Virtual Reality for Demonstrating Tool Pose

Philipsen, Mark Philip; Wu, Haiyan; Moeslund, Thomas B.

Publication date:
2018

Citation for published version (APA):
Philipsen, M. P., Wu, H., & Moeslund, T. B. (2018). *Virtual Reality for Demonstrating Tool Pose*. Abstract from Automating Robot Experiments, Madrid, Spain.

Virtual Reality for Demonstrating Tool Pose

Mark P. Philipsen[1,2], Haiyan Wu[2] and Thomas B. Moeslund[1]

***Abstract*— We investigate the viability of collecting demonstrations, for learning robot control, in Virtual Reality (VR). If acceptable demonstrations can be collected offline using VR, problems with affecting the production line can be avoided and large amounts of realistic data from the actual production environment can be used. We show that the relatively low detail data that can be captured from the production line using a RGB-D camera provides sufficient information for an expert to perform acceptable demonstrations in Virtual-Reality. However, performance is not yet on par with what can be achieved through real world observations. We propose several future improvements that will help close the gap.**

I. INTRODUCTION

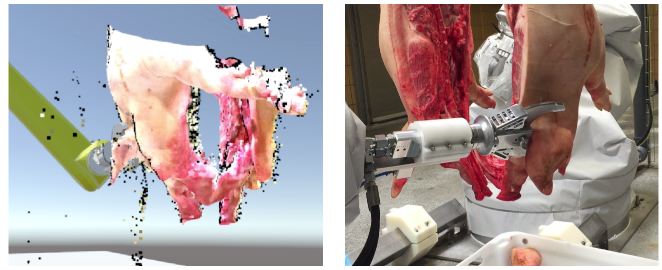
Virtual Reality (VR) provides an intuitive interface for gathering demonstrations and may replace more frustrating methods for programming robots [1]. Besides making the experience more accessible, intuitive, and faster, VR opens up for improvements such as visual aids [3]. The first commercial attempts at VR solutions are beginning to surface, one example is ”VR Robotics Simulator” [5], which lets you build virtual robot cells with robots from a wide range makes.

We are tasked with the automation of processes that are characterized by being destructive and difficult to model. This means that experiments have to be conducted using real and valuable products, making experimentation expensive and slow. For small amounts of data we have the possibility of bringing products from the production line into a controlled environment, but to get the amounts of data that machine learning methods require to be successful, the actual production line is the preferred source of data. The possibilities for experimenting directly on the working production line are limited because production might be disrupted. For this reason we seek to determine how demonstrations in VR compare to demonstrations in the physical world with all the fidelity that brings. Many of the choices that goes into the collection of demonstrations for Learning from Demonstration (LfD) are covered in [2]. With our demonstration setup we strive to limit the differences between the state-action pair that are demonstrated and the states and actions that the robot will respectively experience and perform later. Specifically, we seek to have the demonstrator rely on the same limited data that is available to robot and to use the same tool that the robot will use.

This work was supported by Innovation Fund Denmark and the Danish Pig Levy Fund

[1]Visual Analysis of People Group, Aalborg University, Rendsburggade 14, 9000 Aalborg, Denmark `[mpph,tbm]@create.aau.dk`

[2]Danish Technological Institute, Gregersensvej 9, 2630 Taastrup, Denmark, `[maph,hwu]@teknologisk.dk`



(a) (b)

Fig. 1: Robotic ear cutting from demonstration. (a) Demonstration in Virtual Reality (VR). (b) Resulting tool placement.

Automating the processes we are presented with, requires learning the correct value for one or more of the following parameters for the robot's tool; (1) Position and orientation, (2) Trajectory, (3) Force and torque. Each of these parameters can potentially be learned from demonstrations. We start with processes where demonstration of (1) is sufficient and save more complex processes where (2) and (3) are needed for later. In this paper we use the butchering process of cutting ears of pig heads, as an example of one of the simpler processes where the appropriate position and orientation of a cutting tool must be predicted given the appearance of the pig. We take a current commercial VR system, the HTC VIVE, and stream point clouds from a Kinect V2 RGB-D camera into a virtual environment where an expert demonstrates the wanted position and orientation of the robot's tool. The cutting result of the tool poses demonstrated in VR are compared to cutting results based on demonstrations from outside of VR, where condition are optimal. The quality of the performed cuts are evaluated by an expert butcher. High precision is required since the cutting result can be influenced by the tool being a few millimeters or degrees off. Figure 1(a) shows the view of the demonstrator after having placed the tool by the ear in VR and Figure 1(b) shows physical tool placed according to the demonstration, ready to cut of the ear of the pig.

A. *Contribution*

With this work we seek to determine whether the relatively low detail data captured using a RGB-D camera, provides sufficient information for an expert to perform acceptable demonstrations in VR and how it compares to what can be achieved under ideal conditions.

Workshop on Automating Robot Experiments: Manipulation and Learning
International Conference on Intelligent Robots and Systems 2018, Madrid, Spain

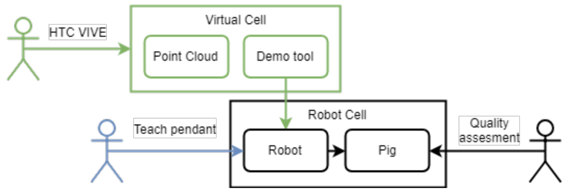


Fig. 2: Experiment setup. (Green) Virtual Reality (VR) demonstration. (Blue) Teach pendant and human vision demonstration.

II. Experiment

Demonstrations are collected using two different methods. Their efficacy is evaluated based on an expert butcher's rating of the results produced with the demonstrated cuts. The result of a cut is primarily determined by the demonstrator's ability to place the tool correctly, which requires experience and visual input. The demonstrator was allowed one hour practise using each method for the ear cutting task prior to the experiment. The demonstrator is reasonably experienced in using VR and TP but has limited limited experience with the butchering task. Time was not limited and the demonstrator was allowed to use as much time as need to place the tool, however most demonstrations took around 5 seconds in VR and slightly more using TP.

Figure 2 shows an overview of the components and actors involved in the experiment.

Method A: Teach Pendant: The robot's teach pendant (TP) is used to place the tool, while standing next to the physical robot and pig. This presents the ideal conditions for the demonstrator to correctly place the tool because the demonstrator can sense the scene directly with his own eyes and from as many view points as needed.

Method B: Demonstration in VR: VR demonstrations are collected in a virtual clone of the physical production cell, build in the Unity game engine [6]. Tool pose demonstrations and point clouds are respectively send and received to and from the virtual cell using the "ros-sharp" software library [4]. The demonstrator is limited to sense the scene through a single camera's view point.

III. Evaluation

Method A and B are evaluated by performing the demonstrated cuts on real pigs. A butcher visually inspects and rates the result on a scale of 1-5, with 1 being unacceptable, 3 acceptable and 5 perfect. Each method was used to cut off 10 ears. Figure 3 shows the ratings given ears cut using respectively TP and VR. The cutting quality through VR results in a median of 3, which is acceptable. However, it is clear that cutting quality suffers compared to the TP approach where the scene can be experienced first hand, from multiple view points. Whereas, with VR the information is limited to a single low resolution and noisy view point.

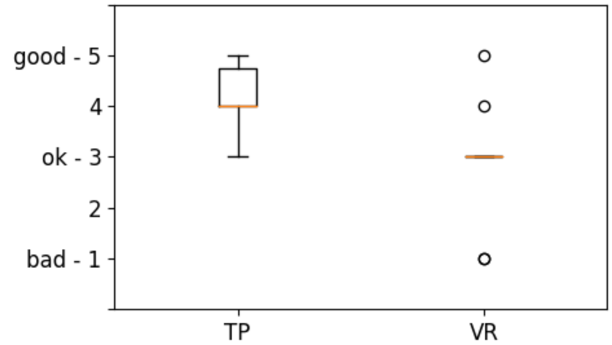


Fig. 3: Score distributions for teach pendant (TP) and Virtual Reality (VR) demonstrations.

IV. Future Work

VR is an intuitive and viable method for collecting demonstrations of tool pose. Tool poses that can be used to train vision systems that can predict appropriate tool pose directly from image data. Further down the road we want to extend the VR demonstration for use with more complex processes, where it is not sufficient to demonstrate a single pose, but where entire trajectories are needed.

We see uses in many areas where it is preferable to utilize the vast amount of data that may be easily and non-evasively collected from the world, as oppose to relying on constructed scenarios for data collection.

The initial experiments have uncovered limitations in relying on a single camera. For this reason we want to add at least one more view point. At the same time further efforts will be put into determining optimal camera placement. Given the need for accuracy in the demonstration, it is a disadvantage that the virtual tool does not accurately reflect the actual robot tool. Additionally, the demonstrator could be aided in locating the transition between head and ear by adding useful visualizations such as overlaying color, showing the orientation of surface normals.

Acknowledgment

References

[1] J. Aleotti, S. Caselli, and M. Reggiani. Leveraging on a virtual environment for robot programming by demonstration. *Robotics and Autonomous Systems*, 47(2-3):153–161, 2004.
[2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and autonomous systems*, 57(5):469–483, 2009.
[3] D. Nathanael, S. Mosialos, and G.-C. Vosniakos. Development and evaluation of a virtual training environment for on-line robot programming. *International Journal of Industrial Ergonomics*, 53:274–283, 2016.
[4] Siemens. ROS#. https://github.com/siemens/ros-sharp, 2018.
[5] M. Technologies. Vr robotics simulator. http://vrrobotsim.com/, 2018.
[6] U. Technologies. Unity. https://unity3d.com/, 2018.