Suhas Bettapalli Nagaraj

EE 597 IoT Review 10

18 October 2020

Injecting Life into Toys

The main objective of this paper is to explore the implications of using the smartphone capabilities such as sensing, learning and communication for use in toys.

The authors explain that toys today are pre-programmed, non-evolving and costly and propose that it might be possible to leverage mobile devices to enable a new ecosystem of toys and apps which can interact and grow with children.

The first step in this process is to Understand the child's language of interaction - the way the child interacts.

Understanding the child's gesture consists of two sub-tasks:

1. detecting the presence of a gesture in the sensor stream, and
2. clustering similar gesture signals for the purpose of classification.

This paper further narrows down the problem scope and attempts to only recognize the child's "motion gesture vocabulary".

Data collection : The data collection was done in a toddler day care facility.

1. Experiments in toddler day care facility Multi-modal sensing app on phones inserted in toys
2. Continuous passive sensing Hand written ground truth (time : activity)

Real world challenges : There are a few real world challenges such as :

1. Low Gesture Density - Sparse interaction with toy.
2. Identifying gesture Signals - Many indeliberate motions exhibiting high energy.

Other challenges :

1. No pre-defined gestures to train on

2. Signals from similar gesture look different
3. Detecting the start and end of the gesture

Thus, there is a need some way to generate more gesture data and we need some way to narrow down search space.

The authors made an observation that children tend to repeat gestures whenever the toy responds to the gesture. The authors try to leverage this opportunity - Suspect gesture in real time (allow for false positives). Generate arbitrary response (e.g., cartoon sounds). Look for repeat gesture.

Back End Tasks :

1. When to respond? Return **signal segmentation**
2. Are they gestures? Return **signal self-similarity / dissimilarity**
3. Collating gestures? **Return clustering**

1. **Segmentation** : Choosing when to respond - Too quickly is annoying and can lead to many false positives and too late does not exhibit responsiveness.

2. **Signal dissimilarity** : Dynamic Time Warping (DTW) - Compute distance between two signal segments.

3. **Hierarchical Clustering** : Assume gestures repeated after a feedback and count the number of neighbor pairs in a subtree. If no neighbor-pair in a subtree, return **non-gesture**. But, If a subtree has at least one neighbor pair, return **gesture**.

Cut the tree such that we can maximize the number of gestures, and then neighbor pairs.

Results : Ground truth: 2 children, 6 gestures (38 instances)

Detected: 4 gestures, precision: 85.1%, recall: 81.6%

However, there were a few points that needed more clarity :

Questions :

1. Small data set - IRB constraints, difficult to work with toddlers. Can we make it more robust with a bigger dataset and possible even extend the scope?

2. Better algorithms such as Hidden Markov models, time frequency analysis, etc. could have been used to both increase the accuracy as well as to predict what the toddler will do next using Hidden Markov Models.

3. Energy consumption - This paper side steps energy but it is possible to optimize the energy issue and help increase the battery usage.

4. Better definition of gestures - Stepping on toy, or flicking it … defined as non-gestures.

More details about the paper

Fan, S., Shin, H., & Choudhury, R. R. (2014, February). Injecting life into toys. In Proceedings of the 15th Workshop on Mobile Computing Systems and Applications (pp. 1-6).