HMM-based Activity Recognition with a Ceiling RGB-D Camera

Daniele Liciotti, Emanuele Frontoni, Primo Zingaretti, Nicola Bellotto, and Tom Duckett

Università Politecnica delle Marche (Italy), University of Lincoln (UK)

{d.liciotti, e.frontoni, p.zingaretti}@univpm.it, {nbellotto, tduckett}@lincoln.ac.uk

Introduction

The main goal is to classify and predict the probability of an analysed subject action. We perform activity detection and recognition using an inexpensive RGB-D camera. Human activities, despite their unstructured nature, tend to have a natural hierarchical structure; for instance, generally making a coffee involves a three-step process of turning on the coffee machine, putting sugar in the cup and opening the fridge for milk. Action sequence recognition is then handled using a discriminative hidden Markov model (HMM).

The innovative aspects are in proposing an adequate HMM structure and also the use of head and hands 3D positions to estimate the probability that a certain action will be performed, which has never been done before, for the best of our knowledge, in ADLs recognition in indoor environments.

ADLs Model

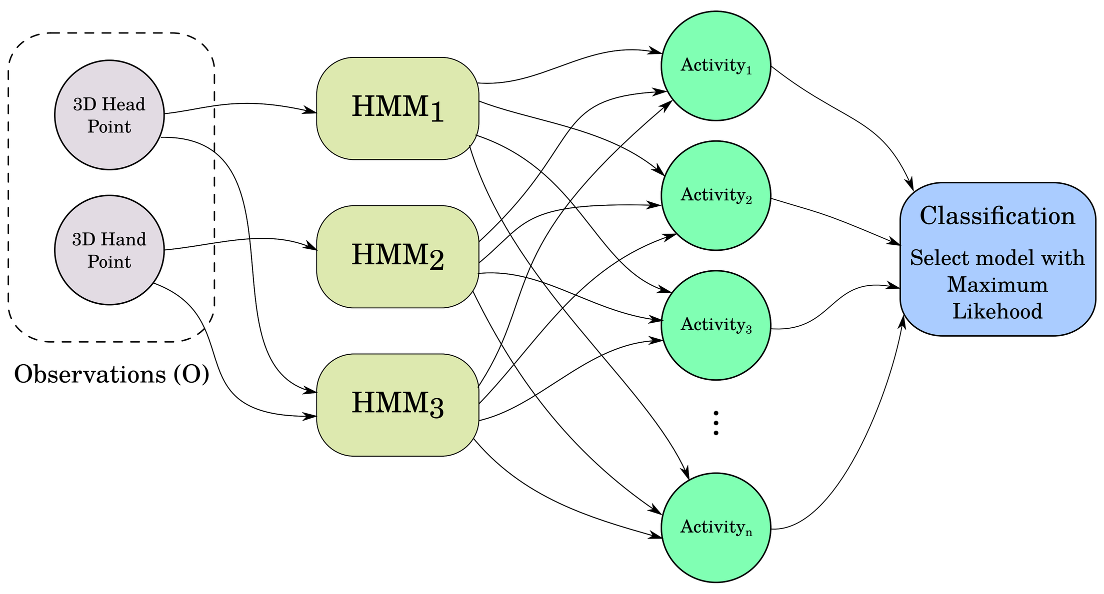


Figure 1: Block diagram of the recognition process.

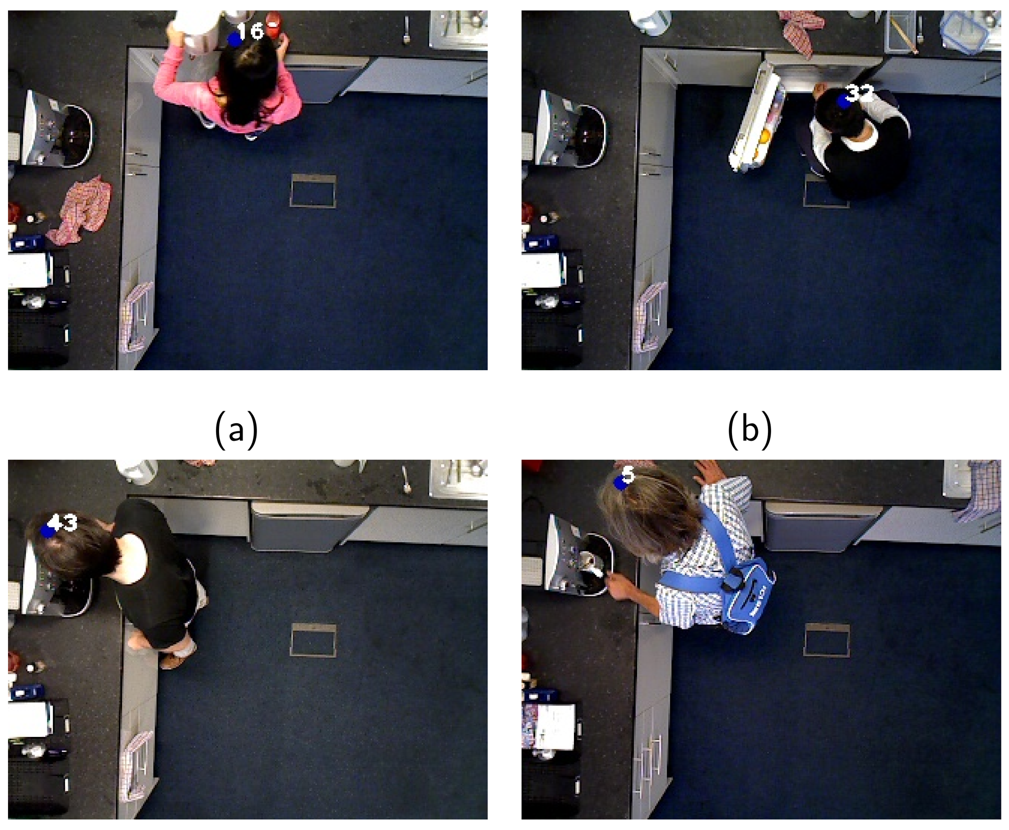
Information provided by head and hands detection algorithms can be used as input for a set of HMMs. Each of these recognise different actions sequence. After training the model, we consider an action sequence

$$s = \{s_1, s_2, \ldots, s_n\}$$

and calculate its probability λ for the observation sequence $P(s|\lambda)$. Then we classify the action as the one which has the largest posterior probability. Figure 1 depicts the general scheme of the recognition process. In particular, we used three different HMMs, which have as observations 3D points of: the head; the hands; both head and hands together. Finally, the classification module provides the action x_j that maximizes P_{HMM_i}. It is the HMM trajectory probability that follows the activity sequence s given the sequence of n observations, i.e.:

$$x_j = \arg\max_i P_{HMM_i}(X_{1:n} \in seq_n(s)|o_{1:n})$$

Setup and Acquisition



(a) (b) (c) (d)

Figure 2: Snapshots of RADiAL session registration.

To evaluate the usefulness of our approach for activity recognition, we built a new dataset (RADiAL) that contains common daily activities such as making coffee, making tea, opening the fridge and using the kettle. The RGB-D camera was installed on the ceiling of L-CAS laboratory at approximatel $4m$ above the floor (Figure 3).

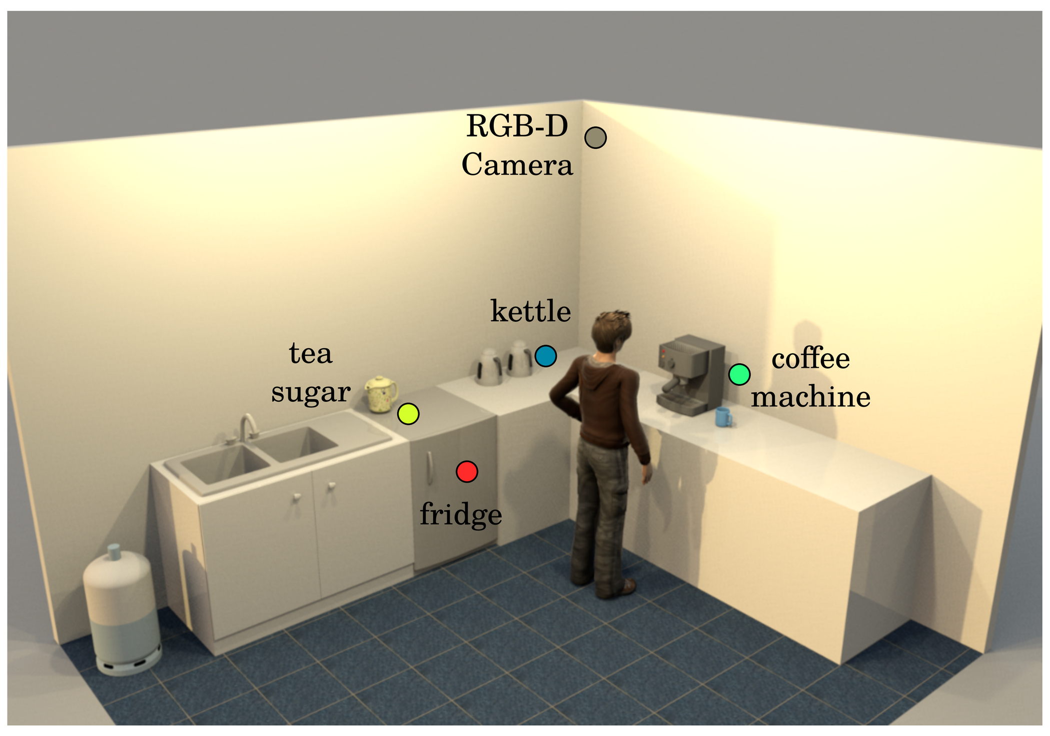


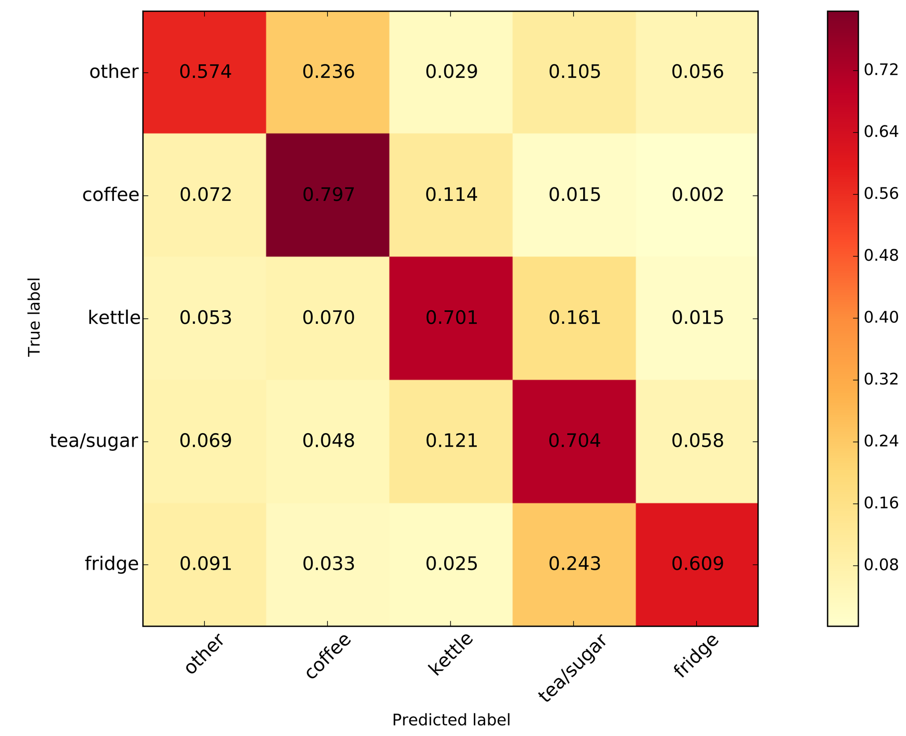
Figure 3: Reconstructed layout of the kitchenette where RGB-D camera is installed.

The RADiAL dataset[1] was collected in an open-plan office of the Lincoln Centre for Autonomous Systems (L-CAS). The office consists of a kitchenette, resting area, lounge and 20 working places that are occupied by students and post-doctoral researchers. We installed a ceiling RGB-D camera (Figure 3) that took a snapshot (with dimensions of 320×240 pixels, Figure 2) of the kitchenette area every second for 5 days, and we hand-annotated activities of one of the researchers over time. Furthermore, the RADiAL dataset contains the 3D positions of the head and hands for each person with a minute-by-minute timeline of 5 different activities performed at the kitchen over the course of days. RADiAL contains 100 trials. Each trial includes the actions related to one person.

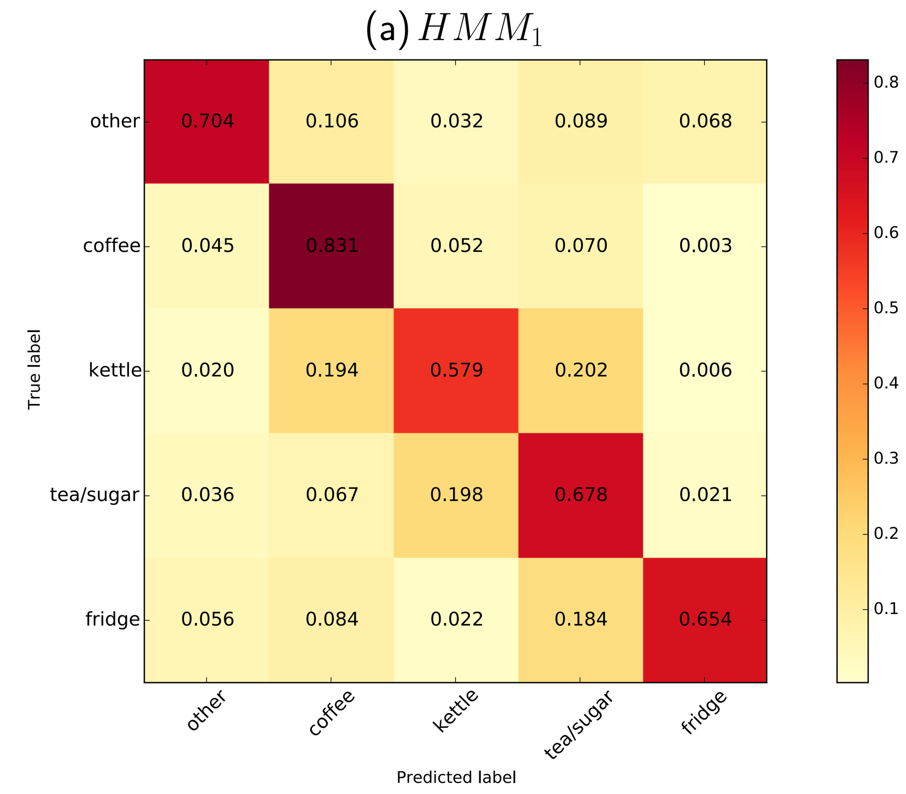
[a] http://vrai.dii.univpm.it/radial-dataset

Experimental Results

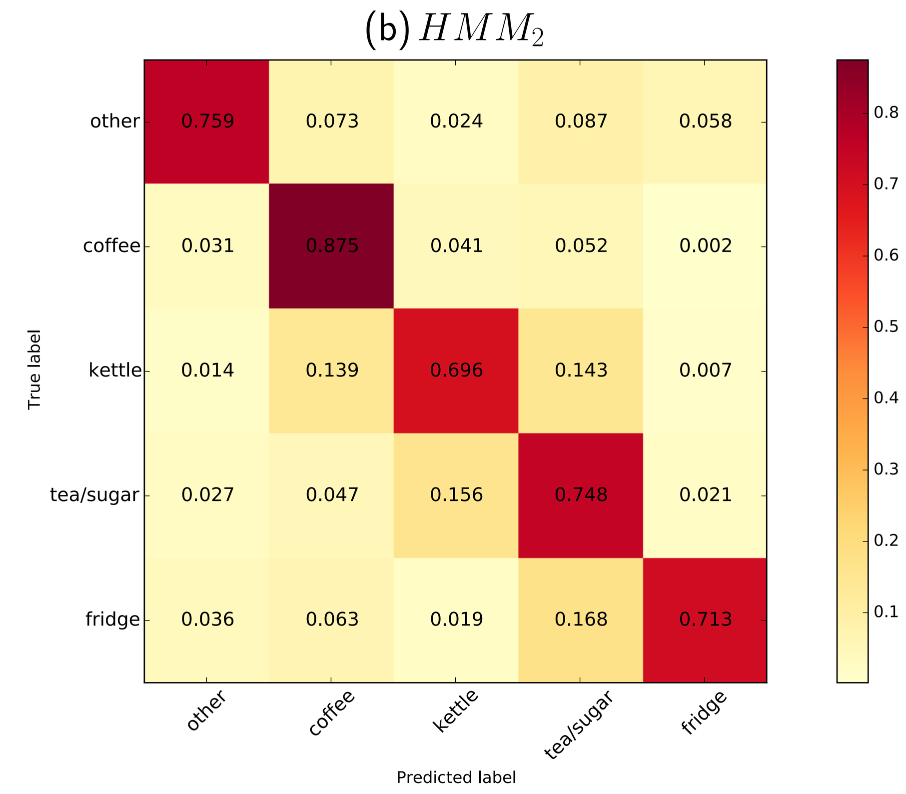
Five models were used to recognize activities in the RADiAL dataset and correspond respectively to the activities "other" (this action contains all the other activities perfomed in a kitchen environment), "coffee" (making a coffee), "kettle" (taking the kettle), "tea/sugar" (making tea or taking sugar), and "fridge" (opening the fridge). The results were obtained using two different validation techniques.



(a) HMM_1



(b) HMM_2



(c) HMM_3

Figure 4: k-fold cross-validation confusion matrices.

Table 1: Classification Results Cross Validation.

	HMM₁			**HMM₂**			**HMM₃**		
	PPV	TPR	F_1	PPV	TPR	F_1	PPV	TPR	F_1
other	0.73	0.57	0.64	0.89	0.70	0.79	0.93	0.76	0.84
coffee	0.67	0.80	0.73	0.69	0.83	0.75	0.76	0.87	0.81
kettle	0.60	0.70	0.65	0.47	0.58	0.52	0.58	0.70	0.63
tea/sugar	0.66	0.70	0.68	0.64	0.68	0.66	0.70	0.75	0.72
fridge	0.74	0.61	0.67	0.74	0.65	0.69	0.78	0.71	0.74
avg	0.68	0.68	0.68	0.73	0.71	0.71	0.78	0.77	0.77