# Cooking task dataset

**Documentation**

20 May 2015

## 1 General information

| | |
|---|---|
| **Project title** | Recognising the actions during cooking task |
| **Project id** | D2011-KTA-KHY |
| **Principal investigators** | Frank Krüger, Albert Hein, Kristina Yordanova, Thomas Kirste |
| **Affiliation** | Mobile Multimedia Information Systems, Computer Science, University of Rostock |
| **E-Mail** | kristina.yordanova@uni-rostock.de |
| **Date** | 2011 |
| **Type** | Measurement |
| **Location** | Smart Appliance Lab, Albert-Einstein-Straße 22, 18059 Rostock, Germany |
| **Keywords** | activity recognition, kitchen task assessment, cooking task |
| **Rights** | CC BY 4.0 |

### 1.1 Objective

Recognise the actions of a user and the manipulated objects while executing kitchen tasks.

### 1.2 Problem Statement

One person is preparing a meal in the kitchen that includes: preparing the ingredients for a soup; cooking the soup; serving the soup; having meal; cleaning the table; washing the dishes. The task is to recognise the fine-grained actions that constitute these tasks and the objects in the environment which the person is manipulating.

## 2 Description

This Dataset was recorded during the diplom thesis of Fabian Wolff. The results from the dataset can be found in the PlosOne paper"Computational State Space Models for Activity and Intention Recognition. A Feasibility Study" by Krüger et al. It contains 7 datasets that describe the execution of preparing and having a meal.

### 2.1 Data format

The `Raw/` folder contains 7 data files, one for each recording. The sensor data was recorded with motion capturing system based on wearable inertial measurement units (IMUs). Each data file contains the 16 most significant features from the original data for each time stamp.

The `annotation/` folder contains the corresponding annotation for the 7 datasets. For each time stamp it contains:

**16 action classes** where the observed class is indicated with 1, the rest with 0;

**4 locations** where the observed location is indicated with 1, the rest with 0;

**6 fixed places** where the observed place is indicated with 1, the rest with 0;

**10 objects** where the observed object is indicated with 1, the rest with 0;

**hands** where 1 indicates that there is object in the hand, 0 that the hands are empty.

**annotation** where the executed action, the place, and the objects being taken are described in the form *action-object-fromLocation-toLocation*.

### 2.2 Dataset recording

The sensor data was recorded with motion capturing system based on wearable IMUs. For each sensor three axis acceleration and angular rates were recorded, with a sampling rate of 120 Hz. The resulting data stream of 5|6 30 signals was segmented into frames using a simple window-based segmentation with a window size of 128 samples and 75% overlap, giving a frame rate of 3.75 Hz. For each frame, mean, variance, skew, kurtosis, peak, and energy were computed for each signal. This stream of 180-dimensional feature vectors at 3.75 Hz was then subjected to dimension reduction by applying principal component analysis to the full set of feature vectors, choosing the loadings of the factors corresponding to the 16 largest eigenvalues as effective observations. The sensor data

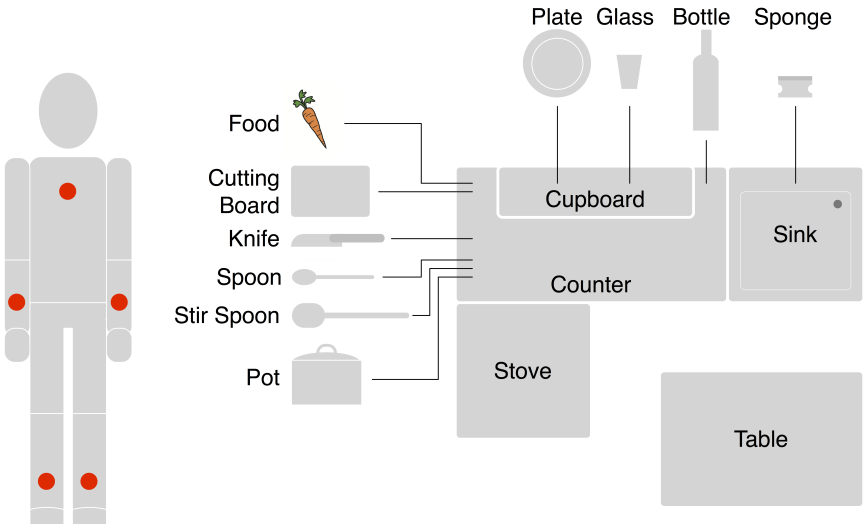


Figure 1: Positions of the sensors on the body (with red). Layout and objects in the experiment.

was then scrambled to reduce order effects and dependencies between observations in durative actions. The experiment layout and the sensors locations are shown in Figure 1. Seven voluntary subjects acted according to a script to generate the observation data for the datasets.