects have been carried out for the purpose of evaluating the performance of the deformable object tracking method in rough scenarios. The program is run and fed with the information obtained from an RGB-D camera in real time. Although the method has been designed to handle dynamic objects and cameras, in the experiment the camera is kept static for the sake of simplicity. For each experiment a sequence of RGB images, depth maps and the 3D point cloud of the resulting tracked object are shown. Sequences were recorded in real time with the Realsense D435 RGB-D camera. The programming

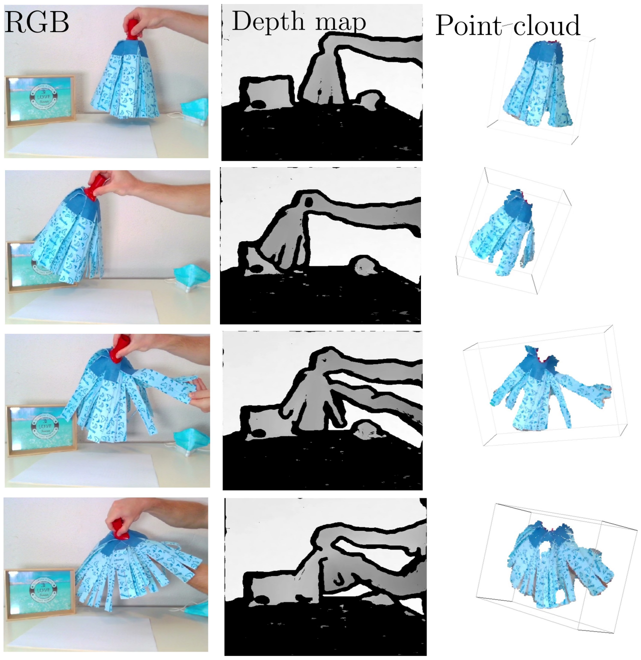

Fig. 2: Experiment 2: mop head. A blue mop head is manipulated avoiding confusion with the background.

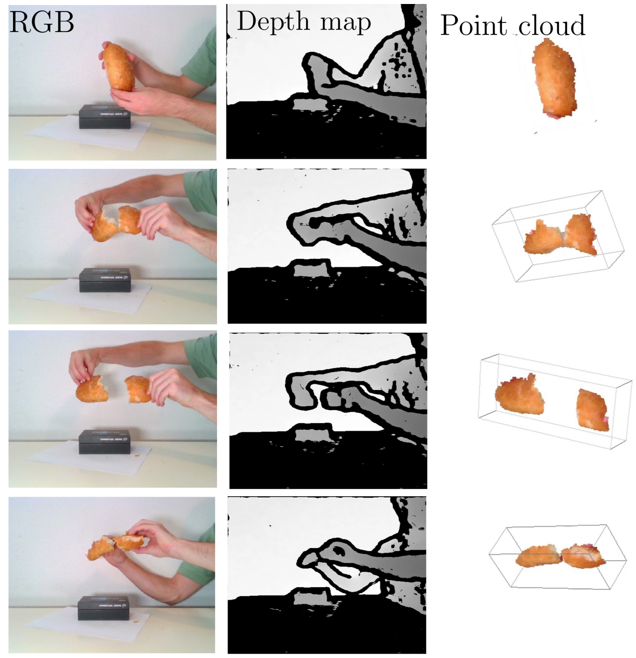

Fig. 3: Experiment 3: piece of bread. A piece of bread is torn in half. After the tearing process, both pieces are smashed.

language used is C++ and all measurements have been recorded on an Intel Core i7 1.8 GHz processor.

1) Elastic band: In this sequence, an orange elastic band is stretched and deformed in multiple manners. This experiment has been selected because the target object is not box-shaped: note how the object's augmented oriented bounding box (see Fig.1) changes its aspect ratio drastically along the sequence. RGB-D information is not as desirably consistent along the surface: noise and depth information gaps can be observed, specially in the third and fourth frame. Although orange is the prominent object colour, given the local analysis of the object, some white numbers and stripes could act as colour perception firewalls. Nonetheless, as the method makes extensive use of depth and surface normals information as well, both the stripes and the numbers are added and tracked properly to the object's point cloud.

2) Mop head: A blue mop head is manipulated in such a way that its ribbons separate and move independently from each other (Fig.2). Similar colour objects (framed picture in the background and mask on the right) have been intentionally placed along the scene. The deformable object tracking method manages to keep track of the separated ribbons, specially noticeable in the second and third frames. In the fourth frame, noisy depth map information leads to gaps and separated surfaces fusion. The use of position and surface normal's information successfully prevents the method from spreading the object location to the blue-coloured picture (second and third frame).

3) Piece of bread: This sequence shows a piece of bread being torn in half (Fig. 3). After being torn, both remaining pieces are smashed. Although the method was not specifically designed for object tearing processes, in [7] a paper-tearing process is performed properly. However, unlike a flat piece of paper, this bread is a more spherical-surface object. It also presents a colour more similar to the skin of the manipulating hands as well as varying texture: clearer tones appear mid-surface when the pieces of bread are smashed in the fourth frame. The method overcomes these difficulties and properly tracks both pieces.

References

[1] R. Herguedas, G. López-Nicolás, R. Aragüés, and C. Sagüés, "Survey on multi-robot manipulation of deformable objects," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2019, pp. 977–984.

[2] A. Petit, V. Lippiello, and B. Siciliano, "Real-time tracking of 3D elastic objects with an RGB-D sensor," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 3914–3921.

[3] J. Schulman, A. Lee, J. Ho, and P. Abbeel, "Tracking deformable objects with point clouds," in *IEEE International Conference on Robotics and Automation*, 2013, pp. 1130–1137.

[4] F. Alqahtani, J. Banks, V. Chandran, and J. Zhang, "3D face tracking using stereo cameras: A review," *IEEE Access*, vol. 8, pp. 94 373–94 393, 2020.

[5] J. Taylor, L. Bordeaux, T. Cashman, B. Corish, C. Keskin, T. Sharp, E. Soto, D. Sweeney, J. Valentin, B. Luff, *et al.*, "Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences," *ACM Transactions on Graphics*, vol. 35, no. 4, pp. 1–12, 2016.

[6] Y. Wu, D. Kong, S. Wang, J. Li, and B. Yin, "An unsupervised real-time framework of human pose tracking from range image sequences," *IEEE Transactions on Multimedia*, pp. 1–14, 2019.

[7] I. Cuiral-Zueco and G. López-Nicolás, "RGB-D tracking and optimal perception of deformable objects," *IEEE Access*, vol. 8, pp. 136 884–136 897, 2020.

[8] J. Papon, A. Abramov, M. Schoeler, and F. Worgotter, "Voxel cloud connectivity segmentation-supervoxels for point clouds," in *IEEE conf. on computer vision and pattern recognition*, 2013, pp. 2027–2034.