Semantic Annotation for the CMU-MMAC Dataset (Version 2)

Documentation

13 July 2018

1 General information

Experiment title	Semantic Annotation for the CMU-MMAC Dataset (Version 2)
Experiment id	D20180105-Semantic_Annotation_CMU-MMAC-KY
Principal investigators	Kristina Yordanova[1], Frank Krüger[2], Thomas Kirste[1]
Affiliation	1 Mobile Multimedia Information Systems Group, Institute of Computer Science, University of Rostock, Rostock, Germany 2 Signal Theory and Digital Signal Processing Group, Institute of Communications Engineering, University of Rostock, Rostock, Germany
E-Mail	kristina.yordanova@uni-rostock.de
Date	13 July 2018
Type	Annotation and models
Location	Albert-Einstein-Straße 21, 18059 Rostock, Germany
Keywords	behaviour annotation, labelling, activity recognition, CMU-MMAC sensor dataset

Notes	This dataset contains the 2nd version of the dataset published as https://doi.org/10.18453/rosdok_id00000163.
Language	English
Rights	CC BY 4.0
Download link	https://doi.org/10.18453/rosdok_id00002273

1.1 Objective

To create semantic annotation of the CMU-MMAC grand challenge kitchen dataset, which is often cited but, due to missing and incomplete annotation, almost never used.

1.2 Problem Statement

Providing ground truth is essential for activity recognition for three reasons: to apply methods of supervised learning, to provide context information for knowledge-based methods, and to quantify the recognition performance. Semantic annotation extends simple symbolic labelling by assigning semantic meaning to the label, enabling further reasoning. We use a novel approach to semantic annotation by means of plan operators [4] to create semantic annotation of the CMU grand challenge dataset [2], which is often cited but, due to missing and incomplete annotation, almost never used.

2 Changes to Earlier Versions

- Changes to the starting times and the actual plan sequences. The previous version contains errors in the plan sequences.
- The dataset now also contains the models that were used to validate the annotation plan sequences. The models were also used for semantic reasoning on the environment properties.

3 Description

3.1 Overview of the CMU-MMAC

The Carnegie Mellon University Multi-Modal Activity Database (CMU-MMAC) consists of five sub datasets (Brownie, Sandwich, Eggs, Salad, Pizza) [2]. Each of them contains recorded sensor data from one food preparation task. The dataset contains data from 55 subjects, were each of them participates in several sub experiments. While executing the assigned task, the subjects were recorded with five cameras and multiple sensors. While the cameras can be used for computer vision based activity recognition [1], the resulting video log is also the base for the dataset annotation. An annotated label sequence for 16 subjects can be downloaded from the CMU-MMAC website[1]. Albeit following a grammatical structure of verbs and objects, the label sequence is still missing semantics which if present would allow the deriving of context information such as object locations and relations between actions and entities.

3.2 Semantic annotation for the CMU-MMAC

To enable the usage of the CMU-MMAC dataset, we followed the process proposed in [4] and annotated three of the five sub datasets (Brownies, Sandwich, and Eggs).

To define the label set, two domain experts reviewed a subset from the video logs and identified 13 action classes (11 for the Brownie, 12 for the Eggs, and 12 for the Sandwich). Table 2 shows the action classes for the three datasets. The action definitions created in

Table 2: Action classes for the three datasets.

Dataset	Action classes
Brownie	open, close, take, put, walk, turn on, fill, clean, stir, shake, other
Eggs	open, close, take, put, walk, turn on, fill, clean, stir, shake, other, turn off
Sandwich	open, close, take, put, walk, turn on, fill, clean, stir, shake, other, cut

this step later enable different annotators to choose the same label for identical actions. In this step the domain experts also identified the entities (30 for the Sandwich dataset, 44 for the Brownies, and 43 for the Eggs). From these dictionaries, in step two, a discussion about the type signature and possible instantiations took place (119 unique labels where identified for the Sandwich dataset, 187 for the Brownies, and 179 for the Eggs. Additionally, 13 state properties were defined.

Two annotators followed the proposed process until all datasets were annotated without gaps and all annotation sequences were shown to be valid plans.

The resulting annotation consists of 90 action sequences. Interestingly, while annotating, we noticed that the experimenter changed the settings during the experiments' recording. In all sub-experiments it can be seen that, before recording subject 28, some objects were relocated in different cupboards.

More information about the annotation process and the evaluation of the quality of the produced annotation can be found in the paper "Providing Semantic Annotation for the CMU Grand Challenge Dataset" [4].

3.3 Data format

The annotation is produced with the help of the ELAN annotation tool [3]. Figure 1 shows a screenshot of the "Brownie" dataset being annotated. The ELAN tool saves the annotation in XML-like format, which contains information about the annotation,

[1] http://www.cs.cmu.edu/~espriggs/cmu-mmac/annotations/

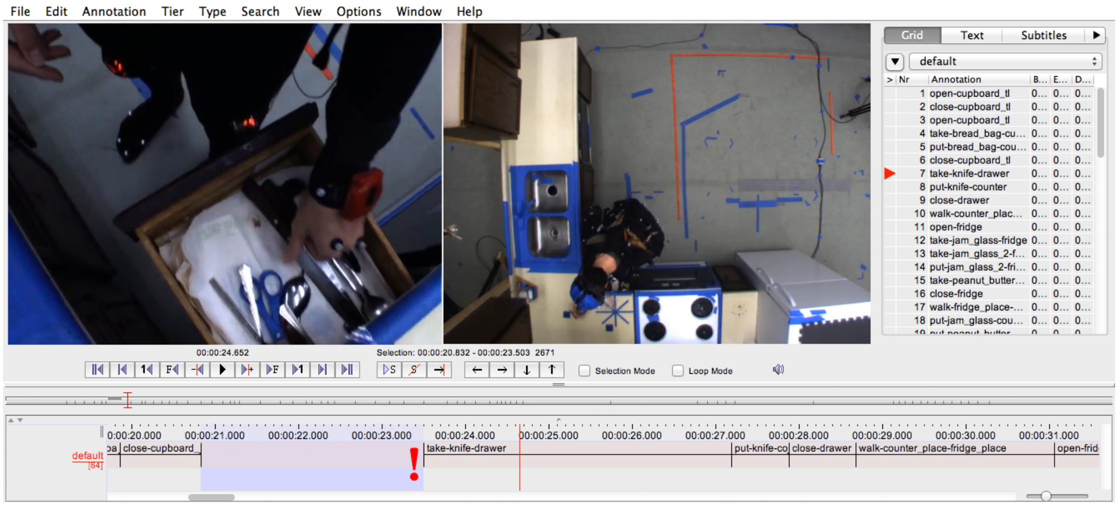

Figure 1: An example of the annotation in ELAN.

the time slot to which each label was assigned and the videos that were used for the annotation.

Beside the ELAN annotation format (.eaf), we also used our own tool to convert the .eaf files into plans, which were later validated with our semantic models. Table 3 shows an extract of the annotated plan for subject S09 from the "Brownie" dataset.

Folder *Annotation* contains three sub-folders for each of the three sub-datasets. In each folder, there are two types of files. The first is an .eaf format, containing the ELAN annotation. These files can be opened with the ELAN tool. The second format is a .txt format and it contains the plans corresponding to the annotation (see Table 3). There the first column contains the time in milliseconds, the asterisk indicates that the action is new, and the last column is the executed action.

The sensor data corresponding to the annotation can be downloaded from `http://kitchen.cs.cmu.edu/`. The names of the annotation files correspond to the names of the subjects on the CMU website.

3.4 Models

The models are implemented in a PDDL like syntax using preconditions and effects. However, since it is not standard PDDL, you cannot use plan validators like VAL or INVAL for plan validation directly. In order to use the model, the computational causal behaviour models (CCBM) is required. In order to get access to this toolbox, please contact the authors via email.

4 Acknowledgments

We would like to thank the students who annotated the dataset. The video data was obtained from `kitchen.cs.cmu.edu` and the data collection was funded in part by the National Science Foundation (EEEC-0540865).

5 Bibliography

References

[1] Ehsan Zare Borzeshi, Oscar Perez Concha, Richard Yi Da Xu, and Massimo Piccardi. Joint action segmentation and classification by an extended hidden markov model. *IEEE Signal Process. Lett.*, 20(12):1207–1210, 2013.

[2] Fernando de la Torre, Jessica Hodgins, Javier. Montano, Sergio Valcarcel, R. Forcada, and Justin Macey. Guide to the carnegie mellon university multimodal activity (CMU-MMAC) database. Technical Report CMU-RI-TR-08-22, Robotics Institute, Carnegie Mellon University, July 2009.

[3] Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. ELAN: a professional framework for multimodality research. In *Proc. of Int. Conf. on Language Resources and Evaluation*, pages 1556–1559, 2006.

[4] Kristina Yordanova, Frank Krüger, and Thomas Kirste. Providing semantic annotation for the cmu grand challenge dataset. In *Workshops Proceedings of the IEEE International Conference on Pervasive Computing and Communications*, Athens, Greece, March 2018.

Table 3: Extract of the annotated plan from the Brownie dataset of subject S09. The overall number of actions is 142.

	Start Time	Action
1	00:00.000	open-cupboard_tl
2	00:03.198	take-brownie_box-cupboard_tl
3	00:04.616	close-cupboard_tl
4	00:05.456	other
5	00:10.735	put-brownie_box-counter
6	00:12.006	open-cupboard_tr
7	00:13.941	take-bowl-cupboard_tr
8	00:17.044	put-bowl-counter
9	00:18.522	take-measuring_cup_s-cupboard_tr
10	00:20.540	put-measuring_cup_s-counter
11	00:21.571	take-measuring_cup_l-cupboard_tr
12	00:24.704	put-measuring_cup_l-counter
13	00:25.682	close-cupboard_tr
14	00:26.708	open-cupboard_br
15	00:29.379	take-oil_bottle-cupboard_br
16	00:30.412	close-cupboard_br
17	00:31.313	put-oil_bottle-counter
18	00:32.758	other
19	00:35.519	take-brownie_box-counter
20	00:36.784	other
21	00:39.069	put-brownie_box-counter
22	00:40.044	walk-counter_place-fridge_place
23	00:41.175	open-fridge
24	00:42.547	open-egg_box
25	00:43.660	take-1-egg_shell-egg_box
26	00:45.432	take-1-egg_shell-egg_box
27	00:46.697	close-egg_box
28	00:47.296	close-fridge
29	00:48.724	walk-fridge_place-counter_place
30	00:50.082	put-2-egg_shell-counter
31	00:50.946	take-1-egg_shell-counter
32	00:52.502	open-egg_shell
33	00:54.224	fill-egg-open_egg_shell-bowl
34	00:56.411	put-1-empty_egg_shell-sink
35	00:57.910	take-1-egg_shell-counter
36	00:58.976	take-1-egg_shell-counter
37	00:58.980	open-egg_shell
38	01:00.668	fill-egg-open_egg_shell-bowl
39	01:03.167	put-1-empty_egg_shell-sink
40	01:05.817	walk-counter_place-fridge_place