Master's Thesis Engineering Technology

2022-2023

Affodances-based recognition of assembly activities through probabilistic modelling

Colson Joren

Engineering Technology Master of Electromechanical Engineering

Setting

This master's thesis is conducted on behalf of the ACRO research group, specialised in automation, computer vision and robotics. This research group belongs to the Department of Mechanical Engineering of KU Leuven and is located in the Technology Centre at Diepenbeek Campus. This research belongs to ACRO's robotics branch, more specifically: **human-robot interaction and collaboration**. In this field, ing. Martijn Cramer is active with his doctoral research in which a robot **detects operator activities** based on skeleton data to provide optima assistance. Another ongoing doctoral research at ACRO is conducted by ir. Yanming Wu in which the goal is to **estimate the 3D translation and rotation of an object**. This master's thesis seeks for an opportunity to link these two doctoral studies.

Problem statement

Multiple state-of-the-art methods already exist for activity recognition i.e., skeleton-based activity recognition, illustrated in Figure 1. This method focuses **only** on the **operator's movements**. This is not ideal since a lot of information is lost by looking only at the movements of the operator and **different operators** perform the same actions in **different ways**.

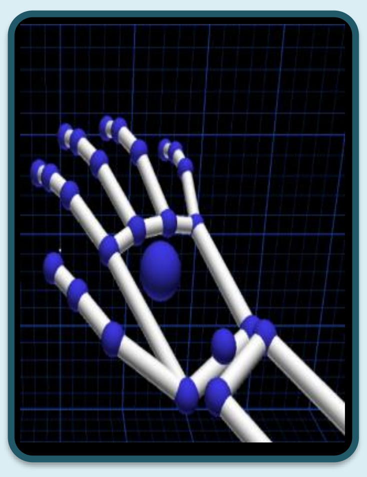

Figure 1: representation of skeleton data [1]

Objective

- **More flexible and robust approach for recognising operator actions**, the approach is depicted in Figure 2
- Higher success rate than similar conducted research : **78 per cent**
- **Free positioning** of the assembly parts and **usable by multiple operators**

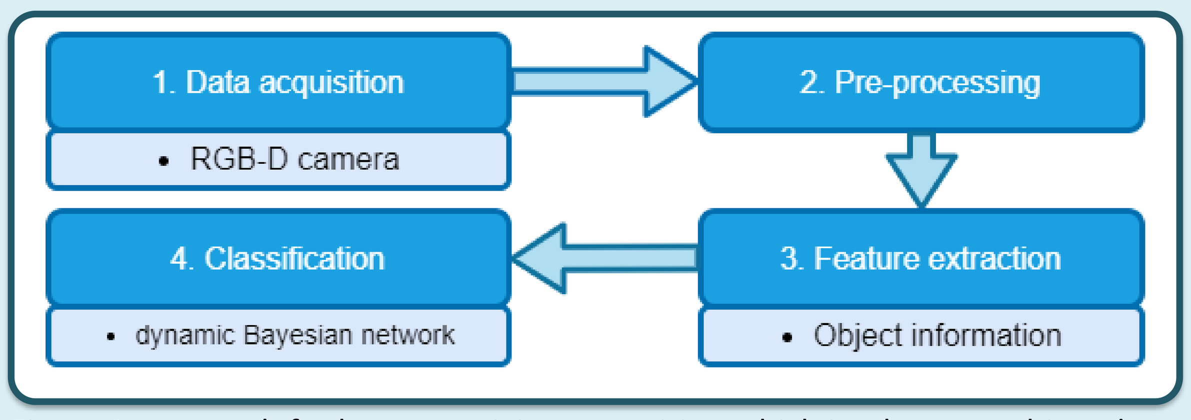

Figure 2: approach for human activity recognition which implements object data

Method

- Study to compare different **probabilistic graphical models**, **probabilistic logic languages**, and **object detection and tracking methods**
- **Dynamic Bayesian network** to recognise the performed assembly activity at each time step based on information obtained from the assembly parts and their movements at that time step. After multiple iteration, the model of the network is finalised and depicted in Figure 3. The parent nodes are:
  - the **main assembly part** that is being manipulated,
  - the **relative distance**, **relative velocity**, and **relative direction** between that part and its relevant parts.

The CPD table between the parent nodes and the child node is constructed based on training data. During the training phase, 35 assembly sequences of the Bourjault ballpoint pen, visualised in Figure 4, were recorded and each frame was manually **labelled** with the **input variables** and the **assembly activity** that was performed in that frame. The possible assembly activities are presented in Table 1. The probability that an activity is being performed based on a set of input variables can then be calculated by applying the following formula:

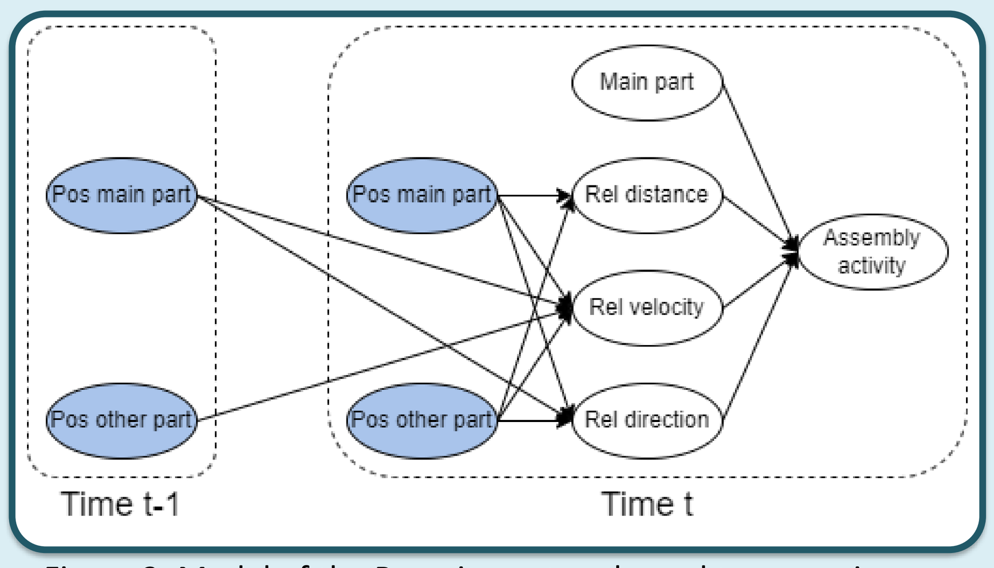

Figure 3: Model of the Bayesian network used to recognise assembly activities

$$P(Activity|\ set\ of\ input\ variables) = \frac{\#frames|activity, set\ of\ input\ variables}{\#\ frames|set\ of\ input\ variables}$$

Table1: possible assembly activities

No assembly activity	Moving parts
Loading ink	Inserting filling
Attaching bottom	Attaching ballpoint
Placing cap	

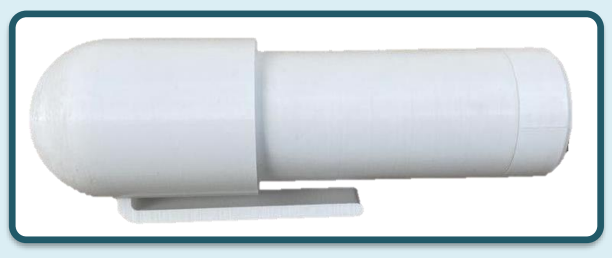

Figure 4: Bourjault ballpoint pen

Results

In order to assess the performance of the model, a 7-fold cross-validation approach was employed. This involved utilising 30 assembly sequences as training data and the remaining five sequences as test data to calculate metrics such as accuracy, precision, recall, and F1-score. The final model yielded

- **an accuracy of 75%,**
- **a precision of 76%,**
- **a recall of 75%**
- **and an F1-score of 74%.**

Furthermore, when the model trained by all 35 assembly sequences was applied to recognise operator activity in an additional assembly sequence, it achieved an accuracy of 77%.

Figure 5 illustrates the result of this final assembly sequence using bar plots, where each bar plot represents the actual activity performed, and each bar in their respective plot represents the amount of times that each activity is recognised. For example, in the third plot corresponding to the "loading ink" activity, six frames were labelled as "no assembly activity", 92 frames were correctly labelled as "loading ink", 19 frames were labelled as "ink loaded", and three frames were labelled as "filling inserted".

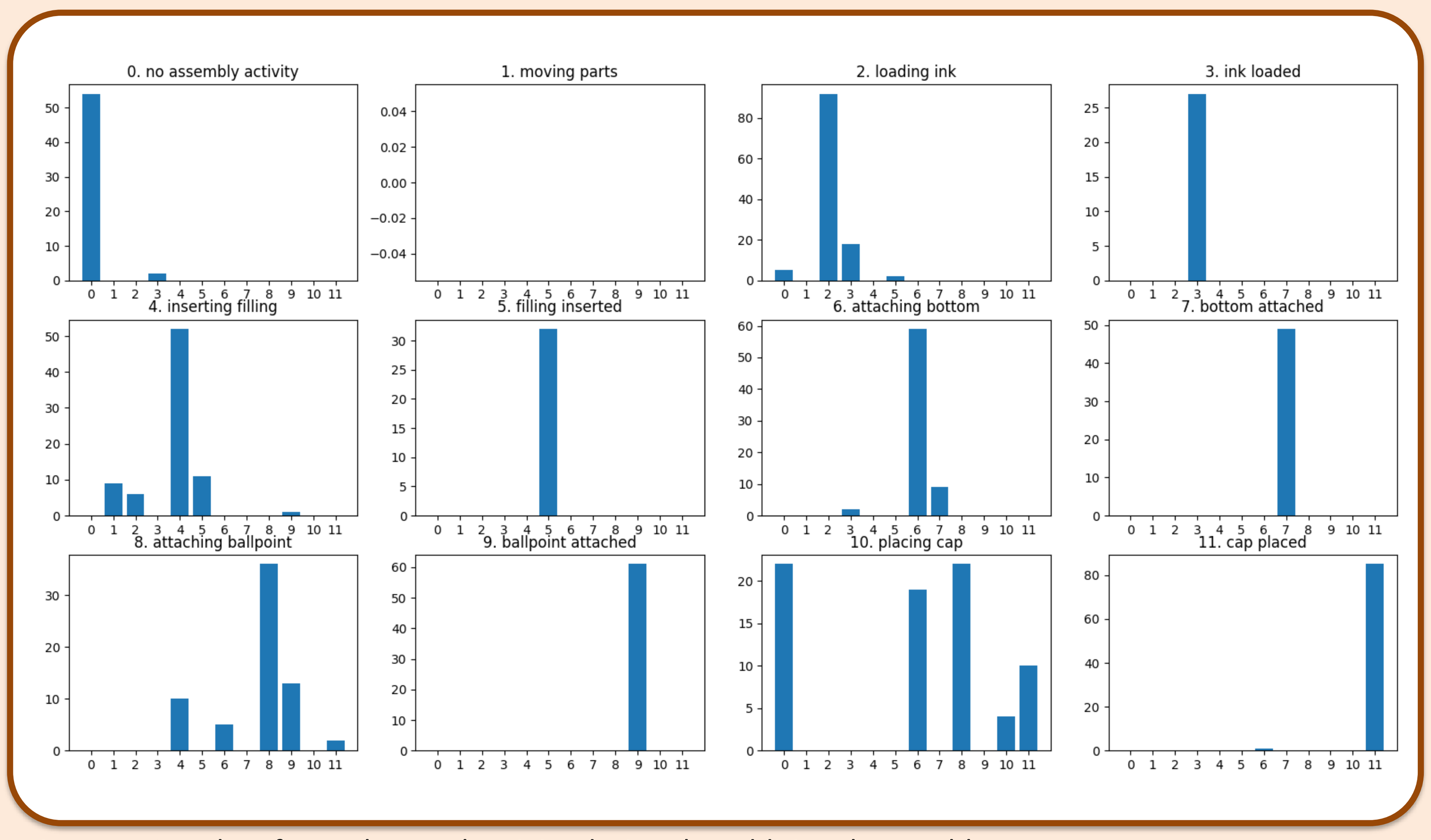

Figure 5: Bar plots for each actual activity during the additional assembly sequence

Supervisors / Co-supervisors / Advisors: Prof. dr. ir. Eric Demeester, Ing. Martijn Cramer, Ir. Yanming Wu

[1] A. Roitberg, N. Somani, A. Perzylo, M. Rickert, and A. Knoll, "Multimodal human activity recognition for industrial manufacturing processes in robotic workcells," in *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, Association for Computing Machinery, Inc, Nov. 2015, pp. 259–266. doi: 10.1145/2818346.2820738.