CADT Beamer Poster Template

Author 1, Author 2, Author 3

Department of Engineering Science, Cambodia Academy of Digital Technology

{author1,author2,author3}@cadt.edu.kh

ACET 2022
RESEARCH | SHARE | CONNECT

1. Introduction

► **Symbol Grounding × Chatting**
► Research on language acquisition and symbol grounding (focus on the acquisition of physically grounded knowledge through utterances that express physical things, such as objects and motions)
► Most of the previous studies have focused on learning without any prior symbolic knowledge
► The problem of how to acquire physically grounded knowledge based on grounded utterances through natural interaction has yet to be explored
► We focus on object-teaching utterances as grounded utterances
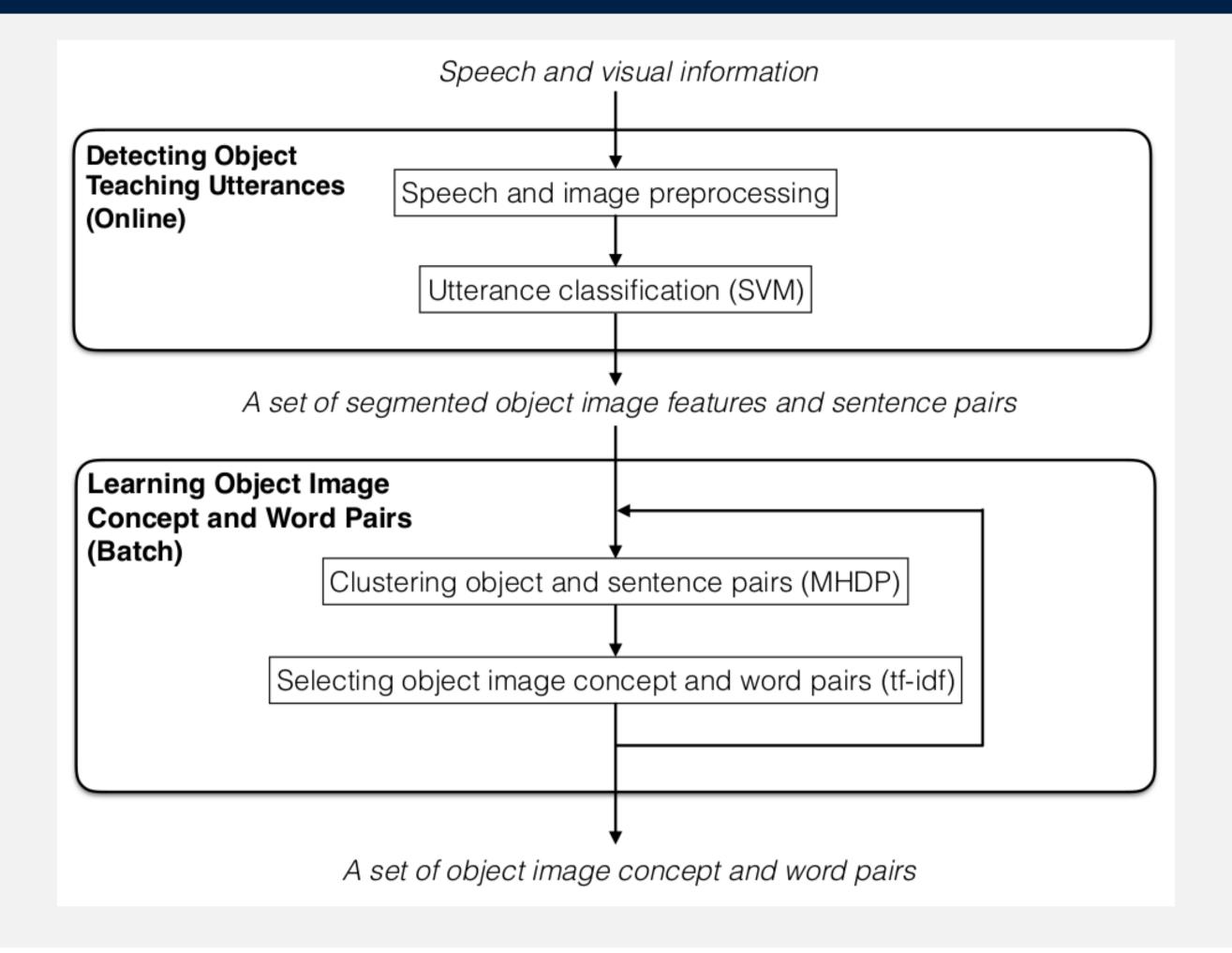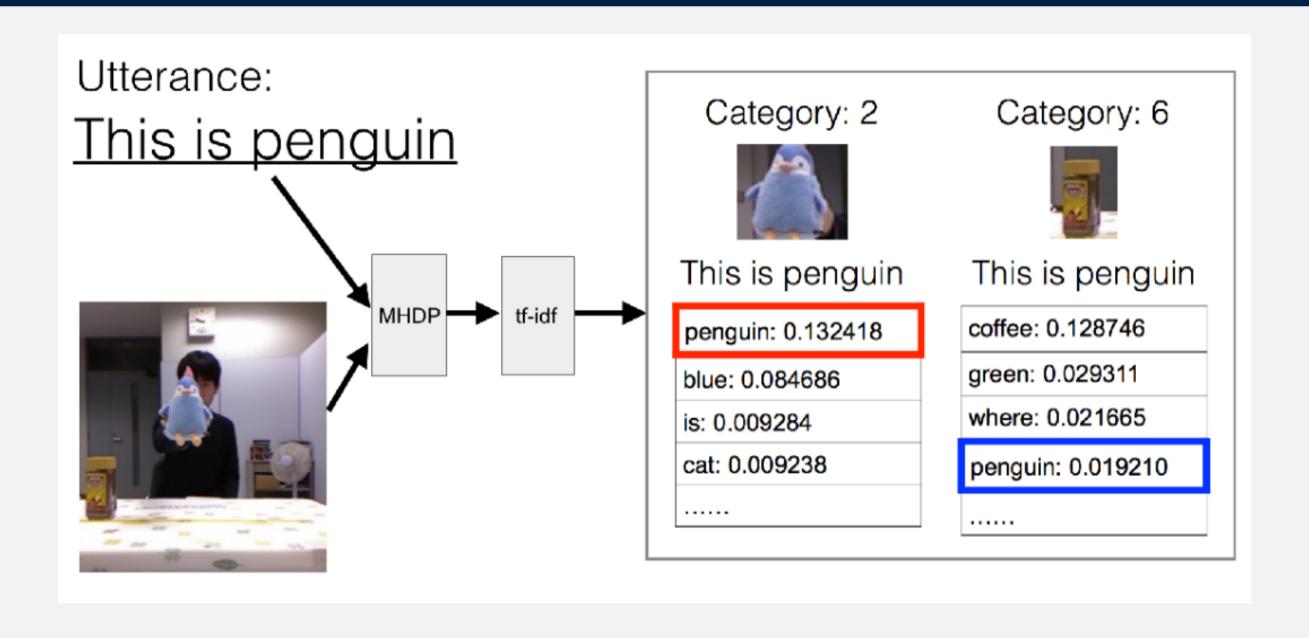
2. Experimental Environment

3. Example Dialogue

► Some examples of dialogue between human and robot:

Do you know any toys?	I am not familiar with toy.
Here is the stuffed toy.	Oh, I see.
Do you like animals?	I like dogs.
I like this penguin.	I got it.

4. Propose Method

Speech and visual information

Detecting Object Teaching Utterances (Online)
- Speech and image preprocessing
- Utterance classification (SVM)

A set of segmented object image features and sentence pairs

Learning Object Image Concept and Word Pairs (Batch)
- Clustering object and sentence pairs (MHDP)
- Selecting object image concept and word pairs (tf-idf)

A set of object image concept and word pairs

5. Learning Method (MHDP+tf-idf)

Utterance:
This is penguin

Category: 2 — This is penguin
- penguin: 0.132418
- blue: 0.084686
- is: 0.009284
- cat: 0.009238
- ......

Category: 6 — This is penguin
- coffee: 0.128746
- green: 0.029311
- where: 0.021665
- penguin: 0.019210
- ......

6. Experimental Setup

The ten objects used in the experiment are: two black stuffed toy cats (small & big), two stuffed toy fishes (red & yellow), and two cups (red & yellow)

7. Results

Results...

Table 1:Results of learning accuracy of object and words (%)

without loop	31% (61/196)	30% (59/196)	10% (19/196)
with loop	35% (69/196)	57% (112/196)	28% (54/196)

Here,
► P_w: probability of selecting correct word in each sentence
► P_c: probability of selecting object image concept for each sentence
► P_{wc}: probability of selecting both correct word and object image concept for each sentence
► Result: without loop $<$ with loop

Acknowledgments

References

[1] Ye Kyaw Thu, Takuya Ishida, Naoto Iwahashi, Tomoaki Nakamura, and Takayuki Nagai. Symbol grounding from natural conversation for human-robot communication. In Britta Wrede, Yukie Nagai, Takanori Komatsu, Marc Hanheide, and Lorenzo Natale, editors, *Proceedings of the 5th International Conference on Human Agent Interaction, HAI 2017, Bielefeld, Germany, October 17 - 20, 2017*, pages 415–419. ACM, 2017.

[2] Ye Kyaw Thu, Takuya Ishida, Naoto Iwahashi, Tomoaki Nakamura, and Takayuki Nagai. Original poster for CADT Beamer Poster Template. https://github.com/ye-kyaw-thu/papers/blob/master/HAI2017/hai2017-poster.pdf, 2017. [Online; accessed 28-June-2022].