PRIP-TR-124

Experiments for Assistive Living in the Smart Room

Tamir Hassan, Rainer Splechtna,
Anton Fuhrmann,
Walter G. Kropatsch

Pattern
Recognition &
Image
Processing
Group

Institute of
Computer Aided Automation

Technical Report

Pattern Recognition and Image Processing Group
Institute of Computer Aided Automation
Vienna University of Technology
Favoritenstr. 9/183-2
A-1040 Vienna AUSTRIA
Phone: +43 (1) 58801-18351
Fax: +43 (1) 58801-18392
E-mail: tam@prip.tuwien.ac.at
URL: http://www.prip.tuwien.ac.at/

PRIP-TR-124 December 2009

Experiments for Assistive Living in the Smart Room[1]

Tamir Hassan, Rainer Splechtna, Anton Fuhrmann, Walter G. Kropatsch

Abstract

This report describes the experiments that were carried out in December 2009 at VRVis to demonstrate how blind (or blindfolded) persons can be assisted by computer-based tracking and feedback systems. The positions of the user are tracked precisely in six degrees of freedom (three dimensions position and rotation) using three infrared cameras. Feedback is generated by the use of a 3D headset. This report describes in how the set-up was created, and suggestions are also included for continuation work.

[1]This work is partially supported by the Austrian Science Fund under grant S9103-N13.

Contents

1 Introduction

As a joint venture between VRVis and PRIP, both institutes invested in hardware for a Smart Room; a large room with a number of video cameras and related equipment, enabling the practical development and trialling of intelligent solutions based on video acquisition. It was planned that experiments would be undertaken at both institutes to demonstrate the effectiveness of such solutions.

Unfortunately, in the weeks prior to the experiments, two of the cameras at PRIP developed a fault and needed to be sent back to the manufacturer for repair. It was therefore no longer possible to continue with the experiments at PRIP during the runtime of the CogVis project. This report therefore describes the experiments undertaken in December 2009 at VRVis to demonstrate how the newly acquired equipment enables such solutions to be prototyped.

This work is based on a previous demonstration in the Smart Room, carried out at PRIP in March 2008 and documented in [1].

2 System setup

2.1 Input

The room is surveilled by three infrared cameras, located along a single wall, as shown in Figures 1–2. All cameras are mounted on the ceiling. A table is also present, which was used as an obstacle in two of the experiments.

Objects in the room are tracked by having a marker plate/optical target (a fixed configuration of typically 4–6 single markers) affixed to them, which is detected by the cameras. Each tracking camera acquires image data within the infrared band and processes it using its integrated Linux-based computer. The extracted 2D marker positions are sent to the tracker server via Ethernet. The tracker server calculates six degrees of freedom data (three dimensions position and rotation) for each optical target from the cameras' 2D marker positions and sends the resulting 3D position and orientation relative to the reference coordinate system (via Ethernet) to the computer which runs the Smart Room software. The Smart Room software uses *OpenTracker* [2] to integrate the data sent by the tracker server. The location of the marker plates is tracked with an accuracy of a few millimetres and an update rate of 60 Hz.

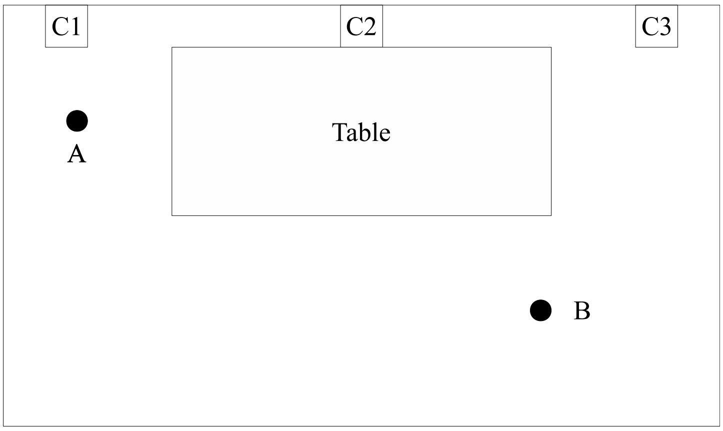

Figure 1: Experiment 1: walking from A to B

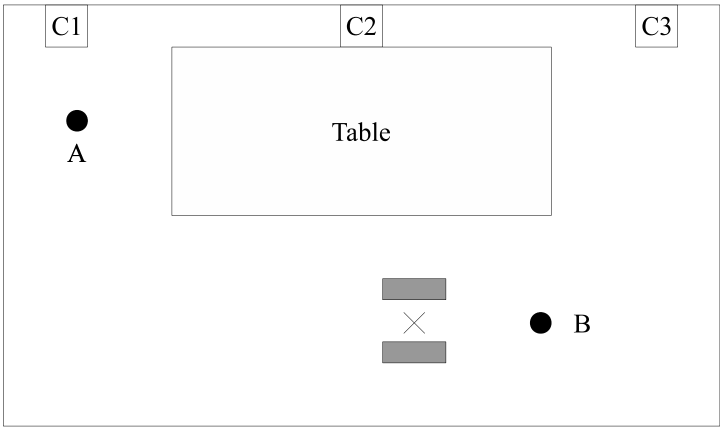

Figure 2: Experiment 2: negotiating a doorway

Three marker plates were affixed to the test person; one on the head and one for each arm, close to the hand. In Experiment 3, an additional marker plate was used as the target.

Since the position and orientation of all tracked objects (for head, hand, moving objects) and static objects is known and each object is represented by a geometric model within the system the distance between head/hand and the different obstacles/objects can be calculated on the fly. Also the direction of an obstacle/object relative to the current orientation of the head can be calculated.

Figure 3: Experiment 2: negotiating a doorway

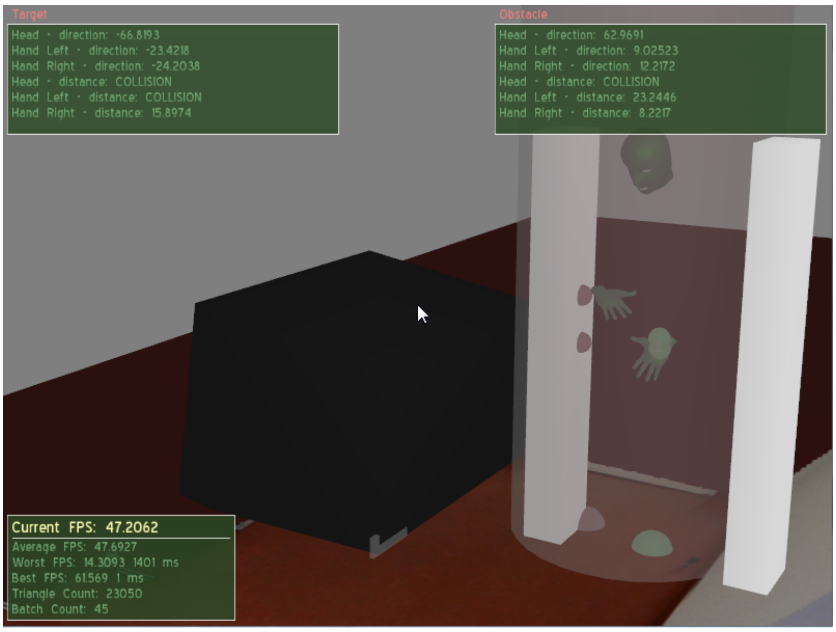

Figure 4: 3D reconstruction of the doorway in Expt. 2

Figure 5: Experiment 3 at the start

Figure 6: Experiment 3 being carried out

2.2 Output

Feedback was given to the user using 3D sound, which was generated with a 3D sound card and output to a set of wireless headphones. For audio processing and spatialization we use the *Cross-platform Audio Creation Tool (XACT)* [3]. XACT is an audio programming library and engine released by Microsoft as part of the DirectX SDK. XACT uses X3DAudio, a spatialization helper library, to calculate the values of the volume level matrix (the set of speaker/volume pairs that give a sounds its perceived position in a 360 degree circle around the listener at unity volume). We are using a sound card with Creative X-Fi sound processor. This sound processor is capable of applying HRTF (head related transfer functions) in real-time (the CMSS-3D mode for headphone has to be activated in the settings of the sound card driver) so that the direction of the different sounds which are calculated by the SmartRoom software can be optimally conveyed via headphone.

Using this set-up, it is possible to communicate the location of 2–5 objects at once, by emitting a given sound repeating approximately every half-second. The following parameters were used to relate the sound to the measured distance and direction: the direction of the sound was used to represent the direction of the object (or obstacle); the frequency of the sound was used to represent the object's proximity, with the frequency increasing as the test person nears the object.

A "tick" sound was used for the obstacle and a "ding" sound was used for the goal in the experiments. Normally, the distance between the obstacle and the test person was measured from the test person's head. However, when the test person's arm got to within approximately 10 cm of the obstacle, the arm's position was used instead, and the sound changed to a recorded voice saying "left" or "right", depending on which arm was used. This enabled the test person to more accurately "feel" his/her way around an obstacle.

3 Experiments

Three experiments were carried out, in order to determine the degree of improvement offered by the system and suggest directions for future work. For each of the experiments, the amount of time taken to complete the experiment was measured; for the first two experiments, in which the test person was expected to move considerably around the smart room, his/her trajectory was recorded.

In order to enable a comparison, all experiments were carried out first without the use of the system and, afterwards, with the use of the system.

All experiments were carried out using three different test persons, one of which was not connected with any work on the smart room project. The test persons were instructed to keep their eyes closed throughout the entire experiment. The three experiments are described in the subsections below.

Each test person was given ample time to practice the experiment and become familiar with the setting, so that no obvious disadvantage was present in the first run compared to the second run.

3.1 Experiment 1: Walking around a corner

In the first two experiments, the test person had to find his/her way between two points, **A** and **B**, in the room, avoiding obstacles. After reaching point **B**, the test person returned to point **A**. This sequence was repeated four times.

Throughout the experiment run, the proximity and direction of the obstacle(s) was communicated to the test person.

For the first experiment, the table was the only obstacle in the room, and was stationary. Point **B** was located around the corner from point **A**. The test person therefore had to go around the corner and back four times to complete the experiment. This set-up is shown in Figure 1.

3.2 Experiment 2: Negotiating a doorway

For the second experiment, the test person also had to move from point A to point B and back four times, but had to negotiate an (open) doorway located just before point B. A virtual target was used inside the centre of the door frame, which was output to the test person. This set-up is shown in Figure 2.

For both of the above experiments, the trajectories of the test person as well as the timings of each test person were recorded.

3.3 Experiment 3: Choosing the correct bucket

In the third experiment, the test person was not required to find his/her way around the room. Four buckets were placed on a table located next to the test person. In one of these buckets, a target object was placed. The target marker was placed behind this bucket, and its location output to the test person. The test person had to find the object in the correct bucket, and the time required to do this was recorded. The test person was only allowed to use one arm (of his/her choosing) for this experiment.

This experiment was repeated ten times for each test person. Before each test run, another (or the same) bucket was randomly chosen, and care was taken to ensure that the test person could not determine which bucket was chosen by the noise that was produced.

Figure 5 shows the experiment at the beginning, and Figure 6 during a test run.

4 Results

The trajectories for Experiment 1 are given in Figure 7 and the timings are in seconds as follows:

	Person 1	**Person 2**	**Person 3**	**Averages**
without system	114.6	126.4	228.6	156.5
with system	116.2	100.6	171.0	129.3

The trajectories for Experiment 2 are given in Figure 8 and the timings are in seconds as follows:

	Person 1	**Person 2**	**Person 3**	**Averages**
without system	146.3	160.2	163.9	156.8
with system	222.9	190.8	202.3	205.3

The timings for Experiment 3 are given in Table 3.

5 Discussion

In the first experiment, a marked improvement in the trajectories of the first two test persons can be observed. The third test person did not show such a significant improvement, and the trajectory with the aid of the system also contains an anomalous point, which is likely a measurement error, possibly caused by another reflective object or occluded markers.

With the aid of the system, the tests were generally concluded faster. The first test person, whose trajectory was very tight, actually required two seconds longer to complete the test as without the system. Clearly, the test person's individual behaviour has a great influence on the results. However, an improvement — either in the time taken or the accuracy of the trajectory — was observed in all three cases.

In the second experiment, improvements to the trajectory are particularly obvious in the area around the target, the door. This would suggest that

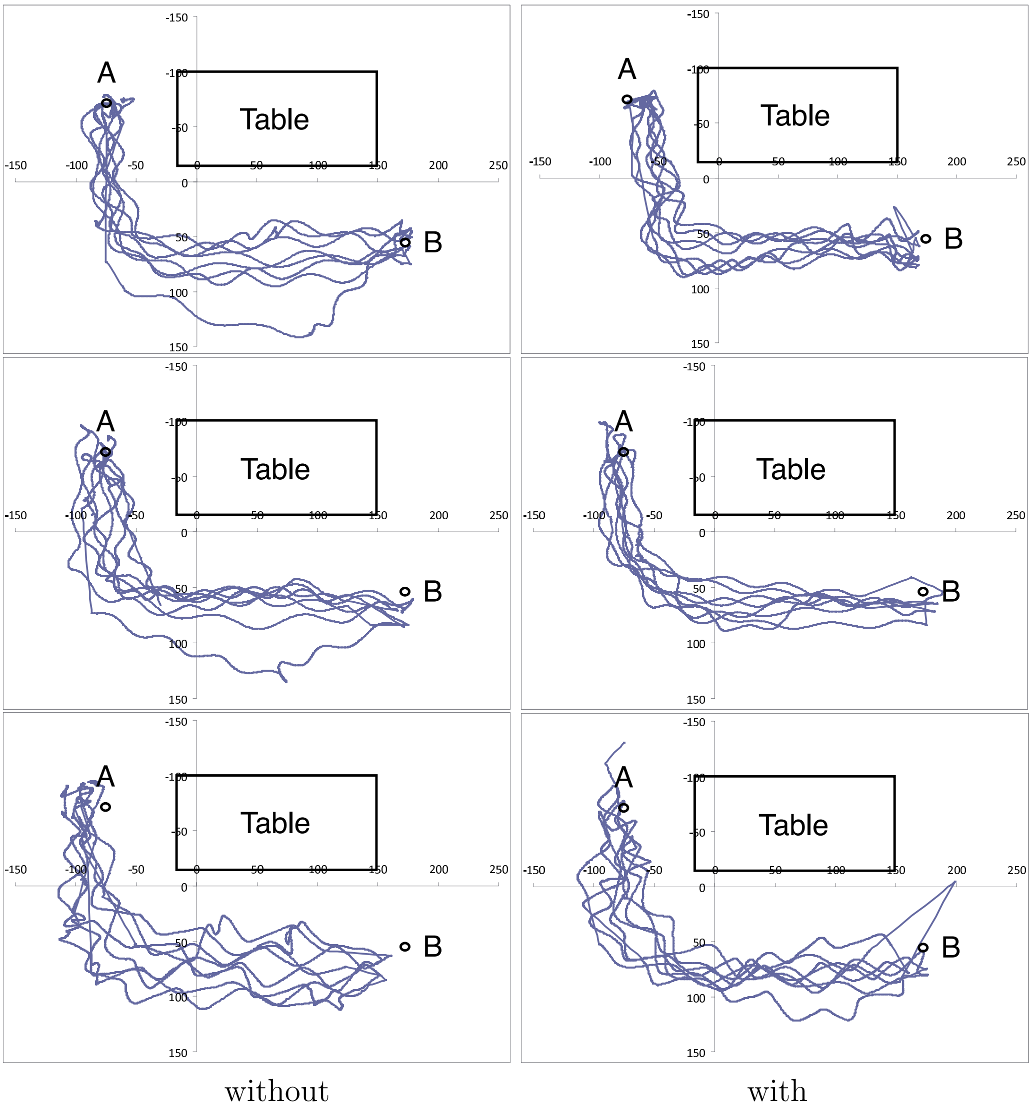

Figure 7: Trajectories of Expt. 1: walking from A to B. The rows indicate the results of different test persons. The left column shows the results without the aid of the system; the right column shows the results with the system being active.

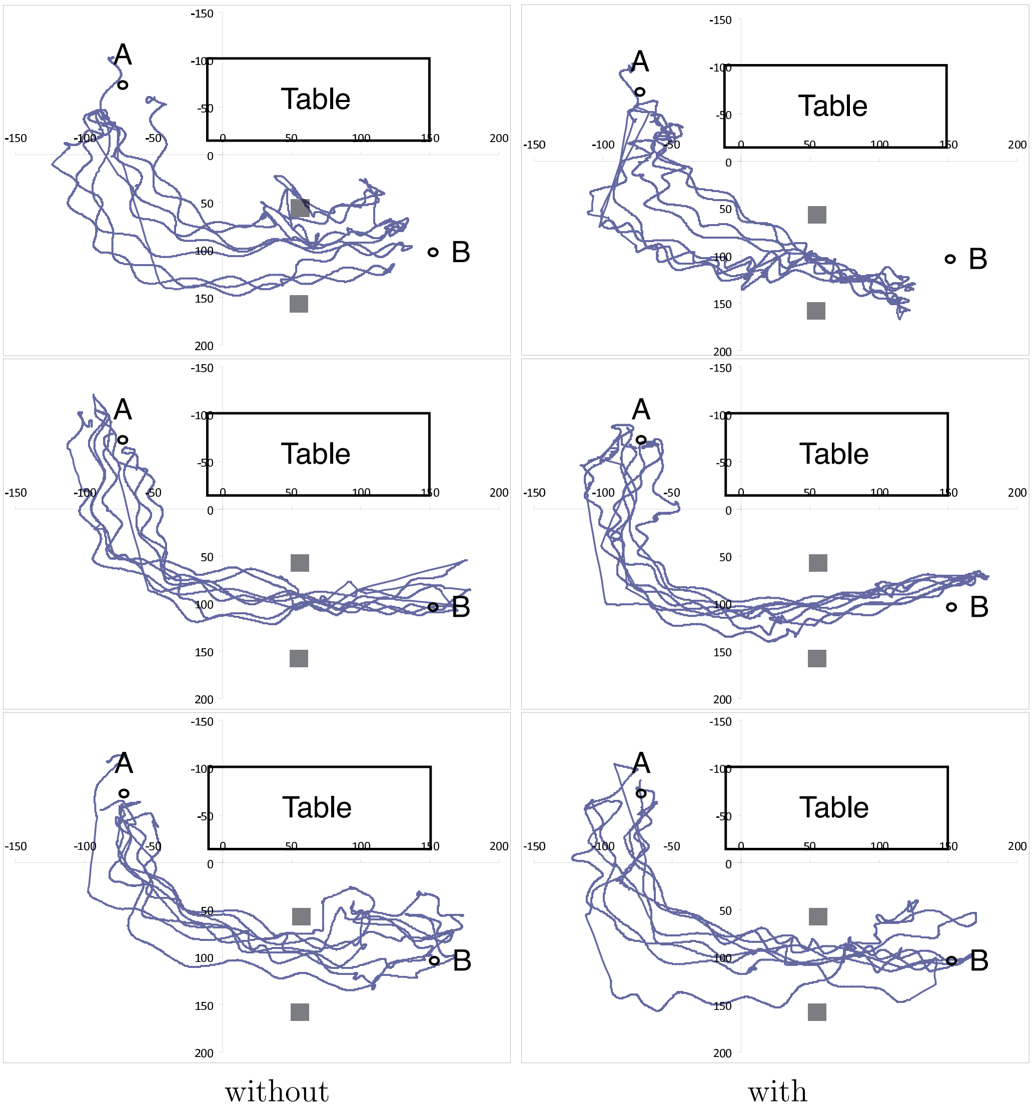

Figure 8: Trajectories of Expt. 2: negotiating a doorway. The rows indicate the results of different test persons. The left column shows the results without the aid of the system; the right column shows the results with the system being active.

cognitive vision systems should perhaps be focussed on communicating to their users where they should be heading to, rather than which obstacles they should avoid. If the user has a specific task to perform, the system should figure out the safest, least cost path and tell him/her what to do, rather than leave the user to explore the entire physical space.

In comparison to the first experiment, all test runs actually took longer with the aid of the system than without. This could be due to the test persons requiring significant time to process the information from the computer and building a mental image of the test arena.

What the results do not show is that no collisions occurred when the system was enabled; without the use of the system, collisions occurred very often. This shows that assistive vision could be of particular importance in situations where the relative “cost” of a collision is high, e.g. in a kitchen scenario, where a collision could lead to injury or harm.

The third experiment also showed a net average improvement in time to complete the task (although the third person again took slightly longer). With the aid of the system, the test persons took slightly longer to reach for the object, but always picked up the correct object first time. With the system switched off, the test persons tried each object randomly until they were told that they had found the correct object. In a real world situation, they would not necessarily have been able to know that they had found the correct object at all, or have required significant extra time to discover this (consider the example of a salt and pepper shaker). This shows that cognitive vision systems could be of great benefit in situations where objects cannot be readily distinguished by their external shape.

6 Future work

These experiments provide a basis upon which future, more elaborate real-life settings could be constructed. The scenario of the home kitchen is an obvious real-life example. But other situations, such as avoiding dangers at home, such a staircase, also present good candidates. The location and orientation of the cameras should also allow for tracking of the user, even when certain markers are occluded.

References

[1] Hassan, T., Ion, A., Kropatsch, W.G.: A Demonstration for the Smart Room. PRIP Technical Report 119. (2008) 2

	without aid of the system			with aid of the system		
Test run	**Person 1**	**Person 2**	**Person 3**	**Person 1**	**Person 2**	**Person 3**
1	4.2	2.4	3.5	3.0	5.3	5.1
2	12.2	7.6	8.1	3.7	5.1	6.1
3	2.9	7.3	6.9	3.7	4.1	4.5
4	4.8	3.3	4.6	6.5	3.4	4.6
5	4.8	3.2	7.0	3.3	3.0	7.3
6	8.6	4.4	3.5	4.1	4.7	6.4
7	10.3	7.1	6.1	3.7	4.6	5.2
8	5.5	8.1	6.8	3.4	2.6	7.6
9	4.5	4.5	3.5	5.4	3.1	6.9
10	6.5	21.3	7.2	3.8	6.3	7.2
Averages	6.4	6.9	5.7	4.1	4.2	6.1

Table 3: Timings (in seconds) of Experiment 3

[2] Reitmayr, G. and Schmalstieg, D.: Opentracker - a flexible software design for three-dimensional interaction. In *Virtual Reality*, 9(1), pp. 79–92, 2005 2

[3] Microsoft Cross-Platform Audio Creation Tool, `http://msdn.microsoft.com/en-us/library/ee416126%28VS.85%29.aspx` 6