A Step Towards Automated Simulation in Industry

Himangshu Sarma, Robert Porzel, and Rainer Malaka

Digital Media Lab, TZI
University of Bremen,
Bibliothekstr. 1
28359 Bremen, Germany
{sarma,porzel,malaka}@tzi.de

Abstract. Automation always plays an important role in industry. Today, its a basic need for industry. To develop faster manufacturing or delivery automation is an important need. Robots always play the main role for automation in the industry. Robots are mainly designed for specific task. But, the main problem is robots are too expensive for one task. Thats why, it is almost impossible to use robots for small industries. Therefore, we are aiming to develop a pipeline to design a multitasking robot, especially for different kinds of packaging tasks. Typical text based instruction sheets are be the main source of these automation robots, that means robots can package different types of shapes using typical text based packaging instructions. In robotics, learning by demonstration in robotics, could benefit from large body movement dataset extracted from textual instructions. The interpretations of instructions for the automatic generation of the corresponding movements thereof are difficult tasks. We examine methods for converting textual surface structures into the semantic representations and explore tools for analysis and automated simulation of activities in industrial and household settings. In our first step, we try to develop a pipeline from textual instructions to virtual actions that includes traditional language processing technologies as well as human computation approaches. Using the resulting virtual actions we will train robots through imitation learning or learning by demonstration for multitasking packaging robots.

Keywords: Industry, robots, textual, instruction

1 Introduction

Automation is a very important need in industry. Already, automation took place in many areas in industry. But, again there are lots of areas which are difficult to automate until now in industries. For example if we look at the packaging in different industry. We can see there are limitations during automation of packaging. The same robot cannot be used for different types of work. Robots are very expensive [1], and that is why researchers are trying to develop a low cost robotic vision system [2]. Therefore, multitasking robots could be beneficial for the real world, which can perform different tasks in industry. Therefore, our main aim is to develop a pipeline by which we can develop a robot who can handle different kind of packaging shapes and also different kind of other industrial tasks too based on typical text based instruction sheets. Where, as input users will provide them with text based instructions for the actions (e.g. packaging). To perform instructions based task by robots we need to teach robots to do so. For that we need lots of data. Using those data we can teach robots to perform different packaging works in industry. To collect those data we are going to use human computation approaches. So, as a first step we are going to generate 3D animations which will automatically be generated as an output from the provided instruction sheet. Using a human computation approach we determined which visualization serves best as an output of the video [3]. Assessing the quality of human body movement performances is an important task in many other application areas other than industry, ranging from sports to therapy, learning by demonstration in robotics, automated systems for generative animation, and many more. For example, the manual transformation of physical therapy exercises into computer-supported playful exercises in the form of so-called exergames requires a lot of time and effort, making it impractical for therapists or smaller practices to transform their preferred sets of therapeutic exercises into exergames to be used by their patients. Motivated by our use-case of automatically generating movement patterns to be used in motion-based games for the support of physiotherapy, rehabilitation, and prevention, we thus set out to explore the potential of crowd-based quality of motion assessments, as a necessary intermediate step in the extraction and validation of motions.

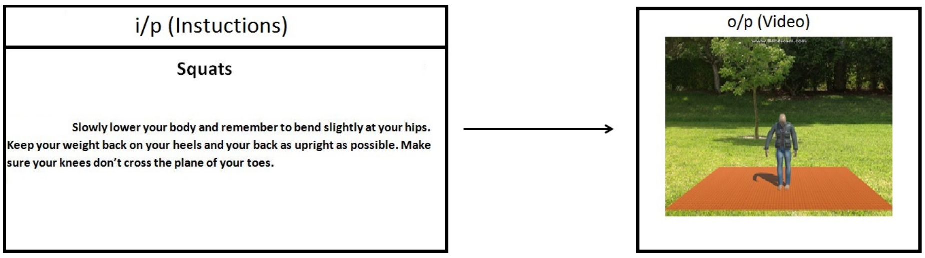

Fig. 1. Expected input and output

2 Background and related work

Based on the main idea to develop an instructions based packaging robot, first we need a lot of human movement data. It is always better to collect this kind of data through some game or some applications rather than paying real humans. Data collections through a game are cost effective and it is also possible to collect lots of data in a small period of time. If we look at the past work we can find some many efforts, where lots of researchers tried to collect data through computer games. Below, we have mention one work where a group of researchers try to collect data to help cooking robots [4].

Tower defense game: There are large sets of data available for imitation learning to robots. It is always difficult to collect such amounts of data. Again, as we mention earlier, it is also very expensive and time consuming to collect this amount of data. In this project researchers mainly tried to collect a large set of data which can use for imitation learning mainly for cooking robots. Here they mainly designed a tower defense game using a human computation approach to gather data on human manual behavior.

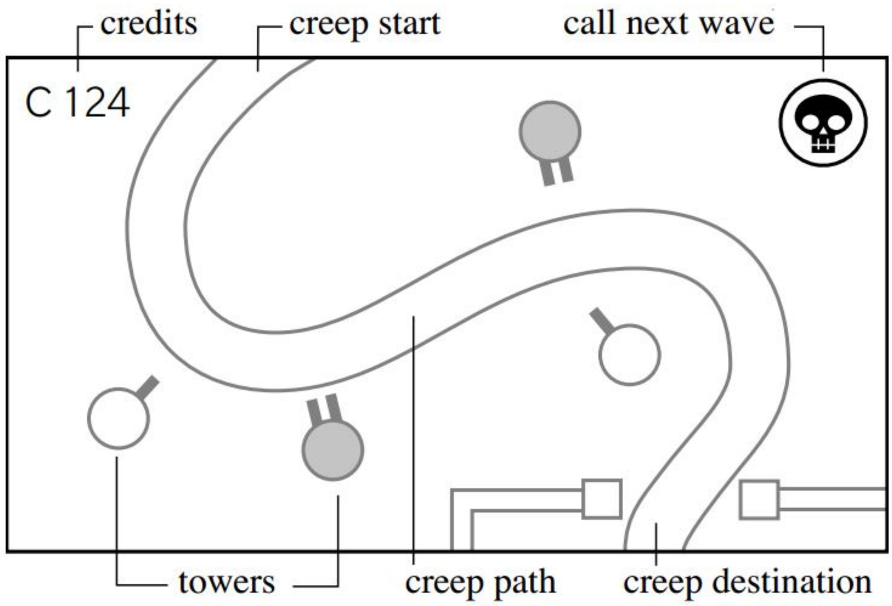

Fig. 2. Kitchen tower defense game

Based on the current state of the art, we set out to establish a human computation based pipeline for extracting validated movements from instruction sheets, with the goal to then explore the potential of further automating the different steps involved in that pipeline, starting with a focus on the step of quality of motion assessment. In our previous work we have found the best visualization category to use in automated 3D animation from typical text based instructions [3]. Below we show how we have found the best visualization, first step towards the autonomous packaging robots.

Data Collection and Results

At the early stage of the work we have developed a Physical Exercise Instruction Sheet Corpus(PEISC) of around 1000 physical exercise instructions. On the basis of different actions we categorize PEISC in different categories, e.g. standing, seating, leaning, lying. We have chosen five exercises from standing where there is no any external equipment used during exercise for our experimental work. The five exercises we have recorded are listed below:

- Squats
- Lateral Lungs
- Standing IT
- Forward Lungs
- Reverse Lunges

With the help of a Kinect [1] we have recorded the performance of those five exercises from seven participants where three male and four female out of them; four are from 15-25 and three from the 25-35 age group.

We have developed four different categories of video from the collected recording data. Below we list the all four categories with specific reason why we selected those:

- *RGB*: its a simple color video, which can easily generate by anyone, people usually watch this kind of video in daily life
- *Depth*: the visualization is likely same as RGB videos, but best part is one can't see the real person, so best when it comes in privacy matter
- *Skeleton*: this category is fully generated by joint tracking sensor of the human body, with this category we can easily find that which joint is moving how much
- *Virtual Reality*: this category is also generated from the skeleton or joint tracking, we can add any virtual character to the skeleton, in case of privacy it is one of the best option for game related area

Using above four categories we have developed a survey application to find the best visualization.

Results

Rather than the above works, there are also certain works which somehow related to our approach. A group of researchers [5] introduced a text to scene engine where user can also change the scene if it's not perfect. Also another multimodal text-to-animation system CONFUCIUS [6] generates animations of virtual humans from single sentences containing an action verb.

[1] https://dev.windows.com/en-us/kinect

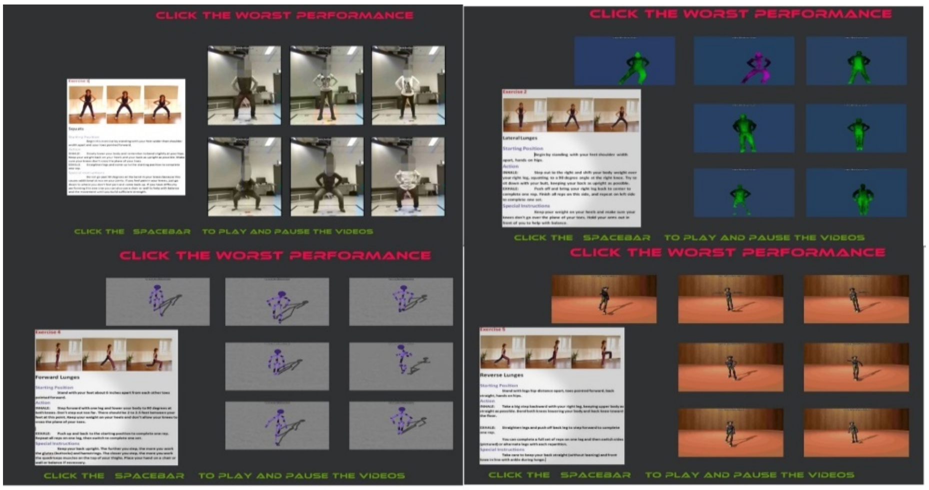

Fig. 3. Screenshots of the survey application showing different visualizations

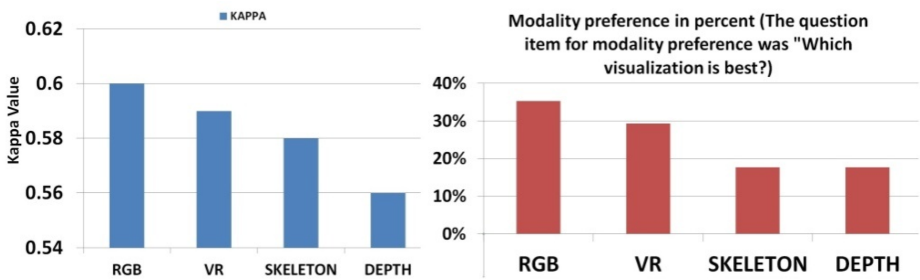

Fig. 4. Best visualization

3 Proposed Approach

Automated packaging simulations in industry are never an easy task. We are trying to achieve the target through different small tasks, which are shown in the pipeline below:

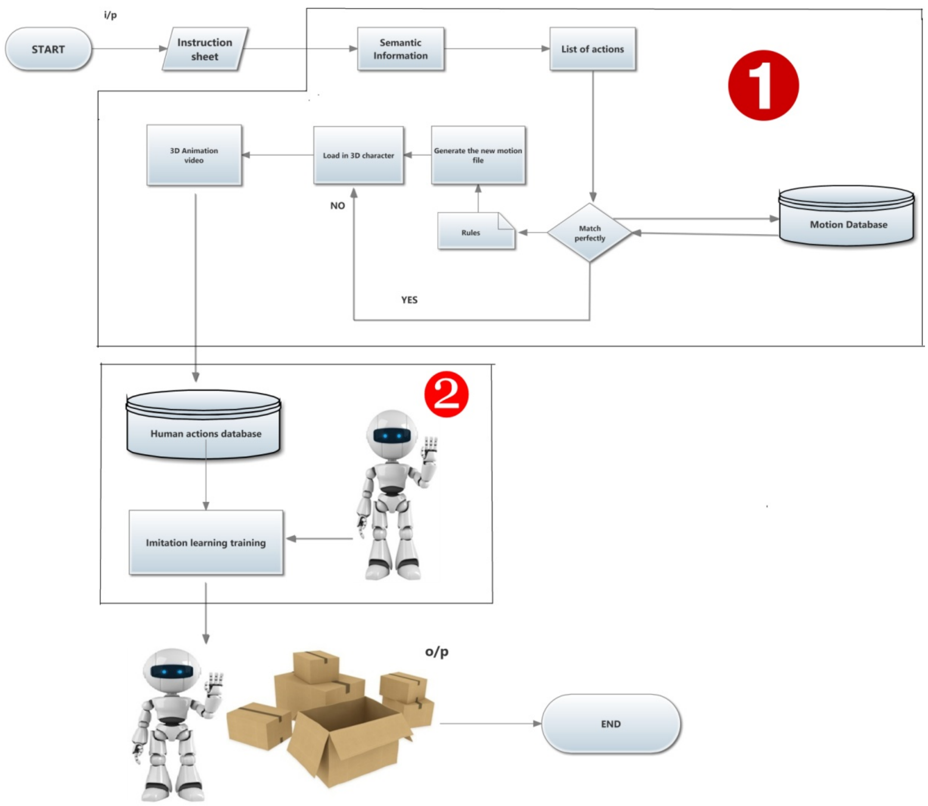

Fig. 5. Automatic packaging simulation pipeline

The whole pipeline of the objectives is mainly divided into two main parts, as shown in Figure 5. Below, we have analyzed the two parts:

- We try to divide this step in three small sub tasks as below:

  1. In this step, our main aim is to extract the list of motions terms from a typical text based instruction. First, we will try to generate the list

of motions extracting the semantic information of the instruction using the Stanford parser [7] and Framenet [8]. Using, human computation approach, crowdsourcing we will verify those motions to get a perfect list of motions.
2. In second step, our main aim is mapping the motion terms to the database. Using, human computation we will label all motions that are stored in the database.
3. This is the last step before we get to our main goal. Here, we try to set up some rules to generate new motion using human computation or crowdsourcing. In detail, if the required motion is not present in the database, then we need to generate new motions using the motion database. To generate new motions from the database, we need different rules. Then, we can use those rules to generate new motions.

– This is the final step to achieve our desired target. After we get the animations from the typical instructions sheet, we will use those data as a database to teach robots. Using this database, we will train robots to perform automatic packaging simulations using imitation learning or learning by demonstration.

4 Challenges

To achieve the desired target as shown in Figure 5, there are some challenges. Below we have described some of those challenges:

– To extract the required motion action terms from the texts based instructions.
– To generate new motions which are not yet available in the database.
– During generation of the virtual actions, it is an important task to ensure clean transitions between two different actions or poses, which is also a big challenge.
– Use this motion database as a training set for imitation training for robots.

5 Conclusion and Future Work

We have presented a pipeline for developing a multitasking robot, for automated simulation in industry, especially for packaging. Here we have divided the whole task into small sub tasks, mainly two major tasks. We tried to show how we will solve these tasks and also we have shown the main challenges to complete these tasks.

References

1. Weinberg, J.B., Engel, G.L., Gu, K., Karacal, C.S., Smith, S.R., White, W.W., Yu, X.: A multidisciplinary model for using robotics in engineering education. In: Proceedings of the 2001 ASEE Annual Conference and Exposition. (2001)

2. Sandhu, G., McGinn, C., Kelly, K.: Investigation into a low cost robotic vision system. (2015)
3. Sarma, H., Porzel, R., Smeddnick, J., Malaka, R.: Towards generating virtual movement from textual instructions: A case study in quality assessment. In: Proceedings of The Third AAAI Conference on Human Computation and Crowdsourcing (HCOMP-2015), AAAI (2015)
4. Walther-Franks, B., Smeddinck, J., Szmidt, P., Haidu, A., Beetz, M., Malaka, R.: Robots, pancakes, and computer games: Designing serious games for robot imitation learning. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM (2015) 3623–3632
5. Chang, A.X., Savva, M., Manning, C.D.: Semantic parsing for text to 3d scene generation. ACL 2014 (2014) 17
6. Ma, M.: Automatic conversion of natural language to 3D animation. PhD thesis, University of Ulster (2006)
7. De Marneffe, M.C., MacCartney, B., Manning, C.D., et al.: Generating typed dependency parses from phrase structure parses. In: Proceedings of LREC. Volume 6. (2006) 449–454
8. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The berkeley framenet project. In: Proceedings of the 17th international conference on Computational linguistics-Volume 1, Association for Computational Linguistics (1998) 86–90