

Dataset Recording Briefing Document

The aim of this document is to define and describe specific tasks in order to generate RGB-D train and test data video-recordings for my thesis. Conceptually, these tasks have in common that a robot R observes a human subject S , who naturally and bimanually performs several tasks. The generated recordings should capture the described tasks from the robot's point of view – the robot is therefore simplified by an RGB-D camera on a tripod at approximately 1.70 m.

This document is structured as follows: The first section will roughly illustrate the context for the tasks, whereas the consecutive section will define, how the subjects should be briefed and what they may or may not know. Finally, last section will specify, how a given task should be acted out and what objects are relevant.

Context Definition

The tasks take place in either a kitchen context or a workshop context. While there is a differentiation between kitchen and workshop contexts in the tasks, it does not make a difference from a conceptual point of view.

We consider top-down-like tasks as depicted in Figure 1, where a human subject S and a robot R stand in front of a counter C , while the subject is performing various activities on it. The counter can either be a kitchen counter or a workbench – the only important aspect is that the subject can conveniently work on the counter's surface while standing. Workbenches and kitchen counters usually have a height of about 90 cm, while desktops or coffee tables are considerably lower.

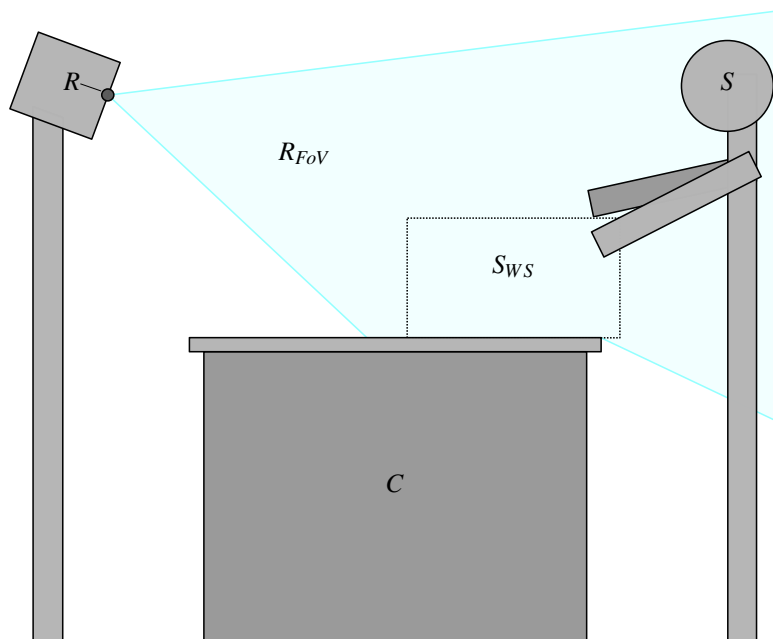


Figure 1: Schematic of the experimental setup. R : Robot, simplified as point (RGB-D camera's lenses); R_{FoV} : Robot's field of view; S : Human subject; S_{WS} : Subject's workspace; C : Counter (height: approximately 90 cm).

The robot's field of view R_{FoV} must include both the subject's workspace S_{WS} and shoulders, and it should include the subject's head if possible. Both the robot's height (reference: 1.7 m) and its distance to the counter may be adjusted to ensure this (i. e. it is preferred to adjust those parameters over not having the subject's head on tape). The subject's workspace is a hypothetical fixed-size bounding box which contains all the objects the subject interacts with, for each given moment in a recording. In other words: For each recording, all objects the subject interacts with should be in the robots field of view and not cut off.

Subject Briefing

The tasks are laid out for 6 subjects (3 female, 3 male) repeating each task 10 times to account for inter- and intra-subject-variations and for different ways of doing certain tasks (to a certain degree). Therefore, it is up to the subject to interpret a task's procedure. The supervisor may intervene and tell a subject to do a certain task differently (e.g. different order), but only after the subject has initially completed a recording of a task on their own (unless they misunderstood something). Subjects must not watch the recording of another subject to avoid them imitating unspecified sub-tasks, and subjects may know the full procedure of a task in advance.

While performing their tasks, the subject can and should move freely and naturally, but should also be aware of the extends of the robot's field of view to not leave it for too long. Said tasks are defined in the following.

Task Definitions

In this section, all tasks will be outlined. There are 9 tasks in total, 5 of which in a kitchen context, and 4 in a workshop.

Task 1 / Kitchen: Cooking with Bowl and Bottle

Context: *Kitchen*

Key objects: *A mixing bowl, a whisk, and a bottle*

In the first task, the subject will prepare a meal. It is assumed that the bowl contains the meal, and the bottle is already open and filled with oil, vinegar, or the like. Procedure:

- The bowl, bottle, and whisk are placed on the counter; The subject stands in front of it
- The subject grasps for the whisk with their dominant hand and for the bowl with their non-dominant hand to hold it
- The subject stirs the contents of the bowl with the whisk
- The subject releases the bowl and grasps for the bottle while still stirring
- The subject pours the contents of the bottle into the bowl while still stirring (in the recording, the bottle should clearly be above the bowl, so the subject should not hold the bottle too low)
- The subject places the bottle back on the counter while still stirring
- The subject grasps for the bowl again and keeps stirring for a short while
- The subject restores the initial state by putting the objects back to where they were at the beginning

Brown glass bottles that are used for oil would be good, since they are easier to detect (that is, they are not transparent), and heavy, so they will not so easily fall over when being handled as empty plastic bottles would do.

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.

Task 2 / Kitchen: Cooking with Bowls

Context: *Kitchen*

Key objects: two *mixing bowls* and a *whisk*

This task is just like Task 1, except for the bottle, which is now replaced with another bowl. So while stirring, the subject uses the new bowl instead of the bottle to add ingredients. It does not matter too much if the bowls are similar or not, regardless in which aspect (colour, shape, model, ...), as long as they can clearly be identified as such.

Task 3 / Kitchen: Pouring Water

Context: *Kitchen*

Key objects: A *cup* and a *bottle*

The subject pours themselves water and takes a sip. The bottle and the cup are placed on the counter, while the subject stands in front of it. The exact procedure is not defined in order to allow the subject to complete the task as they would naturally do it.

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.

Task 4 / Kitchen: Cleaning the Counter

Context: *Kitchen*

Key objects: A *sponge* and 1–3 random objects

The subject cleans the kitchen counter with a sponge. A few (1–3) other random objects are placed on the counter. The objects should be known from other tasks and fit in contextually (like cups, bowls, bottles, ...). Procedure:

- The sponge and the objects are placed on the counter; The subject stands in front of it
- The subject grasps for the sponge with their dominant hand
- The subject begins wiping the counter
- The subject successively lifts each object with their non-dominant hand to wipe underneath it. The objects should be held high. Before lifting the next object, there should be a short phase of wiping the counter with all objects placed on it. Objects should not be relocated, only lifted
- The subject restores the initial state by placing the sponge back on the counter

Task 5 / Kitchen: Preparing a Bowl of Cereals

Context: *Kitchen*

Key objects: A *banana*, a *knife*, a *cutting board*, a *bottle of milk*, a *bowl*, and a *pack of cereals*

The subject will prepare a bowl of cereals with fresh ingredients in this task. A plastic banana may be used which can be assumed to be peeled, and the bottle of milk is already opened. Procedure:

- The cutting board, the knife, the bowl, the bottle of milk and the pack of cereals are placed on the counter. The banana is placed on the cutting board; The subject stands in front of the counter
- The subject pours cereals from the pack into the bowl
- The subject cuts the banana into slices

- The subject takes the cutting board and shoves a few slices of banana from the cutting board into the bowl using the knife
- The subject pours milk into the bowl

The subject may also cut the banana into slices as the first step, and then pour all prepared ingredients into the bowl.

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.

Task 6 / Workshop: Disassembling a Hard Drive

Context: *Workshop*

Key objects: *A hard drive and a screwdriver*

The subject disassembles a hard drive by unscrewing the case screws using a screwdriver. A real hard drive should be used and the subject should actually loosen screws. The screwdriver may be magnetic. The case of the hard drive should not be removed – unscrewing is enough. Procedure:

- The hard drive and the screwdriver are placed on the workbench; The subject stands in front of it
- The subject grasps for the screwdriver with their dominant hand and fixates the hard drive on the workbench with their non-dominant hand
- The subject starts loosening screws on the casing using the screwdriver
- After each loosened screw, the subject should remove the screw from the casing and lay it aside
- After loosening a few screws (2 or 3), the subject restores the initial state by putting back all objects to where they were in the beginning

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.

Task 7 / Workshop: Freely Disassembling a Hard Drive

Context: *Workshop*

Key objects: *A hard drive and a screwdriver*

This task is similar to Task 6, but beginning with the second step, the subject will hold the hard drive in their non-dominant hand instead of fixating it on the workbench.

Task 8 / Workshop: Hammering

Context: *Workshop*

Key objects: *A hammer and a piece of wood*

In this task, the subject will hammer a few nails (2 or 3) into a piece of wood. Real nails should be used. Procedure:

- The hammer, the nails, and the piece of wood are on the workbench; The subject is standing in front of it
- The subject takes a nail with their non-dominant hand and the hammer with their dominant hand
- The subject holds each nail on the wood with their non-dominant hand and hammers it slightly in to fixate it. Then, the subject will go on to fixate the wood with their non-dominant hand and hammer the nail in entirely

- With the nails being hammered in, the subject will restore the initial state by placing the hammer on the workbench

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.

Task 9 / Workshop: Sawing

Context: *Workshop*

Key objects: A *saw* and two pieces of *wood*

In this task, the subject will saw a piece of wood in half. For this it is best to tape two pieces of wood together instead of actually sawing one piece (better reproducible, no dust, faster). Procedure:

- The saw and the piece of wood are placed on the workbench; The subject stands in front of it
- The subject grasps for the saw with their dominant hand and holds the wood with their non-dominant hand
- The subject pretends to saw the wood by slowly cutting the tape with the saw
- Before completely cutting through, the subject lays aside the saw to manually tear the pieces of wood apart

If it is possible to completely saw through the wood without one of the pieces falling off the workbench, prefer doing so (for example by using a clamp or similar).

It is desirable to place contextually fitting objects (0–2) on the table to add authentic noise.